ACCURACY OF PREDICTION OF URBAN AIR POLLUTANT CONCENTRATIONS BY DIFFUSION MODELS

John H. Seinfeld, California Institute of Technology

Projections of future transportation-related air quality impacts require the use of mathematical models that relate emissions to air quality. Whereas the derivation and use of such models have received much attention (at least for inert pollutants, such as CO), much less attention has been paid to questions of the interpretation of the concentrations these models predict and how the predictions relate to real atmospheric quantities. Concepts of validity and accuracy must be carefully defined for any model that is to be used in order that the predictions from the model can be properly evaluated. The purpose of this paper is to formulate the concepts of validity and accuracy for atmospheric air pollutant diffusion models and to suggest numerical experiments that can be used to test both the validity and the accuracy of the models.

VALIDITY OF ATMOSPHERIC DIFFUSION EQUATION

All conventional atmospheric diffusion models are based on the equation of conservation of mass:

\[
\frac{\partial c_i}{\partial t} + u \frac{\partial c_i}{\partial x} + v \frac{\partial c_i}{\partial y} + w \frac{\partial c_i}{\partial z} = D_i \left( \frac{\partial^2 c_i}{\partial x^2} + \frac{\partial^2 c_i}{\partial y^2} + \frac{\partial^2 c_i}{\partial z^2} \right) + R_i(c_1, \ldots, c_n, T) + S_i(x, y, z, t)
\]  

(1)

where \( c_i \) is the concentration of species \( i \); \( u, v, \) and \( w \) are the fluid velocities in the 3 coordinate directions; \( D_i \) is the molecular diffusivity of species \( i \) in air; \( R_i \) is the rate of generation (or the negative of the rate of disappearance) of species \( i \) by chemical reactions at temperature \( T \); and \( S_i \) is the rate of injection of species \( i \) into the fluid from sources.

Because the atmosphere is a turbulent flow, the velocities \( u, v, \) and \( w \) are random functions of space and time. Consequently, the concentration \( c_i \) is also a random function of space and time. Solutions of equation 1 with realistic atmospheric velocities are difficult to obtain, even in the case in which \( R_i = 0 \) (inert species). To render equation 1 solvable, the fluid velocities are decomposed into mean and fluctuating components, \( u = \bar{u} + u' \) and so on. The quantities \( \bar{u}, \bar{v}, \) and \( \bar{w} \) represent the ensemble mean velocities of an infinite number of realizations of the same flow. Correspondingly, we can divide \( c_i \) into \( \bar{c}_i + c_i' \), where \( \bar{c}_i \) is the ensemble mean concentration (23).

Substituting the mean and fluctuating terms into equation 1 and averaging the resulting equation over the ensemble of flows, we obtain the equation governing \( \bar{c}_i \). In atmospheric applications, the molecular diffusion term is negligible when compared to that representing advective transport. Thus, neglecting the contribution of molecular
diffusion, the equation for $\bar{c}_i$ is

$$\frac{\partial \bar{c}_i}{\partial t} + u \frac{\partial \bar{c}_i}{\partial x} + v \frac{\partial \bar{c}_i}{\partial y} + w \frac{\partial \bar{c}_i}{\partial z} = \frac{\partial}{\partial x} \left( K_u \frac{\partial \bar{c}_i}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_v \frac{\partial \bar{c}_i}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_w \frac{\partial \bar{c}_i}{\partial z} \right) + R_i(c_1, \ldots, \bar{c}_n, T)$$

$$+ \bar{S}_i(x,y,z,t) \quad (2)$$

We note the emergence of the new variables $u' \bar{c}_i'$, $v' \bar{c}_i'$, $w' \bar{c}_i'$, which represent the fluxes of species $i$ in the 3 coordinate directions as a result of the velocity fluctuations, $u'$, $v'$, and $w'$. If species $i$ is involved in second-order chemical reactions, then the term $R_i$ will also lead to new dependent variables of the form $c_i c_j'$. For example, if species $i$ decays by a second-order reaction $I + I \rightarrow k$, then the mean rate of disappearance of species $i$ is given by $R_i = -k \bar{c}_i^2 = -k \bar{c}_i^2 - k \bar{c}_i'^2$, where the first term is the contribution to the mean rate of reaction from the mean concentration and the second term is the contribution to the mean rate of reaction from the fluctuating concentration. We see that in this case the new dependent variable $\bar{c}_i'^2$ enters equation 2.

