UNCERTAINTY IN PREDICTIONS FROM AIR QUALITY MODELS

AKULA VENKATRAM

Environmental Research and Technology, Inc., 696 Virginia Road, Concord, MA 01742, U.S.A.

(Received in final form 1 September, 1983)

Abstract. This paper examines the role of the definition of the ensemble in determining the uncertainty in air quality model predictions. We difine the ensemble in terms of the variables included in the model input set. Because observations are also governed by variables not included in this set, we can have an infinitely large set of observations associated with a single model prediction. The variation of these observations about the model prediction corresponds to all possible values of the unknown variables, and it represents the uncertainty in the model prediction. This concept is illustrated by estimating the uncertainty related to model predictions for two dispersion problems.

1. Introduction

In principle, if boundary and initial conditions are completely specified, the equations of motion allow us to predict all the details of turbulent flow. However, this is impossible in practice because of the nature of turbulent flow, and we are forced to treat turbulent motion in terms of a resolved 'mean' flow and an unresolved residual 'turbulent' velocity. Each 'mean' flow field is associated with an infinite set of unknown turbulent velocity fields. Therefore the best an air quality model can do is to provide an estimate of the average of the concentrations measured during different 'realizations' of the flow. An observed concentration is expected to deviate from the corresponding model prediction.

This paper provides a discussion of the concepts underlying the estimation of the expected deviation between model prediction and observations. This is not a review of techniques of calculating this component of model uncertainty, and the reader is referred to Csanady (1973) and Chatwin and Sullivan (1979) for an overview. The main objective of this paper is to emphasize the relationship between model prediction and observation. In doing so, we demonstrate the need to be precise in defining the ensemble of observations being considered. The role of the definition of the ensemble in the estimation of model uncertainty is illustrated through two examples. We then show that it is necessary to formulate a model for the expected deviation (concentration fluctuation) before the model for the ensemble mean can be evaluated.

2. The Concept of an Ensemble

The concept of an ensemble of realizations is central to the understanding of the statistics of concentrations. It is most simply introduced by expressing the observed concentration as follows:

$$C_0 = C_0(\alpha, \beta)\,. \tag{1}$$

Boundary-Layer Meteorology **27** (1983) 185–196. 0006–8314/83/0272–0185$01.80.

In Equation (1), α refers to the set of variables whose values are known, and β denotes those variables which are unknown to us but which also affect C_0. For example, α could refer to the hourly averaged set of specified values of wind speed, mixed-layer height, surface heat flux and distance of the pollutant source from the receptor; β could include the unresolved turbulent velocities. A fixed value of α defines an ensemble of experiments. Because β can take on any value for a given α, the observation $C_0(\alpha, \beta)$ will vary from experiment to experiment. Any observation made during an experiment belonging to this ensemble can be written as

$$C_0(\alpha, \beta) = \overline{C_0(\alpha, \beta)}^{\beta} + c(\alpha, \beta)\,. \tag{2}$$

In Equation (2), the overbar refers to an average over the ensemble defined by a specified set of values α. Since the average is over all possible β, it is only a function of α and we can write

$$C_0(x, y) = C_p(\alpha) + c(\alpha, \beta)\,. \tag{3}$$

As α represents the input information available to an air quality model, the preceding discussion suggests that the model prediction should ideally be the average $C_p(\alpha)$ over all observations for fixed α.

The relationship between a model prediction $C_p(\alpha)$ and observations $C_0(\alpha, \beta)$ is illustrated in Figure 1. Each prediction $C_p(\alpha)$ its associated with is own 'cloud' of observations. This spread of the possible observed values about C_p is σ_c. Since σ_c is a function of α, the statistics of c can be calculated only by sampling observations keeping

