ENVIRONMETRICS
Environmetrics 2000; **11**:351–371

Bayesian uncertainty estimation methodology applied to air pollution modelling

Renata Romanowicz[1]*, Helen Higson[2] and Ian Teasdale[3]

[1] *IENS, Lancaster University, Lancaster LA1 4YQ, U.K.*
[2] *Cambridge Environmental Research Consultants, 3 Kings Parade, Cambridge CB2 1SJ, U.K.*
[3] *Math Engine plc, Oxford Centre for Innovation, Oxford OX2 0JX, U.K.*

SUMMARY

The aim of the study is an uncertainty analysis of an air dispersion model. The model used is described in NRPB-R91 (Clarke, 1979), a model for short and medium range dispersion of radionuclides released into the atmosphere. Uncertainties in the model predictions arise both from the uncertainty of the input variables and the model simplifications, resulting in parameter uncertainty. The uncertainty of the predictions is well described by the credibility intervals of the predictions (prediction limits), which in turn are derived from the distribution of the predictions. The methodology for estimating this distribution consists of running multiple simulations of the model for discrete values of input parameters following some assumed random distributions. The value of the prediction limits lies in their objectivity. However, they depend on the assumed input distributions and their ranges (as do the model results). Hence the choice of distributions is very important for the reliability of the uncertainty analysis. In this work, the choice of input distributions is analysed from the point of view of the reliability of the predictive uncertainty of the model. An analysis of the influence of different assumptions regarding model input parameters is performed. Of the parameters investigated (i.e. roughness length, release height, wind fluctuation coefficient and wind speed), the model showed the greatest sensitivity to wind speed values. A major influence on the results of the stability condition specification is also demonstrated. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: Gaussian air dispersion model; sensitivity analysis; Bayesian uncertainty estimation; likelihood functions; prior and posterior probability density functions; prediction errors; prediction limits

1. INTRODUCTION

This work addresses the problem of the uncertainty of predictions of an air pollution model and their dependence on the uncertainties of observations. The uncertainty in atmospheric dispersion modelling can be a result of both the uncertainty in modelled atmospheric processes and observation errors, and the structural and numerical errors of the mathematical model. Structural

* Correspondence to: R. Romanowicz, IENS, Lancaster University, Lancaster LA1 4YQ, U.K. e-mail: R.Romanowicz@lancaster.ac.uk

Received 1 October 1998
Accepted 7 November 1999

errors of the mathematical model originate from the simplifications involved in the description of the atmospheric processes. Also the scale represented by measured physical variables and the scale of their representation in the mathematical model differ. In the search for the physical representation of variables, scientists often forget that the effective values of model parameters at a certain scale may not be truly represented by the quantities measured in the field. This type of uncertainty can be accounted for, to a certain extent, by the application of parametric uncertainty methodology and conditioning model results on observations. However, the degree to which the uncertainties can be decreased depends on the amount of available information. Model structure limitations can substantially influence the predictions because the formulae used necessarily include simplifying assumptions which differ from reality and have been developed under specific conditions. Moreover, the random nature of atmospheric processes does not allow the model assumptions to be fully met.

The proposed methodology is based on the concept of Bayesian Inference (e.g. Box and Tiao, 1992; Haylock and O'Hagan, 1997). The method presented in this paper has been used in hydrology as the Generalised Uncertainty Estimation Technique (GLUE) (Beven and Binley, 1992; Romanowicz *et al*., 1994), and applies likelihood measures to estimating the predictive uncertainty of the model. Simulated model outputs are compared with available observations of the variables of interest and the distribution of the resulting errors of predictions is used to derive the credibility intervals for the predictions. The value of the credibility intervals (prediction limits) lies in their objectivity. However, as the model results depend on the information included in its input and output measurements, prediction limits will depend on the assumed input and output distributions and their ranges. Hence the choice of these distributions is very important for the reliability of the uncertainty analysis. A sensitivity analysis of the model variables may also be important in gaining a better understanding of the model performance and its internal structure. However, sensitivity analysis on its own is not sufficient to estimate the errors of model predictions and should be followed by uncertainty estimation techniques based on the comparison of model predictions with observations.

When applying the Gaussian plume model described in the next section to the prediction of ground-level concentrations of a pollutant, all input variables must be specified, based on the analysis of the conditions in which the release takes place and, in particular, on the atmospheric conditions and geographical features of the terrain. In the case of regulatory applications, it is crucial to know how reliable the model predictions are and how variability in the measurements of the environmental variables will affect model performance. The amount of information we possess about these input variables will vary, depending on the type of variable, the accuracy of the measurements and on the role which the variable plays in the model.

The resulting prediction limits rely strongly on assumptions regarding the distribution of input variables and parameters. As neither the model parameters nor the input distributions are known exactly, it is important to investigate the influence of different assumptions regarding these variables and parameters on the predictive uncertainty of the model. This analysis will also lead to recommendations about the degree of accuracy of the measurements. The measurements of the variables which have the largest influence on model performance (e.g. wind speed) will be more important than the measurements of global parameters, which do not have a fully physically-based interpretation in the model (such as roughness length). Sensitivity analysis will be used to examine the influence of the input variables and parameter variability on the model results, while uncertainty analysis will provide the analysis of influence of parameter and input distributions on prediction limits.

2. PROBLEM DESCRIPTION

2.1. Short description of the Gaussian plume model structure

The Gaussian plume model for a continuous source originates in the work of Sutton (1932). Pasquill and Smith (1983) and Gifford (1961, 1968). It is obtained as a solution to the Fickian diffusion equation for constant diffusivity coefficient and uniform wind speed. The model is derived as a steady state solution of the basic transport model. The assumption of constant diffusivity is valid only if the size of the plume is greater than the size of the dominant turbulent eddies, so that all turbulence implicit in this parameter is taking part in the diffusion. In fact, diffusivity is seldom constant in time and space. Also the assumption of constant wind speed is very restrictive, as wind speed varies with height throughout the atmospheric boundary layer. Assuming that the meteorological conditions are constant during the travel of the plume limits the application of the model to short time periods (15 minutes to 1 hour) and small distances (up to approximately 30 km). These limitations should be taken into account when applying the model. Even though these conditions are never met in practice, the Gaussian plume model has been widely adopted (see Hanna *et al.*, 1982).

The mathematical description of the Gaussian model used in the present work is given in NRPB-R91 (Clark, 1979), and is hereafter referred to as the R91 model. The basic equation using a Gaussian plume model for an elevated release has the form (Clarke, 1979):

$$C(x,y,z) = \frac{Q}{2\pi u_{10}\sigma_z\sigma_y}\exp\left[-0.5\left(\frac{y^2}{\sigma_y^2}+\frac{(z-h)^2}{\sigma_z^2}\right)\right] \tag{1}$$

where C is air concentration (g m^{-3}) or its time integral (g s m^{-3}); Q is the release rate (g s^{-1}) or total amount released (g); u_{10} is the wind speed at 10 m above the ground (m s^{-1}); σ_z is the standard deviation of the vertical Gaussian distribution (m); σ_y is the standard deviation of the horizontal Gaussian distribution (m); x is the rectilinear co-ordinate along the wind direction (m); y is the rectilinear co-ordinate for cross-wind (m); z is the rectilinear co-ordinate above the ground (m); and h is the effective release height (m). The origin of the co-ordinate system is at ground level beneath the discharge point.

