Journal of Applied Mechanics and Technical Physics, Vol. 48, No. 3, pp. 428–436, 2007

MATHEMATICAL MODELS OF ENVIRONMENTAL FORECASTING

V. V. Penenko and E. A. Tsvetova

UDC 551.51 + 519.6

A technique of environmental forecasting with allowance for climate-related factors is proposed. For this purpose, a set of subspaces ranked in terms of disturbance scales is chosen by means of orthogonal decomposition from multidimensional multicomponent databases containing information on the functions of state that describe atmospheric processes for a long period. The leading part of subspaces that take into account climate-scale processes composes an informative basis for hydrodynamic background formation in calculating the forecast scenarios of changes in atmosphere quality. Calculated scenarios with estimates of atmosphere pollution in the Far East of Russia and at the adjacent territories of China and Korea are presented.

Key words: *long-term forecasting, environment pollution, atmosphere hydrodynamics, climate, mathematical modeling, orthogonal decomposition of multidimensional fields.*

Introduction. Environmental forecasting has certain specific features because one should take into account a wide range of interrelated processes (physical, chemical, and biological) of various space and time scales in formulating the problems to be solved. Long-term forecasting problems are even more difficult because, in addition to insufficient current knowledge of processes, external and internal sources of disturbances, and the current state of the system, there is an uncertainty in estimating the future behavior of the climate system as a background for pollution processes. Therefore, the development of a method for including climate information into the technology of long-term environmental forecasting is an important aspect of improving forecast reliability.

Environmental research normally involves systems consisting of models and data determined in multidimensional phase spaces of state variables and parameters. These systems have a large number of degrees of freedom. A straightforward analysis of evolution of the processes involves significant difficulties. In view of these circumstances, it is reasonable to supplement the system by a certain set of generalized characteristics of the state variables and to develop a technique for qualitative and quantitative analysis of these characteristics. For this purpose, we use variational methods, methods of sensitivity theory generated by the variational methods, and methods of orthogonal decomposition of multicomponent multidimensional fields of observation data and modeling results. Such a concept takes into account that environmental prospects should be estimated within time periods whose scales are commensurable with the lifetime of currently acting and designed objects that are sources of man-caused effects on the biosphere.

In the present paper, we consider one aspect of the concept of solving interrelated problems of ecology and climate, which is developed at the Institute of Computational Mathematics and Mathematical Geophysics of the Siberian Division of the Russian Academy of Sciences [1–5]. In particular, an approach to analyzing and using climate information in environmental forecasting related to atmosphere quality estimating is described.

1. Mathematical Models and Functionals for Environmental Forecasting. Climate and environmental research is based on models of atmosphere dynamics in combination with models of transport and transformation of heat, moisture, and optically and chemically active substances in the atmosphere [5]. The structure of these models can be presented as

---

Institute of Computational Mathematics and Mathematical Geophysics, Siberian Division, Russian Academy of Sciences, Novosibirsk 630090; penenko@sscc.ru. Translated from Prikladnaya Mekhanika i Tekhnicheskaya Fizika, Vol. 48, No. 3, pp. 152–163, May–June, 2007. Original article submitted November 10, 2006.

$$L(\boldsymbol{\varphi}) \equiv B\,\frac{\partial \boldsymbol{\varphi}}{\partial t} + G(\boldsymbol{\varphi}, \boldsymbol{Y}) - \boldsymbol{f} - \boldsymbol{r} = 0. \tag{1}$$

The initial conditions (at $t = 0$) and parameters of the model are formulated as follows:

$$\boldsymbol{\varphi}^0 = \boldsymbol{\varphi}_a^0 + \boldsymbol{\xi}(\boldsymbol{x}), \qquad \boldsymbol{Y} = \boldsymbol{Y}_a + \boldsymbol{\zeta}(\boldsymbol{x}, t). \tag{2}$$

Here $\boldsymbol{\varphi} \in Q(D_t)$ is the real space of the vector state functions, B is the block diagonal matrix, $G(\boldsymbol{\varphi}, \boldsymbol{Y})$ is the nonlinear matrix differential operator whose basic elements are advection-diffusion operators acting on various components of the state functions, $\boldsymbol{f}$ are the source functions, $\boldsymbol{r}$, $\boldsymbol{\xi}$, and $\boldsymbol{\zeta}$ are the functions that describe the errors and uncertainties of the models, initial data, and parameters, $\boldsymbol{Y}$ is the vector of parameters of the model, which belongs to the range of admissible values of $R(D_t)$, $\boldsymbol{\varphi}^0$ is the initial state, $\boldsymbol{\varphi}_a^0$ and $\boldsymbol{Y}_a$ are the *a priori* estimates of the corresponding objects, $D_t = D \times [0, \bar{t}\,]$, D is the range of variation of the spatial coordinates $\boldsymbol{x}$, and $t \in [0, \bar{t}\,]$ is the time interval. The domain D may be the global domain on the spherical Earth or a bounded region of the global system. The boundary conditions for closing the models, which are set depending on problem formulation and are included into the definition of the class of functions $Q(D_t)$, are written as

$$R_b(\boldsymbol{\varphi})_i - f_{bi} = 0, \qquad (\boldsymbol{x}, t) \in \Omega_t, \tag{3}$$

where $R_b(\boldsymbol{\varphi})_i$ are the operators of the boundary conditions and f_{bi} are the source functions for the components of the state vector numbered i at the boundaries Ω_t of the domain D_t.

