Meteorological Data Needs for Modeling Air Quality Uncertainties

W. S. LEWELLEN AND R. I. SYKES

A.R.A.P. Division of California Research & Technology, Inc., A Titan Company, Princeton, New Jersey

(Manuscript received 17 June 1988, in final form 21 March 1989)

ABSTRACT

A probabilistic framework for incorporating uncertainty in air quality models is described. The quantitative dependence of the uncertainty in calculated air quality concentrations on the uncertainty in the input meteorological data is illustrated using a simple Second-order Closure Integrated Model Plume in combination with the EPRI Plume Model Validation and Development Data Set. Evaluation of the model results demonstrate that even though individual hourly samples cannot be deterministically predicted downwind of a powerplant stack, statistical representations of the observed cumulative distribution of the sample values are quite predictable. We discuss the data needed to improve the definition of the range of meteorological uncertainty within an ensemble of flows defined by given meteorological data, and thus provide for improvements in predictability models of the type illustrated. We argue that attempts to collect the data needed to define more precisely the variance within the ensemble of compatible flows will prove more productive than attempts to eliminate meteorological uncertainties in given datasets.

1. Introduction

A critical element of any air quality model is the input meteorological data. Simple models require little more than wind speed and direction, and a stability classification, while more complete models may use spatial and temporal distributions of the mean wind, temperature, and their associated turbulent fluctuation correlations. The accuracy with which a model can be expected to predict the dispersion of a pollutant in the atmosphere is clearly dependent on the input meteorological data. The fact that the data represents an ensemble of possible detailed flows gives an uncertainty in the prediction which is separate from those errors due to inaccurate input data or incorrect representation of the turbulent processes or chemical transformations in the model, and thus creates a component of model error which cannot be reduced by improving model physics or making more accurate meteorological measurements. The resulting variance in the ensemble concentration values has been referred to as the "inherent" uncertainty by the AMS Workshop on Model Uncertainty (Fox 1984), since it is a function of the type of input data, not its accuracy.

We are forced to consider stochastic fluctuations in any air quality prediction, since it is impossible to specify the complete wind distribution in time and space throughout the domain of interest. The meteorology must be described by a finite amount of data, effectively extended by the model to represent the entire flow field. The unknown differences between the actual wind field and the representative wind used by the model can only be represented stochastically in the model. The ability to quantify the uncertainty in air quality modeling as a function of the uncertainty in the input meteorological data is important so that confidence bounds on predictions can be realistically determined. Confidence bounds are necessary for any assessment of model predictions, because of the many cases when the confidence interval may be as wide as the predicted value.

An important part of addressing the problem of the meteorological data needs for air quality models, is the establishment of a framework for dealing with uncertainty in a quantitative way. The next section describes such a framework. This is followed by a specific example of results from a simple model constructed for application to the EPRI Plume Model Validation and Development dataset taken at the Kincaid Power Plant (Bowne et al. 1983). Finally, we speculate about what meteorological data might be taken to assist this type of model representation of air quality probabilities.

2. The stochastic dispersion problem

Clearly, we must deal with uncertainty in terms of probabilistic quantities, such as ensemble mean values, variances, and probability distributions. Virtually all current air quality models, however, can be considered as ensemble mean value predictors only. The models therefore require extension to provide further information on the inherent uncertainty. Lewellen et al.

Corresponding author address: Dr. Stephen Lewellen, A.R.A.P., 50 Washington Road, P.O. Box 2229, Princeton, NJ 08543-2229.

(1988) have attempted this by providing a model for the probability distribution function of the concentration fluctuations, while Irwin et al. (1987) present an alternative approach for incorporating meteorological uncertainty into the standard Gaussian plume models. The concept of a statistical distribution of the concentration observations, however, requires the definition of an ensemble of possible states over which the distributions can be defined. Venkatram (1982) has argued, in accord with the discussion in the previous section, that the ensemble is defined by the meteorological input data used by the model. In essence, the ensemble of concentration fields of interest is composed of all the fields which could be produced by any wind field that is consistent with the model input, observed wind data. The ensemble of concentration values thus depends directly on the observations used by the model.

The ensemble of wind fields can be considered to be made up of a spectrum of fluctuations. There is a high frequency, boundary-layer turbulence component, with eddy scales of the order of the mixed layer depth or smaller. These are the classical turbulent eddies; they are fully three-dimensional, and are virtually unresolvable with conventional meteorological instrumentation. Of at least equal importance is the larger scale variation in the wind field, which may be due to mesoscale circulations induced by temperature contrasts or topographic variations on scales larger than the boundary-layer scale. In principle, the larger scales are more amenable to direct measurement, but an intensive field experiment is still required to define the mesoscale wind field accurately. Rather than seeking to define the possible range of flows more narrowly by greatly increasing the meteorological data inputs, we believe it will generally be desirable to include this uncertainty into the air quality model framework.

The concept of an ensemble of wind fields compatible with a set of observations has also been developed by Lamb (1984) in connection with long-range transport of pollutants. There is considerable difficulty in the specification of the distribution of realizations within the ensemble, i.e., the quantification of the likelihood of any detailed wind field, and this is the area which most needs further meteorological data. The idea of such an ensemble is well defined, however, and allows us to proceed with our framework for the consideration of concentration fluctuations. We shall return to the problem of specifying the meteorological ensemble in section 3.

