

United States
Environmental Protection
Agency

Atmospheric Sciences
Research Laboratory
Research Triangle Park NC 27711

Research and Development

EPA/600/9-87/024 Oct. 1987



PROCEEDINGS

WORKSHOP ON EVALUATION/DOCUMENTATION OF CHEMICAL MECHANISMS

WORKSHOP ON EVALUATION/DOCUMENTATION OF CHEMICAL MECHANISMS

**Dr. Roger Atkinson, Workshop Chairman
Statewide Air Pollution Center
University of California-Riverside
Riverside, CA 92521**

Steering Committee

**Dr. Harvey Jeffries
University of North Carolina-Chapel Hill**

**Dr. Gary Whitten
Systems Applications, Inc.**

**Dr. Fred Lurmann
Environmental Research and Technology**

Project Officer

**Basil Dimitriadis
Atmospheric Chemistry and Physics Division
Atmospheric Sciences Research Laboratory
Research Triangle Park, NC 27711**

**Atmospheric Sciences Research Laboratory
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711**

DISCLAIMER

Although the workshop described in this report has been funded wholly by the U.S. Environmental Protection Agency under contract 68-02-4183 to Systems Research and Development Corporation, the report has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ABSTRACT

Atmospheric photochemists and developers and users of air quality models need to discuss the evaluation and documentation of chemical mechanisms used in air quality simulation models. A Workshop, therefore, was organized and conducted on December 1-3, 1987 by EPA to discuss the latest evidence and viewpoints on the subject and to solicit from experts recommendations on optimum approaches to mechanistic model evaluation, documentation, and further development. Previous practices and underlying issues in the subject areas were reviewed and discussed in background documents prepared and distributed in advance of the Workshop. Participants agreed that smog chamber data provide the most unambiguous test of urban atmospheric photochemistry mechanisms. They also agreed, however, that there are uncertainties associated with the representation of chamber radical sources and of photolytic rates in outdoor chambers, with smog chamber measurement errors, and with the representation of as yet unknown reaction pathways. The participants recommended that task forces and/or review groups be established to discuss and resolve existing smog chamber methodology issues, to assemble a required smog chamber data base for mechanism testing, and to review/evaluate relevant kinetic and smog chamber data and mechanism testing results. Recommendations were also developed on future research needs and on mechanism documentation procedures.

INTRODUCTION

The purpose of this workshop is to solicit from experts input that would help EPA meet its commitment to provide guideline models for use in the development of effective air pollution control strategies. Specific objectives are (a) to develop procedures that could be agreed upon for documenting chemical mechanism modules so as to ensure error-free transfer of the modules from the developer to the user, (b) to critically examine and rank approaches to evaluating the accuracy of gas phase chemical mechanisms for the ozone, aerosol, and acid rain aspects of ambient air quality, (c) to identify all important factors that determine the reliability, comprehensiveness and precision of each mechanism evaluation approach, (d) to define a "standard" data base needed for evaluating a mechanism by a given approach, and finally, (e) to develop, if possible, standard mechanism evaluation procedure(s).

The underlying problem and the need for standard procedures for documenting and evaluating chemical mechanisms were first recognized and discussed during a 1983 workshop on the Empirical Kinetic Modeling Approach (EKMA). At that time, EPA was disturbed by the fact that models of insufficiently documented validity were offered as official Agency guidelines for development of costly control strategies. Furthermore, the plethora of ozone mechanisms in existence, their differences in terms of ozone prediction, and the difficulty in documenting even the relative validities of the various models caused another problem: they encouraged the inappropriate practice of State government officials responsible for development of State Implementation Plans (SIP), to select for application those mechanistic models that happen to support pre-conceived control targets. Therefore, for EPA to be able to defend its control policies and regulations, and for the States to be able to develop objective and effective SIP's, it is imperative that the model-type guidelines issued by EPA be of well documented utility and validity.

The workshop objective to develop standard procedures for documenting mechanistic models is an important and difficult one, but its achievement does not necessarily hinge upon resolution of complex scientific issues. Therefore, while detailed discussions will be needed to derive the collective judgment of the Workshop group, no dichotomy in viewpoint or intensive debate among the workshop participants is to be expected. It should be possible to arrive at a consensus, based on the personal experiences of the participants in dealing with the problem of using someone else's model. In contrast, the objectives related to the evaluation of chemical mechanisms are almost certain to reveal considerable diversity in viewpoint and to incite intensive debate. This is the normal result of the fact that evaluation of chemical mechanisms is fraught with unresolved issues most of which, in the lack of convincing experimental evidence, have acquired a subjective, even philosophical character. For example, questions on the evaluation approach for which there are no consensus answers are: Can field data be used to evaluate chemical mechanisms? Is comparison against smog chamber data the best approach to evaluating mechanisms? Given the conceptual and practical problems of the experimental

approaches, is the theoretical approach (formulation of multi-thousand step "master" mechanism) the preferable approach to developing valid mechanisms? In addressing these questions, there is unavoidable clash between those believing in the unchallengeable validity of real world data and those believing that real world is hopelessly complex; and also between those advocating that theory shall be used as "gospel" and those for whom experimental evaluation is the only option available.

It is indeed unrealistic to expect that the workshop participants will resolve all technical issues and agree, for example, on a single mechanism evaluation procedure or on a "standard" data base of reasonable proportions. It is expected, however, that it will be possible for specific alternatives or options to be identified, clearly explained, thoroughly scrutinized, and, finally, rated based on their relative merits and limitations. It is with these latter expectations in mind that the workshop organizers decided on the format and size of the workshop, and, indeed, on the workshop itself.

Workshop Chairman,
Basil Dimitriades

CONTENTS

Abstract.	iii
Introduction.	iv
Summary and Recommendations	1

Papers

- 1 The Science of Mechanism Development and Testing
Harvey Jeffries
University of North Carolina-Chapel Hill
- 2 How Do You Know a Good Model When You See One?
Joseph Tikvart
U.S. Environmental Protection Agency
- 3 Chemical Mechanism Development for Applications in Atmospheric Simulation
Kenneth Demerjian
State University of New York-Albany
- 4 Development and Testing of Gas-Phase Chemical Mechanisms: Present Status
and Future Research Needs
Roger Atkinson
University of California-Riverside
- 5 Review of the "Strawman" Document for the EPA Workshop on Evaluation/
Documentation of Chemical Mechanisms
Michael Gery
Systems Applications, Inc.
- 6 Uncertainty Analysis and Captive-Air Studies as Aids in Evaluating
Chemical Mechanisms for Air Quality Models
Alan Dunker
General Motors Research Laboratories
- 7 Development and Testing of a Surrogate Species Reaction Mechanism for
Use in Atmospheric Simulation Models
William Carter
University of California-Riverside
- 8 Evaluation and Intercomparison of Chemical Mechanisms
Fred Lurmann
Environmental Research and Technology, Inc.

Appendices

- A. Need for Chemical Mechanism Documentation
Kenneth Sexton and Harvey Jeffries
University of North Carolina
- B. The Science of Photochemical Reaction Mechanism Development and Evaluation
[Strawman Document]
Harvey Jeffries and Jeffrey Arnold
University of North Carolina
- C. Workshop Agenda
- D. Workshop Attendees

EPA Workshop on Evaluation/Documentation of Chemical Mechanisms

Summary and Recommendations

prepared by
Workshop Steering Committee:
Roger Atkinson, Chairman,
Harvey Jeffries, Fred Lurmann,
and Gary Whitten

Introduction

The EPA conducted this workshop to discuss the evaluation and documentation of chemical mechanisms used in air quality simulation models. A chemical mechanism is the set of chemical reactions and associated rate constants which describes the transformation of emitted chemicals into intermediate and final products. In the context of this workshop, the initially emitted chemicals are hydrocarbons and oxides of nitrogen, and ozone is the product species of major interest. From this list of chemical reactions, a set of coupled, non-linear, differential equations which describe the rates of change of chemical species concentrations with time are derived. These equations can be integrated using a computer to model the changing concentrations of these chemical species. The major use of chemical mechanisms is to serve as a component of comprehensive air quality models that predict atmospheric smog formation by simulating the emissions, transport, dispersion, chemical reactions, and scavenging process that affect air pollution.

A chemical mechanism must have a certain level of detail to describe adequately the complex and detailed photochemistry of smog formation. Limitations on the size and speed of computers and in the availability of data for inputs to air quality models place restrictions on the detail allowed in a chemical mechanism. Thus the development of practical chemical mechanisms for use in urban air quality models requires a careful balance between detail and simplicity.

As discussed by Atkinson and Jeffries in their presentations, the goals of this workshop were:

- to assess present practice in photochemical reaction mechanism development and testing for those mechanisms intended for use in urban air quality control calculations;
- to determine if there might be a commonly agreed upon mechanism evaluation procedure; and
- to determine if there might be a standard data base that would be useful in distinguishing among different mechanisms.

A review of previous practice and a discussion of the underlying issues was presented in background documents "The Science of Photochemical Reaction Mechanism Development and Evaluation" by Jeffries and Arnold and "Need for Chemical Mechanism Documentation" by Sexton and Jeffries, which were distributed to workshop participants prior to the meeting.

Four scientists and an EPA user of models were asked to respond to the background documents and to offer their viewpoints on evaluation procedures and testing databases. These were Kenneth Demerjian, Roger Atkinson, Micheal Gery, Allan Dunker, and Joseph Tikvart. These four scientists were in agreement that there is, and has been, a generally accepted procedure for testing the extent of "reasonable agreement" between model predictions and experimental measurements. This procedure involves the use of laboratory kinetic, mechanistic, and product data, the use of smog chamber data, and the use of other test data, such as captive air irradiation measurements and ambient air measurements. Many participants believed, however, that the latter type of comparisons (*e.g.*, model predictions compared to ambient measurements and, for some, even captive air studies) require so many approximations and suffer from such instrumental limitations that the extent of agreement expected would be quite limited and thus such efforts would not be clear tests of our understanding of the chemical transformation processes. All four scientists agreed that, although there certainly were problems with their data, environmental or smog chambers still provided the most unambiguous data for the testing of urban chemical transformation mechanisms. Subsequent discussion confirmed that this approach or method was generally the accepted approach used by the workshop attendees.

The EPA model user (J. Tikvart) strongly supported the draft proposal for mechanism documentation. Other workshop attendees, including William Carter, Fred Lurmann, Gary Whitten, and Gregory McRae described their current practice and recent model testing strategies and results. While various mechanisms have been developed and tested against limited numbers of smog chamber experiments, only two mechanisms, the SAPRC/ERT mechanism and the latest Carbon Bond Mechanism, have been tested against a large number of chamber experiments (ca. 500) from different chambers.

Both of these well-tested mechanisms agree with the large body of experimental environmental chamber data to within about $\pm 30\%$ for ozone maxima and show varying levels of "reasonable agreement" for other measurements (for example, see Figure 1, 2a, and 2b). Discussion revealed that there is still some disagreement among the modeling community over how best to represent the chamber radical source(s), how best to represent the photolytic rates in outdoor chambers, and whether some of the species measurement techniques at the different chambers perform satisfactorily and how these measurements intercompare among different chamber groups. In addition, the unknown reaction pathways after hydroxyl attack on the aromatic species are represented differently in the two mechanisms.

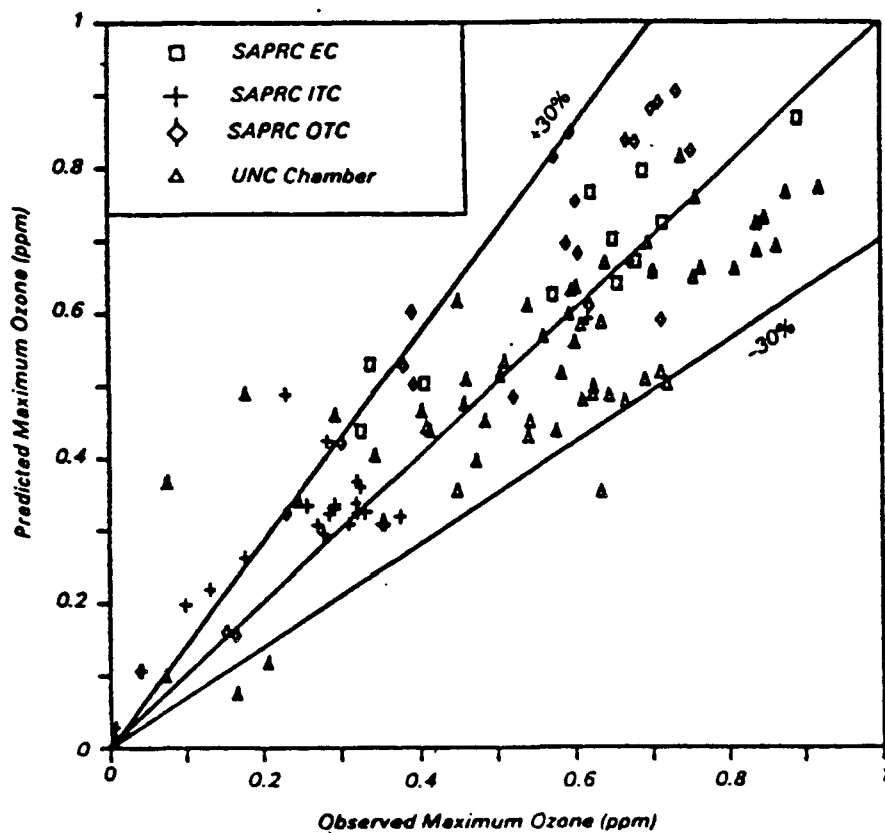


Figure 1. Predicted and observed maximum ozone concentrations for irradiations of complex organic mixtures and NO_x , carried out in the SAPRC evacuable (EC), indoor Teflon (ITC) and outdoor Teflon (OTC), and in the UNC outdoor dual chamber (UNC). From Carter *et al.*, EPA/600/3-86/031, 1986.

Introduction

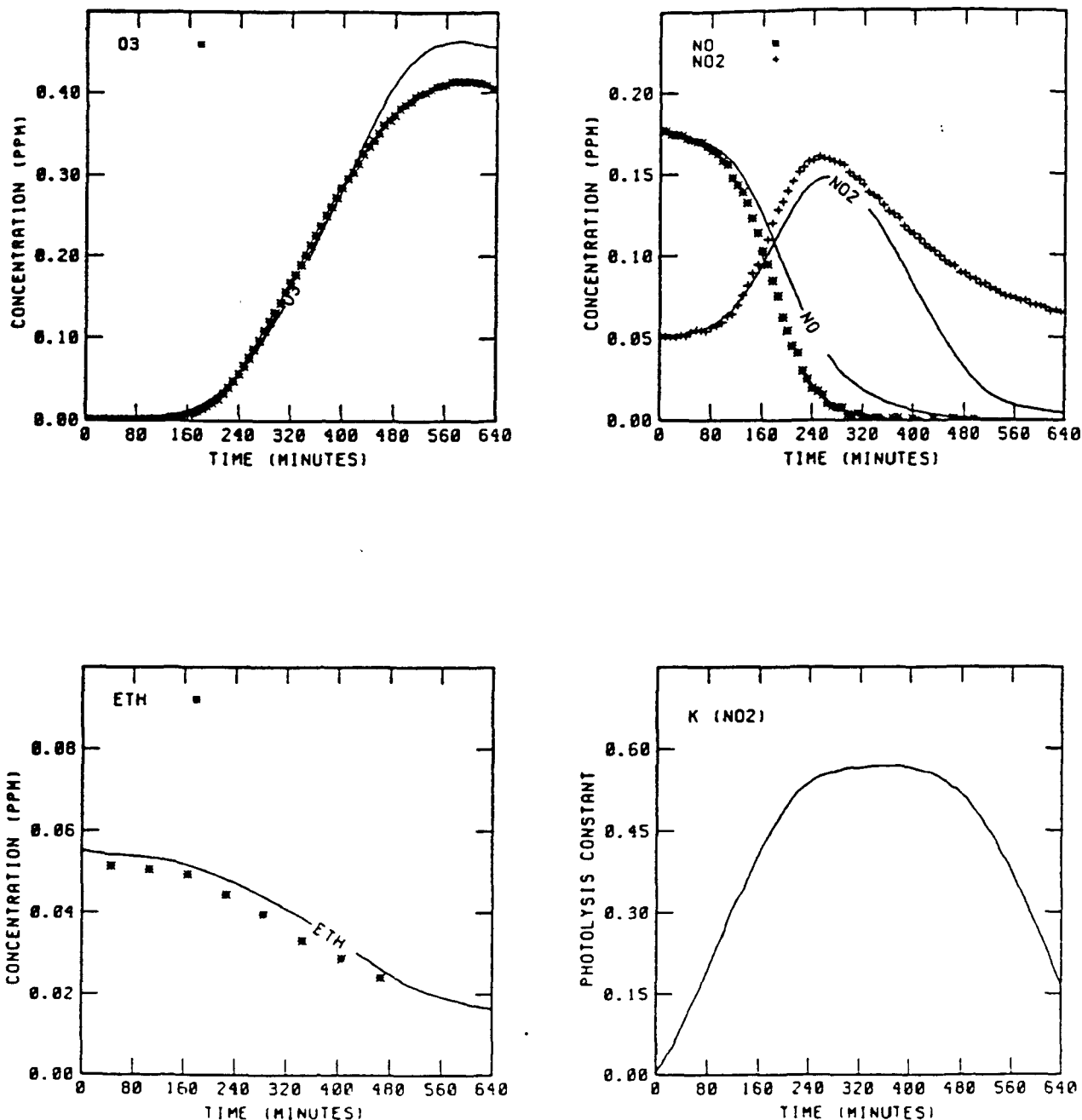


Figure 2a. Example of SAI model performance. The lines are Carbon Bond Four predictions and the symbols are experimental measurements from one side of the UNC outdoor chamber for the Sept. 20, 1981 experiment using a 14-component synthetic exhaust mixture.

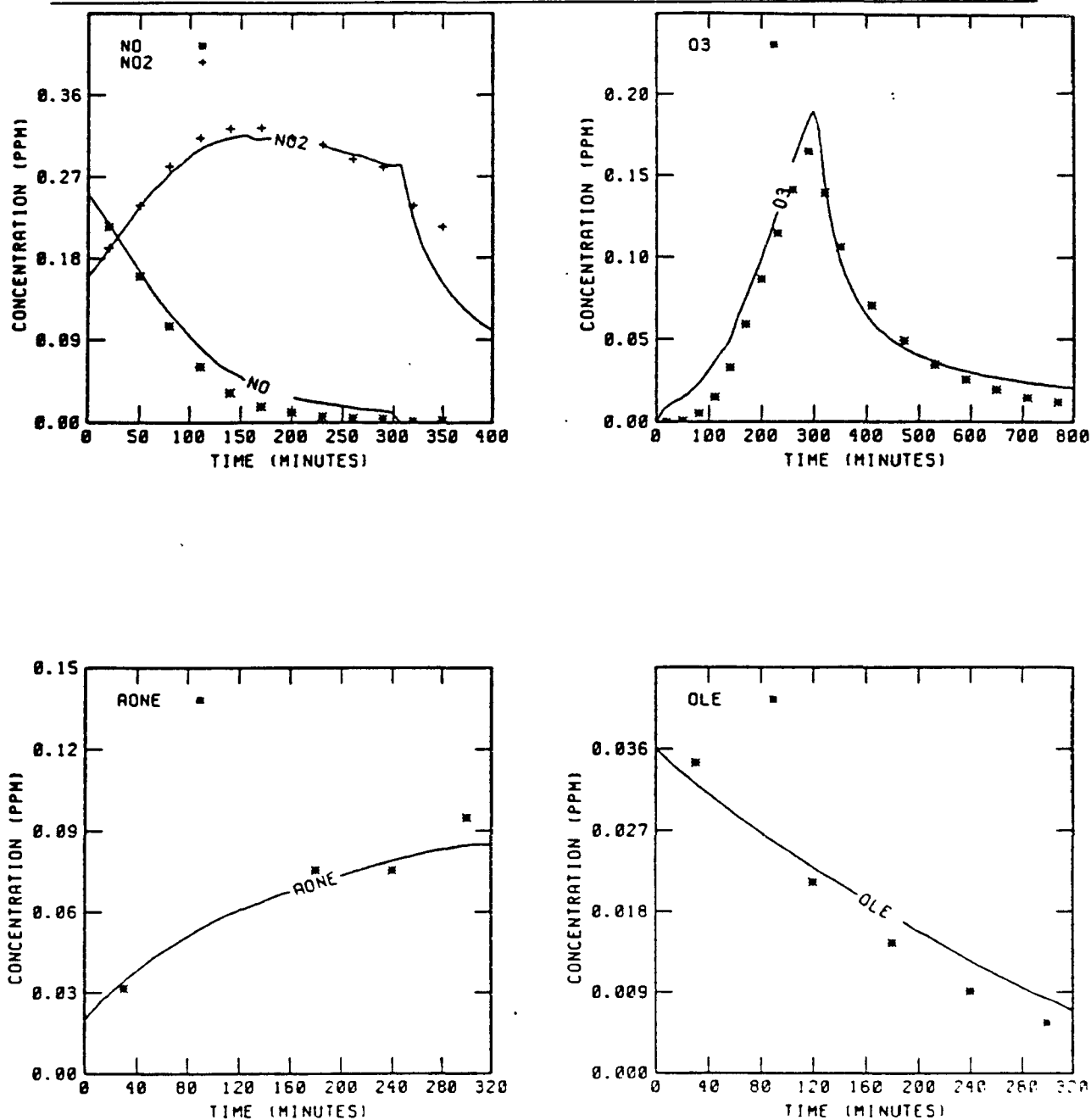


Figure 2b. Example of SAI model performance. The lines are Carbon Bond Four predictions and the symbols are experimental measurements from one side of the UCR Outdoor Teflon Chamber for August 3, 1983 experiment using a 9-component synthetic exhaust mixture.

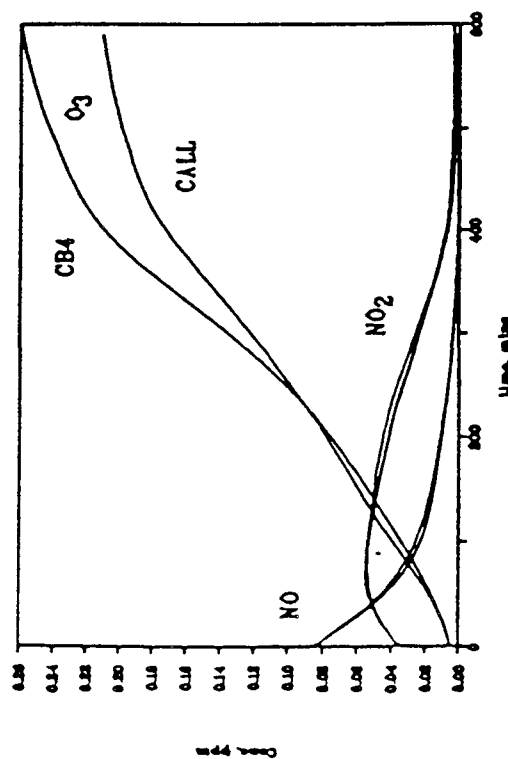
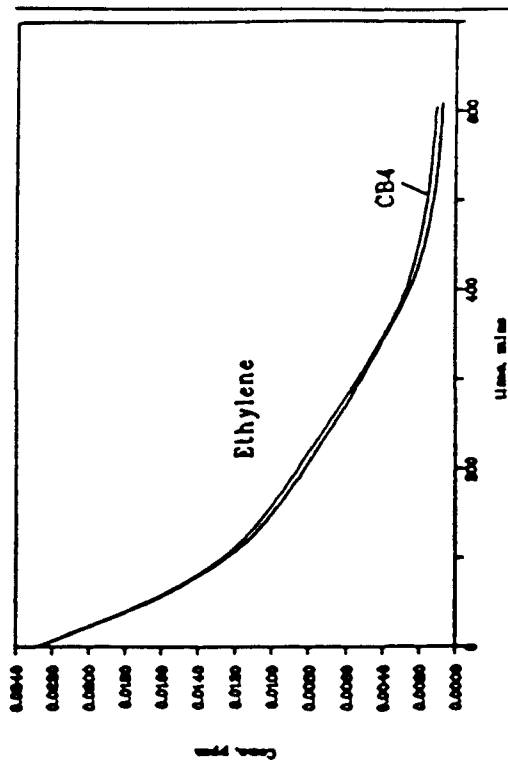
While many of the species predictions (*e.g.*, nitric oxide (NO), nitrogen dioxide (NO₂), ethylene, toluene, nitric acid (HNO₃)) in urban air quality applications agree remarkably well between the two mechanisms, ozone predictions can still disagree by 20 to 30% (see Figure 3), leading to differences in hydrocarbon control predictions. The Steering Committee concluded that the different methods of dealing with the uncertainties described above are the most likely cause of the difference in the mechanism predictions.

The Steering Committee concluded that, without further work, it was not possible to choose one of these mechanisms over the other on the basis of scientific evidence, and that the EPA should be encouraged to use both mechanisms as a method to estimate the present uncertainties in control requirement predictions. Furthermore, similar situations are likely to occur whenever two or more mechanisms are developed over similar time periods.

The Steering Committee also concluded that several review groups should be assembled to review the kinetic and chamber data bases and to review the extent of agreement among the models and these data bases. These recommendations will be described below.

It is clear from the data presented at this workshop, together with work published over the past five years, that a vast amount of progress has been made during the past decade, both with respect to urban-area chemical mechanism development and the data base upon which these mechanisms are based and tested.

Because of these advances, chemical mechanisms are now being used for applications other than urban oxidant control, for example, rural oxidant, long range transport and deposition, and the impact of toxic chemical emissions on the ecosystem. These present and future societal needs make it necessary that the development and testing of even more complex chemical mechanisms continue, and that present chemical mechanisms be updated and/or superseded by mechanisms able to deal accurately with the formation of gas phase precursors to acid deposition and to include or be interfaced to aqueous phase and cloud water chemical mechanisms. It is, however, clear that, despite the great advances made over the past few years, new challenges remain which require a wider fundamental chemistry data base coupled with more detailed and accurate environmental chamber measurements of a wide range of difficult-to-measure species than is presently customary. An integrated experimental and theoretical approach, which is closely coordinated with other proposed or upcoming programs such as the NSF/NASA "Global Tropospheric Chemistry" project is needed to address these needs.



8

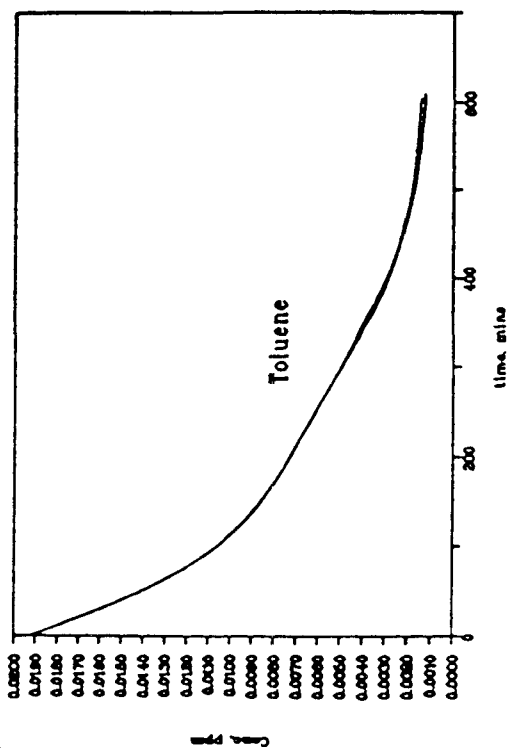
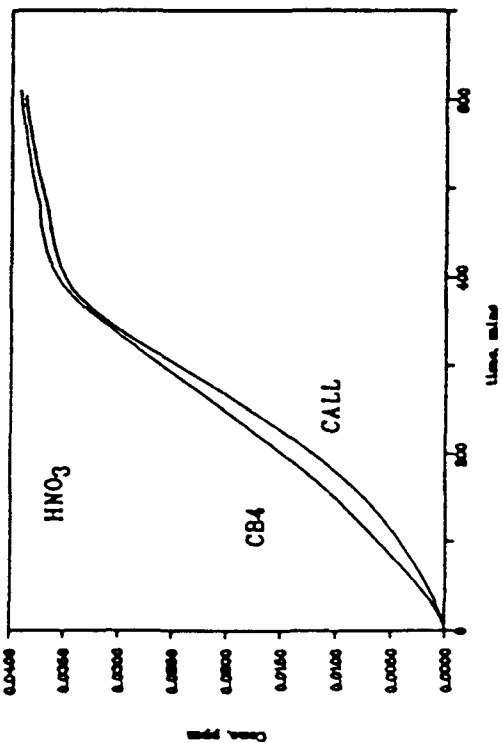
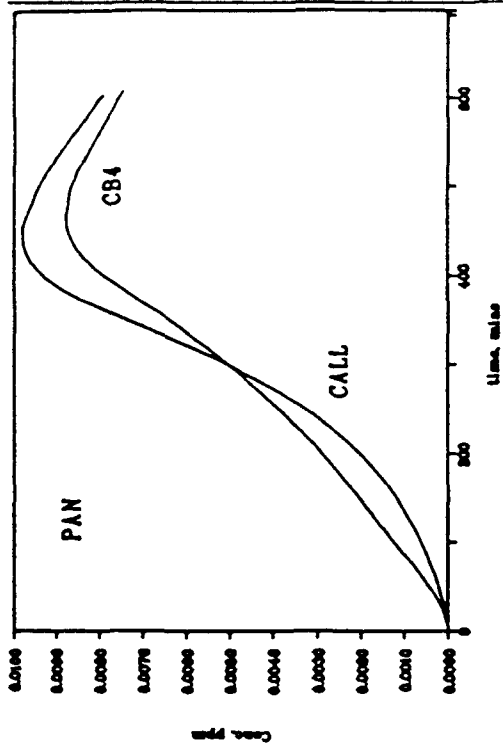


Figure 3a. Selected concentration-time profiles predicted by SAPRC/ERT and Carbon Bond Four Mechanisms used in a simple trajectory model. Starting time was 0800 LDT. Hydrocarbon and NO_x emissions were highest at start and decreased over five hours; initial mixing height was 250 m and final mixing height was 1235 m at 1600 LDT.



6

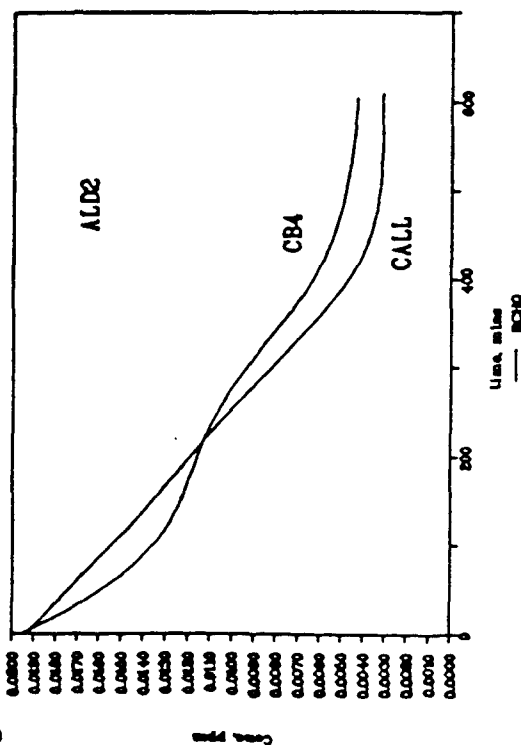


Figure 3b. Selected concentration-time profiles predicted by SAPRC/ERT and Carbon Bond Four Mechanisms used in a simple trajectory model. Starting time was 0800 LDT. Hydrocarbon and NO_x emissions were highest at start and decreased over five hours; initial mixing height was 250 m and final mixing height was 1235 m at 1600 LDT.

Recommendations for future chemical mechanism development, documentation, and testing intended to extend the present urban-oxidant-only mechanisms to the new areas that will confront EPA in the coming years are also given below.

Guidelines for Mechanism Development and Testing

General Approach

Demerjian described what most workshop participants accepted as a general approach to mechanism development and testing. The major components of this approach are shown in Figure 4. Within this generally accepted approach, however, there are differences in the practice among different modeling groups.

The mechanism development process begins with the assembly of reactions and rate parameters which are obtained from a consideration of chemical and thermodynamic theory, the chemical literature, recommendations and evaluations for the simpler species, and from consideration of the available literature data for the more complex portions of the overall mechanism. This step is subject to several limitations:

- not all rate data or mechanistic pathways of interest are available in the literature, and
- laboratory kinetic measurements must often be performed at conditions significantly different from those in the atmosphere, thus frequently requiring extrapolation and introducing uncertainty, and
- important competing processes may not have been present in the laboratory kinetics measurement system leading to either unintentionally ignoring important reactions or, where this problem is recognized, requiring the use of estimation methods to account for their absence.

Therefore, throughout the mechanism development process some form of testing is used to determine the importance of inclusion or exclusion of a given reaction and to assess the sensitivity of the mechanism's predictions to the choice of rate parameters which have some uncertainty either due to the use of estimation techniques or to experimental measurement uncertainty. In addition, some form of testing for completeness of description must be performed.

For most model developers, these tests have included the simulation of smog chamber experiments. Dunker stated that, "No mechanism developed entirely from theory and the literature is satisfactory without testing with chamber or captive air bag data." Such data, however, also have problems with regard to model testing. Some of these are:

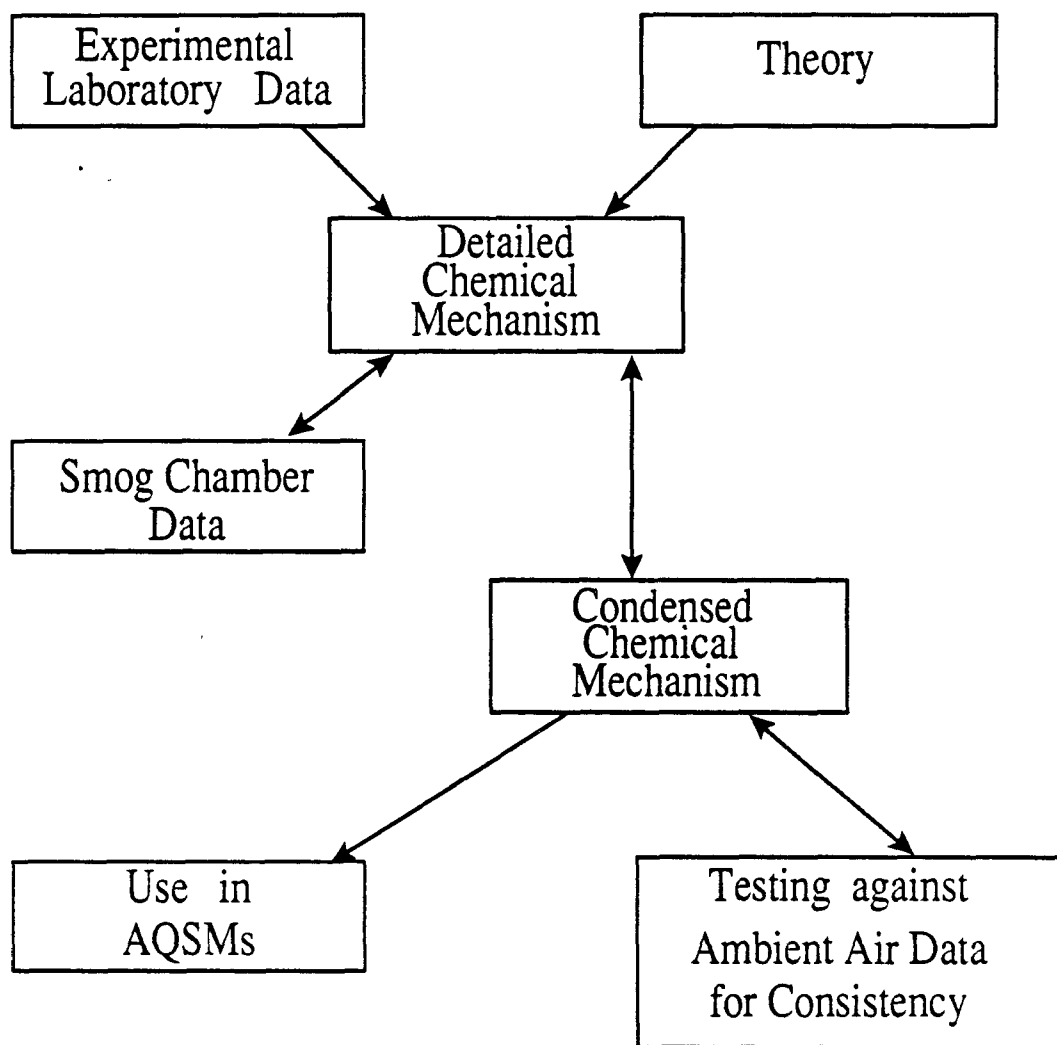


Figure 4. Schematic method for development and evaluation of chemical mechanisms. After Kenneth Demerjian.

- identifying and quantifying the chamber conditions that influence the chemical transformation processes, and
- coping with the sparsity and quality of measurements of reactant species and reaction products.

In the first category are the magnitude and spectral distribution of the photolytic rates in the chamber, the temperature, humidity, and mixing characteristics of the chamber, the magnitude and production mechanisms of chamber radical sources, and chamber walls as sinks for intermediate and final products (*e.g.*, HNO_3 , H_2O_2). In the second category are quality control problems, lack of analytical techniques that are both highly sensitive and specific for important chemical species. Some of the most important species, such as HO^\cdot and HO_2^\cdot , for example, can not be measured in chambers. Also there are differences among the measurements for intermediate and very important species such as formaldehyde.

Both sets of problems, the incomplete kinetics data base problems and the chamber effects and measurements problems, limit the extent to which each source of knowledge is expected to agree with the the other. This expected extent of agreement ultimately causes problems. As Demerjian said, "Some modelers like to have the lines for the model predictions fit every point from the chamber, while others are pleased if they are just close."

Chemical mechanisms that have enough detail to describe the features of several sets of smog chamber data are too detailed to be used in air quality models. Therefore generalization, distortion, and deletion processes are applied to these more complex models to produce a more practically sized mechanism for use in the air quality models without sacrificing too much accuracy of prediction. This step is generally described as producing a "condensed chemical mechanism."

Finally, Demerjian expressed a need to "test against ambient air data for consistency." This step is fraught with problems and was seen by many at the workshop as being ambiguous in terms of decision making about the chemical mechanisms. That is, because so many other factors could be responsible for the disagreements between ambient air measurements and model predictions, that assignment of any part of the error to the chemical mechanism would be very difficult, if not impossible. But it was stated by Whitten that, "If the models and data disagree in the ambient air, then the modeler ought to be required to consider explanations that at least include problems with the chemical mechanism."

Mechanism Evaluation Procedures

This section reviews some specific procedures for mechanism development and evalu-

ation that were presented by Lurmann, Gery, Dunker, Carter, Whitten, and McRae. There was general agreement that

- mechanism testing should be performed according to a “hierarchy of species” as shown in Figure 5,
- data from at least two, and preferably more, chambers should be used in the testing,
- testing should include at least two phases:
 - ▷ testing and refining the representations of chamber dependent phenomena, and
 - ▷ testing and refining the complex organic species and mixtures of species,

and

- as much data as is available should be used in the testing, and at at minimum, 15 experiments should be used in the first phase and 50 experiments in the second phase for each chamber.

In the first testing phase, the so-called “core” mechanism or the “inorganic and carbonyl reaction set,” which includes the reactions for the lower hierarchy species (*i.e.*, the inorganic species, carbon monoxide (CO), formaldehyde (HCHO), acetaldehyde (CH₃CHO) and PAN) should be tested with the following types of experiments:

- NO_x-air,
- NO_x-CO-air,
- HCHO-air,
- NO_x-HCHO-air,
- NO_x-HCHO-CO-air,
- CH₃CHO-air,
- NO_x-CH₃CHO-air,
- NO_x-CH₃CHO-CO-air.

Because the core mechanism is generally based on widely accepted recommendations, such as the NASA and IUPAC evaluations and reviews, the primary purpose of this phase is not to test the adequacy of the chemistry. Rather, the purpose of this phase of testing is to test and refine the representations of chamber dependent phenomena such as solar radiation intensity and spectral distribution, and wall processes including ozone destruction, nitrous acid (HONO) production from NO₂, nitric acid (HNO₃) production from N₂O₅, and possible NO, NO₂, HCHO, HONO,

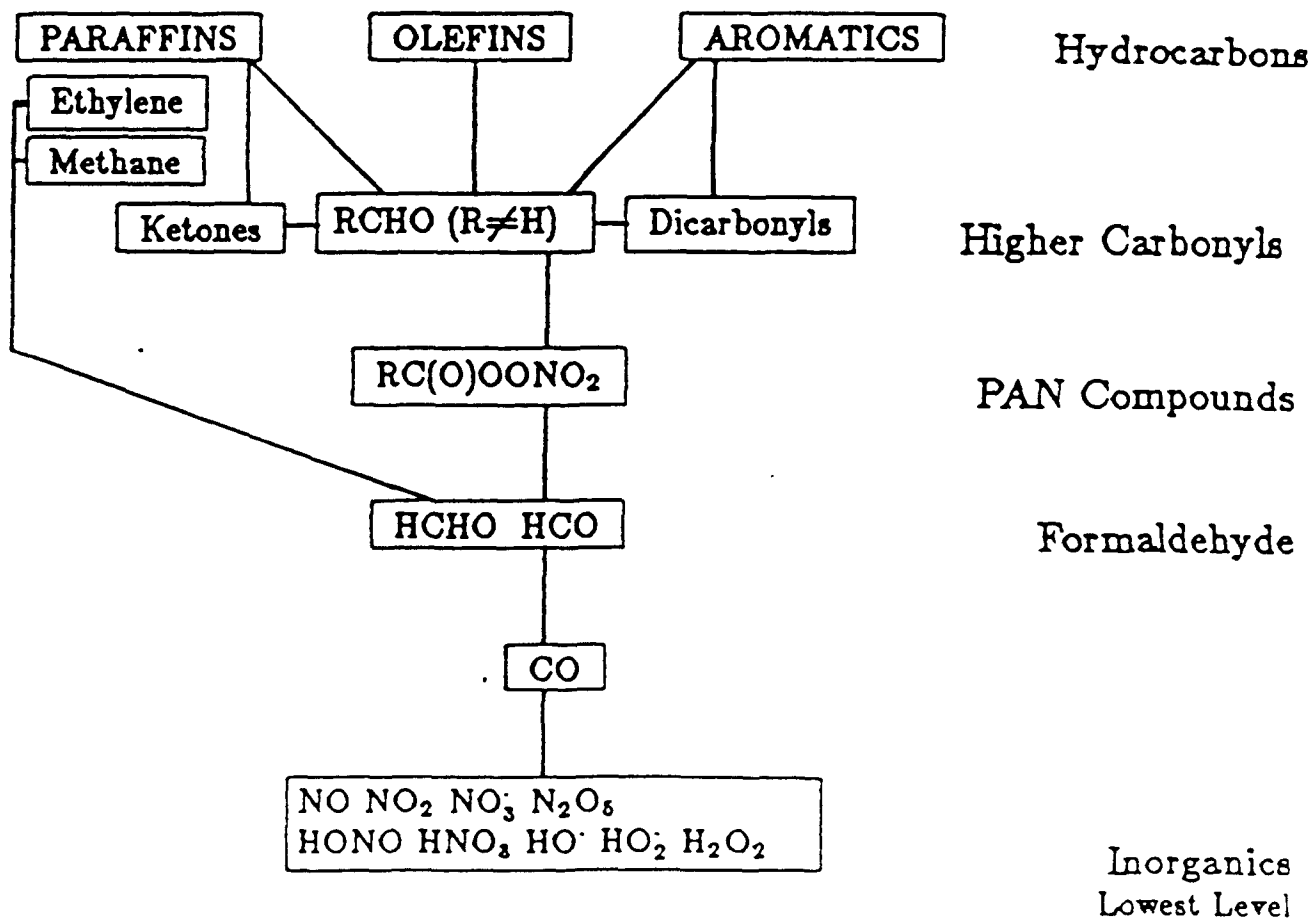


Figure 5. Hierarchy of Species

and HO[•] production from chamber-related processes. There is agreement that a uniform treatment of photolytic processes should be used. That is, the photolytic rates should be derived from species cross-sections and quantum yields and from the spectral distribution of the radiation sources in the different chambers. To the extent that the processes are understood, a uniform treatment should also be used for chamber sources and sinks of reactants. Data from each chamber should be used to determine the parameters for the descriptions of these processes. In some cases, knowledge of prior chamber use may be important in determining the stability of values of parameters.

In the second phase, the mechanism's performance against more complex hydrocarbon (HC)-NO_x-air experiments is tested. The mechanism's reactions for other carbonyls, alkanes, alkenes, and aromatic species are first evaluated independently by testing against experiments with single organic compounds, and parameterization of uncertain reaction rate constants and product yields is generally carried out using these single HC-NO_x experiments. The mechanism's performance for complex mixtures of organics with NO_x is then evaluated using known mixtures with 2 to 20 compounds. Additional evaluation using experiments with automobile exhaust and with urban air (*i.e.*, captive air experiments) are recommended because these mixtures are most representative of ambient hydrocarbon mixtures. For these real very complex mixtures, however, analysis frequently results in a 5 to 20%C unknown hydrocarbon fraction, and therefore tests using such experiments may yield ambiguous results depending upon the speciation of the unknown HC.

As indicated in the general discussion above, assessing goodness-of-fit between the experimental data and the model predictions is a topic that requires further work and one which the Steering Committee recommends the creation of a Task Group. Problems that arise in measuring model performance are:

- the large number of experiments and individual measurements, (*i.e.*, having to manipulate large quantities of information),
- both the time-to-events and the magnitude of results are important,
- the actual path taken to a point is important, (*i.e.*, the "shape of the curves"),
- some species concentrations depend upon the ratio of two other events rather than the absolute levels, thus seemingly correct predictions can arise for the wrong reasons,
- in some cases small errors in representing chamber characteristics or temporal changes in parameters can cause large errors in predictions at the very end of the simulation, but not for most of the experiment (*e.g.*, see Figure 2a in

which the final O_3 is over-predicted in the last 60 minutes, yet the model's predictions agreed well with all data up to that point),

- uncertainties in experimental measurements, especially the initial conditions, can lead to apparently poor predictions,
- lack of correspondence between the surrogate species in the mechanism and the experimentally measured species makes comparison difficult (i.e., PAR in CBM or chemiluminescent NO_2 measurements as in Figures 2 a and b),
- inaccuracies at very low concentrations and sample line losses or conversions make some comparisons difficult,
- for outdoor experiments, the sky condition or cloud cover condition, which may not be well represented in the model's photolytic rates, may be a major source of errors in the prediction.

As an example of one approach to the goodness-of-fit problem, the developers of the SAPRC/ERT mechanism used simple statistical measures to describe the goodness-of-fit for their model. These included:

- for the NO_x -air experiments, plots of observed and predicted total change in NO and NO_2 ,
- for experiments that produced O_3 , scatter diagrams of observed and predicted maximum O_3 , and an average rate of production, $d([O_3] - [NO])/dt$; similar plots were made for PAN and aldehyde maxima,
- frequency distributions of the absolute error in maximum O_3 , and
- tables of average error and bias.

Figure 1 is an example of this type of presentation.

As an example of another approach to the goodness-of-fit problem, the developers of the Carbon Bond Four mechanism used several hundred plots of the type shown in Figure 2a and 2b to illustrate the adequacy of their fit. Their reports also generally contain one or two tables giving the maximum O_3 and time to maximum for the model predictions and the observed values, but no scatter plots or error distribution plots are given.

Some members of the Steering Committee feel that both of these approaches to showing "goodness of fit" are incomplete. For example, in the SAPRC/ERT approach one can not tell if experiments that are over-predicted are also too fast, or if problems such as that shown in Figure 2a are responsible for the over-prediction (e.g., model close except at very end). One can not tell if mixtures that contain high concentrations of ethylene are always too low, and so forth. In the SAI approach,

for example, the typical reader simply can not comprehend the detail presented in so many graphs. No generalization of the information on the graphs is attempted in the SAI approach. Thus, the SAPRC/ERT approach tends to flatten all the dimensions of the problem to a few, while the SAI approach maintains all of the nearly overwhelming dimensionality and does little to provide an overview. Hence, some members of the Steering Committee assert that further work is required for demonstrating goodness-of-fit.

Development of Condensed Mechanisms

In general, the detailed chemical mechanism, developed and evaluated as discussed above, must be condensed for use in air quality simulation models for control strategy assessment purposes. It is strongly recommended that the condensed chemical mechanism be derived from the final tested version of the detailed mechanism with testing and comparison against the detailed mechanism at each step of the condensation process. Procedures for condensing mechanisms and testing the condensed mechanism against simulations of the expanded or explicit mechanisms were discussed in three reports: Whitten, Johnson, and Killus ("Development of a Chemical Kinetic Mechanism for the U.S. EPA Regional Oxidant Model," EPA-600/3-85-026, 1985); Whitten and Gery ("Development of CBM-X Mechanisms for Urban and Regional AQSMs," EPA-600/3-86-012, 1986); and Lurmann, Carter, and Coyner ("A Surrogate Species Chemical Mechanism for Urban-scale Air Quality Simulation Models, Adaptation of the Mechanism," EPA-600/3-86-031, 1986).

The condensing and testing procedures are intended to minimize compensating errors, minimize the number of species simulated, and yet retain the most important features of the original mechanism. A progressive approach is recommended that actually provides a series of condensed mechanisms, the final mechanism being the most condensed. Each of the mechanisms in the series would relate to the less condensed versions above it and thus could not be easily modified independently. Thus, changes or updates should be first introduced into the detailed or explicit (*i.e.*, the first) mechanism, and any adjustments or modifications should be propagated down through the various condensed mechanisms.

Specific methods for mechanism condensation include eliminations of species that are unreactive or do not change in concentration during the reactions and thus should not change the numerical results of simulations using the intermediate condensed mechanism when compared to the detailed mechanism. Species that satisfy the latter are M for pressure-dependent reactions and O₂ when the pressure or altitude are constant. These species can then be incorporated into the reaction parameters. For some applications, water may also be treated in this manner, but

frequently water varies in the experimental situation and must be treated explicitly. An example of unreactive species are product species which are of no interest and for which the mechanism contains no further reactions, for example H_2 , CO_2 , and formic acid.

Other steps include the use of fractional stoichiometric coefficients and intermediate species which have unimolecular branched pathways of reaction. As long as the unimolecular decay lifetimes are reasonably short, the numerical simulation results will not be significantly affected. Tests are required for lifetimes longer than a few minutes. Other techniques include the use of a "mass balance" or "counter species" which are given arbitrarily high decay rates to form "intentional" steady-state-like species. For example, several versions of the Carbon Bond chemistry employ a mass balance species usually called X, which rapidly removes the species PAR. Such species can be eliminated by the use of negative stoichiometric coefficients (*e.g.*, -1 PAR). Both mechanisms use a generalized RO_2 species to act as the sum of all alkylperoxy radicals in radical-radical termination steps, thus eliminating a large number of reactions while still approximately accounting for their effects with only two or three reactions per radical.

Condensation steps beyond those above involve assumptions about typical atmospheric situations, and the condensed versions of the mechanism must be tested to determine the bounds of such assumptions. Techniques have been developed to identify unimportant reactions and species. Elimination of unimportant reactions provides little benefit to simulation costs, because the costs are most sensitive to the number of species. Of course, fewer reactions do make a mechanism easier to understand.

Mechanism Documentation

It was an accepted fact among the Steering Committee that clear and comprehensive documentation of chemical mechanisms by their developers is needed to ensure their proper use. At this Workshop because of other pressing issues, little time was spent discussing what "documentation of chemical mechanisms" means. In his presentation, Joseph Tikvart reviewed and strongly supported the approach proposed in the workshop background document "Need for Chemical Mechanism Documentation," by Sexton and Jeffries. In this document an example outline of a guidance document for the application of a chemical mechanisms was proposed. In his writing for the Steering Committee after the workshop, Fred Lurmann also described several items that need to be documented in a mechanism. These are given in Table 1.

Table 1. Items Needed to Document a Mechanism.

1. A name, version number, and date.
2. A listing of the reactions and species.
3. A listing of all species assumed to be in steady-state, including the diagnostic equations, solution method for computing steady-state species concentrations, and discussion of the range of applicability (*i.e.*, temperature, pressure, concentration, *etc.*).
4. A listing of the rate constants and procedures to calculate their dependence on temperature, pressure, and actinic radiation.
5. A listing of absorption cross-sections and quantum yields as a function of wavelength including specifying averaging intervals and method of integration between points for all photolytic species.
6. The values of all photolytic reaction rates should be given at 10 degree zenith angle increments for a standard set of actinic fluxes such as those reported by Peterson in 1976 for the earth's surface with a standard estimate of albedo.
7. A complete listing of references to the sources of the kinetic and mechanistic data for each reaction and the actual form and value of the literature rate expression should be given (*i.e.*, $6.1 \times 10^{-13}(T/T_{ref})^{-23}$ cc-molecules⁻¹-sec⁻¹).
8. For use in computer codes that only accept Arrhenius expressions for rate constants, the non-standard form of rate constants (including fall off expressions) should be restated in Arrhenius rate form in the units typically used in the mechanism (*e.g.*, ppm⁻¹-min⁻¹) and the value at 760 mmHg and 298 K should be given.
9. The actual values of, or formulas for, all assumed concentrations of reactants, such as M, O₂, and H₂O that have been incorporated into rate constants should be given.
10. Clear identification of any differences in reactions and rate constants for modeling environmental chambers and modeling the ambient air.

continued on next page...

Table 1, cont. Items Needed to Document a Mechanism.

11. Complete documentation on the hydrocarbon representation system, including:
 - a) describing the basic approach,
 - b) describing the number of carbons in each lumped organic species,
 - c) rules for converting organic mixtures into mechanism input,
 - d) tables listing the assignment of commonly occurring compounds to lumped classes,
 - e) complete examples of the conversion of detailed organic mixtures to lumped mixtures.
12. An address where "official" versions of the mechanism can be obtained.

Implementation of a chemical mechanism by non-developers is often difficult not only because of deficiencies in documentation, but also because of software and computer differences. Computed solutions to mechanism test problems are essential to proper implementation of chemical mechanisms by non-developer users. These test problems should be designed to test all reactions in the mechanism. The Steering Committee believes that a minimum of four test problems are needed. These initial test problems should not involve dilution, entrainment, emissions injection, or deposition, but rather be examples of the pure chemical kinetics. Ideally, the solutions should be computed using a high quality algorithm such as the Gear algorithm with tight error control. In these test cases, the algorithm used, the method of computing the Jacobian, the use of absolute or relative error tolerances, and the minimum, initial, and maximum integrator step sizes should be documented. All species should be integrated rather than determined from the steady-state assumptions. The first two test problems should employ constant light intensity and temperature and one should be at a low HC-to-NO_x ratio while the other should be at a high ratio. The second two problems should employ diurnally varying light intensity, spectral distribution, and temperature, again with a low and high HC-to-NO_x ratio. The low HC-to-NO_x ratio test should be designed to test the nighttime chemistry also.

All inputs and methods used in the test problems should be fully documented. It is especially important to document the method(s) of updating the time-varying parameters. Frequent output intervals should be used and the output should also include the values of the time varying parameters.

A final documentation item discussed by the Steering Committee was the need for a standard set of conditions for comparing mechanism prediction of VOC control requirements when using the OZIPM program. It was suggested that the regulatory groups in EPA should be involved in developing these test cases to insure that they cover typical cases of interest for regulatory purposes.

Differences Among Well-Tested Mechanisms

When the guidelines for chemical mechanism development described in previous section are followed by multiple research groups, it is expected that large sections of the mechanisms developed will be very similar, if not identical. This will arise because of a) ongoing NASA and IUPAC evaluations for the inorganic and small ($\leq C_3$) organic species reactions (these have changed little in recent years), and b) because certain reactions and reaction sequences for the more complex organic compounds in the detailed mechanism are sufficiently supported by kinetic and mechanistic data that little uncertainty is associated with their magnitudes and structure. The portions of the detailed chemical mechanism, however, which are either unknown or only poorly understood will have to be assembled using estimations or arguments by analogy, or perhaps simply be parameterized. Thus, different methods of representing these unknown or poorly known sections of the chemistry will arise in different mechanism developments.

The latter process will most likely lead to detailed (and subsequently condensed) overall mechanisms which, based on past experience, may differ in their control strategy predictions, despite the fact that each chemical mechanism may be consistent within the bounds of reasonable agreement with the elementary reaction kinetics, mechanisms, and products databases and with environmental chamber data. Other than further efforts to refine the measure of uncertainty in the both the kinetics and chamber data and new formulations and testing of the mechanisms, there may well be no scientific reason to accept or reject one of these chemical mechanisms over the other.

Given that EPA must proceed with applications of the mechanisms and given that the existing data might not allow a further reduction of the uncertainties among the mechanisms offered as an explanations of the transformation processes, it is recommended that the EPA not select only a single mechanism for making control strategy predictions. Without further experimental data, such a choice would have to be made on other than a scientific basis. The use of even two mechanisms will suggest how the present uncertainties in the experimental databases translate into uncertainties in predicted control strategies. EPA must make the final decision of which strategy to use, and the uncertainty in the predicted control requirements along with other relevant societal values and issues should be considered in this decision.

While the need for regulatory consistency, and therefore a desire to have as long a time between changes in the tools used to make control decisions, is recognized,

our past experience has shown that sufficient changes occur in theoretical ideas and in experimental data in even a two year period as to result in potentially significant reductions in the uncertainties of control predictions. A failure to account for these changes in a timely manner could result in the view that the EPA is not using the best science. This problem may be exacerbated by the fact that the EPA may be sponsoring some of the "best" science being performed in this area. Therefore, the workshop participants recommend that EPA create a set of review processes, coordinated to the regulatory time line, that would use a base time period of two years for assessing the quality of control strategy predictions. A major scientific finding, such as a breakthrough in understanding aromatic reaction product mechanisms, may require a more rapid review and response. The exchange of detailed information in these review meetings is itself expected to accelerate the rate of improvement in the measure of agreement between models and data and therefore to improve the certainty for control strategy predictions.

Task or Review Groups Needed

The workshop participants recommended that several interacting sets of review processes or review groups be established, each with a goal of assessing and documenting the state of knowledge and degree of reasonable agreement to be expected in a given domain. At least four different domains were identified for such activities:

- 1) kinetic and mechanistic data needed in constructing photochemical transformation mechanisms;
- 2) environmental chamber data needed for comparison with mechanism predictions;
- 3) mechanism intercomparisons tests; and
- 4) user or application tests.

The first domain not only includes the traditional review process as exemplified by the Atkinson and Lloyd 1984 review, it also includes statements about what components of models are expected to look alike and where models might differ due to lack of convincing evidence.

The second domain includes all sources of information that influences judgments about the extent to which experimental chamber data contribute to the uncertainty in the agreements with models. This includes the type of information found in QA documents, such as calibration sources, histories, intercomparisons, and special tests. It includes agreement on what constitutes chamber characterization and would involve an assessment on the state of characterization of the chambers in current use and would include historical trends or adjustments needed to recover or use older data. It also includes assessment on the extent to which the various environmental chambers in use agree or disagree on various measures of consistency, *e.g.*, formaldehyde yields from ethylene experiments. An expected product of this group would be the identification and documentation in a uniform format of at least 100 high quality chamber experiments that could be used by the modeling community in general, or by EPA, to intercompare chemical mechanisms offered as explanations of the smog transformation process. It should be clearly recognized that such a data set would be seen as merely a necessary test, not as a sufficient test. That is, if mechanism predictions do not reasonably agree with these data, then the mechanism is not satisfactory for use in urban applications, but such agreement does not imply that a given mechanism would result in a particular set of control requirements in atmospheric applications. Furthermore, the data set presently available is not suitable for testing the predictions of the final acid

generation products such as hydrogen peroxide and nitric acid, although it can most likely test the production of precursors to these species, *e.g.*, formaldehyde.

The third domain encompasses those activities associated with merging information from the first two domains. It includes reaching agreements on the processes for describing how the authors actually constructed the particular mechanism, and how the choices in the areas exhibiting uncertainties were made. This domain includes clearly identifying the experiments used to develop the model (*i.e.*, those that may have resulted in changes in model content) and experiments used to test the model (*i.e.*, ones used to assess accuracy of the predictions). This domain also includes defining what constitutes fit between predictions and data, *e.g.*, how many species must be fitted, the extent of expected agreement expected given the uncertainties caused by uncertainties in rate constants and mechanism components and the uncertainties in the data as described by the assessment group in the second domain. Activities in this domain also include the identification and comparison of the parts of both explicit and compressed mechanisms which are different as the result of differing interpretations in the mechanism development process. Activities to design experiments, either laboratory kinetic or environmental chamber, which have some potential to discriminate among the representations are included in this area. Other efforts to identify additional constraints on the comparison of model and data are included also, for example, the use of dual chamber data to test two different regimes at the same time. Mechanism authors may be asked to provide documents that describe in detail why they believe that their own mechanism is superior to other mechanisms.

The fourth domain includes understanding prediction differences among the models, that is, understanding causes of the differences when there are no comparison empirical data. This domain includes tests of predictions using atmospheric data, the identification of critical points in each mechanism which result in predictions being different, the establishment of the sensitivity of the predictions to various input choices, and the sensitivity of the predictions to "correction" or scaling techniques as examples. This activity should also establish a series of test cases which all mechanisms for use by EPA must perform. Significant or meaningful measures of performance other than control requirements should be identified.

The Steering Committee suggests that EPA's role in these review activities should be one of support: providing funds to hold meetings, produce and disseminate documents. It should be further stressed, however, that each group is likely to make strong recommendations for needed work to resolve uncertainties further and to improve the science. In this case, EPA's role, especially in the area of environ-

Task or Review Groups Needed

mental chamber data production is seen as vital. No other governmental agency is supporting work in this area and EPA is seen as the major client for such data.

Short-Term Recommendations

Task Force on Chamber Light and Wall Process

The EPA has presently in hand two state-of-the-art chemical mechanisms for use in EKMA in the upcoming SIPs revisions. Each of these mechanisms has been tested against a large set of environmental chamber data, and the resulting mechanisms are very similar in many respects. There are several areas, however, in which, because of uncertainty, or lack of data, the model developers diverged in their mechanism construction. There are new experimental data that can help resolve some of the uncertainties and other differences are perhaps resolvable by direct intercomparison of the modelers techniques in a single setting. Thus, the Steering Committee believes that it may be possible to further refine these mechanisms on a short-time scale and without the need for any new experimental data. While this advancement needs no new experimental data, recommendations for further well-defined research may result from this effort.

The first priority is to use the presently available environmental chamber data base to refine the testing of present chemical mechanisms.

The Steering Committee therefore recommends the establishment of a task force of non-EPA scientists to evaluate and resolve chamber effects and light intensity/spectral distribution issues.

The first step will be to use new light intensity and spectral distribution data recently obtained at the UNC chamber to retrospectively better define this highly important chamber characteristic for the UNC chamber. In addition, similar spectral measurements should be made in the UCR Outdoor Teflon Chamber so that a consistent treatment can be applied to all outdoor chamber data. Likewise, an intercomparison of spectral measurements using the UNC spectroradiometer at the UCR chamber facilities would be useful in ensuring a uniform treatment of photolytic rates in all of these chambers.

After agreement has been reached on how to treat the photolytic rates in all of the chambers in use, the next step would be for the group of modelers and chamber operators to review and resolve the remaining divergent opinions concerning the best methods to represent chamber effects.

The Steering Committee believes that the best way to resolve these differences is to have a modeling session involving both modeling groups and both chamber groups at a single location. The group, for example, could meet at the University of

California, Riverside facility, where both of the recent mechanisms would be running on computer facilities that accessed the basic chamber characterization experimental data for all chambers. This would allow the various proposals to be immediately tested and discussed. In addition to the modelers and experimental scientists involved, a knowledgeable and impartial scientist in the atmospheric chemistry field must be appointed to act as moderator and ultimate decision maker.

This group must first concur on the appropriate photolytic rate and spectral distributions to use in all chambers, especially the UNC outdoor chamber. A re-modeling of the chamber characterization data bases must then be performed using *the different methods for treating all chamber's wall processes*. These tests must include:

- a) chamber control and characterization experiments, *e.g.*, NO_x-CO-air irradiations,
- b) NO_x-formaldehyde-air irradiations,
- c) NO_x-acetaldehyde-air irradiations,
- d) NO_x-simple organic compound-air irradiations,
- e) NO_x-multi-organic mixture-air irradiations in order of increasing complexity.

Where side-to-side test data exists, both sides must be modeled with consistent assumptions and simultaneously compared. In addition, recommendations will be made concerning a limited amount of experimental environmental chamber data which would lead to further significant decreases in uncertainties concerning chamber effects. These experimental data will probably involve monitoring of formaldehyde, using analytical techniques capable of monitoring formaldehyde in sub-parts-per-billion levels, during NO_x-air, NO_x-CO-air, and NO_x-propene-n-butane-air irradiations in the UNC and various UCR chambers.

If it is possible, all parties will agree on a single best representation of photolytic rates and chamber wall processes for all chambers, or they will agree and describe the evidence that supports different representations and will suggest experiments or measurements needed to support or refute the evidence offered. As a result of this re-analysis, the respective chemical mechanisms will almost certainly require some refinements and adjustments, possibly leading to a narrowing of the differences between the mechanisms with regard to their control strategy predictions.

The expected benefits from this task force include: a decrease in the uncertainty associated with the chamber data bases; an improvement in the representation of

chamber wall processes in both EPA-sponsored models; and a significant increase in the confidence of the predictions of both models.

The time required for this work is about one month of initial preparation and about one-week of meeting, followed by about two-weeks of analysis and documentation.

Task Force To Select Environmental Chamber Data for Testing

One of the goals of this workshop was to identify a standard data base that would be useful in distinguishing among different mechanisms. The Steering Committee concluded that it was possible to assemble a "necessary" environmental chamber data set, that is, one containing chamber experiments which all mechanisms for use in urban areas would have to simulate. The data base is described as "necessary" because if a mechanism could not successfully simulate these experiments it most likely would be considered unsatisfactory for EPA applications, but the data base would not be "sufficient" to resolve all questions associated with oxidant prediction. Mechanisms that could successfully simulate all experiments in the "necessary" data base might still give different predictions in urban atmospheric applications. Other recommendations given below address this situation.

The Steering Committee recognized that to assemble such a database would require significant input from chamber operators at both UNC and UCR and from model developers at SAI and UCR, as well as other interested parties who might wish to use the data. Furthermore, such a data base could not be assembled until the task force on chamber light and wall processes had finished its work.

It is envisioned that the data base would contain about 100 experiments from several chambers. The data would be maintained in a uniform format. Each experiment would contain detailed recommendations and supporting information on photolytic rates, chamber wall processes, and initial and temporal conditions. These would be the result of review, discussion, and consensus among task force members.

The benefits expected from this effort are a set of consensus data needed to assure a minimum level of performance among all mechanisms that may be potentially used by EPA or by state agencies in meeting EPA requirements.

The time required to accomplish this task is at least one year and would require funding for individuals at both chamber facilities and at the modeling groups, as well as for several working meetings of members of the task force group.

Longer Term Recommendations

Establishment of Review Groups

Review Group for Evaluation of Fundamental Kinetic and Mechanistic Data For Use in Model Development

It is evident that a critical necessity for present and future chemical mechanism development is that a continual effort be initiated to critically review and evaluate the kinetics, mechanisms, and products under atmospheric conditions for organic compounds. Because of the existence of the NASA and IUPAC evaluations panels dealing, respectively, with the stratospheric and tropospheric reactions of inorganic species and of the up to organics (methane, ethane, propane, ethene, propene, and acetylene and their degradation products in the presently ongoing IUPAC evaluation), this evaluation effort will not need to duplicate these present evaluations, but rather focus entirely on the atmospheric chemistry of the greater than C_3 organics. This ongoing evaluation, together with the NASA and IUPAC evaluations, will provide the "recommended" reaction list for use in detailed chemical mechanism development. This evaluation will identify the chemical reactions and reaction schemes which are unknown or only poorly understood, as carried out in the 1984 Atkinson and Lloyd review article, and will provide the impetus for needed laboratory experiments. It must be pointed out that this evaluation is extremely cost and time effective, avoiding the necessity for each chemical mechanism developer to independently review and evaluate the huge amount of literature available.

It is recommended that this evaluation effort be carried out in loose cooperation with the ongoing IUPAC evaluation. A 5-6 member review team is recommended, including at least one international expert in this area. To facilitate literature retrieval, close cooperation with the IUPAC panel with respect to publication must be carried out, and the entire evaluation must be readily available through a computer database system. Dr. Roger Atkinson, who is a present member of the IUPAC evaluation panel and senior author of the 1984 Atkinson and Lloyd review article, is a potential chairman of the EPA-funded panel. Publication and up-dates should be carried out at two-year intervals with the computer data base being updated more often for use by the tropospheric modeling community.

This effort could begin within one year and operate at an estimated yearly funding level of approximately \$100,000 to \$150,000, including publication and travel costs.

Review Group for Evaluation of Environmental Chamber Data

This standing review group would be an out-growth of the Task Force to select the standard or "necessary" data base that was described under "short-term recommendations" above. Its role would be to review and make recommendations with regard to the extent of "reasonable agreement" among various environmental chamber results and to make recommendations for additions or deletions from the standard data base. This effort could begin within one year and operate at an estimated yearly funding level of approximately \$50,000 - \$100,000 including publication and travel costs.

Review Group for Mechanism Intercomparison

This group would be the out-growth of the Task Force on chamber photolytic rates and wall effects that was described under "short-term recommendations" above. The full domain of this group was described above; its role would be to formally compare and resolve differences among different mechanisms in comparison with chamber data.

This effort could begin within one year and operate at an estimated yearly funding level similar to that of the Evaluation of Environmental Chamber review group above.

Review Group for Mechanism Predictions in Applications.

This group would provide scientific support for activities presently conducted by OAQPS staff. The full domain of this group was described above; its role would be to understand and explain differences in predictions from mechanisms in urban control strategy calculations.

Needs for Future Mechanism Development

Data Needs Concerning the Atmospheric Chemistry of Organics

Presently the further scientific refinement of chemical mechanisms is hindered by our lack of knowledge concerning several key areas of the chemical reactions occurring in the atmospheric degradation of certain classes of organics. While a detailed set of data needs will emerge from the data evaluation panel, it is abundantly clear that key areas of uncertainty involve the aromatic hydrocarbon chemistry, the chemistry of the larger ($\geq C_6$) alkanes, and aspects of the ozone-alkene reactions. Further laboratory product (and kinetic) data need to be obtained for these compounds before more accurate chemical mechanisms can be developed.

Specifically, the most important priorities are:

- a) to determine the fates of the HO \cdot -aromatic adducts under atmospheric conditions. This includes the determination of the fraction of these HO \cdot -aromatic adducts which result in the formation of phenolic compounds, and the elucidation of the ring-opened degradation products.
- b) to determine the reactions of the $>C_6$ alkoxy and alkylperoxy radicals under atmospheric conditions and the subsequent reactions of their products.
- c) to determine the radicals formed, and their yields, from ozone-alkene reactions under atmospheric conditions.

In addition, there is an urgent need for absorption cross-section and photodissociation quantum yields and product data for the carbonyl compounds formed as intermediate products in the degradation schemes of organics. Immediate candidate carbonyls are:

- ▷ methyl ethyl ketone
- ▷ propionaldehyde
- ▷ acetone
- ▷ glyoxal
- ▷ methylglyoxal

These laboratory studies must be coordinated with other ongoing or proposed efforts, for example the NSF/NASA "Global Tropospheric Program" and NSF Atmospheric Sciences Division projects, to avoid and unnecessary duplication of effort.

Needed Environmental Chamber Data

It is evident that for any future advances in the accuracy and breadth of a data base upon which detailed chemical mechanisms will be tested, new high quality environmental chamber data will be required. This need will arise from a number of reasons:

- a) The need to discriminate between present divergence approaches to representing chamber effects—i.e., HCHO and HONO offgassing verses heterogeneous formation of HO[•] radicals (probably via the intermediate formation of HONO). Because of the relatively slow rate of HCHO photodissociation, measurement of HCHO with sub-ppb sensitivity in the environmental chambers used will resolve this divergent opinion.
- b) The need for high quality environmental chamber data, together with correspondingly high quality chamber effects (offgassing rates, deposition rates, hydrolysis rates, photolysis intensity and spectral distribution, *etc.*) to provide more stringent tests of detailed chemical mechanisms.
- c) The need for data concerning the production of acidic species (HNO₃, H₂SO₄) and/or of their precursors (H₂O₂, HO₂[•], organic acids, organic hydroperoxides) for testing of acid deposition models. In addition, experimental data obtained under low NO_x conditions will also be needed, especially for more regional and tropospheric model testing.

Concurrent with these data collection efforts is the need for continued development and use of analytical techniques to detect and routinely monitor key labile species such as HONO, HCHO, H₂O₂, NO₂ (not as NO_x-NO), HNO₃, *etc.* Analytical measurement development must be viewed as an integral part of this effort.

Studies of Chamber-Dependent Effects

It is clear that at present a major uncertainty in the use of environmental chamber data for testing of chemical mechanisms concerns chamber-dependent effects. In particular, the role of chamber walls in the production of radicals during irradiations, either from contaminant offgassing or from the heterogeneous formation of such species as HONO from NO₂, with subsequent photolysis to yield radicals, is an area of much uncertainty. While for the short and median terms a more accurate parameterization of these effects and their magnitudes is required, on the longer time scale fundamental studies of these heterogeneous chamber wall effects are needed to elucidate the chemistry/physics of these processes and to investigate the optimum surface materials for use in future environmental chambers. These studies will require the involvement of scientists of other disciplines, for example,

surface chemists and physicists and experts in catalysts systems, as well as material scientists.

This area of research will be crucial to the development of future environmental chamber facilities for the production of high quality data for the testing and development of chemical mechanisms for, for example, acid deposition and long range transport applications where experimental data concerning hydrogen peroxide, organic hydroperoxides, organic acids, and nitric acid are required.

Investigation of the Applicability of the present EKMA Method

When the EKMA method is used with different chemical mechanisms, it has been common that each mechanism predicts a different control strategy. While it is possible that all of the difference in predicted control requirement are caused by differences in the chemical mechanisms themselves, it often appears that the control requirements are more sensitive than the predictions of absolute levels of O_3 , for example. It is possible that the "scaling" process the EKMA technique requires when the whole OZIPM model and its inputs do not correctly predict the observed O_3 might be responsible, at least in part, for some of the differences in predicted control requirements in this situation.

This might be so because in the present EKMA method all of the uncertainty in the model's fit to the world is placed in the absolute emissions level: it is the emissions that are adjusted to get a "fit" between the observed O_3 and the predicted O_3 . (Actually, in OZIPM, the emissions are expressed relative to the initial concentrations in the starting box, and the initial concentrations are increased or reduced to obtain the observed O_3 level. This is equivalent to adjusting the total HC mass in the simulation, and therefore the absolute emissions.) The applicability of this technique has not been demonstrated with modern mechanisms. The original support for this approach was based upon the now obsolete Dodge Mechanism, and essentially every rate constant or species mechanistic pathway has changed since that study was performed.

The initial and final mixing height and the rate of rise (the magnitude of the dilution rate) are probably the real factors that vary from day to day, and these are more likely to be the controlling factors in the real situation. It would not be any more difficult to program a computer search for the set of these conditions that give the observed O_3 than it is now to look for the correct combination of NMOC and NO_x . Then the mixing height parameters would be fixed and the NMOC and NO_x reductions could be explored in an absolute manner. Different mechanisms may

give much closer predictions of control requirements under this situation than they do under the standard EKMA approach.

In any event, the Steering Committee believes that the scientific foundations and assumptions in the present EKMA technique require further examination in conjunction with the adoption of the new mechanisms. This would be a modeling and sensitivity study and may require re-writing of the OZIPM code to implement new scaling methods.

Paper 1

The Science of Mechanism Development and Testing

**Harvey Jeffries
Department of Environmental Science and Engineering
University of North Carolina-Chapel Hill**

The Science of Mechanism Development and Testing

Harvey Jeffries

Department of Environmental Science
and Engineering

University of North Carolina
Chapel Hill, NC 27514

Purpose

This talk is about the practice of science related to photochemical kinetics mechanism development for urban atmospheres.

- ◊ One of its goals is to introduce a *meta-model of modeling*:
 - to describe the *process* of what modelers do to develop and test a mechanism as opposed to the *content* of the models themselves;
 - to examine the components of the process and to understand how these contribute to the advancement of knowledge about atmospheric chemistry;
- ◊ Another goal is to place before you questions that I believe highlight the issues we are gathered to discuss.
- ◊ A final goal is to suggest actions that I believe will maximize the benefits from all of our efforts.

Why Does EPA Need Science?

As described in the Government Accounting Office (GAO) report on EPA models, air quality (AQ) models are used as tools in administering the Clean Air Act.

EPA, therefore, wants models that *predict*, accurately and reliably, the effects of changes in emissions of chemical precursors on the magnitudes of secondary pollutant concentrations formed hours to days after the emissions and 10s to 1000s of kilometers away from sources.

Quantitative observations of such atmospheric transformation processes are nearly beyond our abilities, both from an experimental viewpoint and from a social viewpoint as well.

Because science, as a way to provide growth in knowledge, has been preeminently successful especially in the 20th century, government often turns to science for help in achieving its goals. For many involved in this enterprise science is seen as especially useful because they believe that “science makes possible knowledge of the world beyond its accessible, empirical manifestations.”

Thomas Kuhn has described the aim of science as “to invent *theories* that explain *observed phenomena* and that do so in terms of *real objects*, whatever the latter phrase many mean.”

If this aim of science were fully accomplished, then near-perfect prediction should be possible.

The EPA has supported scientific work with an aim to provide, if not near-perfect, then certainly acceptable prediction. Yet, according to some, the ‘scientific community’ apparently has not accomplished this goal in 15 years, for example:

- ▷ the GAO says EPA’s models are based on “assumptions, approximations, and judgments”;
- ▷ the GAO says that EPA’s most reliable models estimate—they do not even say predict—“actual pollution concentration within a range of minus 50% to plus 200%”;
- ▷ the GAO quotes un-named EPA officials as saying “more precise results are unlikely because of the limitations in the science”;
- ▷ the House Committee on Oversight wants a timetable for improvements.

Are public and Congressional expectations for science unreasonable? One of the things the public expects from science is the capacity to *explain* things and concurrently to enable at least scientists to *predict* what will happen in various combinations of circumstances. One normally explains a phenomenon by citing the *mechanism* that brought it about or sustains it, where the mechanism itself is understood by subsumption under causal, or at least time-dependent laws. When done ‘right’ in the minds of the public, the explainer not only derives the explanation correctly, but he also gets others to understand and *accept* his derivation. These actions, *explain*, *predict*, *accept*, are at the crux of the issues we are here to discuss.

Problems With Our Science.

I believe that our problems arise, in part, because of a lack of agreement. For example,

- ▷ there is no single, generally accepted chemical mechanism to describe the atmospheric transformation of precursors into the regulated and potentially regulated air pollutants.
- ▷ there is such dis-agreement and rapid change in the modeling community that many, both inside and outside, wonder about whether any of the solutions proposed are legitimate.

To many people this is a frustrating situation. After more than a decade and a half of air pollution modeling research, we still cannot predict ozone to many people's satisfaction.

The results of this situation are:

- ▷ The sense of frustration with conflicting and contradictory results has led some administrators at EPA to question the policy of reliance on atmospheric models for predictive use.
- ▷ Some have even doubted the necessity of further attempts to resolve the divisive modeling questions within the scientific community that created them. Throwing good money after bad, they seem to be saying, makes even less sense in science than it does in economics, and so EPA perhaps should not seek a solution to the problems arising from models by funding additional model research.

Users of models want to understand:

1. Why do we have conflicting and contradictory results in the applications of models?
2. Are some models wrong and other models right?
3. Are no models right?
4. Are smog chambers data wrong and theory right?
5. Are atmospheric data the ultimate judge of correctness?
6. What tools or strategies are available to help EPA, and the 'scientific community' of modelers themselves, make decisions about chemical mechanisms?
7. Given our situation, how should the (urban) atmospheric chemistry community *now* achieve progress? That is, how should we further actualize the

promise of science to explain atmospheric chemical transformations, and do so in a manner sufficient to meet EPA's needs?

A little reading of scientific history and philosophy of science will show that, in the broader sense, these questions are *not* specific to the problems of atmospheric chemical models or to EPA's needs for such models, but they address something of core importance in all sciences.

As explained in the background document, "The Science of Photochemical Reaction Mechanism Development and Evaluation," I have relied on the writing of an acclaimed philosopher of science, Thomas Kuhn, for many of the concepts I have used to investigate our situation. Kuhn's careful analysis of many scientific situations provides significant insight into the process of scientific advancement.

Further, he suggests that science *requires* a decision process that which permits rational people to disagree. He is also clear on how the ultimate judgments are made in science; we will discuss this process later.

EPA's Needs Verses Scientists' Needs

We must clearly realize, however, that the needs of EPA and the needs of the scientists do not necessarily go hand-in-hand.

From EPA's short-term perspective, the best outcome of this workshop might be the development of a "standard mechanism evaluation procedure(s)" and the definition of a "standard data base" needed for performing the evaluation.

Such an procedure may not be useful from the scientist's perspective.

- ▷ The scientist is concerned to understand the world and to extend the precision and scope with which it has been ordered.
- ▷ If intense scrutiny of this world reveals pockets of apparent disorder, these challenge him to a new refinement of his observational techniques or to a further articulation of his theories so as to extend the precision and scope of his understanding.
- ▷ Often the theories (and so the mechanisms) that arise in this process must be in complete contradiction to the ones that have gone before them. This is so because the new theory must be more successful precisely where the old theory failed. A new theory may necessitate new experimental evidence and a radically (so to speak) different mechanism.
- ▷ This has certain implications for the concept that there may be a single mechanism evaluation procedure and its concomitant database suitable for all present and future theories and mechanisms. Such a concept ignores the evolutionary character that has brought about the very success it seeks to measure.

But the difference in the need for rules goes beyond the evolutionary character of mechanism development.

Kuhn says that a lack of a standard interpretation or of an agreed reduction to rules will not prevent a 'scientific community' from conducting research that is acceptable to the community. This is true, however, only as long as the relevant scientific community accepts *without question* the particular problem-solutions already achieved.

Who Agrees With Whom About What?

I think it is not it not much of an exaggeration to suggest that every modeler disbelieves major aspects of every other modeler's model. And that some modelers

have serious doubts about all smog chamber data. And that many smog chamber experimenters think that the models fit the data poorly because the models are incomplete or just plain wrong.

Thus rules can become important wherever the underlying guiding principles of a science are successfully challenged or even where the community only perceives that the principles are weakening.

On the other hand, if we have arrived at the point where rules are now required to sort theories and mechanisms, then something fundamental in the science of urban air chemistry modeling is under challenge.