Equation 2 is a rigorously valid equation for $\bar{c}_i$ (neglecting, of course, molecular diffusion); and, if the variables $u' \bar{c}_i'$, $v' \bar{c}_i'$, $w' \bar{c}_i'$, and any of those arising from $R_i$ are known as functions of space and time, it can be solved in principle to yield $\bar{c}_i$. Unfortunately, $u' \bar{c}_i'$ and so on cannot be measured at all points in an atmospheric flow and cannot be predicted exactly because of the closure problem of turbulent flow. Thus, we must resort to models for these terms. The model employed in virtually all cases in which atmospheric flows are involved is that based on the concept of eddy diffusivities:

$$u' \bar{c}_i' = -K_u \frac{\partial \bar{c}_i}{\partial x} \quad v' \bar{c}_i' = -K_v \frac{\partial \bar{c}_i}{\partial y} \quad w' \bar{c}_i' = -K_w \frac{\partial \bar{c}_i}{\partial z} \quad (3)$$

The eddy diffusivities $K_u$ and $K_v$ are postulated to be functions of space and time (and not of $\bar{c}_i$ or any of its gradients).

Although there has been some study of the nature of terms of the form $c_i c_j'$ arising from turbulent chemical reactions, no atmospheric diffusion models for chemically reactive pollutants currently include expressions for these terms. All models employ the approximation

$$R_i(c_1, \ldots, c_n, T) \approx R_i(\bar{c}_1, \ldots, \bar{c}_n, T) \quad (4)$$

The result of using equations 3 and 4 in equation 2 is the so-called atmospheric diffusion equation (ADE):

$$\frac{\partial \bar{c}_i}{\partial t} + u \frac{\partial \bar{c}_i}{\partial x} + v \frac{\partial \bar{c}_i}{\partial y} + w \frac{\partial \bar{c}_i}{\partial z} = \frac{\partial}{\partial x} \left( K_u \frac{\partial \bar{c}_i}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_v \frac{\partial \bar{c}_i}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_w \frac{\partial \bar{c}_i}{\partial z} \right)$$

$$+ R_i(\bar{c}_1, \ldots, \bar{c}_n, T) + \bar{S}_i(x,y,z,t) \quad (5)$$

Equation 5 is the fundamental equation on which most current urban air pollution models for photochemical pollutants are based.
The validity of the atmospheric diffusion equation relates to how closely the predicted mean concentration \( \bar{c} \) corresponds to the true ensemble mean concentration. If the mean velocities, \( \bar{u}, \bar{v}, \) and \( \bar{w} \), are known precisely at all points as a function of time, then, for an inert species, the only source of a discrepancy between the predicted and true mean concentrations is the eddy diffusivity model for the turbulent fluxes. Thus, for an inert species, the only source of invalidity of the atmospheric diffusion equation is, in principle, the eddy diffusivity closure of the turbulent flux terms.

If the true ensemble mean velocities and concentrations are known for an atmospheric flow, then it is relatively straightforward to assess the validity of equation 5 for specified forms of \( K_H \) and \( K_D \). Unfortunately, for any atmospheric flow the ensemble mean velocities and concentrations can never be computed since the atmosphere presents only one realization of the flow at any time. (Of course, for a statistically stationary flow, ensemble averages can be replaced by time averages. The atmosphere is, however, seldom in a stationary condition for any appreciable period of time.) Figure 1 shows a hypothetical relation among the instantaneous concentration measured at a point, the ensemble mean concentration at that point (which is unknown), and the mean concentration predicted by the atmospheric diffusion equation at the same point. Because the true mean velocities that are required to solve equation 5 and the true mean concentration with which the solution of 5 is to be compared are not available in general, an unambiguous measure of the validity of equation 5 for any particular flow cannot be obtained.