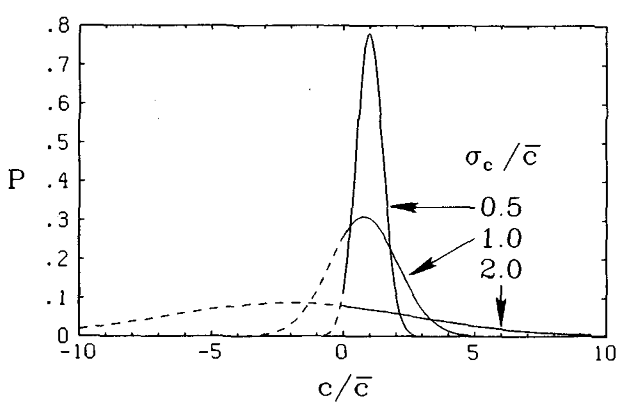

FIG. 1. Sketch of the clipped normal distribution for three values of $\sigma_c/\bar{c}$. The unrealizable negative concentration values are replaced by a delta function at zero, which represents intermittency.

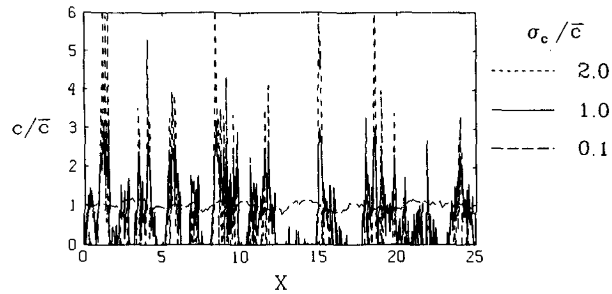

FIG. 2. Possible realizations of the time trace from a clipped normal distribution normalized by its mean value for three different ratios of $\sigma_c/\bar{c}$ as a function of time normalized by the correlation time scale.

Our discussion of the uncertainty problem so far leads us naturally to consider the concentration at a sampler location as a random variable, and hence to require the prediction of its probability distribution from an air quality model as the means of quantifying the uncertainty. Lewellen and Sykes (1983) proposed a method by which a prediction of the variance of the concentration fluctuations might be used for this task, and subsequent work has produced such a model. The variance model is based on second-order closure for the turbulent correlations, and is described in Sykes et al. (1984, 1986). The probability distribution is then estimated from the mean and the variance using a truncated Gaussian function, which replaces any negative tail in the Gaussian with a delta function at zero concentration. This distribution, illustrated in Fig. 1, was found to describe the fluctuations in lidar plume observations quite accurately (Lewellen and Sykes 1986).

If we assume that this clipped normal pdf does represent the random component of atmospheric dispersion, then Fig. 2 demonstrates possible realizations of the time trace of a sampler normalized by its mean value for three different ratios of $\sigma_c/\bar{c}$. To make Fig. 2 as realistic as possible we have included a correlation time scale, normalized to 1, and reduced the relative variance on times shorter than this. Clearly, only when $\sigma_c/\bar{c}$ is small can we expect a direct comparison between a number of samples taken from such a trace and the mean value to be close. Since we know the underlying probability distribution (at least in this assumed case), we can make accurate statistical representations of desired properties. We can generate confidence limits directly from repeated sampling of the pdf.

Two statistics we have concentrated on are the cumulative distributions of the concentration itself and the residuals, the difference between the realization and the ensemble mean value. These statistical comparisons may be exemplified by choosing 100 independent samples from an extended realization of the $\sigma_c/\bar{c} = 1$ case shown in Fig. 2 as a particular set of observed sampler values. This particular observation is compared with the expected cumulative distribution function and the 95% confidence bounds in Fig. 3. The cumulative distribution of the mean values is a step function at unity in this assumed case, since the expected mean for every sample is unity. In spite of the large departures from the model mean value, the "observed" cdf's of this size sample are quite predictable, in that they are close to their expected values, well within the 95% confidence bounds.

When the above statistics are applied to real dispersion data, both the mean value and the variance of the concentration will generally vary in time and space. Thus the expected values cannot be obtained from a simple clipped normal distribution, but still may be

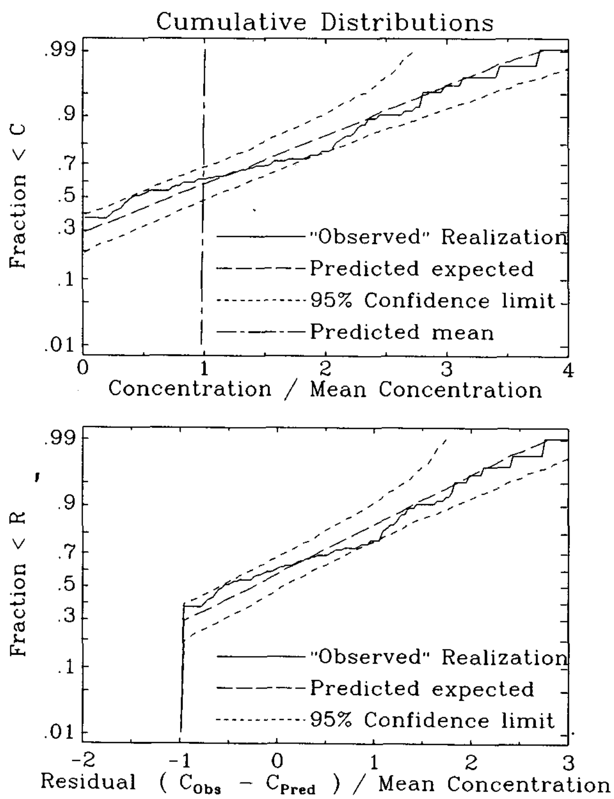

FIG. 3. Statistical comparisons of 100 samples from a single realization of a clipped normal distribution with $\sigma_c/\bar{c} = 1$ with the ideal expected and 95% confidence bounds. (a) Cumulative distribution of sampler concentrations; (b) cumulative distribution of residuals (value of "observed" realization minus predicted mean value).

determined by repeated numerical sampling from the model predicted pdf's.