This challenge arises, I believe, from two situations:

- 1) as a result of the inability to resolve the differences among competing explanations of the chemistry of paraffins and olefins; and
- 2) as a result of the failure to offer acceptable explanations of the aromatics chemistry.

These are failures not only of the experimental arts, but also of theoretical practice.

How is Scientific Knowledge Obtained?

Kuhn says that scientists often work from patterns of behavior (mental constructs) “acquired through education and through subsequent exposure to the literature often without quite knowing or even needing to know what characteristics have given their methodologies and approaches acceptable status in the community that they practice in. If they have learned such abstractions at all, they show it mainly through their ability to do successful research.”

Kuhn further states, “Traditional discussion of scientific method have sought a set of rules that would permit any *individual* who followed them to produce sound knowledge. I have tried to insist, instead, that, though science is practiced by individuals, scientific knowledge is intrinsically a *group* product and that neither its peculiar efficacy nor the manner in which it develops will be understood without reference to the special nature of the groups that produce it.”

- A. *Normal science and revolution.* Once a specific science has been individuated at all, it characteristically passes through a sequence of *normal science—crisis—revolution—new normal science*. Normal science is chiefly puzzle-solving activity, in which research workers try both to extend successful techniques, and to remove problems that exist in some established body of knowledge. Normal science is conservative, and its researchers are praised for doing more the same, better. But from time to time anomalies in some branch of knowledge get out of hand, and there seems no way to cope with them. This is a crisis. Only a complete rethinking of the material will suffice, and this produces revolution.
- B. *Paradigms.* A normal science is characterized by a ‘paradigm’ [(PAIR-uh-dime), a pattern]. There is the *paradigm-as-achievement*. This is the accepted way of solving a problem which then serves as a model for future workers. Then there is the *paradigm-as-set-of-shared-values*. This means the methods, standards, and generalizations shared by those trained to carry on the work that models itself on the paradigm-as-achievement. The social unit that transmits both kinds of paradigm may be a small group of perhaps one hundred or so scientists who write or telephone each other, compose the textbooks, referee papers, and above all discriminate among problems that are posed for solution.

The term “paradigm” enters in close proximity, both physical and logical, to the phrase “scientific community.” A paradigm is what the members of a scientific community, and they alone, share. Conversely, it is their possession of a common

paradigm that constitutes a scientific community of a group of otherwise disparate men.

One thing that binds the members of any scientific community together and simultaneously differentiates them from the members of other apparently similar groups is their possession of a common language or special dialect. Kuhn has suggested that in learning such a language, as they must to participate in their community's work, new members acquire a set of cognitive commitments that are not, in principle, fully analyzable within that language itself. Such commitments are a consequence of the ways in which the terms, phrases, and sentences of the language are applied to nature, and it is its relevance to the language-nature link that makes the original narrower sense of "paradigm" so important.

The paradigm-as-shared-values is a source of commitments for the scientists that share it. These include:

1. Explicit statements of scientific law and about scientific concepts and theories. While these continue to be honored, such statements help to set puzzles and to limit acceptable solutions.
2. A multitude of commitments to preferred types of instrumentation and to the ways in which accepted instruments may legitimately be employed.
3. Quasi-metaphysical commitments that tell scientists what sorts of entities the universe does and does not contain, what ultimate laws and fundamental explanations must be like, and what research problems should be.
4. A commitment to be concerned to understand the world and to extend the precision and scope with which it has been ordered, therefore, says Kuhn, "to scrutinize, either for himself or through colleagues, some aspect of nature in great empirical detail. And if that scrutiny displays pockets of apparent disorder, then these must challenge him to new refinements of his observational techniques or to a further articulation of his theories." Though his concern with nature may be global in extent, the problems on which he works must be problems of detail.
5. More important, the solutions that satisfy him may not be merely personal but must instead be accepted as solutions by a group. This group may not, however, be drawn at random from society as a whole, but is rather the well-defined community of the scientist's professional compeers.
6. One of the strongest, if still unwritten, rules of scientific life is the prohibition of appeals to heads of state or to the populace at large in matters scientific.

Recognition of the existence of a uniquely competent professional group and acceptance of its role as the exclusive arbiter of professional achievements has further implications.

The group's members, as individuals and by virtue of their shared training and experience, *must be seen as the sole possessors of the rules of the game or of some equivalent basis for unequivocal judgments*. To doubt that they shared some such basis for evaluations would be to admit the existence of incompatible standards of scientific achievement. This admission would inevitably raise the question whether truth in the sciences can be one.

Kuhn says, "Normal science is a highly determined activity, but it need not be entirely determined by rules." There are rules for moving the chess pieces; there are few rules for playing a game of chess.

For Kuhn, "Normal science consists in the actualization of that promise [of success offered by the paradigm], an actualization achieved by extending the knowledge of those facts that the paradigm displays as particularly revealing, by increasing the extent of match between those facts and the paradigm's predictions, and by further articulation of the paradigm itself." These are the three areas into which Kuhn says all journal literature of *normal* science can be classified.

The business of science, Kuhn says, is puzzle-solving. Though this may sound demeaning, he certainly intends no offense; he says that this process has required the highest scientific talents. In fact, he choose the metaphor because of the striking similarity between solving jigsaw puzzles and the search for explanations of phenomena and the ways to harness nature.

Kuhn contends that this puzzle-solving activity is a vital part of science so long as the puzzles are interesting and the solutions are accepted by the community. In their day-to-day work, scientists do not try to disprove the theories that direct their research, nor do they seek unexpected and unpredicted results.

To understand what is meant by puzzle solving we must examine a schema for scientific practice:

$$\text{Theory} + \frac{\text{Auxiliary}}{\text{Statements}} = \text{Prediction} \approx \frac{\text{Facts to}}{\text{be explained}}$$

In the daily practice of normal science, a scientists seeks statements of his best guesses (auxiliary statements) about the proper way to connect his own research problem with the corpus of accepted scientific knowledge (Theory).

For example, he may conjecture that a newly discovered product in a chamber experiment is to be understood as an effect of HO[·]-attack on a particular hydrocarbon. The next steps in his research are intended to try out the conjecture or hypothesis. If it passes enough or stringent enough tests, the scientist has made a discovery or has at least resolved the puzzle he had been set. If not, he must either abandon the puzzle entirely or attempt to solve it with the aid of some other hypothesis.

The practitioner must often test the conjectural puzzle solution that his ingenuity suggests. But only his personal conjecture is tested. If it fails the test, only his own ability, not the corpus of current science is impugned. As Kuhn says, "in the final analysis it is the individual scientist rather than current theory which is tested."

Tests of this sort are a standard component of what Kuhn called normal science. In no usual sense, however, are such tests directed to current theory. The scientist does not immediately look for a new oxidizing theory, on the contrary, the scientist must *premise* current theory as the rules of his game. His object is to solve a puzzle, preferably one at which others have failed, and current theory is required to define that puzzle and to guarantee that, given sufficient brilliance, it can be solved.

As normal science advances, the promise of the paradigm is actualized and the paradigm itself is refined, reformulated, and further articulated. This enterprise seems to the outsider (one who does not share the commitments described above) to be an attempt to force nature into the preformed and relatively inflexible box that the paradigm supplies.

Theories are accepted because they have real *explanatory* successes. Although a theory may legitimately be preserved by changes in the auxiliary statements which are, in a sense, *ad hoc* although not unreasonable, its successes must not be *ad hoc*. That we have no new paradigm 15 years after the proposition of HO[·]-attack does not mean that air pollution science has failed to progress in a Kuhnian sense. Rather, the puzzle-solving done in those 15 years to expand and ratify, or modify the theories and auxiliary statements of HO[·]-attack has been the success promised by that paradigm. And it *was* that promise, Kuhn would say and I would agree after looking at the actual documentation of those early years, that led us to accept the HO[·]-attack theory in the first place.

The areas investigated by normal science are minuscule; the enterprise has drastically restricted vision. Kuhn says, "But those restrictions, born from confidence

in a paradigm, turn out to be essential to the development of science. By focusing attention upon a small range of relatively esoteric problems, the paradigm forces scientists to investigate some part of nature in a detail and depth that would otherwise be unimaginable."

Thus the juxtaposition of our theoretical constructs and the observations associated with them leads to an extension of the precision and scope of the ordering of the world.

This progress of the paradigm, the practice of normal science—by its very nature (the finer and finer attention to detail)—encounters anomalies that require distortion to include in the paradigm, or deletion to ignore as important to the paradigm.

The hypotheses of individuals are tested, the commitments shared by his group being presupposed. Group commitments, on the other hand, are not tested, and the process by which they are displaced differs drastically from that involved in the evaluation of hypotheses. This is the crisis situation, in which the anomalies created by the progress of normal science so strain the old theory that many seek revolution, a turning away from the original paradigm to a new theory, a new way of looking at the world, that again offers real explanatory success and is full of promise for the future. We have not had such a revolution in the still young field of atmospheric chemistry: the HO[•]-chain theory still holds much promise and, in our daily scientific activities, we still seek solutions to puzzles *assuming* that an HO[•]-chain process is central to atmospheric chemistry.

Our Situation: Actions Needed.

We currently find many anomalies hidden in the region of “reasonable agreement” between mechanism predictions and chamber data. And because modelers have a real need to succeed in solving the puzzle set by themselves, solutions to puzzles have sometimes been claimed with little vindication. Further, the community has apparently accepted these solutions at the time, only to find themselves in a position of denying the solution a year later.

This makes us appear somewhat less than legitimate or not very serious about how much we really do understand about urban atmospheric chemistry.

Therefore, I believe we must work as a community to re-establish explicit recognition of a set of rules. These rules, in the words of Kuhn “limit both the nature of acceptable solution and the steps by which they are to be obtained.”

If we take ‘limit’ not to mean restrict, but rather to define, then these rules are not unnecessary burdens on science. Also the interpretation of limit as “define” explains why rules can lie unexpressed below the surface of regular scientific activity: definitions, as we have seen, must often be taken for granted if they are to function at all.

These definitional rules do not define what the state of the science, or even what the paradigm, *is*. But rather, they define whether or not solutions fall within the paradigm, require modification or distortion of the paradigm, or fall completely outside the paradigm. Such rules can change in direct response to the input of puzzle- solutions, but clearly they are not infinitely flexible.

Frequently in times of crisis, more solutions are ruled out of the paradigm than are included within it. When the rules function in this way to define the paradigm more and to prevent its unnecessary distortion, the rules can save a still-functional paradigm and work goes on.

When the rules have strengthened and protected the paradigm in these ways, they—now accepted by consensus in the community—recede again from explicit recognition and indirectly guide the community toward interesting puzzles in need of solution. *Thus the definitional rules can be changed in response to the increasingly refined solutions proposed within the paradigm, but they direct research best only when they are sufficiently accepted by the community to be implicit.*

For science to advance at all there must be reasonable agreement between three sets of evidence:

1. between theory and observation;
2. between different sets of observations; and
3. between explanation and prediction.

Without agreement, no detailed comparison could be made between the three sets of evidence: no puzzle-solutions could be evaluated. And the puzzles most meaningful in maintaining or challenging a paradigm are exactly those that lie on the interfaces of the three sets listed above.

Thus the first sense in which reasonable agreement determines scientific progress is in defining what sorts of evidence can be compared. That is to say, that by defining what qualifies as evidence of theories, observations, explanation, and predictions, the community establishes reasonable agreement for quantitative precision.

Once evidential standards of reasonable agreement have been recognized, the puzzles we have described as at the interfaces of theory and observation, between observations, and between explanation and prediction can be defined.

When solutions to the problems are offered, the second function of reasonable agreement is activated: agreement for better quantitative and qualitative precision to discriminate among the solutions. Again, this function is determined by consensus within the community and again as a standard it varies directly with the progression of the paradigm from potential to actualization.

Thus, while I see the need for a "mechanism evaluation procedure" I do not see it as static or as something EPA regulates.

I see the process as consisting of several groups of practicing scientists struggling with what constitutes "reasonable agreement" among the sets of evidence. I see this as requiring face-to-face meetings in which evidence is presented, exchanged, argued and agreed upon or questioned. Part of the current problem is that some of the most important evidence (*e.g.*, plots of model predictions and observational data and even mechanism listings) are sometimes not presented for judgment. Significant operational procedures are omitted for the sake of brevity. Some evidence is treated as 'proprietary.' Clearly, these meetings must be open to the 'scientific community' and must be supportive of those with differing opinions and with different sources of information.

Because the development of knowledge is not static, these assessments must be periodic. The products of these assessments should allow modelers to develop vindication statements for their models.

Finally, it should be our goal not to seek multiple representations of reality—for example, it is not the Carbon Bond Mechanism, nor should it be the Carter, Atkinson, Lloyd, Lurmann mechanism—instead, it should be

- ▷ ‘What is the best representation of what we know with a high degree of certainty?’, and
- ▷ ‘What is an adequate representation for what we probably know?’ and
- ▷ ‘How creative can we be about what we are guessing at?’

Discussion After Jeffries' Presentation

Jeffries: I have passed around some questions that are keyed to certain pages in the background document I wrote. I believe that these questions highlight some of the particular details and issues that we want to discuss. [These questions are included here as Table 1.] Dr. Atkinson's basic proposal made in his opening remarks was that if a model fits smog chamber data, then the mechanism should be satisfactory. Question 10 asks, "Is that really true?" I discussed several situations where mechanisms 'fit' smog chamber data very well, only to later be shown to have a major kinetic parameter wrong. Questions 11 and 12, I think define what this conference is really about. "How do we prevent mechanisms from being mere opinion?"

Dimitriades: Do you intend to provide answers to these questions?

Jeffries: These are questions to stimulate discussion. I do not know if there are "answers" to the questions. I certainly have provided discussion of them in the background document. For example, on page 42 of the document you will see that in the Hecht-Sienfeld-Dodge paper, it was proposed that validation consisted of two steps: 1) validate against chamber data, and 2) validate against atmospheric data. But they never explained why they wanted to validate against atmospheric data. So, what is the real need to validate against atmospheric data? What do you gain by doing that? How would you actually use this process?

Dimitriades: That's fine. I think the discussion can, perhaps, address these questions of yours.

Jeffries: I think it's premature, Basil, I think we should wait for some others to have a chance to reflect on some of the things that have been said.

Atkinson: Maybe we can go through the rest of the day and bring that up tomorrow morning.

Riordan: As I understand it, in the normal evolution of sciences, as stated by Kuhn, there is an adoption of a paradigm, then people apply the paradigm, and over time anomalies develop as people look at the detail in the world. To deal with the anomalies they jerryrigged this paradigm. Then all of a sudden these people realize that you cannot jerryrigged it anymore; then you get someone questioning the basic paradigm and you get crisis, and then revolution. Do you picture what

Table 1. Discussion Questions Submitted by Jeffries

1. Are “more precise results” in applying models unlikely “because of the limitations in the science” as was reported to have been by some EPA official in the GAO report? (page 1)
2. How should a scientist deal with the “delemma of the underdetermination of theory?” (page 32)
3. What are the conditions necessary to believe in a true correspondence between entities in a mechanism and the real world? (page 38–39)
4. How else, other than by the use of smog chambers, can we produce the detailed facts that must be compared unambiguously and directly with the model while preserving as much of the whole chemical transformation process as possible to allow for the evaluation of significant over-generalization and deletion? (page 40)
5. How, or even do, comparisons of model predictions with chamber data deal with the question of under-determination of theory? (page 40)
6. What is the real need to validate in the atmosphere? (page 42)
7. What defines “reasonable agreement?” (page 48)
8. Given the comment made by Whitten on page 76, how can we believe the prediction of the mechanism outside of the conditions on which its explanatory function was demonstrated?
9. Given the ability to *predict* a chamber radical source strength, but not the ability to *explain* the source, can mechanisms be adequately tested with chamber data? (page 81)
10. What does testing mean? What does failure to fit chamber data mean for a mechanism? What does success in fitting chamber data mean? (page 95)
11. How do we prevent mechanisms from being mere opinion? What constitutes “probative force to justify” a modeler’s claims?
12. What is required to make the vindication statement on page 97? Is this statement satisfactory for EPA’s needs?

has been proposed here as a process that is beneficial, as trying to accelerate that, or as being sort of complimentary, or as being contradictory to that. I guess what I am trying to figure out is, if you have this normal evolution of science and it occurs somewhat naturally, what are we doing here trying to talk about rules of evidence that would be appropriate in trying to discover anomalies?

Jeffries: Within the Kuhn structure of things, I think our present field is still in the middle of practicing normal science. We have not come across any of the kinds of crisis that Kuhn has described. In other words, none of us sitting around this room disbelieves that HO[•] exists or that the HO[•]-chain theory can account for observed atmospheric chemistry. We still, I think, fundamentally believe that there are solutions out there operating under our current conceptual framework. And so that's not what is at issue here at all. That is, we are not facing a Kuhnian crisis in the sense of a failure in our paradigm. What is unique about our particular situation, is that we're wrapped up in your regulatory issues. It is you who comes along and imposes a time table on the science. Given long enough time, our problems, the ones that I described as problems of the science, would go away, by themselves. We would just shake things out. So, if someone discovered something that was accepted by the community for a while, and then later we find out that it was not true, we would just change it. The problem is that you take this first result and embed it in law and use it. That's where I think the interface between the science and the regulatory issues cause us to have a problem. You are forced to use tools prematurely. Or, to put it another way, as I said in the write-up, we have 'forced actualization over promise.' We've been forced to achieve an actualization of the promise of the paradigm before its time, in a sense, without adequate foundation. That's where I think the real crux of the problem lies. And, we're forced to do that because 1) we have to succeed in solving the puzzles or we are seen as failures as scientists, and 2) we are forced to conclude that we always have reasonable agreement because we have to meet the terms of your contract or the grant.

Whitten: I have two questions to bring up. First of all I think that the discussion is not talking about engineering. I think that there are science kinds of problems and there is the need for a regulatory agency to have a model. What is needed is an engineer in the middle that uses the information in the scientific community and constructs something based on it. We need some sort of an engineering approach. There has always been a problem historically between engineers and scientists and this is a lot of what we have been talking about here. The second point is that I have problems with your use of the word paradigm in the context of the HO[•] mechanism. The paradigm that I see that is more involved with the production of ozone and smog, is the conversion of NO to NO₂. I see that as a central issue that is currently

been challenged to a certain extent. There are two cycles of smog chemistry. One is the HO[•]-organic cycle and the other one is the inorganic and NO_x, and they are connected with this NO to NO₂ conversion. I would say the paradigm is the NO to NO₂ conversion in connection with these cycles. So that all the organic chemistry is run through the HO[•] mechanism and production of hydrogen containing radicals. I offer that as an alternative to the paradigm structure. It's not inconsistent with what you've said but it's a little different.

Jeffries: I believe that EPA wants a credible, and therefore, 'scientifically sound' approach. And for the second point, I do not see what you have described as being any different at all. The paradigm-as-shared-values is a general acceptance of a set of beliefs as to how the world functions with respect to smog chemistry. I certainly believe the general description that you gave.

Atkinson: Why must we accept a single mechanism for use in an individual air shed or regional model? Provided that mechanisms are all based on the best available data and are tested equivalently, you can still end up with different formulations of those mechanisms. Indeed, they well may differ in their predictions. But that is because the testing against environmental chamber data, or whatever data, really is not testing under conditions which are applicable to the atmosphere. We are quite a ways away from true atmospheric conditions in the testing. So you can easily end up with different predictions and different estimations. But, based upon our present data, they still could be wrong. They do give you, if you have multiple mechanisms, based upon those same data base, some idea of the potential uncertainties in the data. I would not say accuracies, but certainly potential uncertainties.

Jeffries: I was not very clear in making my point, Roger. I do not advocate a single mechanism, by any means. But what has come to be true outside of our own small group, is that users tend to think of the mechanisms as an entity: it is the 'Carbon Bond Mechanism,' it is the Carter-Atkinson-Lloyd-Lurmann Mechanism. It's an entity. It is like one is right and one is wrong. What I am suggesting, is that there are elements of knowledge—as you say—there are kinetic basis for which all of us probably believe that they are true representations of what is really going on....

Atkinson: Well, they are our best representations....

Jeffries: ... Ok, but they are adequate enough for a lot of people to agree that they are probably the best explanations we have at the current time. And

further we would expect, if a mechanism is going to include that information, then it should be the same in the different mechanisms. In other words, there is a set of information which we believe has to be the same in all mechanisms. ...

Atkinson: Right, which means that they're based on the same data base.

Jeffries: Yes, but then there is the next group of information for which we are less certain about what is going on, and for which we do not have adequate data, either kinetically or mechanistically to make distinctions among them. To build a model we have to make estimates of what we believe is true or right or wrong. Different people do make different estimates, and thus come up with different answers that then led to different predictions. That is what I called the second class of knowledge I wanted to identify. That is, 'What is an adequate representation for what we probably know?' And what are the consequences of different people having chosen that information differently. This approach is different in concept you see, than simply saying, 'the Carbon Bond Mechanism is wrong, or the Carter-Atkinson-Lloyd-Lurmann mechanism is right.' Instead of thinking about it that way, which tends to give the impression that there are different representations, we need to move to a three-tiered representation. In simple terms, there is a part of the chemistry which we all believe in, there is an area in which we are uncertain so we make different choices, and there is another whole area which is wide open to speculation.

Atkinson: Yes, but those two mechanisms ultimately, how ever many there may be, will ultimately converge as our knowledge gets better.

Jeffries: Granted.

Atkinson: So, I do not think you should differentiate too much between different mechanisms. I mean, clearly there're all aiming towards the same end purpose.

Jeffries: Yes. But that's not how they're perceived outside.

Atkinson: They serve to define the uncertainties and how best to attack them, if done correctly. They serve to define the uncertainties in our knowledge—the areas where further input data are needed. By having those two or three representations of the same data base.

Jeffries: There is another side of that coin: they serve to confuse the hell out of people who do not know the details of the data base. And I think that is what

leads to part of the problem here. That people see that one mechanism predicts one way and one predicts another way. Then they begin to say, 'Which one is right and which one is wrong?' In truth, aspects of both are right, aspects of both are probably ok, and other aspects of both are probably wrong.

Atkinson: Of course.

McRae: Harvey, since you set your report up as a straw man, I would like to challenge several points you made. First of all, I'd like to challenge a basic premise that you say that there are lots of assertions about model performance and our inability as a 'community to predict outcomes'. I think one of the things that would be very useful in this workshop is simply to define what we think are these uncertainties. Are they scientific uncertainties? Are they educational uncertainties? For example, somebody in the White House or on the Hill may not be able to distinguish between nitric oxide and nitrous oxide. That is an educational issue as distinct from a science issue. I think it would be frequently referred to as what constitutes a good prediction. To whom and why? And what exactly do we mean by that? The third area when it comes to the basic premises that I think is extremely important to distinguish between uncertainties and the effects of those uncertainties. It is very easy to build a laundry list of what we do not know. The critical question is, 'What is the effect of those uncertainties in terms of predicted outcome?'

The second thing I would like to challenge is in use of Kuhn's work as a paradigm for two reasons. One is that it doesn't really reflect what's been going on over the last 25 years in artificial intelligence and common good psychology. There's lots of insights in terms of problem solving, that are helpful in that area. But, my major criticism with the use of Kuhn's description is that in fact it is descriptive. It is not constructive. It does not provide direct guidance in terms of what we should do to improve our behavior in the future. The second point is that I'd like to challenge the view that you stated a number of times, is that the revolutions in science are a group project and I can think of lots of examples in which that's not true. The view that change—most revolutions in science do not easily fit into Kuhn's paradigm. For example, Newton's Law of Gravity or Einstein's General Theory, are not good products. They are step changes. There was no pre-existing science, pre-existing theories, they were just something that was created.

Kuhn doesn't really discuss the role of data in discriminating between alternative theories. One of the difficulties in reading this report is that, it would be interesting to write a historical analysis of the evolution of the quality of smog chamber data. And, to what extent that has influenced the choices that people in

modeling community might make. The final thing about the use of Kuhn's theory that I'm disturbed about is the essentially a post theory view. It is looking back. It doesn't tell you anything about how to identify whether you're in a paradigm or about to undergo the gestalt switch or whatever. You frequently have eluded to the fact that the modelers have prematurely declared victory. I would offer the view that we do not know which race we're running in. In many cases the nature of the problem is evolving. So, the fact that we prematurely claim victory really, I've used the case where a lot more about how to define what victory is at the time you're conducting it.

Another area, which Gary Whitten also raised, which I think is extremely important in terms of this group and the regulatory agencies is the view that between science and engineering, that the science is unraveling the details of the process, the engineering is actually the use of those what you know in a regulatory environment. For example, we do not understand all of the details of materials on the mechanical level but we still build bridges. We need to cross the river. We have an analogous form in air pollution. Thirty percent of the population is exposed to levels of above the Federal Ambient Air Quality Standard. Twenty-five cities are not going to meet the containment deadlines. That's a real problem: an engineering problem. One that we need to suggest approaches to.

The final area, which is a structure to follow what Roger suggested, is that one thing that as a community we could do, is to simply define what we think to be the current uncertainties and nature of the problem. There are lots and lots of assertions from GAO reports that we can't predict anything. What exactly are the problems? I think that we can define them in terms of the science issue or engineering issue. And then I think there's lots of other details which we can call up based on largely what Roger put together as a series of questions. So, I think that Kuhn's idea is an interesting way to write historical perspective, but I do not think that it is particular constructive in this environment.

Jeffries: [These comments were added after the workshop]. I believe that I stated the source of the workshop issues quite clearly in both the background document and in my workshop presentation. I believe that large parts of the paradigm-as-shared-values which I hold, along with others here, are under challenge, in part because of the community's inability to resolve the differences among competing explanations of the chemistry of paraffins and olefins, and as a result of the failure to offer acceptable explanations of the aromatics chemistry. These are scientific uncertainties that lead to problems in application areas, which might be described as educational problems. Papers presented at this workshop will show the effects

of these scientific uncertainties: differences in predicted control requirements that translate into million dollar differences in costs. Thus, the issues discussed here are not just scientific 'laundry lists,' their resolutions are important to EPA's progress in meeting goals established by Congress. Furthermore, those not in the field have questioned whether modelers are honest, or more politely, whether there is any truth in what they present. Why is this so? What is required to prevent mechanisms from being mere opinion?

With regard to the use of Kuhn's work to organize the background discussion, as I said in the document, "There is no one truth about what happened." Kuhn's description is relevant for me—it fits my experience and provides detail in my map of reality. I very much disagree with your statement that Kuhn's view is not constructive, that is, 'does not provide direct guidance in terms of what we should do to improve our behavior in the future.' Kuhn's work provided me the concept of "reasonable agreement" and explained how data and theory function together. His description of normal science is analogous to your concept of the moving target or that 'the nature of the problem is evolving.' As for behavioral guidance, Kuhn's work suggested to me that if this workshop attempted to develop a set of standards and method of testing mechanisms that did not conform with current practice, then the workshop's product would be doomed to failure, it would simply be ignored by the practicing scientists because it would not conform to his needs, in spite of how well it might meet EPA's needs for regulatory consistency. Therefore Kuhn's philosophy certainly influenced my outlook and approach to this workshop and thus has been very constructive for me and also for others here.

Paper 2

How Do You Know a Good Model When You See One?

**Joseph Tikvart
Office of Air Quality Planning and Standards
U.S. Environmental Protection Agency**

How Do You Know a Good Model When You See One?

Joeseph Tikvart
Office of Air Quality Planning and Standards
U.S. Environmental Protection Agency
Research Triangle Park, NC

Introduction

The background document written by Jeffries and Arnold is an interesting historical and philosophical work, but I do not think it is really a "straw man." I am not sure that a group of people should tell those developing or doing the science how to go about the task. They should go at it as they would, based on their training as scientists, and then let others judge how they did. I do not think there should be a test to tell people how to go about that process or to set up a procedure for doing this work. I view the issue that we are dealing with as basically, "How I know a good model when I see one." The background document did not deal with that issue directly.

There was a second background document written by Sexton and Jeffries entitled "Need for Chemical Mechanism Documentation" that I think did attempt to address the issue. This second document considered several points.

First, it expressed the need for guidance from the developer. What should be done about questions where there are other versions of the models or the chemical mechanism developed and where the user modifies that mechanism? How should this problem be handled?

Second, there was discussion of confusion about the support programs for the chemical mechanism; that is, how do we get the input needed to drive the mechanism. I think an example here would be the extent of pre-processors needed to run the SAI airshed model, and how one should deal with these.

Third, the document discussed machine-oriented problems. That is, most codes are written for one type of machine; to apply it to another type of machine might require extensive changes. For example, the simple Gaussian models in UNAMAP are written for the UNIVAC computer and you can not merely take that UNAMAP code and put it on the IBM. There are some minor changes that need to be made. Another example is the air shed model that was written for the PRIME. To run it on the UNIVAC is not necessarily a simple undertaking. So, how should we deal with machine oriented problems?

Next, the document discussed how to deal with errors that are found. How are the changes formally made?

Finally the document proposed what the documentation should consider. For example, what are documentation standards or what are the criteria for a user's guide. The document included a suggestion that there should be EPA guidelines on model use, and there should be EPA distribution of the models. There was a suggestion that there should be EPA review of all new techniques. When dealing primarily with Gaussian dispersion models, EPA has already addressed many of these problems in terms of the documentation and in terms of attempting to answer the question of how you know a good model when you see it.

There are two parts to my presentation. First, I will describe our regulatory modeling guidance. Second, I will discuss some of the qualities of a good model.

Model Guidance

While my personal experience is primarily with Gaussian models, my organization does have some experience with more chemically oriented models. I believe that the present Agency regulatory procedure could be applied specifically to chemical mechanisms or models that include chemical mechanisms. First of all, let us discuss the availability of these techniques. All the Gaussian models that are to be used in regulatory programs are a part of UNAMAP; these cost approximately \$1,000 to obtain. In all cases, there are user's guides for these models. There are codes for the models and those that are most appropriate for regulatory application have a default option. Throw a single switch in the model and it then operates in a specific mode that has been subject to public comment and criticism. In addition, there is a test data set to go with these models to make sure that new users can operate them as the developer intended.

Perhaps I could introduce here the concept of documentation standards for models. Bruce Turner, jointly with a contractor, prepared a handbook on how to prepare user's guides for air quality models. So, in terms of how to document a model, *i.e.*, preparing a user's guide for it, there is guidance already available.

For the models presently applicable to regulatory programs, there are codes, user's guides, and test data sets. I see no reason why this could not be done for chemical mechanisms that this group wishes to entertain.

With regard to guidance for regulatory programs, we do have a guideline on air quality models that specifies data bases and models to be used, and within that

guideline, EKMA and the airshed model are specifically included. Since we can not write down everything in one document that indicates how to use the models for regulatory programs, we have a model clearinghouse, whereby the EPA regional offices and states can ask for clarification or for guidance on how to treat a unique problem. We are in the process of setting up or expanding such a clearinghouse activity to include EKMA.

You should also be aware that in each regional office, and for that matter in most states, there is somebody who is specifically responsible for modeling. In the present situation, when we change the modeling guidance, we have to go through a process of public comment and review as part of the regulatory process.

Model Quality

What was the criteria for the models and data that were included in the modeling guideline? We tried to include the models that were considered to be the best. The first criteria was an evaluation and scientific peer review of the model.

This evaluation or peer review for many of the models did not turn out to be as definitive as we would have liked, since the peer reviewers felt that the models were dated and did not represent the state of the art. The state-of-the-art models had not yet been documented and systematized well enough for a wide variety of people to use them; so, in many cases, we had to include models that were less rigorous, scientifically. They were, however, the best-documented, off-the-shelf items that could be used now. Where the evaluation and peer reviews of these models were ambivalent, recommendations were based on consistency, public familiarity with techniques, and use in the past.

A question was posed before, "Why can't there be more than one model used in regulatory situations?" The answer is that those being regulated have a natural tendency to use techniques that are credible and give them the least control requirements. If that is different from another source downwind, then we find ourselves in a regulatory consistency issue on a national level. Congress has told EPA that we need to be consistent. So, in answer to the question, "Why can't two models be used?," the answer is to be consistent we must use only the best one. Other factors are availability and resources.

There was a question in the background document relating to changes made to the model, whether they are associated with running on different computer, or making a modification to a code, to satisfy particular purposes. In several cases, our

approach has been to set up an equivalency test, where we provide a set of data and say, "If you can run your model with this set of data and come up with essentially the same answers within a certain percentage, then the codes are considered to be equivalent." This applies whether it's a whole different model that you constructed on your own or just changing from one machine to another machine. Thus, we have set a precedent for what we term a equivalency test.

Where the user believes that the model suggested by EPA is not appropriate, there is a mechanism by which they can do an on-site evaluation and thereby show, for their data base and their source, that another model is more appropriate. That approach has been used in a number of cases.

Finally, for situations where there is no model recommended, there are some general criteria that we use. The model used must be scientifically sound, be one for which something about its accuracy is known, and be one that we know does not underestimate the design concentrations for an area.

The steps described above are, in general terms, the program that we have underway; they fit with the needs expressed in the background documents for this workshop. In other words, these are things suggested for consideration that we are successfully implementing.

Where did the models included in our program come from? In 1980, we published a Federal Registry notice that asked for models to be considered for use in regulation. We placed six requirements on those models, however.

First, a model had to be a computerized code unless it was so simple that calculation could be done on the back of an envelope. A model should not be limited to a paper published in a journal with a lot of equations. That might be the basis for a model, but it does not constitute a completely implementable model. It has to be a computerized code. Second, there has to be documentation in a user's guide. Third, there has to be a test data set, so that other people can run the model and test it. Fourth, it has to have an air pollution control application. Fifth, there has to be some comparison with observed data so that we know how good it is. And, sixth, there has to be public assess at reasonable cost.

Most of the models that we received were Gaussian models, for single source or urban applications. We also received reactive models.

A formal model evaluation has been performed for five categories of models: rural, urban, complex terrain, long-range transport, and mobile. All the models

were applied to one or more data bases and a comparative evaluation of the models was conducted for those data bases. The performance was statistically evaluated for each model for the same data. For the first three categories, we had the American Meteorological Society (AMS), under a cooperative agreement with EPA, do a peer scientific review of those models based on the performance evaluation and their scientific assessment of the models.

The peer review for the first three categories is complete. The AMS is now considering peer reviews for the long range transport and for mobile models.

What was the outcome of this peer review? I have already said that, for the rural models, the available models were not state of the art—we could not distinguish among them—they are all 20 years old. I will comment that one of the peer reviewers was a very bright guy, and subsequently submitted a model that fulfilled all the recommendations of the peer reviewers. Unfortunately there was a fatal flaw in one parameter that he chose, thus his model did worst than all the other models. I think he has since corrected that parameter. The point I wish to make is that just because a model looks better scientifically does not mean that it necessarily performs better. This is related to one of the questions which was posed to you, “Is a model a good model because it is scientifically satisfying?” I have had several experiences with scientifically satisfying models that performed worse than models already in hand.

In conducting the performance evaluation, we exercised care. For example, we wanted to avoid any situation where the developer would say, ‘You did not run my model right.’ So, we had the developers first review the data bases, and make any coding changes in the models necessary to accommodate those data bases. Then they verified that we were running the models as the developers would have us run the models with the test data set. Next the developer selected the specific elements and parameters to be used in the model and the options. We wanted to run the model just like the developer thought it should be run. The developers were not always happy in the end, sometimes they said, “Oh, I made the wrong choice.” But for most cases, this procedure worked well. Finally, we allowed the developers to review the findings, comment on the draft final report, and make any final changes, short of changing the way the model was run.

In the urban situation, the reviewers said that the four annual average models had performances that were fairly narrowly grouped, and they again could not distinguish among them. That is, the annual models, which smear out a lot of information, did not perform differently in a statistical sense and had much of the

same theory. For the two short-term models, the peer reviewers essentially said that they have similar abilities as the rural models, but they thought that one was slightly better.

For the complex terrain models, the peer reviewers did feel that one model appeared to perform better than the other models, and, as a result, EPA has requested public comment to include this model in our modeling guidance. And that public comment period is about to close. So we have reacted to getting better science into our guidance.

I have one final comment. Jeffries began his background document with a quote by the GAO of an EPA official that said something to the effect that with models "more precise results are unlikely because of the limitations of the science." I think I am the official quoted there, but the GAO auditors did not interpret me correctly. Essentially, what I said was that there is beginning to evolve concern that even if we had a perfectly accurate model, because of the atmosphere's inability to replicate its turbulent state, one may not know he has a perfectly accurate model. Because of turbulence in the atmosphere, we may have different realizations of the same event. Models calculate a mean for a specific event, while the atmosphere does not. For each replication of the same event, the turbulent nature of the atmosphere may give us a different answer. Therefore, in dealing with the evaluation of models, one has to deal with the accuracy of the data bases, the physical realism of the models, and the natural variability in the atmosphere. Due to natural variability, we may not be able to establish that we have a better model even though, scientifically we might believe we should. The AMS, under the EPA cooperative agreement, is struggling with this problem and I presume that atmospheric chemistry and its problems just compound the problems.

Discussion Following Tikvart's Presentation

Demerjian: Joe, given the fact that EPA has accepted as a guidance criteria modeling technology that is twenty years old, has limited data sets to evaluate the models, and has presumably identified certain kinds of performance statistics based on these, why would EPA want from the chemical models more than, let us say, a factor of two and a half? Actually, I am sure we do better than two and a half.

Tikvart: Tesche made a presentation on the airshed model at the conference here last week. He reviewed ten applications and his conclusion was that the airshed model was accurate to within plus or minus 30 percent. Based on experience we have had, I'm inclined to agree with that. The airshed model does appear to be accurate within plus or minus 30 percent. It does tend to be on the low side, and you can argue about why that is, but that certainly is as good as the Gaussian models. Maybe there is a better chemical mechanism since we have been using Carbon Bond II in the version of the airshed model that has been evaluated. Perhaps there could be a better chemical mechanism, a better way to handle the numerical calculations for dispersion, and so forth; but then there is a need to test that version to find out that the better science does, in fact, give you more accurate answers. I'm not trying to say that we do not need better science but the science we are using appears to be not that bad.

Demerjian: When you looked at the point source models, I remember that basically there was a lot of discussion about the fact that they were pretty much all the same, that they were all mostly based on Gaussian approximations. So people that were dealing in the science itself, whatever science was there, were basically saying, "They are all the same except for a few thousand whistles on how to treat this or that." Then there was another issue about whether the data that was being used for the observations could be even defined within those bells and whistles of those models which were better. But, isn't it true that, if they were used in a regulatory sense, some of those models might require more conservative reductions than others, or do they also do that the same too?

Tikvart: No. Three models, all of which perform the same within your ability to make a statistical distinction, could give three different requirements for control.

Demerjian: Given that, how did you make a decision in terms of which would be the best?

Tikvart: We have reverted to past use and familiarity with codes. That was our position. Scientifically, you cannot distinguish between the models. The peer reviewers essentially refused to distinguish between them.

Demerjian: So, you use precedence then?

Tikvart: It was precedence. In other words, if you have to have a consistent, stable regulatory program, and one of the models—the model that has been the basis for your regulatory program up to now—does not perform differently, and at least as good as other candidate models, why should you switch? Now there are cases where somebody might get a more desirable answer with another model. You do not want to foster inconsistencies by saying go ahead and use whatever model gives the answer you want, unless for that specific plant or facility it can be shown that the model performs better. There have been several cases where somebody has taken on-site data and performed an evaluation for a regulatory model and one or two other models and used the model that performed best. The chosen model was not the regulatory model in these site-specific cases.

Demerjian: Ok, hypothetically we have the EKMA mechanism, we have some version of Carbon Bond, we have a version of the SAPRC/ERT model and there might be some others. Presumably we go through a similar exercise as you have described and we come to the conclusion that within the uncertainties of the data, where ever it may come from, whether it is the Riverside chamber or Harvey's chamber, that they all fit within the limits of uncertainty of the information. And, from what we can tell, they're all based on pretty much the same standard sets of scientific information and they all perform equally well. Now we discover when we exercise them that one requires 20% more reduction in hydrocarbons than another, or 30% more. Now, what is your position at that point?

Tikvart: You've gone through a performance evaluation and a peer review. Right? In a sense we have already established for the regulatory program what to do because the EKMA model is the guidance for a screening type of discussion, although it is good enough to base a control strategy on and the sophisticated model of choice in the modeling guideline is the SAI airshed model.

Demerjian: Let's just talk about the mechanisms and not the airshed model. Let's say we're talking about mechanisms that can plug into EKMA. Even if right now, if there was a consensus opinion that whether it is the Dodge chemistry or any of these three others that they all were indistinguishable in terms of a performance evaluation.

Tikvart: Then why not use EKMA and OZIP the way it is now?

Demerjian: Even if it turned out that EKMA was, let's say 20% or 30% more conservative. Would that ever be factored into your decision?

Meyer: Let me try to get Joe off the hook a little bit here. Our thinking right now is that, we are most probably going to change the recommended mechanism. I think probably the major basis for the recommendation is the more up-to-date experimental work and theory that the newer mechanisms reflect. We're going to try to select one mechanism however, for use in regulatory applications. And, our thinking is basically that this will be used for perhaps a 3 or 4 year period at which time there will once again be a reassessment by the scientific community and the same kind of question will be asked, should we change this mechanism?

Demerjian: Ok, let us move up to this next tier then. We have eliminated the old version mechanism in EKMA. Now we can move to some new mechanism. Now we have the same scenario: three mechanisms, presumably all state of the science, presumably all show reasonable performance credentials in terms of data. But now still showing this difference in terms of policy requirements, in terms of control.

Tikvart: Have you already made the change or are you in the process of selecting the new mechanism?

Demerjian: We are in the selection process and we discovered this unfortunate problem: three mechanisms, all performing within the standards of the data.

Tikvart: All three of them are better than what you have now; you know that?

Demerjian: We as a group are reaching a consensus. We have established that *they are all definite advancements over where we were*. Now we have discovered that even though it is still within the experimental data, there is enough leverage within this to give different control requirements.

Tikvart: The regulatory blackboard in that sense is sort of blank and prepared for the scientific community to write on. We would ask the scientists to give us your best assessment of what this most scientifically credible technique is.

Meyer: One other dimension for this kind of problem is what kind of information is required by the mechanisms that regulators would have to try to come up with in terms of inputs to the mechanisms. I think perhaps one factor that might be

considered is that, although two mechanisms might be equally valid as far as we can make a judgment on, if one requires a lot of information that we have to wave our hands about a lot, presumably the other one would be preferred.

Tikvart: And here you get into terms like familiarity, resource requirements, ease of use, *etc.* If you honestly cannot distinguish between the techniques scientifically, then I think you degenerate to those sort of criteria.

Demerjian: Again, we have a curious situation. We have one mechanism that requires 20% more control. Now, we use this criteria that has just been given and we find out that the one that requires 20% more control actually is a lot easier to use. It turns out that you can pass it out in a form that runs on a PC computer and cities only have to collect a few data points and they can run it. When the people use this tool they find out that that this extra 20% control is costing them \$100 million. Now, what's your justification for making the decision that presumably it is the application part that helps you to decide which model to use?

Tikvart: Then there's another possible answer and that is the models go out for public comment. Since people are going to be regulated by that model, present the facts in the Federal Register notice and say that you are seeking input on what direction you should take. That is one alternative. Since it seems that the blackboard is blank, you can say, "It is your choice, World, go with whatever approach you want to use." However, very quickly I believe you will get into a situation of inconsistency and irregularities being jerked back and forth by what is most convenient or by what is least costly to specific interests. This could result in one model in one part of the country and another model in another part of the country.

Meyer: I think another point too, to raise is, that it is probably unlikely that, across the board, one model would predict 20% less control being needed than the other one. And, indeed if there is a difference, the difference may hinge on some of the inputs to one of the other models, for which there might be a great deal of uncertainty. I think what we would want to do would be to select the mechanism that then has fewer of these very uncertain inputs.

Lloyd: Part of this discussion leaves me very confused. We are meeting, as I understood it, to try to resolve some critical questions in the chemistry area and to help out EPA. Now, I am hearing from Mr. Tikvart that EPA does not necessarily need the best mechanism, but probably the most readily available, best documented mechanism.

Tikvart: No, first of all we want the best mechanism. I do not think there's any equivocation about that.

Lloyd: But your time frame, as Mr. Meyer was saying, is not going to allow you to use the best mechanism. You are talking a 3 to 4 year lag time in terms of identifying mechanisms and having any changes. Our earlier discussion led us to believe that evolvement of a model is a constantly changing thing as you get more input data. If in fact there's no way in which you can allow changes for 3 to 4 years, then by definition you're coming up with something which is outdated.

Tikvart: I have a couple of comments on that. First of all, does the scientific community know when to let loose? This is a complaint I have; I do not think that the scientific community does know when to let loose of a new technique. Frequently, they have to be forced to let loose. I should not say forced, I mean preempted. And then somebody may take it and do something dumb with it. So, the scientific community is oriented to continue improvement; it is difficult to decide when this is the best shot. Another comment is that just because somebody thinks he has an improvement, does not necessarily mean that he actually does; it might be just one interim step to something better. So, I believe there is merit in having some lag between when a new idea pops out and when it is actually implemented. It does need some testing before you go jerking around multi-million dollar programs and just saying, 'Here is a new technique.' A perfect example is work on plume rise equations. At one point, about 10 years ago, you would have been changing the plume rise equation used in Gaussian models every six months, as new information evolved. I do not think that is in anybody's interest. The third point is, that just because somebody thinks they have a better technique, doesn't necessarily mean that they have a better technique until it's evaluated and compared to other existing and new techniques. So, there are some tests I believe that are appropriate before you impose new technology on the regulated community.

Lloyd: One of the points, I think we must make clear here from the chemistry point, is we are not just dealing with the ozone as we have in the past with some of those. We have to look ahead and we are already addressing many of the problems in the acid deposition area. Secondly, one simply question. Supposing with the end of this meeting we came up with a consensus, best mechanism. How long would that be before EPA would use that?

Tikvart: Ok, I believe you would have to address some of the questions I just posed.

Lloyd: Supposing we did.

Tikvart: Addressing them includes a user's guide, codes, documentation in general, evaluation and peer review, all of which take time. Perhaps some of you have data you brought with you. But, let's say all of this has been taken care of and you now feel you have the best technique. First of all, as a regulator what I would want to do is, see what that new technique does to control programs. Is there a major relaxation or tightening of the emissions required? That question is asked first because when you put this out for the public, they need the answer. You would have to go out for public comment and announce that there is a new technique that EPA wants to use. Here is the information on how accurate it is. Here is what the scientific community thinks of it and here is what it will do to emission control requirements. World, comment on this as a regulatory tool.

Lloyd: So, what are we talking about in a time frame?

Tikvart: I think we are talking about 18 months.

Jacob: I am quite uneasy with the distinction between urban and rural models, especially as we move towards more regional problems. What is a rural air quality model? Some rural areas have NO_x conditions which are typical of cities and other times have NO_x conditions which are typical of a clean troposphere, so in such a situation, what would constitute a rural model? I think a much more objective criterion that we have got to have in the model documentation is what is the range of NO_x that it can be applied to. What is the range of hydrocarbons can be applied? What kind of hydrocarbons can be present in the system without making the model go haywire? I think all these questions must certainly be a part of the documentation.

Tikvart: I agree with what you are saying. If the model developer would lay out specific limitations on the applicability of the model that will help the regulator who wants to use this technique in a credible scientific framework. What happens too often is there is a regulatory need for a model, the regulator interrupts a research program and says give me what you have now because I need it. Then we take it and apply it to a problem that it may not have really been intended for but there may be no other choice. So the burden of responsibility for the model is transferred in that case from the developer to the person who wants to use it. I think we would much rather see the burden of applicability be on the person who developed it. So, I agree with what you are saying.

Whitten: I just have a couple of comments. First of all, in the beginning of your talk you wanted to ask the question how do you recognize a good model when you

see one? I think that that is a very difficult question to answer, and we can make more of an answer to a related question anyway, which is, 'How do you know a bad model when you see one?' I think in the archives of data that we have, we have certain types of experimental conditions we know of that present grounds to reject certain types of mechanisms that have been used in the past because they just can *not simulate a certain type of condition*. Of course, you have to ask the question, 'Are those conditions of any general applicability?'

Secondly, at the end of your talk, you mentioned about the randomness of the atmosphere and it is true that meteorologically speaking there is a turbulence and a randomness and that makes it difficult to determine where a certain pollutant might be going, where a certain parcel of air is going. The nature of a chemistry model does not have that random aspect to it. It simulates what happens in a well-mixed parcel of air. So the randomness then applies as to we do not know where that parcel of air might go on a given day because of the turbulence or whether or not it might be perfectly well mixed, so that the chemistry is much more deterministic. It does not have that randomness. What randomness it does have is more in the eighth or ninth decimal place type of thing.

Finally, I am a little troubled by Meyer's comment on discounting models which need more data. I think the general approach has been, as you have outlined, to have sort of a default for various things. And in the case of atmospheric mechanisms, the more recent ones have many more species to define the various aspects of reactivity that we see in the atmosphere. The original EKMA for instance, more or less, fixed the reactivity so there was no way to change it. I think the newer mechanisms can be used for a default type of reactivity and so therefore that is always there. So you still do not need any reactivity data. If you have some reactivity data which shows a given urban area or something like that, is far out from the norm, then more recent mechanisms have the ability to accommodate that, but it is not necessarily a data requirement that use them all in the first place.

Carter: You were talking about wanting to keep on using the existing model if it performs adequately, but what do you do in a case when you get new information that shows that the theoretical underpinnings of the existing model are wrong? Like, for example, the original EKMA model. One of the most important reactions in it and was subsequently shown to have a rate constant that was off by an several orders of magnitude, yet that model is still being used. I was just wondering what the attitudes are. Do you continue using it because it seems to perform adequately on the validation data sets or what?

Tikvart: I guess there are a couple of concerns that I would have. First, how serious is the deficiency? Is there some other better choice? And given the deficiency and the other choice, how accurate are both when you compare them to measurements?

Carter: Let us say that we have the other newer models which have not been used as long but that they have that particular deficiency corrected.

Tikvart: Are they more accurate when compared to observed data?

Atkinson: We know that they are scientifically more correct.

Tikvart: Are they more accurate though?

Carter: They are as equally accurate.

Atkinson: They fit the same environmental chamber data base to the same degree, and one is scientifically more accurate.

Tikvart: I would have to say that the criteria that I have used relative to the Gaussian models, is that if you have something that is scientifically better and it performs as well as what you have now you should take steps to make a change.

Meyer: I agree with that. I think that maybe our concern was that these would not be continual changes that they would be done in discrete periods, three or four year periods so that we are not constantly changing the tools that we are trying to use for regulatory purposes.

Tikvart: But, just to make sure that we are clear, let's assume that if you came to me and said, "There is something wrong with what you are using now and I have something that I know is theoretically better." However, unless you can show me that statistically it performs better, I would say come back later after you know more.

Carter: But if this discovery showed that the existing model was sort of invalid at the core, would it at least accelerate the process of finding a new one?

Tikvart: Perhaps, but this has happened more than once. The scientific community has said, "What you use is twenty years old! We do not think the answers you are getting you are getting for the right reasons! And, Oh by the way, here is

something that we think fulfills everything that you want and you should be using it instead." However, I find out that, because of some misjudgement as to what value to assign a parameter in this new model, it is grossly wrong, then you can understand some caution and some need for proof of the better aspects of the new technique.

Jeffries: There is a hidden assumption here that we have all been making. Ken started out making the assumption that all the models fit the data equally well. A major point of my presentation is that we hide many of the anomalies that are present in the system in the asserted "reasonable agreement" with the data. In reality, if you examine each mechanism in detail it will make unique predictions that can be validated, verified, or tested, as a way to choose one from the other. But that requires that we have an agreement among us as to what the data really are. And so, for example, in a case where there is a disagreement about the formaldehyde prediction of one mechanism verses another mechanism, we presently can not resolve the issue because sometimes Carter's measurements of formaldehyde are different from my measurement of formaldehyde and we don't know which one is right. If we concentrate on improving the data, we can separate one mechanism from the other.

Tikvart: Are you limited just to smog chamber data?

Jeffries: Mostly.

Tikvart: OK, that is a tougher problem I would think. If you could take a dispersion model and test a variety of chemical codes in one framework, that might provide you with a test. But if you don't have the ambient data, then ...

Jeffries: Most of the time the problem is difficult enough trying to unambiguously compare the data with the theoretical predictions. In the atmospheric model, there is ambiguity about what is causing the change in concentration because it is a non-linear interactive process and so even though you may have the measurement out there you do not know that it is the chemistry that is causing you not to predict that number correctly. It could be a transport, or a chemistry, or an emissions problem and this is too complicated. So you must resort, you must go back to a simpler system to juxtapose unambiguously the data, the observations, with the predictions. It does not do any good to compare observations with predictions if the situation is so ambiguous as to cause and effect.

Bradow: I think that the use of chemical models in the acid rain program to predict products of other than ozone may be the key element in creating some

of the differences in discussion here. In other words, it may be possible that all the chemical modules do reasonably well in predicting ozone in those cases that are most important for controlling ozone. This may not be the case if we use these same chemical modules, in the case that we wish to evaluate the importance of the oxidation processes in creating hydrogen peroxide and nitric acid and thus influencing acid rain. A key element that Jeffries speaks of, and others here too, has been how well do these models predict products other than ozone? It is possible that engineering designers figuratively have built a birch bark canoe in order to control ozone and then find themselves faced with transporting Noah's Ark animals.

Atkinson: I would like to make one comment on that and actually it bares on Jacob's comment about the developer putting on the areas of applicability of his model. If you use the data base we presently have, unfortunately I am not sure that you could ever say that you test a model for acid disposition against certain environmental chamber data and I can see that it fits with your comment. We really are extrapolating, I believe and I may be wrong, but I believe we are extrapolating quite a way beyond our data base. I mean we understand hopefully, pretty well, the chemical reactions which occur but we don't really have a data base, a global data base to test them against. Do you have a comment on that Daniel?

Jacob: I think I quite understand the problem with smog chamber data, but I just want to point out that it is extremely dangerous to extend such models that were obtained at high concentrations.

Atkinson: I fully agree. I mean anybody who has to have NO in their chemistry, clearly it has to be applicable to nighttime conditions or multiday conditions. There are two sides of the coin. There are always problems in extrapolating data for the long-range consequences.

Jacob: Yes, I think aside from the fact that the smog chamber data does not cover a wide range of compulsory conditions, there is also problems within the development of the chemical mechanism itself. Based on smog chamber data which is done satisfactory at pretty high concentrations you can say, 'I understand pretty well how the system works and even if the concentrations are lower that is what I should get.' But we all know that there is a big difference between for example, the low NO_x and the high NO_x region and the chemistry imbedded in the chemical mechanism may just not work. Aside from not having the observation data to support it is the fact that based on chemical principles it will not go below a certain range of NO_x.

Atkinson: Or two mechanisms may give very different results under those conditions.

Dimitriadis: I am not sure that we have been reacting to Joe Tikvart's presentation in the right context. I think what Joe has been saying is that what we are doing in the dispersion model part, which would perhaps be applicable in the chemical mechanism area also, is that if we have any questions on the model's performance, or if we want to choose a model among two or three based on performance, we do two things. In essence, we are using existing field data processed according to the American Meteorological Society procedures to derive some statistical measures of the performance of the models. Then we throw all this information into the hands of a group of reviewers and we ask them to look at them and give us their judgments. Out of five reviewers you may have three reviewers that say yeah and two saying nay and the yeahs have it and that is the judgment. That is the outcome of this process. What we are dealing with here in this workshop is a different thing altogether.

I think we are concerned, number one, with the data base that was used to derive statistical measures of the performance. And this data base may be good or bad, of appropriate comprehensiveness or not, *etc.* The other question that we are dealing with is 'How do the reviewers go about making their judgments?' If it is subjective judgments, we do not want them. Therefore, we want to come up with some guidance in essence telling the reviewers how to go about judging the data so that they come up with a valid conclusion. Thus another group of reviewers should come up with the same judgment. Or if the same reviewers are asked the same questions one, two, or three weeks later they will come up with the same judgment, provided they have this procedure by which they go about making their judgments. So, Tikvart's procedure is something which is perhaps acceptable in some respects, but we are dealing with something entirely different. It is not a question of whether his procedure is applicable to the mechanism area or not. We are dealing with an entirely different subject and entirely different questions. So in a way what we are aiming in doing here presumably will be used in the process that Joe Tikvart has described and the two things then should be considered complementary. Would you agree with that assessment?

Tikvart: Fine, I think the key that I did not understand earlier was the importance of the data base question as you just described. There is some question as to how to interpret the basic data used to develop and evaluate the chemical mechanisms themselves. Whereas in the straightforward evaluation we took the data as absolutes, even though we knew there was some error.

Atkinson: I have a further comment which is that on the one hand we are being asked to look into how we get more accurate chemical mechanisms, gas-phase chemical mechanisms, and how we go about that; this in a way is the purpose, or part of the purpose, of this Workshop. On the other hand, what I heard this morning would lead me to believe that even if we do develop them they may well never be used, just because of time scale of the EPA. How is that going to impact things?

Tikvart: I do not think that should be a concern. If there is a better technique to be had, I think the people involved in Research and Development should find it within available funds. And those of us on the regulatory side get enough pressure that we will want to have the best science out there. I would hope that eventually best science would prevail even though there are those with different stances, *etc.*, that may find other approaches better. But I would hope that ultimately the best scientific technique, given that we can demonstrate that it is, would prevail.

Jeffries: Let me see if I understand what you are telling us. A lot of the discussion has been premised on the assumption that the mechanisms all fit the data equally well. If it turns out that we could demonstrate that a new mechanism based on better theory or more generally accepted theory could also be demonstrated to fit new data better than some other mechanism, would it be the choice?

Tikvart: I think the answer is yes, if I understood everything you said.

Jeffries: If we can narrow the range of reasonable agreement between observations and predictions from mechanisms to a point where we believe this one is a better representation than that one, we accept this one over that one and you will essentially be forced to do that because that now becomes the definition of better science. You do not want to go forward with a mechanism that is not good science.

Tikvart: Ned, have you heard anything here that you would find fault with?

Meyer: Well, the only thing I guess I would like to say again is that barring some major breakthrough, our inclination would be to make these changes in discrete increments. That is, to say we would look to see all the developments that have occurred in the last three to four years, try to get some judgment from the scientific community about what the best state-of-the-art is at that time and then try to incorporate those recommendations into our regulatory procedures and then essentially freeze our regulatory procedures for three or four years while science goes on. We would then have another one of these reassessments in three or four years.

The only exception to that I think would be if some major breakthrough came in that showed that what we were doing was just entirely wrong.

Atkinson: Seems like an awfully long change over this time scale, three to four years.

Jeffries: Ned, I think it might be useful for you to say how many SIPS have been done with which mechanisms.

Meyers: They have almost all been done with the Dodge mechanism, which is ten years old. Or more.

Tikvart: But can we say that in the interim there has been a clear agreement that there is something better?

Dodge: Oh, yes!

Jeffries: Yes.

Whitten: I think that perhaps we should be a little more clear in our definition of what data we are talking about here. The Dodge mechanism ran into trouble not so much from smog chamber data but from laboratory data which had shown a particular rate constant or two rate constants were in considerable error and that was based on laboratory data, not smog chamber data. And then, when those new rate constants based on laboratory data were then put into the mechanism and you went back, the procedure used say to adjust the propylene to butane ratio could have been readjusted and that particular type of mechanism could have been updated at that time. But it was felt that the, at least I believe, it was felt that the mechanism wasn't that far off. With its compensating shall we say errors between a rate constant which didn't agree with laboratory data and a percentage of propylene/butane there was an empirically adjusted to fit smog chamber data. So it was felt that the fit of the smog chamber data wouldn't be improved significantly enough if you stayed with the old rate constant and the old propylene/butane split. I just wanted to use this as an example. I think it shows that there is laboratory data that is really different from the smog chamber data and they are used in a different way and I think that we just need to be conscious of it.

Lurmann: I am still curious as to how the EPA decides that one mechanism that is the chosen mechanism to use in their planning and so forth. Is there a formal review process, what evaluation criteria have been used in the past and so forth.

Some of the same issues that we have been want to deal with in this conference today have been dealt with in previous situations and I guess that is really my question for now.

Meyer: Well, at the time we have to make the decision, which again is almost ten years now, we obviously did not know as much about the sensitivity of control estimates and chemical assumptions that go into these models, and from an OAQPS standpoint, we asked for the research people's best judgment, you know, what is the best available mechanism at the time and the recommendation was this propylene butane mechanism which since became known as the Dodge mechanism. I think one of the major purposes of this Workshop as we pointed out this morning is to come up with some more formalized rules that might be used periodically by the scientific community in making/helping EPA make judgments about, all right, as of this point in time what is the most appropriate mechanism for use in subsequent regulatory analyses and I guess the point I wanted to make was that this would probably have to be done on a periodic basis. I suggested every three or four years. In the meantime, of course, there will be intervening research and people's perspectives maybe will change between iterations of this group. I don't really know what all the criteria should be.

Lloyd: Just for clarification here are you saying that the next round of SIPS that the Dodge mechanism will be used and that you are looking for input from this group to see what criteria should be used for replacement?

Meyer: No, I think that for the next round probably we don't know yet but I think it is more than likely that we are not going to be using Dodge.

Lloyd: So then what set of criteria are you using to replace that mechanism?

Meyer: Well, again, I will have to pass the buck and in the sense that we have asked our people in the Office of Research and Development what their judgment is about the best mechanism at this time.

Dimitriadis: Actually, what Ned is going to do is he is going to come to us (ORD) will ask us, 'What is it that you recommend at this time?' This is what the practice has been in the past. We have to look at the existing mechanisms with some help from consultants. We may organize workshops, (had an EKMA workshop three or four years ago if you remember) and we use all this information internally for our own purposes. Then we derive a judgment and we pass it back to QAQPS. This has been the process in the past. Now, in the last two or three years there has

been a lot of emphasis on having everything that we do peer reviewed in a more or less systematic fashion. So I expect that any recommendations and any suggestions that we will pass onto the regulatory program will have to be supported through some kind of peer review process. But we still have all the faults of the previous process in the sense that we don't have a standard way that everybody would agree is the best way of judging a mechanism. That we do not have this we recognize as a lack and this is what we are trying to do at this workshop. But our process so far has been relying mostly on our own judgment which was helped through the use of consultants, workshops, but in these workshops the judgment was not left up to the participants to make. We simply used the advice and took it as best we could and made our own judgments.

McRae: I want to ask Basil and Joe a more general question. Part of what we are doing in this Workshop is to think about the future and one of the things I am concerned about and I am curious about your reaction is as mechanisms becoming more complicated, and data requirements becoming more complicated, the process of peer reviewing models and mechanisms is becoming much more difficult. Have we thought about the question of whether there is enough people out there to carry out these tests?

Tikvart: Are there enough different people from the developers who have the knowledge to do an unbiased review?

McRae: Well, let me say it in a more direct sense. There are twenty-five urban areas that currently violate the Federal Ambient Air Standard for ozone. In each of those regions how many people do you think there are that understand how to pick the difference between two competing chemical mechanisms and translate the evolving knowledge of chemistry into data collection procedures within their local agency?

Meyer: I think that is exactly the reason why we find ourselves in the position of having to recommend a mechanism. To take that prerogative basically out of the hands of the State and local agencies.

Tikvart: The answer is, "No there aren't very many people out there." What typically happens is the area hires a consultant and the consultant imposes his knowledgeable assessment of the situation. That consultant tends to drive it the way he feels it should go and that way generally, because of the interest of his client, will result in lower controls. So I think you're right, there aren't enough people out there to have a lot of freewheeling discussion and assessment and there

are only a few that can do the job, but that is a fact of life. So I am not sure that there is anything that can be done because, as you pointed out, it is getting very complicated. The number of people who have the knowledge to assess it is small.

McRae: Doesn't it argue for things like training programs?

Tikvart: Is this so complicated though that training programs to understand the mechanisms might be of limited use and training programs to apply a basic technique like EKMA might be more appropriate. I would question a training program to teach people how to use the Urban Airshed model because it is so complicated. It takes a staff who have a wide variety of knowledge and abilities to run the model. A prime example is the New York State Department of Environmental Conservation, which does have that ability. But, I would say that New York State is probably one of the best states to take this job on. I won't expect that every state could take on the Urban Airshed model, nor would I want them to, nor would I try and train them to. So yes, perhaps there is a desirability of training, but I am not sure that understanding the chemical mechanisms in the right way at a level you are talking about.

McRae: There is a lot more than just the chemistry. For example, if you look at what's happened to the structure of the hydrocarbon chemistry in most [tape change] ... people who have to give their emission inventories. And I am a little concerned that if we spend a lot of time worrying about getting the mechanism right and not thinking about the future implications of these mechanisms in terms of emission inventories that we will lose the whole ballgame. We have a very accurate, highly characterized, precise mechanism based on smog chamber data but lousy emissions data to drive that.

Tikvart: I am more concerned about that point, about the accuracy of the emissions inventory, than I am in understanding the chemical mechanisms. Ned, to what extent has our guidance addressed the specifics of the emissions inventory requirements? We have attempted to address it to some extent but, whether we addressed it fully enough relative to this comment, I am not sure.

Meyer: Well, I think we have something called the VOC speciation manual which is basically set up I think to subdivide the organic emissions categories suitable for use with the Carbon Bond type of approach. You know, if another kind of approach were made, then obviously that kind of a document would have to be changed so that it would be consistent with the categories needed.

Tikvart: I would be more concerned with the emissions inventory. Perhaps that comment ties into something that Ned said this morning about what mechanism to use. It has to be a mechanism that you could have data sufficiently detailed to drive and the number of species, etc., that you can include in your inventory might be one of those limitations.

Lurmann: Yeah, I think I agree with Greg. The mechanisms are getting more complicated. You have eight, ten, twelve classes of hydrocarbons instead of maybe four or five or six that you had in mechanisms of five or ten years ago. But quite frankly, I don't see the relevance of choosing what mechanism to use just because one is only has five classes of hydrocarbons and your data base happens to be set up, or your program set up for only five. I think the approach that should be used is that is to try to figure out which mechanism appears to be more accurate and then build a program based on that. The data species manual, it gives profiles for speciation based on individual compounds. Those can be run into any chemical class that you want. For example, the State of California keeps computerized files for three different mechanisms so that they are not forced into the situation of having the speciation profile set up for one mechanism because they use more than one and that is the approach that I would recommend here.

Paper 3

**Chemical Mechanism Development for Applications in
Atmospheric Simulation**

**Kenneth L. Demerjian
State University of New York-Albany**

Chemical Mechanism Development
for Applications in Atmospheric Simulation

By

Kenneth L. Demerjian
State University of New York

Presented at the Environmental Protection Agency

Workshop on Evaluation and Documentation of Chemical
Mechanisms Used in Air Quality Models

The Mechanism Development Process

The development of chemical mechanisms for the purpose of representing the transformation of pollutants and background trace constituents in the atmosphere has evolved over approximately a twenty year period and has considered several principal components in the development process. The components that form the basis for the methodology of this development process are illustrated in Figure 1.

Historically, each of these components has assumed a dominant role in the mechanism development process, somewhat reflecting the state of the science during the various development stages. For example, in the early 1970's at the onset of research and development activities in chemical mechanisms for the simulation of atmospheric transformations in polluted environments (Demerjian et al., 1974; Niki et al., 1972;), a majority of the elementary reaction steps were generally theorized from thermochemical kinetic estimates based on the methods introduced by Benson in 1968. Mechanisms were developed using data from smog chamber experiments as a basis set for truth, that is, mechanisms were judged on their success in fitting the concentration - time profiles of the experimental data. The assumption being that the smog chamber experiments provided a representative analogue of the chemical systems operative in real world polluted atmospheres, therefore allowing the extrapolation of the developed mechanisms to simulating the chemical transformations in polluted atmospheres.

Many critical elementary reaction steps were identified in the process described above for which no laboratory chemical kinetic data existed. The importance of these reactions in understanding the mechanistic transformations of pollutant species in the atmosphere created a forcing function which stimulated laboratory chemical kinetic studies. Rate parameters and elementary modes of reaction for a large variety of species and reactions were provided by the chemical kinetic community which introduced significant advances and refinements in mechanisms in the late 70's and 80's.

As the feedback process above was occurring, so also was a process between the model development and smog chamber communities. The mechanistic modelers attempted from the start to develop and test their mechanisms for as many chamber systems and data sets as possible. In doing

so the modeling community recognized certain limitations in the data bases and chamber systems they were utilizing and initiated some guidelines for the smog chamber experimentalists. This resulted in additional enhancements in the mechanism development and represents an important methodological component in the evolutionary process.

The scientific community has, on a continuing basis, been developing and evaluating mechanisms against smog chamber data sets. As the science of chemical mechanism development has become more sophisticated, the community's requirements for quality and performance have also become more refined. The feedbacks and intercomparisons that were components of the methodological development process began to identify limitations in the chamber experiments. Scientist began to question the degree of effort one should make to fit individual runs, series of runs, and runs between different experimental smog chamber systems. The debate arising from the fact that many smog chambers whose data sets had been used for mechanism development had not been adequately characterized with respect to wall effects. The chamber walls which can act as both sources and sinks for important chemical constituents introduce noise in the chamber data. This limits the precision and accuracy claims which might be inferred by precise fits of modeled and observed concentration - time profiles.

In addition these limitations are not always explicitly characterizable and are thought to ultimately contaminate mechanisms which have been developed from them. This contamination can take the form of inherent noise or a systematic bias. This phenomenon represents a significant scientific challenge to the community and is the subject of this discussion and the focal point of this convened workshop.

The utilization of atmospheric observations as a vehicle for mechanism development and evaluation is intuitively the most satisfying scientifically. But until recently, the instrumentation technology necessary to characterize the detailed chemical components of the atmospheric system were beyond reach. Also the complexities introduced by the dynamics of the atmosphere introduced considerable uncertainties which make diagnostic interpretations of mechanisms quite difficult. But even with these caveats it would seem that progress in instrumentation technology and the importance of studying real world chemical systems suggests that atmospheric observations of increasing sophistication become a major component in the development of new generation chemical mechanisms.

Application of Chemical Mechanisms

The Environmental Protection Agency's interest in the research and development of chemical mechanisms of polluted atmospheres stems from their responsibility to manage air quality and its associated environmental effects. The chemical mechanisms are a critical component in modeling techniques which provide a quantitative relationship between the emission of chemical precursors which react in the atmosphere both in the gas- and liquid- phases and in sunlight as well as in the dark, to produce chemical species of environmental concern. Currently these compounds include: ozone, nitrogen dioxide, fine particulate matter (primarily sulfates) and acid bearing substances, and most likely the list will expand with time.

A case in point and relevant to the subject of this workshop is the

modeling of the formation of ozone in urban atmospheres. The air quality simulation model incorporates a chemical mechanism in conjunction with emissions information and some treatment of the transport and diffusion of the chemical species under study. The model is exercised to provide quantitative guidance as to the amount of precursor emission control (non methane hydrocarbons and oxides of nitrogen) that is required to meet a specified concentration of ozone. At issue is the precision and accuracy of this quantitative relationship, methods for its evaluation, and standards of acceptability (or success).

Chamber noise discussed in the previous section has important implications on the limits of precision and accuracy of the quantitative relationship desired in regulatory applications. If the chamber effects are unknown or incorrectly specified, the result is a chemical mechanism that has most likely over- or under- compensated the radical production processes in order to achieve an acceptable fit of the chamber data. The application of such a mechanism in a regulatory model will result in a systematic bias in the quantitative relationship between precursor emissions and ozone production.

Scientific versus Regulatory Success

It is quite apparent that the precision and accuracy requirements which would meet the scientific community's standards for success may not be acceptable to the regulatory community. By this I mean that the regulatory application may require better precision and accuracy performance from the mechanism in order for it to be an effective tool for developing quantitative control strategies.

For example, the scientific community in reviewing the various sources of error associated with the chemical mechanism development process anticipates uncertainties of the order of plus or minus 30% in the mechanisms predictive capability when compared with chamber data. Establishing that this is a reasonable error limit when the mechanism is applied under real atmospheric conditions remains a critical issue which must be demonstrated if these approaches are to have any scientific credibility. But more importantly if the 30% uncertainty is reflected in the control requirement of precursors (e.g., non methane hydrocarbons) to meet the ozone standard in a given city, the associated cost differentials can be both economically and socially prohibitive.

The recognition of this uncertainty and factoring it into the quantitative procedures required in state implementation plans to demonstrate a course of action for attaining the national ambient air quality standard for ozone would seem a logical first step to be taken by the regulatory community. The next step should be to establish realistic precision and accuracy performance standards for mechanisms to be used in quantitative models for control strategy development.

Recommendations: Research Agenda to Address "Control" Precision and Accuracy

Working under the assumption that there is a desire to improve the precision and accuracy of chemical mechanisms and their ability to provide quantitative relationships between precursor emissions and ozone production

the following recommendations are presented. It should be noted that though these may not be all inclusive, they are thought to be the most critical. Finally, the recommendations represent an integrated approach which requires each to be carried out in order to successfully achieve the overall objective.

1. Mechanism Mapping - develop procedures for and perform flow diagrams of all mechanisms under research and development or currently used in applications to identify critical nodes for radical initiation, propagation, and termination.

2. Laboratory Studies - perform laboratory studies to reconcile differences in critical nodes between mechanisms which result from lack of knowledge of the chemical details of a reaction process (e.g., elementary reaction steps, fragmentation channels and yields).

3. Smog Chamber Characterization - develop guidelines and standards for smog chamber operations and basic requirements for chamber characterization. The chamber performance characteristics must be documented and submitted to peer review, preferable through publication in the open literature.

4. Smog Chamber Intercomparison Studies - develop guidelines and perform smog chamber intercomparison studies to determine internal bias among systems, variations in artifact effects and overall reproducibility of results between systems.

5. Atmospheric Observations - initiate a program to systematically build an atmospheric chemical observational data base for the purpose of diagnostic interpretation and evaluation of mechanisms.

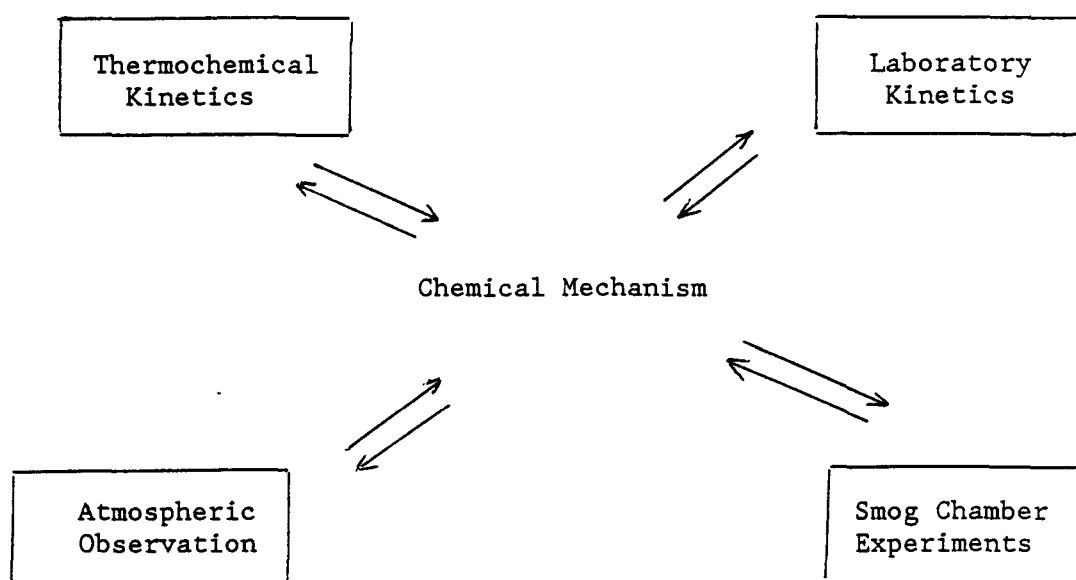
References

Benson, S.W., Thermochemical Kinetics, John Wiley, New York, 1968.

Demerjian, Kenneth L., J.A. Kerr, and Jack G. Calvert, "The Mechanisms of Photochemical Smog Formation", Adv. Environ. Sci. Technol., Vol. 4, 1-262, (1974).

Niki, H., E.E. Daby, and B. Weinstock, Adv. Chem. Ser. 113, 16 (1972).

Figure 1. Components of Chemical Mechanism Development



Paper 4

Development and Testing of Gas-Phase Chemical Mechanisms:
Present and Future Research Needs

Roger Atkinson
Statewide Air Pollution Research Center
University of California-Riverside

DEVELOPMENT AND TESTING OF GAS-PHASE CHEMICAL MECHANISMS:

PRESENT STATUS AND FUTURE RESEARCH NEEDS

Roger Atkinson

Statewide Air Pollution Research Center
University of California
Riverside, California 92521

In a real sense, atmospheric chemistry can be said to date from approximately 1969-1970, when the hypothesis that the OH radical is a key reactive intermediate in atmospheric reactions was proposed (Heicklen et al., 1969; Stedman et al., 1970). It is now recognized that the OH radical plays a major role in initiating organic consumption and, in combination with HO₂ and RO₂ radicals, NO to NO₂ conversion and the formation of ozone.

Since 1970, chemical mechanism development has proceeded in parallel with progress in elucidating the kinetics, mechanisms and products under atmospheric conditions of elementary reactions and in conducting environmental chamber experiments. Table 1 lists a number of chemical mechanism development and review/evaluation efforts which have led to a steady and continual advance in our detailed understanding of the chemical reactions of organic compounds leading to ozone formation. The dates given in Table 1 are those of publication; in most cases these mechanism development or review efforts were initiated several years earlier, as, for example, the Atkinson and Lloyd (1984) review which in its initial form was completed in 1980 and which was used as the basis for formulating the ALW mechanism (Atkinson et al., 1982).

To a large extent only two environmental chamber facilities have been operational over the past few years, at the Statewide Air Pollution Research Center at the University of California, Riverside, and at the University of North Carolina, and only a small number of groups have been active in chemical mechanism development. Despite this and the fact that much of the funding for these programs came from the U.S. Environmental Protection

Table 1. Selected Chemical Mechanism Development Programs

Niki, Daby and Weinstock (1972)	Detailed propene mechanism
Demerjian, Kerr and Calvert (1974)	Review of atmospheric chemical reactions
Hecht, Seinfeld and Dodge (1974) } Durbin, Hecht and Whitten (1975) }	Airshed mechanisms
Falls and Seinfeld (1978)	Airshed mechanism
Hendry et al. (1978)	Detailed alkane, alkene and aromatic mechanisms
Carter et al. (1979)	Detailed propene, n-butane mechanisms
Whitten et al. (1979, 1980a)	Detailed organic mechanisms
Whitten et al. (1980b)	"Carbon-bond" mechanism
Atkinson et al. (1980)	Detailed toluene mechanism
McRae et al. (1982)	Airshed mechanism
Atkinson, Lloyd and Winges (1982)	"Detailed mechanism" for floating-box applications
Killus and Whitten (1982)	Detailed toluene mechanism
Atkinson and Lloyd (1984)	Review/evaluation of the detailed atmospheric chemical mechanisms of eight hydrocarbons
Leone and Seinfeld (1984)	Detailed toluene mechanism
Leone and Seinfeld (1985)	Detailed "master" mechanism and comparison of then current mechanisms
Lurmann, Lloyd and Atkinson (1986)	Detailed and condensed mechanisms, for long-range transport/acid deposition applications
Stockwell (1986)	Airshed mechanism for acid deposition application
Carter et al. (1986)	Detailed chemical mechanisms

Agency, in retrospect the development of chemical mechanisms does not appear to have been a totally coordinated effort.

The fact that the two major reviews concerning the chemical reactions occurring in the atmosphere have led to continued chemical mechanism development points to the crucial need for an ongoing effort to critically review and evaluate the available kinetic, mechanism and product data for the elementary inorganic and organic reactions occurring in the troposphere. This aspect of chemical mechanism development, and of tropospheric chemistry in general, has received short shrift to date. The stratospheric chemistry community have had the NASA and, until recently, CODATA evaluation panels on a continuing basis. CODATA has terminated its evaluation effort, and this has been picked up by IUPAC with a slant towards tropospheric chemistry -- with organics up to C₃ (propane, propene and their atmospheric degradation products) being included.

A major problem in providing a review and evaluation program for organics of the complexity needed for regulatory purposes is the lack of basic data -- I estimate that of the hundreds of elementary organic reactions involved in the latest Carter et al. (1986) mechanism, actual kinetic or product data exist for no more than 10-20% of these reactions. Clearly, speculation and argument by analogy plays a major role, and review evaluation efforts should include "directions" for such data gaps, utilizing thermochemical arguments and any other theoretical techniques available.

PROTOCOL FOR CHEMICAL MECHANISM DEVELOPMENT

In my opinion, the development of a chemical mechanism for use in ambient atmospheric calculations must follow the following path:

(a) Assembly of a detailed chemical mechanism, whether it be of the "carbon bond" approach or the "representative species" approach, which is totally consistent with our current understanding of the elementary chemical reactions which occur under atmospheric conditions. This chemical mechanism, or list of reactions, hence utilizes the best kinetic, mechanistic and product data base available from experimental laboratory and theoretical studies.

After assembling a detailed chemical mechanism (or modifying/updating an existing mechanism) which is consistent with the available kinetic, mechanistic and product data base, this mechanism must be tested against a relevant data base. These

data are not those for elementary reactions, but rather they concern the overall effects of complex reaction sequences.

These observed effects involve, for example, organic consumption rates, NO to NO₂ conversion rates, ozone formation and product formation rates and yields, and may be termed "global" data. In theory, these can be obtained either from environmental chamber irradiations of NO_x-organic-air mixtures or from ambient air measurements. Neither of these data bases are free from problems, but at the present time I think it is clear that the environmental chamber data base involves the least number of variables to be taken into account.

To test (but not validate) a chemical mechanism against environmental chamber data, a number of extra reactions to take into account chamber effects need to be added to the list of reactions in the chemical mechanism. These include:

- Wall decay/offgassing of O₃, NO₂, H₂O₂, HNO₃, N₂O₅ and other species.
- Heterogeneous formation of HONO from NO₂.
- Formation of "chamber-dependent" radicals, probably of OH radicals from the photolysis of heterogeneously formed HONO.
- The spectral distribution and light intensity of each chamber need to be accurately represented.

These "chamber-dependent" effects all introduce uncertainties which affect mechanism testing.

Clearly, there is a need for more work to be carried out concerning these chamber effects, at two levels:

- Characterization or parameterization of chamber effects. For example, the chamber radical source is presently parameterized as a flux of OH radicals, with the magnitude of this flux being determined from associated chamber control/characterization experiments.
- Fundamental understanding of the problem: for example, understanding the detailed chemical formation routes which give rise to the chamber OH radical source.

