United States
Environmental Protection
Agency

Atmospheric Research and
Exposure Assessment Laboratory
Research Triangle Park, NC 27711

Research and Development

EPA/600/S3-89/071 Sept. 1989

EPA

Project Summary

Air Quality Simulation Model Performance for One-Hour Averages

G. E. Moore, T. E. Stoeckenius, Steve Hanna, and Dave Strimitas

If a one-hour standard for sulfur dioxide were promulgated, air quality dispersion modeling in the vicinity of major point sources would be an important air quality management tool. Would currently available dispersion models be suitable for use in demonstrating attainment of such a standard in the vicinity of large, elevated, buoyant point sources such as utility power plants? The results summarized in this report suggest that using these models in connection with an hourly average standard does not present the regulatory community with any significant additional uncertainties that are not already being dealt with in connection with the current 3- and 24-hour standards. Predictions of peak hourly average concentrations unmatched in time or location are slightly more conservative on average than corresponding predictions of 3- and 24-hour averages. Although certain conditions not specifically treated by the models, such as inversion break-up fumigation, are known to occur and to have the potential for producing high ground-level impacts, these conditions have been only rarely observed in model evaluation studies and do not seem to contribute to the bulk of the 50 highest concentrations. In fact, nearly all of the highest concentrations observed in the evaluation studies examined occurred under convective conditions, which are treated by the existing models. Peak model predictions under such conditions match reasonably well with the highest observed concentrations, although large prediction errors can occur in any given hour. Consideration of situations involving impacts on elevated terrain or building downwash phenomena have been explicitly excluded. This review of the results of model evaluation studies was restricted to some fairly simple analyses. Additional analyses are recommended. Recommendations to improve model formulations are made regarding the treatment of plume behavior in the vicinity of the mixing height and the simulation of plume dispersion within the mixed layer during convective conditions.

This Project Summary was developed by EPA's Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

An analysis of the results of previous dispersion model performance evaluation studies has been conducted to ascertain the suitability of currently used models for the prediction of hourly average concentrations. This investigation was motivated by the potential promulgation of a statistical (expected exceedance) 1-hour SO_2 NAAQS. Implementation of such a standard would rely largely on the use of dispersion models that are capable of accurately predicting the whole upper tail of the distribution of hourly average SO_2 concentrations, not just the highest or second highest concentrations.

Our analysis was confined to the study of model performance for elevated, buoyant point sources of the type typically associated with fossil fuel-fired power plants. Such sources represent one of the major areas of model application for the prediction of SO_2 air quality impacts. We conducted analyses designed to address the following issues:

How well do models predict the 1, 2, . . . , N highest concentrations irrespective of time or location matching?

How well do models predict the 1, 2, . . . , N highest observed concentrations when observations and predictions are matched in time but not location?

How well do models predict concentrations occurring under specific meteorological conditions? More specifically, what meteorological conditions produce the highest observed concentrations? What conditions are associated with the largest discrepancies between observed and predicted concentrations?

To answer the above questions, we reviewed the published results of previous model performance studies and, whenever possible, obtained the aerometric and meteorological data from those studies for further analysis.

Model Performance: Ability to Predict Peak Concentrations

Short term (1- to 24-hour average) ambient air quality standards are designed to control the magnitude and frequency of occurrence of peak ground-level concentrations. Demonstrating that the emissions from a major source do not violate such standards often involves the use of a dispersion model to predict the highest concentrations occurring over a one- to five-year period in the vicinity of the source. A model's ability to accurately and precisely predict peak concentrations is therefore very important. However, a model need not necessarily predict the exact time and location of the peak values as this information is not usually of direct relevance to the attainment demonstration. On the other hand, examining a model's ability to reproduce the N highest observed concentrations during the hours in which they occur (i.e., matching concentrations in time) provides insight into the types of meteorological conditions under which models do especially well or poorly and the reasons for such performance. This information, in turn, can be used as a guide to improving models and to the level of performance that can be anticipated from a given model under a given situation.

Concentrations Unmatched in Time or Location

We examined the performance of models in predicting the N highest concentrations (typically around 25) regardless of their time or location of occurrence. In some cases, only comparisons of the highest and second highest observed and predicted concentrations were available. We caution the reader that it is difficult to draw generalizations from such comparisons because the highest and second highest concentrations usually have a large degree of variability.

With the exception of the Maryland Power Plant Siting study (for which a value of N = 50 was used), the CRSTER and equivalent models tend to overpredict slightly the peak concentrations (N less than or equal to 25) or produce no significant bias. TEM, on the other hand, underpredicts at some sites and produces no significant bias at others. The only other models for which evaluation results are available from more than one site are PPSP and MPSDM. PPSP overpredicts consistently at all sites, in some cases by more than a factor of two. MPSDM also produced significant overpredictions.