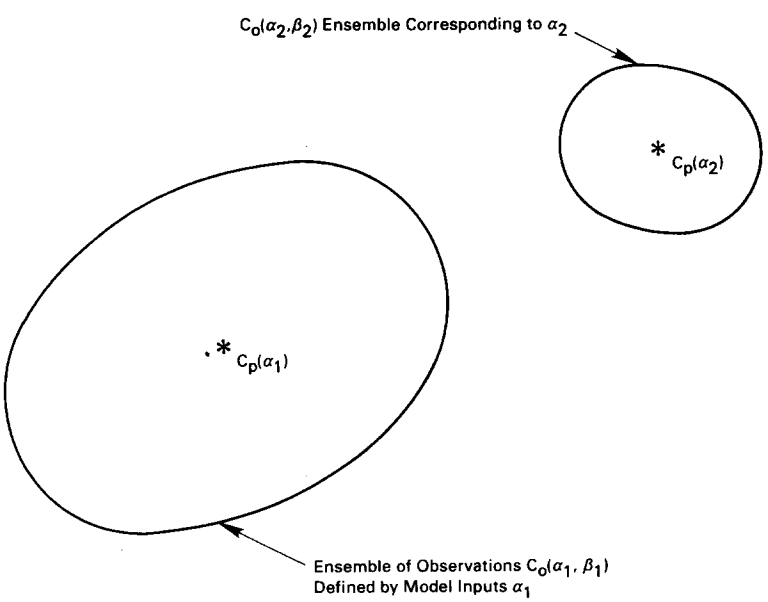

Fig. 1. Relationship between a model prediction and observations.

α fixed. Because this is impossible in practice, we have to transform c to make it independent of α. Proposed methods of achieving this are described in a later section.

The next section discusses the term 'inherent uncertainty' which is often incorrectly applied to σ_c.

3. The Meaning of 'Inherent' Uncertainty

As mentioned earlier, the concentration fluctuation $c(\alpha, \beta)$ is directly related to our lack of knowledge of the set of variables β. In principle, the uncertainty $\sigma_c(\alpha)$ can be made arbitrarily small by including more information in the set α, and therefore, there is nothing 'inherent' or intrinsic in this uncertainty. Our definition of the ensemble rather than 'nature' determines the magnitude of $\sigma_c(\alpha)$. However, turbulence, which governs dispersion, places a severe limitation on our ability to decrease $\sigma_c(\alpha)$. This limitation is caused by the rapid amplification of small disturbances in turbulent flow. Therefore, small model input errors have large consequences on model predictability. To see this in the context of air quality models, let us express the observed concentration as follows

$$C_0(\alpha, \beta) = C_p(\alpha_I) + I(\Delta\alpha) + c(\alpha, \beta)\,. \tag{4}$$

In Equation (4), $I(\Delta\alpha)$ is the random error associated with the input error defined by

$$\Delta\alpha = \alpha - \alpha_I \tag{5}$$

where α_I is the incorrect input.

Figure 2 shows the qualitative behavior of the model errors σ_c and σ_I as a function of the number of variables $N(\alpha)$ in the set α. If we think of α as a vector, $N(\alpha)$ is the number of dimensions of the space used to represent α.

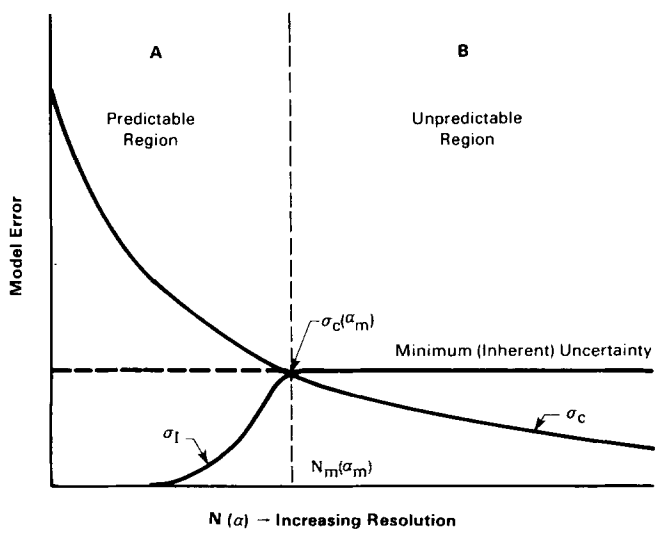

Fig. 2. The variation of model error with model complexity. $\Delta\alpha$ is arbitrarily small.