The R91 model incorporates reflection of the plume from the ground and the top of the atmospheric boundary layer. In this case the model solution has the form:

$$C(x,y,z) = \frac{Q}{2\pi u_{10}\sigma_z\sigma_y}\exp\left[\frac{-y^2}{2\sigma_y^2}\right]F(h,z,A) \tag{2}$$

where

$$F(h,z,A) = \exp\left[-\frac{(z-h)^2}{2\sigma_z^2}\right]+\exp\left[-\frac{(z+h)^2}{2\sigma_z^2}\right]+\exp\left[-\frac{(2A+z+h)^2}{2\sigma_z^2}\right]$$
$$+\exp\left[-\frac{(2A+z-h)^2}{2\sigma_z^2}\right]+\exp\left[-\frac{(2A-z+h)^2}{2\sigma_z^2}\right]+\exp\left[-\frac{(2A-z-h)^2}{2\sigma_z^2}\right] \tag{3}$$

Equation (2) should be used at distances where $\sigma_z \leqslant A$ (where A denotes the depth of the boundary layer). When the value of the vertical dispersion coefficient σ_z becomes much greater than the depth of the boundary layer A ($\sigma_z \gg A$), the vertical concentration distribution effectively becomes uniformly distributed throughout the mixing layer. The concentration is then given by:

$$C(x, y, z) = \frac{Q}{(2\pi)^{1/2} u_{10} A \sigma_y} \exp\left[-\frac{y^2}{2\sigma_y^2} \right] \tag{4}$$

Typical values of boundary layer depth are given in Table 2 of NRPB-R91 (Clarke, 1979) for Pasquill–Gifford stability categories A to G.

The vertical standard deviation σ_z at a given distance from the source is modelled as being a function of the atmospheric stability, downwind distance and ground roughness. Hosker (1974) derived analytical formulae used in the R91 methodology (NRPB-R91) for specific roughness lengths and interpolation is used for intermediate values of roughness coefficient. This may be summarised as:

$$\sigma_z = F(x,z_0) * G(x, CAT) \tag{5}$$

where z_0 denotes roughness length, $F(x, z_0)$ is an expression which varies continuously with distance from the source and has a specific form for particular values of roughness length, CAT denotes the stability condition and $G(x, CAT)$ is a function which has a slightly different form for each stability condition, each of which is a continuous function of the distance downwind from the source.

The horizontal dispersion of the plume, characterised by the standard deviation σ_y, is the result of turbulence processes and fluctuations in wind direction. The formulation used in the R91 model uses an expression for σ_y given originally by Pasquill for very short (3 minute) releases. This may be applied for releases much less than 30 minute duration. For longer time releases some account must be taken of fluctuations in wind direction. The final formula for σ_y is due to Moore (1976) and has the form:

$$\sigma_y^2 = \sigma_{\mathrm{yt}}^2 + \sigma_{\mathrm{yw}}^2 \tag{6}$$

where σ_{yt} represents diffusion due to turbulence (3-minute term) and σ_{yw} is the component due to fluctuations in wind direction.

The values of σ_{yt} are given in Figure 10 of NRPB-R91 (Clarke, 1979), which shows separate curves for each stability condition (CAT), each of which varies continuously with the distance x:

$$\sigma_{\mathrm{yt}} = P(CAT) * x^{0.91} \tag{7}$$

σ_{yw} is evaluated using the following empirical formula (Jones, 1983):

$$\sigma_{\mathrm{yw}} = 0.076 \left[\frac{7T}{u_{10}} \right]^{1/2} x \tag{8}$$

where T denotes release duration (averaging period) in hours and the expression in brackets is non-dimensional.

These formulae for the horizontal distribution of a plume as represented by Equations (2) and (4) can be used for any duration of release longer than about 30 minutes for which weather category and wind direction may be assumed constant.

Explicit variables used in the R91 model are distance from the source, wind speed at 10 m height, effective release height, release rate, roughness length and weather category. All those variables must be specified by the user. Some of them can be treated as independent input variables and some act as parameters of the model.

The standard deviations of the vertical and horizontal Gaussian distribution are implicit model variables. They are given in the form of empirical relations (5) and (6), (7) and (8). These formulae depend on the explicit parameters, but also on some internal parameters based on field experiments, which are treated as fixed by the model. These are, for example, the depth of the boundary layer, which varies with the weather category, and the wind direction coefficient equal to 0.065 in Equation (8). Moreover, different relations apply depending on the weather category (i.e. the dependence is not continuous).

In particular, the vertical standard deviation (5) is a function of downwind distance, with the exact form of this function depending on the specified roughness coefficient and stability condition. Also the relationship describing the influence of wind direction deviations (8) in the definition of the horizontal plume dispersion (6) is very simplistic and may involve some error.

The stability categories for the weather used in R91 are described in NRPB-R91 (Clarke, 1979) and have the form of seven discrete conditions with associated values of wind speed and boundary layer height. When analysing the R91 relations, stability conditions affect the boundary layer height, horizontal dispersion and vertical dispersion.

The uncertainty involved in the specification of model variables influences the uncertainty of the model predictions and hence is essential in this analysis. In order to study this influence, variables are varied within specified ranges, according to the assumed random distribution as described in Section 3 below. In the case of the R91 model these will be: wind speed, stability condition, roughness length, release height, turbulent wind fluctuation coefficient and inversion layer height.

Even though all these variables must be specified to run the model in a particular application, some are treated as input variables and others as model parameters. When applying the model to predict ground concentrations from a pollutant release, some variables will be observed inputs (and therefore independent of the modeller) and some are treated as model parameters, to be calibrated. The distinction between input variables and model parameters is, to some extent, subjective and depends on a knowledge of the physical process being modelled and a detailed analysis of the model structure. This distinction is important when the variable ranges and distributions are estimated. The variations introduced for independent (input) variables should reflect the input observation errors and their ranges should not depend on the model performance. Dependent parameter ranges should be chosen to give the best model performance based on the comparison of model results with the available output observations.

Among the explicit R91 model variables, roughness coefficient should be treated as a dependent model parameter. It is defined as a value representing an integrated effect of the land surface on turbulent mixing; hence it is not measurable. The available point measurements of roughness coefficient will not correspond to the effective roughness length required by the model, as the surrounding area is not homogenous. All other variables can be treated as model inputs.

The next problem arises from the reliability of the measurements of the input variables. We can assume that the wind speed measurements are accurate, but wind speed is very variable, and as in the case of the spatially variable roughness length, some time averaged value of wind speed should be used in the model. The physical release height is a well defined parameter; however, there are a number of effects which may influence the 'effective' stack height, namely the presence of buildings and plume rise due to density/buoyancy driven effects. Hence it was decided to vary the release height as well. In addition to these explicit input variables, there are other hidden parameters in the model, estimated from empirical relations derived from different experiments. These are, among others, inversion layer height and wind direction fluctuation coefficient. Inversion layer height is assumed constant for a particular weather category in the model and is given by values in Table 2 of NRPB-R91 (Clarke, 1979). The turbulent wind fluctuation coefficient [Equation (8)] is represented by the value 0.065, which has been derived empirically from other field experiments; this has been allowed to vary here, to account for variations in horizontal plume spread.

The stability condition is used by the model as a switch leading to different model parameterisation. Information about the stability condition is also uncertain and its uncertainty should be introduced into the model. One way of doing so would be by modifying the model structure to account for the lack of information about the stability conditions. This can be equivalent to setting the stability condition in the model to one category only and using such a model to derive the predictions in different atmospheric conditions. Certain errors will be introduced in this way, which would reflect the lack of information about the correct stability condition. The other method may consist of changing the model structure such that the model chooses the stability condition depending on the wind speed and also on some other parameters, such as roughness length.