Before confidence can be placed in any model's prediction, it must be evaluated in comparison with real world data. Figure 4 shows four consecutive hours of data from the Kincaid experiment (Bowne et al. 1983); the figure shows surface patterns of SF_6 tracer concentration for hour averages during the afternoon of 28 May 1981. This is a light wind, convective boundary layer, and the surface pattern shows a relatively coherent plume toward the northeast. The details of the pattern are not smooth, however, with large variations from one hour to the next at any particular site, although the boundary layer structure and the overall surface pattern do not change much over the period. Any model for the average concentration field would be expected to produce a smooth surface pattern with a slow evolution over the four hours. Our own predictions from SCIMP (Second-order Closure Integrated Model Plume), which will be described in the next section, are given in Fig. 5.

A standard evaluation procedure might compare observations and mean predictions directly, either paired in space or time, or on the basis of some maximum value over a time period, or spatial area. This method ignores the fact that the prediction is an ensemble average, while the observation is a random variable, and direct comparisons paired in space and time almost always suggest extremely poor model performance. A plot of observed concentrations against predicted mean values paired in space and time for 9 hours on the day illustrated by the distributions in Figs. 4 and 5 is shown in Fig. 6.

If all the departures from the predicted–equal–observed line are interpreted as errors, then this would be deemed a very poor simulation. A much more consistent statistical comparison, however, is to determine if the set of observed samples may be considered as a possible realization from our predicted pdf. That is, we may use repeated sampling of the predicted pdf to determine expected cumulative distributions and associated confidence bounds as exemplified in Fig. 3. Figure 7 shows that the observed samples for these nine hours are indeed quite consistent with what should be expected from a single realization of the set of predicted pdf's. The expected cumulative distribution is very close to the observed distribution, although very different from the distribution of the ensemble mean values. The average plume tends to be spread out over a larger surface area than that occupied by an actual realization. Thus the mean distribution contains a large number of small concentration values, while the expected distribution contains a small number of large values and a lot of zeroes. This tendency is quite evident in Fig. 7a. Figure 7b shows that a significant number of residuals are expected to be greater than 50 ppt, and that the observations are consistent with this prediction.

