Formation of atmospheric circulation scenarios for climate and ecological studies

V.V. Penenko and E.A. Tsvetova

Institute of Computational Mathematics and Mathematical Geophysics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk

Received February 12, 2001

A new approach to construction of atmospheric circulation scenarios for making ecological predictions and planning is proposed. It is based on a combined use of hydrothermodynamic models and archived data on climate. The main feature required from the scenarios is that their information content in climatic aspect to be no less than that of climatic models of the general circulation of the atmosphere. In addition, the scenarios should reasonably accurate reconstruct any situation from the database used for its construction according to some preset criteria. Implementation of the approach is based on the factor analysis, measurement data assimilation, and optimal extrapolation on the informative basis. Depending on the aim of a study, multicomponent bases with a given time and spatial structure are constructed. Numerical models are used to fit and assimilate measurement data, to reconstruct missing data, to calculate the state functions, as well as the adjoint and sensitivity functions needed in forming the scenarios.

Introduction

The problem on the formation of hydrothermodynamic background for ecological prediction and planning has no a unique solution until now. Its specificity is in the fact that the characteristic lifetime of the already existing and designed objects as a sources of anthropogenic impact on the climate system are, as a rule, much longer than the characteristic intervals of predictability of the current hydrothermodynamic models forming the basis for general circulation models and the theory of climate. This means that in solving such problems it is preferable to make use of the scenario approach, in which the forecast is being constructed based on a set of solutions for typical situations. In construction of scenarios, mathematical models are used, as rule, along with the measurement data. In this case, it is always possible to compromise the structure of algorithmic description of the processes by combining the best properties of informative, i.e., based on the actual information, and hydrothermodynamic models.

In the proposed technique, the set of scenarios is constructed based on the combination of four methods of system levels, namely, the methods of factor analysis, methods of studying the sensitivity of models and functionals, methods of direct and inverse modeling with assimilation of actual data. The idea of this complex approach and the principles of organization of its algorithmic constructions have been formulated in Ref. 1 in 1981 as early. However, at that time insufficient computer resources and databases available for numerical experiments restricted the possibilities of implementing this approach.

It is clear that scenarios for ecological prediction should reflect actual situations on the climatic scale. Consequently, databases for their construction should contain information compatible with such time scales. Now we have various databases on the state of the climate system, as well as the capabilities of their efficient use in solving research and practical problems. In our study, we use the archive of NCEP/NCAR data reanalysis.[2] It is very convenient for multifunctional applications in the interrelated problems in ecology and climate.[3–6] To make use of these data, we have created a specially developed modeling system, which allows the structure of atmospheric circulation to be reconstructed with the help of hydrothermodynamic models and with the use of the data of reanalysis in the assimilation mode.[7] Our experience shows that for applications we should construct scenarios of the spatiotemporal structure.[3–6]

The proposed algorithms for construction of informative bases and subspaces using the method of factor analysis supplement this modeling system. Various aspects of factor analysis are considered in Ref. 8. It should be noted that factor analysis has already a 100-year history starting from the paper by Pearson dated by 1901. Some of its applications to solving the problems of hydrometeorology as a method for separating natural components of meteorological fields are discussed in Ref. 9.

Statement of the problem

The central place in the methodology of formation of scenarios and analysis of the results obtained by modeling is identification of the main factors that govern the behavior of the climate system. In terms of the main factors, it is also possible to identify the manifestations of the climate system response to anthropogenic impact.

0235-6880/01/06–07 444-05 $02.00

The problem is to construct the algorithm of factor analysis (assuming that the structure of the data set to be analyzed is determined) and to organize their relation to the methods of model sensitivity. Mathematical simulation always involves three components: model, data, and algorithms, among which the main and linking component is the mathematical model of the processes under study. When formulating a mathematical model, we also formulate a suitable variational principle, which is a generalized characteristic of all elements of the modeling system. It not only reflects the physical nature of the processes, but also generates all computational algorithms. In this case, when determining functional spaces, scalar products in them, the structure of input and output data, as well as the basis functions, we proceed from formulation of the class of models, including the models of hydrothermodynamics of the atmosphere and water objects and the model of transport and transformation of pollutants.[1,3–6,10]

When organizing algorithm of factor analysis, the structure of information spaces and arrays is connected with the structure of the objects like databases available, functions of the state of discrete analogs of the model, adjoint functions as solutions corresponding to the purposes of studying the adjoint problems, sensitivity functions of the models and functionals, and functions presenting the results of thematic processing of objects of the above four types.