The main factors influencing the diffusion process in the atmosphere are not easily measurable, e.g. the friction velocity in turbulent motion over heterogeneous surfaces or a measure of the buoyancy generated by internal density or temperature differences. The uncertainty involved in the specification of model variables influences the uncertainty of the model predictions and hence is essential from the point of view of this analysis. The other essential source of errors come from the very limited range of data on the simulated variables, which limits the calibration of the model empirical parameters. In this respect also the observations should be treated as samples from a random process (Lee and Irwin, 1995).

2.2. Description of the Copenhagen data set

The methodology was applied to the results of the Copenhagen experiment (Gryning, 1981). The Copenhagen data consist of 23 sets of cross-wind maximum concentration measurements from 1-hour releases of SF_6 tracer, taken at different times. The experiment was performed in neutral and unstable conditions corresponding to Pasquill–Gifford stability categories C and D. The tracer SF_6 was released without buoyancy from a tower at a height of 115 m. The measurements were obtained at 10 different irregular distances from the source varying from 1.9 km to 6 km, from up to three cross-wind series of tracer sampling units. The value of roughness coefficient was estimated as 0.6 m. Meteorological measurements included vertical profiles of wind speed at 10 m and 100 m height. The character of the data indicates that, using the Pasquill–Gifford classification scheme, only two stability conditions must be analysed. This fact simplifies considerably the R91 options that should be analysed using the Copenhagen experiment data.

Moreover, in category C, reliable observations are present only for four distances, with at the most two samples at each distance. For category D, there are at most four samples at six different distances. This scarity of observations does not enable any statistical analysis of the dependence between the observations to be performed. Deposition processes are neglected in the analysis. Observations of cross-wind-integrated concentrations are also available in the Copenhagen dataset. However, in practical applications, the model is most commonly used to estimate maximum concentrations at a specified location, or to estimate maximum ground level concentrations. Hence, the analysis was limited only to the maximum cross-wind point concentrations.

3. SENSITIVITY ANALYSIS OF R91 MODEL

The sensitivity of the model predictions to the following five parameters will be sought: roughness length z_0, inversion layer height coefficient β, $A' = A + \beta$; where A' denotes the inversion layer height [Equation (2)] and β denotes the variation of inversion layer height around its reference level A; turbulent wind fluctuation coefficient α, such that 0.065 becomes $(0.065 + \alpha)$ in equation (8); wind speed u and release height h. The choice of parameters was based on the analysis of the model equations and a preliminary sensitivity analysis of model responses.

The sensitivity analysis was performed using a statistical analysis of output which utilises Monte Carlo techniques (Helton, 1993). It is based on multiple model evaluations, with input and parameter values of the model selected according to the chosen probabilistic sampling. The results of the simulations are used to determine the uncertainties in model predictions and their relation to the uncertainty of input variables. The sensitivity results are obtained without the use of an intermediate surrogate model by exploring the mapping from model input to model predictions. Among the methods of analysis of the Monte Carlo generated random output variables are scatter plots, sample mean, variance, output distribution function, prediction limits and Spearman-ranked correlation coefficient. Use of only two moments from the sample (mean and variance) to characterise the variability of the output is equivalent to assuming that this variable follows a Gaussian distribution. Obviously, large amounts of information can be neglected in such summary statistics. Another way of summarising the variability in the model output is by the estimation of a distribution function (e.g. Helton, 1993). The Spearman-ranked correlation coefficient (Yong *et al.*, 1993) uses the ranks of sample values rather than the sample values themselves. Plots of standardised regression coefficients or partial correlation coefficients as functions of time or location may also be used.

The ranges and assumed distributions for all the varied parameters and input variables are given in Table I. All the parameter and input variable deviations were assumed to be uniform,

Table I. Parameter ranges for sensitivity analysis used in multiple model simulations.

Parameter	Symbol	Distribution	Lower limit	Upper limit
Turbulent wind fluct. coeffic.	α [ ]	uniform	-0.1	0.1
Wind speed	u [m s^{-1}]	uniform	0.1	8.0
Inversion height coeff.	β [m]	uniform	-100	100
Roughness length	z_0 [m]	uniform	0.0001	4.0
Stack height	h [m]	uniform	105	125

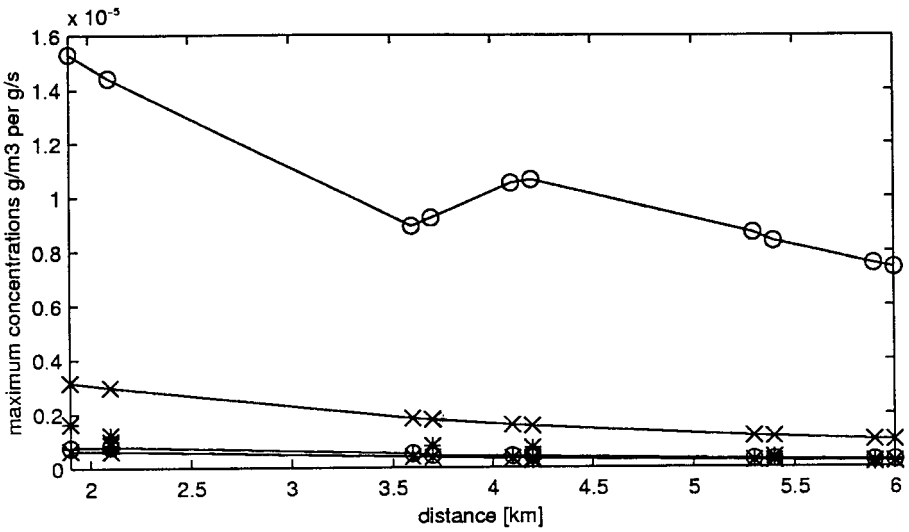

Figure 1. Ninety-five per cent prediction limits for the predictions obtained from the sample output distribution without conditioning (○); 95 per cent prediction limits with the use of the input observations (X); observations from the Copenhagen data set (*).

to obtain a clear (not biased) image of the model sensitivity. The stability category was set to D. From the scatterplots of model results against parameter values, it is not evident which parameter values give the best model performance. The output results were used to evaluate the sample distribution. The 95 per cent prediction limits found using this distribution are shown in Figure 1. As the estimated prediction limits are very wide, the prediction value of these results is rather poor. These predictions would be equivalent to the case when the model is used in an accident risk assessment, under partly unknown meteorological conditions. Under a worst case scenario, it should be assumed that the polluting substance is released from the source at the maximum possible rate in the specified direction of human settlements. The predictions should be performed for all possible wind speed ranges, corresponding to each stability category (A–G). The prediction limits obtained with the use of input information (equivalent to sampling from the normal or log-normal distribution with the mean value given by the observed values and spread estimated from the possible variations) are also shown on Figure 1. They are much narrower than in the case of uniform input distributions.

Spearman ranked correlation coefficients for 10 distances from the source and five parameters are shown on Figure 2. The coefficients for the wind speed and wind fluctuation coefficient have the largest absolute values, indicating that these variables have the greatest influence on the results. Next is roughness length for the first two distances from the source (1.9 and 2.1 km) and release height (for the same distances). This result is consistent with the physical features of the dispersion process under category D stability conditions. For elevated releases, both roughness length and the release height control the plume dispersion before it reaches the ground (i.e. at small distances).