Note that $\sigma_c(\alpha)$ will always decrease with increase in $N(\alpha)$. After we have accounted for all the physical variables in α, an increase in $N(\alpha)$ is equivalent to improving the time and space resolution of the model. Beyond a certain point, this improvement is curtailed by the extreme sensitivity of predictions of small-scale turbulent motion to input errors. In the words of Tennekes 1978, "since this problem grows worse with improving resolution, turbulence theory (with some notable exceptions) tends to give up altogether: deterministic forecasting of the evolution of individual eddies is abandoned, and wholescale averaging techniques are employed". What this means is that beyond a certain α, denoted by $\alpha = \alpha_m$ in Figure 2, *arbitrarily small* errors in additional model inputs lead to a spread in model predictions that is indistinguishable from σ_c at $\alpha = \alpha_m$; the extra information is essentially useless. We see that the model error cannot go below the value $\sigma_c(\alpha_m)$. Therefore, $\sigma_c(\alpha_m)$ is in a sense the 'inherent' uncertainty dictated by the nature of turbulence.

In the figure, A represents the region in which σ_I can be made arbitrarily small by making the input α more accurate. Region B is where we have no control over the magnitude of σ_I. We do not want to imply that the transition between A and B is as sharp as shown in the figure.

There is little theory to guide us on the selection of α_m. The only way of determining the 'uncertainty barrier' is to make the set α as large as possible. At this point, there is no reason to believe that available dispersion models have reached the limits of predictability. In this connection, it is claimed that the long time scales associated with the eddies of the convective boundary layer lead to a large σ_c and consequently low model predictability. This is not true because σ_c can be reduced by including more information in the currently used model input sets. Because the time scale of the dispersing eddies is of the order of ten minutes, there is every reason to believe that model predictions can be improved by using the measured probability density functions of turbulent velocity fluctuations. However, such information is not likely to be available routinely, and we have to live with the relatively large σ_c associated with limited inputs. The point to be made is that there is nothing intrinsic about the large σ_c for dispersion in the convective boundary layer.

In the following sections we discuss some examples that illustrate the estimation of σ_c for dispersion problems.

4. The Physics of $c(\alpha, \beta)$

Here we discuss the processes which determine the statistics of the 'concentration fluctuation' $c(\alpha, \beta)$. To do this, it is useful to consider the concentration time series caused by material released with an initial concentration C^p. If we neglect the effects of molecular diffusion, the instantaneous concentration measured at a receptor will be either C^p or zero as shown in Figure 3. It can be readily shown that (Venkatram, 1979)

$$\langle (C(t) - C_m)^2 \rangle / C_m^2 = C^p / C_m - 1 \tag{6}$$

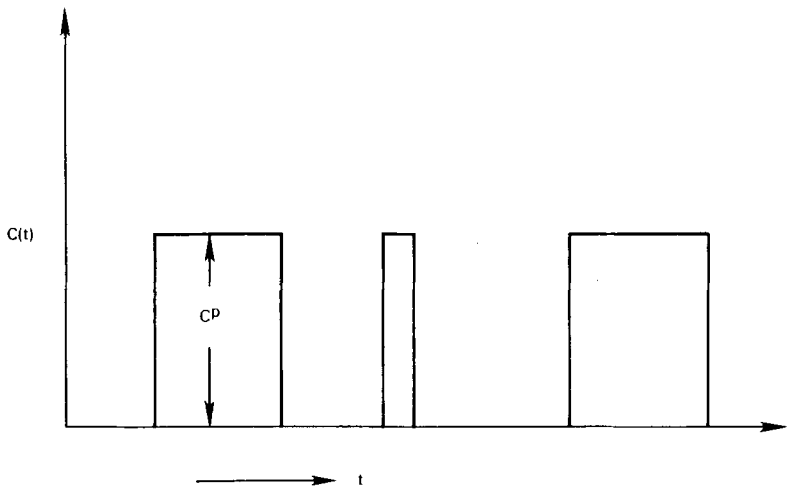

Fig. 3. Concentration time series associated with an initial release of C^p.

where

$$C_m = \langle \overline{C}^T(\alpha, \beta) \rangle \tag{7b}$$

and

$$\overline{C}^T(\alpha, \beta) = \frac{1}{T} \int_0^T C(t, \alpha, \beta)\, \mathrm{d}t\,. \tag{7c}$$