System (1) contains models of atmosphere dynamics, designed for forming the hydrodynamic background of the carrier medium [5]. Moreover, reproduction and prediction of ecological situations involve models that describe transport and transformation of various substances

$$L\boldsymbol{\varphi} \equiv \frac{\partial \pi \varphi_i}{\partial t} + \operatorname{div} \pi(\varphi_i \boldsymbol{u} - \mu_i \operatorname{grad} \varphi_i) + \pi(S\boldsymbol{\varphi})_i - \pi(f_i(\boldsymbol{x}, t) + r_i) = 0, \tag{4}$$

where $\boldsymbol{\varphi} = \{\varphi_i(\boldsymbol{x}, t),\, i = \overline{1, ns}\} \in Q(D_t)$ is part of the state vector function from (1) and ns is the number of various components. The components φ_i are the temperature, water–air mixing ratios for characteristics of humidity in the atmosphere (water vapor, cloud water, rain water, snow, and ice crystals), and concentrations of polluting admixtures in gaseous and aerosol states; $\boldsymbol{f} = \{f_i(\boldsymbol{x}, t),\, i = \overline{1, ns}\}$ are the source functions of heat, moisture, and admixtures; $\boldsymbol{r} = \{r_i(\boldsymbol{x}, t),\, i = \overline{1, ns}\}$ are the functions that describe uncertainties and errors of the models; $(S\boldsymbol{\varphi})_i$ are the nonlinear matrix operators that describe the local processes of transformation of the corresponding substances and do not contain derivatives of the state functions with respect to $\boldsymbol{x}$ and t. The following characteristics of the hydrodynamic background participate explicitly in Eqs. (4): the velocity vector $\boldsymbol{u} = (u_1, u_2, u_3)$, the coefficients of turbulent exchange $\mu_i = (\mu_1, \mu_2, \mu_3)_i$ for the substance φ_i in the coordinate directions $\boldsymbol{x} = \{x_i\}$ $(i = \overline{1, 3})$, and the function π determined via the meteoelements in accordance with the structure of the vertical coordinate in the domain D_t. The structures of models of the processes, domains, and coordinate systems used in (1)–(4) are described in detail in [5].

The basic approaches to constructing mathematical models for studying the atmosphere quality and methods for implementation of these models on the basis of advanced computational technologies have been formulated [1–4]. The quality of predictions provided by these models, however, is not always satisfactory for the present-day society. The success of environmental forecasting depends to a large extent on the quality of predicting hydrometeorological processes. The contemporary models offer a reliable forecast for a period of approximately 10 days, while the question of long-term forecasting is still open. We believe that the main reserve of new ideas and methods of long-term environmental forecasting and, especially, design should be sought by solving the problem of predictability. The classical methods of studying problems similar to (1)–(4) can answer the questions on the existence and properties of their solutions, e.g., uniqueness, correctness, degree of smoothness, summability, etc. The essence of the predictability problem, which is a fundamental part of the prediction theory, is elucidation of the issues of principal importance: Does the mathematical model and the description of the processes based on this model adequately reflect the observed behavior of a real physical or natural system and how the influence of uncertainties can be mitigated?

To find a compromise (from the viewpoint of predictability) under conditions of uncertainties in input data and model parameters, it seems reasonable to use a scenario approach. To form a set of prognostic scenarios, we add new elements to problem (1)–(4): guiding phase spaces (GPSs). The physical meaning of GPSs can be defined as a certain characteristic of the background state of the carrier medium with a behavior obtained as a solution of a corresponding mathematical model being dynamically adjusted in accordance with a certain criterion. These subspaces are calculated on the basis of the entire available *a priori* actual information and monitoring data.

In the present work, we consider a forecasting method based on information about the evolution of the climate-ecological system of the Earth for a period greater than 30 years. By the definition of the World Meteorological Organization, 30 years are a classical period for estimating the basic climate parameters [6].

With allowance for the reasoning given above, we construct the guiding phase spaces participating in formation of the hydrodynamic background by the method of orthogonal decomposition of multidimensional multicomponent databases that describe the behavior of the global climate system or its parts during long periods of time. The decomposition is performed by consecutively ranking the basis subspaces with decreasing characteristic scales of perturbations described by these subspaces (e.g., climatically significant processes and weather noise). In such a decomposition, the guiding subspaces can be defined as a sum of two structural elements:

$$\boldsymbol{\varphi}_d(\boldsymbol{x},t) = \boldsymbol{\varphi}_d^0(\boldsymbol{x},t) + \boldsymbol{\varphi}_d^1(\boldsymbol{x},t), \qquad (\boldsymbol{x},t) \in D_t^{hd}. \tag{5}$$

Here $\boldsymbol{\varphi}_d^0(\boldsymbol{x},t)$ is the large-scale part expressed as a linear combination of the leading basis subspaces within the framework of orthogonal decomposition, $\boldsymbol{\varphi}_d^1(\boldsymbol{x},t)$ are the subspaces constructed on the basis components of smaller scales, and D_t^{hd} is a defined discrete set of points within the domain D_t. The component $\boldsymbol{\varphi}_d^1(\boldsymbol{x},t)$ may be deterministic-stochastic within the range of variability of the corresponding parameters in the database under analysis.