Model Performance for Hourly Averages Compared with Performance for 3- and 24-Hour Averages

Would implementation of a 1-hour average NAAQS for SO_2 result in significantly greater modeling uncertainty than currently exists in demonstrating attainment of the 3- and 24-hour NAAQS? To answer this question, we reviewed an analysis of the variation in model performance with averaging time by Hayes. This study showed that the ratio of the average of the 25 highest predictions to the average of the 25 highest observations (unmatched in time or location) decreases with increasing averaging time for all models included in the EPA Rural Model Evaluation Study, the CEQM model at EPRI PMV Kincaid, and for RAM (but not TEM) in the EPA Urban Model Evaluation Study. Results for ratios of the highest and second highest concentrations were mixed. Thus, models that overpredict peak 3- and 24-hour averages tend to overpredict peak hourly averages by a greater amount, while models that underpredict peak 3- and 24-hour averages may overpredict peak hourly averages. This suggests that, at least as far as bias goes, shifting emphasis from 3- and 24-hour averages to 1-hour averages will result in more conservative predictions of peak concentrations.

The study also examined the correlation between observations and predictions paired in time but not location as a function of averaging time at the EPRI PMV Kincaid site using the CEQM model. The results indicate that correlations for 3-hour averages are only slightly better than those for hourly averages, while the correlation for 24-hour averages is not significantly different from that of hourly averages. Thus, the model's skill in predicting hourly averages does not appear to be very much worse than it is for 3- and 24-hour averages.

Concentrations Matched in Time

A model's ability to predict peak concentrations can also be examined by comparing the N highest observed concentrations with the highest concentrations predicted during the same hours (i.e., compare concentrations matched in time but not location for the hours with the N highest observed concentrations). Such a comparison evaluated the model's ability to reproduce high observed concentrations under the specific meteorological conditions that cause them (exclusive of wind direction).

As has been demonstrated by EPRI and other studies, correlations between time- matched observed and predicted concentrations are generally very low or not significantly different from zero. Correlations for just the hours with the 50 highest observed concentrations are not significantly different from zero except for the PPSP and RTDM models. Also, the number of predictions falling within a factor of two of the observations is less than 32 percent except for RTDM and PPSP.

On the basis of the above findings, we conclude that the poor model skill for time-paired observations and predictions is primarily the result of two factors:

1. Meteorological inputs to the dispersion models do not reliably

represent atmospheric conditions occurring in a particular hour.
2. Model formulations are insufficient to describe the actual plume behaviors that lead to the highest observed concentrations.

Model Performance Under Specific Dispersion Conditions

One concern with using presently available dispersion models to perform attainment demonstrations for an hourly average ambient standard is that these models may not adequately reproduce all of the various types of dispersion phenomena that can lead to high hourly average concentrations. We performed several analyses in an attempt to identify dispersion conditions and plume behaviors that are associated with high observed concentrations and to estimate how well the models perform under such conditions.

Plume Behaviors and Dispersion Conditions Potentially Resulting in High Concentrations

A great deal of meteorological and aerometric data is needed to adequately explain the specific dispersion conditions and plume behaviors that lead to a high observed hourly average concentration. No model evaluation study carried out to date has involved the collection of enough data to allow one to unequivocally group hours with high observed concentrations into a handful of well-defined categories of dispersion conditions and plume behaviors. Enough information from previous theoretical, field, and laboratory studies is now available to make it possible to describe several general types of plume behaviors and dispersion conditions that can, and in some cases have actually been observed to, lead to high ground-level concentrations or large discrepancies between observed and predicted concentrations in the vicinity of elevated, buoyant point sources. Plume behaviors considered are: convective fumigation, moderate wind convection, bumping-looping, limited rise, plume penetration, and inversion break-up or shoreline fumigation.

Applicability of Model Formulations to Important Plume Behaviors

No presently available point source dispersion model contains all of the algorithms needed to make accurate predictions under all of the plume behaviors. There are, however, sufficient differences between the models included in our review to suggest that some may be better equipped to deal with these phenomena than others. For example, plume behaviors such as bumping-looping and partial penetration in which the plume centerline is near the mixing height can prove to be difficult for models to reproduce accurately unless some sort of partial penetration algorithm is included in the model. Neither the CRSTER/MPTER/CEQM or SHORTZ models have such an algorithm. These models assume a zero ground-level impact if the plume centerline is predicted to be above the mixing height and a perfect reflection of plume material off of the inversion base otherwise. This oversimplified approach can lead to large prediction errors.

Accurate estimates of plume height during convectively unstable conditions are needed to produce accurate predictions of ground-level impacts. The plume breakup/touchdown algorithm in the PPSP model represents an improvement over the standard Briggs formulas used in the other models that is designed to meet this need.

Models that rely solely on the PGT stability class to estimate the rate of plume dispersion may have difficulty differentiating between the effects of various convective and mechanically driven turbulence phenomena and may use inaccurate dispersion coefficients for some hours since the PGT classes are based solely on wind speed, cloud cover, and solar zenith angle. Some of the models we reviewed (RTDM, SHORTZ and PPSP) have the capability of using more direct measures of turbulence intensity in calculating dispersion coefficients. If accurate estimates of turbulence intensities are available, these models should produce realistic dispersion coefficients and concentration predictions.