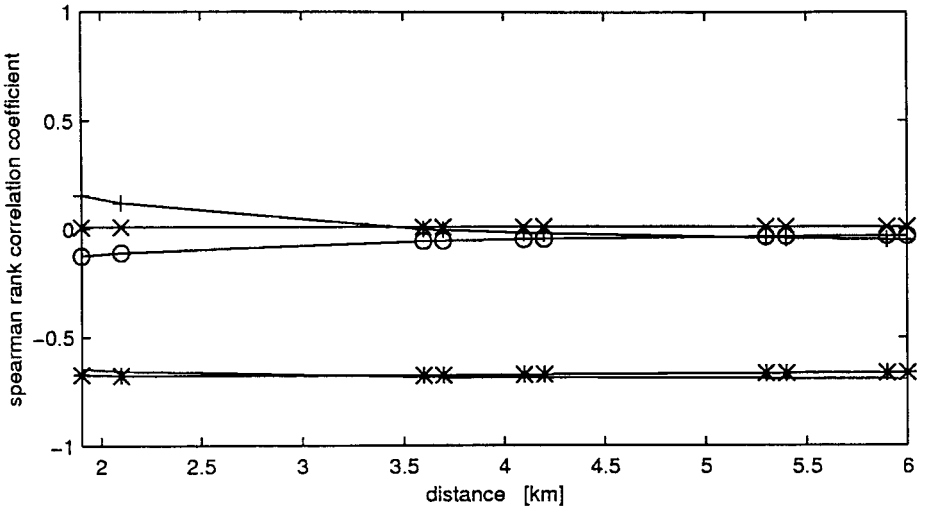

Figure 2. Spearman ranked correlation coefficients versus distance from the source, uniform input distributions (Table III) (✳) wind speed; (+) release height; (×) inversion height; (—) wind direction coefficient; (○) roughness length.

4. APPLICATION OF THE BAYESIAN UNCERTAINTY ESTIMATION TECHNIQUE TO THE R91 AIR DISPERSION MODEL

4.1. Short description of the methodology

Uncertainty analysis based on the statistical analysis of output alone does not give any information about the validity of the model predictions. Uncertainty results are obtained by exploring the mapping from model input to model predictions. In that sense, the predictive uncertainty of the model, understood as the probable error of its predictions, is still not known. Bayesian Inference is a methodology which provides tools for the comparison of model results with observations.

In this study we follow the methodology developed in Romanowicz *et al.* (1994) for the case of hydrological rainfall–runoff modelling. It is assumed that the errors between observed Z_i and simulated variables (maximum cross-wind concentrations at given distances) have the additive form:

$$Z_i = g_i(\theta) + \delta_i \quad i = 1, \ldots, n \tag{9}$$

where $g_i(\theta)$ denotes simulated by the model maximum cross-wind concentrations being the function of θ; θ denotes a vector of model parameters and input variables; δ_i denotes model errors and i denotes observation points at different distances from the source, n denotes a number of measurement distances from the source. Here, δ (dim $\delta = n$) is modelled by the Gaussian model with non-zero mean μ, and covariance matrix Σ.

Due to the sparse amount of Copenhagen observation data, for this univariate case, a stationary process was assumed for δ, with constant mean μI and $\Sigma = \sigma^2 I$ with no correlation between the errors at the observation sites. This assumption may be partly justified by the fact that the

observations were obtained from independently performed series of 1-hour releases. In general, the observations of ground concentration of the release along the distance from the source are correlated. The methodology presented may be applied to this more general case using a correlated noise process δ. Also the assumption about additive noise in Equation (9) should be re-examined when more data are available. A logarithmic form of this equation should be used when the errors are assumed to have multiplicative form (see Romanowicz *et al.*, 1994).

From the error model it is seen that the likelihood function of the predicted concentrations can be expressed as the likelihood of the error variate with parameters $(\phi, \theta) = (\mu, \sigma, \theta)$, depending on the air dispersion model parameters and input variables θ and noise parameters μ, σ.

Under these assumptions the likelihood function is defined as (Romanowicz *et al.*, 1994; Griffith, 1988):

$$f(\mathbf{z}|\theta, \phi) = \prod_{i=1,n} f_{\delta_i}(z_i - g_i(\theta)|\phi) \tag{10}$$

where δ is given by (9) and

$$f_{\delta_i}(\delta|\phi) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{1}{2\sigma^2} (\delta - \mu) \right)^2$$

The above distribution (10) denotes the probability of the observations $\mathbf{z}$ given the input data and model and error model parameters as a function of input variables and model parameters. It will be used to derive the predictive model uncertainty by applying Bayes theorem (Box and Tiao, 1992):

$$f(\theta, \phi|\mathbf{z}) = \frac{f(\theta,\phi) f(\mathbf{z}|\theta, \phi)}{f(\mathbf{z})} \tag{11}$$

where $\mathbf{z}$ is the vector of observations, $f(\theta, \phi|\mathbf{z})$ denotes the posterior distribution of input variables and model parameters θ given the output observations $\mathbf{z}$, $f(\theta, \phi)$ denotes the prior distribution of θ and ϕ. $f(\mathbf{z}|\theta, \phi)$ denotes the likelihood function (10), $f(\mathbf{z})$ denotes the probability of observations and can be treated as a scaling factor. Under the assumption of independence between modelling errors and parameters, this joint distribution may be written as $f(\theta) f(\phi)$ where $f(\theta)$ denotes a prior distribution for the parameters/input variables (see the discussion about its form in the next section), and $f(\phi)$ denotes a prior distribution of parameters of the error model (9) and is assumed uniform.

Equation (11) can be applied sequentially as new data become available and the existing posterior distribution, based on $(n-1)$ calibration sets, is used as a prior distribution for the new data in the nth calibration set. This can be written in the form:

$$f(\theta,\phi|\mathbf{z}_1, \ldots, \mathbf{z}_n) \propto f(\theta,\phi|\mathbf{z}_1, \ldots, \mathbf{z}_{n-1}) f(\mathbf{z}_n|\theta, \phi) \tag{12}$$

where $f(\mathbf{z}_n|\theta, \phi)$ is the information about θ and ϕ from the nth calibration set.

Differences between observations and simulated model output values, together with the

assumed prior distributions of parameters, are used to build the posterior distribution of parameters reflecting the model performance. In this way it is possible to incorporate the information from observations from different time periods and/or sites using the Bayesian updating described by Equation (12).

The cumulative distribution of the error term at a location i, given a particular set of statistical model parameters $\phi = (\mu, \sigma)$, is then given by:

$$P(\delta_i < \delta|\mu,\sigma) = \Phi\left(\frac{\delta-\mu}{\sigma}\right) \quad i = 1,\ldots,n \tag{13}$$

where Φ is a standard normal distribution function $N(0, 1)$. In general, it is a continuous distribution, which is discretised for computational convenience.

The resulting predictive distribution of tracer concentrations Z^* conditioned on the calibration data $\mathbf{z}$ is given by

$$P(Z^* < z|\mathbf{z}) = \sum_{\theta}\sum_{\phi}\Phi\left(\frac{z-g(\theta)-\mu}{\sigma}\right)f(\theta, \phi|\mathbf{z}) \tag{14}$$

From relation (14) one can evaluate the predictive limits for the concentrations at the observation sites. The assumption that the errors are independent was introduced due to the limited number of observations for the Copenhagen data set with individual observations made during different time periods and in different stability conditions for the seven distances from the source. Whenever possible, the correlation of the errors between different observation sites should be checked and the appropriate model should be chosen, e.g. as suggested in Romanowicz *et al.* (1994).