As one of the objectives of the environmental research and forecasting is finding relations between meteorological parameters of the climate system and ranges of environmental risks and vulnerability for particular receptor areas, quantification of the forecasting quality involves a special set of functionals $\Phi_k(\boldsymbol{\varphi})$ ($k = 1, \ldots, Kc$; $Kc \geqslant 1$) determined on a set of state functions. These functionals are generalized characteristics of the behavior of the climate-ecological system, depending on variations of parameters and external sources:

$$\Phi_k(\boldsymbol{\varphi}) = \int\limits_{D_t} F_k(\boldsymbol{\varphi})\chi_k(\boldsymbol{x},t)\,dD\,dt \equiv (F_k, \chi_k), \qquad \chi_k \subset Q^*(D_t), \quad k = \overline{1, Kc}. \tag{6}$$

Here $F_k(\boldsymbol{\varphi})$ are the estimated functions of a given form, which are bounded and differentiable with respect to $\boldsymbol{\varphi} \in Q(D_t)$; $\chi_k(\boldsymbol{x},t) \geqslant 0$ and $\chi_k(\boldsymbol{x},t)\,dD\,dt$ are the nonnegative weight functions and the corresponding Radon measures [if the values of the functions $F_k(\boldsymbol{\varphi})$ are distributed in space] or Dirac measures [if $F_k(\boldsymbol{\varphi})$ is determined on a set of discrete points in the domain D_t] [7]; $Q^*(D_t)$ is the subspace of adjoint functions. The carriers of the weight functions, i.e., the domains of their nonzero values, can be determined as receptor areas in D_t whose configurations are defined as the input parameters in structures (6).

The objective functionals for complexing prognostic solutions from the state functions and guiding phase spaces (5) calculated by the model are defined by analogy with functionals for data assimilation in process models. These are functionals that express the measure of deviation of the sought state function from the vectors $\boldsymbol{\varphi}_d$ under conditions of minimization of the total measure of uncertainties in process models and input data for calculating modeling scenarios. For instance, the objective functional for joining the state functions calculated by the numerical model of processes and corresponding GPS elements (5) in the data-assimilation mode can be written in the form (6). For this purpose, the function $F_k(\boldsymbol{\varphi})$ at the points of the domain D_t is determined as an energy scalar product in the space of various components of the state function

$$F_k(\boldsymbol{\varphi}) = ((\boldsymbol{\varphi} - \boldsymbol{\varphi}_d), C_1(\boldsymbol{\varphi} - \boldsymbol{\varphi}_d))\Big|_{(\boldsymbol{x},t)}, \tag{7}$$

where C_1 is the diagonal matrix with positive elements whose values are found from the expression for the energy of the system with allowance for physical dimensions and the content of the components of the state functions. The functionals for assimilation of measured data are determined in a similar manner. In this case, the functions $F_k(\boldsymbol{\varphi})$ are also defined by energy-type functionals for estimates of residues between the measured values and the calculated images of these quantities.

The required estimates of the functionals and their variations are obtained with the use of sensitivity relations for functionals (6). As the space–time dynamics of these relations in $Q(D_t)$ and $R(D_t)$ is expressed via the sensitivity functions (SFs) of the analyzed functionals to variations of model parameters, a joint analysis of data, models, and SFs requires agreed determination of the energy scalar products in the corresponding functional spaces. These algorithmic structures are based on the variational principle for estimating functionals and models [3, 4]:

$$\tilde{\Phi}_k^h(\boldsymbol{\varphi}) \equiv \Phi_k^h(\boldsymbol{\varphi}) + [I^h(\boldsymbol{\varphi}, \boldsymbol{Y}, \boldsymbol{\varphi}^*)]_{D_t^h}; \tag{8}$$

$$I(\boldsymbol{\varphi}, \boldsymbol{Y}, \boldsymbol{\varphi}^*) \equiv \Big(B\,\frac{\partial \boldsymbol{\varphi}}{\partial t} + G(\boldsymbol{\varphi}, \boldsymbol{Y}) - \boldsymbol{f} - \boldsymbol{r}, \boldsymbol{\varphi}^*\Big) = 0. \tag{9}$$

Here $\boldsymbol{\varphi}^* \in Q^*(D_t)$ are auxiliary functions determined by the specific features of the variational principle and belonging to the space adjoint with respect to the space of the state functions, $\tilde{\Phi}_k^h(\boldsymbol{\varphi})$ is an augmented functional, which takes into account the objective functional of the form (6), (7) and the description of the mathematical model in variational form; the superscript h marks the discrete analogs of the corresponding objects.