In writing Equation (7), we have defined our ensemble with respect to the set $\tilde{\alpha} = \{\alpha, T\}$. It is seen that the deviation of the instantaneous concentration $C(t)$ from the mean C_m is dependent on the way we define our ensemble. To see this more clearly, let us consider the example used by Chatwin and Sullivan (1979). Material with mass Q is initially mixed over a volume L_0^3 and then released into a turbulent flow. To analyze this situation, they consider an ensemble for which Q, L_0^3 and a set of unspecified turbulence statistics are fixed. The origin of the co-ordinate system is attached to the randomly moving center of mass of the pollutant cloud. The basic vectors of this system are fixed in space. We will see that this description of the co-ordinate system is an important component of the definition of the ensemble in question.

The initial concentration in the pollutant cloud is Q/L_0^3. Until the effects of molecular diffusion become important, the concentration of the fluid marked with the pollutant remains at this value. However, the cloud is distorted and stretched as its center of mass moves around. A measure of this distortion caused by the turbulent flow field is $L(t)$ given by

$$QL^2(t) \equiv \left\langle \int |\mathbf{r}|^2 C(\mathbf{r}, t)\, \mathrm{d}V(\mathbf{r}) \right\rangle \tag{8}$$

where **r** refers to the position vector from the origin of the co-ordinate system and the integral is taken over all space. It is easily seen that the probability that a receptor fixed to this co-ordinate system measures a concentration is approximately L_0^3/L^3. The mean concentration at the receptor is Q/L_0^3 times the probability that the receptor measures a concentration. The answer, which agrees with intuition, is Q/L^3. On the basis of previous arguments, the concentration fluctuations about this mean can be expressed as

$$\frac{\langle (C(\mathbf{r}, t) - C_m(\mathbf{r}, t))^2 \rangle}{C_m^2} = \frac{\langle c^2 \rangle}{C_m^2} = \frac{L^3}{L_0^3} - 1 . \tag{9}$$

Since $L(t)$ increases with time, the fluctuations about the mean also increase. Note that the cloud spread is a function of the co-ordinate system used in describing the ensemble. To see this, let us assume that we have sufficient measurements of the turbulent field to allow us to predict the motion of the 'major axis' of the distorted pollutant cloud. We now put a sensor at a specified **r** in a co-ordinate system fixed to the major axis. It is clear that the probability of detecting a concentration is higher in this co-ordinate system than in the previous one. This suggests that L^3 will be smaller in this system and that the concentration fluctuations will therefore be smaller. In principle, if we have enough information, we can make sure that the sensor is always inside the marked fluid. In this case, the concentration fluctuations are zero. This discussion emphasizes the importance of defining the ensemble before calculating the statistics of concentrations. As Chatwin (1982) points out, "the term 'mean concentration' has no meaning unless it has first been made clear what ensemble the mean is taken over".

Our arguments have neglected the effects of molecular diffusion. Chatwin and Sullivan (1979) show that this assumption is justified until the size of the cloud $L(t)$ is L_d given by

$$L_d = L_0^3 \lambda_c^{-2}/\gamma ; \qquad L_d/L_0 = L_0^2 \lambda_c^{-2}/\gamma \tag{10}$$

where λ_c is the conduction length given by

$$\lambda_c = (\nu\kappa^2/\varepsilon)^{1/4} \tag{11}$$

where κ is the molecular diffusivity, ε is the turbulent dissipation rate and ν is the kinematic viscosity of the fluid. Estimates of λ_c are of the order of 1 mm. The value of the 'constant' γ is uncertain. Chatwin and Sullivan (1979) use $\gamma \simeq 10^3$ m. If we take $L_0 = 1$ m, L_d works out to be $\sim 10^3$ m. This simple calculation suggests that we can neglect molecular diffusion in our analysis until the cloud is 'large' compared to L_0. The role of molecular diffusion can be better understood when one realizes that turbulence only distorts the cloud. An initially spherical cloud of pollutant is stretched into strands by the straining turbulent velocity field. The edges of these strands are 'blurred' by molecular diffusion. This blurring has a negligible effect on the initial volume of fluid containing the pollutant. It is only when the thickness of these strands is of the order of λ_c that molecular diffusion becomes effective. At this stage, the rate of diffusive transfer from these strands becomes comparable to the rate at which the strands are