Dispersion Conditions Associated with Highest Observed Concentrations

The highest observed concentrations frequently arise when a buoyant plume is trapped under a strong inversion and convective activity below the inversion mixes relatively undiluted material to the ground. High observed concentrations can also occur when plume heights are well below the mixing height and large-scale eddies rapidly mix plume material to the ground.

Dispersion Conditions Associated with Over- and Underpredictions

During the hours with the 50m highest observations the largest overpredictions tended to occur under unstable conditions whereas underpredictions were the general rule during neutral and stable conditions. The findings suggest that the largest underpredictions for models other than PPSP and RTDM are primarily the result of incorrect plume rise treatment during strong inversion conditions indicative of bumping/looping behavior. The analysis of the largest over- predictions at selected sites shows that overpredictions generally occur under unstable light wind conditions when the mixing height is relatively high, that is, convective fumigation conditions.

Implications for Model Applications in Connection with a One-Hour Sulfur Dioxide Standard

If an hourly average NAAQS for SO_2 were to be promulgated, would currently available dispersion models be suitable for use in demonstrating attainment of the standard in the vicinity of large, elevated, buoyant point sources as utility power plants? The results suggest that using these models in connection with an hourly average standard does not present the regulatory community with any significant additional uncertainties that are not already being dealt with in connection with the current 3- and 24-hour NAAQS. Predictions of peak hourly average concentrations unmatched in time or location are slightly more conservative on average than corresponding predictions of 3- and 24-hour averages, and model precision for observations and predictions matched in time is no worse for hourly averages than it is for longer averaging times.

Although certain conditions not specifically treated by the models, such as inversion break-up and seabreeze fumigation, are known to occur and to have the potential for producing high ground-level impacts from elevated sources, we found that these conditions have been only rarely observed in model evaluation studies and that they do not seem to contribute to the bulk of the 50 highest concentrations. In fact, nearly all of the highest concentrations observed in the evaluation studies we examined occurred under convective conditions, which are treated by most of the models as Pasquill stability class A and B cases. Peak predictions under such conditions match reasonably well with the highest observed concentrations, although large prediction errors can occur in any given hour. Thus, special dispersion situations that might be of concern in modeling peak hourly average concentrations do not appear to occur frequently enough around typical power plant sources to invalidate the use of currently available dispersion models. In reaching this conclusion, we have explicitly excluded consideration of situations involving impacts on elevated terrain or building downwash phenomena.

It should be noted that poor time-matched model performance may be due at least in part to the fact that model predictions represent a mean value over an ensemble of similar meteorological conditions while the observed values represent a single realization of an event from the ensemble. It may be possible to estimate the size of the contribution of this effect to the observed lack of model skill by comparing observed and predicted concentrations unmatched in time that belong to the same ensemble. Ensembles could be defined in terms of simple model inputs (e.g., all hours of A stability, very light winds, and low mixing heights).

Additionally, we recommend that additional data analyses and possibly field studies be carried out to identify situations in which inversion break-up and seabreeze fumigation as well as downwash phenomena occur and to characterize model performance under these conditions.

Recommendations for Improvements in Model Formulations

As pointed out above, one of the principal causes of poor model performance is the apparently unrealistic treatment of plume behavior in the vicinity of the mixing height. In particular, the overly simplified treatment of the vertical profiles of mean wind and stability parameters used in most models can lead to inaccurate estimates of plume height, especially in the vicinity of elevated inversions. This in turn can result in inaccurate estimates of ground-level concentrations. For example, the CRSTER/MPTER/CEQM and SHORT. models assume a zero ground-leve impact whenever the predicted plum height is greater than the mixing heigh Given the inevitable errors present i estimates of both of these quantities an the fact that the plume rise is calculate without regard to the limitation to ris imposed by the elevated inversion, zer ground-level concentrations are ofte predicted when in actuality substantia impacts are occurring. Substantia reduction of the number of cases c underprediction could be achieve through the use of more realisti treatments of the interaction of plume an inversions.

Another area in which mode improvements are needed is in th simulation of plume dispersion within th mixed layer during convective condition: The commonly used Pasquill stabilit classes and dispersion parameters do nc always accurately describe the degree c lateral and vertical dispersion as well a formulations based on directly observe stability parameters, such as u/w_* or th enhanced lateral dispersion of a plum during bumping-looping as discussed b Pierce.

G. E. Moore and T. E. Stoeckenius are with Systems Applications, Inc., San Rafael, CA 94903. Steve Hanna and Dave Strimitas are with Sigma Research Corp., Lexington, MA.
***D. Bruce Turner** is the EPA Project Officer (see below).*
The complete report, entitled "Air Quality Simulation Model Performance for One-Hour Averages," (Order No. PB 89-233 506/AS; Cost: $21.95, subject to change) will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650
The EPA Project Officer can be contacted at:
Atmospheric Research and Exposure Assessment Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

United States Environmental Protection Agency

Center for Environmental Research Information
Cincinnati OH 45268

Official Business
Penalty for Private Use $300

EPA/600/S3-89/071