It should be recognized that chemical mechanism testing is actually carried out under limited NO_x and hydrocarbon concentration regimes/ratios, which are not often representative of ambient atmospheres, as shown schematically in Fig. 1. If at

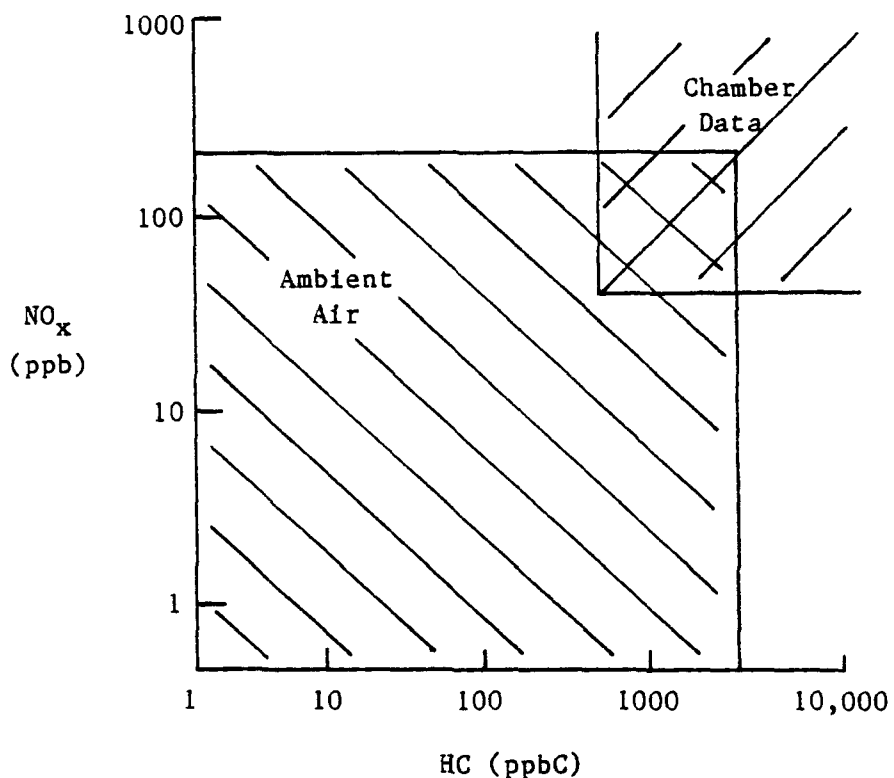


Fig. 1. NO_x and hydrocarbon concentration regimes which are representative of ambient atmospheres and of environmental chamber data.

all possible, the chemical mechanism should be tested against single organic data, for example, against NO_x -air irradiations of HCHO , CH_3CHO , other carbonyls, alkanes, alkenes and aromatic hydrocarbons in a hierarchical manner. Testing should then continue with NO_x -air irradiations of mixtures. This stepwise testing procedure should be carried out since compensating errors can allow complex mixtures to be fit well, but not simple single organics.

Finally, we obtain a detailed chemical mechanism which has been compared/tested against a "global" data base, for example, environmental chamber data. For use in an airshed computer model we must then:

- Condense the mechanism to fit within the limits of the computer resources.
- Remove the chamber-effects reactions.

Condensation of the chemical mechanism should be carried out by a gradual procedure, testing the condensed versions at each step of the process against the fully detailed mechanism over a wide range of organic/ NO_x concentrations and ratios. Retesting against environmental chamber data is not sufficient for any major condensation of the chemical mechanism.

Chemical mechanisms must be well documented, both in terms of the chemistry upon which the detailed mechanism is based as well as the testing against environmental chamber data and subsequent condensation. This includes a complete documentation about how chamber effects and light intensity and spectral distribution were handled.

However, it is not totally obvious that the removal of certain of the chamber-effects is correct. Thus, the major chamber effect in terms of the reactivity of environmental chamber irradiations concerns the "chamber-dependent" radical source. This probably originates from a photo-enhanced heterogeneous wall formation of HONO from NO_2 . The rate coefficient for the dark formation of HONO at 50% relative humidity and 298 K in ~6000 liter volume environmental chambers is surprisingly similar to the HONO formation rate coefficients derived from ambient atmospheric measurements of nighttime HONO. This suggests that there may well be an atmospheric OH radical (or HONO) source which, for a given NO_2 concentration, is similar in magnitude to those needed for chamber simulations. If indeed there is an atmospheric heterogeneous source of OH radicals, this presumably could have significant implications on control strategy decisions.

To summarize, gas-phase chemical mechanism development requires close collaboration and parallel research efforts in

- Basic gas-phase kinetic, mechanistic and product studies carried out under laboratory conditions.
- An ongoing data evaluation and review process of the kinetics, mechanisms and products of elementary reactions of importance to atmospheric chemistry.
- Research studies to elucidate several significant environmental chamber effects.
- Environmental chamber studies.
- Computer modeling and chemical mechanism development, with this research area providing the critical input to identify the basic kinetic, mechanistic and product data and the environmental chamber/chamber effects data required.

Only a cohesive approach will provide advances in the "accuracy" of chemical mechanisms in the next few years.

At the present time, we are data limited with respect to both basic kinetic, mechanistic and product data and high quality environmental chamber data (including high quality characterization data). Indeed, further computer modeling programs utilizing our present data base are a waste of both time and effort.

FUTURE RESEARCH NEEDS

Basic Data

Vast advances in our understanding of the chemistry occurring in the atmosphere have been made over the past 15 years. However, because of funding constraints the rate of advance has markedly slowed since the EPA-sponsored workshop "Chemical Kinetic Data Needs for Modeling the Lower Troposphere" was held in Reston, VA, in 1978. Indeed, for example, the limited amount of experimental work carried out since then concerning the atmospheric chemistry of aromatic hydrocarbons has led to an advance in our actual knowledge of their chemistry, but has shown us how little we still know about the atmospheric chemistry of this class of compounds. The same is true for the long-chain alkanes.

We need (a) basic mechanistic and product data (under atmospheric conditions) for most classes of organics emitted from anthropogenic and biogenic sources. This includes, but is not limited to, aromatic compounds and the larger alkanes, and (b) an integral and ongoing review and evaluation effort, either as a stand-alone effort for chemical mechanism development or as a "piggy-back" onto other presently ongoing evaluation efforts.

Environmental Chamber Data Needs

The environmental chamber data necessary for future chemical mechanism development include the following:

- A review/evaluation of chamber-effects is needed, to determine the nature of chamber effects and how they should be taken into account.
- Further basic research is needed to elucidate the chemistry/physics of certain of these chamber effects.
- High quality chamber data, combined with high quality characterization/chamber effects data, are needed. For the

upcoming generation of chemical mechanisms, such species as HONO, HCHO, NO₂ (not as NO_x-NO), HNO₃ and H₂O₂ must be measured in real-time. Light intensity and spectral distributions must also be measured accurately.

- Chamber data under low NO_x-organic concentration conditions, and in the absence of NO, are necessary for mechanism development for long-range transport and multi-day applications.

In addition, ambient air data are needed to aid in chemical mechanism development. A prime example concerns whether or not OH is formed in the atmosphere from NO₂, as it is in environmental chambers.

At the present time, we are at a critical juncture. Thus, the present generation of chemical mechanisms are close to being as up-to-date with respect to the available data base as is possible, and any further advances in the predictive accuracy of urban and regional airshed computer models will require a renewed, closely coordinated, effort by research groups with diverse interests as well as by funding and regulatory agencies.

REFERENCES

- Atkinson, R., Carter, W. P. L., Darnall, K. R., Winer, A. M., and Pitts, J. N. Jr., 1980, Int. J. Chem. Kinet., 12:779.
- Atkinson, R., Lloyd, A. C., and Wines, L., 1982, Atmos. Environ., 16:1341.
- Atkinson, R. and Lloyd, A. C., 1984, J. Phys. Chem. Ref. Data, 13:315.
- Carter, W. P. L., Lloyd, A. C., Sprung, J. L., and Pitts, J. N., Jr., 1979, Int. J. Chem. Kinet., 11:45.
- Carter, W. P. L., Lurmann, F. W., Atkinson, R., and Lloyd, A. C., 1986, "Development and testing of a surrogate species chemical reaction mechanism," Final Report to EPA Contract No. 68-02-4104, May.
- Demerjian, K. L., Kerr, J. A., and Calvert, J. G., 1974, Adv. Environ. Sci. Technol., 4:1.
- Durbin, R. A., Hecht, T. A., and Whitten, G. Z., 1975, "Mathematical modeling of simulated photochemical smog" EPA-650/4-75-026, June.
- Falls, A. H. and Seinfeld, J. H., 1978, Environ. Sci. Technol., 12:1398.
- Hecht, T. A., Seinfeld, J. H., and Dodge, M. C., 1974, Environ. Sci Technol., 8:327.
- Heicklen, J., Westberg, K., and Cohen, N., 1969, Report No. 115-69, Center for Air Environment Studies, Pennsylvania State University, University Park, PA.

- Hendry, D. G., Baldwin, A. C., Barker, J. R., and Golden D. M., 1978, "Computer modeling of simulated photochemical smog." EPA-600/3-78-059, June.
- Killus, J. P. and Whitten, G. Z., 1982, Atmos. Environ., 16:1973.
- Leone, J. A. and Seinfeld, J. H., 1984, Int. J. Chem. Kinet., 16:159.
- Leone, J. A. and Seinfeld, J. H., 1985, Atmos. Environ., 19:437.
- Lurmann, F. W., Lloyd, A. C., and Atkinson, R., 1986, J. Geophys. Res., 91:10905.
- McRae, G. J., Goodin, W. R., and Seinfeld, J. H., 1982, "Mathematical modeling of photochemical air pollution." Final Report to California Air Resources Board Contract Nos. A5-046-87 and A7-187-30, April.
- Niki, H., Daby, E. E., and Weinstock, B., 1972, Adv. Chem. Series, 113:16.
- Stedman, D. H., Morris, E. R., Jr., Daby, E. E., Niki, H., and Weinstock, B., 1970, 160th National Meeting of the Am. Chem. Soc., Chicago, IL, Sept.
- Stockwell, W. R., 1986, Atmos. Environ., 20:1615.
- Whitten, G. Z., Hogo, H., Meldgin, M. J., Killus, J. P., and Bekowies, P. J., 1979, "Modeling of simulated photochemical smog with kinetic mechanisms." Vol. I, Interim Report EPA-600/3-79-001a, January.
- Whitten, G. Z., Killus, J. P., and Hogo, H., 1980a, "Modeling of simulated photochemical smog with kinetic mechanism." Vol. I, Final Report EPA-600/3-80-028a, February.
- Whitten, G. Z., Hogo, H., and Killus, J. P., 1980b, Environ. Sci. Technol., 14:690.

QUESTIONS

1. HOW DO WE DEVELOP CHEMICAL MECHANISMS WHICH ARE SCIENTIFICALLY RIGOROUS?

- SHOULD BE BASED UPON THE BEST KINETIC, MECHANISTIC AND PRODUCT DATA AVAILABLE.
- SHOULD BE TESTED AGAINST AVAILABLE ENVIRONMENTAL CHAMBER DATA, USING A CONSISTENT METHOD TO TAKE INTO ACCOUNT CHAMBER EFFECTS.
- MUST HAVE ADEQUATE DOCUMENTATION OF THE CHEMICAL MECHANISM AND RESULTS OF TESTING AGAINST ENVIRONMENTAL CHAMBER DATA, INCLUDING TREATMENT OF CHAMBER EFFECTS.

2. WHAT IS THE NEEDED FUTURE RESEARCH WHICH WILL
ENABLE US TO NARROW THE PRESENT UNCERTAINTIES IN
CHEMICAL MECHANISMS?

- NEED FURTHER LABORATORY PRODUCT AND MECHANISTIC DATA, ESPECIALLY FOR AROMATIC COMPOUNDS AND LARGER ALKANES.
- NEED ONGOING CRITICAL REVIEW AND EVALUATION OF THE KINETIC, MECHANISTIC AND PRODUCT DATA BASE.
- NEED FUNDAMENTAL STUDIES INTO ASPECTS OF CHAMBER EFFECTS, SUCH AS THE CHAMBER RADICAL SOURCE.
- NEED EVALUATION OF CHAMBER EFFECTS AND HOW TO DEAL WITH THEM FOR ALL CHAMBERS IN A CONSISTENT MANNER.

- NEED FURTHER MORE DETAILED AND WELL-PLANNED QUALITY ENVIRONMENTAL CHAMBER DATA, WITH REAL-TIME ANALYSIS FOR:

HONO

HCHO

NO₂ (not NO_x-NO)

H₂O₂

HNO₃

AT A MINIMUM.

LIGHT INTENSITY AND SPECTRAL DISTRIBUTION MUST BE ACCURATELY MONITORED. QUALITY, NOT QUANTITY, IS NEEDED.

- ALL OF THESE EFFORTS SHOULD BE COORDINATED, FROM THE VIEWPOINT OF CHEMICAL MECHANISM DEVELOPMENT, WITH THE CHEMICAL MECHANISMS AVAILABLE BEING USED TO DEFINE THE AREAS OF NEEDED INPUT DATA.
- NEED AMBIENT AIR STUDIES - FOR EXAMPLE, DOES THE CHAMBER-RADICAL SOURCE MECHANISM PRODUCING OH RADICALS OCCUR IN AMBIENT AIR AND TO WHAT EXTENT?

Paper 5

**Review of the "Strawman" Document for the EPA Workshop on
Evaluation/Documentation of Chemical Mechanisms**

**Michael W. Gery
Systems Applications, Inc.**

Review of the "Strawman" Document for the EPA Workshop on
Evaluation/Documentation of Chemical Mechanisms

Michael W. Gery

Systems Applications, Inc.
101 Lucas Valley Rd.
San Rafael, CA 94903

Introduction

In the following review of the Jeffries and Arnold document, I shall take a somewhat different approach than that taken by the authors in describing what I feel to be the essential workshop goals concerning areas of uncertainty and the need for documentation in the model development, evaluation and application practices. Initially, I shall present a review of the important points in the strawman document with respect to historical development of our science. Following this I shall describe the current status of what I perceive to be our present modeling methodology, developing out of this a discussion pertaining to uncertainties in approach and data, future needs and options, and possible courses of action.

The Current Tension and the Approach

There appear to be at least three distinct views concerning the degrees of validity in photochemical kinetics modeling, each dependent on the needs of the respective groups who utilize these models in different ways. At the policy level, as exemplified in the authors' quotes from the GAO, legislators want to know how reliable different modeling tools are in an effort to justify the complex decisions regarding issues of the Clean Air Act. The authors point out that since the quantification of model validity based on certainty in well-grounded principles is difficult, vindication may be a better description of the of the actual requirement for model performance at this level. Within the Agency, however, costly control strategy decisions must be made based on predictive estimates. The need for determining the relative validity of different chemical mechanisms based on scientific evaluation is, therefore, a necessity. Unfortunately, because the scientists who develop these models cannot describe all real situations with absolute precision, inherent uncertainty often masks the differences between models. This conflict, between scientists with goals of absolute validity and regulators who need to use the "best" tool currently available, creates a tension resulting from the what the authors describe as a "forced realization of the paradigm." That is, the photochemical kinetics models are needed to make predictions beyond the bounds of certainty within which the scientists presently feel secure. Hopefully, however, because this is a continually improving process, it may be possible to describe the types and measures of uncertainties which exists at present, in an effort to both minimize these error bands and prescribe tests which could detect mechanisms which do lie beyond the realm of what is presently acceptable. I feel that the best way to do this is to examine each process in our current modeling methodology, defining the structure and associated uncertainties in detail. This paper is based on such an approach.

As noted above, Jeffries' and Arnold's description of the historical and technical progression in photochemical kinetics mechanism development has provided us with a review of those works that set the scientific

foundation for the current practice of our science. This description of the development and status of our normal science allows us to look at the structure and rules for various distinct processes which make up the paradigm-as-set-of-shared-values. That is, our methods, standards and ways of solving problems. This is critical to our present needs for two reasons; first, because it is necessary to understand the chemical model development and evaluation processes that we currently employ to modify or enhance them, and secondly, because a review of past work provides us with descriptions of practical aspects of methodologies in our field which are often overlooked or taken for granted because the purely analytical results from the earlier works have been superseded. Therefore, I will briefly review the historical development of the basic paradigm-as-set-of-shared-values within the philosophical descriptors set up by Jeffries and Arnold. I will later describe individual processes or methodologies to isolate the current status in the form of basic rules and current uncertainties of each. It is these processes that involve the subjective decisions by the model developer with respect to process and data uncertainty and are therefore encumbered with uncertainty requiring validation. The approach I take is aided by a diagram (Figure 1) which illustrates the historical development of these processes and can later be used as a basis for discussion. Obviously, such a diagram represents my idealized conception of the current model development methodology.

The Historical Development of Mechanism Development Methodologies Through the Eyes of Jeffries and Arnold

The preparadigm period is described by Jeffries and Arnold as that time when the need to describe the atmospheric chemistry of ozone formulation was apparent but the science of air pollution chemistry was lacking in its ability to explain the reasons for the chemical transformations resulting in the formation of secondary pollutants. Collections of chemical reactions along with some kinetic and mechanistic information existed, but the relative significance of individual reactions or cycles was not yet elucidated. This pool of information, which has

grown significantly since the preparadigm period, is represented by the top box of Figure 1. This data is drawn upon in the initial compilation (and later reevaluations) of a photochemical kinetics mechanism.

The first mechanistic descriptions were developed prior to 1970, but many of the steps were empirical, lacking any real physical or chemical meanings. About 15 years ago, an essential "revolution" occurred with the initial mechanistic descriptions of hydroxyl radical chemistry and its role in organic oxidation through the OH-HO₂ cycle. Jeffries and Arnold coin this the beginning of the "OH-paradigm," and suggest that because this description and its attendant auxiliary hypotheses (coupled with the possibilities of computer simulation) were more successful at solving the problems then recognized, rapid Gestalt switches made quickly by many scientists practicing in different fields. Because of these developments, the first successful mechanisms that were based on "well-founded scientific principles" were put forth. The paradigm had taken a big step in that a paradigm-as-set-of-shared-values was clearly developing. In fact, sets of intuitive rules, many of which still govern aspects of mechanism development today, were postulated at that time. These rules were often put forth as a method of limiting uncertainty through the normalization of scientific process.

This process of compiling a chemical kinetics mechanism through the evaluation of the existing pool of data, combined with needed assumptions is depicted in Figure 1 as the data evaluation process. I will discuss this process more below; however, it is important to note that the formulation of a mechanism from the pool of data is a "mapping process." Hence, some measure of ambiguity is inherited from the generalization, deletion or distortion of reality which accompanies the creation of a map. Of these, deletion, especially deletions which occur because missing reactions were simply unanticipated and therefore unstudied, is the most difficult to identify and deal with.

The authors note that one of the early applications of a mechanism developed through this process was the Niki, Dalby and Weinstock (NDW)

mechanism. In their work, NDW simulated older data from the irradiation of mixtures of propylene and NO_x . From these simulations, which utilized the newly developing tool of computer modeling, it became apparent that additional tools and data were needed to further the science. This was, however, an example of an initial application of such a mechanism. It supported the OH-paradigm as an explanation of the chemical dynamics in the smog chamber experiments. Jeffries and Arnold refer to this occurrence as the first class of Kuhn's "normal science, where the model based on the OH-paradigm is used to reveal the nature of things as described by a paradigm that is becoming more worthy of testing and development.

The first well known example of normal science following the adoption of the OH-paradigm was the Demerjian, Kerr and Calvert (DKC) mechanism development and application. DKC were careful to document their method of mechanism development, and they presented many of the formulation and testing rules that are the basis of our current paradigm-as-set-of-shared-values. They noted the need for a good pool of kinetic and mechanistic information, their evaluation of the existing data was a solid foundation for the mechanistic development effort. One particular extension of the normal science of mechanistic development was the clear utilization of smog chamber data to refine assumptions inherent in the mechanism. This process is shown in Figure 1 as a feedback loop of evaluation and verification of the current mechanism using "real" data; in this case, smog chamber data. It is, in effect, the altering of the mechanistic map through comparisons with a new set of "truth" (the smog chamber data) to minimize distortions and identify unintended deletions in the original map. Thus, while most earlier studies that attempted to simulate chamber data were mainly concerned with verifying the paradigm, DKC also used simulations to articulate and strengthen the paradigm by resolving ambiguities. DKC saw the need to build on the foundation of basic data with new information derived from smog chamber simulations. The authors say that DKC implied the need for smog chamber data to evaluate "alternative mechanisms and reaction rate constants ... even though they [DKC] did not explain how this occurred." In their discussion of the DKC

"asymmetry for the treatment of experimental fact," Jeffries and Arnold state, however, that while the chamber data was used to refine the mechanistic map, it was not considered a requirement that the mechanism "fit" all data. It should be noted that the amount and quality of smog chamber data was limited at this time. While DKC could not yet see the utility of smog chamber validation with the data they had, they did use their model in a way that manifested the need for the better data that now exists, and the validity tests that can now be employed.

The 1974 work of Hecht, Seinfeld and Dodge (HSD) is discussed next. HSD reaffirmed and added to the description of the model development and verification processes described above, and also clearly identified the need for both fundamental kinetic data and smog chamber data. That is, evaluation of basic kinetic data and simulations versus smog chamber data are two very compatible processes which, when treated equally, minimize a great deal of uncertainty associated with the mapping process of mechanism development. The basic kinetic and mechanistic data contains the descriptions of conditional chemical variability upon which atmospheric simulations are based and legitimized. The smog chamber data provides a real, controlled, although somewhat limited, test bed for detecting deletions or distortions in the basic data. Hence, the iterative adjustment of mechanistic assumptions through smog chamber simulation is a way to link the strengths of both types of data (fundamental kinetic and smog chamber) so that uncertainties in each may become obvious through comparison with the other.

A second key feature of the HSD work developed out of their different objective. The authors state that, "While a prime objective for DKC was to have 'a complete reaction scheme' and thus to explain, the prime objective for HSD was "that the mechanism predict the chemical behavior of a complex mixture of many hydrocarbons, yet that it include only a limited degree of detail' which would allow the mechanism to be applied to the 'prediction of both smog chamber reaction phenomena and atmospheric reaction phenomena....This work, therefore, in its conception and definition, set out toward a different end than the DKC work: DKC sought

scientific explanation; HSD wanted prediction." The need for prediction, specifically the eventual need for prediction of atmospheric chemistry, resulted in the formulation of two additional processes: (1) condensation of the mechanism (computers were and still are limited by the large number of calculations necessary to run complete chemical kinetics simulation code), and (2) verification that the predictions of the mechanism are valid in the physical and chemical ranges of the predictive application.

These processes are graphically presented in Figure 1. Note that there are two apparently parallel pathways leading to a verified, condensed mechanism. One describes the condensation of an explicit mechanism that has first been tested versus smog chamber data, while the other implies the production of an explicit mechanism from "first principles," followed by evaluation versus experimental data. In both cases, the evaluation against smog chamber data is still required as an essential test against real data to establish the completeness of the explicit compilation. It is my personal preference to proceed along the path which provides a validated explicit mechanism prior to condensation, since: (1) compensating errors in the mechanism cannot be masked by condensation prior to testing versus real data, and (2) a complete explicit mechanism that has been verified against all available data can be used to produce a number of condensed mechanisms at different levels of condensation and for different applications. The alternate pathway requires a new condensation for every iterative step in the smog chamber evaluation loop, resulting in a more complicated task to produce the same results.

As shown in Figure 1, the need for a simplified mechanism to predict data collected at contaminant monitoring stations in an urban area is a drastic variation on a complete reaction scheme used only for explanatory purposes. Recognizing this, HSD said, "a kinetic mechanism, once developed, must be validated." Hence, the nature of this workshop. However, it is important to note that the validation requirement originally derived from the HSD need for atmospheric simulation and prediction, and this should be the focus of the workshop as well as the initial makeup of and verification of the initial explicit mechanisms.

The Current State of "Normal Science"

In the above section I have developed an idealized structure of the paradigm-as-set-of-shared-values that I believe we tend to follow in our current version of normal science. I would now like to discuss the present status of these processes by isolating each and considering the current "rules" and the remaining sources of uncertainty. If we use Figure 1 as an idealized representation of the mechanism development process, we see that the boxes in the schematic represent real (although sometimes temporary) data, information, compilations, mechanistic versions; in short, they are starting, nodal and stopping points. Listings of data, results and mechanisms can be produced at any time to represent the current status at a point. The ovals are processes, often requiring sufficient ingenuity, test data and validation to proceed through. These processes are the sources of uncertainty which must be clearly described in the formulation of a predictive chemical kinetics mechanism from basic data. Because of this they are also the areas which must be clearly documented in each application. The amount of uncertainty attained at each process, the past methods derived to limit uncertainty, and the future prospects for better data and clear tests are the topics for discussion at this workshop.

The evaluation of existing chemical kinetics and mechanistic data is clearly a necessity in the initial development and periodic updating of chemical kinetics mechanisms. As noted above, this process was established in the comprehensive reviews in the NDW and DKC papers. The authors observed that one of the primary goals of DKC was, "to evaluate the various alternative mechanisms and reaction rate constants proposed in view of the best kinetic data in hand today."..Two of the rules DKC used to normalize this process and minimize uncertainty are reported by the authors as: (1) mechanism reactions must be elementary reactions, and (2) rate constants must be based on experimental determination of rate constant or must be estimated using thermodynamic techniques. In other words, given well founded data and reasonable estimation techniques, use them; rate constants cannot be considered adjustable parameters.

The compilation of existing data is far more objective than another of DKC's goals, "to identify some of those potentially important reactions in the mechanism of photochemical smog formation for which there was insufficient basic kinetic data to allow for a realistic judgment of their importance." This enters the more subjective realm of mechanistic mapmaking as described above). Because the methods of making data evaluations and assumptions are based on individual understanding and intuition, different mechanism developers will make different choices based on similar information or lack of information. Thus, the errors of the mapping process, distortion and overgeneralization of basic information, combined with unanticipated deletion, will occur. As noted earlier, however, the second process (mechanism evaluation versus real data) allows some of these errors to be detected.

Although an excellent compilation of the current pool of information is almost always produced or referred to by mechanistic developers (for instance, the compilation by DKC, the reviews by Atkinson and coworkers, Kerr and Calvert, and the NASA and CODATA evaluations), the information of how or why certain decisions were made in the development of a mechanism from these data is often poorly conveyed. Instead, an explicit mechanistic description is simply asserted as a listing which reflects the results, but not the reasons, for choices made by the authors. I hasten to add that this is not the case in all work, but only a trend apparent from this paper. This occurrence necessitates a third rule; the mapping decisions, as well as the sources of data, must be well documented to provide a more viable base of formulation. Such a description will aid in the future location of anomalies and help to identify the range of applications for which the mechanism was originally based.

If there is such a measure as goodness of fit for this process it is through the peer review process. I see no problem with this method, certainly the reviews published to date have followed a trend of rigorous evaluation, and more importantly, clear delineation between actual measurements and intuitive estimates. What does contribute to the overall uncertainty in the model development is the relative errors associated

with the kinetic and mechanistic data available and the unavailability of certain data. The solution to this problem is primarily related to the amount and direction of research. Periodic reevaluation and compilation of extensive reviews directed at the chemistry of air pollution must be enacted to make the wide range of data generally available and deciphered. These reviews require funding and, to some degree, guidance toward desired goals. They are not necessary, however, if no new data is available to review. Hence, a major method of reducing overall uncertainty in chemical mechanisms, especially in areas of new interest and little information, is to support new research into elementary chemistry in those areas. As more data comes in and the pool of information is enhanced, many of the individual choices which contribute to uncertainty in mechanism development are unnecessary and the map is a more formidable description of reality.

The evaluation/adjustment of the current explicit mechanism using smog chamber data is the second process described in Figure 1. The need for such an approach has already been stated. Basically, smog chamber experiments provide realistic data, which inherently include tests to identify unintentionally deleted chemistry or overgeneralized or distorted mechanisms. As with most evaluation processes, this is iterative in form, resulting at times in what the authors describe as "knob twittering." Because of the large number of variables, data and subjective decisions, this process probably provides the greatest accumulation of uncertainty in the entire mechanistic development and application practice. On the other hand, it must be noted that elimination of such a verification against real data would most certainly yield far greater uncertainty in any resulting mechanism.

Although the process is shown as an iterative loop (Figure 1) with one decision block, the process is really a multi-step operation (shown in Figure 2) involving: (1) the mechanical testing of model predictions using the current mechanism versus data, (2) the more subjective evaluation of comparison results, and (3), the intuitive process of adjustment of the mechanism to provide "better" predictions. This process

has evolved over the last two decades as more and higher quality data has become available. However, we are still confounded by the basic dilemma of when to exit the loop. This dilemma results from the uncertainty in two areas: (1) how good is the data used in the evaluation, and (2), how does one measure "goodness of fit" between mechanism prediction and data (and thus, decide he has reached "reasonable agreement" and terminate the adjustment process)? Clarification of this dilemma has stimulated the formulation of many of the rules and methods that I will discuss next; however, though these methods have improved our ability to measure reasonable agreement (or in some cases, locate unreasonable agreement), the decision as to when the fit is good enough is still ill-defined. Because of this, it is not unlikely that the actual number of iterations may sometimes be based on the level of remaining funding rather than any reasonable method of testing. This is unfortunate and must be a primary focus of the workshop. It is probably not possible to define the necessary requirements for a mechanism to be considered "valid", but it is possible to establish some minimum requirements needed for the claim of reasonable agreement.

It is important to minimize the error associated with data variability in the test data and to clearly define an evaluation program that will enhance data evaluation as well as prove the strength of a mechanism to the scientific community. One way this is done is through the formulation of basic rules of evaluation and adjustment. One such rule, noted early on by DKC and HSD, is that it is essential to establish a strong data set prior to evaluation of and adjustments to the mechanism. This should include: (1) data from different sources, (2) a variety of hydrocarbons studied, (3) single hydrocarbon and mixture experiments, and (4), different initial conditions (including a broad HC/NO_x ratio). When adjustments are made to the mechanism or when test data are missing, these values must be supplemented in a rational way. It is also critical that the conditions and characteristics of each smog chamber be defined or estimated, especially chamber artifacts that differ from ideal. Although I will not discuss specific chamber artifact representations, it should be noted that the uncertainty developed from

different hypotheses concerning these effects must be overcome prior to any meaningful comparison of mechanisms. This should be a primary concern of the workshop. The reactions that are chamber dependent should be consistently applied within their definitions for all simulations attempted in each study, and hopefully, with the future agreement of the modeling community, this consistency will result from a clearer understanding of these processes. Of course, these assumptions, plus all other definitions and estimates, must be documented.

Since I was asked to specifically comment on the hierarchy of chemical species concept developed by Whitten, I will extend this section concerning design of the evaluation process. It is logically intuitive and operationally clear that the process of mechanism evaluation against smog chamber data be based on a stepwise hierarchy of simulations. Such an approach was first suggested by DKC in their study methodology when they noted that the understanding of smog chamber data was best revealed by starting with simple systems and increasing the complexity of the hydrocarbon components stepwise. Whitten later put forth the principles of the hierarchical approach, a method of model development and testing intended to clarify (or identify) the sources of uncertainty in simulations. Figure 3 shows a hierarchy of species for the Carbon Bond X mechanism. The concept is that one should validate mechanisms in a stepwise order starting from the lowest level first. Once acceptable agreement between simulation results and measurements is obtained, changes in rate constants and reaction stoichiometry for the already tested part of the mechanism is prohibited and the simulations increase to higher hierarchical levels. If disagreement between simulation results and measurements occurs at a higher hierarchical level and all lower species data have already been simulated successfully, it is most likely that the new chemistry for the higher level species is in error and not the whole mechanism. Such a methodology minimizes the possibility of fortuitous agreement between simulations and measurements from the effect of compensating errors which can occur in a mechanism created directly from "first principles." Jeffries and Arnold point out that the hierarchical approach narrows the macroscopic facts of the smog chamber so that they

more clearly confront the theory. The opposite method, direct creation of a mechanism to simulate a mixture of hydrocarbons and NO_x in the atmosphere, without the benefit of simulation of the products of the hydrocarbon mixture, is a less certain process because of the possibility of mechanistic compensating errors.

Because this process appears to possess such a large degree of uncertainty, it should be a direct concern of the workshop. Hopefully, we can take initial steps to improve both the standard test data and the measure of "reasonable agreement". We can reduce the uncertainty in the standards of mechanistic comparison (smog chamber measurements) by agreement on a good set of basic test data, consisting of many smog chamber experiments from different facilities for at least some of the lower hierarchical level species. This will provide a minimum requirement for photochemical kinetics models to fulfill. That is, the initial goals of a mechanism should be to handle time varying photolysis rates and the chemistry of basic inorganic and simple organic species. Because of its uniform nature, such a dataset may also allow initial attempts at the clarification of a measure of goodness of fit. More complex test sets could be agreed upon. However, these sets must be carefully conceived to account for the mechanistic tests that the original experimental design intended to develop. Finally, we must also recognize that we are specifically discussing ozone and oxidant chemistry here. As new areas entered, there is usually little or no experimental data to develop or evaluate a new mechanism with. We must consider new areas and decide whether there are substitutes to smog chamber data which could be utilized in the reality evaluation loop.

The need for prediction applications of photochemical kinetics mechanisms coincided with the requirement that a condensed chemical mechanism be used in place of an explicit version because of the high computational costs. The process by which a condensation technique is verified and a condensed mechanism is formulated is again an iterative process similar to the smog chamber evaluation of explicit mechanisms. It is shown schematically in Figure 4. There are different conditions in

this process, however, which have led to more clearly defined methodologies than in the smog chamber evaluation loop. This is because the object of this process is not to simulate smog chamber experiments correctly, nor, in fact to improve the theory at all, but to duplicate the ability of the explicit mechanism with a more applicable configuration. In fact, the condensed mechanism cannot simulate the chamber results more accurately if all it is based on is simplification of the explicit mechanism since, in terms of the mapping process, it is a more generalized or distorted map with additional deletions.

I will not discuss specific condensation methodologies, however, some basic rules have, again, been laid down to help unify or normalize this process in HSD and later works. These are mainly related to the condensation of oxidant mechanisms. Inorganic reactions are usually treated explicitly, although some deletions occur. Hydrocarbon representation must occur through a logical methodology, cannot be overly simple, and must include a specified method for representing complex atmospheric mixtures in the mechanism. The condensation/evaluation process should be performed in a stepwise or incremental fashion, with stoichiometry and mechanism structure derivable from underlying chemistry, and rate constants related to elementary reactions or within realistic range of measured values. In such a process, which differs with every researcher and application, it is clearly beneficial, both to the developer and the user, to document nearly every step.

In terms of the actual conditions used in the comparison of explicit and condensed mechanism simulations results for the purpose of verifying accurate condensation, it is not always necessary to use real measurements (since there is no attempt in this process to verify that the mechanisms simulate real measurements) although such data sets often provide a good foundation. It is a better idea, rather, to not only verify that the condensed mechanism is an accurate map of the explicit mechanism within the range of physical and chemical conditions available in smog chamber data, but also to test the condensation results in more extreme conditions that will be applied in the intended application. Hence, it may be

necessary to fabricate test conditions to perform such a comparison. Recall that the basic tenet of this process is that "after validation" of explicit chemistry, the results of any simulation using the explicit chemistry can serve as data to be used for the derivation and evaluation of a condensed mechanism. Therefore, the explicit chemistry, although encumbered with uncertainty from the previous development, is viewed here as "truth." Because there is no experimental variability in the test data (the test data are the results of the same explicit mechanism every time), a much tighter measure of goodness of fit must apply in the verification decision of Figure 4.

As noted above, the condensation process inherently requires generalization, distortion and deletion choices. For this reason, the decisions as to the specific condensation approach must be oriented towards the application. Of course, these choices have most often been based on the simulation of urban ozone formation. However, other applications may have different chemical requirements or involve new data sets. In addition, new validation in the previous process loop (reverification of the explicit mechanism), may occur. All of these occurrences necessitate a new condensation, and therefore, new documentation.

The production of a photochemical kinetics mechanism that has been developed and tested with smog chamber data and has been condensed to the size requirements for a specific model, does not necessarily make the mechanism valid for the intended applications. Often it takes a significant leap of faith to accept the predictions of such a model. For instance, one might say that even given ideal fit to all smog chamber tests and a very good condensation technique, the chemical mechanism has only been validated for the chemical and physical conditions exhibited in those data. In addition, any attempt at validation against atmospheric data carries the inconsistencies of the entire model into the simulation results. To take this a step further, while smog chamber validation simulations were probably performed with the Gear algorithm, many large air quality simulation models utilize lower order solution methods that

could induce error unassociated with the chemical kinetics mechanism. This is a formidable set of "ifs", but careful mechanism development combined with an understanding of the sources of data involved in that process can minimize some error and aid in the quantification of the remainder.

The extended atmospheric range of physical and concentration differences must be approached from two directions. First, to the best of our ability, we must ensure that any condensation-type processes are valid over the entire range in the intended application. Second, we must recognize that the description of conditional variation in the chemical mechanism is not based solely on smog chamber data. Rather, while the smog chamber simulations are used to test the completeness of the mechanistic representation, the inherent conditional variations are usually based on kinetic studies with much wider physical ranges than the smog chamber experiences. Therefore, although the explicit mechanism is only verified against real measurements within the range of conditions available in a smog chamber, there is no reason to expect rapid deterioration of chemical description in conditions somewhat beyond those of a smog chamber.

The differences between smog chamber and ambient air concentrations are more dramatic. Reactions or products which were felt to be unimportant, or went unnoticed in higher concentrations smog chamber conditions, may become significant in atmospheric applications. Whereas it is trivial to remove chamber artifact reactions from a chemical mechanism to simulate the atmosphere, it is far more difficult to include reactions which go unnoticed in the chamber but are significant in the atmosphere. For example, because of the orders of magnitude in differences between urban reactive hydrocarbon concentrations and smog chamber initial conditions, background species such as urban CO₂ and methane are of little importance in clean chamber simulations but must be included in atmospheric models. Of course, documentation of the changes needed in a mechanism for proper application to a given situation is critical. However, if we are to clearly understand and separate the

chamber artifacts from the mechanism, better smog chamber experiments which provide test data at concentrations nearer to ambient values must be obtained. This will require new or different facilities that limit artificial processes, coupled with better descriptions of incident light and background chemistry.

The question of the error induced by the solution algorithm used to calculate the changes in concentrations with time for a given modeling application is a peripheral issue, but should be addressed briefly. As noted above, the chemical kinetics mechanisms are often developed with a higher order solution technique which can be utilized because of the physical simplicity of the system. Atmospheric simulations requiring simultaneous, multi-cell solutions of chemistry and meteorology usually use a lower order technique which can develop mathematical errors that translate into inaccurate predictions. It should be the responsibility of the model developer to provide and document comparisons between whatever solution techniques were used to develop the mechanism and represent the chemistry in the application model. The performance of the same photochemical kinetics model over a simulated application range should be compared.

Finally, as noted earlier, not all atmospheric applications have smog chamber datasets which can be used to verify the mechanism against real data (locate deletions or distortions in the mechanistic map). This is becoming clear with the new generation of acidic precipitation gas-phase models. At present, comparison with the sparse set of atmospheric measurements is the only evaluation possible. These models will benefit, to a degree, from advances made for all atmospheric models in the meteorological and solar radiation algorithms, but we may have to face the fact that unless some controlled tests similar to smog chamber experiments are forthcoming, the predictions of these models will be viewed with far more skepticism than the urban ozone models.

Conclusions

When I first considered the present state of the mechanism development and evaluation practice, it appeared to be encumbered by uncertainties to the point where we could not distinguish one mechanism from another. That, of course, is a primary goal of the workshop. After reviewing the present state of our science, however, I believe we must consider this apparent shortcoming in conjunction with the progress we have made in this area. Most of the facts and processes reported in this paper were developed at or after the time when the Dodge mechanism was first implemented in the EKMA. Since then, both the amount of information available to model developers and the methodology of mechanistic development have grown significantly. This new knowledge, the gathering of which was largely supported by EPA, has significantly enhanced the science of photochemical kinetics mechanism development. Whereas it was not yet possible to develop a comprehensive mechanism for the description of urban ozone formation from reactive hydrocarbon mixtures at the time of the original EKMA, we now agree on a good deal of this chemistry and can clearly identify specific area of uncertainty that must be addressed experimentally. Thus, the objective of this workshop, the identification and minimization of uncertainties in the photochemical kinetics mechanism development process, must be viewed with the significant advancement of our science in mind. The chemical representations in most of the mechanisms being compared today have converged to the point where, within the bounds of uncertainty, it may be difficult to distinguish the benefits of one over another. Thus, there seem to be two goals for this workshop: (1) to define testing procedures which allow one to verify that a mechanism is within those bounds of uncertainty which with we presently feel can be attained (that it is the state-of-the-science), and (2), to find new ways to limit the uncertainty in tests which we use to distinguish between mechanistic representations.

The uncertainty resides in the test data (mainly smog chamber data) and the measurement technique (our estimate of reasonable agreement between simulation results and data). The foundation of a photochemical

kinetics mechanism is the elementary kinetic and mechanistic data upon which it is based. Support of new data gathering and periodic evaluation of the data base will diminish uncertainty in the basic chemistry. The input of mechanism developers cannot be overlooked in this process since the areas of largest uncertainty sometimes only become apparent after application of a model. We must also strengthen our existing smog chamber data sets. Besides the continuing need for new experiments (possibly from new or different facilities), we must study the artificial processes in existing chambers so that uncertainty in existing data sets can be minimized. Study of the wall radical issue is of immediate importance. In addition, the careful characterization of light sources in every chamber should be carried out and made available with the data. Chamber intercomparisons, intercomparisons of instruments and calibration schemes, and clear descriptions of uncertainty bands in measured concentrations must be made available.

With improvements in experimental data sets such as those noted, it should be possible to define some basic tests with which the validity of a chemical mechanism can be compared to another. As stated earlier, however, such test may not be powerful enough to point out a truly excellent mechanism compared to a good mechanism within the current bounds of our uncertainties. On the other hand, the continuing limitation of uncertainty in the test data will allow such tests to demonstrate some basic weaknesses in poorer mechanistic descriptions. At a minimum, the participants of the workshop should try to agree on a method for producing a set of good smog chamber data regardless of the intended use. Such a set would be a great benefit to model developers. For some types of mechanisms, however, we must recognize that we are in the early phase of normal science. Finally, mechanistic representations for hydrogen and higher molecular weight peroxides, formic and higher molecular weight organic acids, aqueous chemistry, and secondary aerosol formation are under development but lack the test data base available for urban oxidant mechanisms. This will eventually impact on the credibility of these mechanisms and steps to develop this data must now be considered.

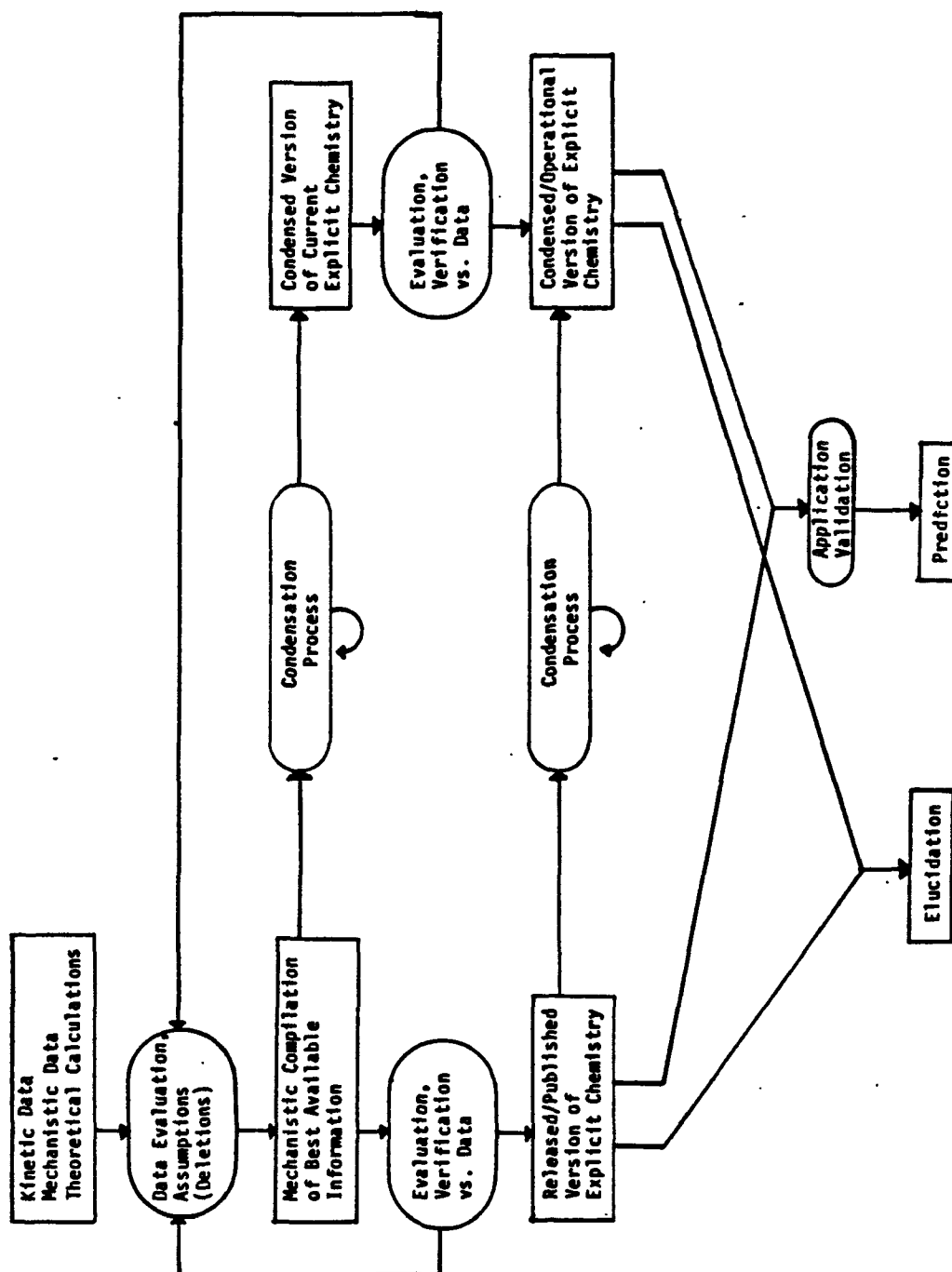


Figure 1. Schematic Diagram of the Process of Mechanism Development and Evaluation.

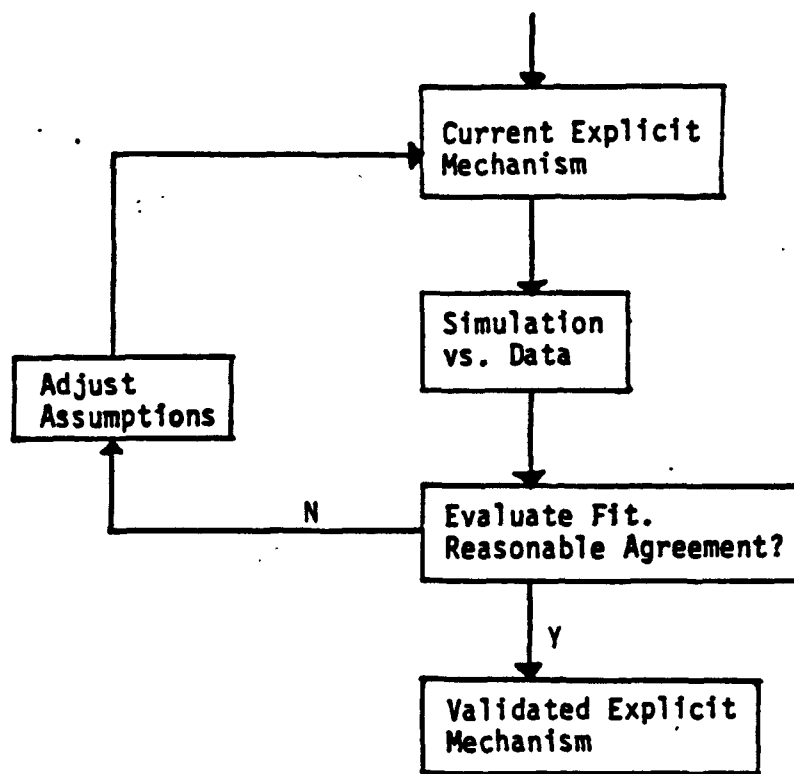


Figure 2. Schematic Description of the Mechanism Evaluation Process.

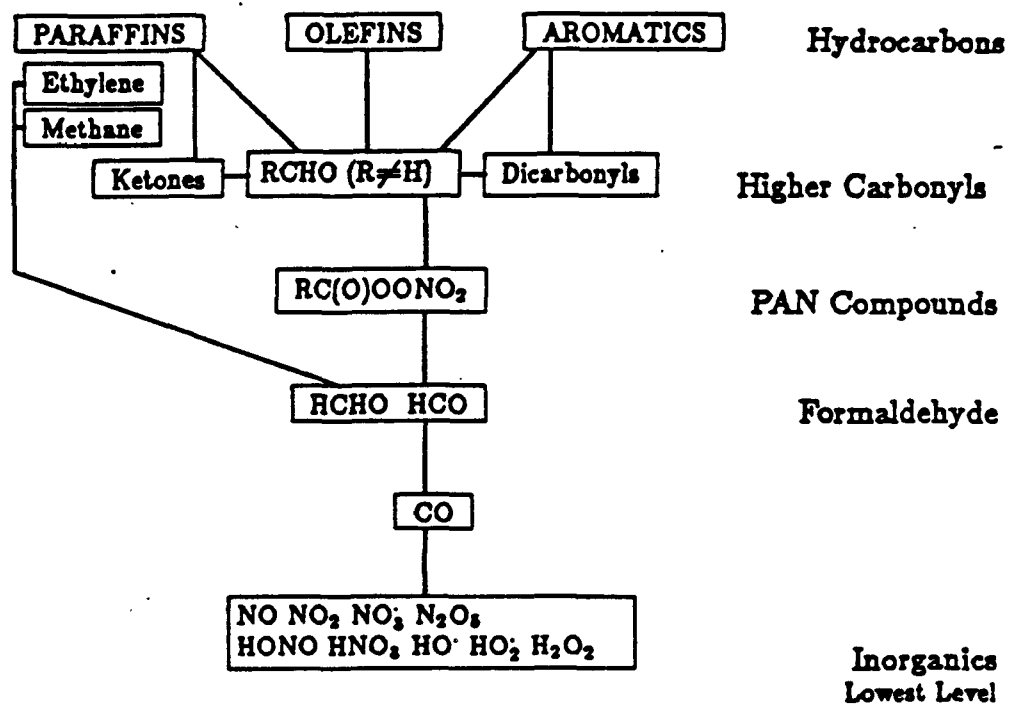


Figure 3. Schematic Description of the Hierarchy of Species for the Carbon Bond Mechanism-IV.

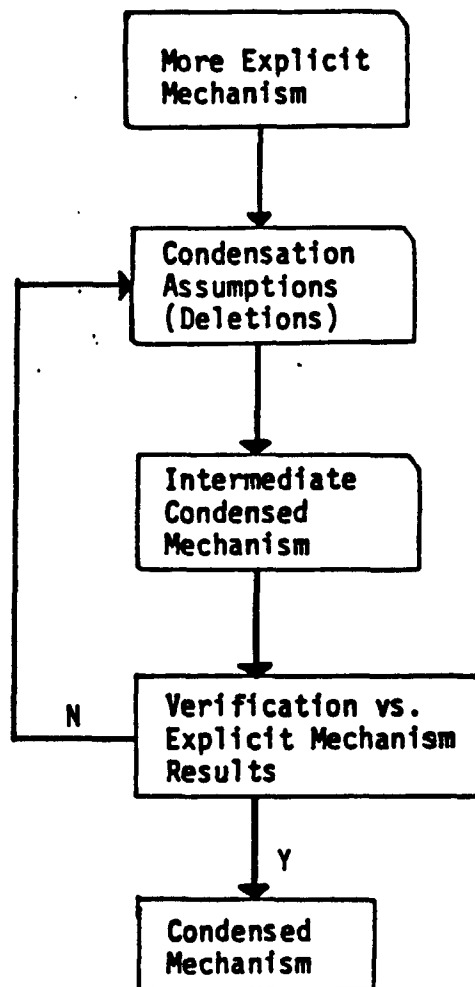


Figure 4. Schematic Description of the Mechanism Condensation Process.

Paper 6

**Uncertainty Analysis and Captive-Air Studies as Aids in
Evaluating Chemical Mechanisms for Air Quality Models**

**Alan M. Dunker
Environmental Science Department
General Motors Research Laboratories**

Abstract

Chemical mechanisms employed in air quality models to describe the formation of pollutants, such as ozone, are simplifications of what actually occurs in the atmosphere, and therefore the mechanisms must be evaluated to determine their accuracy. Concern has arisen regarding how mechanisms are evaluated because different mechanisms can predict different emission control requirements. The procedure for evaluating mechanisms can be improved through the use of uncertainty analysis and the results of captive-air experiments.

In uncertainty analysis, probability distributions are estimated for the parameters in a mechanism, chiefly rate constants and initial concentrations. Using these probability distributions, the variances in the concentrations predicted by the mechanism are calculated. The variances can be related to uncertainty limits, and comparison of the uncertainty limits on the predicted concentrations with the uncertainty limits on the experimental data shows whether differences between the mechanism and the data are significant.

In captive-air experiments, ambient air is loaded into Teflon bags in the early morning and irradiated by sunlight throughout the day. The advantages of such experiments are that an ambient mix of

hydrocarbons at ambient concentrations is used and the photolysis rates correspond closely to those in the atmosphere. Further, multiple bags can be run on the same day with known amounts of clean air and/or emitted compounds added to some bags. The known additions can simulate the effects of emission control strategies and broaden the range of conditions studied. Comparison of the predictions of a mechanism with the data from captive-air experiments provides tests of the mechanism under conditions close to those in the atmosphere.

Introduction

Secondary pollutants, such as ozone (O_3), are formed in the lower atmosphere by a complicated sequence of chemical reactions which involve precursors, such as hydrocarbons and NO_x . Chemical mechanisms, sets of chemical reactions, have been assembled to describe this complex chemistry. Coupled with descriptions of transport and dispersion, chemical mechanisms are used in atmospheric models to develop strategies for reducing the concentrations of the secondary pollutants by controlling emissions of the precursors.

Because atmospheric chemistry is so complex and incompletely understood, chemical mechanisms are by necessity condensations of and approximations to what is actually occurring in the atmosphere. As a consequence, there is not one uniformly accepted chemical mechanism but rather a number of different mechanisms, and as new information on reactions becomes available new mechanisms are developed (Killus and Whitten, 1981a; Atkinson et al., 1982; Lurmann et al., 1986; Stockwell, 1986). Comparisons of different mechanisms have shown that they give qualitatively similar results. However, the comparisons have also shown that quantitative predictions of emissions reductions needed to meet the National Ambient Air Quality Standard for O_3 can be substantially different from different

mechanisms. (Jeffries et al., 1981; Carter et al., 1982a; Dunker et al., 1984; Leone and Seinfeld, 1985; Shafer and Seinfeld, 1986)

This difficulty has focused attention on the process employed to evaluate mechanisms to determine how well they represent what occurs in the polluted troposphere. Jeffries and Arnold (1986) have recently reviewed the history of how mechanisms are developed and tested. In the past, mechanisms have been evaluated by comparing their predictions with the results of smog chamber experiments. Such experiments attempt to simulate conditions in the atmosphere, but the concentrations of the precursors used are generally higher than found in the atmosphere and the mixtures of hydrocarbons used represent only the major species found in the atmosphere. Additional concerns have arisen regarding the effects of chamber surfaces on the observed gas-phase chemistry. (Killus and Whitten, 1981b; Carter et al., 1982a, 1982b; Dunker et al., 1984; Leone and Seinfeld, 1985; Shafer and Seinfeld, 1986). The sources of disagreement between mechanisms and chamber data, errors in the mechanism versus errors in the data, very often are not clear.

This report addresses the question of how to improve the process for evaluating chemical mechanisms, and two suggestions are made. The first suggestion is to use uncertainty analysis when comparing a chemical mechanism to chamber experiments, and the second suggestion is to conduct captive-air experiments. To date

these tools have rarely been applied in evaluating mechanisms for the polluted troposphere. The next section describes how uncertainties arise in chemical mechanisms and what numerical techniques are available to calculate uncertainty limits for the predictions of the mechanisms. This is followed by a section discussing captive-air experiments and why they are a valuable tool in determining whether chemical mechanisms adequately represent the atmosphere. The final section gives concluding remarks.

Uncertainty Limits

A chemical mechanism is evaluated by using it to simulate chamber experiments and then comparing the predictions of the mechanism for the concentrations of different species with the experimental measurements. Both the measurements and the predictions, however, have uncertainties associated with them. Furthermore, the agreement between the measurements and the predictions is never perfect, and the question then arises whether the differences seen are significant. If the differences are significant, the chemical mechanism does not adequately represent the data. If the differences are not significant, the uncertainties in the experimental data and/or in the results of the chemical

mechanism must be decreased to make a more precise evaluation of the mechanism. Uncertainty limits have often been placed on the data acquired in chamber experiments. However, very little work has been done to place uncertainty limits on the results of chemical mechanisms for simulations of the troposphere in general and chamber experiments in particular. This is in contrast to the study of stratospheric chemistry where efforts have been made to place uncertainties on the results of simulations (Stolarski et al., 1978; Ehhalt et al., 1979; Stolarski and Douglass, 1986).

There are several sources of uncertainty in a chemical mechanism. First, some important reactions may be completely unknown and therefore not included in the mechanism. Second, for those reactions included in the mechanism, the rate constants are not known precisely but rather are known to varying degrees of accuracy. Third, the products or the distribution of products for some reactions are in question. In addition, when a chemical mechanism is used, initial concentrations must be supplied for all the species in the mechanism. Some of these initial concentrations may not be measured, and those that are measured will have uncertainties associated with them. Important examples encountered in simulating chamber experiments are the uncertainties in the initial HONO concentration, the background OH radical flux, and the aldehyde photolysis rates.

The first source of uncertainty cited above, complete ignorance of an important reaction, cannot be treated in an uncertainty analysis. One must at least suspect that a reaction occurs, know the reactants, and have some estimate of the rate constant. The other sources of uncertainty can be treated, however.

The different sources of uncertainty cause uncertainties in the predictions of the chemical mechanism which vary with time in a given simulation, vary from one simulation to another, and vary from one chemical species to another. A schematic representation of the effect of errors in initial concentrations and rate constants on the results of a mechanism is shown in Figure 1. The points denote measurements of the concentration $C(t_i)$ of some species in a chamber experiment, and the vertical bars denote the uncertainty associated with each measurement. The curves represent results obtained with a chemical mechanism for simulations of the experiment. In practice, one begins the simulation with the measured initial concentrations, c_0^m , and uses in the chemical mechanism the approximate rate constants, k^m , measured in other experiments. This produces the calculated concentration $c(t; c_0^m, k^m)$ as a function of time t (upper curve). If, however, the true initial concentrations, c_0 , and the true rate constants, k , were known, these could be used in the chemical mechanism to produce the concentration $c(t; c_0, k)$ (lower curve). The difference between the two curves represents the effect of errors in the initial concentrations and rate constants on the

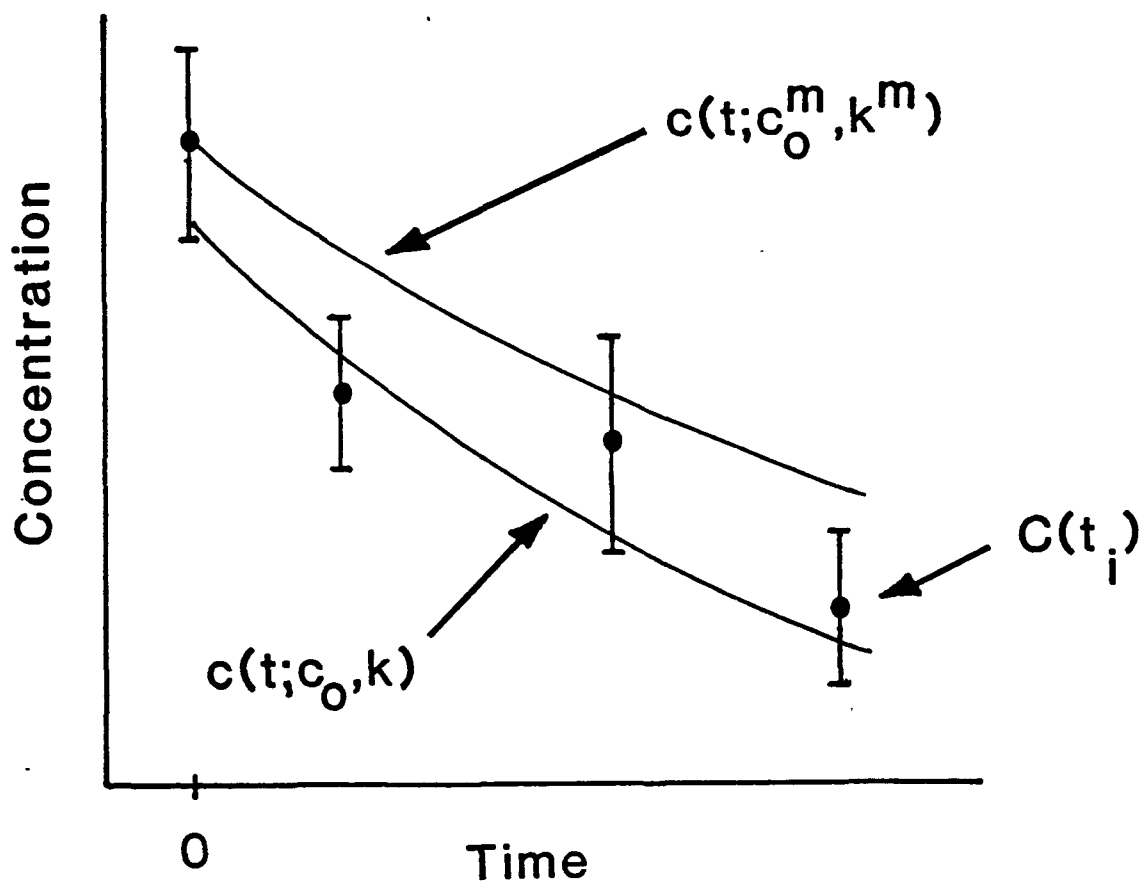


Figure 1. The effect of errors in initial concentrations c_0^m and rate constants k^m on the results of a mechanism. The upper curve is the concentration normally obtained in a simulation of a chamber experiment, the lower curve is the concentration that would be obtained using the true initial concentrations c_0 and true rate constants k , and the points are experimental measurements.

predictions of the mechanism.

Quantifying the effects of possible errors or uncertainties in c_0^m and k^m requires two steps. First, the probability distributions of c_0^m and k^m , $P(c_0^m)$ and $P(k^m)$, must be estimated. Because the true values c_0 and k are not known, estimating the uncertainties in c_0^m and k^m (and thereby estimating $P(c_0^m)$ and $P(k^m)$) necessarily requires some subjective judgments. The uncertainties in the initial concentrations can be estimated based on knowledge of the experimental techniques and the reproducibility of the measurements, and these estimates are probably best made by the experimenters themselves. Estimates of uncertainties in rate constants are available in some reviews of kinetic data (Atkinson and Lloyd, 1984; Demore et al., 1985). Such reviews are very useful in that they provide the most impartial, least subjective assessments of the uncertainties in rate constants.

The second step in the uncertainty analysis is calculating the variance $\sigma(t)^2$ in the predicted concentrations. $\sigma(t)^2$ is defined by

$$\sigma(t)^2 = \langle [c(t; c_0^m, k^m) - \langle c(t; c_0^m, k^m) \rangle]^2 \rangle \quad (1)$$

where

$$\langle f(c_0^m, k^m) \rangle = \int f(c_0^m, k^m) P(c_0^m) P(k^m) dc_0^m dk^m \quad (2)$$

for a function f . The integration here is over all variables for which uncertainties are considered. If uncertainties in a large number of rate constants and initial concentrations are included in the analysis, calculating $\sigma(t)^2$ is not a trivial task. Current chemical mechanisms contain 70 or more reactions involving 30 or more chemical species. For a complete uncertainty analysis, then, calculating $\sigma(t)^2$ would require evaluating an integral of dimension 100 or greater (Eq.(2)).

There are three numerical methods for evaluating integrals of high dimension which can be readily applied to calculate $\sigma(t)^2$. In the Monte Carlo method, the integrand of Eq.(2) is evaluated for random variations of the variables c_0^m , k^m , and the average value of the integrand is used to determine the integral. This method has been applied by Stolarski et al. (1978) and Stolarski and Douglass (1986) to determine the uncertainties in predictions of stratospheric ozone depletion. For applications to chamber experiments, the Monte Carlo method has a disadvantage in that large amounts of computing time may be required. The chamber experiment

must be simulated anew for each of the random variations in c_0^m , k^m to obtain $c(t; c_0^m, k^m)$. Furthermore, $\sigma(t)^2$ will vary from one chamber experiment to another, so the entire analysis will have to be redone for each experiment.

The second approach is the Fourier method (Cukier et al., 1978). Here, c_0^m and k^m are varied simultaneously in a systematic fashion and the resulting variations in $c(t; c_0^m, k^m)$ analyzed by Fourier series to evaluate the multidimensional integrals. The method was originally developed for the sensitivity analysis of complex models, and Falls et al. (1979) have applied the method to a chemical mechanism. While calculations of $\sigma(t)^2$ have not been reported, this quantity can be obtained by the Fourier method. The Fourier method is more efficient than the Monte Carlo method, but the computing time required can still be large.

The third approach is the sensitivity coefficient method in which gradients of $c(t; c_0^m, k^m)$ with respect to c_0^m and k^m are calculated and employed to evaluate the multidimensional integrals. Ehhalt et al. (1979) have used this method, again, to place uncertainty limits on the predictions of stratospheric ozone depletion. To calculate the sensitivity coefficients, Ehhalt et al. (1979) made small perturbations in the variables c_0^m , k^m , one at a time, and applied finite difference formulas. A more accurate and efficient method for calculating the sensitivity coefficients, the

decoupled direct method, has been described by Dunker (1984). McCroskey (1985) and the Acid Deposition Modeling Project (1986) have used the decoupled direct method in calculating uncertainty limits for simulations of some chamber experiments. The sensitivity coefficient method requires less computer time than the Monte Carlo and Fourier methods and thus appears better suited for routine application to chamber experiments. The sensitivity coefficient method does require care in treating cases with large uncertainties in c_0^m , k^m , and for very large uncertainties may require extension beyond present capabilities.

Uncertainty limits corresponding to one standard deviation can be placed on $c(t; c_0^m, k^m)$ using $\sigma(t)$ as illustrated in Figure 2. At the initial time $t = 0$, these uncertainty limits are identical to the one standard deviation uncertainty limits on the measured initial concentration, but for later times $\sigma(t)$ can increase or decrease from $\sigma(0)$. Thus, $\sigma(t)$ provides a quantitative estimate of the effect of uncertainties in c_0^m and k^m on the results of the chemical mechanism. Jeffries and Arnold (1986) have raised the question of what constitutes good agreement between a mechanism and chamber data. $\sigma(t)$ can be employed to define the level of agreement. It is proposed here that a necessary condition for reasonable agreement between simulation results and chamber data is that uncertainty limits two standard deviations wide on the simulation results and the data overlap. Similarly, it is proposed

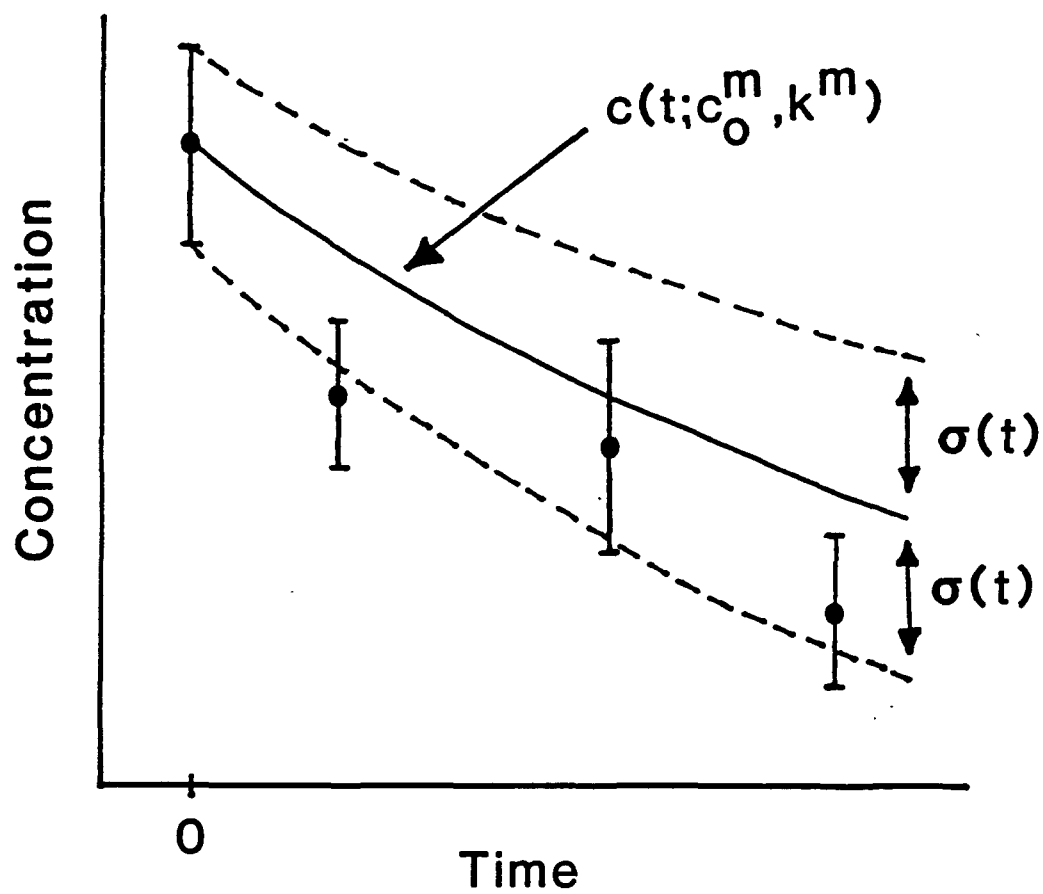


Figure 2. Uncertainty limits (dashed curves) for the concentration predicted by a chemical mechanism (solid curve) using measured (approximate) initial concentrations and rate constants.

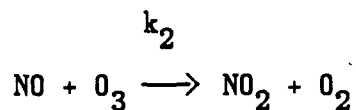
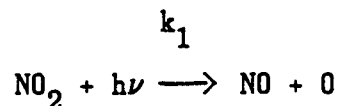
that a necessary condition for good agreement between simulation results and chamber data is that the uncertainty limits one standard deviation wide overlap.

Several additional remarks regarding uncertainty analysis should be made. First, if a complete analysis taking into account the uncertainties in all the rate constants and initial concentrations cannot be made, due perhaps to the complexity of the mechanism or lack of information on uncertainties for some rate constants, a partial analysis can still be useful. For example, an analysis taking into account only the uncertainties in the chamber photolysis rates and the surface sinks and sources will show whether discrepancies between the simulation results and the chamber data can be explained by uncertainties in the chamber operating conditions.

Second, an uncertainty analysis must be done with some care since the uncertainties in the rate constants and initial concentrations may not all be independent. As an example, the NO_2 photolysis rate in chamber experiments is often measured by the photostationary-state method (Wu and Niki, 1975). This method gives

$$k_1^m = \frac{[\text{NO}] [\text{O}_3]}{[\text{NO}_2]} k_2^m \quad (3)$$

where k_1^m and k_2^m are the measured rate constants for the reactions



According to Eq. (3), the uncertainty in k_2^m will cause uncertainty in k_1^m , and the correlation between the uncertainties in these two rate constants should be taken into account in specifying $P(k_1^m)$ and in calculating $\sigma(t)^2$ (Eqs. (1) and (2)).

Finally, when applying a chemical mechanism, one often compares concentrations of the same species obtained in different simulations, e. g., comparison of O_3 concentrations from simulations with different emission rates. In such situations, the variance of the concentration difference should be calculated because it is likely that it will be less than the sum of the variances for the separate simulations. That is, one expects

$$\text{var}[c(t)-c'(t)] < \text{var}[c(t)] + \text{var}[c'(t)]$$

where $c(t)$ and $c'(t)$ are concentrations obtained in different simulations. The reason for this is that the uncertainties in $c(t)$ and $c'(t)$ will be correlated because the same rate constants with the same errors will be used in both simulations. There is hence a tendency for errors to cancel when forming the difference $c(t)-c'(t)$. Since the uncertainty limits are obtained from the variances, using $\text{var}[c(t)] + \text{var}[c'(t)]$ to determine the uncertainty limits for $c(t)-c'(t)$ will likely give limits that are overly pessimistic (unnecessarily wide). The better approach, using $\text{var}[c(t)-c'(t)]$, has been employed in studies of stratospheric ozone depletion, where differences in ozone concentrations are examined (Stolarski et al., 1978; Ehhalt et al., 1979; Stolarski and Douglass, 1986).

Captive-Air Studies

In a captive-air experiment, ambient air in an urban area is loaded into a Teflon bag in the early morning, between 6 AM and 10 AM. The bag is then irradiated by sunlight throughout the day, and the concentrations of the precursors and the secondary pollutants are measured periodically. Such experiments are similar to but have

several advantages over traditional smog chamber studies.

The first advantage is that an ambient mix of hydrocarbons at ambient concentrations is used in captive-air experiments. Traditional chamber experiments, however, use mixtures containing only the major hydrocarbon species found in the atmosphere, generally at higher than ambient concentrations. Second, the photolysis rates in captive-air experiments correspond very closely to those in the atmosphere because thin Teflon film is virtually transparent to solar radiation (Kelly, 1981a). The photolysis rates in traditional chamber studies are not as close to atmospheric rates because the spectrum of artificial light sources does not exactly duplicate the solar spectrum. Finally, captive-air experiments can be run in several bags on the same day with known additions made to some bags. The known additions broaden the range of conditions studied and can mimic the effects of control strategies. Table I describes some of the additions which can be made and their relationship to strategies for reducing O_3 concentrations. Changes in concentrations of the precursors have, of course, been made in traditional chamber experiments, but, again, the hydrocarbon mix to which the perturbations are made is a simplification of the ambient hydrocarbon mix.

Considering these advantages, it is proposed here that captive-air experiments can help bridge the gap between traditional smog

Table I
Variations in Captive-Air Experiments

Addition to Ambient Air	Relationship to Control Strategies	Effect on Hydrocarbon/ NO_x Ratio
Clean air	Equal control of all pollutants	Unchanged
NO_x	Opposite of NO_x control	Decreased
One or more hydrocarbons	Opposite of hydrocarbon control; change in distribution of hydrocarbons	Increased
NO_x followed by clean air to dilute back to the original NO_x concentration	Control of all hydrocarbons	Decreased
One or more hydrocarbons followed by clean air to dilute back to the original total hydrocarbon concentration	Control of NO_x ; change in distribution of hydrocarbons	Increased

chamber experiments and the atmosphere. This is not to suggest that traditional chamber experiments should be discontinued; they are valuable in building and evaluating chemical mechanisms. However, comparison of the results of a mechanism with data from captive-air experiments provides an evaluation of the mechanism under conditions as close as possible to those under which it will be applied.

There are disadvantages to captive-air experiments as well. If experiments are run in several bags simultaneously, the surface effects in each bag must be characterized or the surface effects must be small and the expected variability from bag to bag must be determined. Otherwise one cannot determine whether differences seen between experiments conducted in different bags are due to different precursor concentrations or different surface effects. Second, it can be difficult or impossible to control some parameters in captive-air experiments, such as light intensity, temperature, and humidity. As a result, the failure rate and cost of captive-air experiments are higher.

Lastly, since ambient air contains a complex mixture of hydrocarbons, careful measurements of individual hydrocarbon species must be made at the start of, and preferably throughout, the experiments if they are to be used to evaluate chemical mechanisms. This is the most significant disadvantage of captive-air experiments because it is impossible to identify all the hydrocarbons in ambient

air. In simulating the experiments, then, it is necessary to make assumptions regarding the fraction of hydrocarbons which are unidentified. However, this problem must be faced because it is also encountered when using a chemical mechanism in an air quality model for simulations of an entire urban area. If a significant fraction of the hydrocarbons is unidentified, then a comparison between the results of the mechanism and the data will provide a test of the mechanism combined with the assumptions on the unidentified hydrocarbons.

Several captive-air studies have been conducted to date. The most ambitious study was done by Grosjean et al. (1982) in Los Angeles in the fall of 1981. Both large (70,000 l) and small (4000 l) Teflon bags were used, and individual hydrocarbon species were measured in the captured air. Results from the large bag were compared to results from one small bag to check the effect of chamber size on the results, and additions of NO or NO₂ were made to the other small bags. Predictions of the Atkinson et al. (1982) chemical mechanism and the EKMA (Empirical Kinetic Modeling Approach) mechanism (Dodge, 1977) for O₃, NO, NO₂, hydrocarbons, and aldehydes were compared to the observed concentrations. Unfortunately, experimental problems and uncertainties in the flux of OH radicals from chamber surface reactions prevented a rigorous evaluation of the chemical mechanisms with the data.

Kelly (1981b, 1985) has conducted captive-air studies in Houston (summer, 1977) and Detroit (summer, 1981). In these experiments, small (500 l) Teflon bags were used, and the total concentration of nonmethane hydrocarbons, rather than individual hydrocarbon species, was measured in the captured air. An extensive set of additions to different bags was done; clean air, NO, NO₂, butane, ethene, and propene alone and in combination were added. Using default estimates of the relative amounts of different hydrocarbon species, predictions of the carbon bond III mechanism (Killus and Whitten, 1981a) and the EKMA mechanism (Dodge, 1977) for O₃ were compared to the observed concentrations. While this approach provides an indication of the validity of the mechanisms, detailed measurements of hydrocarbon species are necessary for a rigorous evaluation. As part of the Southern California Air Quality Study in the summer of 1987 (Blumenthal et al., 1986), General Motors Research plans additional captive-air experiments in Los Angeles in which individual hydrocarbon species will be measured. This data should allow more thorough testing of chemical mechanisms.

Conclusion

It has been recognized for some time that a critical element of an air quality model is the chemical mechanism. In designing

strategies for reducing secondary pollutants such as O_3 , inaccuracies in the chemical mechanism lead to inappropriate and ineffective strategies. It has also been recognized that evaluating chemical mechanisms to determine how well they represent the transformations actually occurring in the atmosphere is not a simple task. A review of past procedures for developing and evaluating mechanisms (Jeffries and Arnold, 1986) and consideration of the fact that current mechanisms predict different control requirements (Jeffries et al., 1981; Carter et al., 1982a; Dunker et al., 1984; Shafer and Seinfeld, 1986) indicates that improvements are necessary in how mechanisms are evaluated.

As described above, uncertainty analysis and captive-air experiments are two tools which can aid in evaluating mechanisms. Uncertainty analysis can quantify the agreement between chamber data and the predictions of a mechanism and show whether there is a fundamental disagreement between the data and predictions. Uncertainty analysis can also place error bounds on the predictions of a mechanism in an application to the atmosphere. Further, uncertainty analysis can indicate how the error bounds can best be reduced. That is, the rate constants, product yields, and initial concentrations contributing most to the uncertainty in the predictions of the mechanism can be identified. Captive-air experiments provide tests of a mechanism under conditions a step closer to those in the atmosphere than traditional smog chamber

experiments. Because the complexity of captive-air experiments is greater, greater effort is required to achieve the same quality of data.

Uncertainty analysis and captive-air experiments are not the only tools which can improve the procedure for evaluating mechanisms. They are, however, tools which have received little attention to date.

References

- Acid Deposition Modeling Project (1986) Preliminary evaluation studies with the regional acid deposition model (RADM), National Center for Atmospheric Research Technical Note TN-265+STR.
- R. Atkinson, A. C. Lloyd, L. Wings (1982) An updated chemical mechanism for hydrocarbon/ NO_x / SO_2 photooxidations suitable for inclusion in atmospheric simulation models. Atmos. Environ. 16, 1341-1355.
- R. Atkinson, A. C. Lloyd (1984) Evaluation of kinetic and mechanistic data for modeling of photochemical smog. J. Phys. Chem. Ref. Data 13, 315-444.
- D. Blumenthal, J. Watson, D. Lawson (1986) Southern California air quality study suggested program plan. Sonoma Technology Inc. Report 10 95050-805.
- W. P. L. Carter, A. M. Winer, J. N. Pitts, Jr. (1982a) Effects of kinetic mechanisms and hydrocarbon composition on oxidant-precursor relationships predicted by the EKMA isopleth technique. Atmos. Environ. 16, 113-120.
- W. P. L. Carter, R. Atkinson, A. M. Winer, J. N. Pitts, Jr. (1982b) Experimental investigation of chamber-dependent radical sources. Int. J. Chem. Kinet. 14, 1071-1103.
- R. I. Cukier, H. B. Levine, K. E. Shuler (1978) Nonlinear sensitivity analysis of multiparameter model systems. J. Comput. Phys. 26, 1-42.

- W. B. Demore, J. J. Margitan, M. J. Molina, R. T. Watson, D. M. Golden, R. F. Hampson, M. J. Kurylo, C. J. Howard, A. R. Ravishankara (1985) Chemical kinetics and photochemical data for use in stratospheric modeling. California Institute of Technology Jet Propulsion Laboratory Publication 85-37.
- M. C. Dodge (1977) Combined use of modeling techniques and smog chamber data to derive ozone-precursor relationships. U. S. Environmental Protection Agency EPA-600/3-77-001.
- A. M. Dunker (1984) The decoupled direct method for calculating sensitivity coefficients in chemical kinetics. J. Chem. Phys. 81, 2385-2393.
- A. M. Dunker, S. Kumar, P. H. Berzins (1984) A comparison of chemical mechanisms used in atmospheric models. Atmos. Environ. 18, 311-321.
- D. H. Ehhalt, J. S. Chang, D. M. Butler (1979) The probability distribution of the predicted CFM-induced ozone depletion. J. Geophys. Res. 84, 7889-7894.
- A. H. Falls, G. J. McRae, J. H. Seinfeld (1979) Sensitivity and Uncertainty of Reaction Mechanisms for Photochemical Air Pollution. Int. J. Chem. Kinet. 11, 1137-1162.
- D. Grosjean, A. C. Lloyd, R. J. Countess, F. Lurmann, K. Fung (1982) Captive air experiments in support of photochemical kinetic model evaluation. Environmental Research and Technology Document No. P-A764-500.