A higher resolution model should in principle be

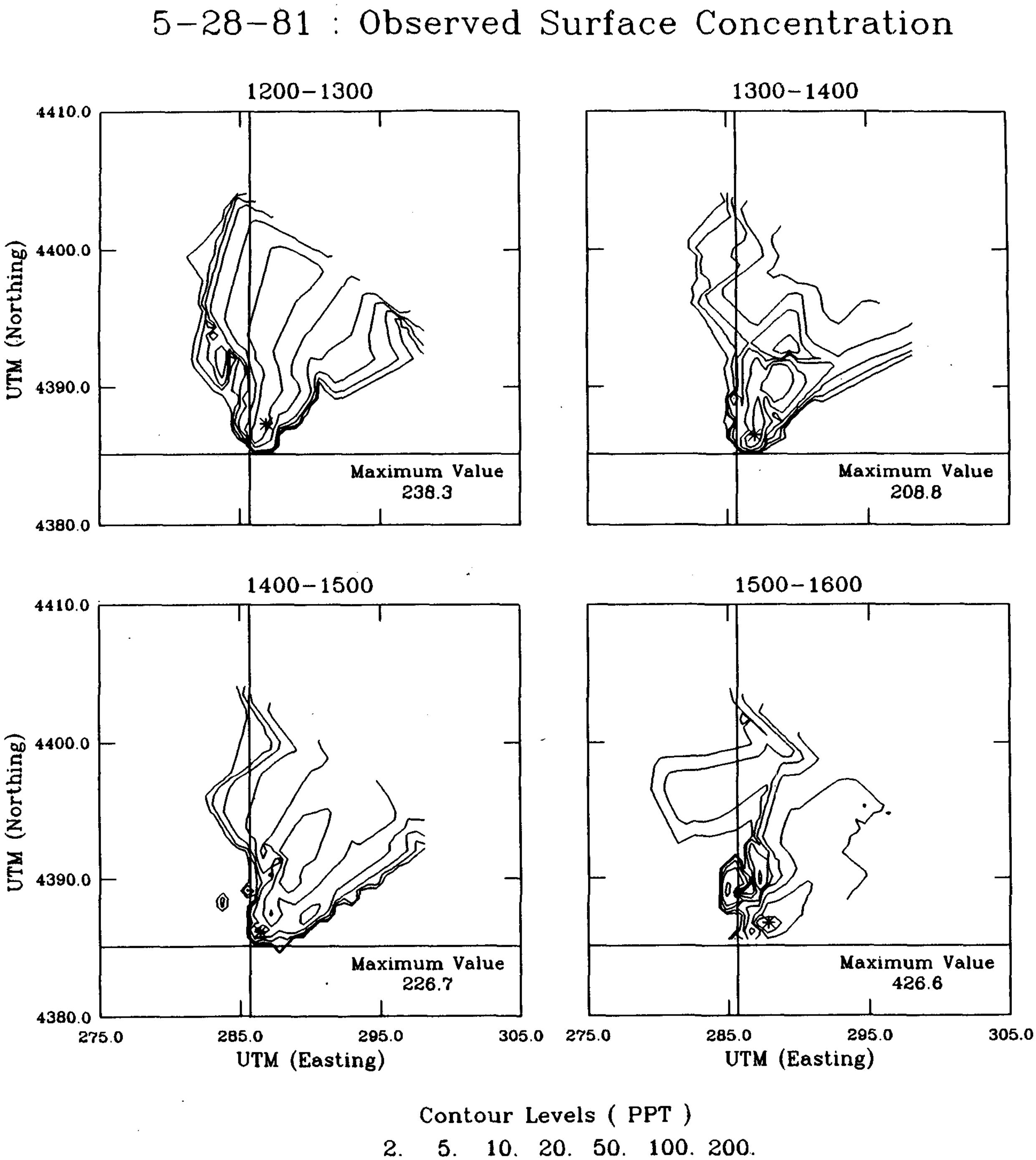

FIG. 4. Hourly averaged surface patterns of SF_6 tracer concentration for 4 hours during the afternoon of 28 May 1981 downwind of the Kincaid power plant stack as taken from the EPRI PMV&D dataset.

capable of reducing the residuals. Attempts at this (Lewellen et al. 1988), however, have so far proved unsuccessful. In spite of eliminating the steady state and simple Gaussian distribution assumptions by using a large number of puffs to represent the plume, SCIPUFF, the moderate resolution version of our dispersion predictability model, showed no consistent performance advantage over the simplest level model. We believe this reflects the importance of the meteorological uncertainty. Improvements in plume dynamics are unable to provide better statistics until the uncertainty in the wind field is more precisely deter-

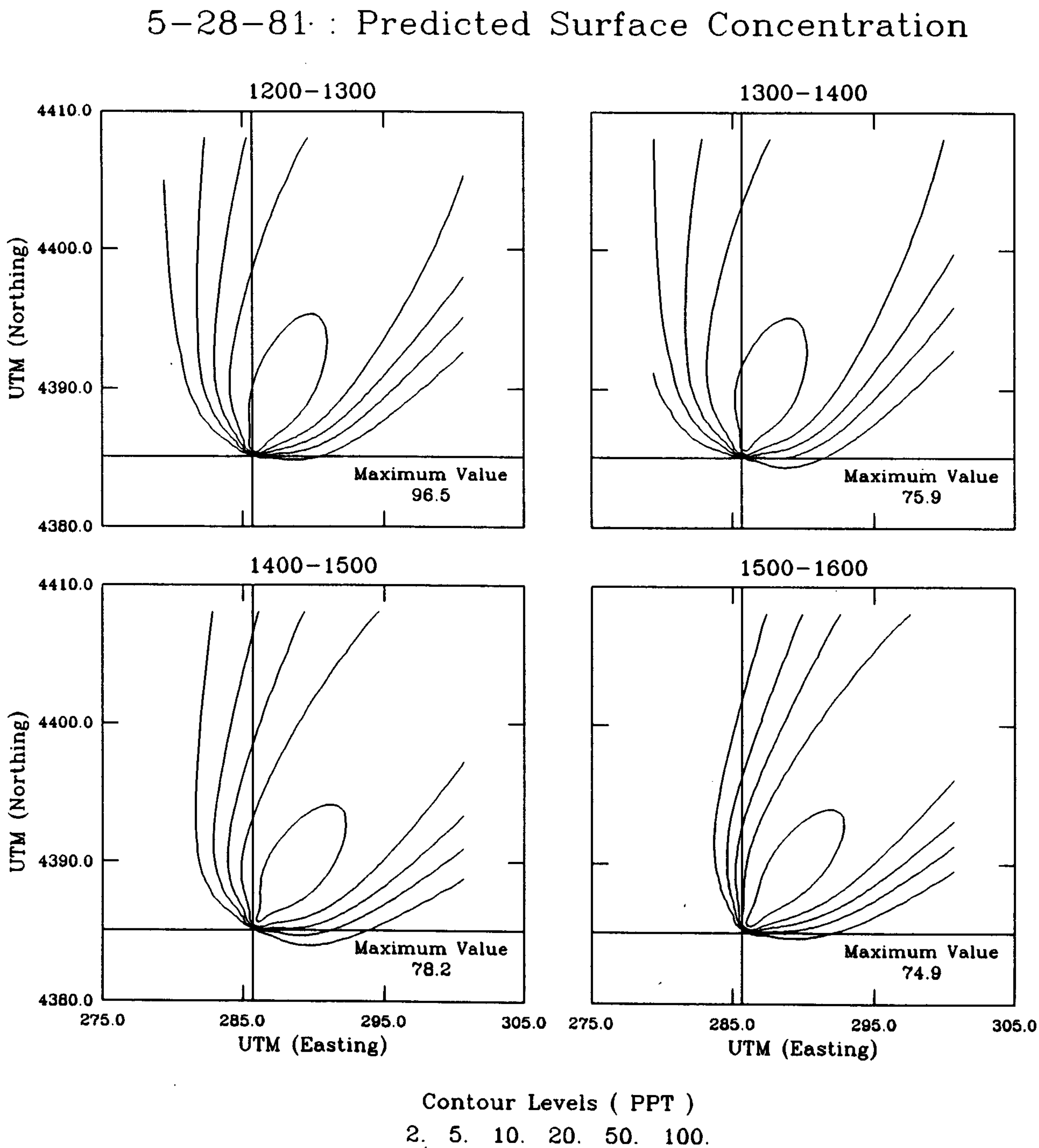

FIG. 5. The average concentration patterns as predicted by SCIMP for the ensemble of conditions compatible with the available meteorological data at the same times as given in Fig. 4.

mined. For purposes of illustrating the model framework, we will deal only with the simplest version of the model here.