The integral identity (9) is a variational formulation of models (1)–(4). The functional is chosen so that relation (9) at $\boldsymbol{\varphi}^* = \boldsymbol{\varphi}$ transforms to the equation of the total energy balance in the system. The required algorithmic structures and the definitions of the scalar products are obtained on the basis of the integral identity (9) and functional (8). For instance, if the model of atmosphere hydrodynamics in the quasi-static approximation is used as the basic model, the energy scalar product for the state functions in (1)–(6) can be chosen in the form

$$(\boldsymbol{\varphi}, \boldsymbol{\varphi}^*)_{Q(D_t)} = \int\limits_{D_t} \Big\{ uu^* + vv^* + \sigma_0 \Big[TT^* + \frac{\gamma(p)}{R^2}\, HH^* \Big] \Big\}\, dD\, dt + \sum_{i=1}^{n} \int\limits_{D_t} \beta_i \varphi_i \varphi_i^*\, dD\, dt, \tag{10}$$

where $\boldsymbol{\varphi} = (u, v, T, H, \varphi_i,\ i = \overline{1, nf})$ is the vector function of the state variables, u and v are the horizontal components of the velocity vector, T is the temperature, H is the geopotential, φ_i are the functions that describe the components of the hydrological cycle and concentrations of admixtures in the gaseous and aerosol states, nf is the total number of substances, R is the universal gas constant, and σ_0, $\gamma(p)$, and β_i are the weight factors chosen so that the operations of summation of terms with disparate components of the state function are valid at $\boldsymbol{\varphi}^* = \boldsymbol{\varphi}$.

To study the sensitivity functions, we introduce a scalar product agreed with (6)–(9) and generated by the right side of the sensitivity relations for the functionals $\Phi_k(\boldsymbol{\varphi})$ [3]:

$$\delta\Phi_k^h(\boldsymbol{\varphi}) \equiv (\mathrm{grad}_{\boldsymbol{Y}}\, \Phi_k^h(\boldsymbol{\varphi}), \delta\boldsymbol{Y}) \equiv \frac{\partial}{\partial \alpha}\, I^h(\boldsymbol{\varphi}, \boldsymbol{Y} + \alpha\delta\boldsymbol{Y}, \boldsymbol{\varphi}_k^*)\Big|_{\alpha=0}; \tag{11}$$

$$\delta\boldsymbol{Y} = \eta\, \mathrm{grad}_{\boldsymbol{Y}}\, \Phi_k^h(\boldsymbol{\varphi}), \qquad k = \overline{1, Kc}. \tag{12}$$

Here α and η are the real parameters, $\delta\boldsymbol{Y}$ are the variations of the model parameter vector, which are chosen to be proportional to the sensitivity functions, $\boldsymbol{\varphi}$ is the solution of the basic problem (1)–(4) with prescribed values of the set of parameters $\boldsymbol{Y}$, and $\boldsymbol{\varphi}_k^*$ are the solutions of the adjoint problems generated by variational principles for estimates of variations of the augmented functionals $\tilde{\Phi}_k^h(\boldsymbol{\varphi})$. For convenience of constructing the algorithmic structures, the functions of sources, initial data, and uncertainty functions of the models are included into the number of components of the parameter vector $\boldsymbol{Y}$. The structure of the phase space of sensitivity functions defined by the right side of Eq. (12) is determined by the structure and dimensions of components of the parameter vector and by the form of the functional of the integral identity.

2. Algorithms of Database Analysis and Forecasting Results. Let us consider available information on the examined processes: observations of the actual behavior of the climate-ecological system, results of diagnostic and prognostic scenarios calculated with the use of mathematical models, and fields of the calculated sensitivity functions for a prescribed set of the generalized characteristics of the system. The state functions and sensitivity functions are multicomponent vector aggregates that describe various aspects of the processes under study and are determined in the general case on four-dimensional space–time domains. The components have dimensions correspondent to their physical contents. With allowance for these circumstances, operation with such vectors involves an energy scalar product constructed, for instance, on the basis of definitions (10), (11), or their combinations.

2.1. *Structuring of the Database.* The set of vectors is defined as follows:

$$\Big\{\boldsymbol{\varphi}(\boldsymbol{x}, t, \boldsymbol{Y}) \in Q(D_t);\ (\boldsymbol{x}, t) \in D_t;\ \boldsymbol{Y}(\boldsymbol{x}, t) \in R(D_t)\Big\}. \tag{13}$$

For organizing the algorithms, we arrange the database in the form of a matrix with a block structure. As the algorithms are universal and the specific features of each problem are determined by the structure of the data matrix and by the form of the scalar products, it is reasonable to define the block structure of the matrix depending on the research goals to ensure efficient operation of these algorithms. To describe the blocks, we introduce two groups of independent variables-indices. The first group describes the external structure of the data: the number of blocks and their order in the overall hierarchy. The second group describes the numeration and location of the components inside the block. Thus, the vectors can be presented in the block form as

$$\boldsymbol{\varphi} = \{\boldsymbol{\varphi}_i(k)\}, \qquad \boldsymbol{\varphi}_i(k) \in R_N, \quad i = \overline{1,n}, \quad n \geqslant 1, \quad k \in K \tag{14}$$

(n is the number of blocks in the external structure and K is the set of values of the multi-indices k of the components of the internal structure of each block). The total number of elements in the internal structure is denoted by N. Hereinafter, all operations are performed in real vector spaces R_N and R_n with scalar products.

We determine the energy scalar product at the block level:

$$(\boldsymbol{\varphi}_i, \boldsymbol{\varphi}_j) = \langle \boldsymbol{\varphi}_i(k), C\boldsymbol{\varphi}_j(k)\rangle, \qquad i, j = \overline{1,n}, \quad k \in K. \tag{15}$$

Here C is the diagonal matrix whose elements contain scale factors and elements of volumes in the discrete representation of functionals of the form (6), (7), (10), (11).