To assess the validity of equation 5 requires a flow field for which all desired statistics can be obtained. Since the measurement requirements needed to obtain such statistics in the atmosphere are so overwhelming as to be totally impractical, one must turn to numerical simulations of turbulent flow to provide a flow field within which to perform atmospheric diffusion "experiments." (Such numerical simulations are not, of course, completely exact representations of turbulent flow since a closure assumption must be invoked at some point in the solution of the equations. In Deardorff's simulation the sub-grid-scale turbulent transport becomes a parameter by an eddy viscosity.) Apparently the only examination of the validity of conventional atmospheric diffusion theories with such a numerical simulation is that of Lamb et al. (17), in which a continuous line source was embedded in Deardorff's planetary boundary layer model (4). The mean concentration of an inert pollutant issuing from the line source was computed by averaging over individual particles released in the nearly stationary and homogeneous flow. The "real" mean concentration was then compared to that predicted by the atmospheric diffusion equation as well as conventional Gaussian puff and plume formulas. With the availability of the planetary boundary layer simulation of Deardorff, a number of numerical experiments are possible to assess the validity of the atmospheric diffusion equation.

1. Theoretical forms for eddy diffusivities, particularly \( K_D \), can be tested by comparing the mean concentration field predicted by the diffusion equation with the "true" mean concentration field. [Experiments of this type were carried out by Lamb et al. (17), and further work by Lamb is still in progress.]

2. Chemically reacting flows can be simulated by direct solution of equation 1. True mean concentrations can be determined by time-averaging the resulting concentrations (since the flow is essentially stationary). Then closure approximations for terms of the form \( c_t/c_f \) can be evaluated by comparing the solution of the atmospheric diffusion equation to the data. (Work is in progress on this aspect by Wengle and Seinfeld.)

In summary, assessment of the validity of the atmospheric diffusion equation is an extremely difficult task owing to the fact that mean concentrations for actual flows are nearly impossible to obtain. However, with the advent of numerical turbulence simulations, such as that of Deardorff, such assessments are becoming feasible and will be important areas for future investigation.
ACCURACY OF ATMOSPHERIC DIFFUSION EQUATION

Whereas validity refers to the extent of agreement of the mean concentrations predicted by the atmospheric diffusion equation with those from an actual flow, accuracy refers to the extent of agreement between the mean concentrations predicted by the model based on the diffusion equation and those predicted by the equation with the true values of all input parameters. Thus, if \( \bar{u}, \bar{v}, \) and \( \bar{w} \) are the true mean velocities, \( \bar{R} \) is the true chemical reaction rate, and \( \bar{S} \) is the true source emission function, a \( \bar{c} \) will be predicted that can be taken as the most accurate concentration achievable with that equation. Errors introduced as a result of the facts that \( \bar{u}, \bar{v}, \) and \( \bar{w} \) are not the actual mean velocities, \( \bar{R} \) may not be based on the actual reaction mechanism, and \( \bar{S} \) is not the true source emission function will serve to reduce the accuracy of the model when compared to the most accurate achievable.

Table 1 gives the errors inherent in most urban diffusion models; errors are classified according to whether they are a source of invalidity or inaccuracy. Table 2 gives the sources of discrepancy between the mean concentration predicted by the atmospheric diffusion equation and the concentration measured at a particular location in an urban area. The sources of errors in the atmospheric diffusion equation are shown in Figure 2.

The inputs needed to solve ADE together with possible sources of error in those inputs are given in Table 3. In this section we discuss each input and the likely level of error associated with that input. In each instance unless the actual value of the input is known, the level of error in that input can only be estimated. From the standpoint of the effect of errors on the predictions of the equation, joint consideration must be given to the level of uncertainty in each input parameter and the sensitivity of the predicted concentrations to the parameter. Uncertainty relates to the possible error in the parameter from its true value, and sensitivity refers to the effect that variation in that parameter has on the solution of the equation. A parameter may have a large uncertainty associated with it but have little influence on the solution. In such a case, effort at reducing the uncertainty in the parameter value may be unwarranted. On the other hand, small uncertainties in a parameter to which the solution is quite sensitive will have a large impact on uncertainties in the predicted concentrations. Thus, both uncertainty and sensitivity must be considered when the accuracy of the atmospheric diffusion equation is evaluated.