3. An example dispersion predictability model

SCIMP (Second-order Closure Integrated Model Plume) provides a simulation of the probability distribution of concentration fluctuations at the level of a steady state Gaussian plume model. It is derived by assuming a steady plume and integrating the turbulent transport equations over the plane transverse to the plume to obtain a set of ordinary differential equations describing the downstream evolution of the plume. A description of a passive version of this model, along with some example results, was presented by Sykes et

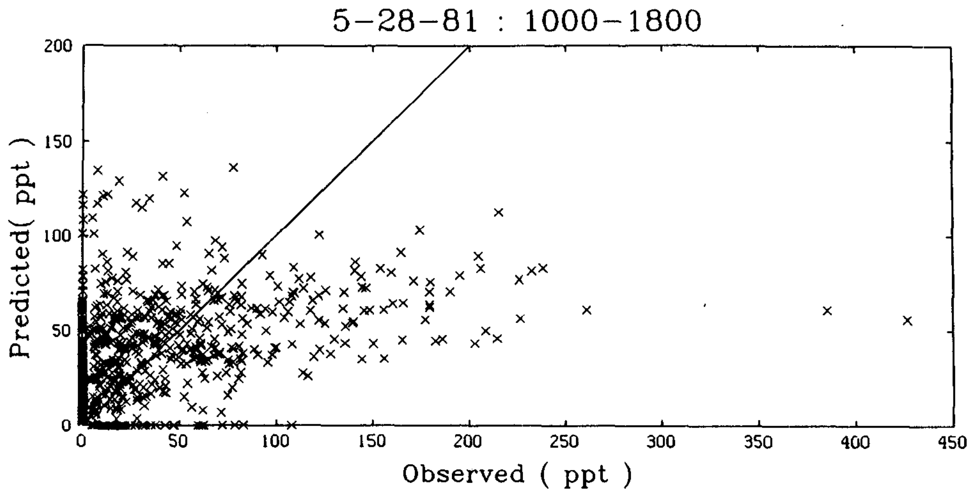

FIG. 6. Plot of observed concentrations against predicted mean values paired in space and time for 9 hours on 28 May 1981, including the distributions in Figs. 4 and 5.

al. (1986). In addition to predicting the mean concentration, the model predicts the variance of the concentration fluctuations. This added statistical information allows us to provide a simple representation of the probability distribution of concentration fluctuations. As noted in section 2, knowledge of the mean and the variance of an ensemble of concentration fluctuations can be used to provide a very reasonable description of the full probability distribution as a clipped normal distribution (Fig. 1).

It does not seem appropriate to repeat the set of ordinary differential equations associated with SCIMP. The set given by Sykes et al. (1986) has been extended to include the dynamic effects of buoyancy and momentum in the power plant emissions (Lewellen et al. 1988). Each hour is treated as a separate steady state simulation. This assumption together with the assumption of a Gaussian profile shape in space provide fairly severe limitations on this model. On the other hand, the simplicity of SCIMP permits repeated runs to be used in assessing the role of meteorological uncertainty.

The meteorological input data required for SCIMP are an estimate of the vertical distribution of the wind, and the turbulent correlations of the wind and temperature fluctuations. The model requires both the average of the ensemble of possible conditions represented by the available meteorological data, and an estimate of the ensemble variance in all three components of wind. Our approach has been to break the variation within this turbulent ensemble into two components: low frequency fluctuations in the wind that may roughly be associated with uncertainty in the mean meteorology, and the higher frequency turbulence associated with that mean meteorology. This somewhat arbitrary division permits us to use different approaches to calculating the concentration variance that arises from these different components of the velocity fluctuations. The contributions from the high frequency turbulent velocity fluctuations are computed directly from the evolution equation for the mean square scalar, which from Sykes et al. (1986) may be written as $(d/ds)\langle u_s \overline{c^2} \rangle = -\langle \overline{c'^2} \rangle/\tau_c$, where s is the coordinate along the plume centerline, the angled brackets represent the areal integral transverse to the plume centerline, the overbar the ensemble average, the prime the ensemble fluctuation, and τ_c a dissipation time scale.

The contributions from the low frequency fluctuations are computed from repeated SCIMP runs from a number of background realizations chosen randomly from the predetermined ensemble of possible meteorological conditions. In simulating the EPRI/PMV&D dataset we found that 10 surface SF_6 sample realizations from each of 50 meteorological realizations (i.e., 500 realizations of each simulated hour) adequately provided a consistent estimate of the cdf of expected concentrations and residuals along with the confidence bounds on these measures.

Even without completely describing the SCIMP equations, the above outline clearly shows that model results will depend critically on the input meteorological uncertainties. We have defined our ensemble as the collection of all wind and temperature fields that are consistent with the meteorological observations used to drive the model. Therefore the task is to deduce what wind statistics are consistent with the observational data.