Assumptions about the form of the error model and the prior distribution of parameters are all rather subjective. Due to the limited amount of input observations and their stochastic nature, assumptions regarding their distributions are also subjective. In particular, as a result of the stochastic nature of observations of concentrations on the ground, the error model will be much more difficult to justify than in the case of hydrological data. However, statistical analysis requires certain assumptions about the nature of the processes to be made and one should be always aware of the limitations of the method used. Discussion on the choice of prior distributions is very extensive in the statistical literature (e.g. Box and Tiao, 1992). From a practical point of view, the prior distribution of the calibrated parameters should be non-informative (see Woodbury and Ulrych, 1993). The analysis of the influence of the form of prior distribution of model parameters on the predictions is one of the aims of this paper. Also the analysis of influence of different assumptions concerning the distribution of input covariates on the resulting model predictions will be performed. The distinction between the input covariates and parameters (i.e. independent and dependent variables) is important only from the point of view of model calibration. From the point of view of uncertainty analysis, it influences the statement of the problem only when the choice of priors is concerned. The prior distribution of input variables will depend, to a greater degree on measurements in the field than the priors for the parameters. The technique outlined will be applied to the evaluation of the uncertainty of the R91 model predictions. The uncertainty analysis is based on the Copenhagen data set, outlined in Section 2.2 and described in detail in Gryning (1981). The uncertainty analysis is applied only to model predictions of near ground-level concentrations. It is assumed that there is no deposition to the ground. Near field

data and far field data are treated separately, which follows from a different process description at different distances (see discussion by Lee and Irwin, 1995). Near field data will be used for the derivation of predictive distributions of model response, while the far field data will be used for the validation of the predictions.

4.2. The analysis of influence of observations of model variables

The analysis will include the influence of: (i) roughness length, (ii) wind speed, (iii) turbulent wind fluctuation coefficient and (iv) release height. The choice of these variables was made using the results of the sensitivity analysis described in Section 3. Marginal posterior distributions will be defined from Equation (11) summed over the noise model parameters, using the likelihood functions based on the error between the simulated and observed values (Equation (9)]. Following the analysis of model structure, it was assumed that the roughness coefficient can be treated as a calibration parameter, while the rest of the variables form the set of independent, input variables. This assumption was tested through the comparison of results obtained without the observations (uniform distribution for the input variables) and with the observations (normal–log-normal distributions). Testing of different standard deviations for the normal-type distribution (normal or log-normal) was also performed for some input variables in order to draw conclusions regarding the influence of the variability of the measurements.

4.2.1. Influence of roughness length. The analysis of the roughness coefficient influence was performed in two stages: (i) uniform distribution in a wide range, with parameter values given in Table II; (ii) log-normal distribution of roughness length with three values of standard deviation (0.1, 1 and 2 m). The analysis was done with all the remaining parameters varying, so as to take into account the interactions between model parameters.

The resulting posterior distributions for the four different parameters as given in Table II (i) (roughness length, release height, wind direction fluctuation coefficient and wind speed) for near field (combined posterior distributions for distances up to 5 km from the source) are given in Figure 3. Updating of the posterior distributions is performed using the Bayes formula (10). Only posterior distributions for the near field distances were updated. The posterior distributions are shown as the projection of points resulting from Equation (10) onto one parameter space (release height, roughness length, wind speed or turbulent wind direction coefficient). Only wind speed shows a pronounced influence on model predictions, with the scatter plots for the other parameters showing no specific regions of better model performance. There is no evidence of a 'best' region for roughness length, within which the measure of model performance (posterior

Table II. Parameter ranges for analysis of roughness coefficient distribution – uniform case (i) and normal/log-normal case (ii).

Parameter	Distribution (i)	Lower limit	Upper limit	Distribution (ii)	Mean	SD
z_0 [m]	uniform	0.0001	4	log-normal	0.6	0.1/1.0/2.0
h [m]	uniform	105	125	normal	115	10
α [ ]	uniform	−0.1	0.1	normal	0	0.1
u [m s^{-1}]	uniform	0.1	11	log-normal	5	0.1/1.0/2.0

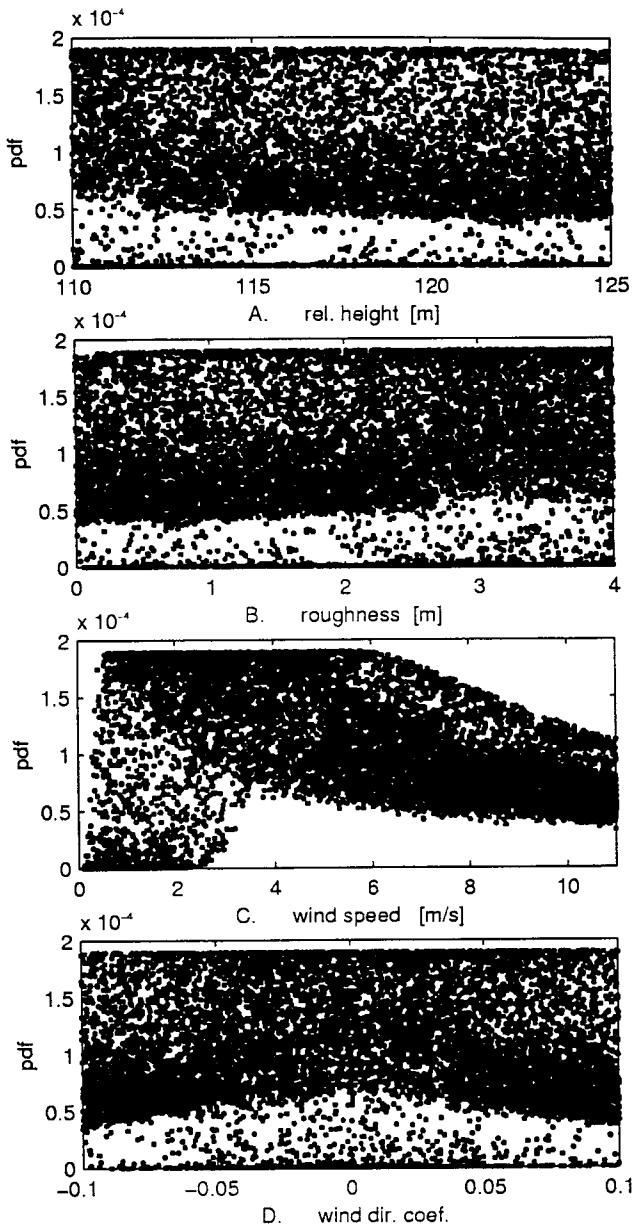

Figure 3. Combined likelihood functions (posterior pdfs) for the distances up to 5 km from the source for parameters as given in Table II(i).

probability distribution) would have its maximum. The 'best' parameter values correspond to the maximum values of the posterior distributions. Hence, the plots suggest that extending the parameter ranges will not result in any better model performance. The experiments were per-

formed with much larger parameter ranges than shown in the figures and the results were not better. When the parameter ranges assumed are too large, there is a danger of missing some particular features of model performance and that is why the parameter ranges are kept as small as possible.

One reason for the existence of the regions of 'non-preference' for some parameters is over-parameterisation of the model. In other words, there are many combinations of parameters giving the same model performance measured with regard to the given set of output observations. From the behaviour of scatter plots alone, it is not possible to say if certain model parameters do or do not have influence on the model output. However, the results of the sensitivity analysis (Figure 2) may be helpful in this respect. For example, they indicated that the roughness length has some influence on model predictions. Going back to the formulae describing the model, it can be seen that roughness length influences the vertical dispersion of the plume [Equation (5)]. This means that observations of vertical dispersion of tracer concentrations (e.g. from wind tunnel experiments) might allow the roughness coefficient to be calibrated. However, the observations from the Copenhagen experiment are not sufficient for this task (see Figure 3).