- H. E. Jeffries, J. Arnold (1986) The science of photochemical reaction mechanism development and evaluation. U. S. Environmental Protection Agency, Workshop on evaluation and documentation of chemical mechanisms used in air quality models.
- H. E. Jeffries, K. G. Sexton, C. N. Salmi (1981) The effect of chemistry and meteorology on ozone control calculations using simple trajectory models and the EKMA procedure. U. S. Environmental Protection Agency EPA-450/4-81-034.
- N. A. Kelly (1981a) The Characterization of Fluorocarbon-Film Bags as Smog Chambers. General Motors Research Laboratories Publication GMR-3735.
- N. A. Kelly (1981b) An analysis of ozone generation in irradiated Houston air. J. Air Pollut. Control Assoc. 31, 565-567.
- N. A. Kelly (1985) Ozone/precursor relationships in the Detroit metropolitan area derived from captive-air irradiations and an empirical photochemical model. J. Air Pollut. Control Assoc. 35, 27-34.
- J. P. Killus, G. Z. Whitten (1981a) A new carbon-bond mechanism for air quality simulation modeling. Systems Applications Inc. Report No. 81245.
- J. P. Killus, G. Z. Whitten (1981b) Comments on "A smog chamber and modeling study of the gas phase NO_x -air photooxidation of toluene and the cresols". Int. J. Chem. Kinet. 13, 1101-1103.

- J. A. Leone, J. H. Seinfeld (1985) Comparative analysis of chemical reaction mechanisms for photochemical smog. Atmos. Environ 19, 437-464.
- F. W. Lurmann, A. C. Lloyd, R. Atkinson (1986) A chemical mechanism for use in long-range transport/acid deposition computer modeling. J. Geophys. Res. 91, 10905-10936.
- P. S. McCroskey (1985) Direct differential methods for kinetic parameter estimation. Carnegie-Mellon University, Master's thesis.
- T. B. Shafer, J. H. Seinfeld (1986) Comparative analysis of chemical reaction mechanisms for photochemical smog II. Sensitivity of EKMA to chemical mechanism and input parameters. Atmos. Environ. 20, 487-499.
- W. R. Stockwell (1986) A homogeneous gas phase mechanism for use in a regional acid deposition model. Atmos. Environ. 20, 1615-1632.
- R. S. Stolarski, D. M. Butler, R. D. Rundel (1978) Uncertainty propagation in a stratospheric model 2. Monte Carlo analysis of imprecisions due to reaction rates, 83, 3074-3078.
- R. S. Stolarski, A. R. Douglass (1986) Sensitivity of an atmospheric photochemistry model to chlorine perturbations including consideration of uncertainty propagation. J. Geophys. Res. 91, 7853-7864.
- C. H. Wu, H. Niki (1975) Methods for measuring NO_2 Photodissociation Rate: Application to Smog Chamber Studies. Environ. Sci. Technol. 9, 45-52.

Discussion Following Dunker Presentation

Sexton: In the discussion of calculating uncertainty, you started with an "if" statement, that if we knew all the reactions properly and all the products then you could go on and discuss the uncertainty from the kinetic, the rate constant, then yesterday Roger mentioned that we actually seem to know less about the proper reactions and the products than we do in the kinetics. Do you have any suggestions for how to deal with, or how to calculate, or estimate the uncertainty, not knowing the chemistry in the products correctly?

Dunker: Consider a situation where you are developing a chemical mechanism and a particular reaction is known to proceed via two pathways to two different products, A and B, but the relative amounts of these products produced is not well known. In such a situation, you can write a single reaction and assign a best guess for the stoichiometric coefficients of the two products. Then, you would place fairly wide uncertainty limits on the two stoichiometric coefficients and determine the effects of these uncertainties on predicted concentrations obtained from the chemical mechanism. (In carrying out the uncertainty analysis, it is important to remember that the stoichiometric coefficients for products A and B in this case are correlated. That is, if 80% of the reaction proceeds to product A, then 20% must proceed to product B; if 30% of the reaction proceeds to product A, then 70% must proceed to product B; and so forth.) If the uncertainty analysis shows that the results of the chemical mechanism have low uncertainty despite the high uncertainty in the stoichiometric coefficients for products A and B, then you need not worry about this reaction further. However, if the uncertainty in the stoichiometric coefficients causes large uncertainty in the results of the chemical mechanism, then you would have to flag this reaction as one which someone should investigate in detail.

The uncertainty analysis can be used to determine where it is most useful to concentrate additional experimental effort. By ranking the contributions of different rate constants and stoichiometric coefficients to the overall uncertainty in the predictions of the chemical mechanism, you can see which are the critical reactions to study. The uncertainty analysis can be done at various levels of detail. I think the simplest analysis would be to consider just those chamber effects which are not well characterized, place uncertainties on these effects, and see what the uncertainty analysis tells you. If it tells you that the uncertainty limits on the predictions of the chemical mechanism for the chamber experiments are large, then, before you do anything else, you should go back and carefully characterize the chamber. Thus, an uncertainty analysis may show that it is impossible to make reliable predictions with a chemical mechanism without further experimental work.

Jeffries: Alan, would you estimate what it might cost to do, let's say a seven or eight simultaneous bag captive air study for producing enough reasonable data. By seven or eight bag, I mean you're looking at seven or eight things at once.

Dunker: I am trying to remember how much the program carried out by ERT in Los Angeles in 1981 cost. I believe the cost was \$200,000 - \$250,000, and, apart from inflation, that is probably a good estimate of the cost of a captive-air study.

Jeffries: Strikes me, looking at what happened, that may not be enough?

Lloyd: I agree. Some of the problems we had indicated that it wasn't.

Lurmann: I think for \$400,000, you should be able to do some very good runs. That would be, the number of runs that were done were on the order of 30, total 30 days 20 to 30 days. You are getting paramount data, in that, if you're doing one run with the ambient air plus seven or lots of different bags, you've got a lot of individual runs on that. If everything worked right, you'd have on the order of 200 or more different conditions. That would be a lot of data, if everything worked.

Dodge: I don't think anyone will disagree that irradiating authentic atmospheric samples is not a preferred route to go. I mean, obviously that's considerably better than irradiating surrogate mixtures, but it just seems from the past work that's done, that it's just fraught with experimental difficulties. You're dealing with concentrations that are extremely, well the results are going to be dominated by surface effects, at those concentrations, the surface effects are variable from run to run, bag to bag. And, I just can't see it really working. Past studies have demonstrated that the data is just not of sufficient quality that you can run test mechanisms.

Dunker: If the bags are loaded during the 6-9 AM period near the center of a city, the concentrations of the precursors should be high enough to conduct meaningful experiments. In the experiments conducted by ERT in Los Angeles, the initial concentration of the nonmethane hydrocarbons (NMHC) exceeded 2.0 ppmC on 15 of 23 days. On four of those days the initial concentration of NMHC was 5.0 ppmC or higher. For the experiments done by Nelson Kelly in Houston, the initial concentration of NMHC exceeded 1.5 ppmC on 9 or 25 days, and for the experiments he did in Detroit the initial NMHC exceeded 0.7 ppmC on 5 of 12 days. On some days the initial concentrations may be so high that it is actually desirable to dilute the contents of the bag during the course of the day to simulate the dilution which occurs in the atmosphere as a result of the increase in the mixing height.

If surface effects in the bags are believed to be a problem, one can compare results from a large bag with results from smaller bags. This was done, for example, in the ERT study in which one 70,000 l bag was used along with several 4000 l bags. I think that most of the experimental problems encountered in conducting captive-air experiments are the same ones encountered in operating an outdoor smog chamber. You have to characterize the chambers and you have to make careful measurements of light intensity. What I see as the big problem in captive-air experiments is that you must make detailed hydrocarbon measurements during the runs. This data is essential to simulate the experiments with a chemical mechanism and to evaluate the chemical mechanism using the experiments.

Dodge: The concentrations that are used in the outdoor chambers are usually considerably higher than even your LA data met. Take the UNC chamber, it's extremely large when you're talking about working with Teflon bags and surface/volume ratio.

Bufalini: In following what Marcia was just suggesting, we've been involved for a number of years doing detailed hydrocarbon analysis and looking over our data as well as Washington State University data, it's not terribly unusual to run into about 20% of the organics as being unidentified. You've got a large number of unknown peaks as Marcia has suggested. Our approach has been to send it through scrubbers, in some cases we haven't been able to identify either aromatic, parafinic, or olefinic, though it's not exactly clear if they're oxygenated. I/we might possibly miss them, and we know approximately what range they're in, whether they are C8's or C9's or C6's, but still there are a number of peaks that are still unidentified. 20% is not unusual and I don't see how you're going to handle them in your model exactly, I mean you could make some sort of approximation of the carbon bond as to what they are, but I still think there's an amount of uncertainty there.

Dunker: I agree that measuring the hydrocarbon species is difficult and it may not be possible to measure all the species occurring in the atmosphere. That is why I feel that making the hydrocarbon measurements is the major problem in captive-air experiments. What we should remember, though, is that we are trying to create chemical mechanisms for application to the real world. If we cannot identify 20% of the hydrocarbons occurring in the atmosphere, then we will have to make some assumptions, either explicit or implicit, when applying any chemical mechanism or model to simulate the real world. For example, the simplest assumption is that the unidentified hydrocarbons have the same relative amounts of alkanes, alkenes, and aromatics as the identified hydrocarbons. If we conduct captive-air experiments and apply whatever assumptions we choose to make for the unidentified hydrocarbons, we will have tested whether the chemical mechanism plus the auxiliary assumptions for the unidentified hydrocarbons adequately represents the chemistry in the real world. If the agreement with the captive-air experiments is poor, then using the chemical mechanism in a sophisticated, three-dimensional atmospheric model with the same assumptions for the unidentified hydrocarbons will very likely not produce good results. Conversely, if we do not evaluate chemical mechanisms using captive-air experiments, we will not have checked an important set of assumptions, namely those regarding the unidentified hydrocarbons.

Dodge: Alan, I think in almost all cases that you would be able to compare because you've got so many uncertain parameters. You can vary their distribution of hydrocarbons in the mechanism because you have that 20 to 40%, you've got all the uncertainty in the wall effects, etc., etc. You've got enough adjustable parameters that you probably could get a good agreement but I'm not sure you could use that as a test discriminator although you're probably correct that you could use that to sort of lock in a default hydrocarbon concentration.

Dunker: If a significant fraction of the hydrocarbons cannot be identified, then the captive-air experiments will only provide a test of the chemical mechanism together with the auxiliary assumptions for the unidentified hydrocarbons and not a test of the chemical mechanism by itself. As you suggest for such a situation, the experiments could be used to develop the best auxiliary assumptions for the unidentified hydrocarbons. One does not have unlimited degrees of freedom in developing the auxiliary assumptions, however. That is, whatever assumptions one develops for the unidentified hydrocarbons should be applied uniformly to all captive-air experiments at a particular location. If a substantial number of captive-air experiments are done, it may

still be possible to conclude that a particular chemical mechanism with its best auxiliary assumptions for the unidentified hydrocarbons performs better than another chemical mechanism with its best auxiliary assumptions.

Jeffries: The real difference is that the two PPMC in the chamber's present 13 compounds which are well characterized, can be analyzed with high precision, and can be followed in great detail. I would have to agree with Marcia in the sense, and Joe too, that when we try automobile exhaust work in our chamber, we couldn't find 20% of the hydrocarbon most of the time, and it turns out that if you look at the distribution and try to match that against all kinds of stuff, there is enough flexibility in the modeler's choices about where he's going to put that 20% that he can make the model fit. It's not as sensitive a test to test the model against automobile exhaust as it is to test it against individual pure compounds where you know in detail what's going on. That's the difference.

Dunker: I agree that in standard chamber experiments using a known mix of hydrocarbons you can accurately follow the concentration of each hydrocarbon. Furthermore, I am not advocating discontinuing such experiments; I think we should continue them. I also think we should face the fact that a surrogate mix of hydrocarbons is not the same as what occurs in the atmosphere and deal with that problem. If 20% of the hydrocarbons in the atmosphere cannot be identified, we should use captive-air experiments to test the chemical mechanisms together with whatever assumptions we intend to make regarding the unidentified hydrocarbons.

Jeffries: I would agree that you would be very unhappy if it turns out that within the range of adjustments, you couldn't fit the captive air experiments. But fitting the captive air experiments doesn't distinguish one from another. By not fitting it, clearly you would be able to distinguish one from all the rest that might be able to fit it, so in a sense, I agree with you it's a necessary but not sufficient condition.

Whitten: I think that the captive air experiments provide a worthwhile supplement to all the other types of experiments, but I don't think they preempt a lot. I think in the future, a very important part of these experiments should include a careful characterization of the chamber surfaces. There are a certain battery of tests that are being developed in various places (I'm going to talk about that later today) that provide some information and this battery of tests could be included in the thing. A half step removed from captive air experiments are two different types of experiments. One is that we put genuine automobile exhaust into a fairly well characterized chamber. I think Harvey Jeffries was a little optimistic in saying that 20% is unidentifiable. Our experience was sometimes a little more than that. Especially with some fairly reactive compounds that were difficult to unequivocally identify with because of analytical problems. Another type of experiment that has been done in Sydney, Australia, where they did a rather careful analysis of the mixture of hydrocarbons and they put together an urban mix, which to the best of their ability, has manufactured all the compounds that they were able to identify. They had as many as 80 or 90.

Jeffries: They actually went to the individual factories and acquired samples of solvents and mixtures and the whole deal, so they were able to put together a mixture of three or four hundred identifiable individual compounds.

Whitten: And then they did smog chamber experiments with this very high number of compounds and they also used those experiments to test some chemical mechanisms and I think that would probably be a little better than your suggestion of holding back 20%. I think going to an entirely different facility with different analytical techniques and a different type of system, perhaps would be a better test of a mechanism. That way they're not being used in the same facility. It is really important that we have data from a group of smog chamber facilities rather than just one or two. It provides much more justification and scientific satisfaction, whatever the word is, for believing that the mechanisms really work in the atmosphere as well.

Dunker: I agree that chamber experiments with complex surrogate mixtures will provide good tests of chemical mechanisms, and I agree that we need chamber data from more than one or two facilities. I view captive-air experiments as a bridge between traditional chamber experiments and the atmosphere. Captive-air experiments can be criticized as dealing with a very complicated mix of hydrocarbon species with no hope of identifying and measuring all the different species. On the other hand, I do not feel that we can develop chemical mechanisms solely from chamber experiments using well-known hydrocarbons and then make the leap to atmospheric modeling. Somewhere in that leap we will have made assumptions regarding the mix of (unidentified) hydrocarbons found in the atmosphere, and those assumptions will never have been tested.

Whitten: You make a good point there. There were some experiments done at the University of Santa Clara where there was a problem with emissions from oil vents in the San Joaquin Valley and less than 10% of those hydrocarbons have been identified. It consisted of a lot of hydrocarbons greater than C10 but yet, right after the vent, the concentration in the air was maybe 20,000 parts per million. So it was easy to fill a cylinder with the emissions from these vents and it was totally unknown as to what these hydrocarbons were. These hydrocarbons were then injected into a smog chamber, subjected to smog chamber experiments and then compared with other known compounds to at least get a relative feeling for the reactivity of this emission source. Even though it had hundreds of compounds that we couldn't identify, that made it possible to include that emission in atmospheric simulation. I think it's very important to have these types of captive air experiments where you can't identify the hydrocarbons. I might add that if you could identify them, I'm not sure we'd know what the individual chemistry of these unknown compounds would be. That's even more scary, but yet we can watch them in a smog chamber and see what their relative reactivity is to other known things. That's an important, useful piece of information.

Demerjian: I had a comment that gets back to some of the things that I talked about yesterday. Alan, you had something on there about comparison with smog chamber data, that if you continue to accept it, you could show that the walls were dominating. I guess one of the problems I have with that is that's sort of one of these kind of statements where, well if no one goes and looks, then how will they ever discover that there's all these problems and therefore we continue to have this cycle that goes on. My feeling is when I was talking about this sort of aggressive approach into trying to really understand what's going on with chambers was across the board, including even these captive air experiments which, granted, would be more complex to characterize just because

of the fact that you're bringing in different ambient conditions in a smaller volume of air. But there is the chance, but I guess my feeling is that what's needed is a very well established procedure in terms of how one actually looks at these effects, how you characterize them. One that's done in a way that gives the people using it information, a very clear understanding of not only the sources of these things but also the potential sinks and as well as the fact of the variability that one might expect, given the type of chamber, if it's an indoor system, one has obviously more control has on the system. If it's an outdoor system, one has less control and one has to understand how it all factors in to this process. It seems to me if there's not going to be an aggressive approach to try to get at those things, then we continue to go around in circles chasing our tail. We've all acknowledged that there's a problem but no one seems to want to take hold of it. It's like, if we were sitting around here planning a manned space flight program, we wouldn't be talking about hot air balloons, okay? And, sometimes I get the feeling that that's what we're doing here. We're talking about using a technology that maybe needs to be worked on in a way that we start to get a handle on some of these problems. We know what the issues are, and yet we don't seem to ever try to get a hold of what they are and solve them. I don't care how many times you keep doing these experiments over and over again, if they still have the inherent sources of error that we're starting to believe are there and we don't seem to want to take the effort to look at and understand it, then I don't see how you can expect to proceed. As far as I'm concerned, it's not enough to say that the captive air experiment is a complicated problem that has this many things. I think we can better look at what the problems are and let's go see if we can get a handle on that. Otherwise, we'll go no place. We'll continue to stay ten years behind.

Dunker: I agree with you completely on the importance of characterizing chambers. I may not have stressed that in my talk, but I feel it is extremely important. Bill Stockwell has done a few calculations in which he put uncertainty estimates on the chamber parameters he felt were not well known, carried through an uncertainty analysis, and found large uncertainties in the concentrations of some species predicted by his chemical mechanism. Unless chambers are carefully characterized, it will be difficult or impossible to use the experimental data to discriminate between chemical mechanisms. I think that characterization of a chamber must be a continual process. That is, you do a series of characterization runs, then do some hydrocarbon/NO_x runs of interest, then return and do the characterization runs all over again. At present I do not believe we know whether or how a series of hydrocarbon/NO_x runs will change the surface characteristics of a chamber, and so characterization runs must be scheduled periodically as part of any experimental program. My one concern is that understanding in chemical detail what is occurring on the chamber surfaces may be very difficult, and the important processes may vary from one chamber to another. We may always be forced to use an empirical (but hopefully accurate) treatment of surface effects.

Demerjian: At least if there were some standard set of procedures that could be carried out through all of them, that would be one thing you could use to make comparisons in terms of surface characteristics.

Bradow: In the general area, one of the things that Alan talked about was the question about extending the ranges of experimentation. One of the places

where I think most of the models, what we're trying to apply to models these days involves, multi-day incidence, long-range transport, regional ozone problems, high ozone to NOx multiples, very low NOx values, all those kind of conditions seem to be almost inaccessible to the current measurements. I don't know that this is necessarily the case, but certainly these very high hydrocarbon/low NOx experiments are very hard to do. Seems to me, I support what Ken says, I think there needs to be some way to either expand the range of experimentation or range of experimental techniques to accomplish these very real conditions that are very, very important in terms of the application of problems.

Barnes: I take it that no chamber walls have been characterized. Do you agree with that?

Carter: Not completely characterized. There have been attempts to characterize.

Jeffries: You can predict the effect but you don't know how to explain it.

Paper 7

**Development and Testing of a Surrogate Species Reaction
Mechanism for Use in Atmospheric Simulation Models**

**William P. L. Carter
Statewide Air Pollution Research Center**

DEVELOPMENT AND TESTING OF A SURROGATE SPECIES REACTION MECHANISM
FOR USE IN ATMOSPHERIC SIMULATION MODELS

William P. L. Carter

Statewide Air Pollution Research Center
University of California
Riverside, California 92521

Introduction

In this presentation, I shall discuss the first phase of a two-phase program we have carried out to develop, test, and adapt an alternative chemical mechanism for use in EKMA models. This program was carried out by Roger Atkinson and myself at the Statewide Air Pollution Research Center (SAPRC), and by Fred Lurmann and Alan Lloyd at Environmental Research and Technology (ERT). The EPA contract monitor for this program is Marcia Dodge, and I also wish to acknowledge her contributions to this work. The first phase of this program was carried out primarily at SAPRC and the second phase was carried out at ERT.

In the first phase of this program, a detailed chemical mechanism for the reactions of representative alkanes, alkenes, aromatics, and their major photooxidation products was developed, using the mechanism of Atkinson, Lloyd and Wings (ALW) (Atkinson et al, 1982) as the starting point. This detailed mechanism was tested against the results of over 490 environmental chamber experiments carried out in the outdoor chamber at the University of North Carolina (UNC), and in two indoor and one outdoor chambers at the Statewide Air Pollution Research Center (SAPRC) in Riverside, California. Testing this mechanism required the development of a consistent chamber characterization model for the four different chambers. This is, I believe, the first time that a single reaction mechanism has been tested against such a large and comprehensive data base of chamber experiments with a consistent set of assumptions regarding chamber effects. This discussion will concern primarily the testing aspect of the mechanism, but I will also briefly summarize the work done on updating and extending the chemical mechanism.

The second phase of this program involved primarily the adaptation of the mechanism developed and tested in the first phase for use in AQSM models. In this phase, the detailed mechanism was condensed, using the "surrogate species" approach, so it is more appropriate for use in AQSM models. The sensitivities of control requirements, predicted using the EKMA technique, to input data were also assessed. Finally, the appropriate

procedures and defaults for using these mechanisms were established and documented. This phase of the program is the subject of Fred Lurmann's presentation, so it will not be discussed further here.

Major Features of the Updated Mechanism

The chemical mechanism which was developed and tested in this study includes detailed reaction schemes for representative alkanes, alkenes, and aromatic hydrocarbons, and for their major oxygenated and organic nitrate products. It includes explicit reaction schemes for all compounds for which single component environmental chamber data are available. This required adding compounds in this mechanism which were not included in the ALW mechanism, such as, for example, 1-butene and 1,3,5-trimethylbenzene, as well as several individual alkanes which are not represented explicitly by ALW.

The detailed mechanism developed in this program is intended to serve as a "master mechanism", against which more condensed mechanisms (which may be more practical to use in AQSM's) can be compared. This is based on our belief that the best approach is to use as accurate and detailed a mechanism as possible when testing against chamber data, so what is being tested is the chemistry and not the lumping techniques, and that the best approach for testing lumping techniques is to compare predictions of lumped models against the detailed mechanism. This is why the mechanism included explicit representations of the species in the chamber experiments used to test it. In addition, where practical, other chemical approximations were minimized. For example, the ALW mechanism ignored reactions of organics with $O(^3P)$ atoms and NO_3 radicals, which limited the range of validity of that model. These reactions were included in this mechanism. On the other hand, condensation techniques which do not include significant chemical approximations, such as "lumping" of complex reaction schemes into a single overall process which has the same overall effect, were employed to keep the mechanism down to a manageable size. In addition, reactions which are obviously of negligible importance under any reasonable range of atmospheric or environmental chamber conditions were excluded from the model.

Although known chemical approximations were minimized in this mechanism, the mechanism had to represent processes whose details are unknown. Other than perhaps the representation of chamber effects (discussed later), the greatest area of uncertainty concerns the ring opening processes in the aromatic photooxidation mechanism, and the nature and reactions of the highly reactive products formed. Recent laboratory data have indicated that previous mechanisms for these processes are incorrect, and that we actually know less than we thought we did previously. In such cases, the development of this mechanism was based on the philosophy that "if you have to represent something in your model that you don't understand, simpler is better". This is our philosophy with regard to representation of chamber effects (discussed below), and it applies equally well to our representation of the formation and reactions of uncharacterized aromatic ring-opening processes. Thus, rather than

attempting to devise some type of detailed speculative mechanism for these processes, the unknown aspects of the aromatic photooxidation mechanism was represented in a parameterized manner, with two surrogate species used to represent the reactions of the of the (probably) many uncharacterized ring-opened products, and with the yields and photolysis rates of these products being adjusted to fit the chamber data. This is obviously an unsatisfactory situation, but until more detailed and quantitative mechanistic and product yield data are available, we really have no other choice.

Another area of chemical approximation which current models in practice cannot avoid is the representation of the peroxy + peroxy radical reactions, which become important in the absence of NO_x or, under some conditions, at nighttime. Test calculations we have carried out have shown that neglecting these reactions entirely is not an acceptable approximation under those conditions. However, representing these explicitly in the model not practical, because of the many different types of peroxy radicals involved, and the fact that such a representation would require all the possible cross-reactions to be included. Thus, in this model, we employed an technique for representing these processes in a relatively efficient manner, which requires the addition of only a few reactions and species to the model. I will not discuss the details of this method further here, except to note that is a similar type of approach to that employed by Lurmann et al. (1986) in the ADOM mechanism.

Another area of approximation which for all practical purposes is necessary is the representation of the many different types of bi- and poly-functional oxygenated products which are expected to be formed in the reactions of the higher alkanes (Carter and Atkinson 1985). These were represented by a limited number of species in the model, according to a modified "surrogate species" approach. For example, "RCHO", or propionaldehyde, was used to represent the reactions at aldehyde groups of such species, and "MEK", or methyl ethyl ketone, was used to represent reactions at other carbonly or -OH groups. This is in line with the representations used in other models, and a more detailed representation of the reactions of these bi- and poly-functional model would significantly increase the complexity (and level of speculation) of the model without significantly affecting its predictions.

I should point out that the current version of the mechanism, which was used for the calculations whose results I will summarize in this presentation, is different in some respects than the mechanism documented in detail in our published report on Phase I of this program (Carter et al. 1986a). The current version of the mechanism has a somewhat more efficient representation of the peroxy + peroxy reactions than employed in the previous model, though this does not have a substantive effect on the results of model predictions. The most substantive difference is that the methods used to represent the unknown aspects of the aromatic photooxidation mechanism was modified to represent them in a more efficient manner. These changes are documented in the Phase II report on this program (Lurmann et al. 1987). However, in most other respects, the

mechanism whose results are discussed here (and in Fred Lurmann's presentation) is the same as that documented in detail in our previous report (Carter et al. 1986a).

Philosophy Used in Testing Mechanism

When testing this mechanism against the chamber data, and in some cases modifying the mechanism as a result of the fits obtained, we attempted to adhere to the following philosophy, which we believe is appropriate in carrying out such a program.

- o Where practical, all reactants in the experiments were represented explicitly. This is because we are testing our best estimate of the chemical processes involved, and not techniques for approximating them. As indicated previously, testing approximation techniques can be done much more appropriately (and with a higher degree of sensitivity) by model calculations using the detailed mechanism as the standard.

- o A consistent set of assumptions was used in modeling all runs. The same chemical mechanism was used for all runs, and a self-consistent set of assumptions regarding chamber effects were employed. If, during the process of model development, we made any change to the chemical mechanism or how we assigned values for chamber-dependent parameters, then all the runs were re-calculated. I might note that complete re-calculations of all the runs were done several times throughout this program.

- o No run-to-run adjustment of parameters was done in order to improve the fits for individual runs. The input data for modeling an individual run was based entirely on experimental measurements (or in some cases the recommendations of the experimentalists) and on assignments of chamber dependent parameters for groups of runs. [By "groups of runs" in this context we mean all runs carried out in a given chamber, or (in the case of light characterization for the UNC chamber) all runs carried out in a given chamber in a given year.] This approach might be questioned by other modelers, since in reality some chamber-dependent parameters may in fact vary from run-to-run in a manner that cannot be determined a-priori. However, in our opinion, run-to-run adjustments of values of unknown parameters invalidate the entire purpose of using chamber data for model testing, since such a procedure might well result in disguising systematic errors in the chemical mechanism being tested. If parameters do indeed vary from run-to-run, this variability will show up as variabilities in the quality of the fit of the model to the experiment, where average or "typical" values for the variable parameters are used in the model. If there are enough experiments, then an examination of the pattern of these discrepancies might reveal whether there are systematic problems with the mechanism, or whether the only problem is variability of conditions. Fortunately, in many cases, there are now enough experiments that such an approach based on distributions of fits, rather than fits to individual runs, can profitably be employed.

o In line with the philosophy that "simpler is better" when representing effects that are not understood, the various poorly-understood chamber dependent parameters were represented in this model as simply as possible, with the absolute minimum number of parameters necessary to describe the data. Limiting the number of parameters also greatly simplified the process of determining the sets of values which best fits the data. When using almost 500 experiments in a 1-2 year mechanism testing program, a detailed analysis of the type of multi-parameter chamber effects model which will result in the very best possible fit for each individual experiment is obviously not practical. Even if it were, at best the resulting speculative model would give the illusion that we understand the details of processes which in fact are poorly characterized, and at worst it may (if there were enough adjustable parameters) have the effect of canceling out errors in the chemical mechanism in the simulations of the chamber experiments.

o Only runs with needed data missing were rejected for use for model testing. The types of data considered to be essential were determined prior to carrying out the test calculations, and included light characterization data in outdoor chamber experiments, spectral distribution data in SAPRC Evacuatable Chamber (EC) runs, and initial concentrations of all reactive species in the experiment. No run was rejected from the model evaluation statistics based on poor fits alone.

o Although we attempted, as far as possible, to base the model on a-priori estimates of unknown mechanistic and chamber dependent parameters, in some cases adjustment of these parameters in order to attain acceptable fits could not be avoided. However, such adjustment was only done on a global basis, and using only the most appropriate type of runs for a particular parameter. In the case of chamber effects parameters, these were adjusted based on fits to characterization runs or types of runs which are the most sensitive to the values of these parameters. In the case of mechanistic parameters, adjustment was done based on simulations of single component - NO_x - air runs only. No adjustment of parameters was done to improve the fits of the model to results of experiments containing mixtures of organics.

Summary of Parameters Adjusted in the Course of Model Testing

The mechanistic and chamber dependent parameters which were adjusted during the course of the development and testing of this model, and the types of experiments used as a basis for this adjustment, are as follows:

Mechanistic Parameters

o The yields and photolysis rates for the species used in the model to represent the uncharacterized ring-opened products formed in the reactions of the aromatic hydrocarbons were adjusted based on the model simulations of selected the benzene, toluene, m-xylene, and the 1,3,5-trimethylbenzene - NO_x - air runs carried out in the SAPRC indoor chambers. Runs with these

compounds in outdoor chambers were not used because of greater uncertainties in characterizing the conditions of such runs, and the much greater computational requirements which would be involved in using such runs in our nonlinear optimization program. The results obtained in the optimizations based on the indoor chamber experiments were evaluated by comparing the adjusted mechanism against the results of the UNC toluene and o-xylene - NO_x - air, and were found to be generally satisfactory in that they did not indicate large systematic discrepancies in the simulations of the outdoor chamber runs.

o The radical yields in the ozone + propene and the ozone + isobutene reactions were adjusted based on fits to propene or isobutene - NO_x - air experiments carried out in the SAPRC Indoor Teflon Chamber (ITC).^x Despite many studies of ozone - alkene reactions, the available data are still inadequate to determine unambiguously this essential aspect of the mechanisms for alkenes other than ethene. The ITC was employed in adjusting these yields because, of the two indoor chambers, it has the lower chamber radical source, making such runs more sensitive to radical input from homogeneous reactions. The radical yields in the ozone + propene reaction had to be adjusted downward from our initial estimate, but it is not greatly different than those used in the ALW mechanism. The adjustments of the ozone + isobutene mechanism, which is totally unknown, is based on simulations of a single experiment, and thus the value obtained must be considered to be highly uncertain.

o The photodecomposition quantum yield in the photoysis of methyl ethyl ketone was adjusted, based on simulations of a very limited number of UNC chamber experiments. Previous models have assumed a unit quantum yield, but that results in significant overprediction of the reactivity observed in these runs. However, in view of the limited number of experiments employed, and the uncertainties in light characterization in outdoor chamber experiments, the results must be considered to be highly uncertain.

o Based on analogy with reactions in the n-butane photooxidation system, it was initially estimated that the OH + butene reactions should result in approximately 10% organic nitrate yields. Assuming this resulted in consistent overpredictions of reactivity in simulations of indoor chamber experiments, so this nitrate formation was removed from the model. However, it should be pointed out that this is based on a limited number of runs which were not well fit in any case, and the model significantly overpredicted the reactivity in the few UNC 1-butene experiments, regardless of which option was used.

o The photooxidation mechanism of alkanes with chains of four or more carbons involve the formation of OH-substituted peroxy radicals, which may react with NO to form alkyl nitrates. The nitrate yields in the reactions of the OH-substituted peroxy radicals are not known, though they are known in the case of unsubstituted peroxy radicals formed in the initial OH + alkane reactions. If it is assumed that the nitrate yields in the reactions of OH-substituted radicals are the same as in unsubstituted radicals, as has generally been the case in previous models for the higher alkanes, we found that the model had a consistent tendency to overpredict

the reactivity in n-hexane, n-heptane, n-octane, and n-nonane - NO_x - air runs. Therefore, the model was modified to assume that the reactions of NO with OH-substituted peroxy radicals does not involve the formation of nitrates. Since this is a radical termination process, the resulting model predicts greater reactivities for these alkanes, which are generally more in line with the results of the chamber experiments, at least on the average. However, as seen later, the model is the most variable in the simulations of the alkane - NO_x - air experiments, which is attributable to the sensitivity of these runs to the variabilities in the chamber radical source. Therefore, any conclusions based on fits of model simulations to results of alkane - NO_x - air experiments are subject to a relatively wide degree of uncertainty, and the possibility that the initially estimated model might be more correct cannot be totally ruled out.

Chamber Dependent Parameters

The chamber-dependent parameters which were adjusted consisted of the NO_x offgassing rates, which were adjusted for each chamber to fit the results of the acetaldehyde - air irradiations, and the background NO conversion rate (represented by OH → HO₂ in the model), which was adjusted for based on fits to the pure air irradiations. The estimate that initial HONO levels are probably minor in Teflon film chambers was also made based on results of preliminary modeling studies. The other chamber-dependent parameters were obtained or estimated based on analysis of results of characterization experiments done using these or similar chambers, characterization data taken during the experimental runs which were modeled. The derivation of the various chamber-dependent parameters used in this model will be discussed below.

Chambers Whose Data Were Used for Model Testing

As indicated previously, data from four chambers, operated by two different research groups, were used for model testing. All four chambers have been operated for a number of years in experimental programs designed to obtain data for this purpose. The major characteristics and distinguishing features of these chambers are as follows:

The SAPRC Evacuatable Chamber (EC) is a rigid-walled 5800-liter indoor chamber. It is routinely evacuated between experiments. This chamber is also thermostated, and experiments can be carried out at varying temperatures, though most of the EC experiments modeled in this study were carried out at ~303 K. It uses a Pyrex-filtered Xenon-arc "solar simulator" as a light source. Its interior walls consist of FEP Teflon coated aluminum, and it has quartz windows on either end, though on the end opposite to the light source aluminum external reflectors are employed to enhance the light intensity inside the chamber. An air purification system supplies purified air for the experiments employing this and the other SAPRC chambers. The purified air used in EC experiments is usually humidified to ~50% RH. The major characteristics of this chamber, and

experimental procedures employed for most of the EC experiments modeled in this study, are described in a previous EPA report (Pitts et al. 1979.)

The SAPRC Indoor Teflon Chamber (ITC) consists of a replacable ~5800-liter flexible bag constructed of 2-mil thick Teflon film which is held in an aluminum frame. The bag is replaced periodically, typically every few months, or between programs involving different types of experiments. The light source consists of banks of blacklights on either side of frame holding the Teflon bag. Although not thermostated, the temperature in the chamber is held relatively constant at ~300 K by means of a cooling system which is employed when the lights are turned on. The same air purification system as employed in the EC is used to supply air for this chamber, which is also usually humidified to ~50% RH. The major features of this chamber, and the the experimental procedures employed, are discussed in various recent SAPRC reports (Carter et al., 1984, 1985, 1986b).

The SAPRC Outdoor Teflon Chamber (OTC) consists of a replacable ~50,000 liter flexible bag also constructed of 2-mil thick film which held on a frame with a network of ropes. The reaction bag is replaced after approximately 1-2 months of use. Natural sunlight is used as the light source. The chamber has no temperature control. The chamber is covered by an opaque tarp between experiments and when the reactants are being injected, and the tarp is removed to begin the irradiation. A typical experiment begins at ~0900 PST and ends at ~1500 PST, and the bag is usually covered overnight in multi-day runs. The bag can be optionally divided in half, to allow sumultaneous irradiation of two mixtures, though it can also be operated in the undivided mode. Dry purified air is used in most experiments with this chamber. This chamber, and experimental procedures employed, are discussed in several SAPRC reports (Carter et. al, 1985, 1986b).

The UNC Outdoor Chamber is constructed of 5-mil thick FEP Teflon film held on a wooden framework. This chamber has two sides, each ~150,000 liters in volume, allowing two different mixtures to be irradiated simultaneously. This is the largest of the chambers whose data are used in this study. Unlike the SAPRC Teflon chambers, the Teflon film walls are rarely replaced. Natural sunlight is used as a light source. Reactants are usually injected before sunrise on the morning before the experiment, and the irradiation begins when the sun rises. The chamber is located in a rural area, and relative clean, unpurified ambient air is used in this chamber. In some, but not all, experiments the humidity inside the chamber is reduced using dehumidifiers to reduce condensation of moisture on the wall. This chamber, and the operating procedures employed, are described in various UNC reports (e.g., Jeffries et al. 1982, Sexton et al, 1987).

Derivation of the Chamber Characterization Model

The testing of a chemical mechanism with chamber data requires appropriate specification of chamber-dependent parameters to represent

effects which may vary from chamber to chamber, and depend on the conditions of specific experiments. The types of effects and chamber dependent input which are represented in the chamber characterization model used when our mechanism was tested are as follows:

- o Intensity, spectral distribution, and (for outdoor runs) time variation of the light source
- o The magnitude of chamber radical source, and its dependences on light intensity and NO_2
- o Initial levels of nitrous acid (HONO), if any
- o Ozone dark decay rates
- o NO_x offgassing rates
- o Rates of heterogeneous hydrolyses of N_2O_5 and of NO_2
- o Excess NO oxidation rates, caused (presumably) by background or contaminant reactive organics.

In this portion of this presentation, I will summarize how these various effects are represented in our chamber characterization model, and indicate how the various parameters employed were derived. For a more detailed discussion of this, interested persons should consult our Phase I report (Carter et al. 1986a).

Light Source Characterization: Indoor Chambers

Characterization of the light source for modeling indoor chamber experiments requires a knowledge of its intensity and its relative spectral distribution. Runs without such information were not used for model testing. In both indoor chambers (the SAPRC EC and the SAPRC ITC), the light intensity is obtained from results of NO_2 actinometry experiments, which are carried out periodically. In the SAPRC EC, the spectral distribution is measured during the course of most runs. For this testing program, the runs were divided into groups based on when the runs were carried out, the lamp employed, and similarities in measured spectral distributions, and the average spectral distribution for each group was used for all runs in the group. The spectral distribution in the EC varies over time (with the most recent runs having significantly less UV intensity), and EC runs carried out before the spectral distribution was routinely monitored were not used for model testing. In the case of the SAPRC ITC, the spectral distribution of the blacklights was measured ~3 times in separate experiments, and does not appear to change significantly with time. One spectral distribution was used in modeling all ITC runs.

The spectral distribution of the ITC, and representative spectral distributions for the EC are shown on Figure 1. For comparison, the $\lambda=40$ solar spectrum is also shown. The degradation of the EC spectral distribution over time, and the relatively low intensity in the ITC at

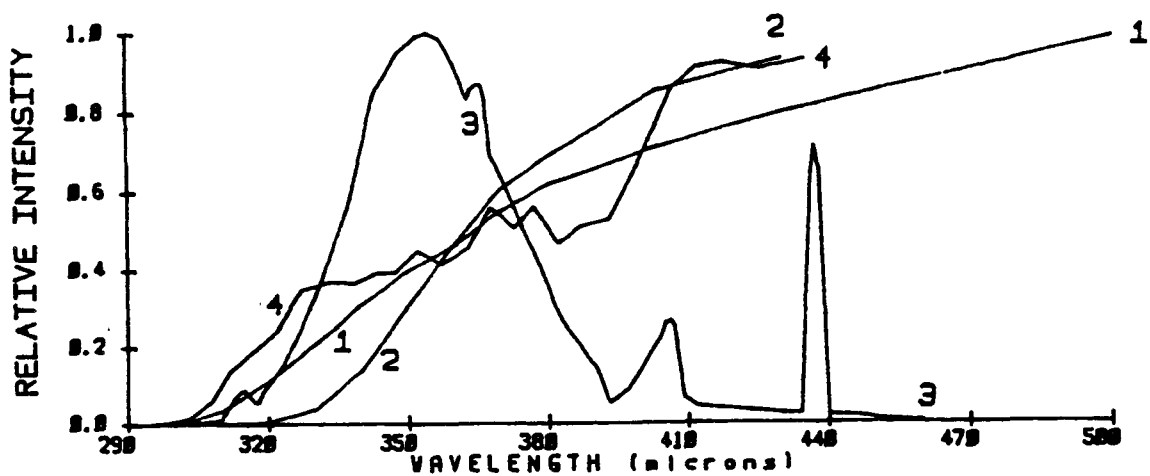


Figure 1. Representative Spectral Distributions for Various Light Sources, Normalized to Give the Same NO_2 Photolysis Rates.

- 1 - SAPRC solar simulator output around the time of run EC-340
- 2 - SAPRC solar simulator output around the time of run EC-900
- 3 - SAPRC ITC blacklight output
- 4 - Tropospheric, ground level solar spectrum for zenith angle = 40° .

wavelengths above ~390 nm (the wavelength region responsible for alpha dicarbonyl and NO₃ photolysis), is apparent from this figure.

I might note that although the ITC spectral distribution at the higher wavelength region is not particularly representative of sunlight, these differences have provided us useful information regarding the wavelength regions which affect the photolysis of the aromatic ring opening products. Aromatic runs carried out in the ITC are much more reactive than predicted by models which assume that the photoreactive products are alpha dicarbonyls (as assumed in most previous models) or compounds which photolyze in similar wavelength regions, indicating that these unknown products must photolyze at lower wavelengths than do the alpha dicarbonyls.

Light Source Characterization: Outdoor Chambers

Characterization of the time-varying light intensity and spectral distributions of outdoor chamber runs is much more difficult, and subject to much more uncertainties than is the case for indoor chamber experiments. We were not able to carry out as comprehensive an analysis of this in the limited time frame of this program as was really required, nor did we have the benefit of the recent work Harvey Jeffries has done in this area. Therefore, I do not think our model for light characterization in outdoor chambers should be taken as the last word on this subject, and, at least with regard to light characterization for the UNC outdoor chamber, is probably superceded in some respects by Harvey's analysis. However, in evaluating the performance of our model in simulating outdoor chamber runs, and understanding the uncertainties involved, it is important to understand how we derived our light characterization model when simulating these runs.

The photolysis rates used when modeling outdoor chamber runs are obtained by multiplying together the following factors:

- o The theoretical photolysis rate for the reaction, calculated using the absorption and quantum yields assigned for it in the homogeneous chemical mechanism, calculated using Peterson's (1976) "Best Estimate" Actinic fluxes. This is a function of (a) the clock time; (b) a clock correction, discussed below; (c) the latitude of the chamber; and (d) the date of the experiment.
- o The ratio of the experimental UV-radiometer measurements to calculated UV-radiometer values which correspond to Peterson's (1976) theoretical light intensity and spectral distribution. The calculated UV values were based on the spectral response of the radiometers used and on empirically adjusted direct vs scattered radiation factors. This is a function of (a) the UV data from the run; and (b) all the factors listed above, which affect theoretically calculated photolysis rates.
- o A constant factor which relates the calculated z=0 UV radiometer readings to the theoretically calculated z=0 NO₂ photolysis rates. The value used, which is 152.9 milliwatt cm⁻² min,

is derived from published results of simultaneous UV and NO₂ photolysis rate measurements. The net effect of this factor is to relate the NO₂ photolysis rates calculated for the outdoor chamber runs to those experimentally measured in the atmosphere, by means of UV data.

- o A correction factor for errors in calibration of the UV radiometer, where applicable. This is a function of (a) the chamber employed and (b) the date of the experiment.
- o A correction factor for the differences in light intensity inside and outside the chamber. Although there may be (and, for the UNC chamber, probably are) wavelength and zenith angle dependences for this factor, these were ignored in this model, since at the time it was developed, no reliable information concerning this was available. This is then a function only of the chamber employed in this model.

The clock correction indicated above is due to the fact that the clocks in neither chamber are not always totally accurate. This error is manifested by the peaks in the UV and TSR readings on clear days being at different times than theoretically calculated. These corrections which, as expected, tended to be constant for runs done around the same time, were obtained by fitting the shapes of the observed and calculated UV and TSR curves. This correction also takes into account corrections for longitude effects.

An analysis of the UV data from the UNC chamber runs carried out over the years indicate a need for a UV correction factor for certain years. A measure of the overall light intensity in a given run, as indicated by the run's UV data, is the extrapolated $z=0$ clear sky UV intensity, which is obtained by curve fitting the experimental clear-sky UV data to calculated UV values, and then extrapolating the calculated values to $z=0$. This factors effects of zenith angle on light intensity. A plot of the calculated $z=0$ UV factors against the date of the run for the UNC and OTC chamber runs on our data set is given in Figure 2. These factors can be seen to vary, and tend to be higher later in the year, but for the UNC values for 1982 and later runs and for the 1983-1984 OTC runs, they are reasonably consistent with each other, and on the average are near the theoretically expected value of 70.2 milliwatt cm⁻². However, the values for previous years in the UNC chamber tend to be anomalously low.

These results suggest probable systematic calibration errors in the UNC UV instrument during this period. This is confirmed by Harvey Jeffries more recent and comprehensive analysis of the UNC UV data, and the history of the UV instrument employed. The correction factors we employed in our model simulations of pre-1982 UNC chamber experiments were estimated by ratioing the averages of the June-September extrapolated $z=0$ UNC UV values to those from the 1982 - 1984 runs. These are: 1.30 for 1978; 1.31 for 1979; 2.06 for 1980; and 1.24 for 1981. These factors are reasonably consistent with the correction factor recently recommended by Jeffries.

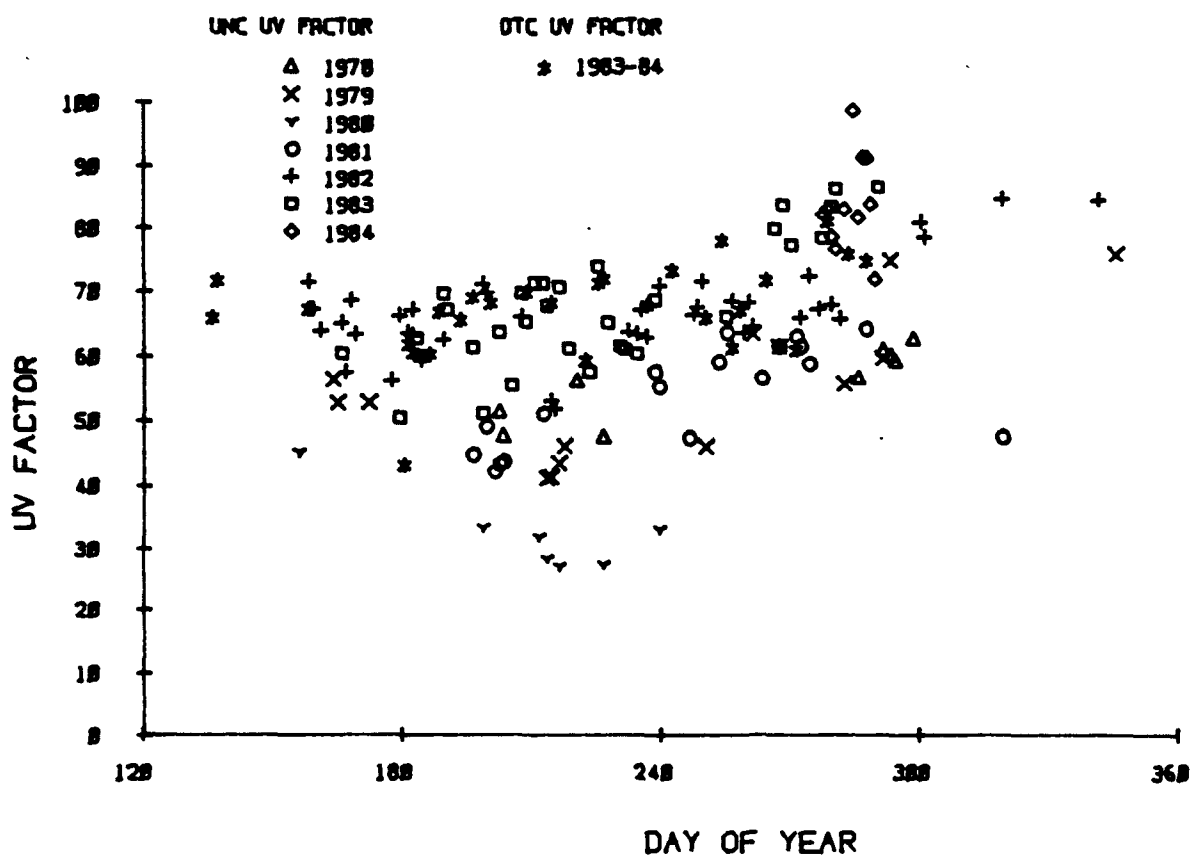


Figure 2. Plots of Extrapolated Z=0 UV Radiometer Readings (in milliwatts cm^{-2}) Against Day of the Year for UNC and SAPRC Outdoor Chamber Runs Modeled in This Study.

However, Jeffries estimates are based on a more comprehensive analysis and should be used in any future modeling program.

The problem of relating the light intensity inside the UNC chamber to the intensity outside (which is measured by the UV radiometer) probably represents a greater uncertainty in the light characterization for that chamber. Jeffries' recent analysis indicates that the walls of this chamber attenuate light in a wavelength-dependent manner, and that the reflective floor enhances light in a manner which probably depends on zenith angle. However, we did not have the benefit of this analysis at the time we had to develop our UNC light characterization model, and essentially no useable data were available concerning these effects. [The NO_2 photolysis rate measurements made inside the UNC chamber by Saeger (1977) are too scattered to be of any use in this regard.] For lack of better data, and in line with our philosophy that "simpler is better" in representing what one doesn't understand, we assumed that the opposing effects of light attenuation by the walls and the enhancement by the reflective floor exactly cancel, and thus no correction factor was used. However, future modeling of UNC chamber runs should re-evaluate this, based on the results of Jeffries' recent evaluation, and any new data that might be obtained in this regard.

The uncertainties in estimating the inside vs outside light intensities for the SAPRC OTC appears to us to be somewhat less, because (a) the SAPRC OTC does not have a reflective floor (the chamber is over a green indoor-outdoor carpet) and (b) the Teflon walls are changed periodically, which should prevent light-absorbing contaminants from building up. Thus we assume that the only effect that needs to be taken into account is the attenuation of the by the walls, and that this attenuation does not depend on wavelength. A correction factor of 0.84 was estimated based on the fact that maximum NO_2 actinometry values obtained from measurements made underneath the chamber (with the light passing in and then out of the chamber) are ~70% of the theoretically expected values. Assuming that the light reaching the actinometer is suppressed by the same factor when it goes in as when it goes out, this corresponds to the assumed correction factor of 0.84, the square root of 70%. Zenith angle effects, if any, are ignored. These estimates are obviously also subject to uncertainty, and may need to be re-evaluated.

At this point, I might note that of all the runs used for model testing, the UNC formaldehyde irradiations are probably the most affected by the uncertainties in our outside chamber light characterization model. This is because these runs are driven by formaldehyde photolysis, which is affected more uncertain lower wavelength radiation, and because in the UNC experiments, the run is initiated by the photolysis of formaldehyde at very high zenith angles, when the uncertainties in the light characterization are the greatest. These uncertainties are less in the SAPRC OTC runs because the irradiations begin later in the day when the sun is higher, and the uncertainties in light characterization are less. Indeed, the results of our model simulations of UNC formaldehyde - NO_x - air irradiations suggest, though do not prove, that there may be problems with our UNC light characterization model. Concentration-time plots for selected species in

HCHO - NO_x - AIR RUN

AU0279B

* EXPERIMENTAL
— MODEL

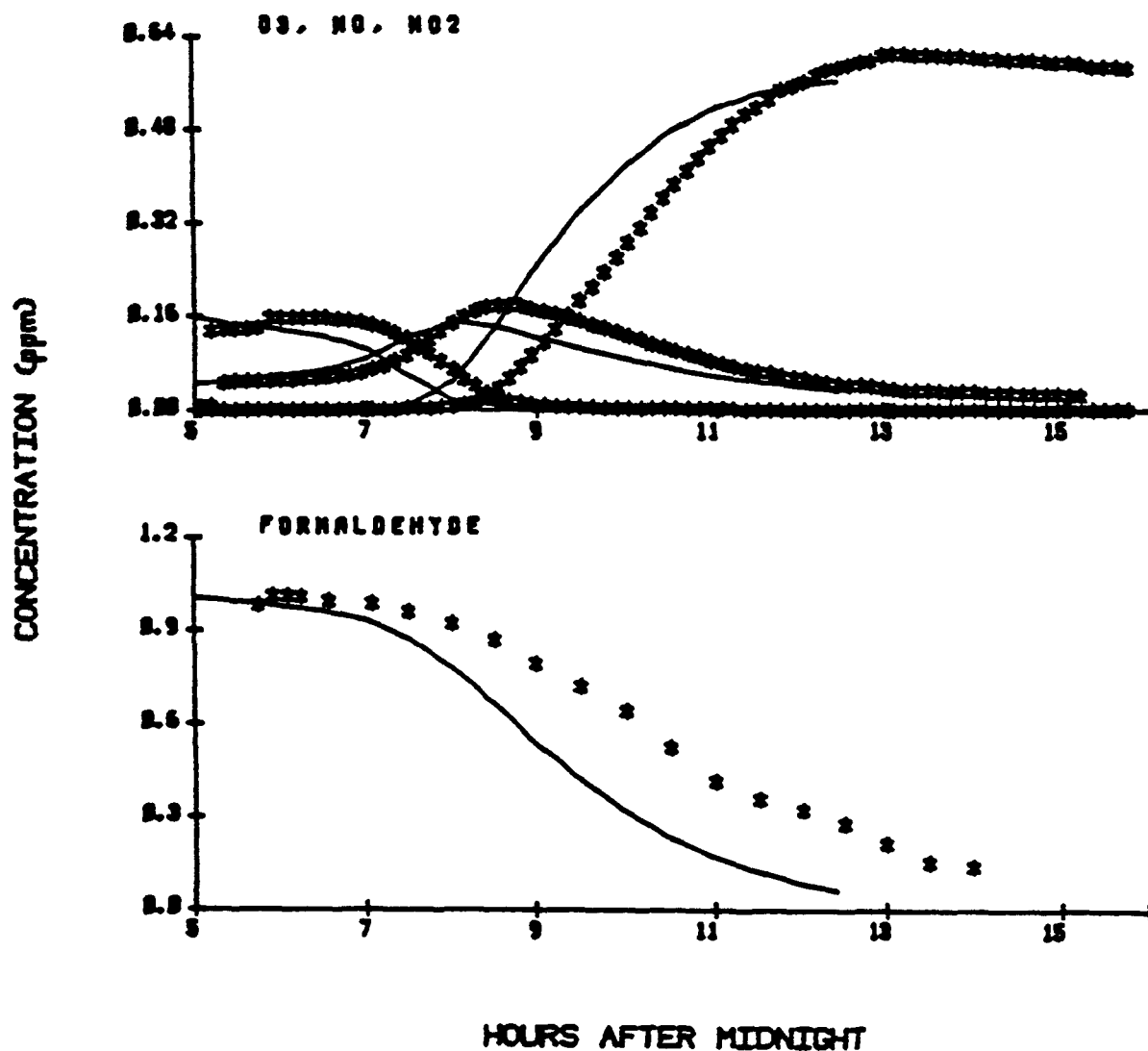


Figure 3. Experimental and Calculated Concentration - Time Plots for Selected Species in UNC Formaldehyde - NO_x - Air Chamber Run AU0279B

HCHO - NO_x - AIR RUN

JL2381B

* EXPERIMENTAL
— MODEL

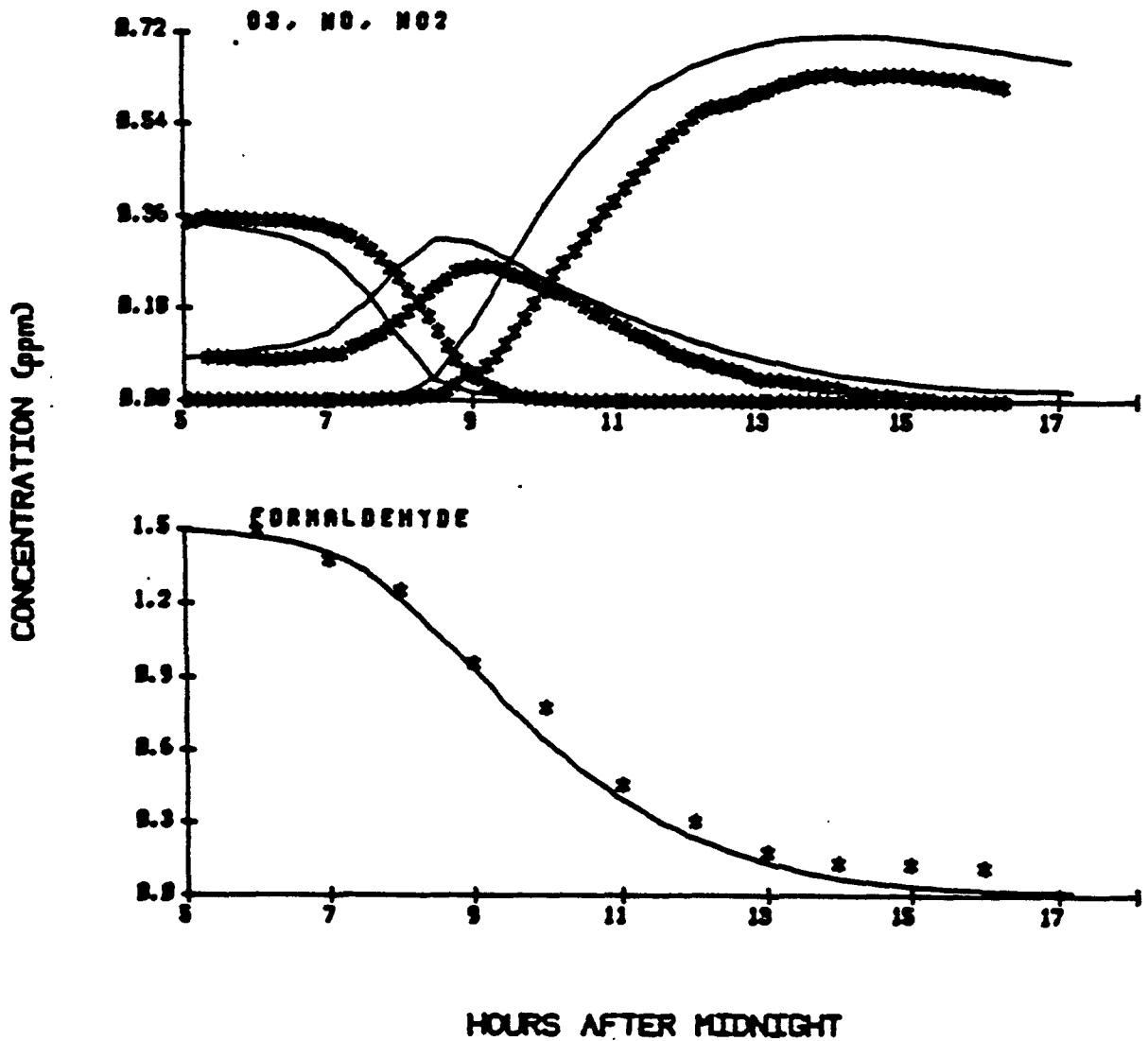


Figure 4. Experimental and Calculated Concentration - Time Plots for Selected Species in UNC Formaldehyde - NO_x - Air Chamber Run JL2381B.

two such runs are shown on Figures 3 and 4. Figure 3 is more typical of most of the UNC formaldehyde runs on our data set, and indicates a significant overprediction of the initial rate of formaldehyde photolysis, and a corresponding overprediction of the initiation of ozone formation. The run shown on Figure 4 is somewhat atypical in that this is the only run in our data set where the model almost fits the formaldehyde decay rate (though not the initiation time for ozone formation). However it is interesting in that it is the same run that Harvey shows as being perfectly fit by his chamber characterization model. It is my understanding that he used the same formaldehyde chemistry as in our model.

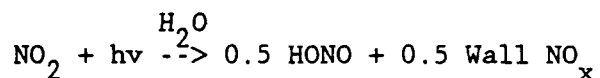
However, the problems we have in fitting the UNC formaldehyde runs do not prove that our UNC light characterization model is necessarily invalid. Formaldehyde is a difficult compound to handle experimentally and to analyze reliably on a routine basis. As shown by the model fits discussed later, there may be problems with the UNC formaldehyde analysis technique, since the UNC formaldehyde data are consistently higher than model predictions in simulations of chemical systems where the model predictions are consistent with SAPRC and Unisearch formaldehyde measurements. Also, the model has the opposite discrepancy in the simulations of the UNC acetaldehyde experiments, where it has a tendency to underpredict the aldehyde consumption rate. The simulations of the other UNC runs do not indicate the types of consistent problems we see with the formaldehyde runs, but these runs in most cases are less sensitive to the assumed light characterization model.

Representation of the Chamber Radical Source

The existence and importance of chamber-dependent radical sources is now reasonably well established, so I will not go into a detailed discussion of this here. In this context, by "chamber radical source", I mean radicals from sources other than from HONO formed from the dark NO_2 hydrolysis reaction (which will be discussed separately with the other known heterogeneous reactions) and other than from HONO which may be initially present in the experiments. In our model, the chamber radical source is represented by NO_2 -independent process, which is represented as a "pure" OH source (since we do now know what the correct representation is)



which is assumed to occur in all four of the chambers. In the SAPRC EC, there is evidence that the chamber radical source has an NO_2 dependence (beyond that expected due to NO_2 hydrolysis), and this is represented in the model by the same type of process as used to represent NO_2 hydrolysis, i.e.,



where the radical source is due to the rapid subsequent photolysis of HONO.

There is no clear evidence for such an NO₂ dependence for the chamber radical source in the other chambers, so this was not used for the chambers other than the EC.

Other modelers represent the chamber radical source as formaldehyde offgassing, but this is not assumed in our model. We find that if this is assumed, then the rates of NO to NO₂ conversion in NO_x - air characterization runs are consistently overpredicted.^x However, I might note that in terms of model simulations of most other runs, it does not really make much difference whether the chamber radical source is represented as formaldehyde offgassing, or as a "pure" OH source as assumed in our model. This is because the excess NO to NO₂ conversions caused by the formaldehyde offgassing radical source model is usually minor compared to the NO to NO₂ conversions caused by the reactive organics present.

The radical source values used in the model simulations whose results are presented here are shown on Table 1. That table also indicates how these values were derived. In all cases, the radical source was assumed to be proportional to the NO₂ photolysis rate, and thus the radical sources shown on that table are given as ratios of radical input rates to the NO₂ photolysis rate.

Table 1. Radical Source Values Used for Model Testing, and a Summary of Their Derivation.

Chamber	RS/kl (ppb)	Derivation
EC	0.6 (NO ₂ =0.2)	Derived from radical source vs NO ₂ regression of Carter et al (1982), derived from tracer - NO _x - air runs at 50% RH and 303 K. The value shown is for a typical NO ₂ value of 0.2 ppm.
ITC	0.15 - 0.6 (0.3 typical)	Derived from averages of radical input rates obtained from tracer - NO _x - air runs carried out in the same reaction bag. Varied from bag to bag. No NO ₂ dependence could be determined and was ignored.
OTC	0.2 - 0.25	
UNC	0.3	Original RS/kl estimate of 0.3 based on average values for ITC and OTC (with similar surface). Not inconsistent with results of UNC n-butane runs.

Initial Nitrous Acid

If nitrous acid is initially present in chamber runs, its photolysis may provide a non-negligible source of radicals. Initial nitrous acid, which is not measured experimentally in any of the chamber runs in our data set (with the possible exception of a few EC NO_x - air irradiations) has been used in previous modeling studies as an adjustable parameter whose value is adjusted from run-to-run to optimize the fits. In this model, such run-to-run adjustment is not done, but instead the initial HONO is assigned based on the chamber employed. In the case of the SAPRC EC, HONO has been measured in a number of tracer - NO_x - air experiments (Carter et al. 1982, Pitts et al, 1983), and the results suggest that under the conditions of the EC runs modeled here, the initial HONO is approximately 7% of the initial NO_2 . This is assumed in this model when simulating EC runs.

For the runs in the other three chambers, we assumed in this study that the initial nitrous acid is negligible. This is based on the fact that studies at SAPRC (Pitts et al, 1984) indicate that HONO formation from NO_2 is much slower in reactors made of Teflon film, as is the case with the ITC, OTC, and UNC chambers, than it is in the SAPRC EC, and thus lower initial HONO levels are expected. Preliminary model simulations of ITC and OTC runs indicate no need to assume initial HONO, nor did we see any evidence that we needed to assume this in simulations of experiments in the UNC chamber. In the case of the UNC chamber, particularly in runs where humidity is not well controlled and water condensation occurs on the walls, the possibility that initial HONO (or HONO offgassed at various times in the middle of the experiment) is playing a role certainly cannot be ruled out. However, the inclusion of such effects in our model simulations requires a detailed run-by-run analysis which was not practical given the number of runs which had to be modeled, and the amount of time available.

NO_x Offgassing Rates

The need to include some type of representation of NO_x offgassing in chamber effects models is evidenced by the fact that ozone $_x$, and in some cases PAN, is formed in environmental chamber runs where no NO_x is injected. This is was represented in our chamber effects model by the following 1-parameter process



where K_N is a zero order rate constant which is assumed to be proportional to the NO_2 photolysis rate. Based on fits to results of acetaldehyde - air irradiations, the values of the proportionality constant, k_N/k^+ , used in our simulations were 0.5 ppb for the SAPRC EC; 0.15 ppb for the ITC; 0.1 ppb for the OTC; and 0.3 ppb for the UNC chamber.

The use of this simple model is preferred over over more complex representations because it requires only one adjustable parameter, and because we have presently no way of knowing whether a more complex

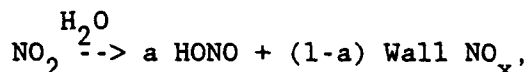
representation is necessarily any more accurate. However, it should be noted that this simple model does not work very well in simulating the acetaldehyde - air irradiations in the outdoor chambers, since in general a different value of k_N/k_1 fits the ozone yields than fits the PAN data. An intermediate value between that that best fit PAN, and that that fit ozone, was used in this model.

It should also be noted that the assumption that the NO_x is offgassed in the form of NO_2 is arbitrary, and the possibility that it is being emitted as NO , or even as HONO , cannot be ruled out. Indeed, this effect may well be related to the chamber radical source, as recently suggested by the SAI modelers. The similarity in the values of the k_{RS}/k_1 and the k_N/k_1 parameters independently derived for the various chambers suggests that such a connection might exist. However, no connection between NO_x offgassing and the chamber radical source was assumed in this model.

Rates of Heterogeneous Hydrolysis Reactions

The heterogeneous hydrolyses of NO_2 and N_2O_5 are also represented in our chamber effects model. The former process contributes to chamber dependent radical input, while the latter amounts to a NO_x sink which may be important in affecting maximum ozone yields. The rates of both of these processes have been measured in the SAPRC evacuable chamber and in a ~4300 liter indoor chamber constructed of the same 2-mil Teflon film as the SAPRC ITC and OTC. The rates observed in the ~4300-liter Teflon chamber were used to estimate their rates in the ITC, OTC and the UNC chamber based on the assumption that these rates are approximately proportional to the surface/volume ratio.

The heterogeneous hydrolysis of NO_2 is represented in our chamber model as the following overall process:

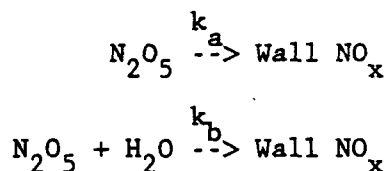


where a is the HONO yield, which is based on the experimental data of Pitts et al (1984) for the EC and ~4300 liter Teflon chamber. The HONO yield parameter for the ITC, OTC, and the UNC chamber was assumed to be the same as that observed in the ~4300 liter Teflon chamber. The data of Pitts et al (1984) indicate that this process is first order in NO_2 , and the representation used in this model is consistent with this. The values of these parameters used in the model for the different chambers are summarized on Table 2.

Table 2. Parameters Used in the Representation of Heterogeneous NO_2 Hydrolysis in the Model

Chamber	Rate Constant (min^{-1})	HONO Yield	Derivation
EC	2.8×10^{-4}	0.5	Data of Pitts et al (1984)
ITC	1.4×10^{-4}	0.2	Measured by Pitts et al (1984) for a 4300-liter Teflon film chamber. K assumed to be proportional to the surface/volume ratio
OTC	0.9×10^{-4}	0.2	
UNC	0.5×10^{-4}	0.2	

Plots of the N_2O_5 hydrolysis rates observed in the EC and the ~4300 liter Teflon chamber against water concentration tend to be linear with non-zero intercepts (Tuazon et al. 1983). Thus N_2O_5 hydrolysis in this model is represented by two overall processes:



The values of the k_a and k_b parameters used in our model are summarized on Table 3.

It should be noted that the representation used for the N_2O_5 hydrolysis shown on Table 4 is based on the assumption that the hydrolysis of N_2O_5 to gas phase HNO_3 , observed by Tuazon et al. (1983) is a homogeneous, and not a chamber-dependent, process. This is in turn based on the assumption the heterogeneous hydrolysis would form HNO_3 on the walls, and once HNO_3 is on the walls, it stays there. Thus the rate constants shown on Table 4 are derived from the difference between the total N_2O_5 loss rates observed by Tuazon et al (1983) and the observed formation rates of gas phase HNO_3 . (The N_2O_5 hydrolysis process forming gas phase HNO_3 is included in our homogeneous inorganic transformation model, and is thus not part of the chamber effects model.) However, more recent data obtained at SAPRC indicate that this assumption that formation of gas phase HNO_3 necessarily indicates a homogeneous reaction may be erroneous, and that this aspect of the model may need to be re-evaluated (Atkinson et al. 1986).

Ozone Wall Loss

Finally, the model must take to account the fact that ozone is lost due to dark decay on the walls. Fortunately, ozone dark decay rates have been

Table 3. Kinetic Parameters Used in the Representation of Heterogeneous N_2O_5 Hydrolysis in the Model

Chamber	k_a^{-1} (min ⁻¹)	k_b^{-1} (ppm ⁻¹ min ⁻¹)	Derivation
EC	4.7×10^{-3}	7.2×10^{-7}	Data of Tuazon et al (1983)
ITC	2.5×10^{-3}	0.5×10^{-7}	Measured by Tuazon et al for a 4300-liter Teflon film chamber. Rates assumed to be proportional to the surface/volume ratio.
OTC	1.6×10^{-3}	0.3×10^{-8}	
UNC	0.9×10^{-3}	0.2×10^{-8}	

measured for all four of the chambers whose data are used in this study. The results of these experiments indicate that this process is a simple first order loss process, and this is how it is represented in this model. The range of experimental values for the ozone dark decay rates, and the values assumed in our model, are given in Table 4. The experimental ozone decay rates are somewhat variable, and the dependences of the ozone dark decay rates on such factors as temperature, humidity, and light intensity are not known. However, the ozone decay rates in these chambers, particularly those constructed of Teflon film, are sufficiently low that these variabilities and uncertainties are not regarded as major problem problems in our chamber characterization model.

Summary of Chamber Runs Used for Model Testing

We will now turn our discussion to the results of the model simulations of chamber experiments based on the chemical mechanism and the chamber effects model we have described. As indicated previously, a large number of chamber experiments were used for this purpose, making this, we believe, the most comprehensive testing of a single model ever carried out as a single study. Almost 500 chamber experiments, carried out in the four chambers, were modeled. The experiments ranged in complexity from pure air and NO_x - air irradiations and other characterization runs, to runs with single^x organics, runs with various simple and complex mixtures, and finally runs consisting of irradiations of auto exhaust. A summary of the types of environmental chamber runs which were used for model testing is given in Table 5.

Table 4. Ozone Loss Rates Used in the Model, and Ranges of Experimental Values From Which They Were Derived

Chamber	Loss Rate Used (min^{-1})	Range of Experimental Values (min^{-1})
EC	1.1×10^{-3}	$(0.7 - 1.6) \times 10^{-3}$
ITC	1.3×10^{-4}	$(0.6 - 2.0) \times 10^{-4}$
OTC	1.7×10^{-4}	$(0.8 - 2.5) \times 10^{-4}$
UNC	1.4×10^{-4}	1.7×10^{-4} (Aug 5-6, 1979) 1.2×10^{-4} (Apr 15-16, 1981)

Table 5. Summary of Environmental Chamber Runs Used for Model Testing

		EC	ITC	OTC	UNC
Characterization	Various	10	14	10	37
Single Organic - NO_x	Oxygenates	8	1	2	15
	Ethene	6	2		6
	Propene	15	7	5	22
	Butenes	6	5		5
	n-Butane	14	5	1	7
	C_5 + alkanes	6	8		6
	Toluene	13	2		5
	Other Aromatics	7	13		4
Known Mixtures	2 to 5 Component	21	25		23
	6+ Component	11	20	62	25
Auto Exhaust	Charger and Volare				25
Dynamic	Propene and Mixtures				9
488 Total		117	102	80	189

The specific UNC chamber runs which were modeled were chosen by Marcia Dodge, our EPA contract monitor, and Harvey Jeffries, the leader of the UNC chamber group, based on their evaluation of what is a comprehensive set of appropriate UNC runs to use in this program. Other than the lack of the recently-completed UNC methanol substitution runs, whose data were not ready for us in time for inclusion in this study, these runs are, I believe, a relatively comprehensive and representative set of the best of the UNC chamber experiments whose data have been processed.

The specific SAPRC chamber runs modeled consisted of almost all of the SAPRC runs which were carried out for the purpose of model testing and which are considered to be sufficiently well characterized for this purpose. The main exception are runs which contain compounds which are not represented in this model, such as biogenic organics, naphthalenes, long chain alkenes, and several heteroatom-containing organics. The set of SAPRC runs also includes results of our recently completed methanol substitution ITC and OTC experiments (Carter et al. 1986b), and several other recent ITC runs. These were not included in the set of runs discussed in our Phase I report (Carter et al. 1986a).

Summary of Results of Model Testing

The performance of this model in simulating the results of these experiments is summarized on distribution plots given in Figures 5 - 12 for the various groups of runs (other than the characterization runs). These figures show the distributions of the discrepancies in the model simulations of various types of experimental observations, where perfect fits are indicated by zero discrepancies (points in the middle of the distribution plots), underprediction by the model indicated by negative discrepancies (on the tops of the plots), and overprediction by positive discrepancies, on the bottom of the plots. (For such a large number of experiments, showing distribution plots such as these is a much more practical method of summarizing model performance than showing fits in concentration-time plots for individual runs. Other ways of showing the performance of the model in simulating large numbers of runs given in figures included with Fred Lurmann's presentation.)

For all types of runs, distribution plots for fits to the maximum ozone yield, and to a measure of the average rate of ozone formation and NO oxidation, are given. The latter quantity is defined as the one half of the maximum change in ($[O_3] - [NO]$), divided by the time required to achieve that change, and is used to measure the "timeing" in the experiment. In addition, distributions of fits to yields of PAN, formaldehyde, and in some cases other aldehydes are shown in runs where these are formed and measured, and distributions of fits to reactant organic half lives are shown for runs containing measured organics which react sufficiently rapidly for its reaction rate to be a useful measure of model performance.

I will not give a detailed discussion of the performance of the model in simulating all of these types of runs. Such a discussion is given in our Phase I report (Carter et al. 1986a), and it can be seen from Figures 6 - 14, and from some of the plots given in Fred Lurmann's presentation, how well the model is performing in simulating the various types of runs. However, the following points regarding the performance of the model in simulating the various types of runs are noted:

- o The fits of the model to the propene runs (Figure 6) are interesting since there are a large number of such runs carried out, and, except for the ozone - propene reaction, the chemistry is believed to be reasonably well characterized. It can be that although there are runs which are not well fit by the model, and the model tends to overpredict ozone in the ITC runs and underpredict it in the OTC runs, in terms of fits to ozone, NO oxidation/ozone formation rates, and propene half lives, the model does reasonably well on the average. This model was developed primarily on results of our simulations of SAPRC experiments, and was not adjusted in any way to fit the UNC runs (which I had never modeled prior to this program), and thus the fact that the fits to the UNC runs is of similar quality as fits to the SAPRC EC runs is encouraging. The quality of the fits to the UNC runs appeared to have no relationship to the "light quality" in those runs; some runs on perfectly clear days were not fit well, and vice-versa.

- o The fits to PAN in the propene runs is of much lower quality in the fits to ozone and the "timeing" parameters. A very wide distribution in the quality of fits is obtained for all four chambers, with the model having a consistent tendency to overpredict PAN yields in the outdoor chambers, though not in the indoor chamber runs. Similar wide distributions are obtained in simulations of other types of runs, such as the complex mixture runs shown in Figure 12, though in that case, there is no consistent tendency for the model to underpredict or overpredict PAN in the outdoor chamber runs; it just has a very wide distribution in the fits, with PAN yields in many runs being fit very poorly.

- o The quality of the fits to the formaldehyde yields is the only area where there is a consistent difference between the UNC chamber and the SAPRC chambers in terms of model performance. The fits in the propene and ethene systems are interesting in this regard, because in both chemical systems the formaldehyde yield is reasonably well established, and in both systems relatively high yields of formaldehyde are expected. The model has a relatively wide distribution in discrepancies of fits to formaldehyde yields in the SAPRC propene and ethene runs, but there does not appear to be any consistent tendency to overpredict or underpredict these yields. On the other hand, the model consistently underpredicts the formaldehyde yields observed in UNC chamber runs when measured by the formaldehyde monitoring technique usually employed in that chamber. However, in a few UNC propene runs, formaldehyde measurements were made by Unisearch Associates, inc. using a different technique, and in those cases, the results were in excellent agreement with model predictions. These results can be explained if the standard UNC formaldehyde monitoring technique

tends to give high values when in routine use in chamber runs. I know that technique was found to give good results in recent special intercomparison studies, but this does not mean that there may not be problems when it is used routinely. However, modeling by itself certainly cannot prove that this is necessarily the case.

- o I should also point out the model's consistent tendency to overpredict acetaldehyde yields in the ITC experiment. In this case, I believe it is most likely a problem with the data. We have had problems recently with acetaldehyde calibrations, and in this case, I definitely think the model is more reliable.

- o The ethene runs (Figure 7) are interesting in that in both the EC and the UNC chamber, there are two groups of runs, carried out around the same time. For each chamber, the model tended to overpredict reactivity in all runs in one group, and to underpredict the reactivity in runs in the other. (The model gave good fits to the reactivities in the two ethene runs carried out in the ITC.) This illustrates how apparent variability in run conditions can have an effect on the performance of the model. It also illustrates the potential for misleading results if only a limited number of runs of a given type are used for model testing, and if all the runs were carried out in the same chamber around the same time. If one looked at only one of those five groups of ethene runs (two each in the EC and the UNC chamber, and the ITC runs), one may conclude that the model is consistently "fast" or "slow" (depending on which group), but looking at all five groups together suggests that there is probably no consistent discrepancy.

- o Of all the types of runs modeled, the alkane - NO_x - air runs have the widest distributions of the fits to ozone yields and NO oxidation and ozone formation rates. This can be attributed to the extreme sensitivity of model simulations of these types of runs to the assumed magnitude of the chamber-dependent radical source. Since the actual radical source tends to vary from run to run, but the model simulations used a single set of assumptions regarding it for all runs carried out in a given chamber, this variability results in a variability in the qualities of the fits of the model predictions to the experimental data.

- o A large number of runs with complex, but known, mixtures were modeled in this study, and the fits to these runs are summarized on Figure 12. As with the simulations of the propene and other types of runs for which there are relatively large numbers of experiments, there is a distribution of discrepancies, but there was no large tendency for systematic discrepancy. The distributions of fits for PAN and formaldehyde yields are much wider than is the case for NO oxidation and ozone formation rates, or reactant half lives.

- o Most of the complex mixture runs carried out in the SAPRC ITC and OTC are runs which employ a single 8-hydrocarbon surrogate, which was designed by the modelers at SAI to represent urban emissions. Many of these runs were carried out as part of our study of the effects of methanol substitution, and thus some of these runs had one third of the hydrocarbon

surrogate replaced by methanol or 90% methanol + 10% formaldehyde. There is no significant difference in the model performance in simulating the runs with methanol than in simulating those with only the hydrocarbon surrogate. I should also note that the EC seven hydrocarbon surrogate runs employed a variety of different relative amounts of the reactants, so they represent more than one single mixture.

o Many of the ITC and OTC complex mixture runs modeled were multi-day irradiations. In general, we found that if the model could satisfactorily simulate the results on the first day, it could usually also simulate the results for the subsequent days. If it performs poorly on the first day, then there is little chance that the subsequent days can be adequately simulated. The distribution plots on Figure 13 show only the fits to the day 1 results. An example of how well the model can perform in simulating a two-day, dual chamber, OTC run is shown on Figure 14.

(o The PAN data shown for Side 1 on Figure 14 are also interesting. The PAN is observed to rapidly increase after $t = 6$ hours, when the chamber is covered for the night. This dark PAN formation on Side 1 is reasonably well simulated by the model, and is due to the reaction of NO_3 radicals with acetaldehyde. This illustrates that these NO_3 + aldehyde reactions, which are generally of negligible importance in the daytime, can be important at night, and must be included in multi-day models.)

o Figure 13 shows that the fits of the model to the ozone yields and the NO oxidation and ozone formation rates in the UNC auto exhaust runs is of comparable, or even slightly better, quality as the fits of the model to the UNC complex known mixture runs. This is reassuring, since auto exhaust runs are obviously important for control purposes, and represent the most complex type of mixture studied. These fits were obtained with no adjusting of the mechanism or modifications of the assumptions made regarding the composition of the exhaust mixtures in order to optimize the fits. Fred Lurmann was responsible for the speciation of the auto exhaust mixture. He did an analysis of the data given in the UNC report on these auto exhaust experiments (Jeffries et al. 1985), and the analytical data obtained during the experiments, and came up with compositions for me to use in the modeling of each of these runs. I used the compositions Fred gave me without modifications, and obtained the fits shown on the figure.