squeezed by turbulence. We can now see how molecular diffusion 'destroys' concentration fluctuations. This rate of destruction becomes large when the cloud 'strand' thickness becomes small or equivalently when the local concentration gradients become large. Note from Equation (10) that the time at which molecular diffusion becomes important is very sensitive to the initial size of the cloud L_0.

Molecular diffusion is probably less important in determining the variance of time-averaged concentrations. In the next section, we describe a simple model for the variance of time-averaged concentrations measured in the convective boundary layer. The formulation of the model again emphasizes the need to define the ensemble to be considered.

5. Concentration Fluctuations in the Convective PBL

This section discusses a method to estimate the variance of concentrations associated with an elevated release into the convective boundary layer. The basic assumption is that we know the 'true' relationship between the variables controlling the *observed* concentration. The observed concentration cannot be calculated exactly only because we have no information on one or more of these variables. Note that the values of the known variables define the ensemble.

In this example, the observed concentration is assumed to be given by an expression suggested by Venkatram (1983). This formulation yields results that compare well with those from the water tank experiments of Willis and Deardorff (1978). Then, the observed cross-wind integrated concentration is

$$\overline{C}^y = \frac{2Q}{x} P(w|z); \qquad w = w_s, \qquad z = h. \tag{12}$$

In Equation (12), Q is the emission rate from a point source with stack height h, x is the downwind distance to the receptor and $P(w|z)$ is the probability density function of vertical velocity fluctuations at z. The velocity w_s is given by

$$w_s = -uh/x \tag{13}$$

and corresponds to the value required to bring plume material from the source to the receptor in a straight line. This physical picture is consistent with the assumption of infinite time scales in the convective boundary layer.

Next we will assume that we cannot predict $\overline{C}^y$ because the precise form of $P(w|z)$ is unknown. Although $P(w|z)$ is known to be positively skewed (Lamb, 1982), the experience of Weil and Brower (1982) suggests that adequate estimates of concentration can be obtained with a gaussian probability density function. We can then write

$$P(w|z) = \frac{1}{\sqrt{2\pi}\sigma_w} \exp\left[-\frac{(w - w_0)^2}{2\sigma_w^2} \right]. \tag{14}$$

In Equation (14), σ_w is taken to be a time-averaged value obtained from measurements.

The mean vertical velocity w_0 is unknown, and is responsible for the variation of concentrations belonging to the ensemble defined by

$$\alpha = [Q, h, x, \sigma_w, u, T] \tag{15a}$$

$$\beta = [w_0]\,. \tag{15b}$$

In Equation (15a), T is the averaging time corresponding to σ_w and u. In the analysis to follow, we will find it convenient to deal with $\ln \overline{C}^y$ given by

$$\ln \overline{C}^y = \ln\left(\sqrt{\frac{2Q}{\pi\sigma_w x}}\right) - \frac{(w_s - w_0)^2}{2\sigma_w^2}\,. \tag{16}$$

The expression for the geometric mean concentration $\overline{C}_g^y$ follows readily from Equation (14):

$$\overline{C}_g^y = \sqrt{\frac{2}{\pi}}\,\frac{Q}{\sigma_w x}\exp\left(\frac{\langle w_0^2\rangle}{2\sigma_w^2}\right)\exp\left(-\frac{w_s^2}{2\sigma_w^2}\right) \tag{17}$$

where the angle brackets refer to an ensemble average. To estimate $\langle w_0^2\rangle$, let us assume that the records of the vertical velocity $w(t)$, corresponding to the different realizations of the ensemble, are derived from a stationary time series with an integral time scale T_w. Then one can show that for $T \gg T_w$ (Lumley and Panofsky, 1964),

$$\langle w_0^2\rangle = 2\sigma_w^2 T_w/T\,. \tag{18}$$

Note that σ_w^2 is not the ensemble variance in the sense that it corresponds to infinite averaging time. In our discussion, σ_w^2 constrains the energy in each of the realizations of the ensemble,