Meteorological Inputs

Practical difficulties of solving the coupled mass, momentum, energy, and species continuity equations for atmospheric pollutant dynamics create a situation in which only the continuity equations of the form of equation 5 are used. The implication of using only the species continuity equation is that a description of the wind field at a resolution equivalent to the grid is required. Unfortunately, the atmospheric measurements required to specify the state of the wind field to this level of accuracy are usually not available. Because the wind field can be specified in terms of only a limited number of observations on a much coarser grid than that on which equation 5 is solved, individual fluid motions with scales smaller than the grid are uncharacterized. A key problem in the use of equation 5 is how to create from the observed meteorology information that is required for the solution of equation 5.

There are 2 basic approaches to obtaining the wind field; the objectives in both cases are to represent the wind field in a way that minimizes irregularities and observational errors. Traditionally meteorologists have drawn isopleths and then used them as a basis for estimating the wind components. An obvious problem with this method is that any one analysis is subject to personal interpretation. The problems of this subjective technique can be avoided if the meteorological variables are objectively related to the space coordinates.

Objective analysis techniques are now used on a routine basis to construct fields of optimized meteorological variables at specified model grid points from irregularly
Figure 1. Hypothetical relations among concentrations.

\[ c_i(x,y,z,t) = \text{instantaneous concentration at a point} \]

\[ \xi_i(x,y,z,t) = \text{ensemble mean concentration} \]

\[ \xi_i(x,y,z,t) = \text{mean concentration predicted by atmospheric diffusion equation} \]

Table 1. Errors in the atmospheric diffusion equation.

<table>
<thead>
<tr>
<th>Source of Error</th>
<th>Type of Source and Error</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>True form of the turbulent fluxes, ( u^'c_i, v^'c_i, ) and ( w^'c_i ), is unknown</td>
<td>Major source of invalidity of atmospheric diffusion equation (ADE)</td>
<td>Higher order closure models will offer improvement over eddy diffusivities in representing these terms</td>
</tr>
<tr>
<td>Chemical reaction mechanism ( R ), does not accurately reflect the actual chemistry</td>
<td>Source of invalidity for chemically reacting species</td>
<td>Continued study of laboratory chemistry and its relation to atmosphere needed to minimize this source of error</td>
</tr>
<tr>
<td>Turbulent fluctuating chemical reaction terms are neglected</td>
<td>Secondary source of invalidity of ADE</td>
<td>Closure models appropriate for turbulent chemistry can be developed</td>
</tr>
<tr>
<td>Mean velocities ( \bar{u}, \bar{v}, ) and ( \bar{w} ) used in solution of ADE are not true ensemble means (usually ( \bar{u}, \bar{v}, ) and ( \bar{w} ) are constructed from data at a finite number of locations)</td>
<td>Source of inaccuracy of ADE</td>
<td>There is no way to determine true mean from data; ( \bar{u}, \bar{v}, ) and ( \bar{w} ) can be calculated from accurate fluid mechanical turbulence model</td>
</tr>
<tr>
<td>Source emission function ( S ), is inaccurate</td>
<td>Source of inaccuracy</td>
<td>Better compilation of emission factors needed</td>
</tr>
<tr>
<td>ADE must be solved on a grid of a size consistent with spatial detail in wind and source emission functions</td>
<td>Source of inaccuracy</td>
<td>Better resolution of wind and source emission fields will allow use of finer grid</td>
</tr>
</tbody>
</table>

Table 2. Discrepancies between predicted and measured concentration.

