# INVESTIGATION OF THE SENSITIVITY OF MODELS OF POLLUTANT TRANSFER IN THE ATMOSPHERE (MODEL AND PROCEDURE)

**V.S. Komarov, S.A. Soldatenko, and S.S. Suvorov**

*Institute of Atmospheric Optics,*
*Siberian Branch of the Russian Academy of Sciences, Tomsk*
*Received December 27, 1994*

*An approach is proposed to the investigation of the sensitivity of models of conservative pollutant transfer through the atmosphere based on stochastic interpretation of the model parameters and the results of measurements. An inverse problem of pollutant transfer has been considered and a procedure for investigating the sensitivity of its solution to variations of the results of measurement of the pollutant concentration induced by a stationary local source has been described. The proposed procedure can be generalized for the case of several local sources operating in both stationary and pulsed regimes.*

Problems of investigation of the characteristics of pollutant transfer models in the atmosphere are of fundamental importance from both theoretical and applied standpoints. The urgency of a solution of such problems is primarily conditioned by a wide use of pollutant transfer models when solving direct and inverse problems of ecological monitoring.

By the direct problem is meant determining the concentration of a certain component at some fixed time with the given initial and boundary conditions for the known function of strength of a pollutant source.

The inverse problem is considered to mean determining the strength field of pollutant sources from the known concentration field, induced by these sources, with the given boundary conditions. Depending on the available *a priori* information about the pollutant sources, the inverse problem can be formulated as the problem for determining the position of emission sources or as the problem for determining the strength of these sources. In general (for the lack of *a priori* information), both characteristics must be determined.

The data obtained as a result of solution of the above problems are used to construct models of social and economic development of regions and to perform ecological expertise of projects of social development. Clearly, the quality of administrative decisions on these problems is determined in many respects by the accuracy and reliability of model results. However, despite a large body of studies devoted to the problems of construction and numerical realization of models of pollutant transfer (see, for example, Ref. 1), the consideration of the model sensitivity to variations of the initial data, that is, substantiation of the degree of reliability and accuracy of the results, remains beyond the scope of the analysis.

In Ref. 2, some estimates of the sensitivity of a solution of inverse problem to the variations of the model parameters were obtained on the basis of a deterministic approach. However, because the true values of the parameters are unknown to us, the variation of a certain model parameter cannot be determined as well. Hence, based on the above approach it is impossible to evaluate the accuracy of the problem solution. We think that we can gain a better understanding of the degree of solution sensitivity by way of stochastic interpretation of the problem, that is, based on the consideration of the model parameters, initial data, and results of problem solution as random variables.

This paper describes a procedure for investigation of the sensitivity of a solution to inverse problem of pollutant transfer in the atmosphere for a single source of pollutant operating in a stationary regime.

## FORMULATION OF THE DIRECT PROBLEM

Let in a certain rectangular region $A$, some tens of kilometers in extent along the $x$ and $y$ axes, at the point $(x_0, y_0, z_0)$, a source of pollutant be located with the strength function $f(t)$ of the form:

$$f = f(t) = \begin{cases} 0, & t < t_0 , \\ Q = \text{const}, & t \ge t_0 . \end{cases} \quad (1)$$

The concentration of pollutant $c$ at a certain time $t > t_0$ is required.

To solve this problem, we use the kinematic model of pollutant transfer proposed in Ref. 3. In this model, the main physical processes, strongly affecting the formation of the pollutant concentration field, have been taken into account. The equations of the model in the $z$–system of coordinates are written as follows:

$$\frac{\partial c}{\partial t} + \mathbf{V} \text{ grad } c - \nabla_s(\mu \nabla_s c) - \frac{\partial}{\partial z} \nu \frac{\partial}{\partial z} + \frac{1}{\rho}(I_c + I_c^*) = f , \quad (2)$$

where $c$ is the specific concentration of pollutant; $\mathbf{V} = (u, v, w)$ is the vector of the wind velocity; $\mu$ and $\nu$ are the coefficients of turbulent diffusion of pollutant in horizontal and vertical directions, respectively; $I_c$ and $I_c^*$ are the functions describing the generation and sink of pollutant in the course of transformation processes; $\nabla_s = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j}$ .

The use of the kinematic model is justified when solving the problem on the transfer of conservative pollutant, i.e., the pollutant which leaves the thermobaric structure of the atmosphere unaffected. Therefore, the coefficients of turbulent exchange and the wind velocity are the input parameters of

0235-6880/95/07 511-04 $02.00 

the model. They can be either determined from the measurement data or calculated with the use of a certain model. The pollutant will be considered produced solely by a source, i.e., $I_c = 0$.

The lower boundary of the solution domain is the underlying surface where the following formalization of interaction of the pollutant with the boundary is justified[4]:

$$\nu \left.\frac{\partial c}{\partial z}\right|_{z=0} = k_s\left(\left. c\right|_{z=0} - c_s\right) + Q_s , \tag{3}$$

where $c_s$ is the concentration of pollutant above the underlying surface, which ensures the equilibrium state of exchange process; $k_s$ is the coefficient of exchange, describing the type of soil, character of plant canopy, and temperature; $Q_s$ is the rate of pollutant flow from the underlying surface to the atmosphere.

The tropopause is chosen as an upper boundary. This choice is physically justified since the tropopause is a thick barrier layer, at the lower boundary of which the following condition can be set[5]:

$$\left.\frac{\partial c}{\partial z}\right|_{z=z_t} = k_c\left(c_{st} - \left. c\right|_{z=z_t}\right), \tag{4}$$

where $z_t$ is the altitude of the tropopause taken to be 10 km in the model; $k_c$ is the exchange coefficient; $c_{st}$ is the equilibrium concentration of pollutant in the stratosphere.

The lateral boundaries of the solution domain are open. Here, the following boundary conditions are laid down. For a region of boundaries where the air inflow occurs, it is assumed that

$$c = 0 , \tag{5}$$

and for a region of the boundary where the air outflow occurs

$$\frac{\partial c}{\partial n} = 0 , \tag{6}$$

where $\frac{\partial c}{\partial n}$ is the derivative of $c$ with respect to the outward normal to the solution domain.

The initial condition is the absence of pollutant in the atmosphere, that is,

$$\left. c\right|_{t=t_0} = 0 . \tag{7}$$

This condition is chosen for simplicity of the analysis of the pollutant transfer from the source under consideration.

Parametrization of the processes on subgrid scale described by the function $I_c^*$, namely, of self–induced raising of aerosol, its washing out, and sedimentation, is realized with the use of the formulas given in Ref. 3.

For numerical solution of equation (2), we use the separation technique for physical processes. We consider only the process of advection of pollutant at the first stage, and turbulent diffusion – at the second stage. At the third stage, the equation is solved describing transformation processes. At the step of advection, the TVD scheme[3] is used, whereas at the step where the turbulent exchange is considered we use the implicit scheme together with separation technique for spatial coordinates.

## FORMULATION OF THE INVERSE PROBLEM

Let a pollutant source be located in the rectangular region $A$ discussed above at a certain point $X_0 = (x_0, y_0, z_0)$ with a known coordinate $z_0$, with the source strength given by Eq. (1). In this case the parameters $Q$ and $t_0$ are unknown. The horizontal coordinates of the source $x_0$ and $y_0$ and its strength $Q$ are to be determined from the results of three measurements of the concentration of pollutant at some points $X_i = (x_i, y_i, z_i)$, $i = 1(1)3$. It should be noted that since the time $t_0$ at which the source starts to operate is unknown, the problem can be solved only when

$$\left.\frac{\partial c_i}{\partial t}\right|_{t_s} = 0, \quad i = 1\,(1)\,3 ,$$

where $c_i$ is the concentration of pollutant at the $i$th point. In addition it is assumed that before the start of the source operation the pollutant was absent at all the points of the region $A$, that is, $c = 0$ at $t < t_0$.

To solve this problem, we use the technique of the adjoint equations. Following Ref. 2, we obtain the model conjugate to Eq. (2):

$$\frac{\partial c^*}{\partial t} + \mathbf{V}\ \mathrm{grad}\ c^* - \nabla\mu\nabla c^* - \frac{\partial}{\partial z}\nu\frac{\partial c^*}{\partial z} + \frac{1}{\rho} I_c^* = f^* . \tag{8}$$

The boundary conditions at the lower boundary are of the following form:

$$\nu \left.\frac{\partial c^*}{\partial z}\right|_{z=0} = k_s\left(\left. c^*\right|_{z=0} - c_s\right) + Q_s , \tag{9}$$

and at the upper boundary they are

$$\left.\frac{\partial c^*}{\partial z}\right|_{z=z_t} = k_c\left(c_{st} - \left. c^*\right|_{z=z_t}\right). \tag{10}$$

In the region of lateral boundary where the air inflow occurs, they are

$$c^* = 0 , \tag{11}$$

and in the region of the lateral boundary where the air outflow from the domain of solution is observed, they are

$$\mu\frac{\partial c^*}{\partial n} + u_n c^* = 0 , \tag{12}$$

where $u_n$ is the component of the wind velocity normal to the boundary of the solution domain.

The initial condition is $c^* = 0$ at $t = t_s$.

Let everywhere, except for a certain point $X_i = (x_i, y_i, z_i)$, be $f^* = 0$. At the point $X_i$ we set $f^*$ as follows:

$$f^* = f(t) = \begin{cases} 1, & t \ge t_s , \\ 0, & t < t_s . \end{cases} \tag{13}$$

Let us integrate Eq. (8) with the indicated boundary and initial conditions until the steady state, that is, until the condition

$$\frac{\partial c_l^*}{\partial t} = 0 , \quad X \in \{A\} \qquad (14)$$

be fulfilled. In this case $c_l^* = c_l^*(x, y, z)$, where $c_l^*$ is the solution of the problem given by Eqs. (8)–(12) provided that $f^*$ is set as shown above.

The function $c^*$ can be treated as the degree of sensitivity of the equilibrium concentration of pollutant at the point $X_l$ to the operation of a point stationary source of pollutant located at a certain point of the region $A$, and the quantity $1/c_l^*(x, y, z)$ can be treated as the strength of a source located at the point $(x, y, z)$ that at the point $X_l$ induces the equilibrium concentration of pollutant being equal to unity. Then, owing to the linearity of the problem,

$$\phi_l(x, y, z) = \frac{c_l}{c_l^*(x, y, z)} \qquad (15)$$

is the strength of a point stationary source located at the point $(x, y, z)$ that at the point $X_l$ induces the equilibrium concentration of pollutant $c_l$.

Since the adjoint problem[2] has unique solution, using *a priori* information contained in the formulation of the problem we obtain that unique point $X_0$ can be found in the region $A$ for which

$$\phi_1(x_0, y_0, z_0) = \phi_2(x_0, y_0, z_0) = \phi_3(x_0, y_0, z_0) = Q. \qquad (16)$$

The coordinates $x_0$ and $y_0$ and the parameter $Q$ are the solution of the formulated inverse problem.

Since the problem described by Eqs. (8)–(12) is solved numerically with the use of the method applied in solving of the direct problem, the condition for finding solution (16) takes the form

$$(x_0, y_0) = \arg\min_{(x, y)}\{(\phi_1(x, y, z_0) - \phi_2(x, y, z_0))^2 + (\phi_1(x, y, z_0) -$$

$$- \phi_3(x, y, z_0))^2 + (\phi_2(x, y, z_0) - \phi_3(x, y, z_0))^2\} , \qquad (17)$$

$$Q = \frac{1}{3}(\phi_1^* + \phi_2^* + \phi_3^*) , \qquad (18)$$

where

$$(\phi_1^*, \phi_2^*, \phi_3^*) = \arg\min_{(\phi_1, \phi_2, \phi_3)}\{(\phi_1(x, y, z_0) - \phi_2(x, y, z_0))^2 +$$

$$+ (\phi_1(x, y, z_0) - \phi_3(x, y, z_0))^2 + (\phi_2(x, y, z_0) - \phi_3(x, y, z_0))^2\} \qquad (19)$$

and $(x, y)$ are the coordinates of grid nodes.

## PROCEDURE FOR INVESTIGATION OF THE SENSITIVITY

In deterministic approach to the investigation of the sensitivity, as a degree of sensitivity of the problem solution to variations of a certain parameter we take the value of the derivative of the solution with respect to this parameter. In this case, the derivative is treated in a classical sense when analyzing its sensitivity to the scalar parameters or in the Gato sense in the case of parameter–functions. However, as has already been noted, it is impossible to make a constructive judgement on the degree of the solution accuracy, since the model parameters and the results of measurements are random variables, whose realization is known with a certain unknown error. In this connection, it is appropriate to consider the results of solution of the problem as a certain random variables (random field) and to use, when analyzing the sensitivity of problem solution, methods of the theory of probability and mathematical statistics.

In the above approach, the dependence of the problem solution on the distribution of the corresponding parameters should be used as a measure of sensitivity. The dependence of a certain characteristic of the solution distribution on the characteristics of the parameter distribution also can be used as such a measure. We propose to use the length of an interval (the volume of the domain for vector solutions) within which the problem solution falls with a given probability for the given distribution of the affecting parameter. With this choice of the measure, it is possible to analyze the sensitivity to an individual parameter and to estimate the comparative degree of influence of different parameters on the accuracy of solution. For an analysis of the influence of errors in determining the parameters on the reliability of solution, the probability of falling of the solution within the interval of a certain length for a given parameter distribution serves as a measure of sensitivity.

If the problem is linear and only one parameter is random, the problem of investigation of the sensitivity considerably simplifies. Actually, in this case the problem solution is a linear function of the parameter and, hence, obeys the distribution of the same type as the parameter does. The distribution characteristics can be determined with the use of the known formulas of the probability theory. With the use of the above characteristics, we can easy calculate the above measures of sensitivity of the accuracy and reliability of the solution. The situation changes abruptly in the nonlinear case, in which we need the apparatus of mathematical statistics, in particular, the method of imitational simulation to perform investigations.

We illustrate the suggested procedure in an example of investigation of sensitivity of the above–formulated inverse problem of pollutant transfer in the atmosphere. The influence of errors in measuring the concentration of pollutant on the accuracy and reliability of determination of the source position and strength will be analyzed.

The first stage is the simulation of realization of solution of the inverse problem. First we simulate the realization of the results of measurements. We write down the result of measurement as follows:

$$\hat{c}_i = \bar{c}_i + \hat{\delta}_c , \qquad (20)$$

where $c_i$ is the mathematical expectation of the measurement result at the $i$th point, and $\delta_c$ is the measurement error.

The mathematical expectation of the measurement result at a certain point is the result of solution of the direct problem described above. When simulating the measurement errors, the following assumptions are introduced:

– systematic error vanishes;

– measurement errors at different points are uncorrelated;

– measurements are equally accurate;

– error is distributed by the normal law.

Since in nominal data of measuring devices only the relative error is indicated, the standard deviation of the measurement error is determined from the rule of three sigma.

For simulating the error we use the standard procedure of statistical simulation of normally distributed random variable. Having simulated the measurement errors at all three points and having added to them the concentration of pollutant derived by solving the direct problem, we obtain the realization of the vector of measurement results. Having solved the inverse problem with these initial data, we derive the realization of solution of the inverse problem. This procedure is repeated until the sample size reaches the required value (conventionally, some hundreds).

An interval variation series is constructed with the use of the above sample, and the law of distribution of solution results is estimated. For this purpose, the following operations are done. We estimate the square asymmetry and the excess of distribution of the solution results. The hypothesis on the distribution law of the solution results is set up on the basis of the asymmetry–excess diagram. This hypothesis is tested with the use of the chi–square criterion. When the hypothesis is true, the parameters of this distribution law are estimated from the sample data.

The next stage is the direct estimate of the solution sensitivity. To do this, the accuracy of the solution is specified by the value of its validity (probability of falling of the solution within a certain interval). The length of this interval is determined using the table of distribution of the solution. In doing so the interval can be selected by two ways, namely, as the interval of minimum length or as the interval located in such a way that the probabilities of solution falling to its right and to its left are the same.

When estimating the sensitivity of solution validity to the accuracy of measurement, the concentration of pollutant is specified by the length of interval within which this solution must fall, and then the probability of this falling is determined. The interval also can be chosen by two ways: as an interval maximizing the probability of solution falling within it or as an interval located so that the probabilities of solution falling to its right and to its left are the same.

The proposed procedure can be generalized for the case of several sources and can be applied to the investigation of the solution sensitivity to variations of the other parameters, for example, the wind velocity. In this case, it is necessary to simulate the random field, i.e., the wind velocity field.

## REFERENCES

1. F. Newstadt and H. van Dock, eds., *Atmospheric Turbulence and Simulation of the Pollutant Transfer* [Russian translation] (Gidrometeoizdat, Leningrad, 1985), 351 pp.

2. G.I. Marchuk, *Mathematical Simulation as Applied to Environmental Problems* (Nauka, Moscow, 1982), 319 pp.

3. S.A. Soldatenko and O.M. Sobolevskii, Atmos. Oceanic Opt. **7**, No. 2, 112–116 (1994).

4. W.S. Broecker and T.H. Peng, Tellus **26**, 21–35 (1974).

5. R.I. Newell, J. Geophys. Res. **82**, 1449–1450 (1977).

6. K. Fletcher, *Calculating Methods in Liquid Dynamics* [Russian translation] (Mir, Moscow, 1991), Vol. 2, 522 pp.