We first estimate the mean wind from time averages of the tower and sounding (balloon-borne sounding with double theodolite tracking for winds) data, assuming that this filters out the small scales and gives a better representation of the ensemble wind for use throughout a 30 km radial domain. The high frequency

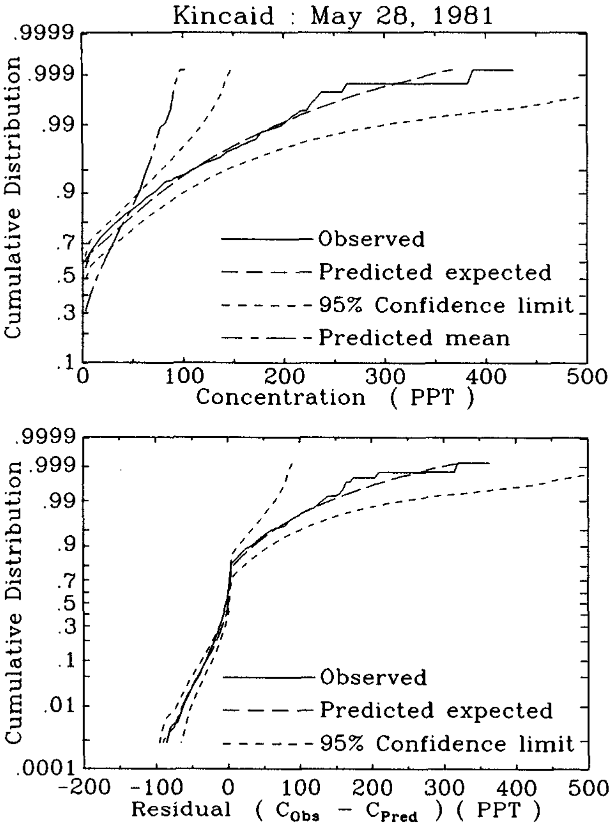

FIG. 7. Statistical comparisons of the SCIMP simulations with the observed surface SF_6 data for 9 hours on 28 May 1981. (Upper) Cumulative distribution of sampler concentrations. (Lower) Cumulative distribution of residuals (observed sampler value minus predicted mean value).

data contain information at the sampler location, but this cannot be expected to correlate over a long distance; we therefore define our input data to be the appropriate time-average and consider the ensemble accordingly. The second-order correlations are obtained consistently from the time-average, and a boundary-layer closure model is used to complete the picture.

A number of assumptions are implicit in our estimation of the ensemble-average fields. We assume horizontal homogeneity, and that our time-averaging procedures yield the correct ensemble average values. Apart from actual data errors, the main problem is one of representativity, i.e., to what extent does a time-average wind from the meteorological tower represent the appropriate large-scale ensemble average. A further problem arises from the use of the soundings, which are our only source of information above 100 m. For these data, we do not even have the luxury of a good time average, because these data are very short samples obtained only once per hour. The averages derived from these data are therefore much more uncertain.

The meteorological uncertainty estimates must account for the range of possible variations in the assimilated profiles, so they should cover time-averaging variations due to limited sampling, spatial variations not represented in the field measurements, and also errors due to instrument inaccuracies. The last category was studied by Smith et al. (1983), and their conclusions generally indicate that these errors are very small. The other two categories are much more difficult to estimate, and we can only make rough estimates on the basis of our examination of the actual data and our experience and understanding of the atmospheric boundary layer. We therefore present here a list of the meteorological variations that we used to represent the uncertainty in this set of predictions.

a. *Mean wind*

We used a standard deviation given by the larger of either 1 m s^{-1} or one-half of the observed standard deviation in the tower data for each velocity component. This is intended to account for nonrepresentativeness of the soundings, which we expect to contain a random component on the order of the observed variance due to their near-instantaneous nature. The lower limit of 1 m s^{-1} accounts for spatial variations which we estimate to be of this order. This estimate is very crude, and indicates an obvious data need. In the absence of any more accurate means of estimating the effect, however, a value of 1 m s^{-1} seems reasonable. It should be noted that this variation in each component of the wind yields a wind direction uncertainty that is a function of wind speed. The resulting direction uncertainty varies from about $\pm 5°$ at 10 m s^{-1} to $\pm 180°$ under calm conditions.

b. *Boundary layer depth*

Our estimate of the boundary layer depth comes from the temperature profiles, which are again based on near-instantaneous sounding measurements. A significant uncertainty exists in these estimates of inversion height because the observations are nonideal, so opinions may vary as to where the temperature gradient changes from near-zero to the positive lapse rate above the boundary layer. Furthermore, the instantaneous inversion moves in response to the turbulent velocity perturbations below, so a perfectly accurate measurement would still contain uncertainty when used as an estimate of the ensemble average. Finally, there is the possibility of spatial variation of the boundary layer depth, so that the sounding location may not perfectly represent the region where the plume reaches the inversion.

Based on our examinations of many sounding profiles, and our estimate of the range of possible inversion heights, together with typical profile variations from one sounding to the next, we have used a Gaussian distribution of boundary layer depths with a standard deviation of 20% of the mean depth. The standard deviation is not allowed to be less than 100 m, because the sounding data generally cannot support a higher resolution, and therefore shallow mixed layers are relatively more uncertain than deep ones. This range of uncertainty is also reasonably consistent with the spatial variation in instantaneous inversion height in laboratory convection layers (Deardorff et al. 1969).

c. *Turbulence intensity*

The data inputs that most influence our estimates of the vertical velocity variance are the direct tower measurement of $\overline{w'^2}$ at 100 m, the heat flux, and the boundary layer depth. We have already discussed the uncertainty in the latter, which leaves the heat flux as the major factor. The heat flux estimate is indirect, whether we use a measurement of $\overline{w'^2}$ or the tower temperature profiles; we also have no information on moisture effects. We were unable to derive quantitative factors to account for this uncertainty, but we used a Gaussian variation on $\overline{w'^2}$ with a standard deviation of 30% of the local profile value. The main problem with this variation is that a situation which produces low turbulence levels cannot generate an ensemble member with any significant turbulence. In most cases this is reasonable, because many low turbulence cases are nocturnal, and the intensity is almost certainly low. The problem cases are near transition times when the boundary layer could contain significant turbulence even though the measurement might suggest otherwise.