<table>
<thead>
<tr>
<th>Source of Discrepancy</th>
<th>Possible Remedy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration measured on a given day is result of random event</td>
<td>None</td>
</tr>
<tr>
<td>Concentration is measured at a point, but ADE is applied to a grid volume</td>
<td>Reduce grid size in calculation</td>
</tr>
<tr>
<td>Measurement errors</td>
<td>Place monitoring point at a location where air quality is representative of a region the order of the grid size</td>
</tr>
<tr>
<td></td>
<td>Develop a sub-grid-scale model</td>
</tr>
<tr>
<td></td>
<td>Better monitoring equipment</td>
</tr>
</tbody>
</table>

Figure 2. Sources of errors in atmospheric diffusion equation.

\[ \text{Spatial averaging inherent in } \xi_i \]

Not the true turbulent fluxes

Not the true chemical reaction mechanism

Turbulent fluctuating reaction rates neglected

Not the true source emission inventory

Table 3. Inputs needed to solve atmospheric diffusion equation.

<table>
<thead>
<tr>
<th>Input</th>
<th>Level of Detail</th>
<th>Sources of Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind velocities</td>
<td>Variation of ( \bar{u}, \bar{v}, ) and ( \bar{w} ) with ( x, y, z, ) and ( t )</td>
<td>Measurements available only at a few locations, generally at ground level, at discrete times</td>
</tr>
<tr>
<td>Eddy diffusivities</td>
<td>Variation of ( K_u ) and ( K_v ) with ( z ) and ( t )</td>
<td>Errors in variation of ( \bar{u} ) and ( \bar{v} ) horizontally, variation of ( \bar{u} ) and ( \bar{w} ) vertically, and determination of ( \bar{w} )</td>
</tr>
<tr>
<td>Chemical reaction mechanism</td>
<td>Rate equation for each ( \xi_i )</td>
<td>No direct measurements available</td>
</tr>
<tr>
<td>Source emissions</td>
<td>Emission rate as a function of ( x, y, z, ) and ( t )</td>
<td>Both magnitude and vertical variation unknown</td>
</tr>
<tr>
<td>Boundary conditions</td>
<td>Location of vertical boundary as a function of ( x, y, ) and ( t )</td>
<td>Inaccurate rates because of inability to simulate atmospheric chemistry in laboratory, unknown reactions, and unknown rate constants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inaccurate knowledge of level of source activity and emission factors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of data or adequate model of temperature structure of atmosphere</td>
</tr>
</tbody>
</table>
spaced observational data. Panofsky (25) introduced objective analysis with a technique of using cubic polynomials to fit observed pressure data and obtaining the coefficients from a least squares analysis. Independent polynomials were obtained for separate areas with the consequence that, at the boundaries, the field was discontinuous. Since the discontinuity has no physical basis, a method that smooths the data along the edges is necessary. Gilchrist and Cressman (14) modified the approach of Panofsky by considering only the data surrounding the grid point and then fitting a "most probable" second order polynomial to the observation. This procedure conforms to conventional synoptic practice in that it does not include data in remote parts of the region as implied by any method using Fourier analysis or polynomial fitting to the whole region.

Some improvements can be made on the interpolation and matching stages (13). One example to create an initial approximation of the field at a point \( (x,y) \) is to use polynomial surface fitting with the data for the least squares analysis weighted by an equation of the following type:

\[
\overline{u}(x,y,0,t) = \sum_{i=1}^{M} f(x - x_i, y - y_i) u_{\text{obs}}(x_i,y_i,t)
\]  

Liu and Goodin (19) and Liu and Perrine (20) omitted the surface fitting and simply used a least squares form of \( f \) with the number of data points, \( M \), dictated by observations within a specified radius around \( (x,y) \). There are various other schemes, the simplest being to use the value of the nearest observation. Wind fields created by these techniques are not, however, necessarily consistent with the condition of mass conservation.