Algorithms of factor analysis

Now we describe the scheme of constructing the algorithms of factor analysis. As an example, we take the spherical coordinate system along the horizontal and the coordinate following the land terrain along the vertical. The main elements of our modeling technology are constructed as applied to this scheme.[3,4,7]

1. Represent the initial data by a matrix set

$$\Phi = \{\boldsymbol{\varphi}_\beta \equiv [\varphi^j_{\beta\alpha imk}]\}, \qquad (1)$$

where β is the number of a vector in the set, $\beta = \overline{1, nk}$, $nk \geq 1$; $\varphi^j_{\beta\alpha imk} = \varphi_{\beta\alpha}(t_j, \psi_i, \theta_m, \sigma_k)$; j denotes time, $j = \overline{1, jk}$, $jk \geq 1$; i denotes parallel (longitude), $i = \overline{1, ik}$; m denotes meridian (addition to latitude), $m = \overline{1, mk}$; k denotes the vertical, $k = \overline{1, kk}$; α denotes the physical dimension of components (in the case of multicomponent analysis, for example, $\varphi_\alpha \equiv \{u, v, \dot{\sigma}, H, p, T, q\}$, u, v, and $\dot{\sigma}$ are components of the velocity vector, H is the geopotential, p is the pressure, T is the temperature, q is the specific humidity of air); $\alpha = \overline{1, \alpha k}$, αk is the number of different components, $\alpha k \geq 1$. If some parameter from the set jk, ik, mk, kk, αk equals unity, then the corresponding dimensionality of the vector $\boldsymbol{\varphi}_\beta$ is excluded from the consideration.

2. Introduce the definition of the scalar product and quadrature formulas for the function of state $\boldsymbol{\varphi}_\beta$ ($\beta = \overline{1, nk}$) as

$$(\boldsymbol{\varphi}_{\beta_1}, \boldsymbol{\varphi}_{\beta_2}) = \sum_{j=1}^{jk} \sum_{i=1}^{ik} \sum_{m=1}^{mk} \sum_{k=1}^{kk} \left\{ \sum_{\alpha=1}^{\alpha k} \langle \varphi_{\beta_1}, \varphi_{\beta_2} \rangle_\alpha \, \eta_\alpha \, \omega_\alpha \right\}^j_{imk} \delta D^j_{imk}, \quad (2)$$

where $\delta D^j_{imk} = \delta D_{imk} \cdot \delta t_j$; $\delta D_{imk} = \delta S_{im} \, \pi_{im} \, \delta\sigma_k$; δS_{im} are the area elements on the spherical surface; $\pi_{im}\delta\sigma_k$ is the vertical dimension of the cells; π_{im} is the function depending on pressure; δt_j is the time step; $\omega^j_{\alpha imk}$ is the weighting function that is chosen from the conditions of normalization of the set vectors to unity. The sums in Eq. (2) follow from approximation of integrals in the scalar product. The scalar product itself in the space of the functions (1) is constructed based on the equation for the integral of the total energy of the system, for which the mathematical model is constructed, and on the definition of the corresponding integral identity.[1,3] To equalize physical dimensions in the multicomponent functions (1) at $\alpha k > 1$, the factors η_α are used. They are chosen according to the definition of the total energy of the system and the functional of the variational formulation of the model and its variations for calculation of the sensitivity ratios of the models.[1,3,4,10] Thus, the methods of factor analysis are related to the methods for studying model sensitivity through the selection of the scalar product.

3. Preparation of the vectors consists in calculation of the mean values, deviations from the mean values, and their normalization. Normalized vectors of deviations have the form

$$\boldsymbol{\varphi}^{\text{norm}}_\beta = \frac{1}{S_\beta} \boldsymbol{\varphi}'_\beta \equiv \left\{ \frac{1}{S_\beta} (\varphi^j_{\beta\alpha imk})' \right\}, \qquad (3)$$

where $\boldsymbol{\varphi}'_\beta = \boldsymbol{\varphi}_\beta - \bar{\boldsymbol{\varphi}}_\beta$ is the vector of deviation and $S^2_\beta \equiv (\boldsymbol{\varphi}'_\beta, \boldsymbol{\varphi}'_\beta)$ is the square norm of the vector of deviation.

4. The Gram matrix for the vectors of the set is calculated by the following equations:

$$R = \left\{ r_{\beta_1\beta_2} \equiv (\boldsymbol{\varphi}^{\text{norm}}_{\beta_1}, \boldsymbol{\varphi}^{\text{norm}}_{\beta_2}), \quad \beta_1, \beta_2 = \overline{1, nk} \right\}. \quad (4)$$

The scalar product in Eq. (4) is calculated by Eq. (2).