Having thus assumed the range of meteorological uncertainty within the ensemble for each hour, we can proceed to determine the effects on the surface SF_6 predictions. Some aspects can be accommodated directly within our second-order closure framework. For example, the uncertainty in the mean wind can be considered as horizontal turbulence with a very large timescale. However, in the case of a dynamic plume, many effects are very nonlinear. The interaction of the buoyant plume with the inversion is a complex process that cannot be described simply enough to calculate the effect of variations in the inversion height. We therefore exploit the simplicity and efficiency of the SCIMP model to calculate the dispersion using an ensemble of meteorological backgrounds explicitly.

The calculation of a Gaussian plume using the SCIMP model takes only a few seconds, so it is feasible to perform many calculations for each case. We therefore choose a number of background realizations, and calculate the plume with a randomly generated background. The procedure is to choose mean wind perturbation, boundary layer depth factor, and turbulence factor from our previously described Gaussian distributions. No attempt is made to impose a dependence between these three factors, but the resulting profiles of each variable are scaled in a physically consistent manner. We generate the appropriate backgrounds by adding the mean wind perturbation to each component at all heights, rescaling the heights on the turbulence profiles, and shifting the temperature profile to produce the change in inversion height. The vertical velocity variance is then scaled, and the turbulent length scale is also scaled to account for the change in mixing depth.

The plume is then calculated in this background profile, and a number of realizations of the surface SF_6 samples are generated from the predicted concentration pdf's. The reason for generating surface sampler realizations from each background realization, rather then accumulating mean and variance statistics over the whole range of background profiles, is that we account for spatial correlations much more consistently. If a particular background realization has no impact on the surface, for example, then we would have an entire set of zero sampler concentrations, which becomes one realization of the distribution. It is difficult to specify the proper spatial correlations for the final ensemble of results.

Figure 8 shows the statistical comparison when the above meteorological uncertainties are incorporated into the SCIMP model framework for the Kincaid Developmental dataset. Some 20 000 hourly samples are included in these observations. The observed cumu-

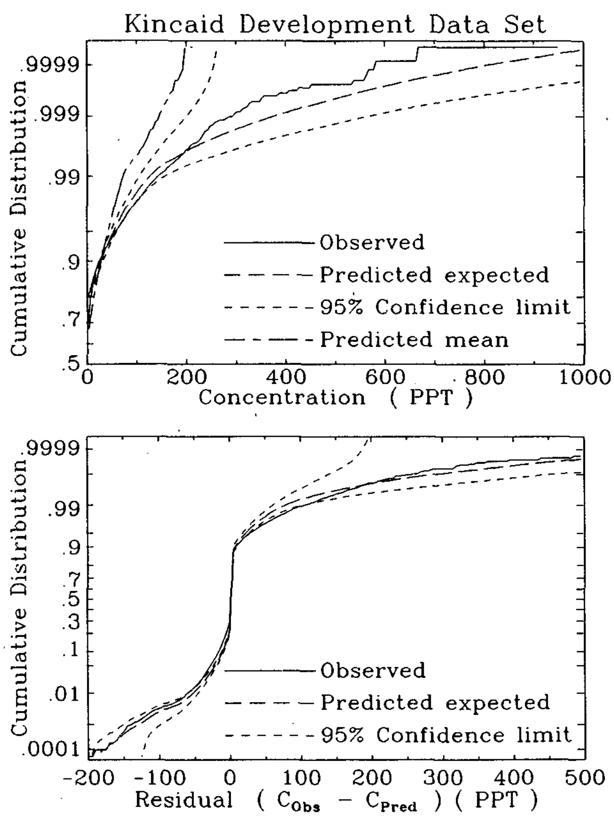

FIG. 8. Statistical comparisons of the SCIMP simulations with the 20 000 hourly samples of the Kincaid Development Data Set. (a) Cumulative distribution of sample values; (b) cumulative distribution of residuals.

lative distributions generally fall within the 95% confidence bounds. If one attempts to condense these comparisons down to a single figure of merit, a possible candidate is the area between the observed distribution and the confidence interval expressed as a percentage of the average observed concentration, namely, 2% for the error in concentration values and 12% for the residual cdf. These very good results may be partially explained by the fact that the model was developed using this dataset. When applied to three other datasets of equal size, the corresponding error measures were increased to 8%, 16% and 18% for the concentration values and 19%, 23% and 24% for the residuals. Considering the roughness with which the meteorological uncertainty was specified, we consider these results to be very good.

4. Desirable meteorological data

The available meteorological data necessarily defines the wind environment for an air quality model. Any wind field that is completely compatible with the existing data may make a possible contribution to dispersion at that time. The more precisely the possible ensemble of wind fields is defined by the data, the more precision can be included in the dispersion simulation. The wind field can be more precisely defined either by narrower confinement of the range of compatible flows, or by more precise definition of the variance within the range, even when the range is relatively wide.

We want to concentrate on the latter, less obvious approach, since we believe it provides greater potential for practical models. There are two difficulties with depending upon a high resolution 4-D dataset to narrowly confine the ensemble. First, it tends to be an expensive way to define the meteorological input for any given episode; and second, the higher the resolution of the four-dimensional data, the more restricted is its applicability. We may think of the problem as five-dimensional: three physical spatial dimensions, time, and ensemble space. In order to narrowly confine the flow in ensemble space, it must be very well defined in space and time. Conversely, if we are willing to accept a wider variance in ensemble space, then the requirements on the space and time resolution can be relaxed. Our assumption is that the data needed to define the wind variance associated with given types of wind measurements and 4-D domain should be more universal than is the detailed 4-D flow field.