Interpolation and surface-fitting techniques can be used as an initial basis for subsequent operations that remove the artificial divergence and convergence introduced by numerical and observational errors. The most common approaches have been to apply various iterative schemes to minimize the divergence of 2-dimensional wind fields in an attempt to satisfy the constraint:

\[
\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} + \frac{\partial (vh)}{\partial y} = 0
\]

where \( h \) is the height of the inversion base above the topography and \( u \) and \( v \) are the \( x \) and \( y \) components of \( \mathbf{u}(x,y,0,t) \). When the approximation schemes are used, a residual other than zero appears on the right side of equation 7, i.e.,

\[
\frac{\partial h}{\partial t} + \frac{\partial (uh)_a}{\partial x} + \frac{\partial (vh)_a}{\partial y} = \epsilon
\]

where \((,)_a\) indicates that the approximate values are used. Sasaki (28) introduced a method based on variational calculus that allows for the adjusted values of the dependent variables to satisfy exactly an imposed constraint (\( \epsilon = 0 \)). Sasaki demonstrated that a Lagrange multiplier \( \lambda \) should satisfy

\[
\nabla^2 \lambda = - \frac{\partial h}{\partial t} + \frac{\partial (uh)_a}{\partial x} + \frac{\partial (vh)_a}{\partial y}
\]

with \( \lambda = 0 \) on the domain boundaries. The adjusted values of \( u \) and \( v \) can be determined.
from a solution of equation 9 by using

\[ uh = (uh)_0 + \frac{\partial \lambda}{\partial x}, \quad vh = (uh)_0 + \frac{\partial \lambda}{\partial y} \]  

(10)

Sasaki (30, 31, 32) introduced a modification of his original formulation by allowing for a reduction of \( \epsilon \) over the grid not necessarily forcing \( \epsilon = 0 \) during adjustment. Dickerson (6, 7) used both formulations to develop mass consistent wind fields for the San Francisco Bay Area (18). Endlich (9, 10) and Endlich and Mancuso (11) developed a point iterative scheme that minimizes the divergence but at the same preserves the vorticity. Liu and Goodin (19) have compared these techniques and their practical validity in creating wind fields for the Los Angeles basin. As a result of their work on the Sasaki and Endlich methods, they proposed a model using the best features of both techniques. The usual method of computing nondivergent and irrotational wind fields is to solve Poisson equations for the stream function and velocity potential (1, 2). There are important practical difficulties with this approach because of the difficulty of specifying satisfactory boundary conditions (15).

Of the approaches described above, most have been applied to the generation of 2-dimensional fields. Lange (18) employed Dickerson's technique to specify a 3-dimensional wind field for a particle in cell formulation of equation 5 for conservative pollutants. Until quite recently most 3-dimensional models of pollutant transport have used subjectively prepared wind fields. Meteorological inputs to the current version of the model of Systems Applications, Inc., have been prepared from observational data by using the methods of Fankhauser (12) and O'Brien (24).

Fankhauser used the horizontal wind field \( u(x,y,t) \) to obtain a vertical profile of divergence at each grid point using

\[ D_k = \frac{\partial u}{\partial x} + \frac{\partial N}{\partial y} \]  

(11)

where the horizontal velocity components are the mean for each grid cell of thickness \( K \). With the use of the continuity equation, integration of the divergence profile with respect to pressure gives the vertical component of velocity \( w_p \) at the top,

\[ w_p = w_{pr\Delta p} + D_k \Delta p \]  

(12)

The vertical velocity computed in this way generally produces reasonable values at low levels, but leads to physically unrealistic results in the upper layers. This is a problem common to many techniques and is primarily attributable to the fact that the quality of wind measurements (as distinct from pressure) deteriorates with increasing elevation angles because of the combination of winds aloft and short sounding duration.

O'Brien (24) developed a scheme used by Fankhauser for altering the original divergence estimates so that the adjusted profiles satisfy some independently defined boundary condition on \( w \) at the inversion base and at ground level. The method is based on the hypothesis that divergence errors increase vertically as a linear function of pressure.

The generation of mass consistent wind fields can be achieved by a number of techniques. Which method produces the best result is largely a question of the quality of the observational data and the degree to which the field is constrained by boundary conditions. Incorporation of the effects of wind shear and different vertical velocity profiles should really be considered only if there are sufficient observational data to determine the power law exponents. (The RAPS study in St. Louis should provide
Experimental data on wind fields aloft to resolve some of the presently unknown boundary conditions.) Until complete planetary boundary layer models can be developed and validated, photochemical modeling must continue to rely on the use of mass consistent wind fields derived from observational data.

Chemical Reaction Mechanisms

Because of the large number of hydrocarbon species in urban air, generalized kinetic mechanisms capable of inclusion in the diffusion equation had to be developed. The source of errors in generalized kinetic mechanisms for photochemical smog are twofold (5):

1. Errors in simulating atmospheric reactions in the laboratory—(a) inadequate or no control and measurement of levels of H₂O in the chamber, (b) impurities in background chamber air, (c) inadequate or no measurement of the spectral distribution and intensity of the chamber irradiation system, (d) inaccurate or ambiguous analytical methods, (e) nonhomogeneity due to inadequate stirring or poor chamber design, (f) adsorption and desorption of reactants and products on chamber walls, (g) chemical reactions occurring on chamber surfaces, and (h) inadequate control and measurement of chamber temperature; and

2. Errors associated with the chemical kinetic mechanism—(a) uncertainties in experimental determinations of specific reaction rate constants, (b) variations of rate constants with temperature either uncertain or unknown, (c) inadequacies in lumping due to the nonrepresentativeness of lumped class reactions relative to species within the class, e.g., reaction rates, products, and stoichiometric coefficients, and (d) inaccuracies in the mechanism due to insufficient verification studies.

As Demerjian points out, most recent smog chamber data have been well characterized with respect to the first category above. A combined uncertainty and sensitivity study of the Hecht, Seinfeld, and Dodge (16) kinetic mechanism for photochemical smog has been carried out by Dodge and Hecht (8). In that study it was found that several reactions previously included in the mechanism have little effect on the concentration predictions and can be neglected. It was also found that the rate constants of 12 reactions should be more accurately determined. It is important to note that this study was concerned not with the validity of the mechanism vis-à-vis atmospheric chemistry, only with the uncertainty of the rate constants and the sensitivity of the concentrations to those rate constants. Thus, assuming that the mechanism is a valid representation of atmospheric chemistry, the study will determine the level of uncertainty in predicted concentrations resulting from uncertainties in rate constants. Although the Dodge and Hecht study was carried out for a batch laboratory system, the combined uncertainty-sensitivity results will apply in atmospheric calculations.

Source Emissions

Source emissions from each class of source can be characterized according to

1. Level of spatial resolution,
2. Level of temporal resolution, and
3. Uncertainty in emission quantity, i.e., in source activity or in emission factor.

The level of spatial resolution achievable is generally as fine as one desires since the locations of all sources can presumably be specified (although traffic count data may not be available on a street-by-street basis). Temporal emission rates will fluctuate some from day to day. For motor vehicles it is safe to use single temporal distributions for surface streets and freeways derived from the analysis of appropriate data. Emissions from some stationary sources may vary with ambient temperature,
but these variations are generally known as a function of temperature. The major problem in properly specifying source emissions is uncertainty in emission quantities arising from uncertainties in source activities and emission factors.

Typical levels of uncertainties in mobile and fixed source activities (e.g., number of vehicle miles traveled and number of units of fuel consumed) should be identified. Then, the typical uncertainties in emission factors (e.g., grams per mile of pollutant emitted per vehicle mile traveled and gram of pollutant per unit of fuel consumed) should be combined with the uncertainties in activities to produce net uncertainties in emissions. Finally, sensitivity studies should be carried out with urban diffusion models to assess the effect of these levels of uncertainties in source emissions on predicted concentrations.

Initial and Boundary Conditions

The initial condition for equation 5 is the concentration field at the time corresponding to the beginning of the simulation, \( c_1(x,y,z,0) \). Simulations are normally begun at sunrise, and the \( c_1 \) field at that time is constructed from the station readings. A ground-level interpolation routine and assumptions regarding the vertical variation of the concentrations are required to generate the full \( c_1 \) field from the station data. To minimize the effect of errors due to inaccurate initial conditions, one would like to simulate several days in succession, starting with a day at the beginning of which the concentrations are relatively low. Because readings are available from which to construct a \( c_1 \) field, we do not expect a high level of uncertainty in the initial conditions when a previous time is simulated.