$$\sigma_w^2 = \frac{1}{T}\int_0^T w^2(t)\,\mathrm{d}t\,. \tag{19}$$

With Equation (18), and using $w_s = -uh/x$ and $\sigma_z = \sigma_w x/u$, Equation (17) becomes

$$\overline{C}_g^y = \sqrt{\frac{2}{\pi}}\,\frac{Q}{u\sigma_z}\exp\left(-\frac{T_w}{T}\right)\exp\left(-\frac{h^2}{2\sigma_z^2}\right)\,. \tag{20}$$

Equation (20) is essentially the gaussian formulation in common use.

Because w_0 is a time average of $w(t)$, we can invoke the Central Limit Theorem and assume that w_0 is normally distributed. Then, it is easy to show that the standard deviation of concentration 'fluctuations' is given by (Venkatram, 1983):

$$\sigma(\ln C(x, y)) = \left[\frac{2T_w}{T}\left(\frac{T_w}{T} + \frac{h^2}{\sigma_z^2}\right) + \frac{2T_v}{T}\left(\frac{T_v}{T} + \frac{y^2}{\sigma_y^2}\right)\right]^{1/2}\,. \tag{21}$$

In Equation (21), T_v is the integral time scale for the horizontal velocity fluctuation $v(t)$ and $\sigma_y = \sigma_v x/u$.

We can specialize Equation (21) for the convective boundary layer by taking $\sigma_z = 0.6 w_*/u$ (Panofsky, 1978) and $T_w \simeq 1.5 z_i/u$ (Kaimal *et al.*, 1976). Here, we have assumed that these expressions are exact and that the release height $\bar{h} > 0.1$. Then, the σ for the centerline concentration is

$$\sigma(\ln C) \simeq 3\left(\frac{z_i}{uT}\right)^{1/2} \frac{\bar{h}}{X}; \qquad X = w_* x/z_i u \tag{22a}$$

$$s \equiv \exp[\sigma] = \exp[3(z_i/uT)^{1/2}\bar{h}/X]\,. \tag{22b}$$

Equation (12) is based on the assumption that the concentration is determined by the flux of *serially* released particles through an elemental area (see Venkatram, 1983 for details). This implies that the averaging time $T \gg T_w$. Therefore, the proposed formulation for s_c is not applicable for T less than or comparable to T_w. The variance of instantaneous concentrations has to be modeled using methods such as that suggested by Durbin (1980).

5.1. Typical results

To get an idea of the magnitude of s_c, we have plotted Equation (22b) for $z_i = 1000$ m, $u = 5$ m s^{-1}, and $T = 3600$ s (1 hr). Figure 4 shows the variation of s_c with the non-dimensional distance X; $\bar{h}$ is the normalized source height. Because our formulation does not account for the limitation of mixing by the inversion lid, results beyond $X = 1$ (especially for $\bar{h} = 0.5$) should be viewed with caution.

To interpret Figure 4, let us assume that the concentrations are lognormally distributed. (The actual distribution, consistent with the assumed normality of w_0, is not lognormal.) Then s_c^2 is approximately equal to the 95% confidence interval for the ratio C_0/C_p. For example, $s_c = 2$ indicates that we expect 95% of the observations to lie within a factor of 4 ($= s_c^2$) of the prediction.

As one would expect, s_c increases sharply with $\bar{h}$, and falls off rapidly with X. If we insist that a 'good' model should predict within a factor of two of the observations 95% of the time, the required s_c is $\simeq 1.4$. We see that for $\bar{h} = 0.1$, the uncertainty meets this requirement for $X > 0.21$. When $\bar{h} = 0.25$, the model should perform adequately for $X > 0.52$, and the corresponding X for $\bar{h} = 0.5$ is roughly 1.0.

It is easily shown that at the location at which the model predicts the maximum concentration, $s_c \simeq 1.8$ for all source heights. This means that we should expect 25% of the observations to lie outside a factor of two of the maximum predicted concentration. Clearly, this should be accounted for in regulatory applications of models. Note that the preceding results depend on the assumed values of the time scales of the vertical and horizontal velocity fluctuations.