Possible Sources of Poor Fits of Model Prediction to Experiment

The results of this evaluation indicate that there are many cases where experimental results are not well fit by the model prediction. While in most cases the discrepancies appeared to be random in nature, there are cases where apparent systematic discrepancies are observed. From the point of view of evaluating a mechanism for use in AQSM's, it is clearly the systematic discrepancies which are of the greatest concern. Possible sources of systematic discrepancies are as follows:

- o Errors in the homogeneous photochemical reaction mechanism
- o Inappropriate representations of chamber effects or light characteristics
- o Systematic measurement errors, especially of initial reactant concentrations or of characterization data
- o analytical interferences

Determining whether errors in the photochemical reaction mechanism exist is clearly the primary goal of this testing program, since such errors would invalidate the AQSM predictions which use this mechanism. While errors in the chamber characterization model would not necessarily affect the validity of AQSM predictions, they could give an indication that a correct mechanism is incorrect, or, worse, that an incorrect mechanism is correct. This is why it is important we make an effort to better understand the fundamental physical and chemical processes which cause such chamber dependent effects.

The possibility that in some cases systematically poor fits of model predictions to experimental data may be due to problems with the data rather than the model also cannot be ruled out. In my experience, in cases where the causes of gross systematic discrepancies in model simulations of a series of experiments have subsequently been determined, it has more often than not been due to problems with the data. Indeed, modeling has frequently been shown to be a useful tool in revealing the existence of experimental problems. Probable problems with the UNC formaldehyde data and known problems with the SAPRC acetaldehyde data are examples of where the experimental data are more suspect than the model predictions. However, it is the experimental problems that we don't know about, rather than those discussed in this presentation, that are of the greatest concern. While environmental chamber experimentalists make every effort at quality control, such experiments involve a very large number of variables, and totally comprehensive quality control is a monumental, if not impossible, task. Significantly improved quality control beyond that currently practiced would require a much greater commitment of funding and personnel to environmental chamber experimental groups than is presently the case.

Random discrepancies are of concern primarily because they limit the level of precision to which the model can be tested. In addition, if there are significant sources of random discrepancies that cannot be eliminated, then a large number of experiments is required for adequate model testing. The most likely sources of such discrepancies, which I suspect might be significant in many of the cases observed in this study, include the following:

- o Random uncertainties or errors in specifying initial reactant concentrations
- o Run to run variation in uncontrolled chamber-dependent effects

- o Inadequate characterization or control of run conditions
- o Incomplete or inappropriate chamber effects model

Improved quality control when carrying out chamber experiments would reduce some, but not all, of these sources of random discrepancy. For example, even if all errors in the analytical and characterization data were somehow eliminated, there may still be uncontrolled variabilities in the chamber radical source. This would result in a corresponding variability of in the fits of model predictions to experimental results, particularly for chemical systems which are highly sensitive to this effect. Reducing this type of source of variability requires understanding what causes and effects such as the chamber radical source, and then either (a) developing experimental procedures or chamber design modifications to eliminate this effect or (more likely) make it consistent and predictable, or (b) developing a quantitative understanding of all the experimental parameters which influence the effect, and using this to develop a predictive model for it. Either alternative requires a major experimental research program.

Conclusions and Recommendations

I will now briefly summarize what I believe are some of the conclusions which can be drawn from this study regarding the performance of this model and the quality of the available chamber data set, and give several recommendations concerning future model testing efforts. These are not the only conclusions and recommendations which can be derived from the results of this study; other conclusions and recommendations, particularly with regard to the problems with using this and other mechanisms in AQSM's, will be discussed by Fred Lurmann in his presentation at this workshop.

Conclusions Regarding Current Model Performance and Data Quality

This study showed that a model based on our current understanding of the atmospheric chemistry and our best estimates concerning chamber effects is, for most types of experiments, usually able to simulate ozone yields and formation rates to within ~30%. However, there are many experiments which are now well simulated by this model, with most of the discrepancies being apparently random in nature. However, there were cases where apparently systematic discrepancies were observed as well, indicating possible systematic problems with the mechanism, our chamber characterization model, or the experimental data. With the exception of the simulations of the formaldehyde runs and the formaldehyde data, there did not appear to be any large difference in the performance of the model in simulating the UNC or the SAPRC runs. This is gratifying, since prior to this program I had no experience with modeling UNC chamber experiments, and those aspects of the model which were adjusted were adjusted primarily based on fits of the model to results of SAPRC runs.

The many instances of wide, apparently random distributions of qualities of fits of model simulations to current environmental chamber data means that fits of a model to a single "typical" run, or limited groups of runs, is of limited significance. Only if the data set for a compound or mixture is sufficiently large or varied can we obtain a meaningful measure of model performance. This measure is obtained on the basis of statistics of fits of the model to the data set as a whole, and not of the results of the simulations of a single experiment.

The available chamber data set is sufficiently large and varied to give us a meaningful test of the model for a number of compounds and mixtures. With regard to single compounds, this is the case for propene, toluene and, to a lesser extent for several other alkenes and aromatics. It is not the case for aldehydes, ketones, butenes and the higher alkenes, any of the alkanes, and aromatics other than toluene and m-xylene. In some of these cases, the inadequacy of the available data set is not due to the lack of runs, but the fact that the available runs are sensitive to poorly controlled or characterized variables. Two examples of this are the alkane - NO_x - air and the formaldehyde - NO_x - air runs. In the case of the alkane runs, their sensitivity to the ^xvariable chamber radical source results in a wide distribution in the quality of fits of the model to the data, and thus a much larger number of runs is required for adequate assesment of model performance than is the case for compounds which are less sensitive to this. In the formaldehyde case, runs with this compound are highly sensitive to uncertainties in light characteristics, and also the analysis of formaldehyde is subject to considerable uncertainty. In this case, it is difficult to asses the extent to which experimental problems contribute to the qualities of the fits observed, and it is doubtful that more experiments carried out using the same procedures as employed previously would significantly improve the situation.

In the current environmental chamber data set, measurements of ozone, NO, and the easily measured and rapidly reacting hydrocarbons give the most reliable indication of model performance. The utility of other types of experimental measurements in the current data set in assesing model performance is much more uncertain. For example, the current model does not perform nearly as satisfactorily in predicting PAN yields as it does for ozone, with much wider ranges of apparently random discrepancies being observed. This wide distribution of fits suggests, though it obviously does not prove, that there may be problems with the quality of PAN data in the current data set. In addition, more confidence in the quality of aldehyde data obtained during chamber experiments is required before these can be used as a reliable indication of model performance.

To some extent, the utility of the existing environmental chamber data base for testing the mechanism can be improved by further characterization of the conditions under which those experiments were carried out. The recent work in studying the light characteristics in the UNC chamber, and our previous work in characterizing the radical source in the SAPRC EC are two examples of this. To the extent that further studies could improve the characterization of the existing chamber data base, these should be

encouraged, since this is probably the most cost effective means to improve the utility of the data base used to test photochemical models.

However, in order to obtain a significantly more comprehensive test of the current chemical mechanisms than was carried out in this study, use of environmental chamber data with more complete characterization and control of chamber-dependent conditions than is possible with the present data set is required. This requires, I believe, studies aimed at determining the most appropriate types of experimental techniques and chamber surface materials to employ, as well as development of improved analytical techniques. For example, fundamental studies of the types of surface reactions which contribute to chamber effects should be considered as part of this effort. Finally, the process of carrying out environmental chamber experiments for use in model testing needs to be given a much greater commitment of manpower and other resources than previously has been possible. This obviously requires a major research program, and may not be consistent with the current funding levels and priorities of the EPA. However, this is required if any major improvement in our ability to test photochemical mechanisms is desired.

Recommendations for Future Model Testing

This study illustrates how we believe photochemical reaction mechanisms should be tested. In particular, the same mechanism should be tested under a wide variety of conditions with a consistent set of assumptions without run to run adjustment. Model evaluation using "typical" runs, or limited sets of runs, or "evaluations" with run-to-run adjustment of parameters to optimize the fits, should no longer be considered to be acceptable for documenting model performance. Using a limited number of runs to evaluate a model might have been appropriate in the past when there were only limited amounts of data, but now we have a large enough data base of chamber experiments, and sufficiently low cost computational capabilities, that this should no longer be necessary. While there might be some areas of disagreement with what we assumed concerning certain chamber effects, and certainly additional or improved methods of evaluating the model performance could be derived, we believe that the general principles and philosophy employed when carrying out this study serves as an appropriate model for future studies.

It is also important that the chemical mechanism which is tested against the chamber data be a detailed mechanism based on our best estimate as to the correct chemistry. In particular, we do not believe it is appropriate to use environmental chamber data to test lumped or condensed mechanisms. Use of mechanisms which are unnecessarily approximate introduce additional and unnecessary uncertainties regarding the accuracy of the model, and the effects of such approximations could be more unambiguously tested by comparison against model calculations using less approximate models. If we obtain poor fits in simulations of chamber runs when using lumped mechanisms, we do not know whether they are due to errors in the chemistry (or the chamber effects model) or to use of an inappropriate lumping technique. Alternatively, if we obtain good fits, there is always the possibility that errors in the chemistry are being

cancelled by errors in the lumping technique employed. The more appropriate approach is to first test the detailed mechanism against the environmental chamber data, and then use model calculations to explore effects of various lumping or condensation techniques, with the detailed, tested mechanism serving as the standard. This is the approach used in this program.

However, this study should not be taken as the "last word" in the evaluation of this mechanism. The assumptions we made regarding chamber and light characterization should be evaluated by other modelers and chamber experimentalists, and possible areas of improvement and alternative sets of assumptions recommended. Then this (and other) model should be re-tested based on the improved or alternative sets of assumptions. At a minimum, the model should be re-tested against the UNC chamber runs based on the new information we have regarding the light characterization in that chamber. It could also be re-evaluated based on alternative sets of assumptions concerning the chamber radical source, such as those employed by the modelers at SAI. Examining how differing assumptions regarding chamber and light characterization affect the model fits would give us a more precise idea of which of these amount to the most critical uncertainties in terms of model testing.

We recommend that other chemical mechanisms which are currently considered to be at the state of the art, such as the most detailed version of the latest "Carbon Bond" mechanism, be tested against these same data using the same set of assumptions. Although various "Carbon Bond" models have at various times been tested against a large number of experiments, as far as I am aware none of these mechanisms has been tested in a systematic manner against such a large and varied data set without run-to-run adjustment of parameters. If such tests have been carried out, they have not been documented. The results of such tests would allow us more direct comparison of the performance of the different models in simulating the available chamber data than is presently possible.

References

- Atkinson, R., A. C. Lloyd, and L. Wings (1982), *Atmos. Environ.*, 16, 1341.
- Atkinson, R., E. C. Tuazon, H. Mac Leod, S. M. Aschmann, and A. M. Winer (1986), *Geophys. Res. Lett.* 13, 117.
- Carter, W. P. L., R. Atkinson, A. M. Winer and J. N. Pitts, Jr. (1982), *Int J. Chem Kinet.*, 14, 1071.
- Carter, W. P. L., A. M. Winer, R. Atkinson, M. C. Dodd, W. D. Long, and S. M. Aschmann (1984), "Atmospheric Photochemical Modeling of Turbine Engine Fuels. Phase I. Experimental Studies," Final Report to the U. S. Air Force, ESL-TR-84-32, September.

- Carter, W. P. L. and R. Atkinson (1985), J. Atmos. Chem., 3, 377.
- Carter, W. P. L., M. C. Dodd, W. D. Long, and R. Atkinson (1985), "Outdoor Chamber Study to Test Multi-Day Effects", EPA-600/3-84-115, March.
- Carter, W. P. L., F. W. Lurmann, R. Atkinson, and A. C. Lloyd (1986a), "Development and Testing of a Surrogate Species Chemical Reaction Mechanism," EPA-600/3-86-031, August.
- Carter, W. P. L., W. D. Long, L. N. Parker, and M. C. Dodd (1986b), "Effects of Methanol Fuel Substitution on Multi-Day Air Pollution Episodes," Final Report on California Air Resources Board Contract No. A3-125-32, April.
- Jeffries, J. E., R. M. Kamens, K. G. Sexton, and A. A. Gerhardt (1982), "Outdoor Smog Chamber Experiments to Test Photochemical Mechanisms", EPA-600/3-82-016A, April.
- Jeffries, J. E., K. G. Sexton, T. P. Morris, M. Jackson, R. G. Goodman, R. M. Kamens, and M. S. Holleman (1985), "Outdoor Smog Chamber Experiments Using Automobile Exhaust", EPA-600/3-85-032.
- Lurmann, F. W., A. C. Lloyd, and R. Atkinson (1986), J. Geophys. Res. 91, 10915.
- Lurmann, F. W., W. P. L. Carter, and L. A. Coyner (1987) "A Surrogate Species Chemical Reaction Mechanism for Urban-Scale Air Quality Simulation Models. Volume I - Adaptation of the Mechanism, Final Report on Phase II of EPA Contract No. 68-02-4104.
- Peterson, J. T. (1976) "Calculated Actinic Fluxes (290 - 700 nm) for Air Pollution Photochemistry Applications", EPA-600/4-76-025, June.
- Pitts, J. N., Jr., K. Darnall, W. P. L. Carter, A. M. Winer, and R. Atkinson (1979), "Mechanisms of Photochemical Reactions in Urban Air", EPA-600/3-79-110, November.
- Pitts, J. N., Jr., R. Atkinson, W. P. L. Carter, A. M. Winer, and E. C. Tuazon (1983), "Chemical Consequences of Air Quality Standard and Control Implementation Programs", Final Report to the California Air Resources Board Contract No. A1-030-32, April.
- Pitts, J. N., Jr., E. Sanhueza, R. Atkinson, W. P. L. Carter, A. M. Winer, G. W. Harris, and C. N. Plum (1984), Int. J. Chem. Kinet., 16, 919.
- Saeger, M. (1977), "An Experimental Determination of the Specific Photolysis Rate of Nitrogen Dioxide", Masters Thesis, University of North Carolina, Chapel Hill, NC.

Sexton, K. G., H. E. Jeffries, J. R. Arnold, T. L. Kale, and R. M. Kamens
(1987), "Validation Data for Photochemical Mechanisms", Report
on EPA Cooperative Agreement 811673, January

Tuazon, E. C., R. Atkinson, C. N. Plum, A. M. Winer, and J. N. Pitts, Jr.
(1983), Geophys Res. Lett., 10, 953.

Figure 5. Summary of Fits for the Aldehyde and Ketone - NO_x - Air Runs.

-- FORMALDEHYDE --				ACETALDEHYDE		PROPALD.	KETONES	
EC	ITC/OTC	UNC		EC	UNC	UNC	UNC	

(Calc-Expt) (ppm)			Maximum Ozone					
(model low)								
< -0.33	:	:	:	:	:	:*	:*	:
-0.33 - -0.27	:	:	:	:	:	:*	:*	:
-0.27 - -0.21	:	:	:	:	:	:	:	:
-0.21 - -0.15	:	:	:	:	:	:*	:	:
-0.15 - -0.09	:	:0	:**	:	:	:	:	:M
-0.09 - -0.03	:	:0	:*	:	:	:	:	:M
-0.03 - 0.03	:-***	:-*	:-*	:-**	:-	:-	:-	:
0.03 - 0.09	:*	:	:**	:	:	:	:	:
0.09 - 0.15	:	:	:	:	:	:	:	:
0.15 - 0.21	:	:	:	:	:	:	:	:
0.21 - 0.27	:*	:	:	:	:	:	:	:
0.27 - 0.33	:	:	:	:	:	:	:	:A
> 0.33	:	:	:*	:	:	:	:	:
(model high)								
(Calc-Expt)/Expt			Average d([O ₃] - [NO])/dt					
(model low)								
< -1.00	:	:	:	:	:	:	:	:
-1.00 - -0.82	:	:	:	:	:	:	:	:
-0.82 - -0.64	:	:	:	:	:	:	:	:
-0.64 - -0.45	:	:	:	:	:	:	:	:
-0.45 - -0.27	:	:	:	:	:	:	:	:
-0.27 - -0.09	:	:0	:	:*	:***	:**	:	:
-0.09 - 0.09	:-**	:-0	:-*	:-*	:-	:-	:-	:-MM
0.09 - 0.27	:*	:*	:***	:	:	:	:	:
0.27 - 0.45	:*	:	:***	:	:	:	:	:
0.45 - 0.64	:	:	:	:	:	:	:	:A
0.64 - 0.82	:	:	:	:	:	:	:	:
0.82 - 1.00	:	:	:	:	:	:	:	:
> 1.00	:*	:	:	:	:	:	:	:
(model high)								
(O = OTC)				(A = Acetone, M = MEK)				

(O = OTC)

(A = Acetone, M = MEK)

(continued)

Figure 5. Summary of Fits for the Aldehyde and Ketone - NO_x - Air Runs (concluded).

(Calc-Expt) (ppm)	FORMALDEHYDE		ACETALDEHYDE		PROPIONALDEHYDE
	EC	UNC	EC	UNC	UNC

(Calc - Expt)/Expt	Reactant Half Life				
(model low)					
< -0.50	:	:	:	:	:
-0.50 - -0.41	:*	:	:	:	:
-0.41 - -0.32	:*	:***	:	:	:
-0.32 - -0.23	:	:***	:	:	:
-0.23 - -0.14	:	:	:	:	:
-0.14 - -0.05	:	:*	:	:	:
-0.05 - 0.05	-:*	-:	-:	-:	-:
0.05 - 0.14	:	:	:	:	:*
0.14 - 0.23	:*	:	:*	**	:*
0.23 - 0.32	:	:	:	:	:
0.32 - 0.41	:	:	:	:	:
0.41 - 0.50	:*	:	:	:	:
> 0.50	:	:	:*	:	:
(model high)					

Figure 6. Summary of Fits to Propene - NO_x - Air Runs.

		EC	ITC	OTC	UNC
(Calc-Expt) (ppm)		Maximum Ozone			
(model low)					
< -0.33	:	:	:	:	:
-0.33 - -0.27	:	:	:	:	:
-0.27 - -0.21	:	:	*	:	***
-0.21 - -0.15	:	:	*	:	*
-0.15 - -0.09	:	:	:	:	****
-0.09 - -0.03	**	****	:	:	*****
-0.03 - 0.03	----	----	----	----	----
0.03 - 0.09	*****	*	:	:	***
0.09 - 0.15	*	:	:	*	**
0.15 - 0.21	:	:	:	*	:
0.21 - 0.27	*	:	:	:	*
0.27 - 0.33	:	:	:	*	:
> 0.33	:	:	:	:	:
(model high)					
(Calc-Expt)/Expt		Average d([O ₃] - [NO])/dt			
(model low)					
< -1.00	:	:	:	:	:
-1.00 - -0.82	:	:	:	:	:
-0.82 - -0.64	:	:	:	:	:
-0.64 - -0.45	:	:	:	:	:
-0.45 - -0.27	*	:	:	:	:
-0.27 - -0.09	*	**	*	:	****
-0.09 - 0.09	----	----	----	----	-----
0.09 - 0.27	*****	**	*	:	*****
0.27 - 0.45	***	**	*	:	**
0.45 - 0.64	***	:	:	:	:
0.64 - 0.82	:	:	:	:	*
0.82 - 1.00	*	:	:	:	:
> 1.00	:	:	:	:	:
(model high)					

(continued)

Figure 6. Summary of Fits to Propene - NO_x - Air Runs (continued)

	EC	ITC	OTC	UNC

(Calc - Expt)/Expt	Maximum PAN			
(model low)				
< -0.50	:*	:*	:	:
-0.50 - -0.41	:	:*	:	:
-0.41 - -0.32	:	:	:	:
-0.32 - -0.23	:*	:	:	:
-0.23 - -0.14	:	:*	:	:*
-0.14 - -0.05	:*	:	:	:*
-0.05 - 0.05	--:**	-:	-:	-:*****
0.05 - 0.14	:*	:*	:*	:*
0.14 - 0.23	:*	:	:	:*****
0.23 - 0.32	:*	:	:	:*****
0.32 - 0.41	**	:	:*	:*
0.41 - 0.50	**	:	:*	:
> 0.50	:***	:**	:**	:**
(model high)				
(Calc - Expt)/Expt	Propene Half Life			
(model low)				
< -0.50	:	:	:	:
-0.50 - -0.41	:	:	:	:
-0.41 - -0.32	:	:	:	:*
-0.32 - -0.23	:	:	:	:*****
-0.23 - -0.14	:*	:	:	:*****
-0.14 - -0.05	:*****	**	:*	:***
-0.05 - 0.05	--:*****	-:**	-:**	-:*****
0.05 - 0.14	**	:*	:*	**
0.14 - 0.23	:	:*	:	:*
0.23 - 0.32	:*	:	:*	:
0.32 - 0.41	:	:*	:	:
0.41 - 0.50	:	:	:	:
> 0.50	:*	:	:	:
(model high)				

Figure 6. Summary of Fits to Propene - NO_x - Air Runs (concluded)

	EC	ITC	OTC	UNC
(Calc - Expt)/Expt .				
Maximum Acetaldehyde				
(model low)				
< -0.50	:	:	:	:
-0.50 - -0.41	:	:	:	:
-0.41 - -0.32	:	:	:	:
-0.32 - -0.23	:	:	:	:
-0.23 - -0.14	:	:	:	:*
-0.14 - -0.05	:*	:	:*	**
-0.05 - 0.05	--:**	--:	--:*	--:*****
0.05 - 0.14	:*****	:	:	:****
0.14 - 0.23	:*	:	:*	:****
0.23 - 0.32	:***	:	:*	:*
0.32 - 0.41	**	:	:*	:
0.41 - 0.50	:	:	:	:
> 0.50	:	:*****	:	:
(model high)				
(Calc - Expt)/Expt				
Maximum Formaldehyde				
(model low)				
< -0.50	:	:	:	:***
-0.50 - -0.41	:	:	:	**
-0.41 - -0.32	:	:*	:*	:****
-0.32 - -0.23	:	:*	:	**
-0.23 - -0.14	**	:	:	**
-0.14 - -0.05	:*	**	:	**
-0.05 - 0.05	--:**	--:	--:	--:UUU
0.05 - 0.14	:***	:	:	:
0.14 - 0.23	:*	:*	:*	: U = Unisearch
0.23 - 0.32	:	:	:*	: HCHO Data
0.32 - 0.41	**	:	:	:
0.41 - 0.50	:*	:*	:	:
> 0.50	**	:*	:	:*
(model high)				

Figure 7. Summary of Fits for Ethene and Butene - NO_x - Air Runs.

				EC	ITC	UNC	
				EC	ITC	UNC	
Calc-Expt (ppm)	Maximum Ozone			(Calc-Expt) /Expt	d([O ₃]-[NO])/dt		
< -0.33	:	:	:	< -1.00	:	:	:
-0.33 - -0.27	:*	:	:	-1.00 - -0.82	:	:	:
-0.27 - -0.21	:**	:	:	-0.82 - -0.64	:	:	:
-0.21 - -0.15	:x	:	:**	-0.64 - -0.45	:	:	:
-0.15 - -0.09	:xx	:	:	-0.45 - -0.27	:**xxxx	:	:*
-0.09 - -0.03	:xxx	:x	:	-0.27 - -0.09	:*x	:	:
-0.03 - 0.03	-:	-:**	-:	-0.09 - 0.09	-:x	-:x	-:***
0.03 - 0.09	:	:xxxx	:*x	0.09 - 0.27	:***	:**	:**
0.09 - 0.15	:*	:	:*	0.27 - 0.45	:	:x	:
0.15 - 0.21	:*	:	:*	0.45 - 0.64	:	:xxx	:
0.21 - 0.27	:*	:	:*	0.64 - 0.82	:	:	:x
0.27 - 0.33	:	:	:	0.82 - 1.00	:	:	:
> 0.33	:	:	:	> 1.00	:	:	:

(Calc-Expt) /Expt	Reactant Half Life			(Calc-Expt) /Expt	Maximum Formaldehyde		
< -0.50	:	:	:	< -0.50	:	:	:**x
-0.50 - -0.41	:	:	:	-0.50 - -0.41	:	:	:*
-0.41 - -0.32	:	:	:	-0.41 - -0.32	:	:	:*
-0.32 - -0.23	:	:	:	-0.32 - -0.23	:	:	:**
-0.23 - -0.14	:	:xxx	:***x	-0.23 - -0.14	:*x	:	:
-0.14 - -0.05	:***	:*	:	-0.14 - -0.05	:**x	:*	:
-0.05 - 0.05	-:	-:*	-:	-0.05 - 0.05	-:***x	-:***x	-:
0.05 - 0.14	:	:	:*	0.05 - 0.14	:	:	:
0.14 - 0.23	:	:	:	0.14 - 0.23	:	:x	:
0.23 - 0.32	:*	:	:*	0.23 - 0.32	:	:	:
0.32 - 0.41	:x	:	:	0.32 - 0.41	:	:	:
0.41 - 0.50	:**	:	:	0.41 - 0.50	:	:	:
> 0.50	:xxxxx	:	:	> 0.50	:*xxx	:x	:

(* = Ethene, x = Butenes)

Figure 8. Summary of Fits to the n-Butane - NO_x - Air Runs

Calc - Expt (ppm)	Maximum Ozone			Calc-Expt /Expt	Avg. d([O ₃]-[NO])/dt		
	EC	ITC/OTC	UNC		EC	ITC/OTC	UNC

(model low)							
< -0.33	:	:	:*	< -1.00	:	:	:
-0.33 - -0.27	:	:	:*	-1.00 - -0.82	:	:	:
-0.27 - -0.21	:*	:	:	-0.82 - -0.64	:	:	:
-0.21 - -0.15	**	:	:*	-0.64 - -0.45	:*	:	:
-0.15 - -0.09	:	:	:*	-0.45 - -0.27	**	:	:*
-0.09 - -0.03	:*	:	:	-0.27 - -0.09	**	:	***
-0.03 - 0.03	-:**	-:*	-:	-0.09 - 0.09	-:	-:	-:
0.03 - 0.09	**	:*0	:*	0.09 - 0.27	:*	:	:*
0.09 - 0.15	:	:*	:	0.27 - 0.45	:*	:	**
0.15 - 0.21	***	:	:*	0.45 - 0.64	*****	:	:
0.21 - 0.27	***	:*	:*	0.64 - 0.82	:*	:	:
0.27 - 0.33	:	:*	:	0.82 - 1.00	:	:*0	:
> 0.33	:	:	:	> 1.00	**	*****	:
(model high)		(0 = OTC)				(0 = OTC)	

Figure 9. Summary of Fits to the C₅+ Alkane - NO_x - Air Runs

		23-DMB	ISO-C ₅ ISO-C ₈	n-C ₅	n-C ₆	n-C ₇	n-C ₈	n-C ₉
(Calc-Expt) (ppm)		Maximum Ozone						
(model low)								
<	-0.33	:	:	:	:	:	:	:
-0.33 -	-0.27	:	:	:	:	:	:	:
-0.27 -	-0.21	:	:	:	:	:	:	:
-0.21 -	-0.15	:	:	:	:	:	:	:
-0.15 -	-0.09	:	:	:	:	:	:	:
-0.09 -	-0.03	:	:	:	:	:	:	:
-0.03 -	0.03	-:	-:UU	-:	-:	-:	-:III	-:
0.03 -	0.09	:EE	:	:	:	:I	:I	:
0.09 -	0.15	:	:	:EU	:	:	:I	:
0.15 -	0.21	:EU	:	:	:EI	:	:	:E
0.21 -	0.27	:	:	:	:	:I	:	:
0.27 -	0.33	:U	:	:	:	:	:	:
>	0.33	:	:	:	:	:	:U	:
(model high)								
(Calc-Expt)/Expt		Average d([O ₃] - [NO])/dt						
(model low)								
<	-1.00	:	:	:	:	:	:	:
-1.00 -	-0.82	:	:	:	:	:	:	:
-0.82 -	-0.64	:	:	:	:	:	:	:
-0.64 -	-0.45	:	:	:	:	:	:	:
-0.45 -	-0.27	:	:	:	:	:I	:	:
-0.27 -	-0.09	:	:	:	:	:	:I	:E
-0.09 -	0.09	-:E	-:UU	-:E	-:I	-:	-:	-:
0.09 -	0.27	:E	:	:I	:E	:	:	:
0.27 -	0.45	:U	:	:	:	:	:I	:
0.45 -	0.64	:U	:	:	:	:	:IU	:
0.64 -	0.82	:	:	:	:	:I	:	:
0.82 -	1.00	:	:	:	:	:	:	:
>	1.00	:	:	:	:	:	:II	:
(model high)		(E = EC, I = ITC, U = UNC)						

Figure 10. Summary of Fits to the Aromatic - NO_x - Air Runs.

Benzene		----- Toluene -----		- Xylenes -			135-TMB	
ITC	EC	ITC	UNC	EC	ITC	UNC	EC	ITC

(Calc-Expt) (ppm)		Maximum Ozone						
(model low)								
<	-0.33	:	:	:	:	:	:	:
-0.33 -	-0.27	:	:	:	:	:	:	:*
-0.27 -	-0.21	:	:	:	:	:	:*	:*
-0.21 -	-0.15	:	:	:*	:	:	:*	**
-0.15 -	-0.09	:	:*	:	:	:*	:	:
-0.09 -	-0.03	:	**	:	:	:*	:	:
-0.03 -	0.03	:-:*****	:-:*****	:-:**	:-:**	:-:*	:-:*	:-:*
0.03 -	0.09	:	***	:	:*	:	:	:**
0.09 -	0.15	:	:	:	:*	:	:	:
0.15 -	0.21	:	:	:	:	:	:	:
0.21 -	0.27	:	:	:	:*	:	:	:
0.27 -	0.33	:	:	:	:	:	:	:
>	0.33	:	:	:	:	:	:	:
(model high)								
(Calc-Expt)/Expt		Average d([O ₃] - [NO])/dt						
(model low)								
<	-1.00	:	:	:	:	:	:	:
-1.00 -	-0.82	:	:	:	:	:	:	:
-0.82 -	-0.64	:	:	:	:	:	:	:
-0.64 -	-0.45	:	:	:	:	:	:	:
-0.45 -	-0.27	:	**	:	:	:	:*	:
-0.27 -	-0.09	:	:*	:	:*	:	:	:
-0.09 -	0.09	:-:**	:-:*****	:-:**	:-:*	:-:**	:-:**	:-:
0.09 -	0.27	:	***	:*	:*	:	:*	:*****
0.27 -	0.45	:***	:*	:	:*	:**	:	:*
0.45 -	0.64	:	:	:	:	:	:	:
0.64 -	0.82	:*	:	:	:	:*	:	:
0.82 -	1.00	:	:	:	:	:	:	:
>	1.00	:	:	:*	:	:	:	:
(model high)								
(continued)								

Figure 10. Summary of Fits to the Aromatic - NO_x - Air Runs (concluded)

Benzene		----- Toluene -----		- Xylenes -			135-TMB			
ITC	EC	ITC	UNC	EC	ITC	UNC	EC	ITC	EC	ITC

(Calc - Expt)/Expt				Maximum PAN						
(model low)										
< -0.50	:*	:	:	:	:*	:	:	:*	:	:***
-0.50 - -0.41	:	:	:*	:	:	:	:	:	:	:*
-0.41 - -0.32	:*	:	:	:	:	**	:	:*	:	:
-0.32 - -0.23	:*	:*	:	:	:	:*	:	:*	:	:
-0.23 - -0.14	:	:	:*	:	:	:	:	:	:	:
-0.14 - -0.05	:	:	:	:	:	:	:	:	:	:
-0.05 - 0.05	:-**	-:	-:	-:***	-:	-:*	-:	-:	-:	-:*
0.05 - 0.14	:	:	**	:	:	:	:	:	:	:
0.14 - 0.23	:****	:	:	:	:	:	:	:	:	:
0.23 - 0.32	:***	:	:	:	:*	:	:	:	:	:
0.32 - 0.41	:*	:	:	:*	:	:	:	:	:	:
0.41 - 0.50	:	:	:	:	:	:	:	:	:	:
> 0.50	:	:	:*	:	:	:	:	:	:	:
(model high)										

Figure 11. Summary of Fits to the Two - Five Comonent Mixture Runs.

		Mixed Alkenes		----- Other Mixtures -----		
		EC	UNC	EC	ITC	UNC
(Calc-Expt) (ppm)		Maximum Ozone				
(model low)						
< -0.33	: **	: *	:	:	:	:
-0.33 - -0.27	: *	:	:	:	:	:
-0.27 - -0.21	: **	:	:	:	:	: **
-0.21 - -0.15	: *	: *	:	:	:	: ****
-0.15 - -0.09	: **	:	:	:	:	:
-0.09 - -0.03	:	:	:	:	:	: *****
-0.03 - 0.03	-: *	-: **	-: **	-: *****000	-: ****	-: ***
0.03 - 0.09	:	:	: *****	: *****00	: **	:
0.09 - 0.15	:	:	: ****	: ****	: **	:
0.15 - 0.21	:	:	:	:	: *	:
0.21 - 0.27	:	:	:	:	:	:
0.27 - 0.33	:	:	:	:	:	:
> 0.33	:	:	:	:	:	:
(model high)						
(Calc-Expt)/Expt		Average d([O ₃] - [NO])/dt				
(model low)						
< -1.00	:	:	:	:	:	:
-1.00 - -0.82	:	:	:	:	:	:
-0.82 - -0.64	:	:	:	:	:	:
-0.64 - -0.45	: *	: *	:	:	:	:
-0.45 - -0.27	: ****	: *	:	:	:	:
-0.27 - -0.09	: **	:	: ****	: 00000	: *****	:
-0.09 - 0.09	-: *	-: **	-: *****	-: *	: *****	:
0.09 - 0.27	: *	: **	: ****	: *	: ****	:
0.27 - 0.45	:	:	:	: *****	: **	:
0.45 - 0.64	:	:	:	: *****	:	:
0.64 - 0.82	:	:	:	: *****	:	:
0.82 - 1.00	:	:	:	:	:	:
> 1.00	:	:	:	:	:	:
(model high)					(O = 1986 Runs)	
(continued)						

Figure 11. Summary of Fits to the Two - Five Comonent Mixture Runs (concluded).

		Mixed Alkenes		----- Other Mixtures -----		
		EC	UNC	EC	ITC	UNC

(Calc - Expt)/Expt		Maximum PAN				
(model low)						
<	-0.50	:*	:**	:	:	:
-0.50 -	-0.41	:**	:	:	:*	:**
-0.41 -	-0.32	:	:	:	:*	:*
-0.32 -	-0.23	:**	:	:	:*	:*
-0.23 -	-0.14	:*	:	:	:****	:
-0.14 -	-0.05	:***	:	:*	:*****	:
-0.05 -	0.05	-:	-:	-:*	-:*****	-:***
0.05 -	0.14	:	:	:***	:	:
0.14 -	0.23	:	:	:*	:00	:**
0.23 -	0.32	:	:	:***	:0	:*
0.32 -	0.41	:	:	:**	:	:*
0.41 -	0.50	:	:	:*	:*	:
>	0.50	:	:**	:	:**00	:***
(model high)		(0 = 1986 Runs)				

Figure 12 Summary of Fits to the NO_x - Air Runs Employing Complex Mixtures.

	EC	"SAI" 8 HC Surrogate		UNC
	7 HC Surg	ITC	OTC	Misc. Mixes

(Calc-Expt) (ppm)	Maximum Ozone			
(model low)				
< -0.33 :	:	:	:*	:*
-0.33 - -0.27 :	:	:	:*	:
-0.27 - -0.21 :	:	:	:	:*
-0.21 - -0.15 :	:	:	:XX	**
-0.15 - -0.09 :	:	:	:XXXX	*****
-0.09 - -0.03 :	:	:XX	:XXXXX*	:*
-0.03 - 0.03 - :***	-:XXXXXXXX	-:XXXXXXXXXXXX	-:XXXXXXXXXXXX	-:*****
0.03 - 0.09 :****	:*	:XXXX	:XXXX	:****
0.09 - 0.15 **:	:	:XX*	:XX*	:**
0.15 - 0.21 **:	:	:X	:	:
0.21 - 0.27 :	:	:*	:*	:*
0.27 - 0.33 :	:	:*	:*	:*
> 0.33 :	:	:	:	:
(model high)				
(Calc-Expt)/Expt	Average d([O ₃] - [NO])/dt			
(model low)				
< -1.00 :	:	:	:	:
-1.00 - -0.82 :	:	:	:	:
-0.82 - -0.64 :	:	:	:	:
-0.64 - -0.45 :	:	:	:*	:
-0.45 - -0.27 :	:	:*	:XXX*	:*
-0.27 - -0.09 :	:	:*	:XXXXXXXXXXXX	:*****
-0.09 - 0.09 - :*****	-:***	-:XXXXXXX	-:XXXXXXXX	-:*****
0.09 - 0.27 :****	:*****	:XXXX	:XXXX	:****
0.27 - 0.45 :*	:*****	:X*	:X*	:*
0.45 - 0.64 :	:	:*	:*	:*
0.64 - 0.82 :	:*	:	:	:
0.82 - 1.00 :	**	:	:	:
> 1.00 :	:	:*	:*	**
(model high)		(X = 2 runs)		
				(continued)

Figure 12. Summary of Fits to the NO_x - Air Runs Employing Complex Mixtures (continued).

	EC	"SAI" 8 HC Surrogate	UNC
	7 HC Surg	ITC	Misc. Mixes

(Calc - Expt)/Expt		Maximum PAN	
(model low)			
< -0.50 :	:	:*****	:**
-0.50 - -0.41 :	:	:	:**
-0.41 - -0.32 :	:**	:*****	:
-0.32 - -0.23 :	:*	:****	:*
-0.23 - -0.14 :	:	:*****	:
-0.14 - -0.05 :	:*	:****	:****
-0.05 - 0.05 -:***	-:	-:*****	-:
0.05 - 0.14 :***	:	:*****	:
0.14 - 0.23 :***	:*	:****	:*
0.23 - 0.32 :**	:	:***	:
0.32 - 0.41 :	:	:	:
0.41 - 0.50 :	:	:*****	:
> 0.50 :	:*	:*****	:*****
(model high)			
(Calc - Expt)/Expt		Maximum Formaldehyde	
(model low)			
< -0.50 :	:*	:***	
-0.50 - -0.41 :	:*	:*****	
-0.41 - -0.32 :	:*	:*****	
-0.32 - -0.23 :*	:	:*****	
-0.23 - -0.14 :*	:***	:*****	
-0.14 - -0.05 :**	:***	:*****	
-0.05 - 0.05 -:***	-:***	-:***	
0.05 - 0.14 :**	:	:*****	
0.14 - 0.23 :*	:	:****	
0.23 - 0.32 :	:*	:*	
0.32 - 0.41 :	:	:	
0.41 - 0.50 :	:*	:	
> 0.50 :*	:*****	:****	
(model high)			
			(continued)

Figure 12. Summary of Fits to the NO_x - Air Runs Employing Complex Mixtures (concluded).

	EC	"SAI" 8 HC Surrogate	
	7 HC Surg	ITC	OTC

(Calc - Expt)/Expt	Propene Half Life		
(model low)			
< -0.50	:	:	:*
-0.50 - -0.41	:	:	:
-0.41 - -0.32	:	:*	:
-0.32 - -0.23	:	**	:
-0.23 - -0.14	:	*****	:*
-0.14 - -0.05	:*	*****	*****
-0.05 - 0.05	-:****	-:***	-:*****
0.05 - 0.14	:***	:	:*****
0.14 - 0.23	:**	:***	:*****
0.23 - 0.32	:*	:	:*****
0.32 - 0.41	:	:	:*****
0.41 - 0.50	:	:	:****
> 0.50	:	:	:**
(model high)			

Figure 13. Summary of Fits to the UNC Auto Exhaust Runs

Calc - Expt (ppm)	Maximum Ozone		Calc-Expt /Expt	Avg. d([O ₃]-[NO])/dt		
	VOLARE	CHARGER		VOLARE	CHARGER	

(model low)			(model low)			
< -0.33	:	:	< -1.00	:	:	
-0.33 - -0.27	:	:	-1.00 - -0.82	:	:	
-0.27 - -0.21	:	**	-0.82 - -0.64	:	:	
-0.21 - -0.15	*	**	-0.64 - -0.45	:	:	
-0.15 - -0.09	***	*	-0.45 - -0.27	:	:	
-0.09 - -0.03	****	*	-0.27 - -0.09	***	*****	
-0.03 - 0.03	*****	*****	-0.09 - 0.09	*****	*****	
0.03 - 0.09	:	**	0.09 - 0.27	**	:	
0.09 - 0.15	:	*	0.27 - 0.45	*	**	
0.15 - 0.21	:	:	0.45 - 0.64	:	:	
0.21 - 0.27	:	:	0.64 - 0.82	:	:	
0.27 - 0.33	:	:	0.82 - 1.00	:	:	
> 0.33	:	:	> 1.00	:	:	
(model high)			(model high)			

Calc - Expt /Expt	Ethene Half Life		Maximum PAN		Maximum Formaldehyde	
	VOLARE	CHARGER	VOLARE	CHARGER	VOLARE	CHARGER

(model low)						
< -0.50	:	:	*****	*	XXXXXX	*****
-0.50 - -0.41	:	:	*	*	:	**
-0.41 - -0.32	:	:	:	*	:	*
-0.32 - -0.23	:	:	:	:	:	*
-0.23 - -0.14	:	**	***	***	:	*
-0.14 - -0.05	***	***	*	*	:	:
-0.05 - 0.05	**	*****	-	**	-	:
0.05 - 0.14	*	**	:	**	:	*
0.14 - 0.23	:	*	:	:	:	:
0.23 - 0.32	:	:	:	*	:	:
0.32 - 0.41	:	:	**	:	:	:
0.41 - 0.50	:	:	*	:	:	:
> 0.50	:	:	:	*	:	:
(model high)						
(X = 2 runs)						

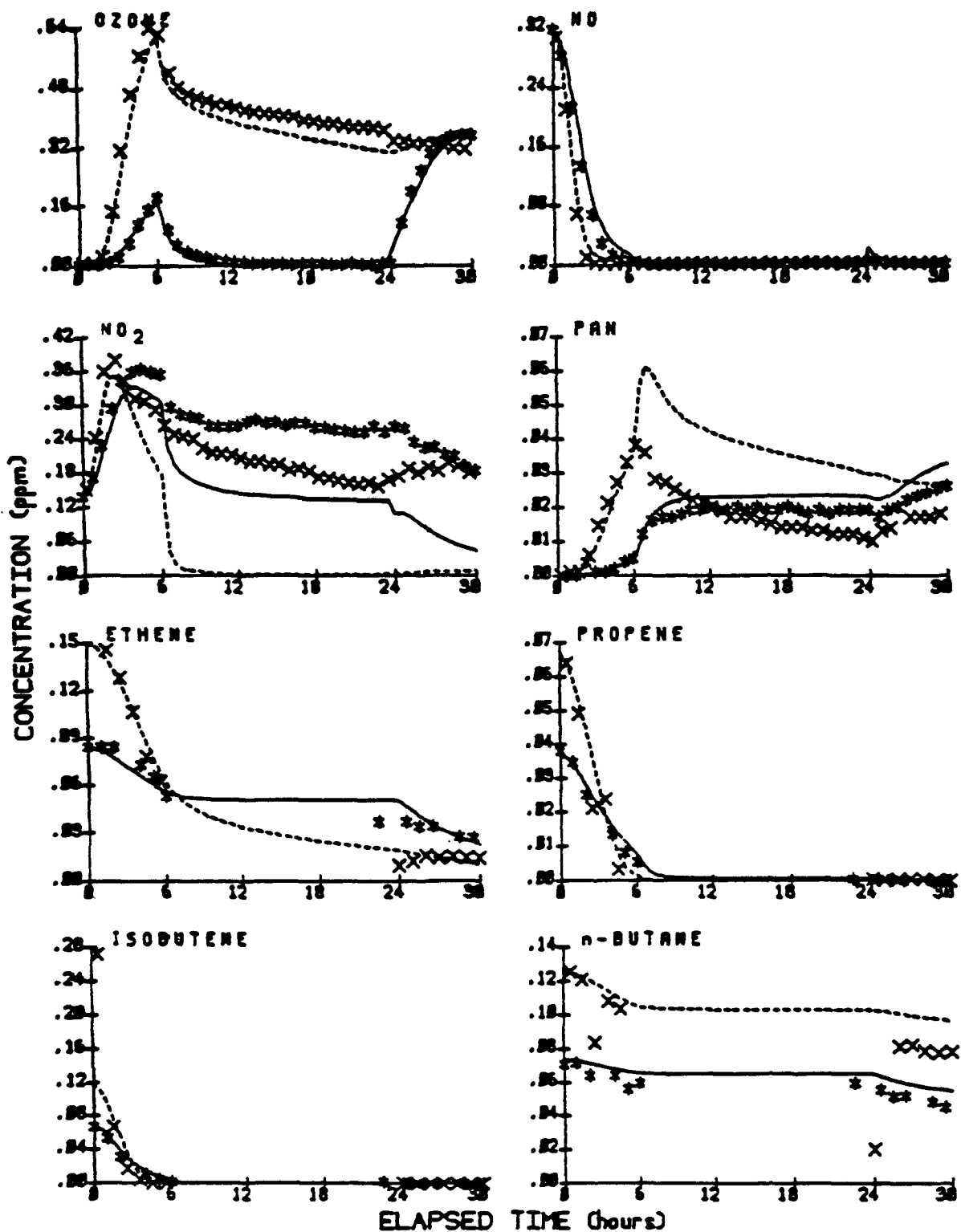


Figure 14. Experimental and Calculated Concentration - Time Plots for Selected Species Observed in the Surrogate - NO_x Air Run OTC-242

- * = Experimental data, side A. (33% Methanol substitution surrogate)
- = Model calculation, side A.
- x = Experimental data, side B. (Base case surrogate)
- - - = Model calculation, side B.

Paper 8

Evaluation and Intercomparison of Chemical Mechanisms

Fred Lurmann
Environmental Research and Technology, Inc.

EVALUATION AND INTERCOMPARISON OF CHEMICAL MECHANISMS

Fred Lurmann
ERT, Inc.
975 Business Center Circle
Newbury Park, CA 91320

Presented At

The EPA/ASRL Workshop on
Evaluation/Documentation of Chemical Mechanisms

Introduction

Various aspects of atmospheric chemical mechanism evaluation and intercomparison are discussed in this paper. The topics covered include the methods used to test chemical mechanisms against environmental chamber data, results from recent mechanism evaluations, and uncertainties in the modeling and evaluation process. Also presented are some preliminary mechanism intercomparison results for the Carbon Bond IV and Carter et al. (1986) mechanisms.

Mechanism Evaluation

The general methods for developing and evaluating atmospheric chemical mechanisms were formulated by the pioneering workers in this field (see e.g. Niki, Daby, and Weinstock 1972; Demerjian, Kerr, and Calvert 1974; Hecht, Seinfeld, and Dodge 1974). While the kinetic and mechanistic data base, and the number of environmental chamber experiments have greatly expanded in recent years, and testing procedures have been refined, the basic methods recommended by the early workers are still followed today. These procedures consist of developing detailed chemical mechanisms that consist of elementary reactions and testing the mechanisms against the best available smog chamber experiments using a consistent protocol (i.e., without arbitrary run-to-run adjustment of uncertain parameters). More recent studies (Whitten and Hogo 1977; Carter et al. 1979; Atkinson et al. 1982; Killus and Whitten 1983; Carter et al. 1986) have shown the importance of testing the mechanism according to the hierarchy of species, testing against data from chambers with different radiation intensity and spectra, testing against a large number of experiments, and testing with a consistent model of the chamber effects (such as heterogeneous wall reactions leading to the offgassing of radical species and NO_x).

Examples of mechanism performance data obtained with a recently developed detailed mechanism are shown in Figures 1-6 (from Carter et al. 1986). The figures show the predicted versus observed maximum ozone concentrations in six different types of chamber experiments carried out in two indoor chambers (SAPRC EC and ITC) and two outdoor chambers (UNC and SAPRC). The experiments involved irradiation of NO_x with individual alkenes, individual alkanes, individual aromatics, mixtures of organics from two groups, complex mixtures of organics from three groups, and automobile exhaust (UNC only). A total of 490 chamber runs were employed in the evaluation. The results show that the mechanism is able to predict the maximum ozone within $\pm 30\%$ in the vast majority of cases. The noticeable exception to this result are the predictions for NO_x - alkane experiments which are very sensitive to the assumed chamber radical source strength. Table 1 shows the average bias and error in the maximum ozone for the different groups of runs employed in this evaluation. These results show that the average bias in the ozone predictions is small, usually less than 10%, and usually positive. The average error is typically 20% to 30%. The performance results for the organic mixtures are particularly important because the mechanisms are generally applied to complex mixtures in atmospheric modeling. Overall, the average bias and error for the mixtures are +4% and $\pm 24\%$. Statistics for timing of the ozone peak and rate of NO_x oxidation are quite similar.

In developing the Carter et al. mechanism, many variations in mechanistic parameters were investigated by testing against chamber data. One of the more interesting and important results was that while alternate mechanistic assumptions affected the average bias for a group of runs (i.e., making the mechanism more or less reactive), the alternate assumptions had only a minor affect on the average error. This suggests that numerous mechanisms can fit the data reasonably well, so it may be difficult to distinguish between mechanisms solely on the basis of their performance relative to chamber data. It also suggests that there are basic uncertainties in the mechanism, chamber data, and chamber characterization which presently limit the performance of the mechanism in this type of testing program. No one has carefully tried to estimate the uncertainties of the data employed in this modeling system, however, I have made some rough estimates. NO_2 photolysis rate are known to within $\pm 10\%$ in indoor chambers and no better than ± 20 in outdoor chambers. Many rate constants have uncertainties of ± 20 or more (see Atkinson and Lloyd 1984; Atkinson 1986). Product coefficients for the organic reactions often have uncertainties of ± 20 or more. Uncertainties in the chamber concentration data, used for initial concentrations and for comparison purposes, are also significant for organics ($\pm 10\%$), PAN ($\pm 20\%$), and oxygenates ($\pm 30\%$). Lastly, the chamber radical source strength could easily be uncertain to $\pm 50\%$ or more. If

these uncertainties are realistic, which I believe they are, then it is not surprising that the mechanism performance evaluation results show $\pm 20\%$ to $\pm 30\%$ error on the average. This level of error in the output is totally consistent with the level of errors in the input data.

If regulatory agencies need mechanisms that are substantially more accurate and precise than the present generation of mechanisms, then higher quality chamber experiments and improved mechanistic and kinetic data must be obtained. The technology is generally available to conduct higher quality chamber experiments. For example, the uncertainty in the NO_2 photolysis rates can probably be reduced substantially by continuous in situ kl measurements in the experiments. DNPH/HPLC methods are available for obtaining high quality data for oxygenated species. The dual organic tracer technique could be employed in many experiments to derive OH radical concentrations and improved radical source strengths. FTIR could be employed to measure nitrous acid early in the runs to get better estimates of its role initiation. Thus, there is clearly substantial room for improving the quality of the data used in the evaluation process.

Mechanism Intercomparison

A sufficient number of chemical mechanisms had been developed by 1980 to allow researchers to begin intercomparing the mechanisms. Carter et al. 1982, and Jeffries et al. 1981 showed that while the mechanisms predict different, but still reasonable results for chamber experiments, they predicted substantially different VOC control requirements when used with the Empirical Kinetic Modeling Approach (EKMA). This topic was explored further by McRae et al. 1983; Leone and Seinfeld 1984, and Schafer and Seinfeld 1985 using slightly more up-to-date mechanisms. Results are presented here which employ even more recent versions of the mechanisms.

Figures 7 show a comparison of the ozone predictions from the Carbon Bond IV mechanism and from the Carter et al. 1986 mechanism for simulations with 1 ppmC of urban VOC and initial NO_x levels 0.05, 0.1, 0.167 ppm. The figure shows that the Carbon Bond IV mechanism consistently predicts higher ozone than the Carter et al. mechanism, and that the discrepancy tends to increase as the initial VOC/ NO_x ratio decreases. Figure 8 shows the NO , NO_2 , and O_3 profiles for a similar simulation carried out with an initial VOC/ NO_x ratio of 3. This shows that both mechanisms appear to oxidize the NO_x at the same rates, however, the ozone prediction is 20% higher on the second day with the Carbon Bond IV mechanism. Figures 9 and 10 compare the OH, HO_2 , HNO_3 , H_2O_2 , toluene, and ethene profiles predicted by the two mechanisms for the initial VOC/ $\text{NO}_x = 3$ case. The agreement on these species is generally better than that found for ozone.

Another set of simulations was carried out with the same two mechanisms in the OZIPM model (Version 3 - Whitten and Hogo 1986). An EKMA analysis was performed for nine cases, representing urban conditions with emissions, an ozone design value of 0.24 ppm, an initial mixing height of 250 m and three maximum mixing heights (500, 1000, and 1500 m), and three initial VOC/NO_x ratios (6, 10, and 20). The OZIPM predicted VOC control requirements to achieve compliance with the 0.12 ppm standard are shown in Table 2. For the 1000 m maximum mixing height cases, the results of the Carter et al. and CBM-IV mechanisms are 29% and 41% at VOC/NO_x = 6, 58% and 68% at VOC/NO_x = 10, and 80% and 85% at VOC/NO_x = 20, respectively. These differences are surprisingly large when you consider that the mechanisms are based on the same basic kinetic and mechanistic data (i.e., Atkinson and Lloyd 1984) and have been testing against many of the same smog chamber experiments. The differences are almost as large as the differences identified five years ago by Jeffries et al. 1982. Thus, while the accuracy and precision of the mechanisms may be improved in the new generation, the problem of different mechanisms giving significantly different VOC control requirement predictions is still very much with us.

TABLE 1

AVERAGE MODEL PERFORMANCE ON MAXIMUM OZONE
IN THE CARTER ET AL. 1986 EVALUATION

RUN TYPE -----	BIAS (%) * -----	ERROR (%) * -----
Formaldehyde	-1	19
Acetaldehyde	-26	26
Other Carbonyls	+4	44
All Carbonyls	-5	25
Ethene	+2	18
Propene	+3	18
Butenes	+4	34
All Alkenes	+3	21
Butane	+31	67
Branched Alkanes	+34	49
Long-chain Alkanes	+83	84
All Alkanes	+46	69
Benzene	+3	5
Toluene	+11	24
Xylenes	-9	16
Mesitylene	-11	21
All Aromatics	+1	19
All Single HC Runs	+12	33
Simple Mixtures	+10	35
Mini Surrogates	+10	22
Full Surrogates	+3	23
Auto Exhaust	-11	15
All HC Mixtures	+4	24
All Run Average	+7	28

* Positive bias indicates model overprediction.

TABLE 2
VOC CONTROL REQUIREMENT PREDICTIONS (%)

VOC/NOx -----	Mixing Height -----	Chemical Mechanism -----	
		CARTER -----	CBM-IV -----
6	500m	26.8%	36.3%
6	1000m	29.1%	41.0%
6	1500m	30.1%	42.7%
10	500m	53.7%	64.0%
10	1000m	57.7%	68.0%
10	1500m	59.2%	69.1%
20	500m	77.4%	82.9%
20	1000m	80.1%	85.1%
20	1500m	81.0%	85.9%

Alkenes

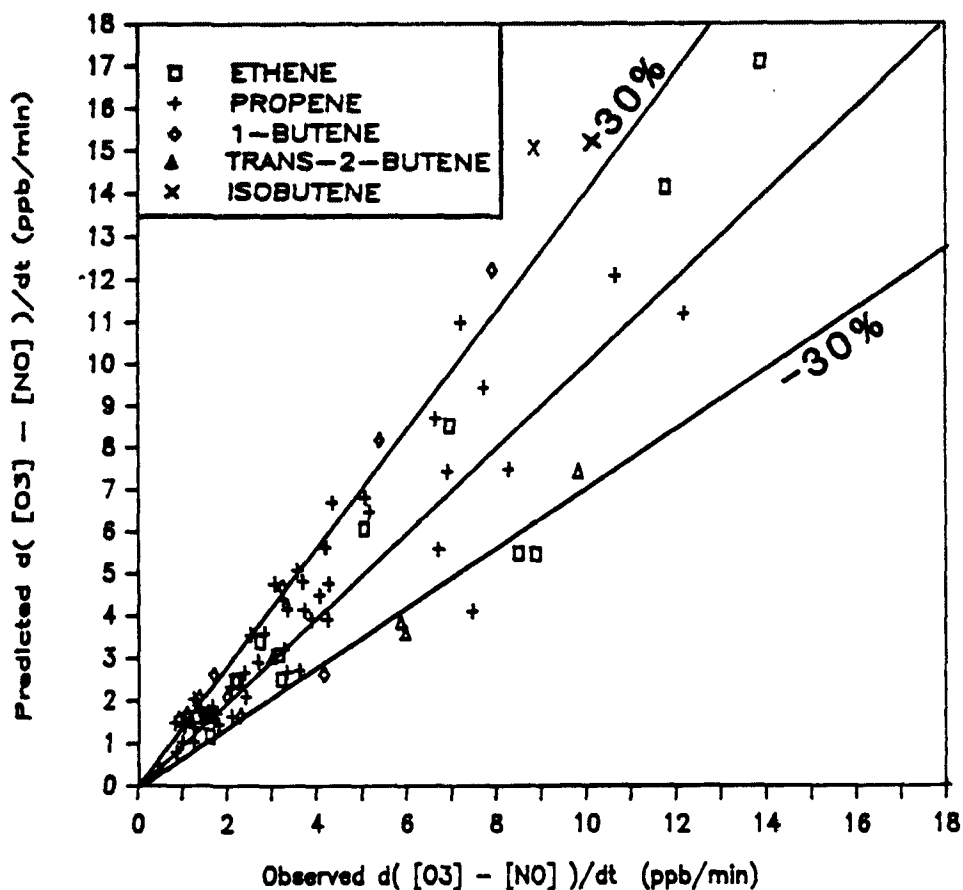
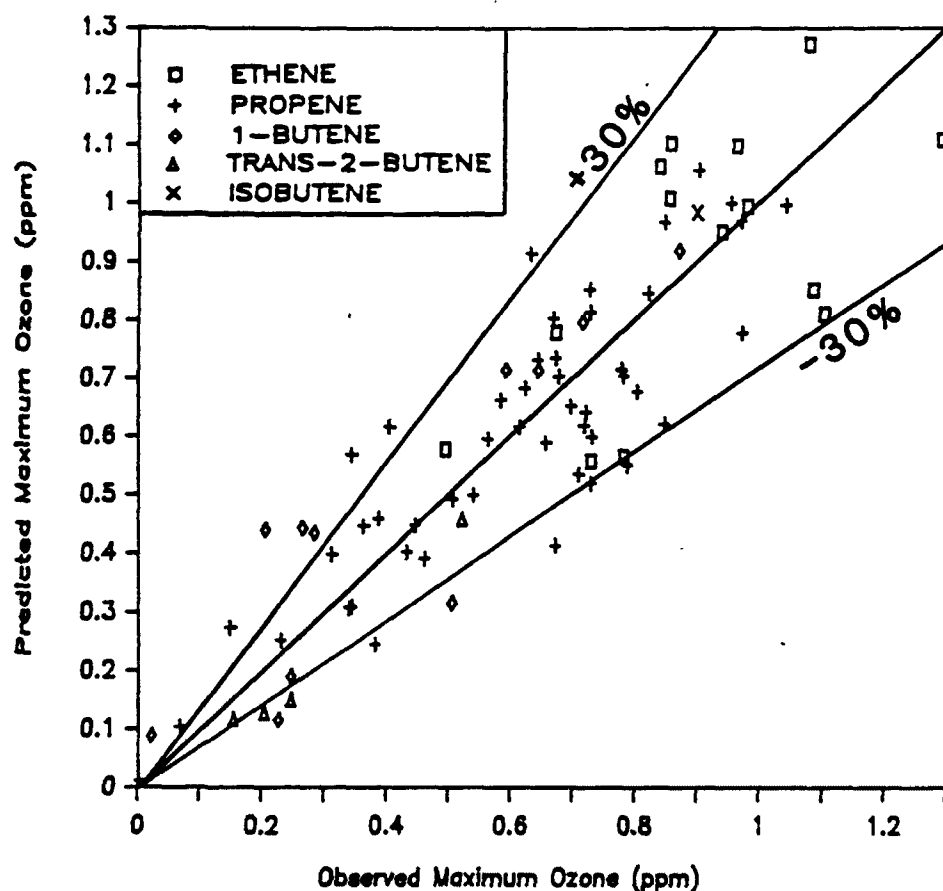


Figure 1. Predicted versus observed maximum ozone concentrations and timing parameter for alkene - NO_x chamber experiments.

Alkanes

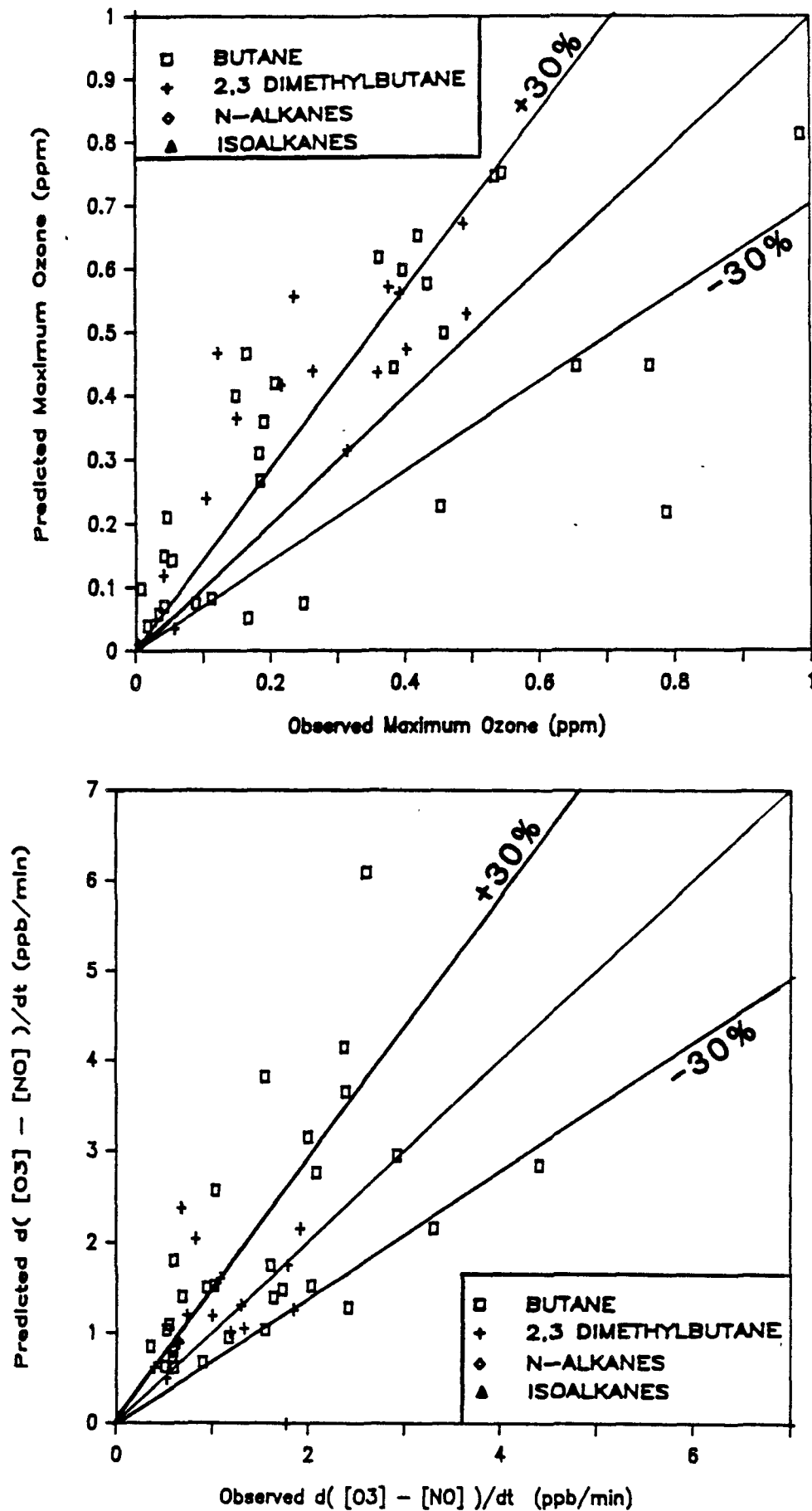


Figure 2. Predicted versus observed maximum ozone concentrations and timing parameter for alkane - NO_x chamber experiments.

Aromatics

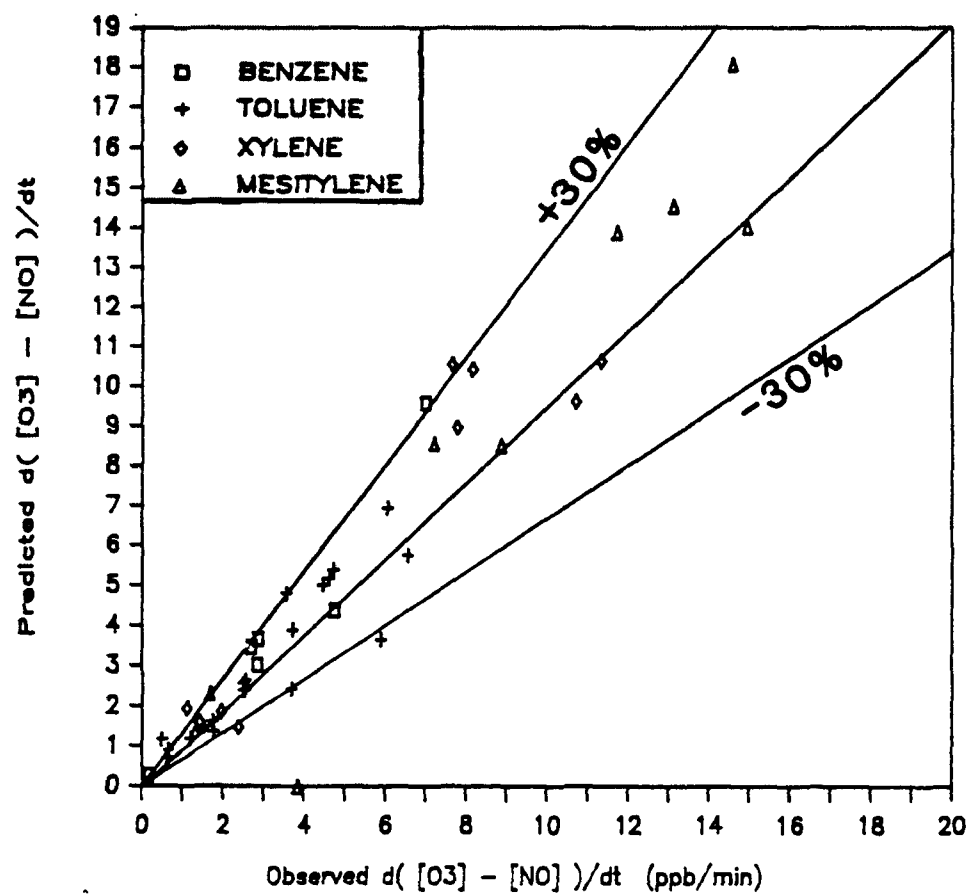
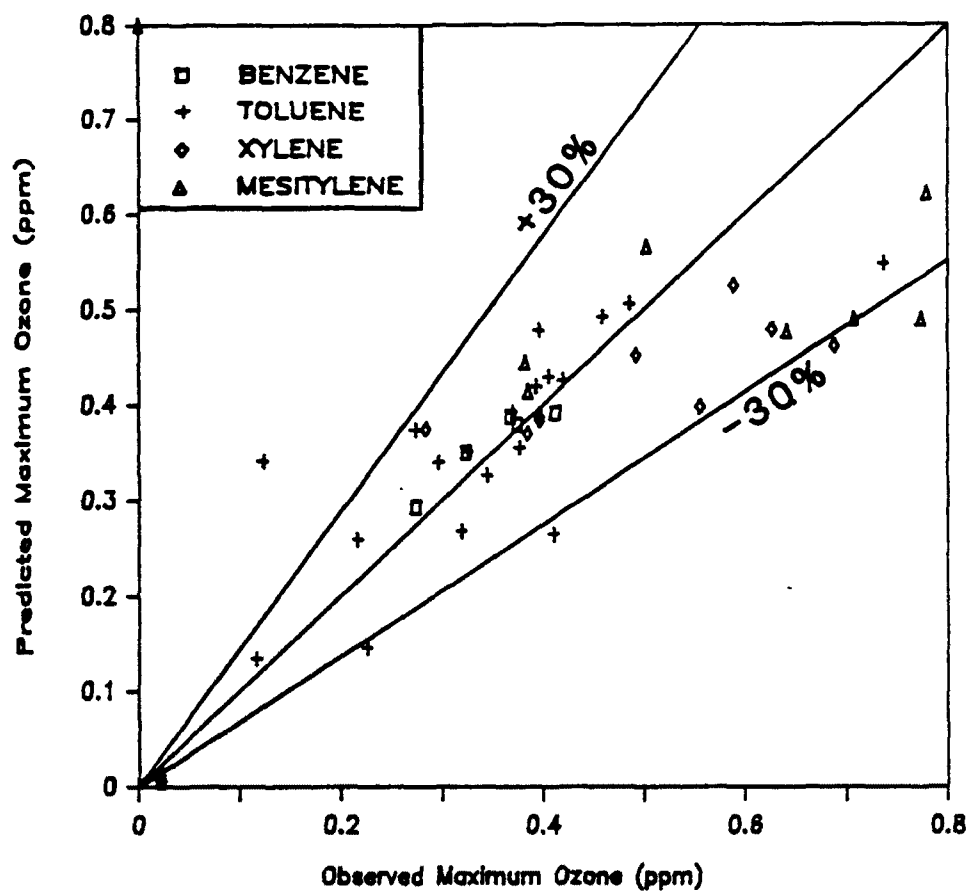


Figure 3. Predicted versus observed maximum ozone concentrations and timing parameter for aromatic - NO_x chamber experiments.

Simple Mixtures

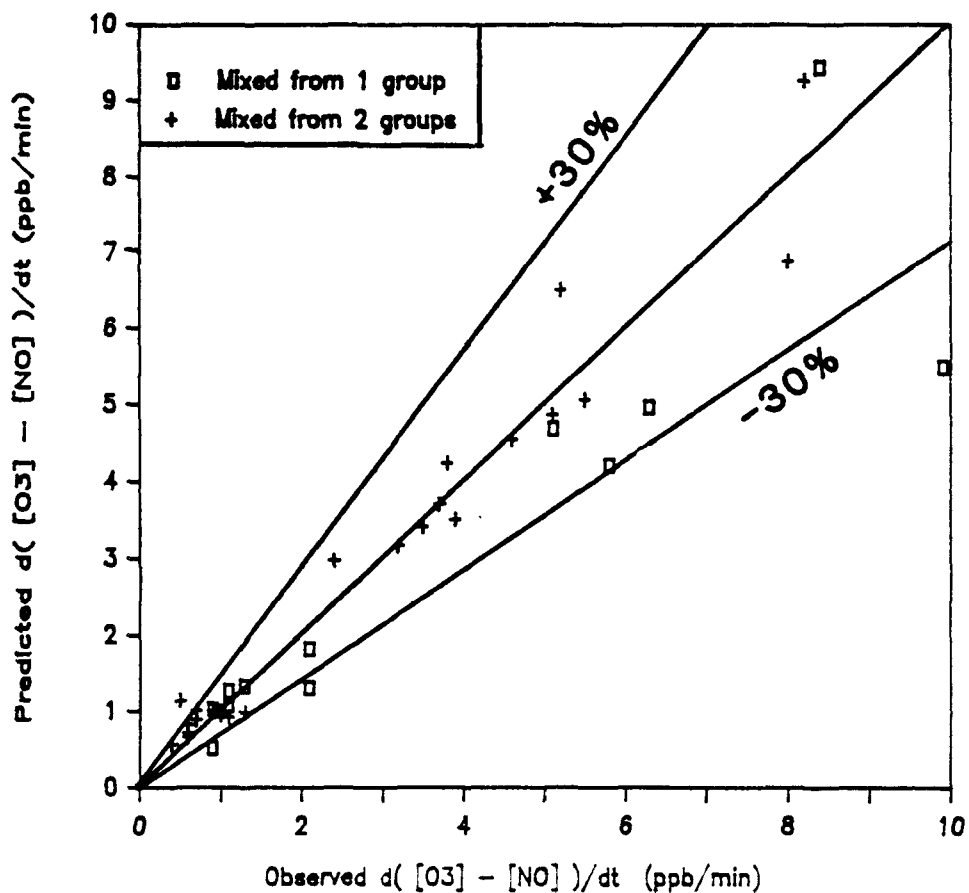
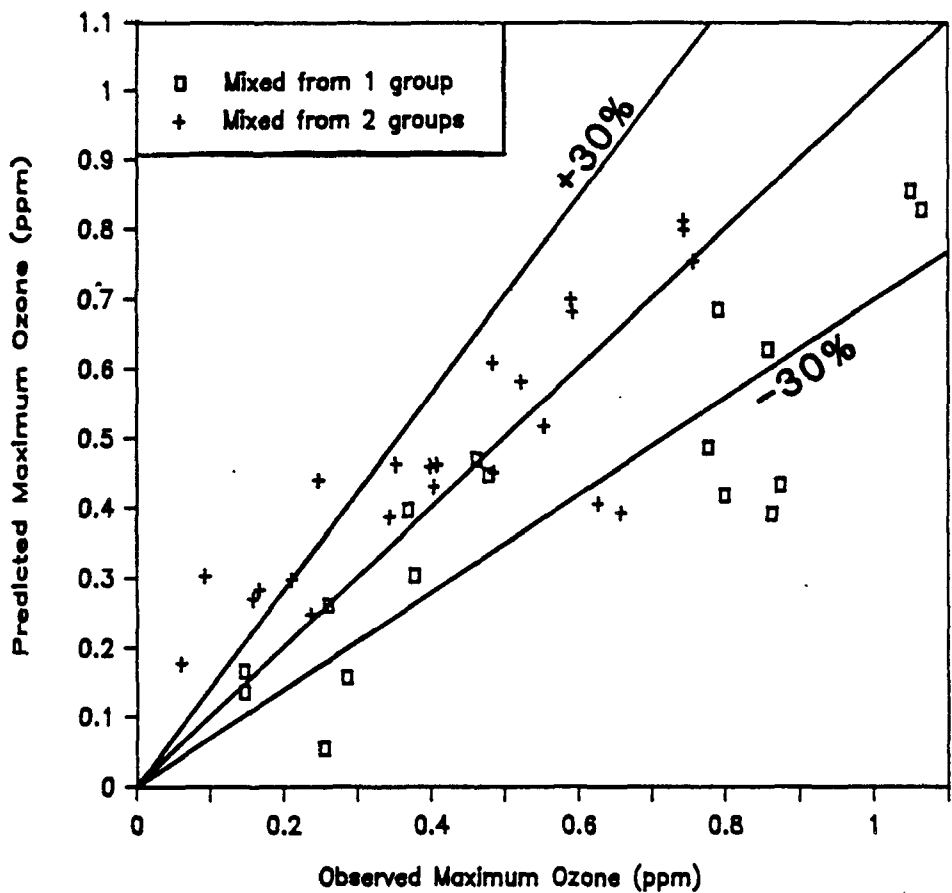


Figure 4. Predicted versus observed maximum ozone concentrations and timing parameter for chamber experiments with simple

Surrogate Mixtures

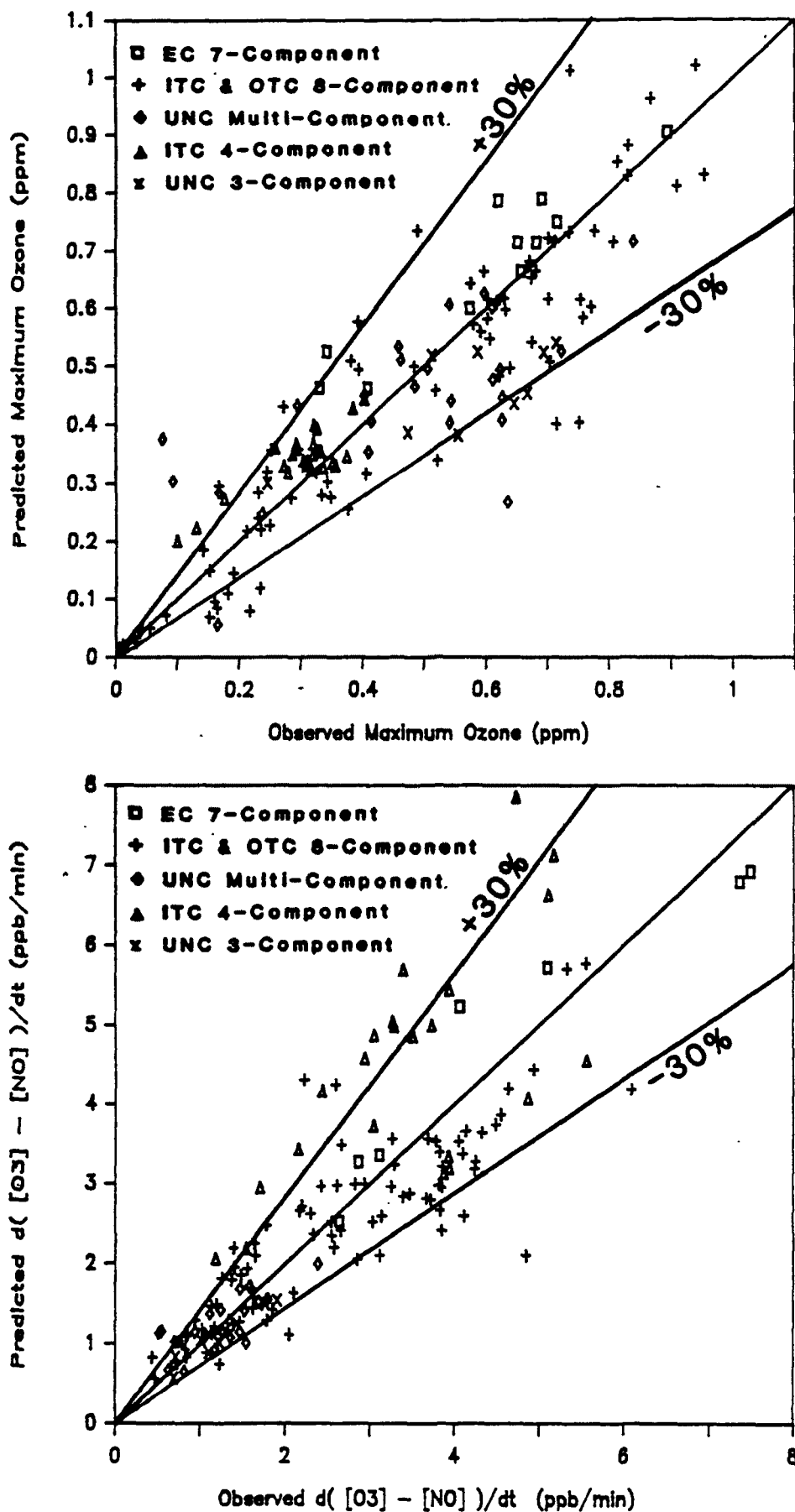


Figure 5. Predicted versus observed maximum ozone concentrations and timing parameter for chamber experiments with complex organic mixtures and NO_x .