The boundary conditions for equation 5 consist of the concentrations upwind of the region, the pollutant fluxes at the ground (the source emissions), and the flux condition at the upper vertical boundary of the region. Concentrations upwind of the modeling region can be estimated if monitoring stations exist at the upwind edge of the airshed. In such a case uncertainties in these concentrations will be low when a previous time is simulated. The major source of uncertainty in boundary conditions generally arises at the upper vertical boundary. First, the temperature structure, for example, the height of the base of an elevated inversion layer, is not known precisely. Second, the pollutant flux condition at the boundary is also not known precisely.

Thus, the major uncertainty in boundary conditions when past time periods are simulated is in specifying the upper vertical boundary conditions, both the location of the boundary and the species flux condition at the boundary. Sensitivity studies should be carried out on the location of an upper boundary of no flux (an inversion base) to determine the effect of uncertainties in inversion base height on predicted concentrations.

A special problem arises in specifying initial and boundary conditions when a model is to be exercised under hypothetical emission conditions, for example, when the effect of emission control strategies is studied. Initial concentrations at sunrise, say, reflect nighttime emissions as well as, perhaps, the previous day's concentration levels if the airshed is not thoroughly flushed at night. For an inert pollutant such as CO it is reasonable to assume that initial and boundary concentrations under future conditions will reduce in direct proportion to emission reductions with account taken of the background levels (the rollback assumption). Although hydrocarbons are reactive, their initial and boundary levels can also be assumed to vary in direct proportion to emissions. The problem arises when NO and \( O_3 \) are considered. Two approaches to specifying initial and boundary conditions for NO, NO\(_2\), and \( O_3 \) were developed by Reynolds and Seinfeld (27).

1. The first method of calculation is based on the assumption that NO and NO\(_2\) concentrations at sunrise and at the boundaries of the airshed will be reduced from the base-year values in direct proportion to the reduction in total NO\(_x\) emissions, i.e.,

\[
[NO] = \phi_{NO_x} [NO]_b
\]
\[ [NO_2] = \phi_{NOx} [NO_2]_b \]

where \( \phi_{NOx} \) is the fractional reduction in NO\(_x\) emissions from the base year, and \( B \) denotes base-year values. Assuming a photostationary state for O\(_3\) concentration (24),

\[ [O_3] \sim \frac{[NO_2]}{[NO]} \]

we see that [O\(_3\)] concentrations will remain the same as in the base year since the ratio of [NO\(_2\)] to [NO] remains the same.

2. The second approach for specifying the NO, NO\(_2\), and O\(_3\) initial and boundary condition is suggested by the rollback concept, in which O\(_3\) concentrations are assumed to be proportional to hydrocarbon emissions and total NO\(_x\) concentrations are assumed to be proportional to NO\(_x\) emissions. These assumptions lead to the following expressions for future-year concentrations:

\[
\frac{[NO] + [NO_2]}{[NO]_b + [NO_2]_b} = \phi_{NOx}
\]

\[
\frac{[NO_2]/[NO]}{[NO_2]_b/[NO]_b} = \phi_{HC}
\]

where \( \phi_{HC} \) is the fractional reduction in hydrocarbon emissions relative to the base year.

Additional study should be directed to the problem of specifying initial and boundary conditions for emission control exercises.

SUMMARY

This paper discusses the concepts of validity and accuracy for air pollutant model calculations and suggests a number of numerical experiments that should be carried out on an urban diffusion model to test its validity and accuracy. These studies are summarized as follows:

1. Perform diffusion experiment in numerical turbulence field and compare data with conventional air pollution models using eddy diffusivities (17, and work in progress by Lamb);

2. Perform diffusion experiment for reactive species in numerical turbulence fields and compare data with conventional air pollution models with various reactive closure models (work in progress by Wengle and Seinfeld);

3. Perform sensitivity analysis of photochemical diffusion model to variations in meteorological parameters (21); and

4. Perform sensitivity analysis of photochemical diffusion model to variations in initial, boundary, and source concentrations (work in progress by McRae, Goodin, and Seinfeld).

ACKNOWLEDGMENT

The author wishes to extend appreciation to Gregory McRae for his help in preparing this paper.
REFERENCES