We have not completed a consistent sensitivity study to determine the separate effects of each of the factors considered in section 3, but based on the results of our limited variations, the horizontal wind variance appears to have the strongest influence. Information to better define this critical variable should be obtainable from the wind profilers discussed in detail by a number of authors in this issue. Such profilers are capable of providing high resolution in two dimensions, time and height. Unless an array of profilers is used, this 2-D high resolution cannot be directly used to provide the high resolution 4-D data needed to narrowly define an ensemble. It should, however, provide information towards defining the wind variance associated with different time filters under different conditions. It is generally recognized that some time filtering of the data is necessary to make the data representative of a given model domain. As the size of the domain the data is asked to represent is increased, the more filtering is required. We suggest that the basic statistics of the wind variance smoothed out in this filtering process be recorded, to help define the uncertainty in the wind for different conditions and domains. This should provide valid statistics on the wind variations under relatively homogeneous conditions when there are little variations in the cross wind direction. We also know, however, that terrain and thermally forced boundary conditions make horizontal stationarity in the atmospheric boundary layer the exception, rather than the rule. Ideally, we would like additional spatial information to determine how this wind variance may be represented as the spatial domain of the data is increased. Can wind data from an array of relatively closed packed profilers be averaged to provide both a mean wind and a mean variance as a function of spatial domain radius? As the domain is increased there will be less correlation between the separate profile fluctuations, and more of the wind energy spectra must be included in the variance rather than as part of the resolved mean wind.

If adequate profiler data can be taken over a closely spaced array to assimilate a detailed four-dimensional representation of the wind field over an interesting domain, then a series of model tracer puffs could be released and tracked in the field to completely define the dispersion statistics for this case. Such field experiments could provide a very synergistic interaction with numerical experiments currently being pursued utilizing large eddy simulations (LES) by Sykes et al. (1988). The numerical experiments can help provide the needed 4-D data assimilation, and the field data can in turn provide a needed validation, and an extension of the range of the LES numerical results. Between the field and numerical experiments, it should be possible to establish the necessary information on temporal and spatial correlations needed to provide more accurate descriptions of the variance associated with data representing a given domain. This will permit improvements in predictability models of the type presented in section 3.

Tennekes (1988) has presented an essay "in praise of the limits of predictability" and on the challenges to meteorologists of "learning to live with the mysteries of nature." Our plea here is in much the same spirit, that much more research on the predictability of the wind environment is needed. We believe that attempts to collect the data needed to define the range of uncertainty are likely to be more productive than attempts to eliminate the uncertainties in a given dataset.

Acknowledgments. This paper is based on work supported by EPRI under Project No. RP1616-28.

REFERENCES

Bowne, N. E., R. J. Londergan, D. R. Murray and H. S. Borenstein, 1983: Overview results and conclusions for the EPRI PMV&D Project: Plains site, EPRI EA3074, RP1616-1.

Deardorff, J. W., G. E. Willis and D. K. Lilly, 1969: Laboratory investigation of non-steady penetrative convection, *J. Fluid Mech.,* **35,** 7–31.

Fox, D. G., 1984: Uncertainty in air quality modeling, *Bull. Amer. Meteor. Soc.,* **65,** 27–36.

Irwin, J. S., S. T. Rao, W. B. Petersen and D. B. Turner, 1987: Relating error bounds for maximum concentration estimates to diffusion meteorology uncertainty, *Atmos. Environ.,* **21,** 1927–1937.

Lamb, R. G., 1984: Air pollution models as descriptors of cause–effect relationships, *Atmos. Environ.,* **18,** 591–606.

Lewellen, W. S., and R. I. Sykes, 1983: On the use of concentration variance predictions as a measure of natural uncertainty in observed concentration samples, *Sixth Symposium on Turbulence and Diffusion,* Boston, MA, Amer. Meteor. Soc.,

——, and ——, 1986: Analysis of concentration fluctuations from lidar observations of atmospheric plumes, *J. Cimate Appl. Meteor.,* **25,** 1145–1154.

——, ——, S. F. Parker, D. S. Henn, N. L. Seaman, D. R. Stauffer, T. T. Warner and H. R. Chang, 1988: A hierarchy of dynamic plume models incorporating uncertainty, Volume 1, EPRI EA-6095, Project 1616-28.

Smith, F., C. E. Decker, R. B. Strong, J. H. White, F. K. Arey and W. D. Bach, 1983: Estimates of uncertainty for the Plume Model Validation and Development Project field measurements—Plains site, EPRI EA-3081, Project 1616-10.

Sykes, R. I., W. S. Lewellen and S. F. Parker, 1984: A turbulent-transport model for concentration fluctuations and fluxes, *J. Fluid Mech.,* **139,** 193–218.

——, —— and ——, 1986: A Gaussian plume model of atmospheric dispersion based on second-order closure, *J. Climate Appl. Meteor.,* **25,** 322–331.

——, —— and D. S. Henn, 1988: Large eddy simulation of concentration fluctuations, *Eighth Symposium on Turbulence and Diffusion,* San Diego, CA, Amer. Meteor. Soc.

Tennekes, H., 1988: The outlook: scattered showers, *Bull. Amer. Meteor. Soc.,* **69,** 368–372.

Venkatram, A., 1982: A framework for evaluating air quality models, *Bound.-Layer Meteor.,* **24,** 371–385.