We introduce transformation of the state variables of the form $\boldsymbol{Z}_i = C^{1/2}\boldsymbol{\varphi}_i$ so that the components of the new vectors have identical dimensions and the energy properties and dimensions of the scalar product (15) and norms remain unchanged. Finally, the database for solving the problem can be presented in the form of an $n \times N$ matrix $Z = [\boldsymbol{Z}_i]$ ($i = \overline{1,n}$), where n is the number of vector columns, each column containing the entire internal structure with the total number of components being equal to N. The quantities n and N and the structure of the set of multi-indices K are the input parameters for structuring of the database Z and forming the scalar product (15).

The data matrix Z can be considered as a set of n vector columns of dimension N from R_N and as a set of N vector rows of dimension n from R_n; hence, we can use two Gram matrices: $n \times n$ matrix $\Gamma = Z^{\mathrm{t}}Z$ and $N \times N$ matrix $M = ZZ^{\mathrm{t}}$, respectively (the superscript "t" indicates transposition). Taking into account that $r \equiv \operatorname{rank}(\Gamma) = \operatorname{rank}(M) \leqslant \min(n, N)$, to make the algorithms efficient, we always structure the initial database so that the inequality $n \ll N$ holds and organize the basic computational scheme with the $n \times n$ matrix Γ. In this case, the value of the parameter n can be limited only by the capabilities of procedures of solving the full spectral problem for a symmetric, nonnegatively determined $n \times n$ matrix Γ. The dimension N of the vectors of the internal structure may be as large as desired.

2.2. *Quadratic Forms and Methods of Decomposition.* In addition to the Gram matrices, an effective instrument for studying linear transformations of vector fields and corresponding databases are bilinear and quadratic forms [8, 9]. Based on this fact, we put the following quadratic form into correspondence with the matrix Z:

$$S(\boldsymbol{v}) = (\boldsymbol{v}^{\mathrm{t}}Z^{\mathrm{t}}Z\boldsymbol{v}) \equiv (\boldsymbol{v}^{\mathrm{t}}\Gamma\boldsymbol{v}). \tag{16}$$

This form is determined in the space of vectors $\boldsymbol{v} \in R_n$. Using methods for studying extreme properties of the form $S(\boldsymbol{v})$, we perform orthogonal decomposition of the space of vectors from $R_N \times R_n$, which compose the matrix Z. Omitting the description of intermediate stages, we obtain the solution of the problem as a set of orthogonal subspaces

$$\left\{ \begin{array}{c} \lambda_p, \quad \boldsymbol{v}_p \in R_n, \quad \boldsymbol{\Psi}_p \in R_N, \\ \boldsymbol{v}_p^{\mathrm{t}}\boldsymbol{v}_p = \lambda_p\delta_{pq}, \quad \boldsymbol{\Psi}_p^{\mathrm{t}}\boldsymbol{\Psi}_p = \delta_{pq}, \quad p, q = \overline{1,n} \end{array} \right\}. \tag{17}$$

Here $\lambda_p \geqslant 0$ and $\boldsymbol{v}_p$ are the eigenvalues and eigenvectors of the $n \times n$ Gram matrix Γ in descending order, δ_{pq} is the Kronecker delta function, and $\boldsymbol{\Psi}_p$ are the subspaces formed by projections of the vector rows of the matrix Z onto the basis $\{\boldsymbol{v}_p\}$; the internal structure of the vectors $\boldsymbol{\Psi}_p$ is the same as that of the vector columns from the matrix Z in the subspaces of R_N.

Normalization of the vectors $\boldsymbol{v}_p$ by λ_p arranges the expansion of the initial space of vectors of the matrix Z into a set of subspaces $\{\boldsymbol{v}_p, \boldsymbol{\Psi}_p\}$ in descending order of the scale of perturbations characterized by λ_p.

By analogy with conventional definitions of the theory and methods of the principal factors and principal components [10, 11], the orthogonal vectors $\{\boldsymbol{v}_p\}$ can be interpreted as the principal components for presenting the vectors of the database Z in R_n, while the orthogonal basis vectors $\boldsymbol{\Psi}_p$ can be interpreted as natural or empirical functions for presenting the vectors of the database Z in R_N.

2.3. *Constructing Orthogonal Subspaces by an Example of the Reanalysis Database.* For climate information to be used for environmental forecasting, all databases that contain characteristics for a long time period are suitable. In the present study, we use the reanalysis database [12], which is a well-structured universal-purpose information system containing the basic set of characteristics of the atmosphere of the global climate system. The computational experiments involved the database for a period from 1950 to 2002 (53 years). The chosen time interval is greater than the averaging period normally used in climatological estimates. By the example of these data, we make the problem formulation more definite and identify the main elements of the technique used. To organize the study and to ensure efficient operation of the algorithms, we form a working database as a subset of the entire information reanalysis database [12] and determine the matrix Z and the quadratic form (16) in a suitable manner to have a generalized presentation of this database. First, it is necessary to study the long-term behavior of the global climate system and identify the annual and seasonal phenomena with detailed presentation of various elements of circulation systems in the space–time domain being decomposed in terms of perturbation scales at the level of orthogonal subspaces.