Actually, factor analysis consists in the following.[8] A representation of the vectors $\boldsymbol{\varphi}^{\text{norm}}_\beta$ of the set (1) and (3) is sought in the form

$$\boldsymbol{\varphi}^{\text{norm}}_\beta = \sum_{p=1}^{nf} a_{\beta p} \, \mathbf{F}_p, \quad nf \leq nk, \quad \beta = \overline{1, nk}, \qquad (5)$$

where $\mathbf{F}_p$ are the orthonormal basis functions, and $a_{\beta p}$ are the expansion coefficients.

The elements of the expansion (5), in their turn, are sought among the given set of values of the vectors (3) based on the conditions of sequential minimization of the following parameters:

$$V_p = \sum_{\beta=1}^{nk} a_{\beta p}^2, \quad p = \overline{1, nf} \tag{6}$$

under the restriction

$$r_{\beta_1\beta_2} \equiv \sum_{k=1}^{nf} a_{\beta_1 k}\, a_{\beta_2 k}, \quad \beta_1, \beta_2 = \overline{1, nk}, \tag{7}$$

i.e., each of the $\mathbf{F}_k$ components in the order as in Eq. (5) should account for the maximum of the total variance of the initial set of vectors in the subspace resulting from subtraction of the sum of vectors of the preceding factors $\mathbf{F}_i$, $i = \overline{1, k-1}$, from the vectors (3). At $k = 1$ the entire initial space (1), (3) is taken into account.

The parameter V_p is the sum of the contributions of the basis function $\mathbf{F}_p$ estimated in the metric of space Φ (1) generated by the scalar product (2) to the sum of norms of the deviation vectors (3), which is equal to nk. The sought coefficients $a_{\beta p}$ in Eq. (5) are obtained from solution of the complete spectral problem for the Gramm matrix (4).

5. The complete spectral problem for the Gram matrix is formulated as

$$R\alpha = \alpha\Lambda, \tag{8}$$

where $\alpha \equiv \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_{nk}\}$ are eigenvectors; $\Lambda \equiv \text{diag}\,\{\lambda_1, \lambda_2, \ldots, \lambda_{nk}\}$ are eigenvalues; $\boldsymbol{\alpha}_\beta \equiv \{\alpha_{j\beta}, j = \overline{1, nk}, \beta = \overline{1, nk}\}$ is the structure of an eigenvector. The first subscript, j, is the number of a component of this vector, the second one, β, is the number of the vector. The eigenvalues are given in the order of the decreasing λ value, $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_{nk}$, and the corresponding vectors $\boldsymbol{\alpha}_\beta$ are given in the same order. To completely solve this complete spectral problem, we use procedures from the LINAL library.[11]

6. The expansion coefficients in Eq. (5) are obtained at normalization of eigenvectors

$$\mathbf{a}_k = \frac{\sqrt{\lambda_k}\, \boldsymbol{\alpha}_k}{(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_k)^{1/2}} \equiv \left\{ a_{jk} = \frac{\sqrt{\lambda_k}\, \alpha_{jk}}{\sqrt{\sum_{j=1}^{nk} \alpha_{jk}^2}} \right\}, \tag{9}$$

$$j = \overline{1, nk}, \; k = \overline{1, nf}.$$

From this, we have

$$\sum_{j=1}^{nk} a_{jk}^2 = \lambda_k = V_k, \quad k = \overline{1, nf}. \tag{10}$$

7. Finally, the sought basis vectors are calculated by the equations

$$\mathbf{F}_p = \sum_{\beta=1}^{nk} \frac{a_{\beta p}}{\lambda_p} \boldsymbol{\varphi}_\beta^{\text{norm}}, \quad p = \overline{1, nf}.$$

All the initial normalized vectors (3) are used in calculations. The function $f_m(\lambda) = \sum_{k=1}^{m} \lambda_k$, $m \le nk$, can be considered as a function of the basis information content on the initial set (1); $f_{nk}(\lambda) = nk$. The number of the basis functions nf is set based on the decreasing rate of eigenvalues and the behavior of the function $f_m(\lambda)$. Fulfillment of the condition that the basis is orthonormal is checked directly.

The thus constructed bases are used for analysis of the initial set (1) to identify and interpret the main factors determining the behavior of the system under study. In the problems of ecological prediction, the hypothesis of relative stability of climate is accepted. On these assumptions, the space-time bases are constructed with the use of reanalysis data for the periods longer than 45 years. These bases are then used to form the subspace directrices, from which the detailed behavior of the functions of state in the modes of diagnosis, forecasting, and extrapolation is reconstructed with the use of mathematical models.