UNC Auto Exhaust

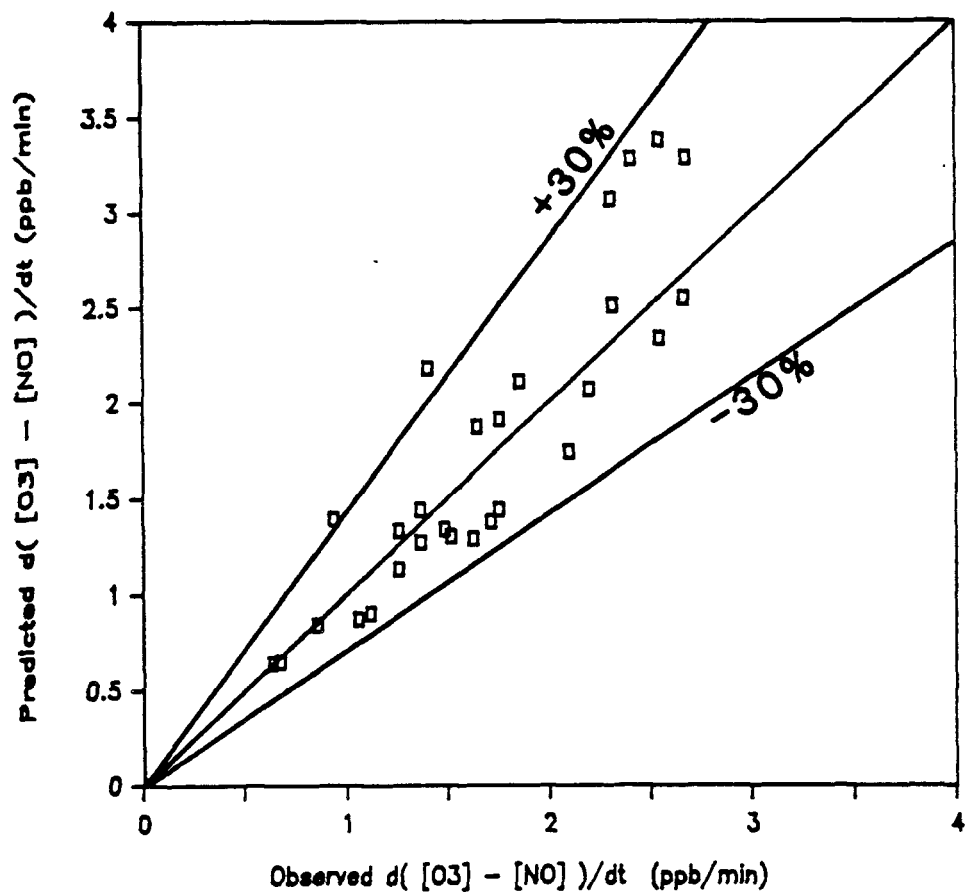
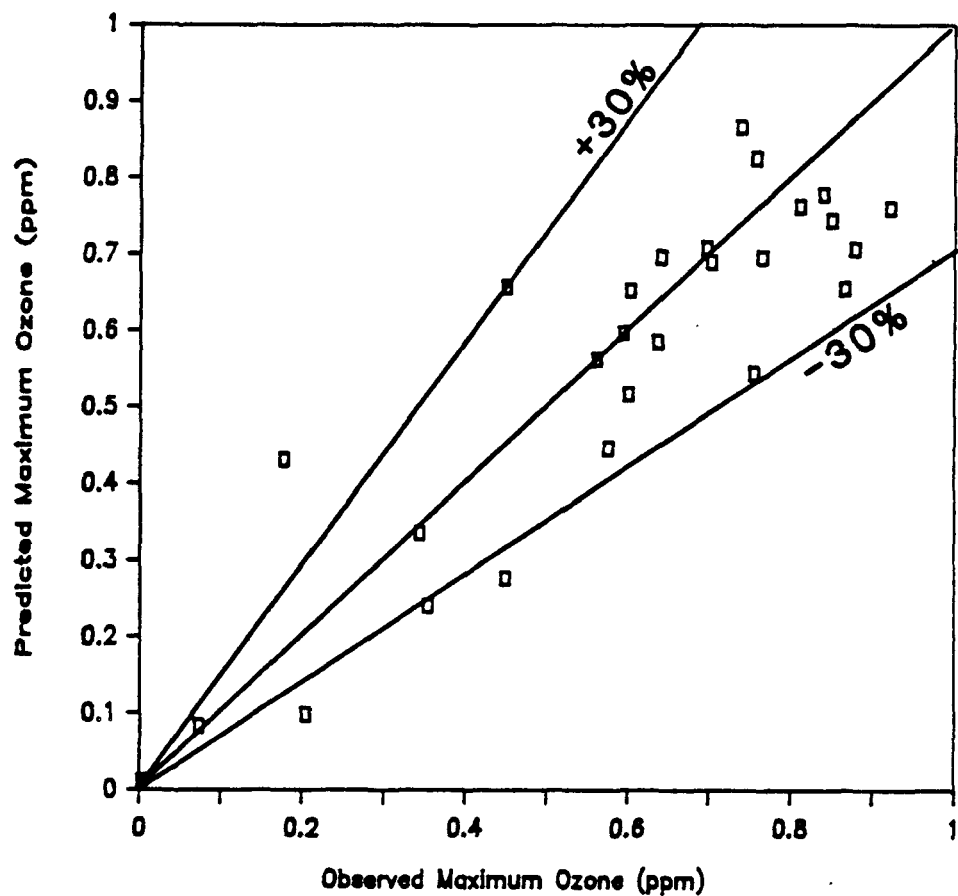
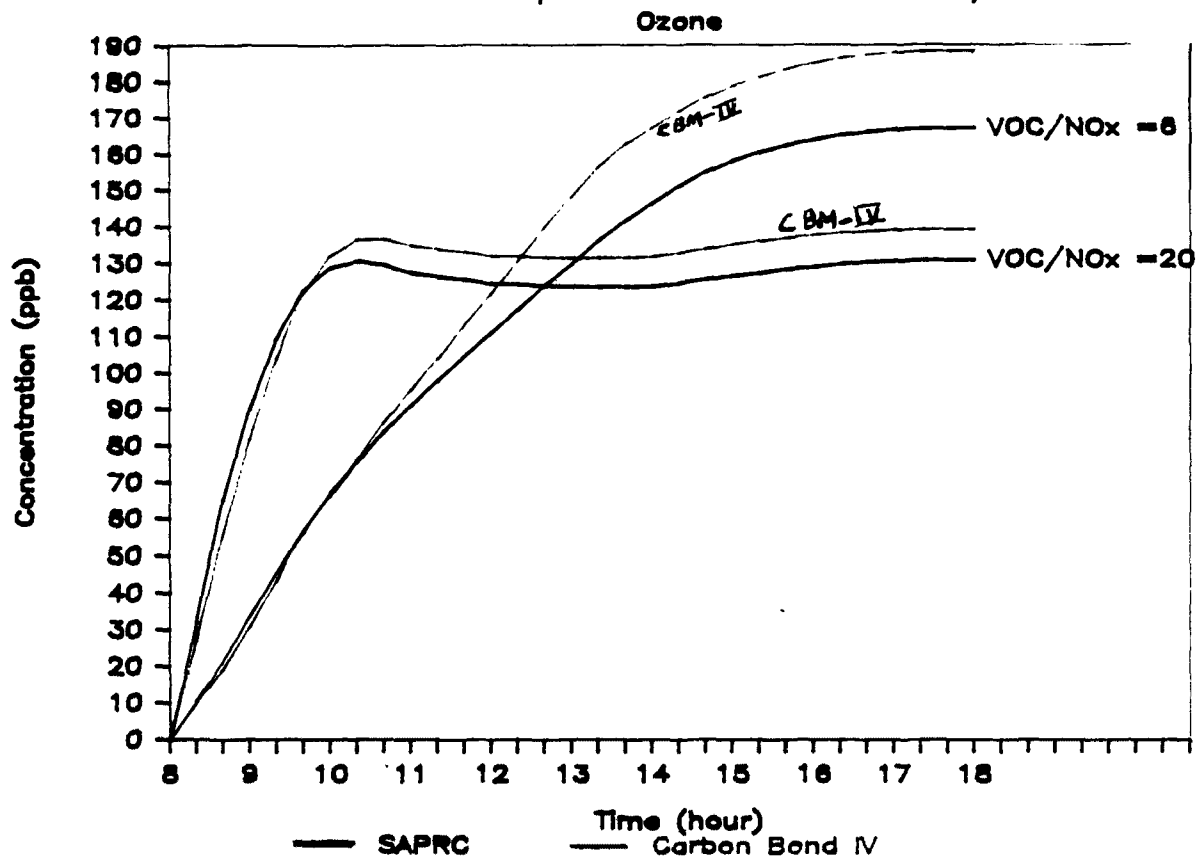


Figure 6. Predicted versus observed maximum ozone and a timing parameter, average $d([O_3] - [NO])/dt$ for auto exhaust- NO_x smog chamber experiments conducted in four chambers.

Mechanism Comparison at VOC/NO_x=6,20



Mechanism Comparison at VOC/NO_x=10

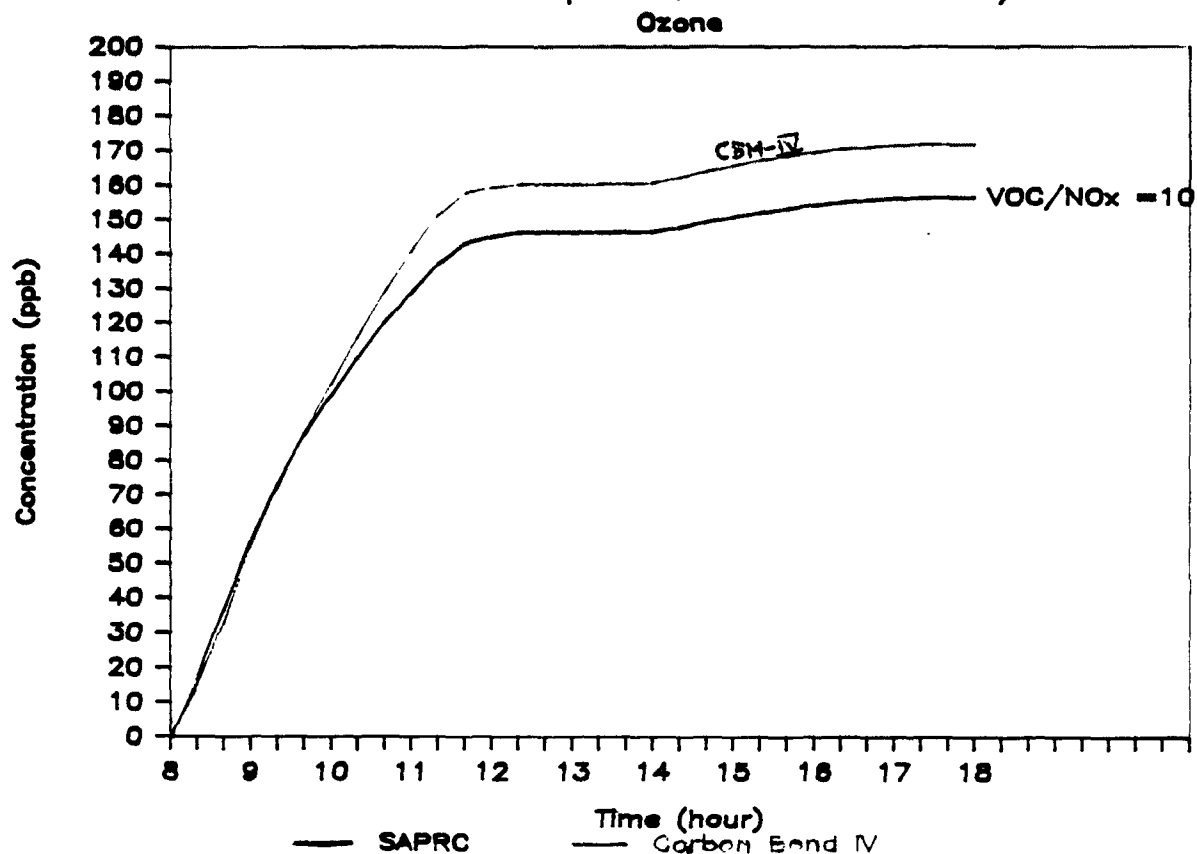


Figure 7. Ozone predictions from the SAPRC/ERT and Carbon Bond IV chemical mechanism for case with initial VOC to NO_x ratios of 6, 10, and 20.

Mechanism Comparison at VOC/NO_x = 3

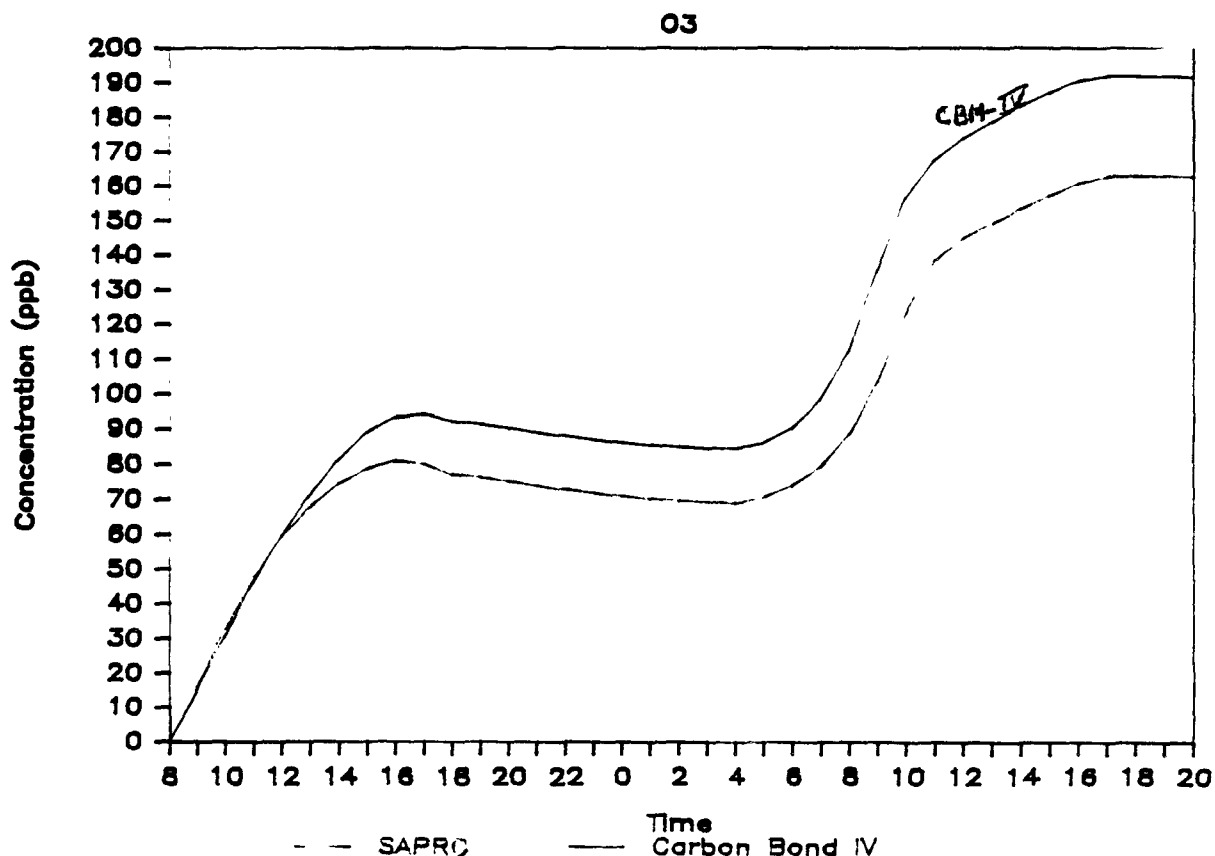
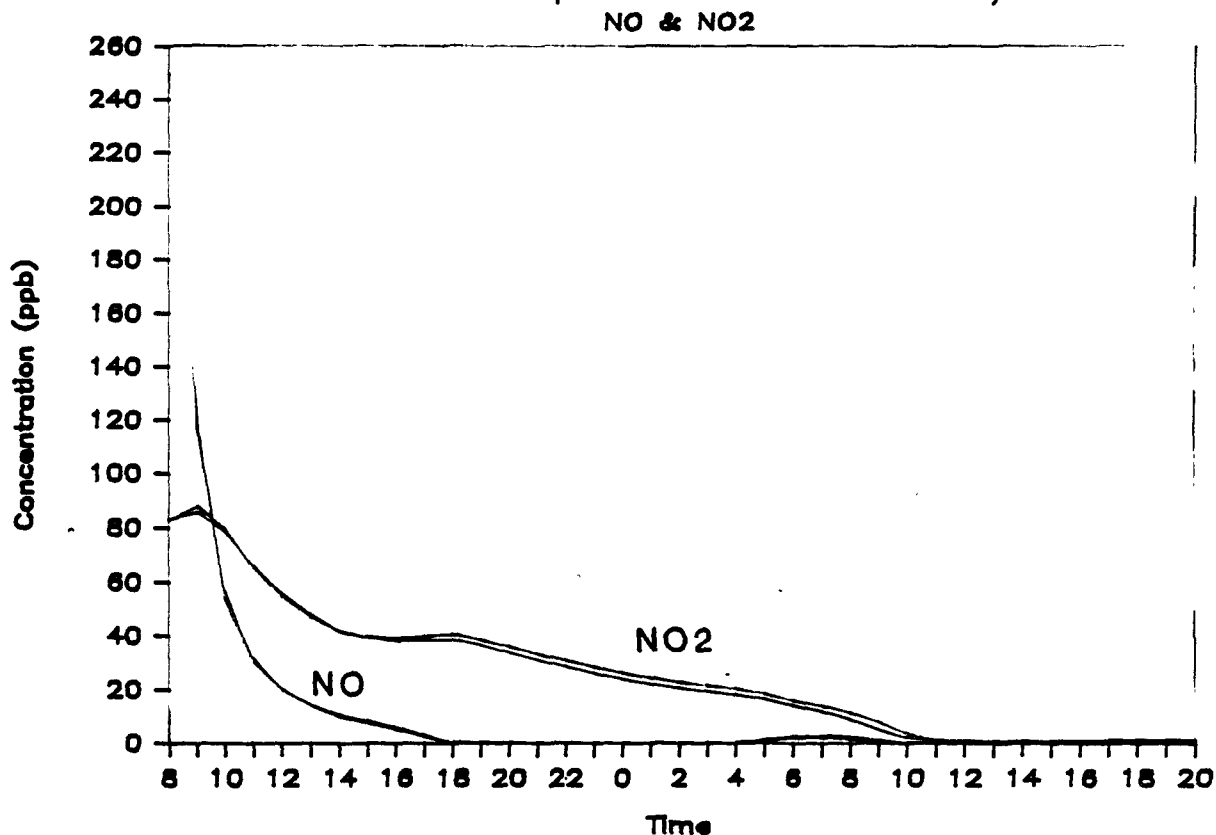


Figure 8. NO, NO₂, and ozone predictions from the SAPRC and Carbon Bond IV chemical mechanisms for a case with an initial VOC to NO_x ratio of three.

Mechanism Comparison at VOC/NO_x = 3

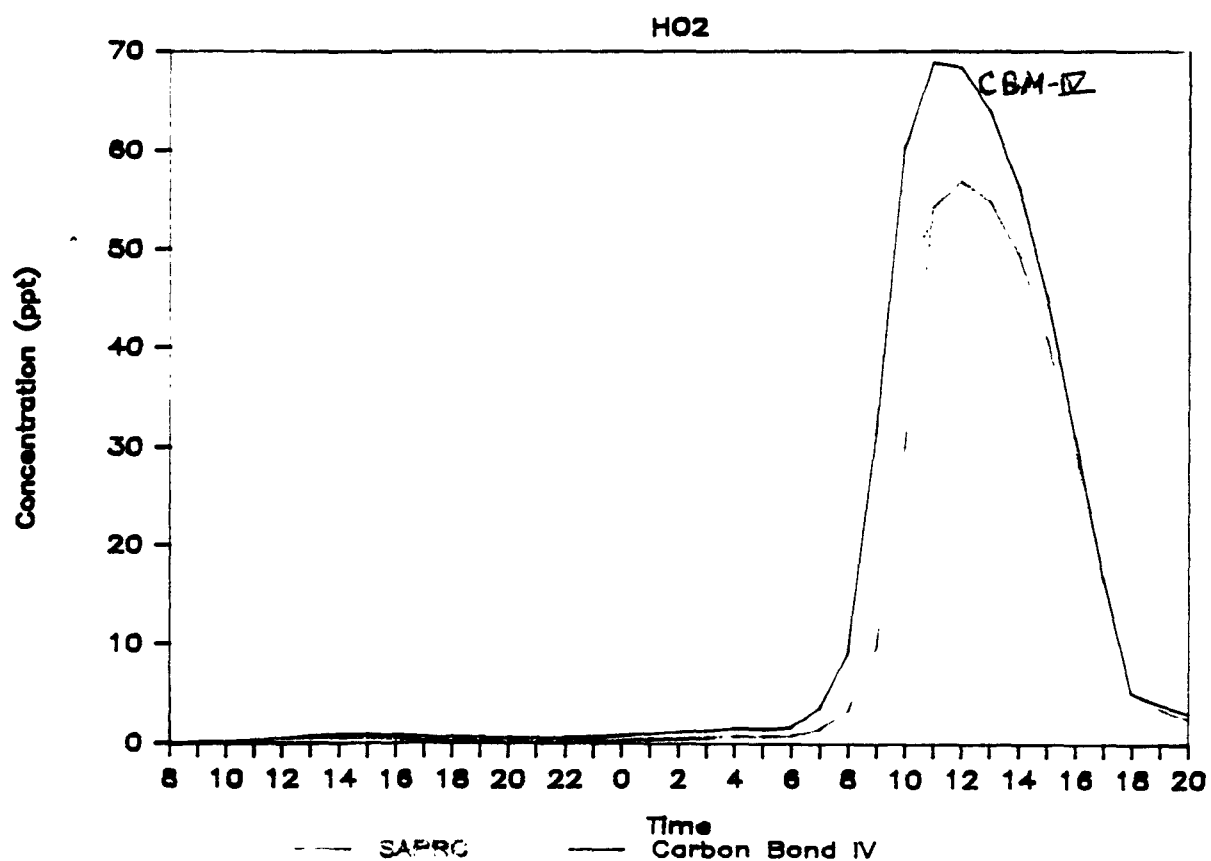
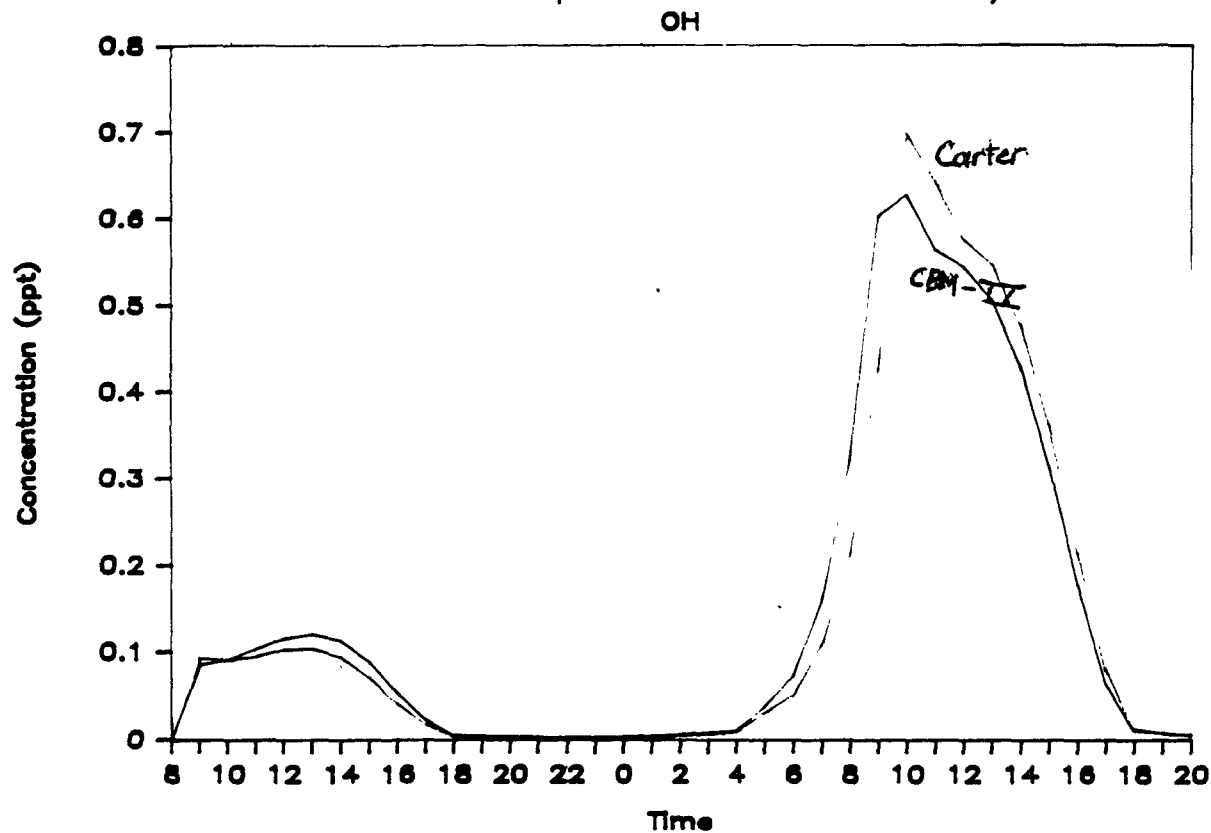


Figure 9. OH and HO₂ radical predictions from the SAPRC and Carbon Bond IV chemical mechanisms for a case with an initial VOC to NO_x ratio of three.

Mechanism Comparison at $\text{VOC}/\text{NO}_x = 3$

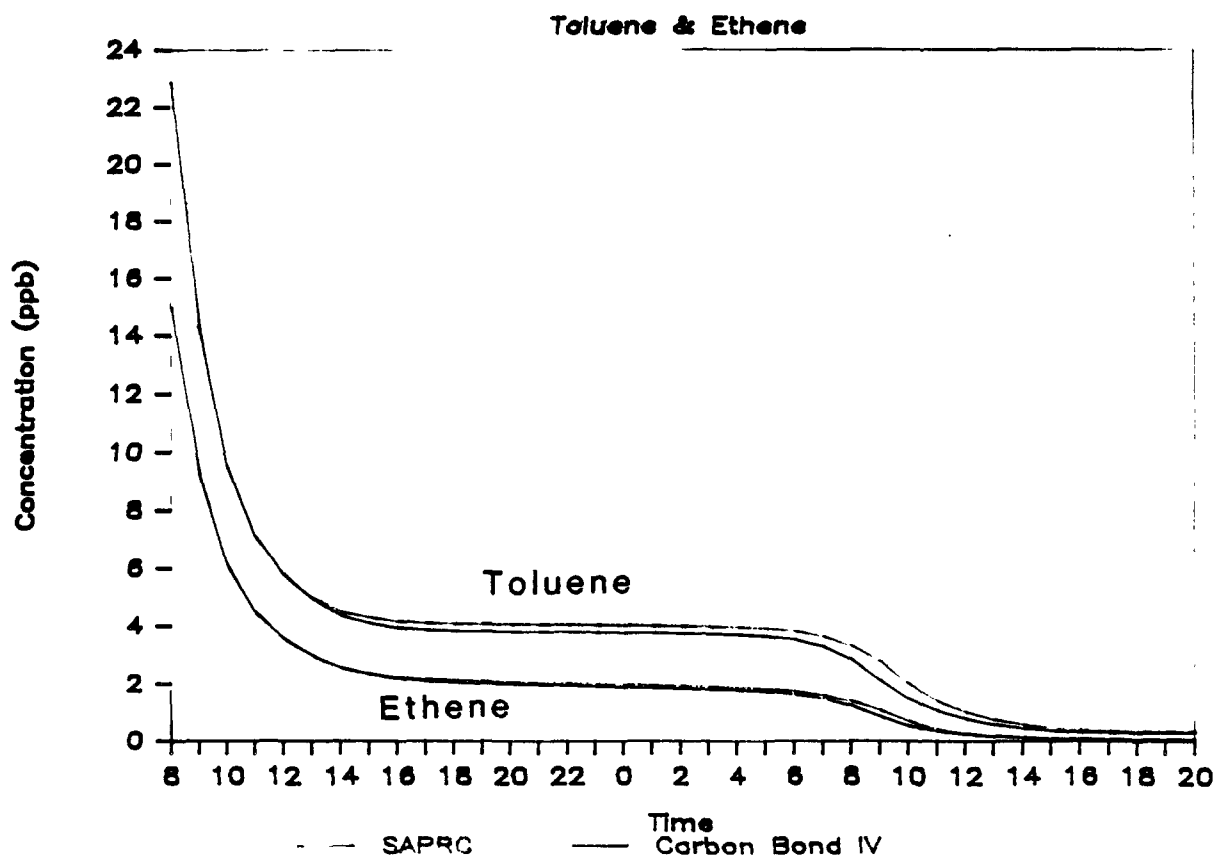
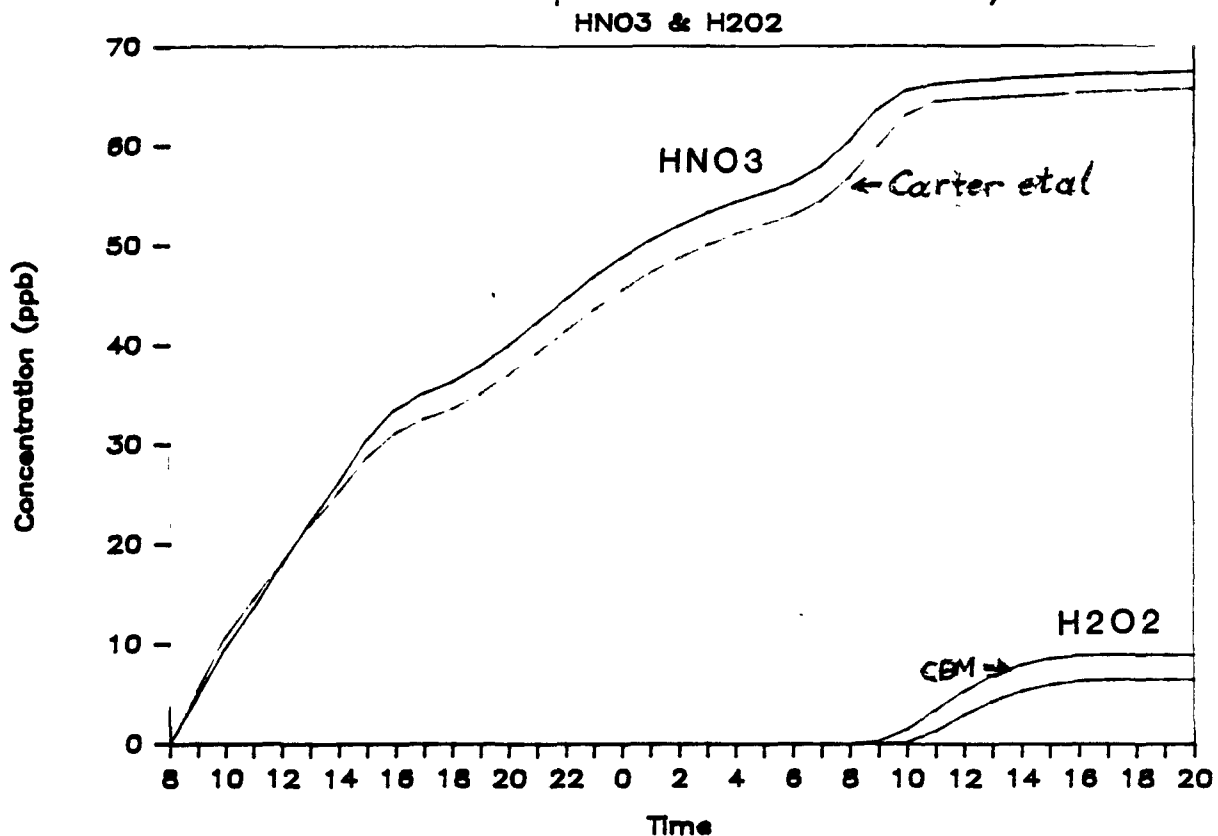


Figure 10. HNO_3 , H_2O_2 , toluene, and ethene predictions from the SAPRC and Carbon Bond IV chemical mechanisms for a case with an initial VOC to NO_x ratio of three.

REFERENCES

- Atkinson, R. Chem. Rev., 85, p. 69, 1986.
- Atkinson, R., A. C. Lloyd, J. Phys. Chem. Ref. Data, 13, pp. 315-444, 1984.
- Atkinson, R., A. C. Lloyd, L. Wings, "An Updated Chemical Mechanism for Hydrocarbon/Nitrogen Oxides Photooxidations Suitable for Inclusion in Atmospheric Simulation Models", Atmos. Envir., 16, pp. 1341-1355, 1982.
- Carter, W. P., A. C. Lloyd, J. L. Sprung, J. N. Pitts, "Computer Modeling of Smog Chamber Data: Progress in Validation of a Detailed Mechanism for the Photooxidation of Propene and n-Butane in Photochemical Smog," Int. J. Chem. Kinetics, 11, p. 45, 1979.
- Carter, W. P. L., A. M. Winer, J. N. Pitts, "Effect of Kinetic Mechanism and Hydrocarbon Composition on Oxidant-Precursor Relationships Predicted by the EKMA Isopleth Technique," Atmos. Envir., 16, pp. 113-120, 1982.
- Carter, W. P. L., F. W. Lurmann, R. Atkinson, A. C. Lloyd, "Development and Testing of a Surrogate Species Chemical Reaction Mechanism," Final Report EPA Contract No. 68-02-4104, ASRL, U. S. Environmental Protection Agency, Research Triangle Park, NC, 1986.
- Demerjian, K. L., J. A. Kerr, J. G. Calvert, "The Mechanism of Photochemical Smog Formation," in Adv. in Environ. Sci. and Tech., 4, John Wiley and Sons, NY, 1974.
- Hecht, T., J. Seinfeld, M. Dodge, "Further Development of Generalized Kinetic Mechanism for Photochemical Smog," Environ. Sci. Tech., 8, p. 327-339, 1974.
- Jeffries, H. E., K. G. Sexton, C. N. Salmi, "The Effect of Chemistry and Meteorology on Ozone Control Calculations Using Simple Trajectory Models and the EKMA Procedure," EPA-450/4-81-034, U. S. Environmental Protection Agency, Research Triangle Park, NC, 1981.
- Killus, J. P., G. Z. Whitten, "Effects of Photochemical Kinetics Mechanisms on Oxidant Model Predictions," EPA-600/3-83-111, U.S. EPA, RTP, 1983.
- Leone, J. A., J. H. Seinfeld, "Evaluation of Chemical Reaction Mechanisms for Photochemical Smog, Part II: Quantitative Evaluation of the Mechanisms," Final Report CA 810184, Atmospheric Sciences Research Laboratory, U. S. Environmental Protection Agency, Research Triangle Park, NC, 1984.

Niki, H., E. Daby, B. Weinstock, "Mechanisms of Smog Reactions," in Photochemical Smog and Ozone Reactions, Adv. in Chemistry Series, 113, p. 13, 1972.

Whitten, G. Z., H. Hogo, "Mathematical Modeling of Simulated Photochemical Smog," EPA-600/3-77-011, U. S. Environmental Protection Agency, Research Triangle Park, NC, 1977.

Appendices

- A. Background Document - Need for Chemical Mechanism Documentation
Kenneth Sexton and Harvey Jeffries**
- B. Background Document - The Science of Photochemical Reaction
Mechanism Development and Evaluation
Harvey Jeffries and Jeffrey Arnold**
- C. Workshop Agenda**
- D. Workshop Attendees**

Appendix A

Need for Chemical Mechanism Documentation

Kenneth Sexton and Harvey Jeffries

Need For Chemical Mechanism Documentation

Kenneth Sexton and Harvey Jeffries

Department of Environmental

Sciences and Engineering

University of North Carolina

Chapel Hill, NC 27514

Introduction

Documentation is needed from both the mechanism developer and the mechanism users to assure that implementation of a chemical mechanism is error-free. Documentation from the developer serves as guidance, and documentation from the user assures that no mistakes were made in its implementation, and that all assumptions, methods, and data are properly identified.

Implementation of a chemical mechanism usually requires more than representing a set of reactions in a model for simulation of the chemistry. The mechanism developer, consciously or not, uses the chemical model with certain methods, assumptions and conditions. Without guidance from the developer, users may use other methods, make other assumptions, with other conditions. The mechanism developer might disagree with these choices, resulting in denial from the developer that the chemical mechanism was properly implemented. Implementation of a chemical mechanism, therefore, requires that knowledge of the supporting methods, assumptions, and conditions, as described or suggested in guidance from the developer, be demonstrated or perhaps implemented, as well as the mechanism. If a user wishes to deliberately implement a user-modified version of a chemical mechanism or if he wishes to use it in a manner which does not conform to guidance from the developer, then the user should carefully identify the modifications (to either chemistry or procedures for use) as part of the description of "the chemical mechanism as modified by...".

Implementation Problems

There are several problems which can hamper implementation of a given chemical mechanism:

- communication problems resulting from lack of detailed specification and guidance;

- confusion between the mechanism and other computer programs used to support it or use it.
- problems encountered when using a different computer/system and/or supporting code due to unrecognized assumptions; and
- less obvious problems which may result in misuse and predictions different than those which would be generated by the mechanism developer.

We will discuss each of these below.

Lack of Detailed Specification

Many of the problems encountered are due to the lack of a clear description/definition of "the mechanism" by the developer. This, in turn, affects how it is implemented, used, modified, and reviewed by others in different situations.

Proper implementation and demonstration of proper implementation requires adequate documentation by both developer and the user if the work and results are to be adequately reviewed.

Ideally, proper implementation implies that the mechanism implemented is as published or distributed personally by the developer or the EPA. Versions obtained from another user may be outdated or modified incorrectly. Formal written guidance with examples are also needed to insure proper implementation. These are often lacking, or incomplete, or are only obtainable from the developer. There are often different versions of a given mechanism in use. The "version problem" results from continuing development and mechanism changes. Ideally, one should obtain the chemical mechanism and guidance directly from the developer or the official distributor (such as the EPA). Guidance has changed after a mechanism was adopted. Examples should be used to verify that implementation of both the mechanism and guidance.

The important assumptions, limitations, procedures, or conditions which are by definition of the mechanism developer an inherent part of the mechanism, should be identified. These might include approach and structure, treatment of light, temperature, pressure, water, radical initiating sources (general or specific to a chamber or an atmosphere), empirical chemistry, background reactivity, composition treatment, treatment of secondary functional groups of hydrocarbons, and rate or stoichiometry modification with composition variation (state whether to do so or not). For example, some mechanisms ignore temperature effects and give a rate constant evaluated at some temperature (often 298K) and more commonly A-factors and activation energies are used. In some mechanisms, photolytic rates are expressed

as ratios to the NO_2 photolysis rate and in others the rates are calculated explicitly. The sources of rate constants including photolytic rate constants should be identified. Quantum yields, cross-sections, and light intensity assumptions as a function of wavelength should be given.

The developer should indicate which reactions are thought of as explicit and which are simplifications. He should describe which reactions can be updated as new kinetic information becomes available and which are to be left "as is" because of other compensating chemistry.

Confusion About Support Programs

Confusion and errors arise from the concept of "the mechanism" and some larger model which may use it. Also the distinction should be made between the theoretical description of the chemical mechanism, and any computer code currently used to represent and implement the representation, or code used to support it. The next section discusses some of the problems which can be avoided with the awareness of these concerns.

Using Different Computer Systems

An awareness of the purpose and assumptions of supporting programs for a given chemical mechanism is necessary in implementation on a different computer/system or with different programs. Depending on the type of example simulations which a new user may be trying to duplicate, different support programs or procedures will be used. One of the main problems in reproducing example simulations are the dependence on supporting programs which have characteristics which may be difficult to reproduce or may have their own peculiarities or errors. For example different programmers have implemented the "Characteristic Curve Mixing Height Profile" for estimating dilution of an air parcel. Different numerical techniques can be used to generate time-varying emission rates for input to a photochemical system from hourly averaged emission estimates. Calculation of photolytic rates from even the same source of input data can be affected by the algorithm and numerical methods used. Only by duplicating these routines or by using the specific output of these routines can the examples be reproduced accurately.

The descriptions of supporting programs (both the original and new) including their purpose, assumptions, and limitations should be given. Ideally the listings of all code should be available. This is to allow for comparison to assure that the code is error-free, can be executed, and adheres to the developers guidance.

Often the "appearance" of the mechanism must change when it is implemented and solved with different code on a different machine. Such changes allow for several types of problems. Examples are needed to allow the new user to verify proper implementation.

Often there are insufficient examples provided by the developer to assure a new user of proper implementation or proper use of the chemical mechanism and/or programs used. Often only printouts (no plots) are provided with results listed for only a few large time- steps. Enough data should be printed to allow for careful comparison with the timing of important chemical events.

Mechanisms appear or can be obtained in different forms. Some forms are not clear or may be incomplete descriptions. Mechanisms are often not implemented in the form that would be used in a journal publication. Journal articles often do not provide enough information to reproduce the examples shown. Sometimes the mechanism representation file is distributed. This mechanism file read by the computer might appear very different. Sometimes the simulation printout lists a representation of the mechanism. Space limitations might require the developer to round-off stoichiometry, rate constants, A-factors and activation energies. The reactions might be rearranged to conform to the syntax requirements and limitations of a computer program which reads it. Reactions might be condensed to save space. Rate constants might be evaluated at some given temperature or for some assumed water concentration.

Clear unabbreviated listings of the chemical mechanism are therefore useful in implementing by a new user. Otherwise, new users may proceed making their own choices and assumptions for aspects not addressed in an incomplete discussion or description of a chemical mechanism model and thus create a different representation than the mechanism authors intended.

When a particular algorithm is essential to the application of the mechanism, for example, a particular way to treat the time variation in hourly averaged emissions, problems may arise. Different machine/code environments can present problems for either the direct transport of code between machines/systems or for the translation of listings of code to new code.

These effects may not be obvious. Very often there are differences between compilers on different machines. The compiler or system treatment of numerical errors may be different. There may be machine limitation differences (eg numerical accuracy). Compiler or language differences may cause translation problems. Poor

coding practices such as dependence of “tricks” (taking advantage of a particular odd trait of a language or compiler) may cause translation problems. An example would be that some compilers perform an auto-zero of uninitialized variables. Problems of this type highlight the need for simulation examples from the developer. The examples provided and used for comparison should demonstrate and test all portions of the code. Ideally, the source code (at least listings) should be provided, with code documentation of high standards. The code documentation should include description of above concerns and identification of dependencies on peculiarities of compiler and canned procedures. References should be given for sources of algorithms and code fragments.

Given the possible causes of disagreement in attempting to reproduce an example simulation, there needs to be a determination as to what is considered “Acceptable Accuracy”.

The Original Errors Problem

Sometimes errors are made by the developer which are subsequently detected by another user. There is often a problem with determining which errors can be corrected by a user without “changing the mechanism”. Publication and distribution errors have been found. Typographical errors by the user are most common during implementation. Typographical errors in species names, for example, can be very difficult to detect. Bad copies or wrong versions or examples are sometimes distributed by the developer. Implementation and misuse errors by new users often result from insufficient guidance from the developer.

A Proposal

The chemical mechanism developer should provide a guidance document which includes examples. This document should begin with a written general description of the chemical mechanism including the approach and type of representation. Assumptions and limitations inherent in the mechanism formulation should be identified to help explain the general use procedures. Describe which “species” in the mechanism represent the different chemical species in atmosphere. Identify which species and chemistry are explicit. Identify which reactions if any are modified (rate or stoichiometry) by composition and explain how with examples, identifying assumptions. Explain how water, oxygen, pressure, M, light, and temperature are treated and represented and incorporated into the mechanism (rate constants). All reactions that are dependent on temperature, water, and light, should be identified. Identify the source of photolytic rates and explain treatment of light. Outline pro-

cedures for determining light for different locations or conditions. Explain how (if allowed) to update the mechanism if photolytic rates need to be changed later.

For implementing a given mechanisms representation into new users code, the new user must have a listing of the mechanism, preferably directly from the developer (recently) which was also used to generate the example also provided. This should be included in the guidance document. To minimize chance for typographical errors by a new user, the mechanism should ideally be provided in a computer file in ASCII which would allow others to read directly.

The new user should obtain example simulations, preferably from the user at the same time the mechanism is supplied. Examples should be as complicated as the intended use (ie try to show examples using variation of several parameters, and options). The examples should provide prediction results with plots and printouts.

The examples should provide initial conditions (both original "real" and for the mechanism) including chemical species, light, temperature, and water, especially if these effect rate constants. Descriptions and examples of input condition calculations from composition speciation and non-speciation (ie from detailed HC analysis and from none or incomplete HC speciation) must be described. The compounds that are considered non-reactive by the mechanism developer should be identified. If these assumptions affected the rate constant of a reactions, these should be identified so that changes can be performed by the new user (if allowed by definition). Any assumptions about default compositions should be included for emissions, surface transport, and aloft materials. This would include aldehydes, hydrocarbons, NO, NO₂, O₃, *etc.*

If heterogeneous processes are included in the mechanism, these should be identified and for which types of simulations these are intended. Any guidance on adjustment of these rates for different conditions should be given.

The new user should be able to duplicate the results provided in all of the examples. Inability to duplicate the results obviously indicates an implementation problem.

An outline of an example guidance document is given in Table 1.

Table 1. Example Outline of Guidance Document

1. Structure of Chemical Mechanism

A. General Description

i. intended use

- ▷ basic input
- ▷ basic output

ii. representation schemes

iii. basis

- ▷ sources of reactions and kinetic information
- ▷ reaction simplification techniques
 - condensation techniques
 - approximation techniques
 - verification
- ▷ techniques for rate constant approximations

iv. assumptions and limitations

v. disclaimer

B. Inorganic chemistry

i. explicit chemistry

- ▷ rate constants
- ▷ assumptions and omissions
- ▷ references

ii. simplifications

- ▷ assumptions
- ▷ user adjustments
- ▷ references

iii. adjustments for heterogeneous processes

- ▷ representations and assumptions
- ▷ measurements and experimental data required
- ▷ procedure for estimation of heterogeneous rates

C. Organic chemistry

i. explicit chemistry

- ▷ rate constants
- ▷ assumptions and omissions
- ▷ references

ii. simplifications

- ▷ assumptions
- ▷ user adjustments
- ▷ references

iii. natural background reactivity

- ▷ assumptions
- ▷ representation

iv. adjustments for heterogeneous processes

- ▷ representations and assumptions
- ▷ measurements and experimental data required
- ▷ procedure for estimation of heterogeneous rates

D. Light and Photolytic Rates

i. Assumptions

ii. Treatment

E. Temperature

i. Assumptions

ii. Treatment

F. Deposition

i. Assumptions

ii. Treatment

2. Representation of Chemical Reactions

A. syntax

B. rates and activation energy units

C. light dependent reactions

D. water dependent reactions

E. oxygen assumptions

F. "M" and pressure dependent reactions

G. description of file of mechanism representation

3. Photolytic Rates

- A. Theory, approach and assumptions
 - B. values needed for mechanism species and units
 - C. Method of Calculation
 - D. Sources of data
 - E. Procedures for Determination of Photolytic Rates
 - F. Description of programs and data files
 - G. examples
4. Determination of Initial Concentrations
- A. NO_x
 - B. Hydrocarbon Calculations
 - i. Types of input data
 - ▷ detailed site-specific analysis: criteria for HC analysis to be considered complete
 - ▷ partial analysis: guidelines for estimating composition of non-measured VOC
 - ▷ no detailed data and default compositions
 - ii. Methods of Calculation
 - iii. Sources of Data
 - iv. Procedures for ...
 - v. Description of programs and data files
 - vi. examples
 - C. O₃
 - D. CO
 - E. other compounds
 - F. examples
5. Hydrocarbon Calculations for Reaction Modifications
- A. Types of modifications
 - i. rate constants
 - ii. stoichiometry
 - B. Method of Calculation
 - C. Sources of Data
 - D. Procedures for ...

- E. Description of programs and data files
- F. Examples
- 6. Suggestions for Support Programs
 - A. Numerical Solution of Chemical Kinetics
 - B. Emissions entrainment
 - C. Meteorology (dilution)
 - D. Sun's Position
 - E. Photolytic Rates
- 7. Example Simulations
 - A. Smog Chamber
 - B. Urban Atmosphere
- 8. Listings of Programs
- 9. References
 - A. for citations made above
 - B. for all available literature that describes the scientific basis and testing of the reaction mechanism.

Guidance from the EPA for the proper general use of photochemical models should be developed. "Use" would include documentation by both mechanism developers and users. Such documentation would require demonstrating the ability to reproduce an example and a full explanation of all assumptions and preparatory calculations. This is needed to assure proper implementation, use, and review. Guidance, policy and definitions of model terms and procedures need to be developed. The EPA should define the format of a "guidance document" to be written for a chemical mechanism and a "user modeling documentation report". The format of the guidance document should include at least one example calculation for every part of the mechanism input and use. The documentation reported by the user should cover the same concerns of the mechanism developer. If the user followed the suggested guidance then the documentation required should actually be easily prepared. User specific input data of course should be listed with example calculations.

Chemical mechanisms and photochemical models are implemented with computers because of the complexity of the problem. The implementation of chemical mechanisms and photochemical models in general would be much more successful if those in this field agree to conform to the documentation standards used in computer programming.

Official policy and guidance on modeling, the requirement of a guidance document from the developer, documentation from the mechanism/model user, and insistence of good documentation in general could form the basis of a model quality assurance program which would help assure proper implementation of chemical mechanisms and models. Some important policy questions pertaining to implementing chemical mechanisms are given in the next section.

EPA should have responsibility of being the official source of chemical mechanisms and photochemical models. Also the EPA should have responsibility for reviewing and determining the adequacy of guidance for each. Ideally, part of the review of modeling for EPA would be that EPA personnel could reproduce the results, ideally with their own code implementing the chemical mechanism with their own support programs.

An Example of Current Guidance and Mechanism Documentation

It is interesting to examine a recently guidance document for a new mechanism developed for the EPA and to compare it to the outline presented above and some of the ideas presented here.

"Guidelines for Using OZIPM-3 with CBM-X or Optional Mechanisms, Volume I," is mostly an instruction manual for using OZIPM as the title states. It starts with a concise overview of the model. Definitions are given and the technical basis of the model is discussed. Necessary calculations are explained, the mechanism is listed, and example simulations are provided. The explanation of how the programs work and the code are presented in Volume II. However, some important items seem to be missing.

The document explains that the photolytic rate constants are represented with a multiplication factor to the NO₂ photolytic constant, and that the user can change these. However, there is no explanation as to the source of any of these rates, making it difficult for a user to determine how much one should change these to attain another absolute rate. There is also the note that these rates have been increased to represent an altitude of 600 meters. These new rates are not supported by any additional discussion or references. Since a user may be interested in modifying these rate constants, it would be helpful to know what assumptions and techniques were used to estimate the existing values.

The mechanism structure and chemistry are only discussed in a few paragraphs. Several references are given to earlier reports for a more complete discussion of the development of the mechanism, technical issues, and associated calculations. It is in some of these earlier reports that show the simulation of smog chamber experiments results.

It appears that at least in this case, several reports are needed to document and describe fully a chemical mechanism and how to use it in a photochemical model. It illustrates the fact that a listing of a mechanism is incomplete documentation for implementation in a new environment.

Questions for Discussion

1. Does the mechanism developer have complete control on the definition of the mechanism and how it is used?
2. Does the validation database or procedure become part of the definition or description of a chemical mechanism? Does it effect the use of the mechanism? Should it limit the application?
3. What aspects of a mechanism or model can be "updated" or "corrected" (with more scientifically defensible information) and by whom? What level of documentation of these changes should be required? What should EPA's response be to a SIP to modeling performed using adjusted photolytic rates from different assumptions of actinic flux from different assumptions of albedo or cloud cover? Is the user using "the validated mechanism" once such a change is made?
4. What changes can be made to "validated" chemical mechanisms/models without invalidating? Are there different answers to this question from the scientific and policy and regulatory groups?
5. Should guidance for model use recognize and distinguish between regulatory modeling vs research modeling and provide different guidance for each?
6. Are there circumstances in which it is satisfactory for a model/mechanism user to "correct" known "incorrect" or "scientifically unsound" aspects of a mechanism before its use?

Appendix B

The Science of Photochemical Reaction Mechanism Development and Evaluation [Strawman Document]

Harvey Jeffries and Jeffrey Arnold

The Science of Photochemical Reaction Mechanism Development and Evaluation

Harvey Jeffries and Jeffrey Arnold

**Department of Environmental
Sciences and Engineering
University of North Carolina
Chapel Hill, NC, 27514
(919) 942-3245/966-5451**

November 17, 1986

**Prepared for
U.S. Environmental Protection Agency
Workshop on
Evaluation and Documentation of
Chemical Mechanisms Used in
Air Quality Models
December 1-3, 1986**

Preface

The role of this document is to stimulate discussion. To accomplish this goal, we have

- ▷ defined the scientific process, both in a general sense and more specifically for the field of air pollution modeling;
- ▷ provided evidence that our view is accurate;
- ▷ argued the implications of this view;
- ▷ made recommendations for continuing the advancement that has occurred over the last 15 years.

To accomplish the first step we have relied extensively on the work of Thomas Kuhn, a science historian and Professor of Philosophy at the Massachusetts Institute of Technology, whose work has been described by the journal *Science* as "A landmark in intellectual history."

We believe that dialogue within the community of scientists practicing in the field of air pollution modeling and testing is needed to plan an action to counter a developing preception, one that questions the efficacy of modeling for EPA's needs and thus has a potential impact for the support the field has enjoyed for nearly two decades. We believe that a clarification of process would be beneficial to the community. Clarification must include examination, and examination might led to new insights that could produce new processes, or perhaps to a reaffirming of methodology upon which progress had been made. In any case, a strategy for EPA operations in the application of models is necessary to avoid a dichotomy of explain-predict.

It is hoped that you will help in this process by examining your own philosophies, comparing and contrasting them with others in our rather small community. And that you then come prepared to discuss and explore the critical issues in an open and amicable manner.

Perhaps EPA will learn something that will allow them to:

- ▷ improve the design and structure of their programs;
- ▷ maximize the benefits (increased knowledge) from all of our efforts;
- ▷ encourage the continued participation in their programs; and
- ▷ promote cooperation and communication.

Contents

1	The Need for Science in EPA Models	1
2	Nature of Science	8
	Textbook Science	8
	A Kuhnian View	8
	A Psychological Interpretation	16
	The Rest	19
3	The New Paradigm	20
	The Beginning of Explanation	20
	Niki, Daby, and Weinstock	23
	The Practice of Normal Science.	25
	Early Normal Science Work	27
	First Class Work	27
	Tools Needed to Advance the Theory	27
4	First Explanation	29
	Demerjian, Kerr, Calvert	29
	Mechanism Development	30
	Realism	31
	DKC's Rules	33
5	First Reformulation	41
	Hecht, Seinfeld, Dodge	41

6	Second Explanation	50
	Durbin, Hecht, Whitten	50
	Late 1970s Explicit Models.	55
	The Beginning of the First Crisis	57
7	Second Reformulation	58
	The Dodge Mechanism	58
	Falls and Seinfeld, McRae Mechanism	60
	Carbon Bond Mechanism	63
	Atkinson, Lloyd, Wings Mechanism	67
	A Contrast	69
	Mechanisms Compared	69
8	Response to Uncertainties	73
	Whitten's Hierarchy of Species for Testing	73
	Carter's Chamber Radical Experiments	79
9	Third Explanation	82
	Atkinson and Lloyd Review	82
	Leone and Seinfeld Master Mechanism	84
10	The Essential Tension	88
	Potential and Actualization	88
	Explanation and Prediction	93
	Testing and Fit	95
	Verification and Vindication	95

11 Methodologies	98
Mechanism Vindication	98
Classes of Scientific Work	99
Agreement	100
References	105

The Need for Science in EPA Models

The Government Accounting Office recently told Congress,¹ that the Environmental Protection Agency uses air pollution models “as tools in administering the Clear Air Act ... and often [models] are the only option available.” For EPA then, models *are* more than useful representations of selected real-world events. According to the GAO, however, the foundations of these air quality models “are based on assumptions, approximations, and judgments, all of which affect a model’s accuracy, validity, and reliability.” Furthermore, EPA’s most reliable models, according to the GAO, “estimate actual pollution concentrations within a range of minus 50% to plus 200%.” The GAO said EPA officials indicated that “more precise results are unlikely because of the limitations in the science.”¹ The GAO implied that improvement *was* possible, however, by quoting the American Meteorological Society to the effect that “the scientific community must improve model physics, calculation techniques, and model input in order to obtain more precise results from models.”

These GAO ‘findings’ led the House Subcommittee on Oversight to ask EPA Administrator Lee Thomas to determine “which model is the most reliable,” and to “explain the range of reliability, and identify the uncertainties of each.” Further, the subcommittee asked Thomas: “When is the ‘scientific community’ going to make the model improvements mentioned and what is the timetable?”²

Though these questions are directed specifically to the science in air pollution modeling, they lead on to policy issues, and thus to the overarching regulatory function of EPA. Basil Dimitriadis³ said that EPA was disturbed by the fact that

models of "insufficiently documented validity" were being used to make costly control strategy decisions. Further, even the "relative validities" of the various models were difficult to document. This led Dimitriades to say that "it is imperative that model-type guidelines issued by EPA be of well-documented utility and validity." He added, "EPA is committed to expend the necessary resources to meet this requirement." By examining and ranking approaches to evaluation—not the models themselves—the EPA workshop described by Dimitriades is part of that effort.

The majority of the GAO comments were about air quality simulation models, especially dispersion models with little or no atmospheric chemical transformation processes. EPA has had a long-standing cooperative agreement with the American Meteorological Society to review and recommend procedures to evaluate the performance of air quality models.⁴ Nothing in this work, however, has ever addressed chemical transformation. While the new "Guideline On Air Quality Models (Revised)"⁵ includes discussion on accuracy and uncertainty of models and makes recommendation on model use, it only contains a recommendation to use the Urban Airshed Model or city-specific EKMA. It directs the user's attention to a number of user's guides for these models, yet none of these discusses the "utility, reliability, and uncertainties" in the various available chemical mechanisms.

The focus of our discussion is the chemical component of those air quality models intended to simulate situations in which the air pollution problems are those caused by chemical compounds formed in the atmosphere. In the past, these models have been ones of ozone formation: OZIPM/EKMA models, the Urban Air Shed Model, and the Regional Oxidant Model. When alternative chemical mechanisms of ozone formation have been used in air quality models, they generally have given different predictions.^{6,7} Non-scientists who must by law use these models want to know "Which prediction is right?" Because the definition of adequate model testing—and this is one of our goals—has not been sufficiently addressed, the policy maker is not alone; even the scientist working in the field has difficulty dealing with this question. And unresolved issues in the use of these air pollution models are no longer restricted to ozone formation chemistry, but now become even more important as this chemistry is taken over for new applications. Increasingly, secondary aerosol formation and acid rain chemistry models are being requested for regulatory use. For the most part, these latter models also must include the chemistry leading to ozone and other oxidant formation, because our current understanding suggests that the formation processes of all the secondary pollutants are interlocked through the hydroxyl-hydroperoxy cycle.

Thus, chemical models *are* being extended to new applications (new models are being developed), and our past experience suggests that this might be the case for some time. Successful model guidelines then must independently evaluate every mechanism's fit to our current understanding, yet they must balance this fit with the power to predict new events. This environment of different mechanisms and conflicting predictions is a classic confrontation of science theory and practice; it is the everyday workplace of the scientist. While the scientist engaged in this modeling activity has clear motives for his actions, such actions are often not well understood by the policy maker or regulator who merely wants a tool to conduct his business. He asks, "When is the scientific community going to make the changes that will give me an accurate, valid, and reliable model and what is the timetable for this effort?"

This question naturally arises because one of the things the public expects from science is the capacity to *explain* things and concurrently to enable at least scientists to *predict* what will happen in various combinations of circumstances. When done 'right' in the minds of the public, the explainer not only derives the explanation correctly, but he also gets others to understand and *accept* his derivation. One normally explains a phenomenon by citing the *mechanism* that brought it about or sustains it, where the mechanism itself is understood by subsumption under causal, or at least time-dependent laws.⁸ At present, there is no single, generally accepted chemical mechanism to describe the atmospheric transformation of precursors into the regulated and potentially regulated air pollutants. To many people this is a frustrating situation. After more than a decade and a half of air pollution modeling research, we still cannot predict ozone to most people's satisfaction. This sense of frustration with conflicting and contradictory results felt by atmospheric scientists—as well as the public—has led some administrators at EPA to question the policy of reliance on atmospheric models for predictive use. Some have even doubted the necessity of further attempts to resolve the divisive modeling questions within the scientific community that created them. Throwing good money after bad, they seem to be saying, makes even less sense in science than it does in economics, and so EPA perhaps should not seek a solution to the problems arising from models by funding additional model research.⁹

Dimitriadis³ suggested that resolution of some of these issues could be obtained, at least for EPA's purposes, by having the scientific community critically examine and rank approaches to evaluating the accuracy of gas phase chemical mechanisms. To stimulate discussion, he suggested that there were at least three different approaches to evaluating the accuracy of mechanisms:

- we could compare model predictions against field data;

- we could compare model predictions against smog chamber data;
- we could evaluate the accuracy of basic kinetic data and construct the most detailed mechanism that theory allows—this mechanism must be close to the truth and we should accept it at face value because of conceptual and practical problems of the experimental approaches.

But before such evaluations could be applied to models currently in use or under consideration at EPA, Dimitriadis suggested, evaluations themselves must be "critically examined and ranked." Identifying in this examination all the factors that determine the reliability, comprehensiveness, and precision of each of the three approaches would insure, from EPA's perspective, that the model evaluation process would be amenable to the scientists and useful to the policymakers. We believe that the 'scientific community' of modelers will welcome this examination and evaluation of approaches. In fact, many have said that such an examination is overdue and EPA's effort to provide it are applauded. Examining approaches and determining their influential factors, however, is *not* equivalent to ranking the models themselves. And it is here, in the use of these to-be-established evaluation procedures, that the differing perspectives of the scientist and the policymaker become of critical importance.

From EPA's perspective, the best outcome of this process of examination and ranking would be the development of a "standard mechanism evaluation procedure(s)" and the definition of a "standard data base" needed for performing the evaluation. It should be recognized, however, that such an outcome may not be useful from the scientist's perspective. That is, the scientist is concerned to understand the world and to extend the precision and scope with which it has been ordered. If intense scrutiny of this world reveals pockets of apparent disorder, these challenge him to a new refinement of his observational techniques or to a further articulation of his theories so as to extend the precision and scope of his understanding.¹⁰ Often the theories (and so the mechanisms) that arise in this process must be in complete contradiction to the ones that have gone before them. This is so because the new theory must be more successful precisely where the old theory failed. A new theory may necessitate new experimental evidence and a radically (so to speak) different mechanism. I believe that this is how science progresses. This has certain implications for the concept that there may be a single mechanism evaluation procedure and its concomitant database suitable for all present and future theories and mechanisms. Such a concept ignores the evolutionary character that has brought about the very success it seeks to measure.

If mechanisms, necessarily different, require necessarily different evaluations, then how are we ever to determine rules? In fact, as will be discussed below, some science historians believe that a lack of a standard interpretation or of an agreed reduction to rules will not prevent a 'scientific community' from conducting research that is acceptable to the community. This condition often contributes directly to rapid and wide-ranging progress in our science. This is true, however, only as long as the relevant scientific community accepts *without question* the particular problem-solutions already achieved. Thus rules can become important wherever the underlying guiding principles of the science are successfully challenged or even where the community only perceives that the principles are weakening.

If we have arrived at the point where rules are now required to sort theories and mechanisms, then something fundamental in the science of air pollution modeling is under challenge. This challenge arises, we believe, from two situations: 1) as a result of the inability to resolve the differences among competing explanations of the chemistry of paraffins and olefins; and 2) as a result of the failure to offer acceptable explanations of the aromatics chemistry. These are failures not only of the experimental arts, but also of theoretical practice. For these reasons, Dimitriadis was certain that achieving his goal would "hinge upon the resolution of complex scientific issues." He said, "This is the normal result of the fact that evaluation of chemical mechanisms is fraught with unresolved issues, most of which, in the lack of convincing experimental evidence, have acquired a subjective, even philosophical character." This, he suggested, is likely to lead to an unavoidable clash between those believing in the unchallengeable validity of real world data and thus for whom experimental evaluation is the only option available, and those believing that the real world is hopelessly complex and thus advocate that theory shall be used as "gospel." As we shall see, however, the nature of the problem is not quite this black and white.

I agreed to assist Dimitriadis by describing and discussing various approaches to model evaluation. Upon reflection on the issues, however, I realized that much of what we want to discuss amounts to a "meta-model" of chemical mechanism modeling, that is, "What is it that modelers do when they develop and then test a model?"

To begin this discussion though, it is necessary to agree on the definitions of terms we all use in developing, describing, and testing our theories and mechanisms. For the substance of this paper let us agree on these terms:^{11,12,13}

evidence: empirical generalizations, restricted in their application, arrived at through observation, and intended to prove or provide grounds for belief; easily refuted with other evidence. Example: the sun rising.

theory: views about the ultimate structure and lawlike qualities of the world. They make some claim for permanence, and do not admit simultaneous consideration of other proposals; that is, they are mutually exclusive. Example: the theory of evolution.

hypothesis: tentative candidates for theory; something not proved but taken for granted for the purpose of argument or inference. Example: electrons are in orbits.

auxiliary statements: simplifying assumptions made to allow application of a theory to a given situation. Example: To apply Newton's theory of universal gravitation we might assume for the purposes of calculation: no bodies exist except the sun and the earth; these exist in a hard vacuum; they are subject to no forces except mutually induced gravitational forces. From the *conjunction* of the theory and these auxiliary statements we can predict Kepler's observations.

model: a set of assumptions about some object or system drawn from existing theory, hypothesis, auxiliary statements, and experimental evidence. Though a model may be associated entirely with a single theory, it makes no claim for isolated permanence; that is, it presents itself as one set of descriptive simplified assumptions in a universe of many such sets. Example: Bohr's model of the atom.

mechanism: That inner structure or composition attributed by a model to the object or system it describes, intended to explain various properties exhibited by that object or system. Example: the mechanism attributed by the Bohr model to the hydrogen atom to explain radiation of discrete wavelengths emitted when hydrogen is excited.

truth: correspondence with experience, facts, or reality (whatever these may be). It need not be a statement about true-belief or about permanence, and it is not to be confused with a) coherence theory of truth in which consistency equals truth; b) instrumentalist theory of truth in which usefulness equals truth; or c) evidence theory of truth in which "know to be true" is taken for "true."

facts: the assertion of something as existing or done; (facts are sometimes things).

I will present, for discussion purposes, an abbreviated history of how I understand the process by which mechanisms have been “developed and verified” over the last 15 years by citing books, journal papers, progress reports, EPA reports, conversation, and even personal experience. From this description of practice, I hope that we might obtain insight into the essential underlying assumptions, methodology, and philosophy of our field, and that we might find a commonality of approach and discover particularly revealing examples of practice. Some philosophers of science believe that scientists work from patterns of behavior (mental constructs) “acquired through education and through subsequent exposure to the literature often without quite knowing or even needing to know what characteristics have given their methodologies and approaches acceptable status in the community that they practice in. If they have learned such abstractions at all, they show it mainly through their ability to do successful research.”¹⁰ This is not to say that there are specific “rules of the game,” but it is meant to imply that there are in every field a set of recurrent and quasi-standard illustrations of various theories in their conceptual, observational, and instrumental applications. By studying these and by practicing with them, the members of the corresponding community learn their trade.

Because all modelers whom I know consider themselves to be scientists, this “meta-model of modeling” must bear some kinship to current philosophies of science. Therefore it will be useful to look first at work in this area for insight and guidance. The application of these more general descriptions of scientific activity can provide a framework—and provide us with a ready-made vocabulary—to organize our thoughts on the *process* of mechanism model development and evaluation rather than on the *content* of the modeling work. I am hopeful that in turn we can use these outlines from other sciences and from the philosophical reflections on them to advance both the content and practice in our own field. This goal is admittedly beyond drawing up guidelines for model evaluation, but I hope to show that questions about the nature of modeling and testing strike to something of core importance in all of the sciences. These are especially important in our field where models are “more than useful representations of real-world events” and are “tools for regulation.”

*These are the days of miracle and wonder,
and don't cry, baby, don't cry. Don't Cry.
Paul Simon
Boy in the Bubble*

Nature of Science

Textbook Science

Prior to about 1960, a textbook image of science involved some combination of the following points (taken whole from Hacking's book *Scientific Revolutions*¹⁴):

1. *Realism*. Science is an attempt to find out about one real world. Truths about the world are true regardless of what people think, and there is a unique best description of any chosen aspect of the world.
2. *Demarcation*. There is a pretty sharp distinction between scientific theories and other kinds of beliefs.
3. *Cumulative*. Although false starts are common enough, science by and large builds on what is already known.
4. *Observation-theory distinction*. There is a fairly sharp contrast between reports of observations and statements of theory.
5. *Foundations*. Observation and experiment provide the foundations for and justifications of hypotheses and theories.
6. Theories have a *deductive structure* and tests of theories proceed by deducing observation-reports from theoretical postulates.
7. Scientific concepts are rather *precise*, and the terms used in science have fixed meanings.
8. *The unity of science*. There should be just one science about the one real world. Less profound sciences are reducible to more profound ones. Psychology is reducible to biology, biology to chemistry, chemistry to physics.

A Kuhnian View

In 1963, Thomas Kuhn, trained as a theoretical physicist, but later a professor of History and Philosophy of Science at MIT, published *The Structure of Scientific Revolutions*¹⁰ in which he presented an alternative picture of science. This work has been described as one of the three essential works in the field and has itself revolutionized current philosophical thought about the nature of science. Again from Hacking's book, I quote his summary of Kuhn's alternative view of science:

- A. *Normal science and revolution.* Once a specific science has been individuated at all, it characteristically passes through a sequence of *normal science—crisis—revolution—new normal science*. Normal science is chiefly puzzle-solving activity, in which research workers try both to extend successful techniques, and to remove problems that exist in some established body of knowledge. Normal science is conservative, and its researchers are praised for doing more the same, better. But from time to time anomalies in some branch of knowledge get out of hand, and there seems no way to cope with them. This is a crisis. Only a complete rethinking of the material will suffice, and this produces revolution.
- B. *Paradigms.* A normal science is characterized by a 'paradigm' [(PAIR-uh-dime), a pattern]. There is the *paradigm-as-achievement*. This is the accepted way of solving a problem which then serves as a model for future workers. Then there is the *paradigm-as-set-of-shared-values*. This means the methods, standards, and generalizations shared by those trained to carry on the work that models itself on the paradigm-as-achievement. The social unit that transmits both kinds of paradigm may be a small group of perhaps one hundred or so scientists who write or telephone each other, compose the textbooks, referee papers, and above all discriminate among problems that are posed for solution.
- C. *Crisis.* The shift from one paradigm to another through revolution does not occur because the new paradigm answers old questions better. Nor does it occur because there is better evidence for the theories associated with the new paradigm than for the theories found in the old paradigm. It occurs because the old discipline is increasingly unable to solve pressing anomalies. Revolution occurs because new achievements present new ways of looking at things, and then in turn create new problems for people to get on with. Often the old problems are shelved or forgotten.

- D. *Incommensurability*. Successive bodies of knowledge, with different paradigms, may become very difficult to compare. Workers in a post-revolutionary period of new normal science may be unable even to express what the earlier science was about. Since successive stages of a science may address different problems, there may be no common measure of their success—they may be incommensurable.
- E. *Noncumulative science*. Science is not strictly cumulative because paradigms—in both senses of the word—determine what kind of questions and answers are in order. With a new paradigm old answers may cease to be important and may even become unintelligible.
- F. *Gestalt switch*. ‘Catching on’ to a new paradigm is a possibly sudden transition to a new way of looking at some aspect of the world. A paradigm and its associated theory provide different ways of ‘seeing the world.’

The contrast between the image presented in points 1–7 and Kuhn’s points A–F is not so much in specifics as in a different concept of the relation between knowledge and its source. Kuhn said,

Traditional discussion of scientific method have sought a set of rules that would permit any *individual* who followed them to produce sound knowledge. I have tried to insist, instead, that, though science is practiced by individuals, scientific knowledge is intrinsically a *group* product and that neither its peculiar efficacy nor the manner in which it develops will be understood without reference to the special nature of the groups that produce it.

Thus Kuhn’s terms deal more directly with the scientist than their science, with the theorists as a group than with their stated theories. He has examined science and its community. He says, for example,

The hypotheses of individuals are tested, the commitments shared by his group being presupposed; group commitments, on the other hand, are not tested, and the process by which they are displaced differs drastically from that involved in the evaluation of hypotheses.

Reflecting on his book 10 years later, Kuhn said,¹⁵

The term “paradigm” enters in close proximity, both physical and logical, to the phrase “scientific community.” A paradigm is what the members of a scientific community, and they alone, share. Conversely, it is their possession of a common paradigm that constitutes a scientific community of a group of otherwise disparate men.

One thing that binds the members of any scientific community together and simultaneously differentiates them from the members of other apparently similar groups is their possession of a common language or special dialect. These essays suggest that in learning such a language, as they must to participate in their community’s work, new members acquire a set of cognitive commitments that are not, in

principle, fully analyzable within that language itself. Such commitments are a consequence of the ways in which the terms, phrases, and sentences of the language are applied to nature, and it is its relevance to the language-nature link that makes the original narrower sense of "paradigm" so important.

But how is this community formed and maintained? Hacking said,

The paradigm-as-achievement is commonly taught, not by giving axioms and making deductions, but by giving examples of solved problems and then using exercises in the textbook to get the apprentice to catch on to the method of problem solution.

Often this early 'catching on' results in the first of many intellectual Gestalt switches that a scientist experiences in his career and prepares him to expect such events throughout his work. Many of you may well recognize the form of your own training in science in Kuhn's paradigm-as-achievement definition. For me, Kuhn's explanation of the history of science education is striking valid.

But mere identity of a community's model and the world of events is not the aim of science. Kuhn described the aim of science as "to invent theories that *explain* observed phenomena and that do so in terms of *real* objects, whatever the latter phrase may mean."¹⁵ Kuhn therefore is a realist. For him, however, it is

a reality for which we construct different representations. Realism is compatible with incommensurability, for representations arising from attempts to answer different problems need not mesh well with each other—perhaps the world is too complicated for us to get one comprehensive theory. Even if our theories are plural and incommensurable, we could still think of them aiming at different aspects of one totality.

For this reason, Kuhn has been described as a "constructionist," one who constructs many realities to account for many views. Hacking says, "... the spirit of Kuhn's approach runs counter to the unity of science. There is a plurality rather than a unity in representations of the world, and successive representations address different problems which need have very little common subject matter." But a careful reading of the passage cited above reveals that Kuhn has already dealt with this problem and so refuted the potentially disastrous consequences of scientific plurality; what seems plural, he says, are only aspects of one totality. We are comfortable with Kuhn's delineation of aspects of one reality if only because it means that the last 15 years of air pollution modeling work (in which we have taken many different assumptions and arrived at contradictory conclusions) have not been describing different realities. If Hacking were right and Kuhn's view necessitated construction of many realities, then scientific conclusion could never be made. So if different mechanisms are describing only different aspects of one reality, then even completely independent evaluation must say something about the mechanism, the theory that generated it, and the fit of both to reality.

Note that I do not yet accord any weight to *comparison* of evaluations, only state that independent evaluations would be valid measures of all mechanisms' fit to the one reality.

Because, as we have seen above, a mechanism often reflects only the new theory (hypotheses and auxiliary statements) that was directly responsible for its generation, and not much of the more commonly assumed theory, it is necessary to examine "theories" before looking at approaches to evaluation. Kuhn says that he does "not believe that there are rules for inducing correct theories from facts." Instead he views theories

as imaginative posits, invented in one piece for application to nature. And though we point out that such posits can and usually do at last encounter puzzles they cannot solve, we also recognize that those troublesome confrontations rarely occur for some time after theory has been both invented and accepted.

In other words, for Kuhn, theories have a natural obsolescence. They cannot be expected or forced to survive past the point where they contribute in a meaningful way to the puzzles posed by the paradigm that directs any particular scientific venture. It is important to note, however, that Kuhn's definition requires at least some time for reflection on the theory and that during this time the developing theory is not directly challenged by troublesome confrontations. This aspect of slow theory-growth and protection by the scientists who brought it about has been explored in great depth by another philosopher.¹⁶

For Kuhn, there is an "intimate and inevitable entanglement of scientific observation with scientific theory."¹⁵ Continuing from Hacking's book:

Kuhn rejects a sharp observation-theory distinction because the things that we notice, and the ways we see or at least describe them, are in large part determined by our models and problems. There is no 'timeless' way in which observations support or provide foundations for theory. The relations between observation and hypothesis may differ in successive paradigms. Hence there is no pure logic of evidence or even of testing hypotheses, for each paradigm, in its own day, helps fix what counts as evidence or test.

Just as Kuhn's definition of incommensurability and noncumulative science helped to justify my claim for independent mechanism evaluation, so here the recognition that there is no observation-theory distinction, leads to significant conclusions. In this case, it is that the process of mechanism evaluation—and so comparison of the theory with facts—is not static, but must evolve along with the scientific content of air pollution chemistry.

There are several issues tied to this notion of evolutionary evaluation the conclusions to which have been directing both the theory and practice of modeling

working since 1975. Here I will describe some of these issues—puzzle-solving, explicit standards, testing, and fit—and point out the effects of their conclusions. Later (in Chapters 3–9) I will introduce specific instances where decisions about testing and fit have considerably changed how modelers model and how they treat chamber data. In Chapter 10, I will talk about how these decisions and their consequences have affected progress in our field. Lastly, I will draw together these assumptions, definitions, and conclusions into suggestions for model evaluation I hope will be the outcome of my work here.

The business of science, Kuhn says, is puzzle-solving. Though this may sound demeaning, Kuhn certainly intends no offense; he says that this process has required the highest scientific talents. In fact, he chooses the metaphor because of the striking similarity between solving jigsaw puzzles and the search for explanations of phenomena and the ways to harness nature. Kuhn contends that this puzzle-solving activity is a vital part of science so long as the puzzles are interesting and the solutions are accepted by the community. In their day-to-day work, scientists do not try to disprove the theories that direct their research, nor do they seek unexpected and unpredicted results. So, according to Kuhn, although science proceeds by the extraordinary episodes of shifts of professional commitments, paradigm revolutions, it cannot advance directly from one revolution to the next. In fact, since Kuhn's structure for science denies linear progression completely, it would not admit progress *from* any paradigm *to* any other in any direct manner at all. That we have no new paradigm 15 years after the proposition of HO'-attack does not mean that air pollution science has failed to progress in a Kuhnian sense. Rather, the puzzle-solving done in those 15 years to expand and ratify, or modify the theories and auxiliary statements of HO'-attack has been the success promised by that paradigm. And it *was* that promise, Kuhn would say and I would agree after looking at the actual documentation of those early years, that led us to accept the HO'-attack theory in the first place.

While others have merely noted an asymmetry with regard to testing theories, Kuhn takes a special view of the process. The asymmetry can be stated as: it is a logical truism that a scientific theory cannot be shown to apply successfully to all its possible instances, but it can be shown to be unsuccessful in particular applications. Some have used this concept to suggest that scientists directly test theories and reject them on the basis of experiment. Karl Popper, for example, in *Logic of Scientific Discovery* written in 1959, said

A scientist, whether theorist or experimenter, puts forward statements, or systems of statements, and tests them step by step. In the field of the empirical sciences, more

particularly, he constructs hypotheses, or systems of theories, and tests them against experience by observation and experiment.

Kuhn says Popper is historically mistaken, and that his mistake "misses just that characteristic of scientific practice which most nearly distinguishes the sciences from other creative pursuits." Kuhn says

There is one sort of "statement" or "hypothesis" that scientists do repeatedly subject to systematic test. I have in mind statements of an individual's best guesses about the proper way to connect his own research problem with the corpus of accepted scientific knowledge. ... He may conjecture that a newly discovered spectral pattern is to be understood as an effect of nuclear spin. The next steps in his research are intended to try out the conjecture or hypothesis. If it passes enough or stringent enough tests, the scientist has made a discovery or has at least resolved the puzzle he had been set. If not, he must either abandon the puzzle entirely or attempt to solve it with the aid of some other hypothesis. ... Tests of this sort are a standard component of what I have called normal science. In no usual sense, however, are such tests directed to current theory. On the contrary, the scientist must *premise* current theory as the rules of his game. His object is to solve a puzzle, preferably one at which others have failed, and current theory is required to define that puzzle and to guarantee that, given sufficient brilliance, it can be solved. ... the practitioner must often test the conjectural puzzle solution that his ingenuity suggests. But only his personal conjecture is tested. If it fails the test, only his own ability, not the corpus of current science is impugned. ... in the final analysis it is the individual scientist rather than current theory which is tested.

There comes a time, of course, when a paradigm's potential for suggesting new and interesting puzzles has been fully actualized. At times of such paradigmatic crisis, Kuhn claims, scientists are reluctant to give up the theories and their auxiliary support statements which have allowed them to make progress. In fact, it is my experience that scientists often work hardest to protect a theory (or even a paradigm) that is devoid of potential. In such cases and even throughout the practice of normal science when theories and paradigms have reached their obsolescence, scientists must confront the anomalies left unexplained (even maybe unpredicted) by the old for only the hope of progress with the new. Kuhn is quite clear, though, that such change is never brought about by the direct refutation of theory by physical evidence. He even extends this argument to apply to the practice of all normal science. Thus, Kuhn denies that scientists practice "falsification" of theory—showing an instance wherein the theory fails in an attempted application, and hence to argue a reason to compel assent from any member of the relevant professional community as to the theory's falseness. He says,

Where a whole theory ... is at stake, arguments are seldom so apodictic. All experiments can be challenged, either for their relevance or their accuracy. All theories can be modified by a variety of ad hoc adjustments without ceasing to be, in their main lines, the same theories. It is important, furthermore, that this should be so, for it

is often by challenging observations or adjusting theories that scientific knowledge grows.

Kuhn explicitly realizes that a Gestalt switch is needed to accept a new theory, but is difficult for scientists to make. He suggests that scientists, "though they may begin to lose faith and then to consider alternatives, do not renounce the paradigm that has led them into crisis." They do not treat anomalies as counter-instances. If an anomaly cannot be assimilated within the existing theory, then the scientist deletes the problem (discounts its existence). And the shift is even more difficult to make if the theory under attack is actually the paradigm, for once a theory has achieved the status of a paradigm, it is declared invalid only if an alternative candidate is available to take its place. Kuhn declares that "No process yet disclosed by the historical study of scientific development at all resembles the methodological stereotype of falsification by direct comparison with nature." Instead, Kuhn asserts that

... the act of judgment that leads scientists to reject a previously accepted theory is always based upon more than a comparison of that theory with the world. The decision to reject one paradigm is always simultaneously the decision to accept another and the judgment leading to the decision always involves comparison of both paradigms with nature and with each other."

Truths turn into dogmas
the moment they are disputed.
—Chesterton

A Psychological Interpretation

Some philosophers and scientists strenuously object to Kuhn's description of scientific progress by saying that it is far too psychological—that it gives the individual more power and importance than the theory. For them Kuhn has said that science is *only temporary agreement within the community of scientists*. In fact, Kuhn's image of how we practice science is very similar to descriptions arising in sociological fields, for example, one "meta-model" for human communication and interaction is neuro-linguistic programming (NLP).¹⁷ This model was developed to explain how some therapists are more successful than others at getting people to change. It was described by Gregory Bateson as "making explicit the syntax of how people avoid change." According to this system, we live in one real world. We do not function directly in this world, however, but through a series of "maps" of the world that we develop from our experiences (of course, not all maps are verbal). Maps, like all models, are created by three general process:

- *Generalization*. To simplify our representation, we treat objects that are different in the real world as if they were the same in our map.
- *Deletion*. To simplify our representation further, we omit from our map objects or details in the real world that we think are not important or we do not have a representation in our map of objects in the real world of which we have no awareness.
- *Distortion*. To fit some facts into our map, we modify the representation of real world objects so as to be consistent with our map. Something must be distorted to make any map.

A possible consequence of this mapping is that one confuses a useful map with the territory it represents and acts as if his map were the world. Successful maps are more likely to bring about this mistake precisely because they are more useful to the map maker. A person with a highly distorted map or otherwise impoverished representation (usually formed under stress and without adequate adult information) often has difficulty coping with the real world; he never sees the options available to him, and so sometimes acts in what seems to others as inappropriate ways.

For example, a person complains that he is not worthy of love and that his wife does not love him. Since he was not loved in his childhood and since he is not loved now, it must be that he is not worthy of love (generalization) and he can prove that because his wife does not love him (the map is real). When it is

pointed out to him that frequently his wife has been observed making loving remarks to him, he replies that he does not recall any such statements (deletion). When they are pointed out in real-time, he then distorts what he hears by saying, "But she is not sincere—she only wants something."

Maps more free from distortion, those sufficiently rich in detail, provide options precisely in those places where others have failed to include all the facts. Such facts, present but not accounted for by some, allow the person with a rich map to choose his response more appropriately, and so to determine the outcome of an event in a way that is more pleasing to him.

A person having difficult life-experiences can often benefit by modifying their maps. It is rare that this person, without careful training, would be capable of "stepping outside" of his map, and thus he usually requires the help of a therapist. According to the NLP model, a successful therapist first deduces the structure of a person's map through his actions and statements (a form of Kuhn's puzzle-solving). Next the therapist creates situations that cause the person to confront the flaws in their maps (non-functional generalizations, deletions, and distortions) and eventually to resolve the discrepancies. The map adjustment is usually accomplished by a 'Gestalt switch' that suddenly allows the person to see solutions to his problems. Successful therapy thus results in a richer map and more options for coping in the world, and therefore, to have more successful outcomes in real-life situations.

Kuhn's concept of the paradigm can be easily and usefully translated into these terms. Kuhn's map is scientific realism, the idea that the world is unitary and explainable. Progress is made within a paradigm, Kuhn would say, in one sense by the scientist who generalizes information about hypotheses and observations; this is problem solving. In a second sense, the same researcher may also group individual representations to generalize more precisely about their behavior or content: this is the basis for prediction. These together constitute the business of normal science.

Similarly, normal science requires the scientist to delete information and narrow his field if he is to make progress at all. Kuhn says,

Though science clearly grows in depth, it may not grow in breath as well. If it does so, that breath is manifest mainly in the proliferation of scientific specialties, not in the scope of any single specialty.