With allowance for the content of the objective functionals (6), scalar products (10)–(12), and functionals (15) for calculating the Gram matrices, we choose a "physical" set of components. Thus, we have seven parameters to characterize each element of the database (13), (14): number of the year, number of the month, number of the data field in terms of the physical content, and four parameters of the space–time presentation of data in a four-dimensional global domain on a sphere or its part. The first two parameters determine the external structure of the database and, correspondingly, vector rows in its presentation in the form of the matrix Z, and the remaining five parameters determine the internal structure. The number of the year is used as the leading variable of the external structure of the problem. In terms of Eq. (14), this is the index parameter $i = \overline{1, n}$. The number of the month is used as an input parameter of the algorithm. In the computations, n is the main parameter governing the algorithm efficiency and computational cost, the latter being mainly dependent on the possibility of effectively solving the full spectral problem for the Gram matrix at large values of n.

The overall structure of the working database (the one used for numerical experiments) and sought basis subspaces is organized with allowance for long-time annual, seasonal, and daily periods of the processes. To take into account the seasonal behavior, the original database is divided into 12 parts (according to the number of months in a year). The annual and daily behaviors are described by two time scales: external time scale in the global time with a total duration $n = 53$ years and a time step of 1 year and internal time scale with a duration of 1 month and with a time step defined parametrically, depending on the database capabilities. In the present work, we used a 12-hours' time step. Thus, we obtain 12 sets of factor spaces (in accordance with the number of months in a year); for each month, we construct a basis complex consisting of 53 orthogonal elements-subspaces (in accordance with the number of years).

The dimension of the vector columns of the matrix Z (14) is determined by the following parameters: 1) in terms of time, $2m$ (m is the number of days in a month, two measurements taken every day); 2) in terms of space, by the size of the global domain in spherical coordinates with a resolution of $2.5° \times 2.5°$ for the horizontal variables and the number of vertical levels set parametrically; 3) in terms of the number of physically different block components, which is also introduced parametrically, based on the structure of the state function in models (1)–(5) and the composition of elements in the scalar products (10), (11). In the regional version, the configuration and location of the region on the globe and the space–time resolution are also set parametrically. Data sampling in time within monthly intervals is a procedure, which is, on one hand, a compromise in terms of the informativeness and efficiency of computations and, on the other hand, convenient for the content analysis and organizing modeling scenarios.

Thus, using the results of orthogonal decomposition, we can perform a fast qualitative and quantitative analysis of the behavior of complicated dynamic systems and deliberately form scenarios based on prescribed criteria for solving diagnostic and prognostic problems.

Let us consider an example of the climate system analyzed for a period of 53 years (1950–2002). Figure 1 shows one of 60 fragments of the first (leading) June basis vector corresponding to the field of horizontal components of wind velocity at a level with a pressure of 500 hPa. As a whole, the vector is normalized in accordance with Eq. (17). The ordinate is the function of the latitude y such that the South and North Poles are located at 0 and 180°. The abscissa is the longitude x from the Greenwich meridian. The informativeness of the first basis vector is 12.3% in fractions of the trace of the Gram matrix for the entire set of data. Climatically significant vectors ($\lambda_p \geqslant 1$) are 18 out of 53.

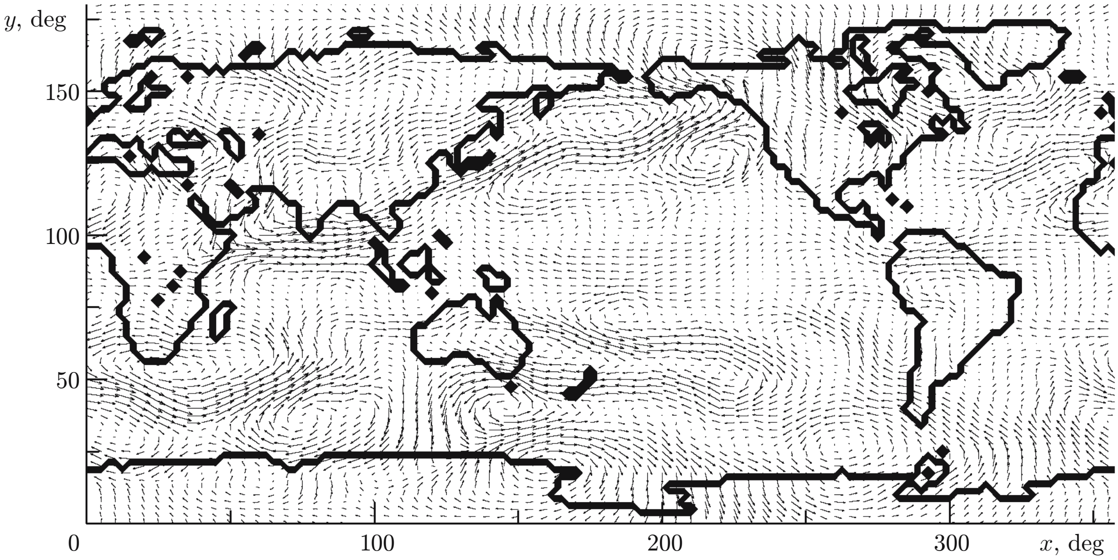

Fig. 1. Component of the first basis vector corresponding to the fields of horizontal components of wind velocity (arrows) at a pressure of 500 hPa in June 1950–2002 (time moment at 00:00 GMT, June 15): the solid curves are the continent boundaries.