As an example, let us present the results of a typical modeled scenario for studying atmospheric processes by the methods of factor analysis. In this example, the task was to evaluate the role of Baikal region, as a source and sink of distortions in the climate system, from the manifestations in geopotential fields at the height of 500-mbar surface, i.e., to assess the scope of Lake Baikal action as a climate-forming factor on the global and local aspects. In particular, the factor components of sets of the geopotential fields were calculated and their information content was estimated using, as an example, December 1998. Figures 1 and 2 show the measure of information content of the basis functions at the dimension of the analyzed set equal to 62 (62 hemispheric geopotential fields for the month according to the data of reanalysis). Figure 1 shows the diagram of the eigenvalues of the Gramm matrix for this set. The eigenvalues are given in the decreasing order; they determine the relative contribution of the corresponding basis function in the series expansion of the whole set of fields in terms of the basis. The sum of contributions from all the basis functions is equal to 62.

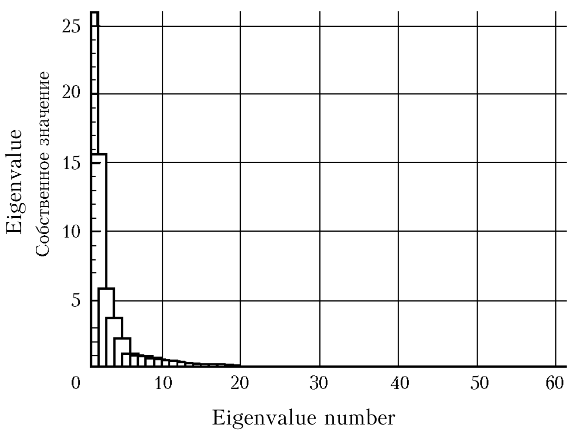

Fig. 1. Eigenvalues of the Gramm matrix for the set of geopotential fields (December 1998).

Figure 2 demonstrates how the information content of the basis series increases depending on the number of functions as it varies from 1 to 62.

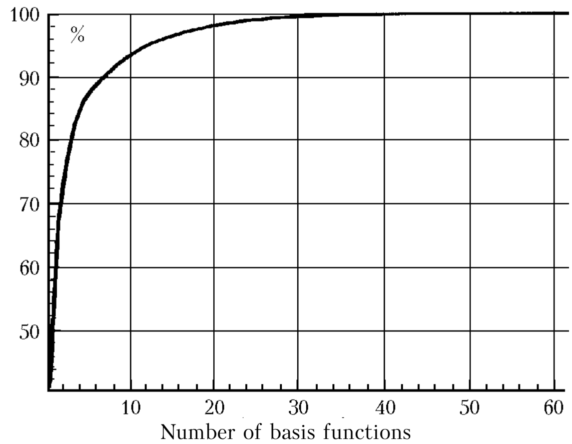

Fig. 2. Function of the information content of basis vectors depending on their number.

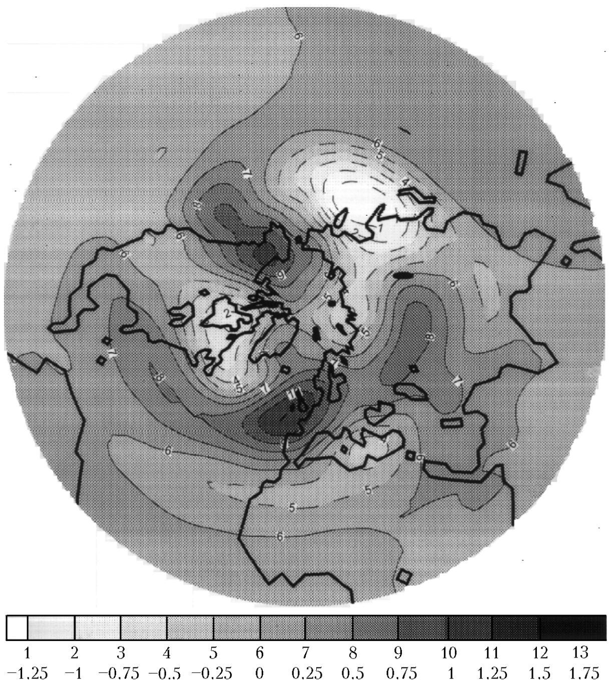

Fig. 3. The main basis function for the set of geopotential fields corresponding to the height level of 500 mbar pressure, December 1998, $n = 1$.