The resulting 95 per cent prediction limits are shown in Figure 4. The model predictions are based on near-field observations (<5 km from the source), hence the resulting prediction limits for the far-field ($\geqslant 5$ km) may be treated as the validation stage of the model performance. Prediction limits are wide enough to enclose most of the observation points. However, there is a visible model over-prediction and the observations lie on or near the lower prediction limit boundary. In spite of the fact that near-field observations were used for the building of posterior probabilities of the predictions, the largest uncertainty in the predictions occurs at small distances; this is an inherent feature of the Gaussian plume model.

In the second stage (not shown in the figures), the assumption regarding the lack of information about the roughness coefficient was tested by using the log-normal prior distribution for the

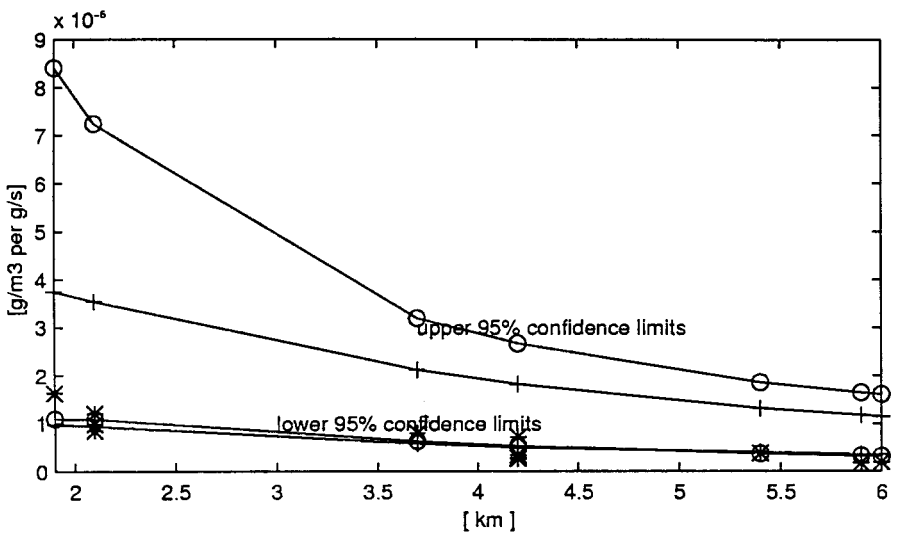

Figure 4. Ninety-five per cent prediction limits based on predictive posterior probability density functions derived from those shown on Figure 3 (○); 95 per cent prediction limits for the roughness length uniform (0.001–4 m) and normal height distribution (+); the observations are marked by stars (★).

roughness length, with mean equal to that indicated by the Copenhagen experiment description ($z_0 = 0.6$ m). After conditioning on near-field observations, the prediction limits for the log-normal roughness length were narrower than in the case of the uniform distribution for roughness length, but the lower prediction limit was too high to enclose some of the observation points. This indicates that the recommended mean value is not adequate for the R91 model in this application. This might be interpreted as a confirmation of our assumption of treating the roughness length as a parameter.

4.2.2. Influence of release height. Release height has been treated as an input variable for an inter-comparison with the Copenhagen data, as it should be generally well-defined in this case. However, there are other cases in which the concept of effective stack height is used (e.g. buoyant plume or to account for building effects), and, consequently, the release height used in the model (i.e. effective stack height) differs from the measured physical stack height. To represent the possible prior information about release height, we shall assume that it is normally distributed with a mean value equal to the measured release height (115 m) and standard deviation equal to 10. The resulting posterior predictive limits, with the remaining parameters as given in Table II (i), are shown on Figure 4. The figure illustrates that information about release height decreases the predictive limits considerably and the new predictions are more 'central' with respect to observations. This means that the information about the release height is very important in the case of the Copenhagen experiment. The prediction limits obtained for release heights varied by a standard deviation of 1 m (not shown) were only very slightly narrower than those shown in Figure 4. In conclusion, conditioning on release height is important and improves the predictions, but only to a certain extent.

4.2.3. Influence of turbulent wind fluctuation coefficient. The turbulent wind fluctuation coefficient is an internal model parameter derived empirically from earlier experiments not related to the given model application. The prediction limits obtained from the simulations with uniform distribution of this coefficient were compared with the results obtained when the coefficient was varied normally around its experimentally derived value. The results in prediction limits did not show any pronounced differences, which means that the observations are not giving any additional information about the parameter value. More detailed studies are necessary to determine what distribution this parameter should take.

4.2.4. Influence of wind speed. The wind speed values are the basic R91 input variables. In this section, the requirements regarding the accuracy of their measurements will be examined. Wind variations will be altered, with the rest of the parameters set to the values shown in Table III. The stability condition will be set as neutral (category D). It is assumed that the wind speed

Table III. Parameter ranges for sensitivity analysis for wind speed distribution.

Parameter	Distribution	Lower limit	Upper limit	Mean	SD (1)	SD (2)	SD (3)	SD (4)
z_0 [m]	uniform	0.0001	4	–	–	–	–	–
α [ ]	uniform	−0.1	0.1	–	–	–	–	–
u [m s^{-1}]	log-normal	–	–	5.0	0.1	0.5	1.0	2.0
h [m]	normal	–	–	115	10.0	10.0	10.0	10.0

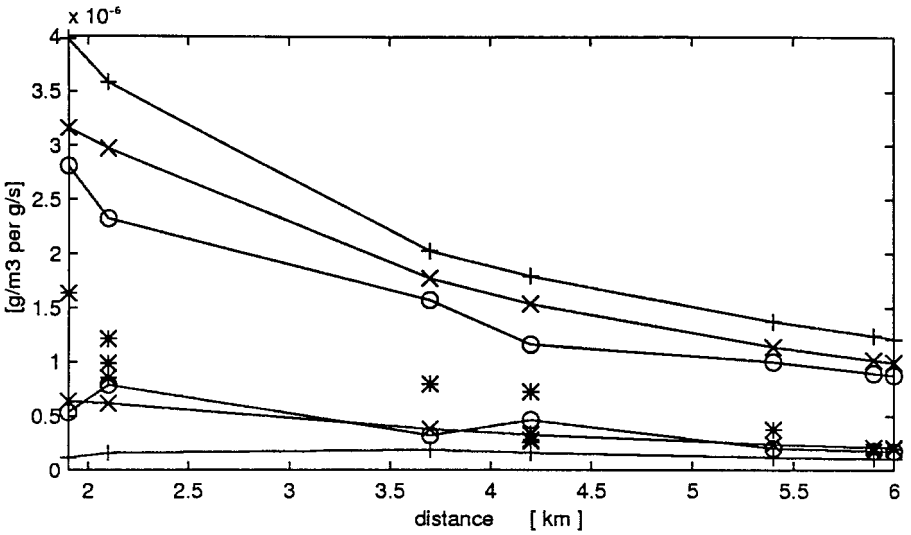

Figure 5. Ninety-five per cent prediction limits for the predictions estimated using posterior updated predictive probabilities for log-normal wind speed distribution and parameters given in Table III with wind speed standard deviation equal to (−) SD = 0.1 m s^{-1}; (×) SD = 0.5 m s^{-1}; (○) SD = 1.0 m s^{-1}; (+) SD = 2.0 m s^{-1}.

follows a log-normal distribution with mean value equal to 5 m s^{-1}. The effects of assuming different values of standard deviation were examined.