The summer type of circulation is established above Eurasia in June under the influence of the Arctic basin. Typical circulation structures characterizing the active zones of the global climate system are clearly visible in Fig. 1. It should be noted that the leading subspace has a quasi-stationary character during the month, i.e., the main circulation structures have fairly stable localization in space.

3. Estimate of Environmental Risks in the Far-East Region. Below we consider an example demonstrating some aspects of environmental forecasting in the region with the use of information on the global climate system. We consider the Far-East region of Russia and adjacent territories of China, Korea, and Japan. In accordance with the estimates of the active zones and regions of risk (vulnerability) in the global climate system, this region is located in the zone of increased environmental risks [4, 13]. From the ecological and geopolitical viewpoint, it is an important object for environmental research and forecasting. The hydrometeorological regime of this region is formed by processes that occur at the ocean–continent interface and processes generated in the Altai–Sayan energy-active zone.

The Altai–Sayan cyclogenesis area is classified as one of the most active zones in the North Hemisphere [14], its most active part being located in the region of 40–50° North latitude and $95 \div 125°$ East longitude from Altai and Sayan on the west to Great Xinganling on the east. (In Fig. 1, this zone is within the band of 130–140° on the ordinate axis.) A large number of interacting factors are responsible for the high space–time variability of the processes of energy and mass transfer in the regional atmosphere. In turn, owing to transport of pollutants from regions with high anthropogenic loading, situations with increased environmental risks arise. Let us consider two fragments of a typical summer scenario and estimate the environmental risks of atmospheric pollution above the region territories. The receptor areas are assumed to be some large cities: Khabarovsk, Vladivostok, Beijing, Shenyang, Harbin, Dalian, Seoul, and Pyongyang. Objective functionals of the form (6), which yield an estimate of the total amount of pollutants accumulated in the atmosphere of the receptor cities, were used in the computations. The scenario is implemented in the regime of inverse modeling on the basis of a four-dimensional model of pollutant transport (4) adapted to regional conditions. The hydrodynamic background is formed on the basis of reanalysis data [12] with the use of the system described in [15]. The sensitivity functions and the regions with a high risk of transportation of pollutants to the receptor areas are calculated, which are structurally defined by the carriers of non-zero values of the weight functions in the definition of the quality functional (6). The risk functions identify areas within the region, where the acting and potentially possible sources of pollution are located, which "send" a signal to the quality functional.