We have assumed that w_0 is the only unknown variable that is responsible for the variance of concentrations. We can readily extend our analysis to a case in which the ensemble is specified in terms of a measured velocity u which differs from the 'true'

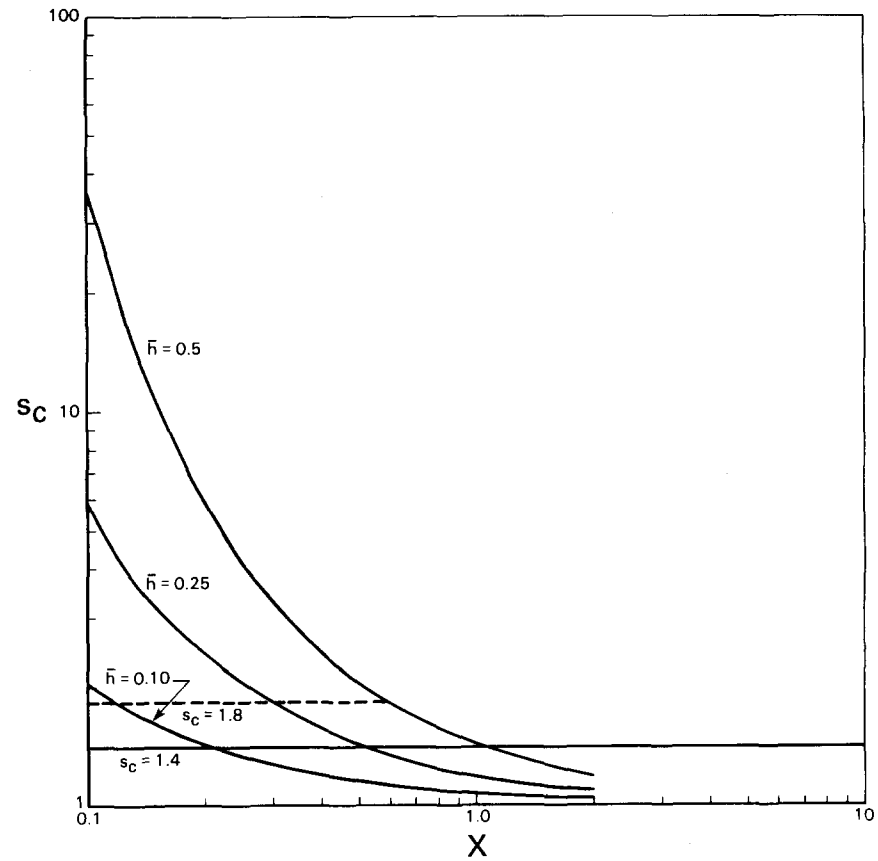

Fig. 4. Variation of the geometric standard deviation of concentration fluctuations s_c with $X \equiv w_* x/z_i u$.

transport velocity by an expected error ε_u. This error could be the difference between the estimated and the actual wind speed in the mixed layer. Then it is easy to show that

$$\sigma(\ln C(x, 0)) \simeq \frac{h}{\sigma_z}\left[\frac{2T_w}{T} + \left(\frac{\varepsilon_u}{u}\right)^2\left(\frac{h}{\sigma_z}\right)^2\right]^{1/2}. \tag{23}$$

If we take $\varepsilon_u/u = 0.1$, we find that except for large h/σ_z (close to the source), the first term in the parenthesis is much larger than the second.

6. The Estimation of Concentration Variance from Data

The uncertainty estimates provided by the preceding models account for only part of unknown variables determining $c(\alpha, \beta)$. However, they identify the contribution to the concentration variance that is correlated to the model inputs α. It always necessary to construct a model for this component of the concentration variance before we can estimate the total variance from data. To see this, let us assume that we have an adequate

model for the ensemble mean. Then the deviation between a model prediction $C_p(\alpha)$ and the corresponding observation $C_0(\alpha, \beta)$ is

$$c(\alpha, \beta) = C_0 - C_p \,. \tag{24}$$

In order to estimate $\sigma_c(\alpha)$ empirically, we need to have a sample of $c(\alpha, \beta)$ for the same α. However, this type of sampling is impossible in practice because each observation corresponds to a different ensemble. The only way to get around the problem is to normalize $c(\alpha, \beta)$ such that it is independent of α. To achieve this, we have to assume that $c(\alpha, \beta)$ can be written as

$$c(\alpha, \beta) = \mu(\beta) f(\alpha) \,; \tag{25}$$

then

$$\sigma_c(\alpha) = a f(\alpha) \tag{26a}$$

where

$$a^2 = \overline{\mu^2(\beta)}^{\beta} \,. \tag{26b}$$

These equations suggest that we have to formulate $f(\alpha)$ in order to estimate $\mu(\beta)$ which is independent of α. Then a^2 corresponds to the variance of the normalized residual $(C_0 - C_p)/f(\alpha)$.

Note that the manipulation of residuals between observations and predictions to generate model performance statistics is meaningful only if $c(\alpha, \beta)$ is independent of the model inputs. Binning of residuals into 'similar' categories before calculating statistics is not useful because the definition of 'similar' is arbitrary. This points to the necessity of formulating models for the mean and the variance of concentrations before evaluating a model.

7. Summary

This paper proposes a framework to estimate the uncertainty in predictions of concentration from air pollution models. We show that a prerequisite for this estimate is a precise definition of the ensemble in terms of the known variables determining the concentration. The variation of the variables not included in the ensemble definition is responsible for concentration fluctuations or model uncertainty. This concept is illustrated through two examples.

Our analysis suggests that it is necessary to formulate a model for concentration fluctuations before one can determine the expected deviation of model predictions from observations.

References

Chatwin, P. C.: 1982, 'The Use of Statistics in Describing and Predicting the Effects of Dispersing Gas Clouds', *J. Hazardous Materials* **6**, 213–230.

Chatwin, P. C. and Sullivan, P. J.: 1979, 'The Basic Structure of Diffusing Contaminant', in *Mathematical Modeling of Turbulent Diffusion in the Environment*, Academic Press, N.Y., pp. 3–32.

Csanady, G. T.: 1973, *Turbulent Diffusion in the Environment*, D. Reidel Publ. Co., Dordrecht, Holland.

Durbin, P. A.: 1980, 'A Stochastic Model of Two-Particle Dispersion and Concentration Fluctuations in Homogeneous Turbulence', *J. Fluid Mech.* **100**, 279–302.

Kaimal, J. C., Wyngaard, J. C., Haugen, D. A., Coté, O. R., Izumi, Y., Caughey, S. J., and Readings, C. J.: 1976, 'Turbulence Structure in the Convective Boundary Layer', *J. Atmos. Sci.* **33**, 2152–2169.

Lamb, R. G.: 1982, 'Diffusion in the Convective Boundary Layer', in *Atmospheric Turbulence and Air Pollution Modelling*, D. Reidel Publ. Co., Dordrecht, Holland, pp. 159–230.

Lumley, J. L. and Panofsky, H. A.: 1964, *The Structure of Atmospheric Turbulence*. Interscience Publishers, John Wiley and Sons, New York.

Panofsky, H. A.: 1978, 'Matching in the Convective Boundary Layer', *J. Atmos. Sci.* **35**, 272–276.

Venkatram, A.: 1979, 'The Expected Deviation of Observed Concentrations from Predicted Ensemble Means', *Atmos. Env.* **13**, 1547–1549.

Venkatram, A.: 1983, 'On Dispersion in the Convective Boundary Layer', *Atmos. Env.* **17**, 529–533.

Venkatram, A.: 1983, 'The Uncertainty in Estimating Dispersion in the Convective Boundary Layer', *Atmos. Env.* (in press).

Weil, J. C. and Brower, R. P.: 1982, *The Maryland PPSP Dispersion Model for Tall Stacks*, Martin Marietta Report No. PPSP-MP-36.

Willis, G. E. and Deardorff, J. W.: 1978, 'A Laboratory Study of Dispersion from an Elevated Source within a Modeled Convective Planetary Boundary Layer', *Atmos. Env.* **12**, 1305–1311.