Figure 3 shows the main component of the geopotential field. Analysis of the whole set of basis functions shows that the effect of Lake Baikal becomes noticeable in the climate system on the characteristic scales of distortions described by the sixth to eighth basis functions. This is seen, for example, in Fig. 4, which shows the sixth basis function. Here Lake Baikal is in the domain of characteristic distortions on the regional scale. The relative contribution of the sixth to eighth functions has the value of the order of unity, i.e., it makes up roughly 1/25 of the contribution of the main factors shown in Fig. 3. From the comparison of the first basis function (see Fig. 3) and the sixth function (see Fig. 4) with the allowance made for their contributions, we can judge on the relation between the global and regional climate-forming factors.

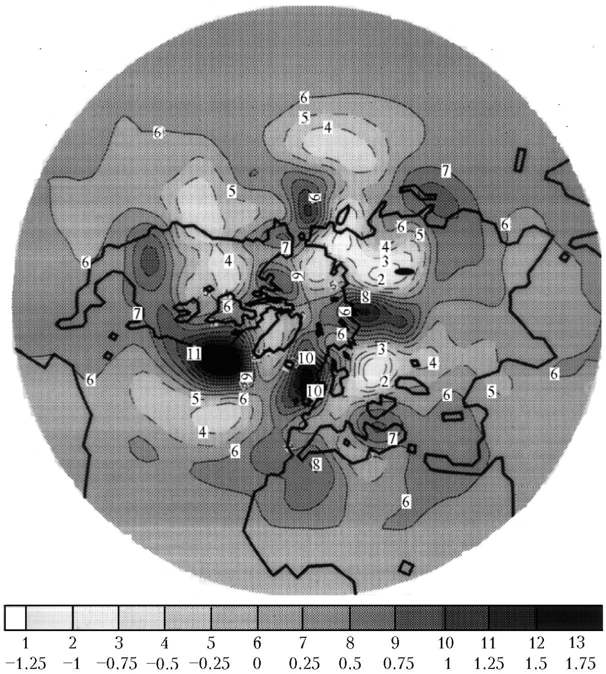

Fig. 4. Sixth basis function reflecting the effect of regional-scale factors, December 1998, $n = 6$.

Thus, it follows from the calculations that the information content at the level of the currently available general circulation models (reaching now 90%) is achieved, in this particular situation, with the use of less than 10 basis functions. This confirms the hypothesis on the possibility of separating centers of action in the climate system needed for regional estimates in the aspect of global changes. That is, we can follow up at what level of the main factors (scales) action the effect of regional distortions begins to manifest itself against the background of the global changes. In the example considered here this ratio is 25 : 1.

We believe that the development of the research field presented in this paper based on the combined use of the methods of sensitivity theory and factor analysis in the modeling technology is of high priority for the further theoretical and practical climate and ecological studies.

Acknowledgments

The work was partially supported by the Grant "Leading Scientific Schools" No. 00–15–98543, of the Russian Foundation for Basic Research (Grant No. 98–05–65318), Program of Promising Computer Technologies of the Russian Ministry of Industry, Science, and Technologies (Project 0201.06.269/349) and Integration Grants of the Siberian Branch of the Russian Academy of Sciences–2000 (No. 64 and No. 56).

References

1. V.V. Penenko, *Methods for Numerical Simulation of Atmospheric Processes* (Gidrometeoizdat, Leningrad, 1981), 351 pp.
2. E. Kalney, M. Ranamitsu, R. Kistler, et al., Bull. Am. Meteorol. Soc. **77**, 437–471 (1996).
3. V.V. Penenko and E.A. Tsvetova, Prikl. Mekh. Tekh. Fiz. **40**, No. 2, 137–147 (1999).
4. V.V. Penenko and E.A. Tsvetova, Prikl. Mekh. Tekh. Fiz. **41**, No. 5, 161–170 (2000).
5. V.V. Penenko and E.A. Tsvetova, Atmos. Oceanic Opt. **12**, No. 5, 447–449 (1999).
6. V.V. Penenko and E.A. Tsvetova, Atmos. Oceanic Opt. **13**, No. 4, 361–365 (2000).
7. V.V. Penenko and E.A. Tsvetova, Atmos. Oceanic Opt. **12**, No. 6, 462–468 (1999).
8. G. Kharman, *Modern Factor Analysis* (Statistika, Moscow, 1972), 486 pp.
9. A.V. Meshcherskaya, L.V. Rukhovets, M.I. Yudin, and N.I. Yakovleva, *Natural Components of Meteorological Fields* (Gidrometeoizdat, Leningrad, 1970), 199 pp.
10. V.V. Penenko and A.E. Aloyan, *Models and Methods for Environmental Protection Problems* (Nauka, