MATHEMATICAL MODELS FOR PROBLEMS OF EXPERIMENTAL DESIGN AND ENVIRONMENTAL QUALITY CONTROL

V.V. Penenko

Computing Center,
Siberian Branch of the Russian Academy of Sciences, Novosibirsk
Received January 22, 1997

Problems of organization and construction of mathematical models for climatic and ecological monitoring, forecasting, environmental quality control, and design of observations are discussed. Interconnections of the models with experimental data, variations of the model parameters, and goal functions are realized based on the variational principle.

1. INTRODUCTION

Organization of modern climatic and ecological monitoring systems intended to study natural phenomena under anthropogenic effects is impossible without an active use of the methods for mathematical modeling and joint use of models of the examined phenomena and data of field experiments. From a social viewpoint, the objective of these investigations is elucidation of premises for ecologically unfavorable and catastrophic situations in specific regions and numerical evaluation of permissible levels of anthropogenic loads based on criteria and limitations of ecological safety. Naturally, the problem on the design of observations by the given criteria for their optimization or increase of their information content also arises. Thus, a new class of problems occurs connected with the estimation of ecological prospects for specific regions under the joint effect of natural and anthropogenic factors. To solve this class of problems, we need complex mathematical models and adequate methods for their implementation together with the observational data for diagnostics, identification, and forecasting of environmental changes as well as for the refinement of methods to incorporate the feedback to control over the anthropogenic loads that affect the environmental quality into the economic and climatic regional system.

For the approach considered here models of hydrothermodynamic processes in the climatic system of industrial regions taking into account anthropogenic effects, models of transport and transformation of pollutants, and models of interaction of air masses with underlying surface elements[1–3] provide a basis for modeling technology.

The implementation of mathematical models for monitoring and ecological forecasting includes their optimization.[4] For this reason, basic models can be conveniently represented in variational form with the use of the integral identity that considers in the main functional the model description in the form of a system of differential equations, boundary and initial conditions, external impacts, and input parameters. The functional of the integral identity is specified by the energy balance equation and other balance relations for examined processes. By way of example, we describe here only a part of the model complex, namely, the model of transport and transformation of pollutants in the atmosphere; the variational model of the atmospheric hydrothermodynamics was described in Ref. 5.

2. MODEL OF TRANSPORT AND TRANSFORMATION OF POLLUTANTS

In accordance with purposes of our investigations, we use the dual description of the model in differential and variational forms.

1) Basic model equations in the differential form can be written as follows:

$$\frac{\partial c_i}{\partial t} + \operatorname{div} c_i \mathbf{u} - \operatorname{div}_s \mu \operatorname{grad}_s c_i - \frac{\partial}{\partial \sigma} \nu \frac{\partial c_i}{\partial \sigma} + (B\mathbf{c})_i = f_i(\mathbf{x}, t),$$
$$i = \overline{1, n}\,. \tag{1}$$

2) Integral identity for model transport can be written as

$$I_c(\varphi, \mathbf{Y}, \varphi^*) = \sum_{i=1}^{n} \left\{ \int_{D_t} \left\{ \frac{1}{2}\left(\frac{\partial c_i}{\partial t} c_i^* - \frac{\partial c_i^*}{\partial t} c_i \right) + (\Lambda c_i, c_i^*) + \right.\right.$$
$$\left. + \mu \operatorname{grad}_s c_i \operatorname{grad}_s c_i^* + \nu \frac{\partial c_i}{\partial \sigma} \frac{\partial c_i^*}{\partial \sigma} + (B\mathbf{c})_i c_i^* - (f_i\, c_i^*) \right\} \mathrm{d}D\mathrm{d}t -$$
$$- \int_{\Omega_t} \mu \frac{\partial c_i}{\partial n} c_i^* \,\mathrm{d}\Omega\mathrm{d}t - \int_{S_t} \left(\nu \frac{\partial c_i}{\partial \sigma} c_i^* \right) \Big|_0^1 \mathrm{d}S\, \mathrm{d}t +$$
$$\left. + \frac{1}{2} \int_D c_i\, c_i^* \,\Big|_0^t \mathrm{d}D \right\} = 0\,. \tag{2}$$

0235-6880/97/06 354-05 $02.00

Here $\varphi \equiv \mathbf{c} = \{c_i, i = \overline{1, n}\}$ is the vector function with components specifying the values of the concentration of pollutants; n is the number of pollutants; $\varphi^* \equiv \mathbf{c}^* = \{c_i^*, i = \overline{1,n}\}$ is the vector function with arbitrary sufficiently smooth components; $\mathbf{f} = \{f_i, i = \overline{1,n}\}$ is the source function; $(B\mathbf{c})$ is the operator of transformation of the pollutants; $\Lambda\varphi$ is the operator of transport of the substance φ in the air flow having the velocity $\mathbf{u} = (u, v, w)$; μ and ν are the horizontal and vertical turbulent exchange coefficients, respectively; the subscript s denotes the operators on horizontal variables; $\mathbf{Y} = (\mu, \nu, \mathbf{u}, \mathbf{c}^0, \mathbf{f})$ is the vector of the model input parameters; $\mathbf{c}^0$ defines the initial state; $D_t = D\times[0, \bar{t}]$, where $[0, \bar{t}]$ is the time interval; D is the domain of spatial variables $\mathbf{x} = (x, y, \sigma)$; σ is the vertical coordinate that keeps track of the Earth's relief; x and y are horizontal coordinates; Ω_t specifies the side boundary; S_t specifies the lower boundary of the domain D_t. The coefficients of the velocity vector $\mathbf{u}$ are related by the continuity equation being a part of the atmospheric hydrothermodynamics model. The rate of gravitational sedimentation of pollutants is also considered in the vertical component of the velocity w. The form of the functional in identity (2) is defined by the energy balance equation of the model. The integrals in Eq. (2) along Ω_t and S_t and over the domain D at the moment $t = 0$ are closed with the use of the corresponding boundary and initial conditions

a) $c_i(\mathbf{x}, 0) = c_i^0(\mathbf{x})$ at $t = 0$,

b) $R_{1i}(\mathbf{c}) = f_{1i}(x, y, t)$ for $\sigma = 1$ $(z = Z_s(x, y))$,

c) $R_{2i}(\mathbf{c}) = f_{2i}(x, y, t)$ for $\sigma = 1$.

On the side boundary Ω the concentration of pollutants is assumed to take its background value. Here, $Z_s(x, y)$ is the function that specifies the surface relief, R_{1i} and R_{2i} are the preset operators; f_{1i} and f_{2i} are sources and sinks of pollutants. The operators $R_{1i}(\mathbf{c})$ describe the interaction of different substances with each other and with the Earth's surface including exchange processes of the air with water, soil, vegetation, etc. and the operators $R_{2i}(\mathbf{c})$ describe the behavior of pollutants on the upper boundary. Forms of the operators $B_i(\mathbf{c})$, R_{1i} and $R_{2i}(\mathbf{c})$ as well as of the functions f_{1i} and f_{2i} are specially assigned for each specific problem and therefore can be considered as generalized input parameters of the model. Boundary and initial conditions for the function c_i^* are the consequences of variational model formulation and of corresponding optimization problems.

It is assumed that the functions φ and φ^* belong to the corresponding functional space and the vector $\mathbf{Y}$ belongs to a set of permissible values of the parameters, i.e.,

$$\varphi \in Q(D_t)\ , \quad \varphi^* \in Q^*(D_t)\ , \quad \mathbf{Y} \in R(D_t)\ . \qquad (3)$$

3. CONTROL AND DESIGN FUNCTIONALS

To formulate the problems of monitoring, forecasting, design and construction of the algorithms for solving these problems, we introduce a set of functionals of the form

$$\Phi_k(\varphi) = \int_{D_t} F_k(\varphi)\, \chi_k(\mathbf{x}, t)\, \mathrm{d}D\, \mathrm{d}t\ , \quad k = \overline{0, K}\ , \qquad (4)$$

where $F_k(\varphi)$ are prescribed functions of φ; $\chi_k(\mathbf{x}, t)$ are the nonnegative weighting functions defined in D_t by the conditions or design of the experimental observations, conditions of estimation of the state either in the domain D_t or units subdomains or at discrete set of points $D_t^m \subset D_t$ containing at least one point; and, $\chi_k(\mathbf{x}, t)\mathrm{d}D\mathrm{d}t$ are the corresponding Radon or Dirac measures in the domain D_t. From this set of functionals, we select the functionals of four types that differ in their structure and purpose: 1) functionals of general estimation of the system behavior; 2) quality functionals that characterize deviations of the measured and calculated parameters; 3) restricting functionals of the state functions; 4) goal functionals for control of the anthropogenic loads.

In this case, we consider restrictions that follow from the conditions of ecological safety and stability and the goal functionals consider in addition social and economic factors like the cost of environmental losses or the cost of salvaging our environment and reducing these losses.

Restrictions on the state functions usually have the form of inequalities. They may be global or local with respect to the domain of definition of the state function and independent variables. Because numerical models have the large number of the internal degrees of freedom, local restrictions in the form of inequalities are inconvenient to consider in iterative optimization algorithms. Therefore, we replace them by equivalent global limitations in the form of inequalities

$$\Phi_\alpha(\varphi) = 0\ , \quad \alpha \in \{k = \overline{0, K}\}\ . \qquad (5)$$

The functions F and that define functionals (3) will be specially represented so that to eliminate violations of local restrictions with the help of inequalities (3).

Let the local restrictions imposed on the state function φ be really specified in the form of inequalities

$$\Psi_\alpha(\varphi, \mathbf{x}, t) \le 0\ , \quad \varphi \in Q(D_t)\ , \quad (\mathbf{x}, t) \in D_t\ , \qquad (6)$$

where Ψ_α are the prescribed functions differentiable with respect to φ. To write down integral restrictions (5) equivalent to local restrictions (6), it will suffice to take F_α in the definition of functional (4) in the form

$$F_\alpha(\varphi) = |\Psi_\alpha + |\Psi_\alpha||. \quad (7)$$

The observational functionals are defined as follows.

Let us suppose that observations are carried out in a discrete set of points $D_t^m \subset D_t$ and denote by $\eta^m = \{\eta_k^m, \ k = \overline{1, k_0}\}$ a set of the observable parameters and by $\mathbf{H}(\varphi) = \{H_k(\varphi), \ k = \overline{1, k_0}\}$ the corresponding models of observations, where k_0 is the number of observations and $H_k(\varphi)$ is the functional description of transformation of the state function into the observable parameters. By $[H_k(\varphi)]_m$ we denote the calculated transformations of the observable parameters η_m at the points of the set D_t^m. Taking into account the above designations, we define the discrepancy functionals for the measured and calculated values of the parameters in the form

$$\Phi(\varphi) = \int_{D_t} \sum_k [(\eta_k^m - [H_k(\varphi)])_m^2 \, \chi_k^m(\mathbf{x}, t)] \, \mathrm{d}D \, \mathrm{d}t , \quad (8)$$

where $\chi_m^k \mathrm{d}D\mathrm{d}t$ are the Dirac measures concentrated at the points of the set D_t^m. To acquire observational data and to carry out diagnostic investigations, we will take into account all observations in functionals (8). To solve the problems of experimental design and diagnostic investigations, functionals for individual observations should be used in addition to the functional that considers all observations. Positions of the observational points refer to a set of the input parameters of the data acquirement and experimental design model.

4. PROBLEMS OF CONTROL OVER THE SOURCES

Let us consider the formulation of problems of control over sources of pollutants. In this case, the source functions in Eq. (1) are represented in the form

$$f_i(\mathbf{x}, t) = \sum_{k=1}^{M_i} (1 - e_{ki}) \, q_{ki}(t) \, \omega_{ki}(\mathbf{x}) , \quad i = \overline{0, n} , \quad (9)$$

where q_{ki} are the functions specifying the power of the sources, $\omega_{ki}(\mathbf{x})$ is the function specifying their positions in the domain D, M_i is the number of sources of the ith pollutant, $\mathbf{e} = \{e_{ki}, \ k = \overline{1, M_i} , \ i = \overline{1, n}\}$ are the relative power adjustment coefficients of the sources. It is assumed that the adjustable parameters satisfy the following conditions:

$$0 \le e_{ki} \le E_{ki} \le 1 , \quad k = \overline{1, M_i} , \ i = \overline{0, n} , \quad (10)$$

where E_{ki} are the preset maximum values. Here, $e_{ki} = 0$ corresponds to the initial state of the sources and $E_{ki} = 1$ means that the corresponding source may be completely switched off in some situations.

The design problem is to determine the permissible levels of anthropogenic loads by the given goal criterion on conditions that the state functions are related with the parameters and sources via mathematical model (1) in discrete form and satisfy the given set of restrictions.[4] Its solution is reduced algorithmically to the estimation of the components of the parameter vector $\mathbf{e}$ for the optimum goal criterion with limitations. The numerical model equations and the functionals of ecological and climatic restriction are used as these limitations.

Because of nonlinearity of the models and functionals, the problem is solved by the gradient iterative methods. Algorithms describing the sensitivity relations and algorithms for calculating the sensitivity functions themselves for each functional from the set of equations (4), (5), and (8) hold a central position here.

Their definitions and calculation formulas have the forms

1) for the sensitivity relations

$$\delta\Phi_i^h(\varphi) = (\mathrm{grad}_{\mathbf{Y}} \Phi_i^h, \delta\mathbf{Y}) \equiv$$

$$\equiv \frac{\partial}{\partial \xi} I^h(\varphi, \mathbf{Y} + \xi \delta\mathbf{Y}, \varphi_i^*) \Big|_{\xi = 0} , \quad i = \overline{0, K} ; \quad (11)$$

2) for the sensitivity functions

$$\mathrm{grad}_{\mathbf{Y}} \Phi_i^h \equiv \frac{\partial \Phi_i^h(\varphi)}{\partial \mathbf{Y}} \equiv$$

$$\equiv \frac{\partial}{\partial \delta\mathbf{Y}} \frac{\partial}{\partial \xi} I^h(\varphi, \mathbf{Y} + \xi \, \delta\mathbf{Y}, \varphi_i^*) \Big|_{\xi = 0} , \quad i = \overline{0, K} . \quad (12)$$

Here and below the superscript h denotes a discrete analog of the corresponding object; ξ is the real parameter; δ denotes variation of the corresponding object, for example, $\delta\mathbf{Y}$ is the vector of variations of the parameter $\mathbf{Y}$. The conditions $\mathbf{Y}, \mathbf{Y} + \xi\delta\mathbf{Y} \in R^h(D_t^h)$, are assumed to be valid, where D_t^h is the grid in D_t. In Eqs. (11) and (12), φ and $\varphi_i^* (i = \overline{0, K})$ are solutions of the basic and conjugate problems for unperturbed values of the parameter $\mathbf{Y}$. These problems are formulated for discrete analogs of the functionals

$$\tilde{\Phi}_i^h(\varphi, \mathbf{Y}, \varphi_i^*) \equiv \Phi_i^h(\varphi) + I^h(\varphi, \mathbf{Y}, \varphi_i^*), \quad i = \overline{0, K} . \quad (13)$$

In Eq. (13), in contrast with Eq. (2), the function φ^* acquires a more specific meaning; namely, here it plays the role of the Lagrangian distributed multiplier to consider the discrete equations of the numerical model as restrictions under examination of the ith functional given by Eq. (4) and therefore it is labeled by the subscript i in Eqs. (11) – (13). In this case, the basic problem is formulated based on stationarity conditions for functionals (13) attendant to arbitrary and independent variations of the components of function φ_i^* at the nodes of the grid D_t^h and the conjugate problems are formulated based on stationarity conditions for the same functionals attendant to variations of the components of the state function φ at the nodes of the grid D_t^h. The number of conjugate problems to be solved is equal to the number of functionals in Eq. (4). The functions $\partial\Phi_i/\partial\varphi$ are selected as sources for the conjugate problems. From here a demand arose for differentiability of functionals (4), (5), and (8) with respect to the state functions. Various applications of the conjugate problems for investigation of environmental systems were discussed in Refs. 1, 2, 4, and 6.

Solutions to the basic problem and conjugate problems allow us to eliminate the internal degrees of freedom of the numerical model itself from sensitivity relations (11) and thereby to relate directly the variations of the functionals with the variations of the parameters with the help of the sensitivity functions. Analysis of the sensitivity functions allows us to reveal the regions with the enhanced sensitivity of the examined functionals to the parameter variations. As already noted, the functions of the pollutant sources refer to the generalized model parameters. Thus, the sensitivity functions for the parameter variations $\delta\mathbf{f}$ yield the information required for regionalization of the territory D by the degree of sensitivity of the prescribed functionals to variations of sources of anthropogenic impacts and for incorporation of feedback to control over the sources.

When the functions $q_{ki}(t)$ and $\omega_{ki}(x)$ in representation (9) are fixed and only adjustable parameters $\mathbf{e}$ are varied, the sensitivity relations provide the way for determining the values of these parameters on the basis of the iterative gradient algorithm. In each iteration step, corrections for these parameters are calculated from the values of the sensitivity functions of the goal functional to be optimized and the restrictions are taken into account by the method of projecting the gradient of the goal functional in the direction of the desired parameter vector onto a linear manifold created by the sensitivity relations for restriction functionals (5)–(7).

The projection operator is formed with the help of the matrix of this manifold. Its elements consist of the values of the sensitivity functions attendant to variations of the desired parameters.

The source parameters so obtained provide numerical estimates of the permissible levels of the anthropogenic loads that keep the state of the climatic system within ecological safety standards. If the functions of geographical position of the sources $\omega_{ki}(\mathbf{x})$ entering into Eq. (9) also should be refined, the sensitivity functions attendant to variations $\delta\omega_{ki}(\mathbf{x})$ also should be used to estimate the source parameters.

For convenience of presentation, we consider here the algorithmically open version of the model pollutant transport with respect to the hydrothermodynamic models assuming that the atmospheric state is known and is described by the velocity field $\mathbf{u}$ and the turbulent exchange coefficients μ and ν that form the set of the input model parameters. Under real conditions, the model should be considered in combination and should take into account direct and inverse relations between the hydrothermodynamic processes and the atmospheric pollution because the atmospheric behavior and conditions of formation of mesoclimates determine in many respects the permissible levels of anthropogenic loads.[2,7] In essence, the problem is to find the conditions of climatic and ecological stability given that natural and anthropogenic factors interact. Therefore, it is very important to examine the degree of sensitivity of the goal functionals and restricting functionals to the entire set of the input parameters and external impacts in an analysis of the results of modeling.

5. DATA ACQUIREMENT AND EXPERIMENTAL DESIGN

Problems on acquirement of observational data, model diagnostics, and experimental design are solved with the joint use of the models and observational data.[8] To implement the algorithm for establishing the direct and inverse relations between observations and models, we should suggest that at least one element of the modeling procedure, that is, either the model itself, or its parameters, or initial and input data contain errors. Errors also can be present in the observational data.

To construct algorithms for solving the problems of data acquirement and model diagnostics, we define the quality functional

$$J^h(\varphi, \mathbf{r}, \mathbf{Y}, \varphi^0) = I^h(\varphi, \mathbf{Y}, \varphi^*) + \{\Phi_0(\varphi) + (\mathbf{r}^{\mathrm{T}} M_0\, \mathbf{r}) +$$

$$+ (\varphi^0 - \varphi_a^0)^{\mathrm{T}} M_1(\varphi^0 - \varphi_a^0) + (\mathbf{Y} - \mathbf{Y}_a)\, \Gamma(\mathbf{Y} - \mathbf{Y}_a)\}^h \,. \quad (14)$$

Here $\Phi_0(\varphi)$ is the observational functional; $\mathbf{r}$ is the error function of the model; φ_a^0 and $\mathbf{Y}_a$ are *a priori* estimates of the initial state φ^0 and of the parameters $\mathbf{Y}$; M_0, M_1, and Γ are the prescribed weighting matrices; the superscript T denotes transposition.

The algorithms are constructed based on the condition of stationarity of functional (14) with respect to variations of the components of functions φ^*, φ, $\mathbf{r}$, $\mathbf{Y}$, and φ^0 at the nodes of the grid D_t^h. They incorporate procedures for solving the basic problem and the conjugate problems and formulas for calculating functions $\mathbf{r}$, φ^0, and $\mathbf{Y}$ in terms of the corresponding sensitivity functions. These algorithms are based on the gradient iterative methods.

Now we consider experimental design that is primary determined by the purposes and chosen criteria for estimating the information content of the observations. Therefore, there are many practical approaches to a solution of this problem.[9] Taking into account distinguishing features of climatic and ecological monitoring and the large number of internal and external degrees of freedom for models of examined processes, it is convenient to use the experimental design procedure based on the sensitivity functions separating out the regions in which these functions attain their maxima and to place observation points in these regions. This procedure is efficient when observations are aimed at estimating the functionals or the model parameters. Position of the observation points is the input parameter of the design problems. By virtue of nonlinearity of the models, we will use the tactics of sequential design correcting the distribution of observers taking into account the sensitivity function of functionals to be estimated to variations of the desired model parameters including the position of the observers.

In this case, the motion of the observation points in the domain D_t may be parameterized in the form

$$\mathbf{x}_{k\tau} = \mathbf{x}_k + \tau\mathbf{V}(\mathbf{x}_k), \quad \mathbf{x}_{k\tau}, \mathbf{x}_k \in D_t , \tag{15}$$

where $\mathbf{x}_k$ and $\mathbf{x}_{k\tau}$ describe the position of the starting and designed points of observations, $\mathbf{V}(\mathbf{x}_k)$ is the velocity in the design space, and τ is the parameter. The function $\mathbf{V}(\mathbf{x}_k)$ is calculated in terms of the sensitivity functions of the goal functional to the variations of the position of observational points and the parameter τ is estimated from the condition of optimizing the goal functional.

6. CONCLUSION

The algorithms for optimizing the problems of monitoring and experimental design can be used to calculate various sensitivity functions and to incorporate the feedback for control of the anthropogenic factors and position of the observational points. This is the principal difference of the procedure for modeling of the problems of monitoring of the ecology of stable development from the convenient methods of direct modeling. With such design techniques, the researcher ceases to be simply the observer and the recorder of environmental changes and has a good chance to seek ways to preserve climatic and ecological stability in the design and realization of economic activity. The work was supported in part by the Russian Foundation for Basic Researches (project No. 94–05–16105).

REFERENCES

1. G.I. Marchuk, *Mathematical Modeling as Applied to the Environmental Problem* (Nauka, Moscow, 1982), 314 pp.
2. V.V. Penenko and A.E. Aloyan, *Models and Methods for Problems of Environmental Protection* (Nauka, Novosibirsk, 1985), 256 pp.
3. V.V. Penenko and G.I. Skubnevskaya, Usp. Khimii **59**, No. 11, 1757–1776 (1990).
4. V.V. Penenko, Obozr. Prikladn. Prom. Matem. **1**, No. 6, 917–941 (1994).
5. V.V. Penenko and E.A. Tsvetova, Bull. NCC, Series Num. Model. Atmosph. Ocean and Environmental Studies, No. 2, 53–74 (1995).
6. G.I. Marchuk, *Adjoint Equations and Analysis of Complex Systems* (Nauka, Moscow, 1992), 326 pp.
7. V.V. Penenko and A.E. Aloyan, Izv. RAN, Fiz. Atmos. Okeana **31**, No. 3, 372–384 (1995).
8. V.V. Penenko, Bull. NCC, Series Num. Model Atmosph. Ocean and Environmental Studies, No. 4, 32–51 (1996).
9. S.M. Ermakov and A.A. Zhiglyavskii, *Mathematical Theory of Optimal Experiment* (Nauka, Moscow, 1987), 320 pp.