Indeed, Kuhn would require that the working scientist delete from his map all potentially interfering information. He calls narrowmindedness a "characteristic" of

scientific communities; it is the concern with detail that ultimately allows comparison of theory and observation on the level necessary to solve the puzzles of normal science.

The third map-generating process, distortion, is yet more powerful and more interesting in two ways: 1) as an essential, directing behavior that creates the map; that is, some things must be distorted to make any map; and 2) as a functional result of the misapplication of the first two processes as when overgeneralization or excessive deletion results in a map untrue to its territory.

In the first sense, distortion allows the scientist to explain arrestingly novel events by distorting their place (significance) on his map (in his paradigm); thus allowing him to solve puzzles and thereby strengthen the paradigm (map). In the second, destructive sense, distortion prevents a scientist from making the break to a new paradigm even when the existing one is repeatedly bettered at explanation and prediction. He has generated his map in response to input taken directly or indirectly from the actual world of events. This map has been successful at the puzzle-solving science has required and this success has re-affirmed his commitment to the map. And, as in therapy, this is the sort of distortion that requires a Gestalt switch for the individual believer to abandon his map. Such distortion is especially disastrous to a practicing scientist in a community undergoing crisis because at that point he is in danger of mistaking his map, his distorted map, for the world he tries to explain. The new map, the old believer claims, cannot be closer to the world of events because his existing map *is* that world. And he believes in that identity because his map has been validated through repeated success in explanation and prediction. Thus, in the terms of our comparison, Kuhn's scientific realism has become in the distorted map both scientific *and* physical reality.

The Rest

Now that we have agreed on terms and laid the philosophical groundwork for a discussion of what it means to theorize, test, and fit reality we can next show where these terms and constructs fit in a history of air pollution modeling and evaluation. It was necessary to define these terms and ideas because they are ones that we all use everyday. Because of that familiarity they have implicit meanings for each reader which must be reconciled if we are to have meaningful conversation. Similarly, it is necessary now to recount a history of mechanism development using the terms and concepts described so that the philosophy I have described will have a real connection to our own science. Lastly, it will be necessary to draw out of this interpretation of our past just how we might proceed.

*It is the customary fate of new truths
to begin as heresies
and end as superstitions.
—T.H. Huxley*

The New Paradigm

To understand the paradigm-as-achievement we must understand what has been achieved. Kuhn said that science historians (or scientists themselves) can agree in their identification of a paradigm without agreeing on, or even attempting to produce a full interpretation or rationalization of it. What follows must be the product of my own generalization, deletion, and distortion. Gerald Weinberg, writing¹⁸ about his profession, system analysis, said,

“If we study history even for a short while—long enough to read more than one source—we learn perhaps the most important lesson:

There is no one truth about what happened.”

Some of you, no doubt, will have a different specific image of what happened; other works, not mentioned here, will have more relevance for you. Nevertheless, I think the historical examples I have selected illustrate many of the essential features of the science of model development, therefore, they are useful for our purposes.

The Beginning of Explanation

Prior to 1970, the science of air pollution chemistry was lacking in major respects; it was in a pre-paradigm stage. In the period 1961–1970, there was no generally accepted view about the exact causes of the chemical transformations that produced secondary air pollutants. The primary source of theoretical information was Phillip Leighton’s 1961 book *Photochemistry of Air Pollution*.¹⁹ Many researchers have commented on how surprised they were to find that Leighton had discussed some topic that later became central to the paradigm. Kuhn says,

In the absence of a paradigm or some candidate for paradigm, all of the facts that could possibly pertain to the development of a given science are likely to seem

equally relevant. As a result, early fact-gathering is a far more nearly random activity than the one that subsequent scientific development makes familiar. Furthermore, in the absence of a reason for seeking some particular form of more recondite information, early fact-gathering is usually restricted to the wealth of data that lie ready to hand. The resulting pool of facts contains those accessible to casual observation and experiment together with some of the more esoteric data retrievable from established crafts like medicine. Because of this ... technology has often played a vital role in the emergence of new sciences.

Some of this pool of facts included Calvert and Pitt's book *Photochemistry* and the smog chamber work of Paul Altshuler and his colleagues. Atkinson *et al.* suggested²⁰ that, in the case of air pollution chemistry, the technological source of information was combustion chemistry. Chain reaction theory was developed in 1934 and the hydroxyl radical (HO^\cdot) was proposed as an important intermediate in high temperature oxidation systems.

In the air pollution field, a major anomaly arose at the juncture of theory and observation when it was noted that the high rates of conversion and the large O_3 production demonstrated in smog chambers could not be accounted for using the prevailing chemical theories. Although the air chemistry textbooks and papers were full of chemical reactions, rate constants were not commonly given and reactions rarely appeared in clusters of more than three or four. Only a few air pollution scientists could successfully argue the consequences of a chemical reaction sequence that was longer than three steps. Their arguments sometimes were so confusing that the listeners were totally lost and unable to engage in discussion.

Along with a large amount of other potential radical chemistry, Leighton had described HO^\cdot and had postulated several reactions: "Hydroxyl radicals are produced by the photodissociation of hydrogen peroxide and possibly by the photodissociations of nitrous and nitric acids." Further, he even had suggested the reaction

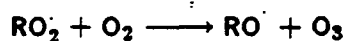


in the middle of a list of four other potential HO^\cdot source reactions, saying, however, "Unfortunately, little may be said as to the rate at which hydroxyl radicals are produced in polluted air." Leighton had described what we now know to be key reaction in alkoxyl radical reactions,



as being "purely speculative." He had concluded that "... certainly there is no evidence that the reactions of hydroxyl radicals with nitric oxide or nitrogen dioxide are of any importance in air." Thus the HO· radical was only one of several ideas, part theory, part observation competing with each other in an effort to explain early smog chamber results. As such they had the suggestive force of a pre-paradigm theory or observation, but not the coercive power of a paradigm, already accepted by many. Nevertheless, for a young researcher entering the field, this was hardly a cause to investigate HO· reactions. As was said above, "the things we notice, and the ways we see or at least describe them, are in a large part determined by our models."

As another example, the first photochemical smog mechanism for use in a computer program that I recall seeing was Lowell Wayne's 15 or 16-step mechanism which appeared in one of the blue covered Air Pollution Foundation Reports about 1967. This mechanism did not use HO·, and to account for the observed rates, Wayne assumed, as others had, that organic radicals reacted with oxygen to make ozone,



Although this approach allowed Wayne to predict O₃ production, it was met with little acceptance, because the reaction must be endothermic,¹⁹ and "... many of the proposed steps are not elementary reactions, and they have little or no physical meaning."²¹

Using Kuhn's terms, there was essentially a "revolution" during 1969-70—"a new way of seeing the world"—and it has led to today's normal science in chemical mechanism development. Greiner²² made the first determinations of the rate constants for HO· with alkanes in 1967. Based on his results, he postulated that these reactions would be important in photochemical smog chemistry. Two groups soon after put forth the supposition that it was a chain process involving HO·-HO₂ that was responsible for the "excess" rate of hydrocarbon consumption and NO oxidation. In a paper presented at "Chemical Reactions in Urban Atmospheres" in Michigan in 1969, Heicklen, Westberg, and Cohen²³ "suggested that HO·, postulated as intermediates in the reaction of hydrocarbons with NO, serve as chain carriers to oxidize NO to NO₂ in the presence of CO." At the same meeting in a discussion, Bernie Weinstock (who, according to Joe Bufalini, was a combustion chemist at Ford Motors), described how he calculated the chain length in propylene/NO_x smog chamber experiments performed by Altshuller. Less than a year later, Steadman, Morris, Daby, Niki, and Weinstock described "The Role of OH

radicals in Photochemical Smog Reactions." Others, too, were rushing to print with ideas based on the HO·-chain concept (see the next chapter, for example). The rapid conversion to the "HO·-paradigm" suggests that Gestalt switches were made quickly by many scientists practicing in somewhat different fields. And the faith required to make such a switch was great because no one had ever determined that HO· radicals existed in the ambient air.

The mere recognition that an HO·-chain was involved, however, was not sufficient to serve as a paradigm. Demerjian, Kerr, and Calvert²¹ say, "The first attempts to describe the individual chemical steps important for the irradiation of mixtures containing propylene, NO, and NO₂ in air were reported by Westberg and Cohen (1969) and Wayne and Ernest (1969)." The former, consisting of 71 reactions, was an Aerospace Corp. publication, and the latter, consisting of 40 elementary reactions, was presented at an APCA meeting. Demerjian *et al.* go on to say, "However some of the 'elementary' reactions ... involved such extreme structural rearrangement or unfavorable energy requirements for reactants to form products, that they can not be important in real systems." Therefore their solutions to the puzzle were not acceptable to the larger community and they were not candidates for a paradigm.

Niki, Daby, and Weinstock

By 1970-71, the HO·-HO₂ chain concept was sufficiently developed that Niki, Daby, and Weinstock had produced a paper, "Mechanisms of Smog Reactions," that solved the puzzle in an acceptable manner. This paper, therefore, received wide circulation²⁴ before eventually being published in 1972.²⁵

Over 150 elementary reactions had been examined and 60 elementary reactions were selected for the mechanism of about 40 species. The mechanism was generated under the HO·-chain theory. But the theory at the heart of this paradigm needed auxiliary assumptions for the model to explain adequately and predict correctly. So the mechanism proposed was HO· attack on the propylene double bond, leading to the formation of an hydroxy-alkyl radical which subsequently reacted with oxygen to produce an hydroxy-alkyl peroxy radical. This radical reacted with NO producing NO₂, HO₂, and aldehydes. Photolysis of the aldehydes supplied a continuous source of new HO₂ radicals which in turn produced HO· radicals. Hydroxyl radicals were lost continuously by reaction with NO₂ to form nitric acid. Other radical-radical reactions also terminated the chain. The only source of O₃ in this mechanism was the photolysis of NO₂ followed by the reaction of O +

O₂ to make O₃. Each reaction had an assigned rate constant supported by multiple literature citations. The work included 123 references. That these rate constants values were so prevalent in the literature is yet another proof that the field was in a pre-paradigm stage.

The model was applied to a propylene/NO_x smog chamber experiment and could account for the principle features of propylene chemistry. Further, the model was used to illustrate the effects of changes in model inputs and experimental conditions and a mechanistic interpretation was given for each outcome. This fulfills the explanation function of a paradigm. These examples of the model in practice helped to establish the basic acceptance of the mechanism concepts. "Theory" had been used to construct a complex description that not only exhibited the right kinds of response, it was consistent with a large body of older scientific information, and it had the "feel" of a good solution (i.e., there was a certain sense of congruity or recognition). Thus the model as theory might also predict accurately. Its predictions only had to resemble the smog chamber results, they did not have to fit it perfectly. Further, problems that could attract a range of researchers were clearly obvious from the work.

This work exemplified the new paradigm-as-achievement. Kuhn says,

We must recognize how very limited in both scope and precision a paradigm can be at the time of its first appearance. Paradigms gain their status because they are more successful than their competitors in solving a few problems that the group of practitioners has come to recognize as acute. To be more successful is not, however, to be either completely successful with a single problem, or notably successful with any large number. The success of a paradigm... is, at the start, largely a promise of success discoverable in selected and still incomplete examples.

Putnam²⁶ says of the new paradigm, "It is important that the application—say, a successful explanation of some fact, or a successful and novel prediction—be *striking*; what this means is that the success is sufficiently impressive that scientists—especially young scientists choosing a career—are led to try to emulate that success by seeking further explanations, predictions, or whatever on the same model."

Prior to the use of such computer models, theoretical air pollution chemistry was mainly descriptive. After 1971, we had a clear example of how it could be quantitative. We had in the paradigm of HO·-chain and its attendant auxiliary hypotheses, coupled with the solution possibilities of the computer, the potential to explain, with increasing precision, exactly how atmospheric chemical processes happen.

As a methodology, computer simulation could make clear the consequences of a tremendous series of "what if" hypothesis in a very short time. It aided in the initial articulation of the theory. Most of this early exercise consisted of exploring the theoretical consequences of basic choices in the theoretically possible chemistry. It was not concerned with direct comparison of model prediction with observations (*e.g.*, smog chamber data) so early emphasis was on explanation and not on testing. In fact, much of the chamber data were not in an appropriate form to compare with the computer predictions. Older descriptive concepts lost meaning (*e.g.*, tabulations of relative reactivity of hydrocarbons for making ozone). New information was in demand: rate constants, reaction pathways, and chamber operating conditions. The paradigm suggested new tools were needed (new chambers), new instruments, and it offered direct guidance on what experiments should be the most revealing. It also suggested what new 'facts' would be critical (rate constants for specific processes). That is, theory began to drive experiment and vice versa.

The Practice of Normal Science.

Kuhn says, "Normal science consists in the actualization of that promise [of success offered by the paradigm], an actualization achieved by extending the knowledge of those facts that the paradigm displays as particularly revealing, by increasing the extent of the match between those facts and the paradigm's predictions, and by further articulation of the paradigm itself."

In other words, normal scientific research is directed to the articulation of those phenomena and theories that the paradigm already supplies and is not concerned with calling forth new sorts of phenomena nor is it normally concerned with inventing new theories. By focusing attention upon a small area, the paradigm forces scientists, Kuhn says, "to investigate some part of nature in a detail and depth that would otherwise be impossible." This attention leads to increasingly finer adjustments and additions to the theory and its auxiliary assumptions. Eventually, it leads to facts that cannot be explained by the theories—*anomalies*. A sufficient number of anomalies, coupled with an alternative explanation, leads to crisis and possibly to revolution if a new theory is selected to form the basis of a new paradigm. So far the HO⁺-chain paradigm has served air chemistry well. Intense scrutiny has been directed to questions of both the theory and practice. Much detailed work has been in the area of testing the auxiliary hypotheses generated under the HO⁺-chain paradigm.

Kuhn says that during the time the paradigm is successful the problems of normal science occur in three classes—determination of significant fact, matching of facts with theory, and articulation of theory. Kuhn claims that these exhaust the literature of normal science, both empirical and theoretical. Because of the complex interaction of theory and observation, it is difficult to separate these activities. Descriptions of the classes follow:

- The *first class* is the production of facts (experimental work) that the paradigm has shown to be particularly revealing of the nature of things. By using these facts in solving problems, the paradigm has made them worth determining both with more precision and in more situations (*e.g.*, measuring reaction rate constants). A small part of theoretical work consists of using theory to predict factual information of intrinsic value. This, however, is often viewed as “engineering” work and not of much real interest to scientists (*e.g.*, producing astronomical ephemerides, making predictions of air pollution control requirements). It is rarely published in journals.
- The *second class* of experimental work is that class of factual determinations that can be compared directly with predictions from the paradigm theory. These often demand theoretical and instrumental approximations that severely limit the agreement to be expected (*e.g.*, measuring the ambient air concentration of HO^\cdot ; wall effects in smog chambers). The second type of theoretical work is the manipulation of the theory so as to produce predictions that can be confronted directly with experiment. These predictions may or may not have any intrinsic value. The purpose of these activities is to display a new application of the paradigm or to increase the precision of an application that has already been made. Precision in theory application concerns the type and nature of the simplifying assumptions, testing assumptions of presumptive generalities, or eliminating anything that is likely to restrict the agreement to be expected with actual measurements.
- The *third class* of empirical work is that undertaken to articulate the paradigm theory, resolving some of the residual ambiguities and permitting the solution of problems to which it had previously only drawn attention. This work includes the determination of quantitative laws. Sometimes, in the early applications of the paradigm, this work includes exploration experiments to see how the paradigm fits in some new area of application (*e.g.*, aromatics chemistry). These might be useful elaboration of the various possibilities and distinguishing among them. Theoretical problems of paradigm articulation are often those of re-formulation both to achieve clarification and to allow extension of application (*e.g.*, creation of condensed

mechanisms for use in air shed models). A large share of the theoretical problems of paradigm articulation, however, are accounting for changes arising from the empirical work undertaken to articulate the paradigm (*e.g.*, revising and updating mechanisms with new rate constants, pathways, refining auxiliary assumptions). Often, when working in this class of work, a researcher will use both facts and theory to adjust and modify the paradigm.

Early Normal Science Work

First Class Work

As you are aware, a large amount of first class (in both senses) experimental work occurred in the 1970's. Chief among this work was the determination of the rate constants of HO \cdot with almost every hydrocarbon. In addition, the rate constants of the most critical inorganic reactions were determined repeatedly with increasing precision. How did we know which ones were the most critical? The models, based on the HO \cdot -paradigm told us!

Tools Needed to Advance the Theory

The mathematical expressions arising out of reaction kinetics present significant solution difficulties. No analytically solutions of these complex systems are possible, although some have been tried for simpler systems. Therefore, numerical integration methods are needed. For the kinetics systems arising in air pollution chemistry and which are described as "stiff," most of the numerical methods are unstable—that is, the so-called solution produced by the method, which is only an approximation, rapidly diverges from the true solution and the predicted values may approach infinity. In the early 1970's these problems limited the effectiveness of computer simulations and were surely a major difficulty in the work described in the next chapter.

In 1971, however, William Gear, a computer scientist, described a stable and efficient method and produced a FORTRAN program suitable for the practical numerical solution of large kinetic reactions systems.²⁷ The wide-spread availability (after about 1974) and efficiency of code using the "Gear" technique allowed many air pollution researchers to take up computer modeling. That is, general computer codes were produced that accepted "chemical reactions" and initial species concentrations as input and produced tables of species concentrations verses time as their output. The user of these programs did not have to know anything about how the method worked and the codes were robust enough that

numerical difficulties rarely arose. The researcher was therefore free to think only about the collection of "chemical reactions" that were used with these "turn the crank" solution methods.

These Gear codes are still too demanding of computer resources for use in three-dimensional grid air shed simulations, however, and no single solution approach has dominated these programs. Thus simulations of smog chambers can be very complex and detailed (because they use the Gear codes), while even the simplest three-dimensional grid atmospheric simulation that couples transport with chemistry requires quite large computer resources (memory and time). Because the chemistry in these latter simulation accounts for about 80% of the computer time, for practical applications, simpler chemical representations must be used. This has created a certain kind of problem for the air pollution modeling field: a difference between explain and predict.

*He wasn't the first to explore the region,
but he returned with maps
that could be followed with confidence by others.
—The Mechanical Universe,
PBS*

First Explanation

The air pollution modeling literature, including books, journal articles, and EPA reports can be divided into two broad categories: those works whose primary aim was *explanation* of the transformation process, and those works whose primary aim was *prediction* of atmospheric change. Initially these were products of two different sets of practitioners; later the two goals were integrated into one body of practice, but with two expressions. I will conveniently describe the development and testing of models by using these two categories as a framework in which to present the process.

Demerjian, Kerr, Calvert

The first well-known example of normal science work following the adoption of the new paradigm was "The Mechanism of Photochemical Smog Formation" by Ken Demerjian, Alistair Kerr, and Jack Calvert²¹ which appeared in 1974. This was a large piece of work, totaling 262 pages, and although it did not appear in its complete form until 1974, it had actually been developed over the time period 1970 to 1972 (the manuscript was submitted to the editor in 1972). Alistair Kerr was a summer visitor in 1970 and '71 to Jack Calvert's group at Ohio State where Ken Demerjian was a PhD student. Demerjian's Master's Thesis, completed in 1970, had been "Computer Simulation of Smog Chamber Studies." Because of the publisher's delay in printing the article, and similar to the Niki *et al.* case, many of the active workers in the tropospheric chemistry area were sent preprints of the article. In addition, a series of articles based on the full manuscript appeared in *Environmental Letters* in 1972 and 1973. Probably of more significance in shaping the careers of many workers in the area of chemical modeling, however, was the "School on the Fundamental Chemical Basis of Reactions in the Polluted Atmosphere" which was held at the Battelle Conference Center in Seattle on June

18–29, 1973 in which Jack Calvert and Ken Demerjian played key roles. Xeroxed copies of the manuscript were essentially used as a ‘textbook’ and a computer link to the Battelle computers allowed Ken Demerjian to offer a ‘hands-on,’ first-time simulation experience for many of the hundred or so attendees. There can be little doubt that this work shaped the early paradigm that was “limited in both scope and precision” into a full example of how to practice science. For this reason, it guided a lot of people who wanted to start modeling—it became a ‘bible.’ The work embodies a whole range of explicit and implicit commitments about how science should be practiced. We can only touch on a few here.

In the introduction Demerjian, Kerr, and Calvert (DKC) said that the purposes of their work were

... to evaluate the various alternative mechanisms and reaction rate constants proposed in view of the best kinetic data in hand today.

... to quantify our own thinking and that of our colleagues.

... to identify some of those potentially important reactions in the mechanism of photochemical smog formation for which there was insufficient basic kinetic data to allow for a realistic judgment of their importance.

... to predict theoretically the effects of atmospheric composition, as yet experimentally unobserved, on the rates of formation of specific secondary products of some interest.

... to stimulate the research activities related to this area of chemistry.

I think it lends support to Kuhn’s description of scientific activity to note that not only do the goals and programs set out by Demerjian *et al.* fit well into what Kuhn has called the classes of normal work, but that the very terms themselves echo Kuhn’s. Truly, as Kuhn has said, it is not necessary for researchers to agree on a paradigm, or even on its existence, to be influenced by it. The relevance of Kuhn’s description is clearly evident in this case at least. Identity of terms insures, I believe, that Kuhn’s careful description about scientific progress and radical change have at least nominal relevance in our specific science. And if they have nominal relevance, then we can profitably look to them to sort out problems of theory, theory-testing, and observational fit. These are the very problems that EPA has raised for this workshop.

Mechanism Development

The *process of mechanism development* was clearly described by Demerjian, Kerr, and Calvert,

We feel that the only real hope of defining the reactions which are important in the real atmospheres lies in the elucidation of the simpler smog chamber experiments.

We will consider first the simplest possible photochemically active polluted atmospheres. Then the effects ... of adding ... reactants common to the real atmosphere will be determined.

We attempt to explain the product formation and reactant disappearance of a particular smog chamber study with respect to a rational series of elementary chemical reactions. Given the initial conditions and the rate constants for the elementary reactions involved, the computer will generate by a numerical solution the reactant and product concentrations versus time curves for the kinetic mechanism proposed.

In recent simulations ... more complete and seemingly accurate reaction rate schemes have been employed. However, as with all such complicated systems for which accurate reaction rate data are not available for many of the reactions, a reasonably large degree of chemical intuition has been employed in the choice of rate constants. In reality many of the chemical reactions which are believed to be involved in photochemical smog have been studied experimentally, and their rate constants must be employed in kinetic matching attempts; they should not be considered as variables which can be adjusted outside of the experimental error limits in the attempted fit by computer. On the other hand there are many reactions of potential importance in these systems for which no experimental data or very limited information exists. In these cases it is necessary to estimate the rate constants 'theoretically,' avoiding any completely arbitrary choice of values. Such estimates must be consistent with the present knowledge of chemical kinetics and thermodynamics.

Although we have made what we believe to be reasonable estimates for the rate constants for these reactions, these estimates may be in considerable error. Reliable values for them are necessary either to confirm our suggestions or to force a reevaluation of the importance of some portions of the smog mechanism outlined here.

... without certain key information among the smog chamber data, it is impossible to develop a reaction scheme which is completely free from adjustable parameters which are varied in an arbitrary fashion to match the kinetic simulation to the experimental data. Simulations which are built on such flexible schemes cannot add significantly to the quantitative scientific understanding of the atmospheric reactions.

... there is often relatively poor reproducibility between chamber product rate studies run by different laboratories, and the results of one individual study may be rather inaccurate and misleading. We have attempted to overcome this problem and to develop a consistent set of chemical reactions related to photochemical smog systems through the treatment of several very different systems studied by a variety of investigators.

Realism

Let us begin our discussion of the philosophy behind DKC's statements by examining the manner in which theories are seen to relate to reality. This topic is a subject of current debate; for example, a conference on scientific realism was held at UNC, Greensboro in 1982. There are two basic philosophical positions (with a great deal of intravariability):

- ▷ *instrumentalism* or *antirealism* posits that the nonobservational portions of even "true" theories are not really assertions about the world, but are devices of some sort for codifying relations between observable variables. This becomes the claim that any theory, however well confirmed and widely accepted in scientific practice, might well be and might eventually be revealed to be false.⁸
- ▷ *realism* posits that physical theories make meaningful assertions about the real world (i.e., the terms in a theory are genuinely referential); they are at least approximately true. The approximate truth of a scientific theory is a sufficient explanation of its predictive success. Theories which are regarded as genuinely explanatory can be used to predict facts other than those for which they were first invented (or they can be modified in a simple way to do this) because the entities that are described by the theories really exist.

In its most common form, realism is the belief that "science makes possible knowledge of the world beyond its accessible, empirical manifestations."²⁸

The antirealist sees major objections to the realist's position:

Whatever continuity may be discerned in the growth of empirical knowledge, theoretical science has been radically discontinuous. Scientific views about the ultimate structure and lawlike organization of the world have frequently been overthrown and replaced by incompatible views. Much of the discarded science was, for an appreciable time, eminently successful by the standards we employ in assessing current science. The inference seems inescapable that the evidence available to support current science is by nature unreliable and systematically underdetermines what ought to be believed about the world beyond our experiences. Scientific theories, however well secured by observation and experiment are inevitably fallible. Nor is there any basis for expecting the future evolution of scientific standards and methods to provide a more secure foundation for scientific knowledge. For methodological developments that have occurred thus far, whatever improvements they have generated at the level of human interaction with nature, have failed utterly to resolve the basic dilemma of the underdetermination of theory.²⁸

The second major area of objection to realism concerns "What kind of explanation is truth?"⁸ That is, what kind of mechanism is truth? How does the truth of a theory bring about, cause or create, its issuance of successful predictions? Briefly, "the truth of a theory is no more likely a mechanism for sifting out false predictions the theory might have made than it is likely as the mechanism by which the theory made true predictions."⁸

An instrumentalist is of course an instrumentalist about all theories. A realist, however, can be a realist about some theories (those which he believes are true) and an instrumentalist about others (which he believes to be useful, though not true).²⁹ In much of the scientific literature the term "model" has been used for the latter sort of theory. (See Box 'Theories and Models' for example).

This latter position has been stated in the form:

If you believe that guessing based on some truths is more likely to succeed than guessing pure and simple, then if our earlier theories were in large part true and if our refinements of them conserve the true parts, then guessing on this basis has some relative likelihood of success.

DKC's Rules

With that background, let us examine Demerjian, Kerr, and Calvert's rules and rationale for mechanism development and testing. These include:

1. Mechanism reactions must be elementary reactions.

These are believed to be the actual steps in which the transformation takes place. This representation is most commensurate with kinetic measurements and thermodynamics. Every other representation involves an approximation (*e.g.*, steady-state, no competing reactions, *etc.*) that may limit the adequacy of the representation under some conditions. Direct generalized reactions "provide no clue as to the real intermediates or transient species involved."

2. Rate constants must be based on experimental determination of rate constant or must be estimated using thermodynamic techniques.

Failure to use all known information to eliminate possibility of mere "curve fitting" results in a description for which there is "no evidence of the uniqueness or the correctness of the chosen mechanism."

3. The mechanism must explain smog chamber data.

If we can not explain the simpler systems in smog chambers we have no reason to infer that we have correctly chosen among the alternative mechanisms and rate constants.

4. The understanding of smog chamber data is best revealed by starting with simple systems and increasing the complexity of the hydrocarbon components stepwise.

Starting with simple systems focuses attention on particular aspects of the theory. Otherwise there is "no evidence of the uniqueness or the correctness of the chosen mechanism."

5. Rate constants of important reactions that were estimated 'theoretically' must be determined experimentally. This, in turn, may force a reevaluation of the smog mechanism.

Theory can predict only an approximate value for the rate constant. The application of thermodynamic theory to calculate a rate constant requires numerous auxiliary assumptions which contribute to an uncertainty in the prediction.

6. When there are "missing" chamber experimental information (e.g., photolytic rates), the experimental data must be supplemented by a rational process for estimating the missing data.

Smog chamber data inevitably have missing information that must be estimated by the theorist. This should be done in as mechanistic a manner as possible, but may involve "fitting" exercises to make final choice. This, of course, contributes to uncertainty, leading to the next requirement.

7. Must model more than one chamber. The basic core of the model must not be changed from species to species simulation or chamber to chamber simulation.

This derives from a long standing heuristic of observation science. An individual chamber and experiment may be inaccurate and misleading, so use repeated observations by different investigators using different tools.

From this abstraction of DKC's approach, we can deduce that DKC are, in some important respects, realists. They believe that there are "real intermediates"; that there are unique and correct mechanisms for the transformations; that some mechanisms are "seemingly accurate"; that there is a close correspondence between their theoretical constructs and the real world so that it is possible to "predict theoretically the effects of atmospheric composition, as yet experimentally unobserved, on the rates of formation" of interesting secondary products—even if they had to construct "imagined" reactions and rate constants from theory. They criticize their more instrumentalist colleagues who adjust models to fit chamber data by saying,

'Theories and Models'

We all learned Boyle's Gas Law,

$$PV = \text{constant}$$

and found out that, although useful, that it was not really correct for real gases. Boyle's law can be derived from a 'model' of gases which supposes that gas molecules are point masses. Later we learned that Van der Vaal's Law,

$$(P + \frac{a}{V^2})(V - b) = \text{constant}$$

was better, but still not quite correct, for real gases. Van der Vaal's law can be derived from a 'model' of gases which supposes that gas molecules are elastic spheres with real volume.

The best basic theoretical knowledge about real gases derives from the calculation of the intermolecular potential between gas molecules from quantum mechanics by solving the Schrödinger equation. In practice, this is virtually impossible, and such calculations have been made only for the very simplest molecules. In the absence of such complete calculations, it is necessary to approximate the potential by a semi-empirical formula, containing one or more adjustable constants, which is chosen on the basis of physical plausibility and mathematical convenience.³⁰

Theoretical laws are not generalizations of observations, but are postulates of a theory, and are tested indirectly by comparing the consequences of the theory, *as a whole*, with experimental or observational facts. Anomalies, by themselves, are insufficient reasons to reject theories.²⁹

It is evident that a good computer match of smog chamber rate data obtained using arbitrarily chosen rate data for the extremely complex systems involved in photochemical smog is no evidence of the uniqueness or the correctness of the chosen mechanism. In fact, conclusions and predictions based on the extrapolation of such a mechanism to conditions not employed experimentally can be completely misleading, since the *actual* intermediates involved and their competitive reaction steps may not be included.

The implication being that their method does not suffer from this problem, because they have used their fundamental belief in chemical kinetics and thermodynamics to calculate rates 'theoretically.' In so doing they opposed their fundamental belief to the "large degree of chemical intuition" that had been employed by others in the "choice of rate constants."

There is a remarkable omission in the DKC work: there is no description of the methodology for choosing the 490 elementary reactions that made up their mechanism from the thousands that they recognize must occur. There are phrases such as "rational series of elementary chemical reactions" and "complete reaction

scheme" and "seemingly accurate" schemes, and an expression of "hope of defining the reactions". It is clear that a near complete literature review was undertaken (there were 219 references, some of which were in turn reviews). A great number of the reactions and their measured rate constants used by DKC came from this review. DKC were also clear that, for rate constants that had not been measured, they had made theoretical estimates. In the latter case, there are two types of reactions: 1) those that have been *observed* to occur in the laboratory, but which have not had quantitative measurements allowing rate calculations; and 2) those that have never been observed but are *supposed* to occur, that is, which were hypothesized to exist to explain some other observation. In this case, we need to distinguish two sub-classes, the first is when the reaction does occur in the real world (*i.e.*, it was a correct guess) and the second is when the reaction does not occur in the real world (*i.e.*, it was a wrong guess).

While DKC implied that they needed smog chamber data to evaluate "alternative mechanisms and reaction rate constants proposed," they did not explain how this occurred. Recall the statement "We feel that the only real hope of defining the reactions which are important in the real atmosphere lies in the elucidation of the simpler smog chamber experiments." They did not make explicit a strategy for, to use Kuhn's terms, "elaboration of the various possibilities and to distinguish among them." In the discussion of each reaction, they occasionally use a practice of making choices on the basis of mostly qualitative comparisons of alternative prediction pathways but mainly with kinetic experimental data and not with chamber data (*e.g.*, this path can not be important because we do not observe its product). The incommensurability between the smog chamber data available to DKC and the needs associated with the theory articulation, however, may have prevented much more, so this strategy was not articulated in their approach (see below). DKC also did not offer a description of the methodology for the elaboration of the theoretical possibilities they considered.

There is a third class of reactions whose rate constants have not been determined. These are reactions that occur and have not been observed in the laboratory because no one has ever known to look for them. The theory never said that they are likely to exist or be important. Thus the only way they are likely to be found is through serendipity, a common enough happening in science. Note, however, that this requires a broadminded observer, one who does not discount the observation as "noise"—everything that is not predicted by the theory. The possibility of important undiscovered reactions must present a double bind for one who is trying to build a "complete reaction scheme": on the one hand, he *will not delete* anything from the scheme because it might be important in some way he

might not yet understand, and on the other, he probably *has deleted* hundreds of reactions that are important! Coming to grips with this situation has got to be uncomfortable for those seeking a “complete reaction scheme” for the “truly scientific understanding of air chemistry.”

How should a scientist deal with what was described above as the “basic dilemma of the underdetermination of theory?”

Another observation that can be made about the DKC approach is that it exhibits an asymmetry for the treatment of experimental fact. Contrasting points 6 and 7, with point 5, it is evident that failure to fit smog chamber data is not a sufficient reason to reject a mechanism. However, the experimental measurement of a rate constant that changes the value already being used in a mechanism may require a complete re-evaluation of the mechanism. By this analysis, chamber data constitute neither a sufficient nor necessary condition for testing theory-generated mechanisms. That is, success in modeling chamber data does not measure a mechanism's accuracy, and failure in modeling chamber data is not acceptable for mechanism modification. In this situation, the conclusions (based on the definition and significance of “sufficient and necessary”) about chamber data fits by mechanism predictions is that no mechanism is ever dependent in any way on chamber data, regardless of how the mechanism's authors construe their defenses.

I think this position troubles few modelers. That is, modelers have traditionally explained their reluctance to use chamber data by pointing to the massive uncertainties inherent in any such measurement system. This I have suggested is in part due to their basic realist philosophies. Such an explanation though is somewhat awkward when we recognize that the foundation that modelers take as unimpeachable—kinetic rate constants—are themselves measurements, often conducted in a highly distorted environment compared to the ambient air, and are subject to problems similar to ones that confront chamber data. It is true that chamber data are artificial, but I would contend that they might be considered less artificial than the conditions used to produce rate constants. In addition, chamber data, used as an environment to test real and non-real world situations, are distorted in the first sense [see Chapter 2] so that their usefulness can fit with the paradigm. They are less distorted in the second sense—that is, as a consequence of generalization and deletion that led to map-making. Kinetic rate data, however, are distorted in both the first and second senses: in the first just as chamber data are, and in the second sense precisely because the model itself

determines which rates are important and which are not, and therefore, may determine which rates are measured.

I believe that it is just easier to juxtapose the fact to the theory for the rate constant, but the fact of the smog chamber is just as real and should carry similar weight. The problem appears to be the difficulty of confronting theoretical detail with the more macroscopic fact from the smog chamber system. Furthermore, as Kuhn says, scientists who have spent their life's work with a particular paradigm can be quite comfortable not treating anomalies as counter-instances, and, therefore, are very reluctant to modify their theories.

In spite of their carefully worded effort to avoid it, DKC were somewhat hard on the smog chamber experimenter compared to the laboratory kinetist. The straight interpretation of what was said is:

- ▷ we had significant and recurring problems simulating smog chamber data, mostly due to incomplete information provided by the smog chamber team;
- ▷ this required that we guess about certain key information; therefore,
- ▷ we could not develop a mechanism free from arbitrarily adjustable parameters; therefore
- ▷ simulations using these mechanisms cannot add significantly to scientific understanding of atmospheric reactions.

And so "chamber-bashing" was added to the modeling folklore.

On the other hand,

- ▷ numerous kinetists did not measure numerous rate constants or even determine the existence of key reactions; therefore,
- ▷ we had to make 'theoretical' estimates of reactions that did not exist; therefore,
- ▷ we could not develop a mechanism free from arbitrarily adjustable parameters; therefore
- ▷ simulations using these mechanisms cannot add significantly to scientific understanding of atmospheric reactions.

'Famous Theoretical Rate Constants.'

While the use of thermodynamic techniques to calculate rate constants can be very useful and sometimes represents the only way to 'guess,' this methodology does not carry a guarantee of valid prediction. For example,



was calculated to have a rate

$$200 \text{ ppm}^{-2}\cdot\text{min}^{-1}$$

and DKC said this value was "consistent with all of our simulations." Later experimental measurements place the true rate at

$$12000 \text{ ppm}^{-2}\cdot\text{min}^{-1}$$

The following cases illustrate that it is possible to 'thermodynamically' estimate rates for reactions of entities that do not exist (e.g., HC(O)O_2) and for reactions that do not occur, as with $\text{PAN} + \text{NO}$. For a complete discussion of the PAN reaction see box 'Explicit Mechanism Formulation' later in this essay.

$\text{HC(O)O}_2 + \text{NO} = \text{NO}_2 + \text{HC(O)O}$ $9.1\text{E}2$ DKC's reaction 51a

$\text{PAN} + \text{NO} = 2 \text{NO}_2 + \text{CH}_3\text{C(O)O}$ $1.6\text{E}-1$ DKC's reaction 53b

The real issue here is the nature of Kuhn's second class of normal scientific work, that is, the work in experimental areas to produce facts that can be compared directly with theory and the manipulation of theory so as to produce predictions that can be confronted directly with experiment. Without close juxtaposition of the experimenter and the theorist, experimental measurements are not likely to include all the input needed to attempt a quantitative simulation of the rate data. The experimenter does not just randomly measure everything, he measures what he thinks is important according to his understanding of the system. If the theorist produces a new explanation, this will of necessity accord a different importance to different facts. Without new measurements suited to this new importance, the theorist must construct the needed "missing" facts if he wishes to use the existing data to test his theory. This is in small part what was recognized by Kuhn when he described the incommensurability of successive bodies of knowledge.

Likewise, one experimenter may have a different understanding of the importance of any explanation than another experimenter, and so may use different methodologies. In the evaluation of models this means that the model testing ability of a particular set of data may be somewhat incommensurate with the particular needs of the theorist while other aspects of the same data may fit other

needs of the theory and the auxiliary assumptions quite satisfactorily. This, then, is a particularly clear example of another way that underdetermination of theory has troubled air pollution chemistry.

This need to combine experimental and theoretical work was recognized by DKC in the last paragraph of the conclusions to their work,

We hope that the reader has become convinced that there is a real benefit to be gained by both parties in the close association between smog chamber experimenters and those who model such systems in terms of theoretical mechanisms. A truly scientific understanding of the chemical changes which occur in the various complex chemical mixtures which constitute our polluted atmosphere may never be possible. However our only hope of achieving this end lies in the understanding of the simpler smog chamber experiments. Eventually the knowledge gained from these systems will allow the atmospheric scientist to exercise the necessary degree of chemical sophistication in computer simulation of atmospheric reactions so that he can develop scientifically sound and useful predictive atmospheric models and control criteria.

Actually the science has advanced as suggested by DKC and the original theory augmented by strong auxiliary assertions has greatly improved the explanation of air chemistry. Likewise, the the chamber data have advanced, and today's best chamber data can offer so many counter-instances for major parts of DKC's mechanism that there would be little chance anyone would judge it an adequate representation of the truth. But in that that was itself one of DKC's goals, we can easily see that their approach certainly must be considered a significant, and primary contribution to the paradigm for the development of explanatory models.

But problems do remain. How else, other than the use of smog chambers, can we produce the detailed facts that must be compared unambiguously and directly with the model (theory+auxiliary statements) while preserving as much of the whole chemical transformation process as possible to allow for the evaluation of significant over-generalization and deletion? How (or even do) such comparisons deal with the question of the underdetermination of theory? And if the uncertainties associated with chamber data will always cause the model to underdetermine, then should we abandon the whole set of them?

... important characteristics of maps should be noted.

*A map is not the territory it represents,
but, if correct, it has a similar structure to the territory,
which accounts for its usefulness.*

*—A. Korzybski,
Science and Sanity*

First Reformulation

Hecht, Seinfeld, Dodge

As its title, "Further Development of Generalized Kinetic Mechanism for Photochemical Smog," suggested, the 1974 paper by Tom Hecht, John Seinfeld, and Marcia Dodge³¹ (HSD) was an effort to articulate further the paradigm's major theory by a reformulation to allow extension of its application to the atmosphere (Kuhn's third class of normal science work). While a prime objective for DKC was to have "a complete reaction scheme" and thus to explain, the prime objective for HSD was "that the mechanism predict the chemical behavior of a complex mixture of many hydrocarbons, yet that it include only a limited degree of detail" which would allow the mechanism to be applied to the "prediction of both smog chamber reaction phenomena and atmospheric reaction phenomena." Thus they sought "a reasonably rigorous, yet manageable, description of the photochemistry of air pollution." This work, therefore, in its conception and definition, set out toward a different end than the DKC work: DKC sought scientific explanation; HSD wanted prediction.

In part I suppose, because it would be difficult to argue from a position of 'completeness of representation' in their case, HSD said,

A kinetic mechanism, once developed, must be validated.

Reading this, I realized that the word "validate" never appeared in the DKC work! Continuing from HSD,

[Validation] is commonly conceived as consisting of two parts: validation in the absence of transport processes and validation in their presence. In practical terms we are speaking, respectively, of comparison of the model's predictions with data collected in smog chamber experiments and with data collected at contaminant monitoring stations in an urban area.

Something is not clear about this argument. We understand why the chemistry must be 'validated' in the absence of transport—this situation sharpens the point of comparison between the theories of transformation and the species measurements. Hecht³² said

In a system as complex as the atmosphere, isolation and characterization of the chemical processes that occur are virtually impossible. Thus, two fundamental approaches have been devised for controlled study of the chemistry of smog formation: irradiation of known concentrations of pollutants in a large reactor (smog chamber) and determination of the rates and mechanisms of elementary reactions thought to occur in polluted air. ... [In the chamber,] one is unable to gain insight into the details of the chemical transformations taking place: only the macroscopic effects of the overall chemical process is observed. On the other hand, detailed information concerning the rate and course of a single reaction, such as that obtained in a kinetics study, reveals little concerning the nature of the overall smog formation process. Indeed, one has difficulty knowing whether all the important reactions have been identified.

Kinetic simulation provides a means of comparing the results of these two experimental approaches for investigating the chemistry of smog formation. ... kinetic simulation takes the results of elementary reaction studies as input data and produces results of a form identical to the data obtained in smog chamber experiments as output. By comparing the predictions of the mechanism with actual chamber data, one is able to determine the degree of agreement between the two experimental approaches.

Therefore, unlike in the DKC approach, Hecht *et al.* recognize *a priori* that deletion is a significant problem in mechanism construction and they accord equal consideration to both kinetic data and smog chamber data.

With regard to the need to validate in the presence of transport processes, however, the wording in the HSD paper might be taken to imply that transport processes somehow modify the chemistry beyond the changes that effects the concentrations of reactants. What is the real need to validate in the atmosphere? Hecht said in the quote above that it was virtually impossible to isolate the chemical process in the atmosphere. In any case, HSD finessed the question by saying they were only comparing "predictions and experiment based on smog chamber studies."

After listing a series of mechanisms (including the DKC) developed by others, HSD made an important point:

For none of the specific mechanisms listed above has there been reported a program of validation over a range of initial reactant concentrations. Thus all five can only be considered at this point as detailed chemical speculations.

To be clear on this point, Hecht said,

... [Mechanism] uses are contingent upon the accuracy of the mechanism.

Smog chamber data constitute a standard against which the accuracy of predictions of a kinetic mechanism can be measured. But because of wall effects and other operating characteristics of the system, the observed time-varying chemical distribution in the chamber cannot be accounted for entirely in terms of the reactions occurring in the gas phase. Thus, the use of chamber results as a standard or reference for judging the accuracy of a mechanism must be carefully qualified.

He went on to say that because a chemical mechanism represents the chemistry that presumably would be observed in an infinitely large, uniformly irradiated, well mixed reactor and because a typical smog chamber does not meet these criteria, then, before predictions of the mechanism and chamber observations can be compared, terms must be added to the mechanism to represent the characteristics of the chamber that differ from ideality.

Further elaboration of the nature of validation was presented by HSD. For example

The significance of the validation results for a kinetic mechanism is to a large degree dependent upon the diversity and reliability of the experimental data base.

Hecht³² said that the rates of numerous reactions were being determined in a continuing fashion and that speculative reactions were being scrutinized, and known reactions were being confirmed. He added

In such a dynamic scientific environment, it would be presumptuous to suggest that any given set of reactions and rate constants is *the* chemical mechanism for smog formation.

A bit later in his discussion he said that most changes in mechanisms come about because of the need to "replace overly general representations with more detailed descriptions and to substitute sound experimental measurements—as they become available—for speculative estimates."

HSD made a clear statement about what they meant by diversity in the chamber data:

- a variety of hydrocarbon systems of high and low reactivity had to be included;
- the experimental data had to be for single reactants and for mixtures;
- the initial conditions had to cover a broad range of HC-to-NO_x ratios.

The latter condition was seen as especially critical for control applications in urban atmospheres.

By reliable data base, HSD said they meant:

- certain experimental variables had to be specified accurately;
- potential chamber wall effects had to be quantified;
- experimental measurements had to be specific, accurate, and repeatable.

An interesting remark made by HSD with regard to the data used by them was “in spite of [*some problems*], the data are in general reproducible, were carefully taken, and are as suitable as any currently available for validation purposes.” This was a recognition of a problem that will always plague modelers. The chamber and laboratory data require 2-3 years to be produced, be refined, and be communicated. New models, therefore, are likely to incorporate ideas that are of newer origin than those that drove the design of the chamber and laboratory experiments. The irony, of course, is that when the data can be made available, they often conflict remarkably and repeatedly with the model predictions. I think that the ability of chamber data to present sincere tests 2-3 years after they were observed—and observed sometimes for different purposes—confirms the importance of the testing for which HSD called. That such testing confirms ‘validation,’ though, has not proved convincing.

By validation of the mechanism, HSD meant following the process of:

1. obtaining estimates of the various input parameters to the mechanism:
 - ▷ the reaction rate constants,
 - ▷ parameterized stoichiometric coefficients,
 - ▷ initial concentrations of reactants,
 - ▷ chamber operating parameters—dilution rate, light intensity, *etc.*
2. carrying out sensitivity studies for these parameters—*i.e.*, establish the effect of controlled variations in the magnitude of the various parameters on the concentration-time profiles for major species;
3. predicting concentration-time profiles for the various reactant mixtures using the initial conditions;
4. comparing predictions with experiment results by plotting both predicted and measured concentrations on the same plot.
5. assessing the “goodness of fit” and identify the sources of uncertainty
6. reaching a conclusion with regard to the mechanism’s adequacy for the particular application.

I had not read this paper for nearly 10 years, but its effects had undoubtedly shaped some of my work. In studying the paper for this essay, I was surprised to

find how accurately it describes my current practice: I used this exact procedure in recent tests of the Carbon Bond Four and the RADM mechanisms.

As to the mechanism development process, Hecht³² provides more insight than the HSD paper. He says:

The primary criterion used to decide whether a reaction involving reactants present in photochemical smog should be included in a chemical mechanism is the rate of the reaction. ... To be sure, the complete kinetic analysis of a reaction entails more than the measurement of the rate constant: the mechanism and products of the elementary reaction must be determined as well.

While many of the reactions in his mechanism were thought to be well characterized, Hecht recognized that there were significant exceptions. He said,

Complete omission or inaccurate specification of products of key reactions, along with inaccuracies and uncertainties in the values of many rate constants, are probably the causes of a major portion of the disagreement between predictions and experimental data. Resolution of these types of uncertainties is a slow and painstaking process. In the absence of experimental measurements, provisional rate constant values can be estimated and elementary reactions and reaction mechanisms can be hypothesized on the basis of thermochemical principles (*e.g.* Benson). However, because of the large uncertainty associated with most estimates, we depend mainly on experimental studies to justify the changes introduced into the mechanism.

That is, while experimental work was underway, Hecht would hypothesize a particular reaction mechanism and rate constants so as "to augment the agreement between the predictions and the data for the smog chamber"; this hypothesized mechanism must be tested extensively. He said

The development of the mechanism during this project followed two complementary paths. The first approach involved testing the quality of predictions in relation to smog chamber data. By adding, modifying, or deleting reactions and then varying the values of rate constants within their bounds of uncertainty, we sought to increase the accuracy of the prediction of the mechanism over a wide range of organic compound-to-NO_x ratios.

The second approach focused on quantifying the influence of uncertainties in the values of individual reaction rate constants on the predictions of the mechanism. ... [The uncertainty] index reflects the importance of obtaining more accurate rate constant measurements as a means of reducing the uncertainties in the predictions of the model. In addition, ... we were able to point out insensitive reactions that can possibly be deleted from the mechanism without significant loss of accuracy.

Along the latter lines, HSD introduced some compression techniques that are still in use in today's mechanisms. For example, alkyl radicals are assumed to be converted to peroxyalkyl radicals exclusively and thus never need specifically enter into the mechanism, and acylate radicals (RC(O)O[•]) are assumed to decompose

exclusively to from alkyl (hence peroxyalkyl) radicals and CO_2 , thus they do not have to be included in mechanism, and so forth. These steps by themselves remove much detail from the mechanism without any significant loss of accuracy and therefore are still used today, even in so-called explicit mechanisms.

HSD listed a number of criteria that had to be met by a mechanism suitable for atmospheric prediction:

- A. it must rigorously treat the inorganic reactions; *i.e.*, explicit elementary chemistry must be used for these reactions;
- B. the detail associated with specific hydrocarbon mechanisms, *i.e.*, ones in which each species in a reaction represents a distinct chemical entity, is not practical in developing a mechanism for the atmosphere; therefore a "lumped" representation must be used, *i.e.*, "one which contains certain fictitious species that represent entire classes of reactants";
- C. on the other hand, the mechanism structure must not depend upon difficult-to-specify stoichiometric coefficients or parameters; such parameters adversely affect the "confidence in the predictions of the mechanism for cases not explicitly validated"; therefore, the stoichiometric coefficients and mechanism structure must be derivable directly from the underlying chemistry of each elementary reaction.
- D. rate constants must be related to elementary reactions; *i.e.*, "While the kinetic mechanism is written in a general fashion, we have striven to formulate it in such a way that all important features of the detailed chemistry are retained. Thus our goal has been to include each elementary reaction thought to contribute to the overall smog kinetics."
- E. the validation value of the rate constants must be within the range of values recommended by a consensus of investigators;
- F. the representation of the atmospheric hydrocarbon mix must not be oversimplified; *i.e.*, the reactivity and change in reactivity must be preserved in the generalized representation; therefore, the mechanism must be "sufficiently detailed to distinguish among the reactions of the various classes of hydrocarbons and free radicals"; the mechanism must include in its specification a method for how to represent the complex atmospheric mixture in the input.

The HSD mechanism was a conscious attempt to produce a representation that contained carefully described, nevertheless fictitious entities for the purpose of simplicity. Look back at the box "Theories and Models." Molecules are not

point masses, yet we find that treatment a useful fiction. By generalizing the important reactions and deleting the unimportant reactions, the HSD mechanism became much more understandable in a qualitative sense. It certainly is easier to understand the overall transformation processes using this mechanism than, for example, using the 47 pages of reactions in the DKC mechanism.

Clearly the emphasis of the HSD mechanism is on *prediction* and less on *explanation*. If one really wanted to know, for example, why substituting one HC for another in a mixture did not change the O_3 , the HSD mechanism as an explanation is fairly useless. One would have to say, "Because the stoichiometric coefficients for the HC that was substituted were the same as those of the original mixture." Not much of an explanation! This is, I believe, part of the reason for the requirement expressed in point C above: the stoichiometric coefficients and mechanism structure must be derivable directly from the underlying chemistry of each elementary reaction. This keeps HSD in line with DKC and with a realist viewpoint in that there was a common belief that the fundamental chemistry is the same for test, explain, predict. From this point of view, HSD were a new tack on the same course. It is the extension of the fundamentals, however, that separates HSD from DKC. The last sentence of point F illustrates this separation: HSD start with an input representing the "complex atmospheric mixture" and then include in their model exactly what DKC had not even finished describing!

After comparing their model predictions with chamber data, HSD said

... the data and predictions are not always in good agreement for all species over the full period of irradiation. These discrepancies can be attributed to at least five [sic..only four listed] possible sources of uncertainty:

The sources were:

- The mechanism may be incomplete.
- The lumping process may introduce error.
- There are uncertainties in the experimental data used for validation.
- Not all the rate constants and pathways are known with a high degree of certainty.

And HSD conclude:

As a consequence of these uncertainties, we have not yet reached the point in model validation where we are in a position to assess the "goodness" of the proposed mechanism, or for that matter, to draw unequivocal qualitative conclusions regarding its merits. Yet, the mechanism appears capable of predicting the concentration-time behavior of a variety of reactant systems over a wide range of initial conditions.

Did this mean that the mechanism was validated? To be validated means³³ to have "strength or force from being supported by fact," to be "well grounded on principles or evidence," and therefore, to be "able to withstand criticism or objection." Because of uncertainties in both the mechanism and the chamber data, HSD were not willing to say that the mechanism was quantitatively accurate, but they were willing to say that the mechanism was valid for "predicting ... behavior." That is, it was capable of predicting external appearance or action. This means that HSD were asserting that the mechanism was satisfactory for prediction, but not satisfactory for explanation.

HSD were making a bold claim! They were announcing to the "scientific community" that they had reasonable enough agreement between theory and facts that subsequent use of their model could be believed as to the predictions of the external appearance or action of smog chemistry.

How could they do this? Kuhn gives us an explanation in his "The Function of Measurement in Modern Physical Science."¹⁵ While describing tables of theoretical predictions verses experimental measurements which are supposed show agreement, he says,

... scientists seek not "agreement" at all, but what they often call "reasonable agreement." Furthermore, if we ask for a criterion of "reasonable agreement," we are literally forced to look in the tables themselves. Scientific practice exhibits no consistently applied or consistently applicable external criterion. "Reasonable agreement" varies from one part of science to another, and within any part of science it varies with time.

For example, in spectroscopy, "reasonable agreement" means agreement in the first six to eight left-hand digits in the numbers of a table of wavelengths. In the theory of solids, by contrast, two-place agreement is often considered very good indeed.

Kuhn further says that the actual function of measurement must be sought in the journal literature, which displays not finished and accepted theories, but theories in the process of development. When a journal article is accepted by the profession as showing "reasonable agreement" it becomes the definition of "reasonable agreement." Kuhn says that is why the tables are there. By studying the tables and graphs a reader learns what can be expected of the theory. An acquaintance with the tables is part of an acquaintance with the theory itself. "Without the tables," says Kuhn, "the theory would be essentially incomplete."

The quantitative comparison of theory and observation is Kuhn's second type of normal scientific work and he claims that this work has required the highest

scientific talents to invent apparatus, reduce perturbing effects, and estimate the allowance to be made for those that remain. In so far as their work is quantitative, most scientists do this type work most of the time. He says, "Its objective is, on the one hand, to improve the measure of 'reasonable agreement' characteristic of the theory in a given application and, on the other, to open up new areas of application and establish new measures of 'reasonable agreement' applicable to them." This is exactly what HSD did. Since there was no refutation of their statement by the 'scientific community', there *was* "reasonable agreement."

We shall see this process at work many times in this brief history.

Second Explanation

Durbin, Hecht, Whitten

Until about 1975, mechanisms had been compared with smog chamber data that were, in effect, “left over” from the pre-paradigm stage. That is, the data were not produced using the new insights provided by the HO[•]-paradigm. As described by Kuhn, this resulted in an incommensurability between the old facts and the new theory because the experimental work before modeling practice began was directed at different goals and needs than the facts needed to compare with models. The pre-paradigm experimenter had not measured things the HO[•]-paradigm modeler considered from his new prospective to be critically important. In addition, as Kuhn suggests,¹⁵ without some expectation as to what the results “ought” to look like, the experimenter was satisfied with what he got.

In the period 1972–75, the Statewide Air Pollution Research Center (SAPRC) at the University of California at Riverside (UCR) under the leadership of Jim Pitts, and with funds from both the state of California and the U.S. EPA, designed and built a chamber facility specifically to produce data for model validation. By 1975, the first of these data became available to modelers and a new round of improving the characteristic “reasonable agreement” between theory and experiment began. Meanwhile, the sensitivity analysis work done by Hecht *et al.* provided guidance to kinetists as to what reactions were critical, and so, new kinetics information was also becoming available. In addition, the stratospheric ozone and SST issues generated a large interest in kinetic data needed to produce credible models of the stratosphere; much of this work was directly applicable to the air pollution models, especially for the inorganic reactions. As described in the previous chapter, however, the production time for such “facts” is 2-3 years and therefore there is always a lag between new mechanism concepts and experimental data.

The first victim of this new data was the generalized HSD mechanism. Whitten, in the introduction to his 1977 EPA report³⁴ said,

Some problems were encountered, however, in attempting to apply the [HSD] mechanism to situations other than ... the range of concentrations and hydrocarbon mixes upon which it was based.

These problems stemmed largely from the use of parameters that were based on smog chamber data, not on the fundamental chemistry. Incorrect fundamental chemistry and chamber-dependent phenomena could be compensated for or masked by these parameters. No single set of parameters would fit all smog systems, and there was little theoretical guidance for adjusting the parameters for systems for which no experimental data existed.

If you think you have read that before, you are right. Demerjian, Kerr, and Calvert made essentially the identical statement (quoted on page ?? of this text).

Durbin, Hecht, and Whitten (DHW) suggested an alternative in their 1975 report.³⁵ That is,

Ideally, a kinetic mechanism would be simply the assemblage of results from kinetics studies of all reactions that occur. In reality, not all of the reactions that could occur have been studied, and often orders of magnitude of uncertainty may be associated with those that have been studied. In addition, the number of possible reactions in smog is very large. Hence, a complete mechanism is neither practical, because it would include an enormous number of reactions, nor feasible, because the needed kinetic information is unavailable. The best one can do to obtain a closed kinetic system of chemical equations and rate constants is to use available kinetic information and methods for estimating other information, to draw analogies, and to make other simplifying assumptions. The kinetic mechanism that results may be fairly accurate, albeit simplified, description of reality.

The mechanism's complexity is then dictated by the criterion for "important."

DHW therefore compared the explicit mechanism's predictions with *detailed product yields from chamber experiments* (the new data) to determine if they have all the "important" reactions. The power of the explicit mechanism, Whitten later explained, was

Because explicit mechanisms are based on the fundamental chemistry, a poor fit between predictions and measurements for a given species can sometimes be traced to uncertainties in chemical reactions or inaccuracies in smog chamber experiments. For example, poor fits between predictions and measurements for some propylene/ NO_x experiments in the evacuable chamber at [UCR led DHW] to hypothesize that the intensity from the UV light source in the chamber was decreasing more rapidly at

Explicit Mechanism Formulation

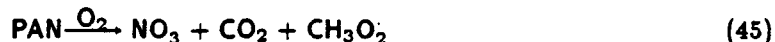
The following example is taken from Durbin, Hecht, and Whitten's 1975 EPA Final Report. It illustrates typical reasoning used in constructing explicit mechanisms.

PAN Chemistry

The formation of peroxyacetylnitrate (PAN), and its homologs, occurs by another radical-radical reaction. Reaction (44),



was discussed in last year's final report. It was speculated there that PAN might hydrolyze on the walls of the UCR chamber. Although this undoubtedly could occur, an analogy to N_2O_5 suggests that gas phase collisional destruction could be several orders of magnitude faster than surface reactions under ambient conditions. Thus, we presently propose that PAN may undergo a thermal decomposition reaction, resulting in the rupture of the peroxy and carbon-carbon bonds:



Based on data contained in Benson (1968) and Domalski (1971), this reaction is exothermic by about 14 kcal/mole.

The occurrence of Reaction 45 is supported by the experiments of Schuck *et al.* (1972) and recent PAN decay experiments in the Riverside chamber. In the former study, PAN was found to oxidize NO to NO_2 . The reaction was first order in PAN and zeroth order in NO. The ratio of CO_2 produced to PAN consumed was nearly 1. The ratio of NO_2 formation to PAN consumption was approximately 2 in a nitrogen atmosphere, but was much greater than 2 in an oxygen atmosphere. This is further evidence for the occurrence of Reaction 45, followed by $\text{NO}_3 + \text{NO} \rightarrow 2\text{NO}_2$ in the nitrogen atmosphere, and the same reaction plus $\text{CH}_3\text{O}_2 + \text{NO} \rightarrow \text{NO}_2 + \text{CH}_3\text{O}$ and $\text{CH}_3\text{O} + \text{O}_2 \rightarrow \text{HCHO} + \text{HO}_2$ and $\text{HO}_2 + \text{NO} \rightarrow \text{NO}_2 + \text{HO}$ in an oxygen atmosphere. Schuck *et al.*'s rate constant, $2.06 \times 10^{-2} \text{ min}^{-1}$, is 10 times that obtained from the half-lives observed in the UCR chamber. This difference may be due to additional wall decomposition in Schuck's reactor. The Riverside half-lives of 5.7 ± 0.1 hours in the light and 5.5 ± 0.4 hours in the dark provide further confirmation that PAN does not photo-decompose at an appreciable rate (Leighton, 1961).

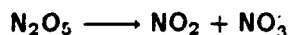
Continued on next box.

short wavelengths than at long wavelengths. Subsequent measurements on replacement light sources at UCR were consistent with this hypothesis.

For an example of uncertainties in chemical reactions see the box 'Explicit Mechanism Formulation.'

Explicit Mechanism Formulation, continued

The N_2O_5 analogy referred to by DHW in the box above was



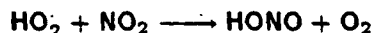
In 1977, based on detailed kinetic experiments, Hendry and Kenley published³⁶ what is now known to be the correct reaction for PAN decomposition:



which is just the reverse of its formation reaction. Thus, the OO-N bond is the weaker one, not the CO-ON bond as was suggested in the DHW analysis. Sometimes analogy works and sometimes it does not.

Subsequent understanding of the PAN system suggested that wall decomposition in Shuck's reactor had little to do with his measured life-times. This example of a counter-instance to a theory-based hypothesis is consistent with Kuhn's description of how scientists often deal with such anomalies.

Hendry and his co-workers were more explicit than most modelers in their 1978 EPA report³⁷ saying, "During this year, the following four major developments in laboratory data set our effort apart from the earlier modeling programs." These were: 1) The observations that



does not occur at a significant rate. 2) A much higher rate constant for



than had been previously used (e.g., 12000 vs. 2000). 3) That PAN is thermally labile. And 4) A breakthrough in the modeling of toluene has resulted from data on the initial products of the reaction of HO^\cdot and toluene under conditions applicable to the atmosphere.

Whitten described the advantage of the explicit mechanisms as follows:

The explicit mechanisms predict smog chamber data better than the HSD mechanism, and without any adjustments of parameters they fit a much wider range of concentrations than does the HSD mechanism. They provide more detailed insight into the smog formation process. Because they are not as empirical, there is a theoretical justification for applying them outside the range of concentrations and hydrocarbon mixes used in smog chamber experiments. Furthermore, if chamber-dependent reactions are removed and appropriate atmosphere reactions are added, an

explicit mechanism can be used as a component of a regional air pollution simulation model.

This is equivalent to saying the the explicit mechanism is, for the most part, "real" and therefore it can be believed to map correctly the real world. This is also a move toward the position of Demerjian, Kerr, and Calvert.

The mechanism developed by DHW and described in their 1975 report³⁵ was used by Dodge as the basis of the EPA control strategy calculation method. This mechanism and the re-formulations produced by Dodge to use it for prediction will be described in the next chapter. That work overlapped the mechanism developments described next.

Late 1970s Explicit Models.

In his 1977 EPA report, which was mostly concerned with developing explicit mechanisms for propylene and n-butane, Whitten says,³⁴

The objective of [this work] is to produce simulations based on scientific knowledge that also fit the measurements. How well this objective was achieved can best be seen by examining the many figures in this report that show UCR measurements and our simulations. The overall fit for most species is good.

As an example of Whitten's concern with fitting, I will use his discussion of his propylene model for the UCR data,

In this report we present what we consider our best simulations of these runs, using the same overall mechanism throughout and varying only the light spectrum and initial HONO concentration.

In our final fitting procedure we began with propylene runs EC-95, EC-96, and EC-121, because detailed spectral measurements were available for these runs. We used our photolysis constant program to calculate the carbonyl photolysis constants. We then varied the percent of ozonide formation in the mechanism until the total maintenance source of radicals from aldehydes, propylene-O atom reactions, and ozone-propylene reactions gave a good fit to the UCR data. Note that the percent of ozonide formation is only a fitting parameter. If the quantity is measured, many other parameters or added reactions could be used as fitting parameters. Some of these are as follows:

- ▷ Reactions of ozonides with radicals to act as sinks, or reactions that produce or release radicals.
- ▷ A slower or faster rate constant for the ozone-propylene reaction.
- ▷ Larger or smaller quantum yields for carbonyl photolysis
- ▷ Fundamental changes in the mechanism, such as PAN chemistry, NO_x loss chemistry, and aldehyde formation.

During the fitting procedure for propylene runs, we discovered that the photolysis constants calculated from the spectral data for runs EC-95 and EC-121 differed. Of course, it was still possible to produce good fits by arbitrarily varying the photolysis constants in the simulations of these runs. ... Thus, we had to lower radical production from the ozone-propylene reaction by setting the tuning parameter, the fraction of ozonide formation, to a high value, 0.33.

[And so forth.]

Because he was so concerned with fitting the measurements, at that time Whitten was often accused of being too much an instrumentalist and too little a realist.³⁸

To understand the dilemma, however, let us contrast Whitten's discussion with that of Carter *et al.*³⁹ for the same data and time period (i.e., late 1977),

We present here a detailed model for the NO_x-air photooxidation of n-butane and/or propene which is consistent with most of the extensive smog chamber data obtained at SAPRC for propene and n-butane, and detailed comparisons between model predictions and experimental data are shown. These efforts to produce a satisfactory model for the n-butane + propene systems have led to the identification of a number of experimental problems and mechanistic questions in need of further study, and these are discussed here. ...

Fits were usually attained to within $\pm 20\%$ or better ...

The good fits to experimental data were attained only after adjusting several rate constants or rate constant ratios related to uncertainties concerning chamber effects or the chemical mechanism. The largest uncertainty concerns the necessity to include in the mechanism a significant rate of radical input from unknown sources in the smog chamber. Other areas where fundamental kinetic and mechanistic data are most needed before a predictive, detailed propene + n-butane-NO_x-air smog model can be completely validated concern other chamber effects, the O₃ + propene mechanism, decomposition rates of substituted alkoxy radicals, primary quantum yields for radical production as a function of wavelength for aldehyde and ketone photolyses, and the mechanisms and rates of reaction of peroxy radicals with NO and NO₂.

In his "Normal Science as Puzzle-solving" chapter, Kuhn says,

If it is to classify as a puzzle, a problem must be characterized by more than an assured solution. There must also be rules that limit both the nature of acceptable solutions and the steps by which they are to be obtained.

If we can accept a considerably broadened use of the term 'rule'—one that will occasionally equate it with 'established viewpoint' or with 'preconception'—then the problems accessible within a given research tradition display something much like this set of puzzle characteristics. The man who builds an instrument to determine optical wave lengths must not be satisfied with a piece of equipment that merely attributes particular numbers to particular spectral lines. He is not just an explorer or measurer. On the contrary, he must show, by analyzing his apparatus in terms of the established body of optical theory, that the numbers his instrument produces are the ones that enter into theory as wave lengths. If some residual vagueness in the theory or some unanalyzed component of his apparatus prevents his completing that demonstration, his colleagues may well conclude that he has measured nothing at all.

I believe that within the 'scientific community' there is something subtly more acceptable about how Carter said essentially the same thing that Whitten said. Whitten appears to have 'knob-twiddled' while Carter used 'scientific uncertainty.'

While some in the modeling community may have found subtle differences between these two, the conclusions reached by many scientists and policy makers in the air pollution community after being exposed to such discussion was “We understand nothing at all!” For example, even in 1984, Leone and Seinfeld said, “It is our feeling that with the current controversy regarding the magnitude of chamber radical sources, it is difficult to consider any mechanism validated simply because it can simulate chamber data.”

The Beginning of the First Crisis

So there was an impression that one group of modelers were adjusting the model in such a way that it was not clear that the values were the correct ones to enter the theory, while the other group of modelers had a vagueness in their theory and a unanalyzed component in their apparatus.

This perception began to undermine the promise that had come with the adoption of the HO₂-chain paradigm and to dampen the exhilaration that arose from the initial successes of its application.

*There's no sense being precise about something
when you don't even know what you are talking about.
—John von Neumann*

Second Reformulation

The Dodge Mechanism

While the field was turning to explicit mechanisms for their explanatory powers there was a pressing need for practical prediction. To deal with this situation, Marcia Dodge produced an interesting re-formulation of the existing paradigm.⁴⁰ She used the explicit mechanism for propylene and butane developed by DHW and Dimitriadis' BOM automobile exhaust smog chamber data to produce an isopleth diagram that was taken to be representative of the "worst case" ambient polluted urban atmosphere. If that seems like a major extension of the paradigm to allow an application, it was. The formulation of this concept developed in the interaction between Dimitriadis' desire to extend his smog chamber results to the atmosphere and Dodge's acute awareness of the limitations of the 1975 models.

To highlight the latter point, Dodge had commissioned the Durbin, Hecht, and Whitten (DHW) report. In this report, DHW had presented forty pages of chemical kinetics discussion to support their formulation of the propylene/butane mechanism. They said that smog chamber studies done at UCR and at Battelle "serve as the data base for validating the kinetic mechanisms." These consisted of five propylene and four butane chamber runs from the UCR chamber, and five propylene runs in the Battelle chamber. DHW had described the comparison of prediction and measurements for the propylene experiments as being "fairly good" for NO, NO₂, and O₃, while carbonyls were "consistently low" and PAN predictions were "some times high and some times low." They remarked that "The chemistry of PAN is still largely unexplored, and this is very likely a large part of the problem." They said,

The butane mechanism was considerably less successful than the propylene mechanism, even though they are similar.

Of course, the traditional scapegoats, uncertainties of surface and photolysis reactions, can also be blamed for the disparity between model predictions and experimental results. Although they probably exacerbate the problem, they are not the sole culprits.

Although the accuracy was not very good, the propylene mechanism was still able to follow the behavior of each species.

The mechanism's application utility was demonstrated in a study of hydrocarbon reactivity and ozone formation. Thus the mechanism can be a useful tool to investigators of photochemical air pollution.

Thus the claims were less bold this time. Nevertheless, there was the suggestion that the mechanism might be useful in prediction, if not entirely acceptable for explanation.

Dodge accepted this proposition. She chose the DHW propylene/butane mechanism as the starting point for her own modeling tests. I think, however, that Dodge was skeptical about both the utility of the DHW formulation and the quality of the nearly 10-year old BOM data until she was successful in fitting the data with the mechanism. Kuhn says,¹⁵

When measurement is insecure, one of the tests for reliability of existing instruments and manipulative techniques must inevitably be their ability to give results that compare favorably with existing theory.

With a very reasonable set of assumptions about the photolytic rates and the chamber characteristics of the BOM chamber runs, Dodge was able to achieve a "goodness of fit" between the predictions and the data that was less (ca. 10%) than the BOM replicate run agreement (ca. 15%) while only using a single adjustable parameter: the relative concentration of the initial butane and propylene. Dodge discovered that the reactivity of auto-exhaust in the BOM chamber could be well represented by the DHW propylene/butane model by assuming that the initial NMHC in the auto-exhaust was represented as 75% butane and 25% propylene. Thus, the mechanism—now known as the Dodge Mechanism—was able to *predict* ozone formation, but could not *explain* how it occurred. That is, the ozone formation in the BOM chamber certainly did not occur because of reactions exclusively for butane and propylene. Dodge was careful to avoid the word *validate*; it was only used once in the whole paper and then to refer to the chamber data, not the model.

Dodge had not *validated* this kinetics mechanism, she had *vindicated* its use in atmospheric applications. Vindication means³³ "a fact or circumstance that justifies" and justify means "to show an adequate reason for something done."

Vindication does not carry that sense of "well grounded on principles" which is associated with validation; it just implies that one can assert a claim. So, based on her ability to predict the ozone concentrations in the BOM chamber, and on the fact that there were mechanistic representations of processes in the model, Dodge asserted that she could adjust the model conditions to be "representative of the polluted urban atmosphere." This she did.

This methodology proved to be such a useful strategy that OAQPS has considered using the Dodge Mechanism 10 years after it was formulated and vindicated. In view of the enormous success EPA has had with this strategy, and its demonstrated scientific utility, one wonders about the utility of vindication verses the hope of validation.

Falls and Seinfeld, McRae Mechanism

In their 1978 paper,⁴¹ Falls and Seinfeld (FS) continued the approach used by Hecht, Seinfeld, and Dodge. That is, they constructed a mechanism containing fictitious entities and stoichiometric coefficients for representing the chemistry of these entities. FS described two approaches for atmospheric mechanism design:

- ▷ surrogate mechanism—a mechanism in which organic species in a particular class, *e.g.*, olefins, are represented by one or more members of that class, *e.g.*, propylene;
- ▷ lumped mechanism—a mechanism in which organic species are grouped according to a common basis such as structure or reactivity.

FS concluded that "because of the computational requirements associated with calculating chemistry and transport simultaneously, surrogate mechanisms are inappropriate for use in atmospheric models that include an adequate treatment of meteorology." FS further concluded that in 1978, "insufficient information is available to include aromatics species" in the mechanism under development, and therefore, the mechanism only included olefins, paraffins, and aldehydes.

With regard to mechanism formulation, FS said,

In view of the complexity of the HO[•] and O₃ reaction paths for individual hydrocarbons, the major problem in developing a lumped mechanism is representing the chemistry of a mixture of alkanes, alkenes, and aldehydes. We choose to partition the organic species into five classes: ethylene, higher olefins, paraffins, formaldehyde, and higher aldehydes....

Because both rate constants and mechanisms differ among the species that comprise a lumped class, lumped mechanisms, by necessity, involve parameters that must

be chosen *within theoretically defined bounds* to provide the best fit of data and predictions. The values of such parameters cannot be defined precisely ... The art of developing a mechanism is to deal with the inherent uncertainties in an optimal manner, while adhering to certain important constraints, such as maintaining carbon, nitrogen, and oxygen balances...

We have opted to concentrate the inherent uncertainty of the lumped mechanism in the RO chemistry. ...



Thus, the lumped olefin-HO \cdot and paraffin-HO \cdot reactions are written as producing RO $_2$ radicals that, when reacting with NO give NO $_2$ and a lumped RO radical. The stoichiometric coefficients obey the following restrictions $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1$.

FS were very careful not to use the word *validate* when describing the tests of their mechanism; instead they *evaluated* the mechanism using smog chamber experiments. Evaluation means³³ "to determine the worth of; to appraise." FS said,

In evaluating a mechanism, the customary procedure is to compare the results of smog chamber experiments, usually in the form of concentration-time profile, with simulations of the same experiment with the proposed mechanism. In all mechanisms a sufficient number of experimental unknowns exist such that the predicted concentration profiles can be varied somewhat by changing rate constants (and perhaps mechanisms) within accepted bounds. The inherent validity of a mechanism can be judged by evaluating how realistic the parameter values used are and how well the predictions match the data. Since the mechanism will generally not be able to reproduce every set of data to which it is applied, two factors must be considered: identification of the major sources of uncertainty, such as inaccurately known rate constants or mechanisms of individual reactions, and evaluation of so-called "chamber effects", phenomena peculiar to the laboratory system in which the data have been generated.

Recalling that *validity* means "strength or force from being supported by fact," in effect, FS said, "The strength of a mechanism that arises from being supported by fact can be appraised by determining how well the parameters conform to 'real' values and how much 'reasonable agreement' there is between the model predictions and the facts." In this particular case, I find it hard to accept this assertion. There are no "real" values of the parameters in the chemistry of the chamber; these are fictitious entities. Therefore, such parameters cannot be specified before hand for a particular mixture, but must be obtained by trial and error fits to the data, that is, they are empirical parameters and are thus limited to the range of conditions used to select their values. For example, FS said, "For

the propylene/butane system, good simulations results were obtained when the RO decomposition step was taken to give RCHO and HO₂ exclusively." In other words, in spite of the accepted fact that propylene yields one HCHO for every HO + propylene reaction, the FS model fits the chamber data best when this fact is ignored even when large concentrations of propylene were present.

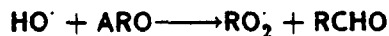
With objection to FS's first criterion, I looked to the second, reasonable agreement between prediction and data. For demonstration of "reasonable agreement" FS simulated

- ▷ three n-butane experiments,
- ▷ eight propylene experiments,
- ▷ five propylene/n-butane experiments.

The evaluation data base barely exceeded the number of smog chamber experiments used to test the original Hecht, Seinfeld, and Dodge mechanism four years before. There were no plots of data and predictions shown in this paper, only a table of O₃, NO₂ and PAN maximum value predictions and chamber values were given as demonstration of "reasonable agreement". Failure to use all the chamber data available (more than 44 experiments) and to use a higher level of reasonable agreement criterion meant that FS had applied insincere tests to their model. Perhaps this is why the paper ended with

Although the revised mechanism contains significant new information on rate constants and some mechanisms, there still exist, in addition to aromatic chemistry, key areas of uncertainty that require experimental study.

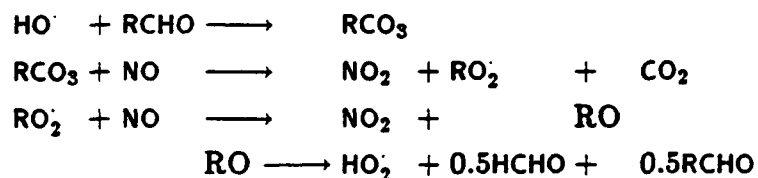
While in the FS paper there was no claim or assertion of adequacy for prediction, this did not prevent McRae from later claiming in Chapter 8 of "Mathematical Modeling of Photochemical Air Pollution"⁴² that the FS mechanism was selected because it "has been validated against a wide range of smog chamber experiments." Furthermore, in McRae's description of the mechanism, aromatics had been added as a lumped species and their chemistries were represented by one reaction



With similarly little justification or explanation, the values of the parameters for RO were assigned $\alpha = 1$, $\beta = 0.5$, $\gamma = 0.5$.

The choice of β and γ has a certain physical significance as was described by Killus in discussing a similar reaction in the Demerjian Photochemical Box Model

mechanism. It is the only choice that preserves carbon balance. Killus pointed out that the cycle



only preserves mass balance *if the coefficients in the last reaction are exactly 0.5!* This is because of the infinite series:

$$1 = 0.5 + 0.25 + 0.125 + 0.0625 + 0.003125 + \dots$$

If γ was larger than 0.5, then the system would gain carbon mass each cycle. This un-realistic production of RO_2 was recognized by Falls, McRae, and Seinfeld⁴³ but they suggested with no further elaborations that it might have a small effect when the initial HC-to- NO_x ratio was large and "other modes of radical production besides RO were available."

Seven SAPRAC smog chamber experiments using multi-component urban mixtures were simulated with the McRae version of the FS mechanism (which was later called the CalTech Model Mechanism). Plots were shown for these simulations. There was no demonstration of the accuracy of representation of the different HC classes in the mechanism (certainly the new aromatics chemistry was not justified). Nor was there any attempt to address compensating errors among the HC classes (*e.g.*, olefins under-reactive and aromatics over-reactive). The same set of stoichiometric coefficients were used for all mixture compositions which clearly violated the formulation concept described by FS and quoted above. No evidence was presented to show that the McRae version of the FS mechanism could adequately represent the daily composition change expected in urban simulation conditions. Again the model had been insincerely tested, now even ignoring its own call for multifaceted explanation and justification.

At the same time the mechanism testing was described, McRae *et al.* were stating,

Model validity refers to the essential correctness of the model in terms of its representation of the basic chemistry and physics as well as to its accuracy of numerical implementation as measured by adherence to certain necessary conditions such as conservation of mass.

The CalTech Model was criticized by Killus and Whitten in part on the basis that it did *not* adhere to conditions of conservation of mass.

Carbon Bond Mechanism

To deal with the need to have a 'practical' mechanism to predict in the atmosphere, Durban, Hecht, and Whitten suggested in their 1975 report that a new approach, different from that of HSD, and based on explicit chemistry be developed. They said,

After validation, the explicit mechanism can then serve as a basis for rederiving a generalized mechanism and for checking the accuracy of lumping techniques. We should emphasize that the explicit approach is an interim step. We hope that it will serve to further clarify smog kinetics and thus to lay the groundwork for the more practical generalized approach.

Whitten later (1977) added, "A condensed version of the explicit mechanism would combine the advantages of a basis in elementary chemical reactions and speed of computation."

In describing the condensation method—which was called the carbon bond approach—Whitten said,

But the most important difference is the scientific basis: whereas the HSD mechanism was of necessity somewhat empirical, the carbon bond mechanism is derived from the fundamental chemical reactions in smog as represented by explicit mechanisms. Thus the carbon bond mechanism is a condensation of our understanding rather than a parameterization of our uncertainty.

The basic idea put forth by Whitten was that a practical predictive mechanism *ought* not be created directly. He suggested that we have no evidence that such a mechanism is likely to succeed in prediction in applications that exceed its validation data base; in fact, we have evidence that this approach does not work. On the other hand, according to Whitten's approach, an explicit mechanism is supposed to be based on scientific theory. And because it has a scientific representation for each important process thought to occur in air, he believes we have a basis for believing that it can function adequately outside the range of its validation database, *because these mechanistic representations are supposed to have a true correspondence to the real world*. Therefore, if we base our practical predictive mechanism on the explicit mechanism, we can, in turn, model this correspondence to the real world with fewer entities. The basis for testing the latter mechanism is the explicit mechanism from which it was derived, not the real world. A potential dilemma here is that the compressed mechanism may fit the real world data better than the explicit mechanism, but Whitten would have to say, "for the wrong reasons."

Whitten, Hogo, and Killus state in their 1980 paper,⁴⁴

The carbon-bond mechanism is a generalized kinetic mechanism in which most carbon atoms with similar bonding are treated similarly, regardless of the molecules in which they occur....

The carbon-bond mechanism is termed a "condensed" mechanism because it is a condensation of explicit mechanisms developed to simulate smog chamber experiments that were initiated with only one or two hydrocarbon species present....

With the exception of aromatics chemistry for which no explicit mechanism exists, the reactions in the carbon bond mechanism represent essential features of four explicit mechanisms developed by Whitten and Hogo.³⁴

... the aromatic oxidation reactions are assumed to produce the same products as the olefin oxidation reactions ... The mechanism for slow double bonds is an interim solution to a difficult problem, that of aromatics oxidation.