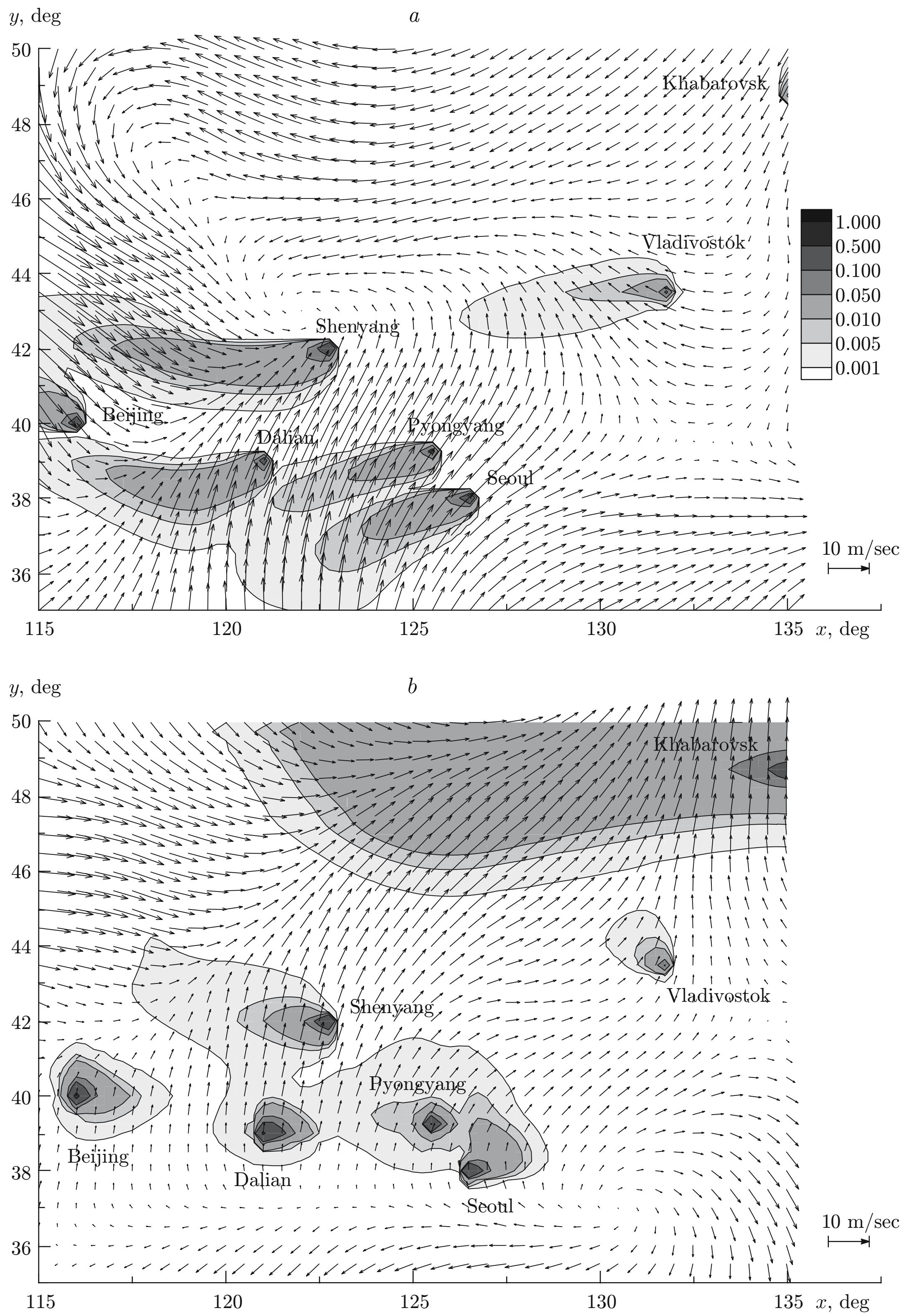

Fig. 2. Two-dimensional sections of the velocity fields (arrows) and isocontours of the risk function for receptor cities at the level corresponding to the surface layer height: (a) time moments at 9 : 00 GMT, June 6, 2000; (b) 8 : 00 GMT, June 3, 2000 (b).

The fields of velocity and sensitivity functions are four-dimensional structures. Figure 2 shows their two-dimensional sections at a computational level of the model corresponding to the surface layer height. The risk function is normalized to its maximum value during the entire scenario period.

A comparative analysis shows that the space and time variability of the hydrometeorological regime and sensitivity (risk) functions is really high. The character of the time moment on June 6 (Fig. 2a) is determined by the north-east background flow of air masses. The city of Khabarovsk at that time is in the zone of comparative ecological wellbeing, i.e., none of the sources in the region (outside the receptor city) affects the quality of the atmosphere in the city. The situation on June 3 (Fig. 2b) is substantially different from the previous one. It is determined by the background flow from the Altai–Sayan region. Being affected by the Great Xinganling Ridge, the air flows start intensively moving along the valleys of the Amur and Sungari rivers. These air flows can bring pollutants from sources located on a considerable area of the region to the atmosphere of Khabarovsk, which significantly deteriorates the ecological situation in the city. A high danger of adverse situations owing to exchange of pollutants is also observed for the cities of China and Korea.

Thus, the synergy of the effects caused by the high activity of the climate system and high anthropogenic loading on nature is observed in the region considered. As a consequence, the regions of potential environmental risks have considerable space and time scales.

Conclusions. The main challenge of environmental forecasting in planning economic activities is reduction of risks and adverse consequences of natural and man-caused catastrophes. The modeling systems with the maximum possible use of available data information may serve as effective tools for the development of environment-protection strategies.

This work was supported by the Mathematical Department of the Russian Academy of Sciences (Program No. 1.3), Russian Foundation for Basic Research (Grant No. 04-05-64562), European Commission (Grant No. 013427), and Program for Basic Research of the Presidium of the Russian Academy of Sciences (Program No. 16).

REFERENCES

1. G. I. Marchuk, *Mathematical Simulation in Environmental Problems* [in Russian], Nauka, Moscow (1982).
2. G. I. Marchuk, *Adjoint Equations and Analysis of Complex Systems* [in Russian], Nauka, Moscow (1992).
3. V. V. Penenko, *Methods of Numerical Simulation of Atmospheric Processes* [in Russian], Gidrometeoizdat, Leningrad (1981).
4. V. V. Penenko, "Methods of inverse modeling and estimating environmental risks due to anthropogenic effects," *Obozr. Prikl. Prom. Mat.*, **10**, No. 1, 26–38 (2003).
5. V. V. Penenko and E. A. Tsvetova, "Mathematical models for the study of interactions in the system Lake Baikal–atmosphere of the region," *J. Appl. Mech. Tech. Phys.*, **40**, No. 2, 308–316 (1999).
6. R. T. Watson, "Climate change 2001," Synthesis Report, IPCC, Geneva (2003).
7. L. Schwartz, *Analysis* [Russian translation], Mir, Moscow (1972).
8. R. Courant and D. Gilbert, *Methods of Mathematical Physics*, Wiley (1981).
9. F. R. Gantmacher, *The Theory of Matrices*, AMS (1998).
10. H. H. Harman, *Modern Factor Analysis*, Univ. of Chicago Press, Chicago (1976).
11. R. W. Preisendorfer, *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, Amsterdam, New York, Tokyo (1988).
12. E. Kalnay, M. Kanamitsu, R. Kistler, et al., "The NCEP/NCAR 40-year reanalysis project," *Bull. Amer. Meteorol. Soc.*, **77**, 437–471 (1996).
13. V. V. Penenko and E. A. Tsvetova, "Methods of sensitivity theory and orthogonal decomposition for studying climate dynamics and pollution," *Proc. SPIE*, **6522**, 652223 (2006).
14. S.-J. Chen and Y.-H. Kuo, "Synoptic climatology of cyclogenesis over East Asia, 1958–1987," *Mon. Weather Rev.*, **119**, No. 6, 1407–1418 (1991).
15. V. V. Penenko and E. A. Tsvetova, "Preparing data for environmental research with the use of Reanalysis," *Opt. Atmos. Okeana*, **12**, No. 5, 463–465 (1999).