Ninety-five per cent prediction limits obtained from the updated posterior predictive distribution for all four values of the variance are presented in Figure 5. The results show that the prediction limits for the smallest standard deviation of wind speed (0.1 m s^{-1}) are wider than the results for the standard deviation equal to 1 m s^{-1}. The largest standard deviation (equal to 2 m s^{-1}) shows wider prediction limits, as expected. Physically low wind speeds give unreliable and highly variable results as the Gaussian model breaks down for low wind speeds.

These results indicate that wind speed plays an important role in the predictions. This clearly illustrates one of the limitations of the Gaussian plume model, which assumes that the wind speed is constant with height. In reality the wind speed varies significantly with height, and model predictions will be highly dependent on the value of wind speed used and how much the actual wind speed varies with respect to this value. The prediction limits obtained assuming a log-normal variation in wind speed are more central to the observations than the results obtained before, for uniform wind speed variations.

The information about wind speed and weather category are not equivalent, hence both should be taken into account simultaneously, when all the information about the experiment is to be used.

4.2.5. Influence of weather category. Analysis of the influence of weather category information on the predictions was performed by introducing uncertainty in the choice of weather category. The choice of Pasquill–Gifford stability category is always rather subjective and may vary from person to person. To investigate the effect of this further, the model was allowed to select the weather category based on some initial assumptions. This is equivalent to introducing a different

Table IV. Parameter ranges for sensitivity analysis for wind speed distribution.

Parameter	Distribution	Lower limit	Upper limit	Mean	SD
z_0 [m]	uniform	0.0001	4		
α [ ]	uniform	−0.1	0.1	–	–
u [m s^{-1}]	uniform	0.1	20		
h [m]	normal			115	10

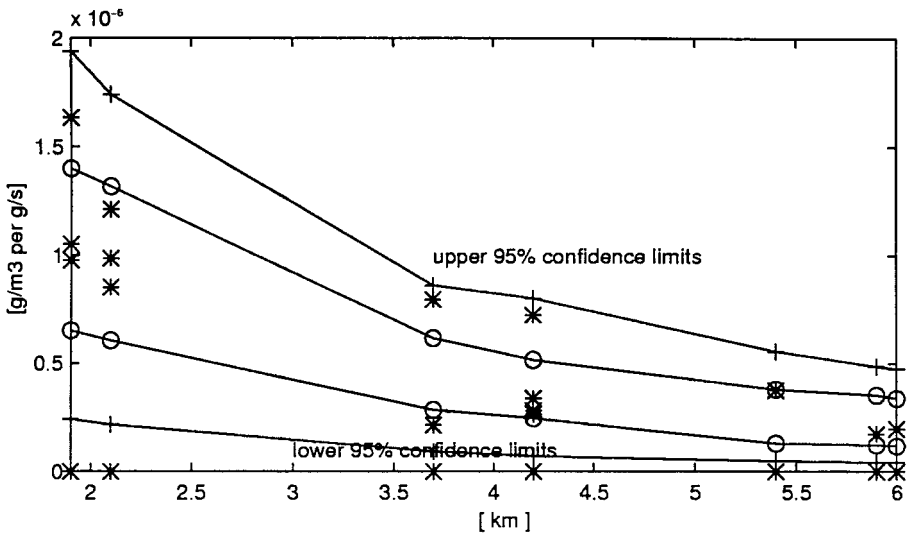

Figure 6. Ninety-five per cent prediction limits for the predictions estimated using updated posterior predictive probabilities for wide uniform wind speed range (0.1–20 m s^{-1}) and stability condition set to: (+) category C, and (○) category D; stars denote field observations.

model structure, as we are setting the rules for the model. It was decided that both the C and D categories are equally probable and the model was run for each category separately, with wind speed uniformly distributed over a very wide range of values. We treat the results as one numerical exercise and both results are combined together. An alternative would be to choose the stability condition randomly, or to allow the possibility of selecting categories other than C or D. The parameters for the simulations are given in Table IV.

Ninety-five per cent prediction limits evaluated using both model run results, but with the same combined probability of predictions for near-field distances, are shown in Figure 6. The prediction limits for category D are lower (denoted as circles) than category C prediction limits (denoted as pluses). All available observations for the Copenhagen data set are shown as stars. These results indicate that differences in model structure introduced through the different stability condition parameterisation might have a significant effect on the resulting predictive uncertainty. The resulting uncertainty of the predictions is set as the lowest and highest of the prediction limits for C and D category. All the observations except one are enclosed within these prediction limits which are also 'central' with regard to the observations. This confirms the importance of

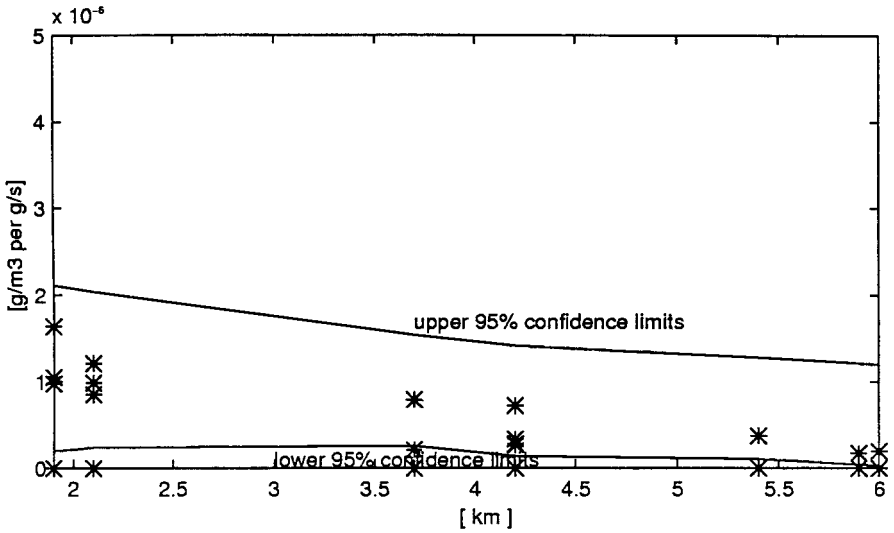

Figure 7. Ninety-five per cent prediction limits for the predictions estimated using updated posterior predictive probabilities with included errors for the observations and parameters as in Figure 5; SD = 2 m s^{-1} (Table IV).

the uncertainty in the choice of weather condition category and indicates the necessity of further work in this direction.

4.3. *Influence of output observation errors*

Application of the Bayesian Inference technique enables introduction of the observation errors into the likelihood function through the error model (9). It is equivalent to assuming that errors have some additional spread and a mean value not equal to zero. Also the correlation between the observations at different distances from the source may be introduced, as described in Romanowicz *et al.* (1994) for the correlation between neighbouring sites. However, in the case of the Copenhagen data, the correlation between neighbouring sites was neglected. These two (or three, when correlation is also taken into account) additional noise parameters are sampled within the assumed ranges, thus giving the posterior distribution sampled on extended parameter space. In the analysis presented in Section 4.2, the posterior distributions were derived in the form of marginal probabilities (summed over the noise ϕ parameter space), i.e. the probability distributions (10) depended only on model parameters and input variables). These marginal probabilities were similar to a least square estimation approach, with the difference that we were not aiming to find the best parameters but rather the best ranges for the parameter sets. However, in order to take into account the observation errors, the whole posterior distribution specified on both model parameter and noise parameter space should be used in the derivation of prediction limits. The parameters for the simulations were set as in Table III, with the exception of wind speed, which was varied according to a log-normal distribution with mean equal to 4.9 m s^{-1} (observed in the field experiment) with the standard deviation equal to 2 m s^{-1} and stability condition D. The derived prediction limits are shown in Figure 7. As expected, the prediction limits are wider and enclose all the observations (for both stability conditions D and C). The

information about the measurement errors would be useful in specifying of the lower boundaries of a combined error model variance (consisting of both structural and observation errors).