The carbon bond concept eliminates the concept of average molecular weight because carbon atoms undergo reaction independently, so there is no stepwise oxidation of a hydrocarbon chain and there is explicit carbon mass balance.

Whitten *et al.* said, "The carbon bond mechanism was developed and validated with the predictions of explicit mechanisms for smog chamber experiments performed at UCR during 1975 and 1976." This was true for the paraffin and olefin portion of the mechanism. For the aromatics portion, the model was tuned directly against chamber data and Whitten *et al.* said that this portion of the mechanism was "more properly an empirical mechanism than a condensed one." A table of 44 SAPRAC chamber simulations was presented, listing four measures of model agreement with data. Fifteen plots showing model agreement with data were also shown.

Whitten *et al.* made a strong point with regard to the comparison of the mechanism with data,

The carbon bond mechanism is intended to produce predictions similar to those of explicit mechanisms in less computing time. Consequently, the carbon bond mechanism is properly validated with data from simulations of smog chamber experiments using explicit mechanisms rather than with data from the smog chamber experiment itself. In other words, the standard for judging the carbon bond mechanism is the explicit mechanism. A fit between the predictions of the carbon bond mechanism and experimental data that is closer than the corresponding fit for an explicit mechanism must be regarded as fortuitous.

In his grant report, however, Whitten said,

Periodic updating of generalized mechanisms such as the CBM is to be preferred to continuous updating. Changes in one reaction may require compensating changes in other reactions to maintain the overall predictive accuracy of simulations using

the mechanism. Consequently, after a change the mechanism should be tested with an entire set of smog chamber data to ensure that no special problems have been created that would create difficulties in atmospheric applications. The costs of such testing makes it desirable to test the effects of several changes at once.

So here he seems to be suggesting that direct testing against chamber data is used to ensure that no special problems have occurred in modifying the CBM mechanism. I would assume from the former statement, that explicit mechanisms would have to be re-done, and the CBM re-derived, *not* that the CBM would be adjusted against chamber data. I believe that what actually happens is that Whitten initially *creates* the CBM from explicit mechanism, but then makes subsequent adjustments *directly* to the CBM, with the implicit assumption that these adjustments really reflect changes to the explicit mechanism. Therefore the two mechanisms, explicit and CBM, are slightly out of step most of the time. This is actually substituting the model of the model for the model.

A recurring problem for the evolution of CBM to CBII to CBIII was the unknown nature of the aromatics chemistry. Whitten said,

Our need for a mechanism that adequately simulates the fate of the aromatics precludes the option of waiting for an accurate explicit description of the chemistry. Since a validated explicit aromatics mechanism similar in accuracy to those available for propylene and butane does not exist, our efforts to produce a condensed kinetic mechanism for aromatic compounds must be viewed as conjectural.

We analyzed the explicit mechanism proposed by Hendry as an explicit description of toluene oxidation chemistry. The explicit mechanism does not satisfactorily describe ozone formation and limitation. Thus, we are forced to rely on empirical relationships combined with our best speculations as to the true nature of the chemistry involved.

By the 1980 final report, ARO chemistry had been modified several times to account for new information.

Approximately 73% of the oxidation pathway leads to glyoxal photolysis in both the explicit and CBII mechanism. One notable difference has been a prediction, using the new chemistry, that the addition of aromatics to a mixture of olefins at high HC-to-NOX ratios would suppress ozone formation.

In spite of the successful prediction of the suppression effect of adding high concentrations of toluene to propylene-NO_x systems, we now know that the aromatic chemistry in CB-II or CB-III is not correct; we still do not know the correct chemistry.

In his EPA report Whitten said, "... results suggest that more of the uncertainty in a CBM simulation of a smog chamber experiment was caused by deficiencies in knowledge of smog chemistry or inaccuracies in smog chamber data

than by the approximations and assumptions on which the CBM is based.” What has been claimed here is that CBM adequately reproduced the explicit mechanism. We should carefully note that *the explicit mechanism may not itself be an adequate model of the data*. Also, you should remember this paragraph when we later discuss the Leone and Seinfeld rationale for producing their “master mechanism.”

In their conclusion in the 1980 paper, Whitten *et al.* said,

We believe that the carbon bond mechanism offers advantages in realism and ease of application. These advantages include: carbon mass conservation, elimination of assumptions concerning average molecular weights, reduction of the span of rate constants covered by each grouped species, and reduced computing requirements.

... we believe that presently available photochemical kinetic mechanisms are sufficiently developed for most applications to the atmosphere.

We must be careful to interpret the term “realism” used above correctly. Certainly the carbon bond mechanism is *not* a realistic representation of atmospheric chemistry—there are no OLE, PAR, ARO, or CARB species in the atmosphere! The carbon bond mechanism is a model of a model and therefore it suffers any inaccuracies and imperfections of the original model plus additional generalizations, deletions, and distortions caused by its own modeling process. Conceptually at least, the latter should be easy to understand and access because both models are available for testing and comparison. We should also recognize that the modeling of a model process is similar to the original model construction in that pieces are assembled to make an image of the whole. Each explicit mechanism may be developed independently of each other, and while each may work on the relatively small range of chamber conditions used to develop it, the combined effect of all pieces must be tested using experimental mixtures that vary the composition in ways that approximate the urban atmospheric variations. Even at the conclusions of such carefully advancing testing, only those reactions described in the explicit chemistries would have been examined. And so the real world (and all the reactions that could or do exist) still awaits.

Atkinson, Lloyd, Winges Mechanism

Roger Atkinson, Alan Lloyd, and Linda Winges (ALW) in their 1982 paper⁴⁵ present another alternative to the need for a practical predictive mechanism:

It is hoped that an intermediate approach such as this, which is in-between a “very lumped” mechanism and a multi-hundred reaction explicit mechanism, may provide a better “fidelity” as regards the predictions of time-concentration profiles for

a wide range of organic-NO_x photooxidations. Additionally, it is hoped that this intermediate approach will be compatible with the more recent detailed ambient atmospheric analyses, and sufficiently flexible to be readily used for explicit single hydrocarbon/NO_x irradiations simulations.

The "very lumped" mechanisms referred to above were the ELSTAR model; the McRae or CalTech model; and the Carbon Bond models.

ALW said, "This chemical model is based on kinetic, mechanistic and product data from the literature, and the reaction rate constants and mechanisms are generally consistent with the recent evaluations of NBS and CODATA for the simpler species, and with the recent evaluation of Atkinson *et al.*, 1981 for the more complex organics." In addition, ALW made extensive use of the Carter simulations in 1979 explicit mechanisms for propylene and n-butane and of Atkinson's 1980 aromatic simulations. As we shall see in the next chapter, the evaluation of Atkinson *et al.* referred to in this paper would have a significant influence on future mechanism development.

The ALW mechanism had *twelve* HC classes: two types of alkanes; three types of alkenes; three types of aromatics; three types of aldehydes; and one type of ketone. The low reactivity in each class was an explicit compound, *e.g.*, propane, ethylene, formaldehyde, benzene, while the higher classes tended to be average representations of several species.

The mechanism was "tested"—validate or evaluate was never used—against 21 UCR smog chamber irradiations. These tests included: one each of six individual species butane, ethene, t-2-butene, propene, toluene, m-xylene; a single four-alkene mixture irradiation; three seven-hydrocarbon mixture irradiations; seven multicomponent hydrocarbon irradiations; and four hydrocarbon-NO_x-SO₂ irradiations. A continuous HO· radical source that varied with each run was used in the simulations. Fifty-seven comparisons of data with prediction plots were included in the paper. ALW said,

As can be seen, [from the plots] the fits are generally very good, being typically within $\pm 20\%$ of the experimental values for O₃, intermediate carbonyls and PAN.

...

In general, the fits of calculated to experimental data for the evacuable chamber multihydrocarbon-NO_x-air irradiations were good, as may be expected from the generally excellent fits of the single hydrocarbon-NO_x-air irradiations. For the All Glass Chamber, the fits were again judged to be reasonably good, since the hydrocarbon, initial NO and NO₂ and O₃ profiles are accurately predicted. ...

Thus, the reaction mechanism has been tested against 21 smog chamber irradiations, yielding good fits between the experimental and calculated time-concentration profiles.

No assertions were made based upon the last sentence quoted above. In the conclusions, ALW said,

The new mechanism has been tested extensively against a variety of smog chamber data for single hydrocarbons, binary, seven component and surrogate hydrocarbon data. The mechanism gave good agreement between predicted and measured time-concentration profiles for all measured species. *[A set of simple reductions]* produces a viable chemical mechanism for inclusion into air quality simulation and reactive plume models.

What is a "viable chemical mechanism"? Viable means³³ "able to live; specifically at that stage of development that will permit it to live and develop under normal conditions." If we take 'live and develop' to mean 'function' then we can interpret the conclusion to mean "The mechanism is at a stage of development that will allow it to function under normal conditions in an air quality simulation." Compared to some other modeler's conclusions this can be seen as a modest statement.

A Contrast

I am somewhat surprised at how the outcomes of testing the mechanisms discussed in this chapter were described as being so good and at the strengths of claims made by the authors for effectiveness in atmospheric prediction in contrast to the "uncertainty" and lack of explanation attributed to the detailed explicit mechanisms in the last chapter.

Mechanisms Compared

By 1980, EPA policy makers were aware that they were depending on "known incorrect chemistry" in the Dodge Mechanism (yet it predicted correctly, didn't it?), and with so many mechanisms making claims of goodness of prediction, they wanted to know how to treat new State Implementation Plan revisions that used different mechanisms.

Until this time, each mechanism had only been tested by its authors and had never been compared with each other. Jeffries, Sexton, and Salmi (JSS) at UNC

were given a contract⁶ to "examine the effects of input choices and modeling methods likely to be used in 'Level II' and 'Level III'-type SIP applications on the predicted ozone precursor control requirements." A Level II analysis required a simple trajectory model incorporating effects of pathway, emissions, mixing height changes, entrainment from aloft, and chemistry be used to simulate days of high O₃ formation.

Four different chemical mechanisms were examined by JSS (Dodge Mechanism, Carbon Bond II Mechanism, Demerjian Box Model Mechanism, CalTech Mechanism). In setting the scope of work, it was argued by EPA that, because the Dodge Mechanism had been successful in simulating the BOM data base, the other three mechanisms would likewise have to be tested with the same data base. First, the BOM data base was the only reasonably complete auto-exhaust data base. Second, because of the empirical nature of the 75:25 butane:propylene split, the Dodge Mechanism could not be used to simulate other HC mixture systems. Yet, the newer mechanisms all made at least an implicit claim that they could simulate various HC mixtures, including urban atmospheric mixtures, which EPA assumed were well represented by dilute auto-exhaust such as that in the BOM chamber.

The 31 BOM irradiations were simulated with each of the four mechanisms. A single consistent set of chamber assumptions were used for all mechanisms (*e.g.*, the same photolytic rates, the same auto-exhaust composition, but uniquely represented in each mechanism). From a plot of observed O₃ vs. predicted O₃ for each mechanism, it appeared that each mechanism showed "reasonable agreement" within $\pm 25\%$, although the CalTech mechanism had several low predictions. Comparison of cross-sections of isopleth diagrams with all four mechanisms plotted on one cross-section, however, revealed differences among the mechanisms. For example, for the CalTech mechanism, predictions for higher NO_x were very low compared to the other mechanisms and to the experimental data. Inspection of the CalTech mechanism revealed several major differences in its use of coefficients for olefin and aromatic products compared to the other mechanisms. The Carbon Bond and Dodge mechanisms' predictions converged at low HC, while the Carbon Bond and Demerjian mechanisms' predictions converged at high HC. JSS said, "Compared to the experimental points, however, it would be difficult to say which mechanism is a better description of the BOM results." They further concluded that "the 'noise' in the experimental results is greater than the differences shown among the Dodge, Demerjian, and Carbon Bond II mechanism simulations which have large differences in construction assumptions." And finally,

Given the range of uncertainty in the BOM data (e.g., O₃ maxima and initial aldehydes) and the sensitivity of mechanisms to necessary assumptions, it is concluded that the Demerjian, Dodge, and Carbon Bond II mechanisms were all capable of providing adequate descriptions of the O₃ formation in the BOM chamber; the Carbon Bond II mechanism is probably slightly better at describing the other measured secondary products.

This was the good news. Now for the bad. In atmospheric simulations, the models predicted significantly different control requirements for the three days on which the meteorological conditions were judged to be reasonably represented (i.e., of the 10 RAPS days selected by EPA, the trajectory calculation methods were so poor on seven of the days that CO could not be reasonably predicted using a meteorology and emissions-only trajectory model, and thus a comparison of chemistry predictions for these days would have been meaningless). In some cases for the three good days, the atmospheric O₃ isopleth diagram produced by the Carbon Bond and CalTech mechanism did not allow an EKMA-type solution to the control requirement. That is, HC-to-NO_x ratio line for the day did not intersect *any* isopleth at all! This was a failure of the EKMA approach itself. For the conditions in which the EKMA approach did produce a solution, the different mechanisms exhibited very different $\Delta O_3 / \Delta HC$ for the different atmospheric simulations and therefore they predicted very different HC control requirements: the range of control requirements was from 15% to 70% HC reduction needed depending upon the day, the mixing height profile, and the mechanism used. The most important factor was the day with its O₃ aloft assumption; the second most important factor was the choice of the mechanism.

JSS conclude, in part,

The chemical mechanisms used in this study are probably inadequate descriptions of the real atmospheric processes because they were not developed by comparison with data bases which included the type and range of variations which apparently occur in the atmosphere. That is, none of the mechanisms have been tested against data bases in which the hydrocarbon composition, including the aldehyde mole fraction, had been systematically varied, and in which dynamic dilution and injection of new reactants had occurred, and in which dynamic addition of O₃ such as from entrainment aloft had occurred, or in other factors similar to atmospheric processes had been used. These data bases do not exist ... It will require years to develop the needed databases to insure that the chemical mechanisms are giving adequate descriptions of the atmospheric chemical processes.

In a Kuhnian interpretation, this would be equivalent to saying that an insufficient amount of the second class of normal scientific work had been done to justify the extension of the paradigm to the atmosphere. That is, the "solution

to the puzzle" had been forced—yet, until this point, accepted by the community. JSS also suggested that the method used by EPA to 'correct' the model's predictions when they did not agree with the real world, was a major cause of the prediction problem. That is, EPA used the mechanisms' predictions in a relative way, assuming that they would perform better that way. In fact, JSS said it made them perform worst.

It is little wonder then, that the JSS study had an impact on the air pollution community. Reviewers at the EKMA conference in 1981 were so affected by its implications that they suggested that EPA abandon the modeling approach. These folks surely felt that the modeling paradigm had failed to provide puzzles with solutions. According to Kuhn, this should have led the modeling community to "either abandon the puzzle entirely or attempt to solve it with the aid of some other hypothesis." Yet several research groups were attracted to the puzzle presented by the fact that the mechanisms could apparently fit chamber data, and still they gave different atmospheric predictions.^{46,47,48,49,50} Some researchers sought explanations in the mechanism details as to why they should give different predictions.

For example, in their EPA report⁶ Leone and Seinfeld state,

These mechanisms are all based on the same body of experimental rate constant data, and each mechanism has been evaluated against data from various smog chamber facilities. In each mechanism, the detailed atmospheric chemistry has been greatly simplified by a process referred to as lumping. However, because this lumping has been carried out in different ways, no two mechanisms are the same.

Their approach was therefore to compare each mechanism with "the detailed atmospheric chemistry," as represented by a fully explicit mechanism, as a way of "understanding the fundamental reasons for the discrepancies."

I believe, based on the conjunction of "the same body of experimental rate constant data" and "the detailed atmospheric chemistry has been simplified," from the quote above that there was a presupposition to the Leone and Seinfeld hypothesis. They presupposed that if asked to produce a "full explanation," the model developers would have produced the *same* explicit mechanism, and therefore it was only the various lumping methods that were the cause of different predictions. I believe that this presupposition was incorrect and that furthermore, it led them to make the same mistake whoes cause they were seeking. It was *not* lumping that was responsible for the mechanisms not being the same. Even the incomplete and brief history presented here shows that "the same body

of experimental rate constant data" only represents a small part of the knowledge needed to construct a mechanism—the modeler's creativity and ingenuity are fully operational in constructing his mechanism. A great deal of what appears in the mechanism is the modeler's *personal* conjecture; if not, mechanisms could be constructed by computers. Leone and Seinfeld presumed that there was a *single* explicit mechanism available in the "body of experimental rate constant data" that represented the truth (they say "the detailed atmospheric chemistry"), and that this represented all that modelers had started with in their model construction. From the subsequent actions of Leone and Seinfeld, it is clear that for them this truth was represented by information given in the Atkinson and Lloyd survey (which we will discuss in Chapter 9). Thus, they confused the Atkinson and Lloyd survey of kinetics information with a mechanism. Leone and Seinfeld therefore created yet another mechanism using their own 'creativity and ingenuity' and called it the "master mechanism"—their version of the truth against which others were measured. Their solution to different mechanisms predicting different O_3 for the same input was to change all the mechanisms to be consistent with their "master mechanism." Then the difficulties mostly go away.

*What is true by lamplight
is not always true by sunlight.
—Joubert*

Response to Uncertainties

At the end of Chapter 6, Second Explanation, I suggested that the late 1970's brought model development and testing to what might be called the HO-paradigm's first crisis—the impression that some modelers were adjusting their models in such a way that it was not clear that the values they produced were the correct ones to enter into the theory, and that other modelers had a vagueness in their theories and an unanalyzed component in their apparatus. I suggested that this perception began to undermine the promise that had come with the adoption of the HO-paradigm.

The two major modeling groups responded to these impressions in different ways. Whitten at SAI put forth principles of model development and testing that “clarifies the sources of uncertainty in simulations” and therefore restricts the amount of ‘knob-twiddling’ possible. And Carter at UCR performed a large number of chamber simulations directed at explaining the large HO background source in the SAPRAC EC chamber.

Whitten's Hierarchy of Species for Testing

In his 1982 EPA report, Whitten said, “A principal goal in computer modeling of smog chemistry is to develop a set of reactions and rate constants that provides the closest possible agreement between simulations and measurements for a series of experiments.” That is, the basic concept is to decrease the margin of “reasonable agreement” between theory and facts. Whitten then described this process as:

- A. Use measurements or estimates for all important reactions, products, and rate constants known or expected to occur in the system of interest, to formulate a kinetic mechanism. While published data on reactions and rate constants should be used where possible in constructing the mechanism, we

should also recognize that all mechanisms contain hypothetical reactions or estimated rate constants.

- B. Estimate the physical conditions that applied for the experiments performed.
- C. Combining these two, simulate the smog chamber experiments using a computer.
- D. Modify the mechanism by adding reactions, products, and rate constants within their limits of uncertainty until satisfactory agreement between simulations and measurements is achieved. However, there are many constraints that must be met:
 - ▷ common reactions must have the same rate constants in all experiments.
 - ▷ chamber-dependent effects should be consistent.
 - ▷ precursor decay must be simulated correctly.

I believe that Whitten's most important concept was that a complex mechanism should not be constructed in one operation, that is, the building of a mechanism should proceed in a stepwise fashion, and the order of building is given by the concept of a *hierarchy of chemical species*. That is, each species in oxidant formation can be assigned to a hierarchical level on the basis of the number of HC/NO_x systems in which it occurs, with the most ubiquitous species occupying the lowest level. Thus CO is lower than HCHO, which is lower than CH₃CHO. A slightly different version of Whitten's hierarchy is shown in Figure 1.

Whitten said,⁵¹

The value of the hierarchical concept is the principle it suggests in validating explicit mechanisms for a series of HC. In brief, one should validate mechanisms for the species at the lowest level first. ... Validating mechanisms in the stepwise order suggested by the hierarchical levels clarifies the sources of uncertainty in simulations.

In validating each successive mechanism, one can focus attention on the added reactions and rate constants because the other reactions and rate constants have already been validated. Following this procedure reduces the possibility that a complex mechanism, such as that for propylene, contains errors that compensate for each other in simulations of a set of smog chamber experiments. ... Another advantage for this approach for validating each mechanism is that it minimizes the possibility of fortuitous agreement between simulations and measurements. A valid kinetic mechanism, unlike a mere curve-fitting exercise, should give reasonable predictions when used in applications such as atmospheric modeling that are outside the range of conditions and smog chamber experiments for which it was developed.

In following this guideline, one would first make sure he could simulate experiments that only contained the inorganic species, *e.g.*, NO_x -air and NO_x -air-CO experiments. In the most ideal case, one would use NO_x -hydrogen peroxide (H_2O_2)-air systems in which the photolysis of H_2O_2 served as a radical source to further test the inorganic reactions. Then, *without any changes in rate constants or reaction content for the already tested part of the mechanism*, he would add and adjust only those reactions and rate constants needed to simulate NO_x -HCHO-air and NO_x -HCHO-CO-air experiments. If, in simulating these latter experiments, difficulties were encountered that required the addition of new inorganic reactions, say for example, O-atom reactions that might have been lumped out, then the NO_x -air and NO_x -air-CO experiments would have to be simulated over again to verify that no fortuitous agreement between simulations and measurements had occurred. Next, one would simulate experiments of NO_x -PAN-air without adjusting the HCHO, NO_x , and CO mechanism. PAN makes HCHO, so all the reactions of HCHO and the inorganic reactions are needed to simulate PAN correctly. Likewise, PAN must be correctly simulated to successfully simulate CH_3CHO . The mechanism is thus built-up in a stepwise fashion, one new species at a time, and only allowing the new chemistry to be adjusted. If the mechanism fails to fit a given species, then either the kinetics data for this species is lacking or some error exists in the experimental data for this species, *but, since all lower species data have already been simulated successfully, it is most likely just this species that is in error and not the whole mechanism.*

We should contrast this approach with how the original HO^\cdot -paradigm smog chamber data were produced: propylene and n-butane were the first species for which new smog chamber data were available. We have already seen that several models were only tested on a few propylene and a few n-butane experiments. Propylene is quite high in the species hierarchy. To model propylene successfully, you have to model not only the HO^\cdot and O_3 attacks on propylene, you have to model acetaldehyde, PAN, formaldehyde, and the inorganic reactions correctly. How did these modelers know that they did not have off-setting errors between say the HO^\cdot -propylene rate constant and the photolysis rate of HCHO? The answer is "They did not." It was not until after the original explanatory models 'failed,' that smog chamber runs using the higher species's products as initial reactants were performed, thus allowing a hierarchical development approach to occur.

In Kuhn's terms, testing by Whitten's hierarchy of species method clarifies the juxtaposition of fact and theory, it narrows somewhat the macroscopic facts of the smog chamber so that they more clearly confront the theory. In other words,

it leads to an improvement of "the measure of 'reasonable agreement' characteristic of the theory in a given application."

I think, however, that modelers, who think of themselves as good puzzle solvers, while expressing support for this concept, do not use it as described. That is, a) having "solved" the chemistry of, say a particular chamber's wall-effects once and having moved on to successfully model HCHO , CH_3CHO , and propylene, and b) upon being told by the chamber operator that new evidence is available on wall effects which suggest that the technique used by the modeler to represent these is "not quite right," the modeler is not interested in solving the puzzle again, (that is, in starting the whole cycle over again and re-testing the current mechanism in successive stages). Instead, assuming that, since the original model worked there cannot be any real problems with the current version, and so he just makes a "patch" to the existing mechanism and proceeds to solve some new puzzle, which is more fun anyway. In addition, the modeler has used the previous mechanism to produce several hundred simulations and graphs for his report and is three-fourths through the total project time and money—he literally can not afford to repeat all the tests! He also can not conclude that his model, the one that he is sticking with, may be wrong. Thus, the more testing that is done with a given mechanism, the less likely it is that a flawed mechanism will be changed, especially if the flaw is in a lower hierarchical species such as formaldehyde, the product of all organic oxidation. The very success of the mechanism rules against its being adjusted in spite of potential anomalies in the lower hierarchical species.

Whitten must have been thinking along these lines when he said,

In computer modeling studies such as this one, many ideas are tried, and large quantities of computer output are produced. In the description of the activity that produced the current closest agreement between simulations and observational data, the implicit conclusion is that the steps taken were both unique and necessary. However, experience has shown that equally close agreement is possible from several combinations of adjustments to physical conditions and mechanisms. Hence, the conclusions presented here must be qualified with the caveat that the results are subject to change in accordance with new data and further modeling efforts.

Whitten also offered guidance on how to proceed within a given species simulation as well. He said that his approach was based on the following:

- The first measurement that must be reproduced with acceptable accuracy are those related to the consumption of the initial precursors. Ozone development and other manifestations of the experiment must depend on the products that result from decay of the precursor hydrocarbons and NO_x .

Figure 1. A Hierarchy of Species

Good agreement between measurement and simulated ozone concentrations, coupled with poor agreement for hydrocarbon decay, is indicative of compensating errors in the kinetic mechanism. Errors that compensate one another under the conditions of a particular smog chamber experiment are not likely to do so under atmospheric conditions.

- In simulating a series of experiments in the same chamber, chamber-dependent effects must be treated consistently. If a light source is assumed to emit progressively lower amounts of short wavelength radiation over a period of several months, the photolysis rate constants for the series of experiments must diminish in accordance with the order of performance of the experiments. Arbitrary adjustments for such effects must be avoided.

Whitten's clear understanding of the cause-and-effect in the mechanisms leads to this advice. First the HC and NO must be converted at the correct rates to the intermediate products that then must maintain the reactivity of the system to produce the final products, usually PAN, HNO₃, O₃, CO, and some aldehydes and ketones. As we have seen, however, "acceptable accuracy" and "good agreement" are ill-defined concepts. More importantly, in the SAPRAC EC chamber, obtaining correct rates for the initial HC lost rate requires assuming a relatively large "unknown" chamber source of HO[•]. Thus, an 'ad hoc' technique was needed to obtain "good agreement" between theory and data. This brings us to the second response to the crisis.

Carter's Chamber Radical Experiments

In Carter *et al.*'s 1979 mechanism paper,³⁹ they said, "The largest uncertainty concerns the necessity to include in the mechanism a significant rate of radical input from unknown sources in the smog chamber." In the UCR models this input was 0.2 to 0.7 ppb/min of HO[•]; this is 0.2 ppmV of direct reactivity in a simulation (i.e., 0.2 ppmV out of say 0.5 ppmV of toluene could be converted into radicals throughout a 400 minute irradiation by an 'ad hoc' specification). To complicate the problem, different modelers using the UCR EC data were representing the needed radical source in different ways: Falls and Seinfeld and Whitten used only initial HONO as a radical source; the UCR group were using a constant source of HO[•], and Hendry *et al.*³⁷ were using a combination of the two methods. Initial HONO will decrease rapidly in the first 30-60 minutes and will make little contribution later in the experiment, whereas the constant HO[•] source will make a relatively smaller contribution initially, but will make a much larger contribution later in the experiment. Carter *et al.*⁵² said

Clearly, aspects of the photochemical mechanism relating to radical initiation and termination processes cannot be unambiguously validated using smog chamber data until this presently poorly characterized radical source is elucidated.

In three short pages, Carter *et al.* presented their 'fit' to a series of chamber experiments designed to investigate the radical source. They said, "detailed descriptions of the experimental procedures will be given in a subsequent publication." The radical source was described as

$$\text{radical flux} = k_1(0.30 + 2.9[\text{NO}_2]) \quad \text{ppb/min}$$

They went on to say,

It is significant to note that this equation predicts radical fluxes very similar to those used in previous computer modeling studies from this laboratory, thus further validating the chemical models developed in those studies.

The results reported here show conclusively that radical input from unknown sources is an important process in smog chamber systems, and that initial HONO is, at most, a minor contributor to this process for typical chamber irradiations of ≈ 6 -h duration. Thus it is clear that photochemical smog models validated against chamber data assuming only initial HONO as the radical source must be reevaluated. However, it is probable that use of a constant radical source during an irradiation is also an oversimplification, and it is clearly essential to understand the chemical and physical processes producing this effect in environmental chambers, and whether similar processes also occur in the atmosphere.

In comments on the Carter *et al.* paper, Killus and Whitten⁵³ first argued that a large number of factors had not been addressed in the work and that Carter

et al. had over-estimated the magnitude of the radical flux. In addition, Killus and Whitten refuted the assertion quoted above,

Carter *et al.* call for a "reevaluation" of only those modeling studies, like ours, which have used initial HONO. Carter *et al.* do not recommend reevaluations of those modeling studies which have used a constant radical source, even though they do admit this is "probably an oversimplification." We suggest that a reevaluation of prior modeling efforts using a chamber source that varies continuously with NO₂ concentration would result in a much larger discrepancy than can be subsumed under the word "oversimplification."

This exchange was followed by an extensive paper by Carter *et al.*⁵⁴ This paper had undergone a number of revisions caused by reviewer comments and successive refinements in the data treatment procedures. The final radical flux in the EC was

$$\text{radical flux} = k_1(0.39 + 1.37[\text{NO}_2]) \quad \text{ppb/min}$$

which is nearly a factor of two less dependence on NO₂ concentrations compared to the first equation proposed above. Most importantly, Carter *et al.* found that the indoor and outdoor Teflon chambers, and new Teflon bags produced data that suggested much lower radical sources (only 30–40% of the EC chamber) and showed little dependence upon NO_x.

Killus and Whitten⁵⁵ again argued that some of the UCR data were suspect because of other chamber background or newly characterized wall processes such as the heterogeneous formation of HONO from NO₂ as reported by Sakamaki *et al.*⁵⁶

The following year, Pitts *et al.*⁵⁷ confirmed the results of Sakamaki *et al.*, and extended the work to both the atmosphere and other UCR chambers such as the indoor Teflon chamber. They confirmed by *in situ* spectroscopic measurements that all chambers had a continuous conversion of NO₂ to HONO, but that the rate and gas phase yield of HONO was different in the EC and the Teflon chambers with the Teflon chambers being less reactive than both the SAPRAC and NIES EC-type chambers, that the formation in the EC chamber was partially a function of its previous history, and that this process was a factor of 10 less than the source determined by Carter *et al.* for the EC chamber. Pitts *et al.* say that

These facts strongly suggest that the interaction between NO₂ and water and the chamber walls, which are responsible for dark HONO formation, could also be involved in the production of radicals during the irradiations. However, the details of this process remain elusive at present.

In their latest work, Carter *et al.*⁵⁸ show that their preferred treatment for the EC chamber is still to use a large radical source that depends upon NO_2 (e.g., for $k_1 = 0.3$ and $[\text{NO}_2] = 0.5$, the source would be 0.44 ppb/min plus 0.14 ppb/min from NO_2 +walls), but to use much smaller radical sources that do *not* depend upon NO_2 for the UCR Teflon chambers and the UNC Teflon chamber (e.g., 0.09 to 0.06 ppb/min plus 0.014 ppb/min from NO_2 +walls).

Thus, Carter is comfortable that he can at least predict the radical source even if he can not completely explain it. He showed that there are experimental alternatives, that is, other chambers with lower radical formation rates that can be used to check results obtained in the EC chamber, thus the models do not have to ultimately rest upon "unknown sources."

The image has a certain dimension, or quality, of certainty or uncertainty, probability or improbability, clarity or vagueness. Our image of the world is not uniformly certain, uniformly probable, or uniformly clear.

Messages, therefore, may have the effect not only of adding to or reorganizing the image. They may also have the effect of clarifying it, that is, of making something which previously was regarded as less certain more certain, or something which was previously seen in a vague way, clearer.

Messages may also have the contrary effect.

—Kenneth Boulding

Third Explanation

Atkinson and Lloyd Review

What started as a review article to support the development of the Atkinson, Lloyd, and Wings mechanism in 1982, became in and after 1984 the standard for all explicit mechanisms in the third phase of explanation.

Although the review article was referenced in the ALW mechanism paper as "in press 1981," it did not actually appear until 1984. According to Marcia Dodge (EPA paid publication costs), the Journal of Physical Chemistry Reference Data had difficulty publishing a 330 page article in one issue. During the delay in both review and in finding an opening in the journal, new and important information became available which required revisions, and so publication was delayed for these revisions. Typesetting such a complex manuscript—and proofing the accuracy of the numerical information—was extremely difficult. In fact, only the first two pages and a few tables were typeset, the rest were photocopied from a typewriter-style printer. Some modelers had pre-print versions of the review at least two years before it was published.

The purpose of the review was to perform "a critical evaluation of the rate constants, mechanisms, and products of selected atmospheric reactions of hydrocarbons, nitrogen oxides, and sulfur oxides in air" using the literature through early 1983. In the Introduction, the authors said,

It is evident that, along with laboratory investigations, ambient atmospheric measurements, and computer modeling development, there must also be a continuing effort to critically evaluate the rates, mechanisms, and products of the relevant chemical reactions and to update these evaluations as new experimental data become available. Such a continuing evaluation for stratospheric reactions has been very successfully carried out mainly through NBS and more recently NASA. However, such evaluations, although dealing very thoroughly with the inorganic reaction systems,

generally consider only the relatively few organics of specific interest to stratospheric problems. ...

Thus there is a critical need for a comprehensive evaluation of the gas phase chemistry associated with the more complex organics encountered in photochemical air pollution systems.

... The results of this evaluation are intended to aid modelers in developing chemical mechanisms to describe smog chamber experiments and to ensure that the best available data are used in the modeling study and to limit any tendency to "curve fit" the data. While we consider this report to be as current as possible, we believe it is very important that the report be updated at regular intervals and expanded to other chemical compounds so that maximum benefit to model development is provided.

... the majority of this report deals with reactions which have not been previously covered, but which are critical for model development. ... estimates are needed if the modeling technology is to improve. Thus, in the more complex chemistry ... much use has been made of thermochemical arguments and analogies with other related systems, since in many cases no unambiguous experimental data are available. Hence in such cases only discussions of related chemistry and suggested reaction rates, pathways, and mechanisms can be presented, rather than firm recommendations. These suggested rates and mechanisms should be regularly evaluated and upgraded as new experimental data becomes available.

In our philosophical terms, Atkinson and Lloyd were saying,

"Here are the auxiliary statements, including some speculations, that are needed to connect the theory of air pollution chemistry to the facts *as we know them in 1983.*"

It is also important to identify what was not said by Atkinson and Lloyd. They did *not* say, "Here is a mechanism for explaining smog chamber data." No simulations were included. There is no assertion as to completeness or adequacy of the description. This can only be done by comparison of the theory with facts and a demonstration of "reasonable agreement." As Kuhn said, without tables (or in chamber modeling, the plots) comparing the theory and the data, "the theory would be essentially incomplete." The Atkinson and Lloyd review can be seen as representing a first class effort to do Kuhn's first class of normal science work, that is, the production of both empirical and theoretical information that the paradigm has shown to be particularly revealing of the nature of things. It is necessary work that must be done before any significant second class of normal scientific work can be undertaken, that is, the direct comparison of factual determinations with predictions from the paradigm theory, in the same way that smog

chamber experiments must be done before they can be compared with the theory predictions.

Thus the Atkinson and Lloyd review may be seen as presenting us with the puzzle parts, but, without chamber data, we do not know what the assembled puzzle parts are supposed to look like—there may be some pieces still left in the box.

Leone and Seinfeld Master Mechanism

At the end of Chapter 7, I described the reasons given by Leone and Seinfeld⁷ (LS) for wanting a detailed explicit mechanism—it was seen by them as representing the truth (LS said “the detailed atmospheric chemistry”) that all modelers had started with in the construction of their individual models used for air quality predictions, and thus the explicit mechanism was seen as a sound basis of comparison of different model’s predictions. Recall that I denied that the lumping process was the cause of different mechanism predictions. I asserted that each modeler held a slightly different image of how to represent the real world given the uncertainty in both the theoretical and experimental data and thus each modeler would not have produced the same explicit mechanism as was assumed by LS.

Although it is clear that the review of Atkinson and Lloyd was used to construct the LS “master mechanism,” the actual formulation process was never described, therefore we have no insight into how LS turned the “many cases where no unambiguous experimental data are available” and were therefore discussed, not recommended, by Atkinson and Lloyd into a whole mechanism. They simply asserted the mechanism in a listing. LS then say,

The evaluation of mechanisms with smog chamber data is not the object of this report. In fact, it is our feeling that with the current controversy regarding the magnitude of chamber radical sources, it is difficult to consider any mechanism validated simply because it can simulate chamber data. However, because the master mechanism plays an important role in this study, we feel it is useful to include the results of simulating a series of experiments performed in the evacuable chamber at SAPRAC. ...

In nearly every case the agreement among experiment and mechanism predictions is excellent. One should keep in mind the master mechanism contains no adjustable parameters. Similarly, the chamber dependent terms are held constant and not varied as the initial conditions of the experiments change.

I think that this comes close to saying,

We do not have the time or effort to evaluate the master mechanism with smog chamber data. But the modeling community will never accept the simple assertion of theory for a mechanism so we are going to do a token evaluation by showing agreement between predictions from our mechanism and two (and after review, 11) irradiation experiments with 7-component HC mixtures. On the one hand, we have real doubts that mechanisms can be validated with chamber data, and on the other, you can believe our mechanism because with no adjustments it fits the chamber data.

Within a few months of finishing the EPA report, LS published a paper⁵⁹ on the toluene mechanism that was a part of the "master mechanism." In this paper LS used comparisons of predictions with chamber data "to evaluate this mechanism." Although they mentioned the "controversy in previous attempts at simulating aromatic hydrocarbon" systems due to chamber derived radical sources, they did not express doubts that they could evaluate their mechanism with such data as they had suggested might be the case in the EPA report. They accepted Carter's fit to the radical source in the EC chamber without modification.

Nine toluene and two cresol experiments were simulated. No experiments were simulated for any of the theoretically predicted fragmentation products (*e.g.*, glyoxal, methyl glyoxal, pyruvic acid); in other words, a hierarchical testing scheme was not followed and the mechanism was used to fit a few pieces of data for the whole system (*e.g.*, NO_x , O_3 , PAN, HC-decay). In the paper, 39 plots of model predictions and experimental data were given. From these one can ascertain a some measure of "reasonable agreement" for the major species only.

LS conclude,

The chemical mechanism for toluene photooxidation which is developed here contains the latest available kinetic and mechanistic data, and has been extensively tested against smog chamber data. The new mechanism gives good agreement between the predicted and observed concentration-time profiles for all the major species in both toluene- NO_x and toluene-benzaldehyde- NO_x experiments ... The agreement between theoretically predicted and experimentally observed oxidant yields is especially good.

The new mechanism has been used to gain insight into areas of toluene photooxidation which are not completely understood. At present the major uncertainties in the mechanism center around the methyl glyoxal and benzaldehyde photolysis products, the exact ring fragmentation pathways, and the aerosol forming ability of some of the oxidation products. Although further experimental information on any of these items would be very useful and should be given a high priority, we feel that the mechanism presented here can be used confidently in its current form, and can also serve as a useful starting point toward the development of accurate lumped mechanisms for modeling gas phase aromatic hydrocarbon photooxidation.

To recommend something confidently means,³³ "with strong assurance; without doubt or wavering of opinion." I have great difficulty reconciling this conclusion with the position taken a few months before in the EPA report, "In fact, it is our feeling that with the current controversy regarding the magnitude of chamber radical sources, it is difficult to consider any mechanism validated simply because it can simulate chamber data." Apparently LS somehow resolved their concern over the magnitude of chamber radical sources, so that it became possible for them to recommend use of a mechanism validated with smog chamber data without doubt or wavering of opinion.

Yet, if this conclusion is correct, then the rationale Leone, Flagan, Grosjean, and Seinfeld present in their 1985 paper⁶⁰ contradicts it. In this paper, Leone *et al.* say,

The SAPRAC chamber has been a prime facility for generating data for evaluating photochemical reaction mechanisms. However, it has been shown that reactions carried out in this chamber are influenced by a significant source of free radicals from chamber walls. For this reason, it is useful to continue to evaluate our understanding of toluene chemistry with data taken from another smog chamber. In this work we present the results of such an evaluation using experimental data from a recently constructed outdoor smog chamber facility that has been carefully characterized with regard to all of the important chamber parameters, including wall radical sources.

So in spite of their confident recommendation there must have been lingering doubt—that is they *wanted* to conclude that the mechanism was correct, but they were not *sure* that it was because of the chamber radical source. As we shall see shortly, there was reason for doubt, but it had nothing to do with the chamber radical source problem.

In this 1985 paper, simulations of six toluene-NO_x outdoor irradiations were shown. A total of 23 plots showing model predictions and experimental data were presented. The fits of model predictions to observations (data included NO_x, O₃, PAN, CO, HC-decay, benzaldehyde, and cresol) were as good as those for the UCR EC chamber experiments.

At the end of the paper, after the quality of the model had been demonstrated, Leone *et al.* said,

Since the mechanism shown in the Appendix closely simulates the data from both the SAPRAC evacuable chamber and our outdoor chamber, one might conclude that it represents a fairly accurate description of the actual toluene photooxidation process. However, Tuazon *et al.* have recently reported yields of methyl glyoxal from toluene that are much smaller [10–17%] than those predicted by our mechanism [72%] or, for that matter, by any of the existing toluene mechanisms.

... Even though the mechanism developed in Leone and Seinfeld has now been evaluated successfully against experimental data from two environmental chamber facilities, there is concern because this mechanism is not consistent with some very recent experimental findings concerning dicarbonyl yields. However, no reasonable sequence could be established that makes use of these new data, and, at the same time, gives reasonable predictions of the observed concentration-time profiles in toluene-NO_x irradiations. Thus, it is clear that additional experimental work is needed ...

Return to Chapter 4 and read the anti-realist objections. We have an example here of "Scientific theories, however well secured by observation and experiment, are inevitably fallible." While I know that there are methodological improvements we possibly could use to avoid the types of problems encountered by Leone *et al.*, I, too, do not know how "to resolve the basic dilemma of the underdetermination of theory." Their conclusion was that additional experimental work was needed.

Could it be that other parts of the "Master mechanism" might suffer a fate similar to the toluene part?

In fact, there are hints that the ethylene chemistry described in the Atkinson and Lloyd review is incomplete (underdetermined). First, in July of 1984, Lurman, Lloyd, and Atkinson updated the information in the review and updated the ALW mechanism for use in ADOM. They tested this updated-ALW mechanism with a series of 15 irradiations from the EC chamber comprising single HC as well as mixtures. They were able to simulate successfully all but two irradiations with *half* the Carter wall source strength. The two irradiations were ethylene experiments (at 2 and 4 ppmC) and these required 1.5 times the Carter wall source strength and 10 ppb HONO to speed up the simulations so as to have the timing and O₃ yields fit the data. In addition the HCHO predictions were not only poor, they were biased opposite on the two runs. Lurman *et al.* said,

... the requirement for the higher wall source strength to simulate the experimental NO₂ and ethene data suggest the chemistry of ethene may not be as well understood as once thought. ... the ethene chemistry incorporated into the mechanism may be slower than the actual chemistry and further research is needed on the fate of the product species.

In UNC's testing of the SAI explicit ethylene, Carbon Bond X, and Carbon Bond Four mechanisms, we independent came to exactly the same conclusions as Lurman *et al.* To fit our outdoor irradiations, Whitten used large wall radical source strengths and unreasonable values of initial HONO. When we repeated the simulations with values that were consistent with our chamber characterization experiments, the simulations were too slow and un-reactive. In addition, HCHO is

greatly underpredicted. We have good agreement between the model and measurements for a nighttime O_3 + ethylene experiment, including HCHO.

In his 1986 study,⁵⁸ Carter modeled six EC ethylene irradiations and six UNC ethylene irradiations. The variability in HCHO predictions compared to the data was the worst of any of his 400 simulations. He tended to underpredict UNC HCHO measurements by 50%. He suggested that our instrument was subject to interferences and was therefore over-reporting the HCHO data. In the summer of 1986, an intercomparison study of HCHO was conducted at the UNC chamber using four different methods. Preliminary results suggest that the standard UNC method does not over-report HCHO data, and specifically in the case of ethylene irradiations.

Thus there is increasing experimental evidence that we not know all of the chemistry of ethylene. Atkinson and Lloyd said,

“...there must be a continuing effort to critically evaluate the rates, mechanisms, and products of the relevant chemical reactions and to update these evaluations as new experimental data become available.”

I would suggest that this is also part of a way to ensure that the needed experimental data *do* become available. If we all believe that there are no puzzles to solve in the ethylene chemistry, then we have no cause to look for interesting and exciting problems there.

The Essential Tension

In the beginning I said that from a description of practice in our field I hoped that we might obtain insight into the essential underlying assumptions, methodology, and philosophy of our field, and that we might find a commonalty of approach or discover particularly revealing examples of practice. I think you can agree that much of this goal has been achieved. The title of this chapter is stolen from one of Kuhn's books wherein he referred to the essential tension of scientific tradition and change. Our brief history here reveals sets of essential tensions that form the crux of the issues that we should consider at this workshop. These are discussed here.

Potential and Actualization

We adopted a new paradigm around 1970. This paradigm has guided both theoretical and experimental development. Like most new paradigms, in the beginning it was mainly one of potential—a promise of success discoverable in selected and still incomplete examples. The paradigm leads to a criterion for choosing problems that, while the paradigm is still taken for granted, can be assumed to have solutions. That is, the paradigm must be a source of puzzles, the solutions of which add to the scope and precision with which the paradigm can be applied. Puzzles are that special category of problems that can serve to test ingenuity or skill in solution. Puzzles must have an assured solution (which itself may have no intrinsic importance). Kuhn says, "One of the reasons why normal science seems to progress so rapidly is that its practitioners concentrate on problems that only their own lack of ingenuity should keep them from solving." Puzzles must have rules that limit both the nature of acceptable solutions and the steps by which they are obtained.

The paradigm-as-shared-values is a source of commitments for the scientists that share it. These include:

- Explicit statements of scientific law and about scientific concepts and theories. While these continue to be honored, such statements help to set puzzles and to limit acceptable solutions.
- A multitude of commitments to preferred types of instrumentation and to the ways in which accepted instruments may legitimately be employed.
- Quasi-metaphysical commitments that tell scientists what sorts of entities the universe does and does not contain, what ultimate laws and fundamental explanations must be like, and what research problems should be.
- A commitment to be concerned to understand the world and to extend the precision and scope with which it has been ordered, therefore, says Kuhn, "to scrutinize, either for himself or through colleagues, some aspect of nature in great empirical detail. And if that scrutiny displays pockets of apparent disorder, then these must challenge him to new refinements of his observational techniques or to a further articulation of his theories."

Kuhn says, "Normal science is a highly determined activity, but it need not be entirely determined by rules." There are rules for moving the chess pieces; there are few rules for playing a game of chess.

Kuhn's analysis leads him to say, "Normal science consists in the actualization of that promise [of success offered by the paradigm], an actualization achieved by extending the knowledge of those facts that the paradigm displays as particularly revealing, by increasing the extent of match between those facts and the paradigm's predictions, and by further articulation of the paradigm itself." These are the three areas into which Kuhn says all journal literature of *normal* science can be classified. (There is extraordinary science—those transitions between paradigms, but few scientists ever get to participate in such events. Those of us who were practicing in the 1970's are lucky to have seen the adoption and growth of our own paradigm, and we know how exciting that time was.)

As normal science advances, the promise of the paradigm is actualized and the paradigm itself is refined, reformulated, and further articulated. This enterprise seems to the outsider (one who does not share the commitments described above) to be an attempt to force nature into the preformed and relatively inflexible box that the paradigm supplies. The areas investigated by normal science are minuscule; the enterprise has drastically restricted vision. Kuhn says, "But those restrictions, born from confidence in a paradigm, turn out to be essential to the

development of science. By focusing attention upon a small range of relatively esoteric problems, the paradigm forces scientists to investigate some part of nature in a detail and depth that would otherwise be unimaginable." Thus the juxtaposition of our theoretical constructs and the observations associated with them leads to an extension of the precision and scope of the ordering of the world.

This progress of the paradigm, the practice of normal science—by its very nature (the finer and finer attention to detail)—encounters anomalies that require distortion to include, or deletion to ignore. As these mount, the promise offered by the paradigm begins to fade—the paradigm ceases to be a source of puzzles with assured solutions. Scientists committed to the paradigm begin to realize that they may fail to solve a puzzle within the confines of the paradigm, but, for a long time, this is likely to be seen as an individual failure, not a failure of the corpus of current science. Individuals respond to this developing crisis in different ways: some leave the field, some stop solving problems, some escalate the distortion and the deletion and force the system to work anyway; these actions stresses the community. Ultimately the paradigm becomes non-functional, it ceases to be a source of solvable puzzles. The scientist committed to understanding the world must then move on to another paradigm or quit. In some rare cases, the promise of a paradigm may be fully actualized, that is, all the puzzles may be successful solved (this happened in optics, for example). In this case, the science becomes a craft, and scientists, committed to increasing the understanding of their world, move on to other more productive areas. I suspect that the EPA policy makers would be very happy if air pollution modeling became a craft.

Functioning with maps that are highly distorted or that suffer from serious deletion must ultimately be painful. This leads to either resolving the crisis and modifying the map or to non-functional behavior from those using the maps.

And this brings us to our current situation. I believe that we need to face the anomalies, not force the fit, and, on the one hand, tighten the standards for "reasonable agreement" while on the other, making it a little safer to take a risk and fail.

Because, I believe, the standards for acceptance of agreement between fact and theory have been set too low by the current modeling community, easy moves occur from one viewpoint to another. I have described the mechanism of "reasonable agreement" in the history above. In Kuhn's view, in so far as their work is quantitative, most scientists attempt to improve the measure of "reasonable agreement" characteristic for a theory in a given application and to open up new

areas of application and establish new measures of "reasonable agreement" applicable to them. He has shown, however, that there is no consistently applied or consistently applicable external criterion for "reasonable agreement" and has suggested that reasonable agreement is what is accepted by the community; usually this is what is accepted for journal publication that is not refuted in a very short time.

We have seen in the example here, that for modeling, the measure of "reasonable agreement" is often established by claim and is based on a ridiculously limited set of evidence. If the reviewers of the papers are not themselves modelers, this evidence may seem convincing, or if they are modelers, they may object—but not too strenuously—because they see themselves in the same position with respect to publishing their own modeling results! Some modelers have gone elsewhere or have found other avenues of expression, for example, EPA reports. These, however, are always associated with funding and contractual commitments that confound the science and the peer review process. These too are most likely to "confirm" or "support" the model developed. In its present situation, EPA is often helpless when it comes to enforcing a measure of quality assurance with regard to the characteristics of "reasonable agreement." Therefore, here too, we find low standards accepted for reasonable agreement.

On the other hand, we must be cautious not to set the standards too high, for then "no one satisfying the criterion of rationality would be inclined to try out the new theory, to articulate it in ways which showed its fruitfulness or displayed its accuracy and scope."¹⁵ Kuhn says, "What from one viewpoint may seem the looseness and imperfection of choice criteria conceived as rules may, when the same criteria are seen as values, appear an indispensable means of spreading the risk which the introduction or support of novelty always entails."

Therefore, I believe that we must work as a community to reestablish an explicit recognition of a set of rules. These rules, in the words of Kuhn "limit both the nature of acceptable solution and the steps by which they are to be obtained." If we take 'limit' not to mean restrict, but rather to define, then these rules are not unnecessary burdens on science. Also the interpretation of limit as "define" explains why rules can lie unexpressed below the surface of regular scientific activity: definitions, as we have seen, must often be taken for granted if they are to function at all. These definitional rules do not define what the state of the science, or even what the paradigm, is. But rather, they define whether or not solutions fall within the paradigm, require modification or distortion of the paradigm, or fall completely outside the paradigm. Such rules can change in direct

response to the input of puzzle-solutions, but clearly they are not infinitely flexible. Frequently in times of crisis, more solutions are ruled out of the paradigm than are included within it. When the rules function in this way to define the paradigm more and to prevent its unnecessary distortion, the rules can save a still-functional paradigm and work goes on.

When the rules have strengthened and protected the paradigm in these ways, they—now accepted by consensus in the community—recede again from explicit recognition and indirectly guide the community toward interesting puzzles in need of solution. *Thus the definitional rules can be changed in response to the increasingly refined solutions proposed within the paradigm, but they direct research best only when they are sufficiently accepted by the community to be implicit.*

Sharing risk is a traditional approach in times of crisis. This I believe is one of the reasons the Atkinson and Lloyd review has acquired such dominance: using it not only avoids having to repeat or re-examine work, it is safer than sticking your own neck out; the user of the Atkinson and Lloyd review does not have to defend it. It has become a kind of standard. A problem with standards, as EPA well knows, is that they become outdated. The force that the Atkinson and Lloyd review has exerted on the field clearly suggests one obvious action EPA can undertake to resolve some of the conflicts: make a commitment to, and provide support for, regular review processes like the Atkinson and Lloyd work.

But agreement as to what constitutes kinetic knowledge is *not* sufficient. There are other categories of empirical and theoretical knowledge that must be agreed upon. We have described this agreement, this standard, as “reasonable agreement,” defined as the fit of various kinds of evidence measured according to consensus, or in some cases, unchallenged assertion. And this in part has been the problem: solutions have been asserted and accepted that later have been shown to have been established on weak evidence. When this becomes ordinary, then crisis is not far away. That is, we want to say that “reasonable agreement” is one type of rule that needs to surface in times of paradigmatic crisis.

For science to advance at all there must be reasonable agreement between three sets of evidence:

1. between theory and observation;
2. between different sets of observations; and

3. between explanation and prediction.

Without agreement, no detailed comparison could be made between the three sets of evidence: no puzzle-solutions could be evaluated. And the puzzles most meaningful in maintaining or challenging a paradigm are exactly those that lie on the interfaces of the three sets listed above.

Thus the first sense in which reasonable agreement determines scientific progress is in defining what sorts of evidence can be compared. That is to say, that by defining what qualifies as evidence of theories, observations, explanation, and predictions, the community establishes reasonable agreement for quantitative precision. Once evidential standards of reasonable agreement have been recognized, the puzzles we have described as at the interfaces of theory and observation, between observations, and between explanation and prediction can be defined. When solutions to the problems are offered, the second function of reasonable agreement is activated: agreement for better quantitative and qualitative precision to discriminate among the solutions. Again, this function is determined by consensus within the community and again as a standard it varies directly with the progression of the paradigm from potential to actualization. Here again, commitments from EPA to support a regular process to examine reasonable agreement in the three sets of evidence above will be beneficial to the modeling community.

Explanation and Prediction

That we are in, or approaching a crisis, brought about by forcing actualization over promise, I believe can be seen in the tension between explanation and prediction revealed in the brief historical examples present here.

Looking back at the work that has been done in the last 15-years, it appears that, first, modelers attempt a rather full explanation of the observations in smog chambers using new kinetics data and producing new hypotheses about pathways and important reactions. It is no surprise, according to Kuhn, that in the end, all these theories+auxiliary statements have "agreed with the data" more or less. If there was not the chance that the puzzle had a solution, it would not have been undertaken. And, no one is going to conclude that he failed to solve the puzzle he set for himself, even if he has to discount facts. If we see demonstrations of "reasonable agreement" at all, we only see those that can be the basis of a claim. Yet other modelers often make very different conclusion from the evidence presented to show "agreement with the data" and so doubt the model's

verisimilitude. Second, the mechanism, that more or less agreed with the data becomes the source of a simpler description, which some compare not with reality, but with the original model—the idea being that the original model is the correct explanation—while others compare with a smaller set of chamber data and also announce “agreement with data.” History shows, however, that both the original modeler and other modelers find “pockets of apparent disorder” which “challenge him to a new refinement of his observational techniques or to a further articulation of his theories.” and they soon present us with a new model.¹⁰ Because, I think, the evidence presented for the former model is too weak, this new challenge immediately undermines the basis of belief in both the explanation and the prediction of the former model. This can be seen to result from a process of “false actualization” and insufficient vindication of the work accomplished.

Viewed by one outside the scientific process, the situation described above closely associates in daily experience with charlatanism or outright fabrication—“Yeah, that’s the ticket! Or, ah, maybe . . .” Viewed from a historical perspective, the theories associated with the HO-paradigm are *new* theories. Kuhn says most new theories do not survive. And when they do, much work, both theoretical and experimental is required before the new theory can display sufficient accuracy and scope to generate widespread conviction. Before a group accepts it, a new theory has been tested over time by the research of a number of men, some working within it, and others within a rival. Such a mode of development *requires* a decision process which permits rational men to disagree, and if a scheme attempts to eliminate this mode of development it will not produce science.¹⁵

So if there is to be science in photochemical modeling, EPA has to learn to live with a process whereby rational men may disagree. Coupled with this recognition are other characteristics of these members of professional scientific groups. Kuhn identified several striking characteristics,

The scientist must be concerned to solve problems about the behavior of nature. In addition, though his concern with nature may be global in extent, the problems on which he works must be problems of detail. More important, the solutions that satisfy him may not be merely personal but must instead be accepted as solutions by many. The group that shares them may not, however, be drawn at random from society as a whole, but is rather the well-defined community of the scientist’s professional compeers. One of the strongest, if still unwritten, rules of scientific life is the prohibition of appeals to heads of state or to the populace at large in matters scientific. Recognition of the existence of a uniquely competent professional group and acceptance of its role as the exclusive arbiter of professional achievements has further implications. The group’s members, as individuals and by virtue of their shared training and experience, *must be seen as the sole possessors of the rules of the game or of some equivalent basis for unequivocal judgments.* To doubt that they

shared some such basis for evaluations would be to admit the existence of incompatible standards of scientific achievement. This admission would inevitably raise the question whether truth in the sciences can be one.

Thus while there clearly may be a role for EPA in affirming standards, it must be one of supporting actions of the scientific community as a group.

Testing and Fit

So we can agree that in times of paradigmatic crisis, standards for definitinal rules and for reasonable agreement must be made explicit, and that with these tools the paradigm is strengthened or successfully challenged. But how are standards employed? And if we really believe that examination of these standards leads to useful guidelines, just what are the lines guiding? The answer is *testing*.

There are at least three different fields for testing and they coincide with the three sets of evidence listed above for the standards of reasonable agreement. That is, theory must be tested with observation, observations must be tested together, and explanation must be tested with prediction. The 'tests' in each of these areas are actually the puzzles set by the paradigm or a theory within the paradigm for the practice of normal science and we have seen several examples in the history. Here we are concerned with how tests of these three areas determine the use of any particular theory, the model generated in response to the theory, and the mechanism that underlies the model. I shall emphasize that, although a model and its mechanism derive from common theory and are its representatives for testing, successful testing of any one component does not convey even the hope of success to the others. Thus it is entirely possible that a useful theory is entirely unsuccessful in model and mechanism generation and so is untestable—and thus does not become the subject of normal science and will remain a mere speculation until it does become testable. More importantly, however, I would argue that models and mechanisms which seem to be well-tested may in fact have not been sincerely tested at all, and that in any case—pass or fail, sincere or not—such tests say nothing about the theory that the models and mechanisms represent. *Then what can testing mean?*

Verification and Vindication

We have seen that scientists apparently do not test the paradigm theory when testing the theories' models, mechanisms, and auxiliary statements. Failure to fit is assumed to be a problem in the mechanism, or the auxiliary statement, not in

the the theory itself. Successful fitting, on the other hand, means just that. What about next time? In testing you see, we rely on inductive logic. Inductive logic is simply the belief that the world will function in the future as it has in the past, based on our best experience of it. In that it relies on the uniformity of nature, such reasoning is not objectionable on its face, but it is by definition invalid for proof. This is because the assertion that the world is uniform is made and weakly confirmed solely through observation—and we have seen that to a large extent, our observations are filtered through our expectations, through our theories so to speak. The sun rises in the East each morning. “You mean it is not moving and we are?”

To the historian, at least, it makes little sense to suggest that verification or validation is establishing the agreement of fact with theory. Invoking realism also has its problems. For example, there is the belief that a model that offers explanation has inherent ability to predict because its ability to explain is taken as proof that it is truth, that the entities in the model have a true correspondence with the real world. No support accrues to realism by showing that realism is a good hypothesis for explaining scientific practice. Need we take a good explanatory hypothesis to be true? A complete acceptance of realism commits one to an unverifiable correspondence with the world. We have many examples of validated mechanisms that are no longer used. Is today’s scientific environment any less “dynamic” than when Tom Hecht said “In such a dynamic scientific environment, it would be presumptuous to suggest that any given set of reactions and rate constants is *the* chemical mechanism for smog formation.”?

Does this mean that all models are just opinion? If we insist upon the criteria of certitude, incorrigibility, and immutability, then yes. If we relax those criteria and recognize that these opinions we can affirm on the basis of evidence and reasons that have sufficient probative force to justify our claim *at the time* that the opinion affirmed is true. We must stress *at the time* because we must be prepared to have the opinion we now claim to be true on the basis of evidence and reasons now available turn out to be false or in need of corrections when new evidence and other reasons come into play.⁶¹

We may agree that induction can not be “validated,” that it is logically impossible to show that the conclusion of an induction follows logically from the empirical evidence. Nevertheless, we can perhaps “vindicate” induction. That is, establish a form of argument “If any method of predicting the future works, then induction works.” Or if induction does not work, then no method will work. In our terms, this translates to being able to make the statement

"If this model does not explain or predict in this situation then I do not know how to produce another model that does. Therefore, we judge that it is sensible to adopt this model for this use at this time."

What is required to make such a statement?

Methodologies

Mechanism Vindication

One purpose of this workshop is to describe a method that would allow a modeler to make a vindication statement about his model and would allow the modeling community and model users to have confidence that it was true.

I believe that some of the elements of this method would be:

A. We could believe that a mechanism is the best description we can obtain about an atmospheric chemistry at this time when:

- 1) there is a high degree of confidence in the existence and behavior of the majority of the entities that are in the mechanism—meaning they have been measured and characterized by more than one research group and the community has reached a consensus on the information, or
- 2) in the case of explicit mechanisms, there is confidence that the remaining entities described in the model exist and behave as described; this requires acceptable processes for
 - ▷ elaboration of the alternatives;
 - ▷ distinguishing among alternatives;
 - ▷ estimating the consequences of good and bad choices;
 - ▷ making unique predictions that can be verified experimentally.
- 3) in the case of simple models, there is confidence that the simple model agrees with both the explicit model and data.

B. There is consistence among the various forms of evidence:

- 1) the mechanism agrees with chamber data and the agreement is established in a stepwise manner;
 - 2) different chambers produce data that agree;
 - 3) the mechanism not only explains data, it predicts correctly in situations not used in its development.
- C. Before a mechanism is declared suitable for use in atmospheric simulations, it has been tested over the range of conditions expected to occur in the atmosphere, *e.g.*, mixture composition (both static and dynamic), large dilution, dynamic injection, entrainment of material from aloft.
- D. Others.

Classes of Scientific Work

Normal scientific work associated with mechanism development and testing comprises:

1. Production of significant facts that the theory of HO \cdot -chain attack has shown to be particularly revealing of the nature of things.
 - A. The determination of the existence, the rate, and the mechanism of elementary reactions of importance to the initiation, propagation, and termination of the chain.
 - B. The organization and evaluation of elementary reactions and the theoretical estimation of the critical reactions that have not yet been determined in A.
2. Production of factual determinations that can be compared directly with predictions from mechanisms based on the HO \cdot -chain theory, and the manipulation of the theory or the production of auxiliary statements to allow the theory to make predictions that can be confronted directly with measurements.
 - A. Producing smog chamber experimental data in a form suitable for the direct comparison with predictions.
 - B. Creating mechanisms that adequately describe the experimental situations.
 - C. Producing simulations and comparing predictions and observations.
3. Production of empirical fact to articulate the HO \cdot -chain theory, to resolve some of the residual ambiguities, and accounting for changes in auxiliary statements and basic theory that arise from the empirical work.

- A. Performing chamber characterization experiments; quality assurance measures, intercomparison of different instrumental techniques; field studies; making measurements of predicted key intermediates.
- B. Formulation of new model programs and approaches; testing of prediction methods, making unusual or "risky" or unique predictions that can be tested.
- C. Comparisons of predictions from different mechanisms; comparisons of observations from different chambers.
- D.

We have seen that essentially none of these activities moves forward without the others.

Agreement

The form that a science puzzle often takes is to deduce the form of auxiliary statements that, when combined with generally accepted theory will show "reasonable agreement" with facts or finding ways to produce facts that can be shown to be in "reasonable agreement" with predictions.

On the empirical side of the puzzle, it has been a long standing heuristic that multiple measurement systems must be used to verify the facts and to estimate the accuracy and bias of their statement. This means that different chambers must be used to test models, perhaps we should even say different chamber groups. At times there may need to be directly related experimental work in different facilities to obtain answers needed to resolve some of the puzzles—*e.g.*, UCR needs to compare their HCHO measurement method with, say the EPA cartridge system, so that the differences in reported HCHO between UCR and UNC can be resolved.

On the theoretical side of the puzzle and in the case of models, I suggest that there is contamination among the classes of knowledge that are used to construct the model. For example, I might identify these classes as:

- ▷ accepted theory
- ▷ new theory
- ▷ quasi-theory
- ▷ current hypothesis
- ▷ speculation

Having understanding and agreement as to what parts of the model belong to which class of knowledge and having a demonstration as to the sensitivity of the model to these parts is essential in judging its verisimilitude. This process was more common in the earlier days of model development, suggesting that recently it has been assumed that there is some implicit agreement for recent models as to what is accepted as true. We have already suggested that Atkinson and Lloyd's review has been accepted by some as defining accepted knowledge and thus, they assert, can be used without defense. We know, however, that Atkinson would be the first to deny the status attributed to his now 3- to 4-year old work.

Therefore, reviews and classification of knowledge, both empirical and theoretical, are needed on a regular basis which would be similar to the reviews of rate constants, but which must include all available experimental data and all current hypotheses in use in all candidate models. To be acceptable to the scientific community, these reviews must be conducted by the group of scientists participating in the creation activities. EPA's role should be one of facilitator—arranger, publisher, and disseminator of results.

We have a good example of a review for the types of evidence described under 1B above. I know of no examples of reviews of the types of evidence described in 2A and 3C above. That is, a review effort is needed to assess the agreement among different chamber groups. Another review is needed to assess the degree of agreement among different modeling groups for representations of known and hypothesized entities in the mechanisms.

Kuhn's analysis of how scientist adopt a new theory strongly suggests that the comparison of a single model with facts will never be the source of rejecting that model. Therefore, at least two models must be compared with facts at the same time to achieve progress. We have already seen, however, that each model helps fix what counts as evidence or test. This means that the resolution of this comparison may need specific facts; such facts often require 2-3 years to produce.

A final type of regular review, then, would be the comparison of models compared to chamber data.

It should be our goal not to seek multiple representations of reality—for example, it is not the Carbon Bond Mechanism, nor should it be the Carter, Atkinson, Lloyd, Lurman mechanism—instead, it should be

- ▷ 'What is the best representation of what we know with a high degree of certainty?', and

- ▷ 'What is an adequate representation for what we probably know?' and
- ▷ 'How creative can we be about what we are guessing at?'

References

- 1 "Improvements Needed in Developing and Managing EPA's Air Quality Models," GAO/RCED-86-94, April 1986.
- 2 Letter from Chairman John Dingell to EPA Administrator Lee Thomas, June 9, 1986.
- 3 Letter from Basil Dimitriadis to participants in workshop on Evaluation and Documentation of Chemical Mechanisms Used in Air Quality Models, July 29, 1986.
- 4 Fox, D. "Judging Air Quality Model Performance," *Bull. Amer. Metero. Soc.*, **62**, pp 599-609, 1981.
- 5 "Guideline On Air Quality Models (Revised)," EPA-450/2-78-027R, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1986.
- 6 Jeffries, H.E, K.G. Sexton, C.N. Salmi, "The Effect of Chemistry and Meteorology on Ozone Control Calculations Using Simple Trajectory Models and the EKMA Procedure." EPA-450/4-81-034, U.S. Environmental Protection Agency, Research Triangle Park, NC. 1981.
- 7 Leone, J.A, J.H. Seinfeld, "Evaluation of Chemical Reaction Mechanisms for Photochemical Smog, Part II: Quantitative Evaluation of The Mechanisms," Final Report CA 810184, Atmospheric Sciences Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC. 1984.
- 8 Levin, M. "What Kind of Explanation is Truth?" in *Scientific Realism*, ed. J. Leplin, University of California Press, Berkeley, 1984.
- 9 Hidy, G.M, "Future Needs in Atmospheric Research and Modeling," Final Workshop Report, ASRL, U.S. Environmental Protection Agency, Research Triangle, NC, 1986.
- 10 Kuhn, T.S., *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, 1962.
- 11 Feyerabend, P.K., "Explanation, Reduction, and Empiricism," in *Minnesota Studies in the Philosophy of Science*, ed H. Feigl and G. Maxwell, Minneapolis, 1962.
- 12 Atchinstein, P., *Concepts of Science: A Philosophical Analysis*, John Hopkins Press, 1969.
- 13 Popper, K., *Objective Knowledge* Oxford Press, 1972.
- 14 Hacking, I. *Scientific Revolutions*, Oxford University Press, NY, 1981.
- 15 Kuhn, T.S., *The Essential Tension, Selected Studies in Scientific Tradition and Change*, University of Chicago Press, Chicago, 1977.

-
- 16 Lakatos, I, in *Minnesota Studies in the Philosophy of Science*, Minneapolis, 1972.
 - 17 Bandler R., J. Grinder, *The Structure of Magic*, Science and Behavior Books, Inc., Palo Alto, CA, 1975.
 - 18 Weinberg, G., *Rethinking System Analysis and Design*, Little, Brown and Company, Boston, 1982.
 - 19 Leighton, P.A. *Photochemistry of Air Pollution*, Academic Press, NY, 1961.
 - 20 Atkinson, R., K.R. Darnall, A.C. Lloyd, A.M. Winer, J.N. Pitts, "Kinetics and Mechanisms of the Reactions of the Hydroxyl Radical with Organic Compounds in the Gas Phase," in *Advances in Photochemistry*, 11, John Wiley and Sons, 1979.
 - 21 Demerjian, K.L., J.A. Kerr, J.G. Calvert, "The Mechanism of Photochemical Smog Formation," in *Adv. in Environ. Sci. and Tech.*, 4, John Wiley and Sons, NY, 1974.
 - 22 Greiner, N.R., *J. Chem. Phys.*, 46, p 3389, 1967
 - 23 Heicklen, J., K. Westberg, N. Cohen, "The Conversion of NO to NO₂ in Polluted Atmospheres," *Symp. Chem. Reactions in Urban Atmospheres*, Research Laboratories, General Motors, Warren, Michigan, Oct. 1969.
 - 24 Niki, H., E. Daby, B. Weinstock, "Mechanisms of Smog Reactions," Ford Motor Co. Technical Report, ca. 1970.
 - 25 Niki, H., E. Daby, B. Weinstock, "Mechanisms of Smog Reactions," in *Photochemical Smog and Ozone Reactions, Adv. in Chemistry Series*, 113, p 13, 1972.
 - 26 Putnam, H. "The 'Corroboration' of Theories' in Scientific Revolutions, ed. I. Hacking, Oxford University Press, 1981.
 - 27 Gear, W. *The Numerical Solution of Initial Value Problems*, Printice Hall, 1971.
 - 28 Leplin, J., "Introduction," in *Scientific Realism*, ed. J. Leplin, University of California Press, Berkeley, 1984.
 - 29 Smart, J.J., *Between Science and Philosophy*, Random House, NY, 1968.
 - 30 The Encyclopedia of Physics, 2ad Edition, ed. R.M. Besancon, Van Nostrand Reinhold Co. NY, 1974.
 - 31 Hecht T., J. Seinfeld, M. Dodge, "Further Development of Generalized Kinetic Mechanism for Photochemical Smog," *Environ. Sci. Tech.*, 8, p 327-339, 1974.
 - 32 Hecht, T. M.K. Liu, D.G. Whitney, "Mathematical Simulation of Smog Chamber Photochemical Experiments," EPA-650/4-74-040, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974
 - 33 Webster's New Twentieth Century Dictionary, Unabridged, Second Edition, Simon and Schuster, New York, NY, 1983.

References

- 34 Whitten, G.Z., H. Hogo, "Mathematical Modeling of Simulated Photochemical Smog," EPA-600/3-77-011, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1977.
- 35 Durbin, P., T.A. Hecht, G.Z. Whitten, "Mathematical Modeling of Simulated Photochemical Smog," EPA-650/4-75-026, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1975.
- 36 Hendry, D.G., R.A. Kenley, *J. Amer. Chem. Soc.* **99**, p 3189, 1977.
- 37 Hendry, D.G., A.C. Baldwin, J.R. Barker, D.M. Golden, "Computer Modeling of Simulated Photochemical Smog," EPA-600/3-78-059, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1978.
- 38 Jeffries, H., personal recollection.
- 39 Carter, W.P., A.C. Lloyd, J.L. Sprung, J.N. Pitts, "Computer Modeling of Smog Chamber Data: Progress in Validation of a Detailed Mechanism for the Photooxidation of Propene and n-Butane in Photochemical Smog," *Int. J. Chem. Kinetics* **11** p 45, 1979.
- 40 Dodge, M., "Combined Use of Modeling Techniques and Smog Chamber Data to Derive Ozone-Precursor Relationships," in Proceedings of Inter. Conference on Ozone, Raleigh NC, 1977.
- 41 Falls, A.H., J. Seinfeld, "Continued Development of a Kinetic Mechanism for Photochemical Smog," *Environ. Sci. Technol.* **12**, pp 1398-1406, 1978.
- 42 McRae, G.J., W.R. Goodin, J.H. Seinfeld, "Mathematical Modeling of Photochemical Air Pollution," State of California Air Resources Board EQL Report NO. 18, 1982.
- 43 Falls, A.H., G.J. McRae, J.H. Seinfeld, "Sensitivity and Uncertainty of Reaction Mechanisms for Photochemical Air Pollution," *Inter. J. Chem. Kinetics*, **XI**, pp. 1137-1162, 1979.
- 44 Whitten, G.Z., H. Hogo, J.P. Killus, "The Carbon Bond Mechanism: A Condensed Kinetic Mechanism of Photochemical Smog," *Environ. Sci. and Technol.*, **14**, pp. 690-700, 1980.
- 45 Atkinson, R., A.C. Lloyd, L. Wings, "An Updated Chemical Mechanism For Hydrocarbon/Nitrogen Oxides Photooxidations Suitable for Inclusion in Atmospheric Simulation Models," *Atmos. Envir.* **16**, pp. 1341-1355, 1982.
- 46 Killus, J.P., G.Z. Whitten, "Effects of Photochemical Kinetics Mechanisms on Oxidant Model Predictions," EPA-600/3-83-111, U.S. EPA, RTP, 1983.
- 47 Carter, W.P.L., A.M. Winer, J.N. Pitts, "Effect of Kinetic Mechanism and Hydrocarbon Composition on Oxidant-Precursor Relationships Predicted by the EKMA Isopleth Technique," *Atmospheric Environment* **16**, pp. 113- 120, 1982.
- 48 Leone, J.A., J.H. Seinfeld, "Evaluation of Chemical Mechanisms For Photochemical Smog:

- Part II. Quantitative Evaluation of the Mechanisms," Final Report, EPA Co-operative Agreement 810184, 1984.
- 49 Dunker, A.M., S. Kumar, P.H. Berzins, "A Comparison of Chemical Mechanisms Used in Atmospheric Models," *Atmo. Environ.* **17**, 1983.
- 50 Dennis, R.L., M.W. Downton, "Evaluation of Urban Photochemical Models for Regulatory Use," *Atmo. Environ.* **18**, pp. 2055-2069, 1984.
- 51 Whitten, G.Z., "The Chemistry of Smog Formation: A Review of Current Knowledge," *Environment International* **9**, pp. 447-463, 1983.
- 52 Carter, W.P.L., R. Atkinson, A.M. Winer, J.N. Pitts, "Evidence for Chamber-Dependent Radical Sources: Impact on Kinetic Computer Models for Air Pollution," *Intr. J. Chem. Kinet.* **13**, pp. 735-740, 1981.
- 53 Killus, J.P., G.Z. Whitten, "Comments on Paper 'Evidence for Chamber-Dependent Radical Sources,'" *Intr. J. Chem. Kinet.* **13** pp. 1101, 1981.
- 54 Carter, W.P.L., R. Atkinson, A.M. Winer, J.N. Pitts, "An Experimental Investigation of Chamber-Dependent Radical Sources," *Intr. J. Chem. Kinet.* **14**, p. 1071, 1982.
- 55 Killus, J.P., G.Z. Whitten, "Background Radical Initiation Processes in Smog Chambers: A Perspective," System Applications Paper, San Rafael, CA, 1982.
- 56 Sakamaki, F., S. Hatakeyama, H. Akimoto, "Formation of Nitrous Acid and Nitric oxide In the Heterogeneous Dark Reaction of Nitrogen Dioxide and Water Vapor in a Smog Chamber," *Intr. J. Chem. Kinet.* **15** pp. 1013-1029, 1983.
- 57 Pitts, J.N., E. Sanhueza, R. Atkinson, W.P.L. Carter, A.M. Winer, G.W. Harris, and C.N. Plum, "An Investigation of the Dark Formation of Nitrous Acid in Environmental Chambers," *Intr. J. Chem. Kinet.* **16**, pp. 919-939, 1984.
- 58 Carter, W.P.L., F.W. Lurmann, R. Atkinson, A.C. Lloyd, "Development and Testing of a Surrogate Species Chemical Reaction Mechanism," Final Report EPA Contract No. 68-02-4104, ASRL, U.S. Environ. Prot. Agency, RTP, NC, 1986.
- 59 Leone, J.A., J.H. Seinfeld, "Updated Chemical Mechanism for Atmospheric Photooxidation of Toluene," *Intr. J. Chem. Kinet.* **16**, pp.159-193, 1984.
- 60 Leone, J.A., R.C. Flagan, D. Grosjean, J.H. Seinfeld, "An Outdoor Smog Chamber and Modeling Study of Toluene-NO_x Photooxidation," *Intr. J. Chem. Kinet.* **17**, pp.177-216, 1985.
- 61 Adler, M.J. *Ten Philosophical Mistakes* MacMillan Publishing, NY, 1985.

Appendix C

Workshop Agenda

Appendix C

Workshop on Evaluation/Documentation of Chemical Mechanisms

Sponsored by EPA's Atmospheric Sciences Research Laboratory
Research Triangle Park, North Carolina

The Raleigh Inn
Raleigh, North Carolina

December 1-3, 1986

AGENDA

12/1/86

9:00 AM	Registration	Workshop Chairman and Moderator: R. Atkinson
9:30 AM	Call to Meeting	B. Dimitriadis, EPA
9:35 AM	Welcome-Introduction of Dr. Courtney Riordan	A. Ellison, EPA
9:45 AM	Opening Remarks	C. Riordan, EPA
10:00 AM	Introductory Remarks by Workshop Chairman	R. Atkinson, UCR
10:15 AM	Presentation	H. Jeffries, UNC
	Discussion	
12:00 NN	Lunch	
1:30 PM	Presentation	J. Tikvart, EPA
	Discussion	
2:30 PM	Presentation	K. Demerjian, SUNY
	Discussion	
3:30 PM	Break	
4:00 PM	Presentation	R. Atkinson, UCR
	Discussion	
Evening:	Steering Committee meets to formulate questions and organize discussions for the next two days.	

Appendix C

12/2/86

9:00 AM	Presentation	M. Gery, SAI
	Discussion	
10:00 AM	Presentation	A. Dunker, GM
	Discussion	
11:00 AM	Presentation	W. Carter, UCR
	Discussion	
12:00 NN	Lunch	
1:30 PM	Presentation	F. Lurmann, ERT
	Discussion	
2:30 PM	Presentation	G. McRae, Carnegie Mellon
	Discussion	
3:30 PM	Presentation	G. Whitten, SAI
	Discussion	
4:30 PM	Report on Steering Committee Deliberations and Plans	R. Atkinson
	Continuing Discussions	
5:00 PM	End of Workshop	

12/3/86

9:00 AM	Final meeting of the Steering Committee
---------	---

Appendix D

Workshop Attendees

Workshop Attendees

Dr. Roger Atkinson
Statewide Air Pollution Research
Center
University of California-Riverside
Riverside, CA 92521

Dr. Marcia Dodge
Atmospheric Sciences Research
Laboratory (MD-84)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Mike Barnes
Atmospheric Sciences Research
Laboratory (MD-57)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Alan Dunker
Environmental Sciences Dept.
General Motors Research Laboratories
12 Mile and Mound Roads
Warren, MI 48090-9055

Dr. Ronald Bradow
Atmospheric Sciences Research
Laboratory (MD-59)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Jack Durham
Atmospheric Sciences Research
Laboratory (MD-57)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Joseph Bufalini
Atmospheric Sciences Research
Laboratory (MD-84)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Alfred Ellison, Director
Atmospheric Sciences Research
Laboratory (MD-59)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. William Carter
Statewide Air Pollution Research
Center
University of California-Riverside
Riverside, CA 92521

Dr. Michael Gery
Systems Applications, Inc.
101 Lucas Valley Road
San Rafael, CA 94903

Dr. Kenneth Demerjian
Atmospheric Sciences Research Center
Earth Science Building - Room 324
State University of New York-Albany
1400 Washington Avenue
Albany, NY 12222

Dr. Daniel Jacob
Department of Earth and Planetary
Sciences
Harvard University
29 Oxford Street
Cambridge, MA 02138

Dr. Basil Dimitriadis
Atmospheric Sciences Research
Laboratory (MD-59)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Harvey Jeffries
Dept. of Environmental Sciences and
Engineering
School of Public Health
University of North Carolina
Chapel Hill, NC 27514

Workshop Attendees

Mr. William Keith
Office of Research and Development
(RD-680)
U.S. Environmental Protection Agency
401 M Street, SW
Washington, DC 20460

Dr. Courtney Riordan
Office of Research and Development
(RD-680)
U.S. Environmental Protection Agency
401 M Street, SW
Washington, DC 20460

Dr. Alan Lloyd
Environmental Research and Technology
975 Business Center Circle
Newbury Park, CA 91320

Mr. Kenneth Schere
Atmospheric Sciences Research Laboratory
(MD-80)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Fred Lurmann
Environmental Research and Technology
975 Business Center Circle
Newbury Park, CA 91320

Dr. Kenneth Sexton
Dept. of Environmental Sciences and
Engineering
School of Public Health
University of North Carolina
Chapel Hill, NC 27514

Mr. Edwin Meyer
Office of Air Quality Planning and
Standards (MD-14)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Jack Shreffler
Atmospheric Sciences Research Laboratory
(MD-59)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Gregory McRae
Department of Chemical Engineering
Carnegie Mellon University
Shenley Park
Pittsburgh, PA 15213

Mr. Joseph Tikvart
Office of Air Quality Planning and
Standards (MD-14)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dr. Hiromi Niki
Science Research Laboratory
Ford Motor Company
P.O. Box 2053
Dearborn, MI 48121-2053

Dr. Gary Whitten
Systems Applications, Inc.
101 Lucas Valley Road
San Rafael, CA 94903

Dr. Deran Pashayan
Office of Research and Development
(RD-680)
U.S. Environmental Protection Agency
401 M Street, SW
Washington, DC 20460