5. SUMMARY

The uncertainty analysis of an air dispersion model described in the paper consisted of an application of the Monte Carlo based sensitivity analysis and the Bayesian Inference technique which was used to condition the model predictions on field observations. Comparison of the results of the modelling with different amount of observation used (assuming different prior distributions about input variables) was performed using a comparison of model prediction limits with the observations of maximum near ground concentrations obtained from the Copenhagen experiment (Gryning, 1982). The prediction limits were derived for the output concentrations without conditioning on output observations and with conditioning on output observations using the Bayesian Inference technique. The major cases analysed were: (i) uniform parameter and input variable distributions; (ii) input variable distributions normal/log-normal following information from measurements; (iii) uniform parameter distribution and normal/log-normal input distributions; (iv) uncertainty in the choice of stability conditions; (v) influences of observation errors.

The results indicate that both sensitivity and predictive uncertainty analysis should be used simultaneously and that additionally they should be combined with analysis of the model structure. The analysis also gave recommendations on the best possible use of the available information about the model variables. It showed that: (i) assumed uniform priors give very wide prediction limits; (ii) assuming normal/log-normal priors gives more constrained prediction limits but does not enclose all the observations; (iii) introducing Bayesian conditioning on output observations narrows the prediction limits but in the case of z_0, for example, an assumed log-normal prior does not appear to adequately reflect the required effective values; uniform prior does better in this respect; (iv) the use of output measurements for the conditioning of the predictions and uncertainty in the choice of weather category gives the prediction limits most central when compared with observations which were not used for the model conditioning; (v) taking into account observation errors widens the prediction limits.

The variables analysed were roughness length, wind speed, wind direction fluctuation coefficient, release height and stability condition. Sensitivity analysis eliminated one parameter (variability in inversion height) from the chosen parameter set as not influencing model predictions for the conditions corresponding to those under which the Copenhagen data were obtained (i.e. within 6 km of a 115 m stack in slightly unstable and neutral conditions). Moreover, it indicated that both wind speed and turbulent wind fluctuation coefficient have the greatest influence on the model results.

The results of the uncertainty analysis using Bayesian Inference indicated that roughness length should be treated as a model parameter and its distribution should follow a uniform distribution over a physically feasible range. Conditioning the predictions on the release height observations significantly improves model predictions by decreasing the prediction limits. The same is true of wind speed measurements, which are the most important model variable. The influence of the turbulent wind fluctuation coefficient was not confirmed by the Bayesian uncertainty analysis. More information is required to draw conclusions for this parameter. The analysis of the influence of the wind speed distribution spread indicated that the prediction limits decrease with decrease

of spread to a certain extent. This is connected with the fact that model predictions will depart from the observations if the spread of wind variations is too narrow. An introduction of uncertainty in the choice of weather category gave the prediction limits, which were the most central with respect to the observed concentration values. This procedure required slight modification of the model structure and more detailed studies are needed to recommend the best way of representing this type of uncertainty.

Analysis of the distribution of model output without the use of available observations, or some independent expert judgement of the possible model results, cannot increase the model reliability. Comparison of model results with observations can indicate how the model structure may be modified to give an improved model performance and provides an insight on the influence of input variables and their uncertainty on prediction errors. More work should be done on weather category uncertainties, vertical and horizontal dispersion representation and representation of output observation errors in model predictive uncertainty.

ACKNOWLEDGEMENTS

This work was carried out at the Westlakes Research Institute, Whitehaven, Cumbria, supported by British Nuclear Fuels Ltd. Jonathan Tawn and Keith Beven (Lancaster University) are thanked for their help. We would also like to thank the anonymous reviewers for their valuable comments.

REFERENCES

Beven KJ, Binley A. 1992. The future of distributed hydrological modelling. *Hydrological Processes* **6**:279–298.
Box GEP, Tiao GC. 1992. *Bayesian Inference in Statistical Analysis*. Wiley: Chichester.
Clark RH. 1979. A model for short and medium range dispersion of radionuclides released into the atmosphere. NRPB report R91.
Gifford Jr FA. 1961. Use of routine meteorological observations for estimating atmospheric dispersion. *Nuclear Safety* **2**(4):47–57.
Gifford Jr FA. 1968. An outline of theories of diffusion in the lower layers of the atmosphere. In *Meteorology and Atomic Energy – 1968*, Slade DH (ed.). USAEC Report TID-24190. U.S. Atomic Energy Commission, NTIS; 66–116.
Griffith DA. 1988. *Advanced Spatial Statistics*. Kluwer: Dordrecht; 17–19.
Gryning SE. 1981. Elevated source SF_6-tracer dispersion experiments in the Copenhagen area. Riso report R446, Riso National Laboratory.
Hanna SR, Briggs GA, Hosker Jr RP. 1982. *Handbook on Atmospheric Diffusion*. National Oceanic and Atmospheric Administration: Oak Ridge, TN, DOE/TIC-11223.
Haylock R, O'Hagan A. 1997. Bayesian uncertainty analysis and radiological protection. In *Statistics for the Environment 3*, Barnett V, Turkman F (eds). Wiley: New York.
Helton JC. 1993. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. *Reliability Engineering and System Safety* **42**:327–367.
Hosker RP. 1974. Estimates of dry deposition and plume depletion over forests and grasslands. In *Proceedings of the Symposium on Physical Behaviour of Radioactive Contaminants in the Atmosphere*, Vienna, November 1973. IAEA: Vienna; 291.
Jones JA. 1983. Models to allow for the effects of coastal sites, plume rise and buildings on the dispersion of radionuclides, and guidance on the deposition velocity and washout coefficients. NRPB report – R157.
Lee RF, Irwin JS. 1995. Methodology for a comparative evaluation of two air-quality models. *International Journal of Environment and Pollution* **5**:723–733.
Moore DJ. 1976. Calculation of ground level concentration for different sampling periods and source locations. In *Atmospheric Pollution*, Sturges WT (ed.). Elsevier: Amsterdam; 5160.
Pasquill F, Smith FB. 1983. *Atmospheric Diffusion*, 3rd ed. Ellis Horwood: Chichester.
Priestly MB. 1981. *Spectral Analysis and Time Series*. Academic Press: London.
Romanowicz R, Beven K, Tawn J. 1994. Evaluation of predictive uncertainty in nonlinear hydrological models using a

Bayesian approach. In *Statistics for the Environment 2*, *Water Related Issues*. Barnett V, Turkman F (eds); Wiley: New York, 297–315.
Sutton OG. 1932. A theory of eddy diffusion in the atmosphere. *Proceedings Royal Society London A* **135**:143–165.
Woodbury AD, Ulrych TJ. 1993. Minimum relative entropy: forward probabilistic modelling. *Water Resources Research* **29**:2847–2860.
Yong NK, Jong KK, Tae WK. 1993. Risk assessment for shallow land burial of low level radioactive waste. *Waste Management* **13**:589–598.

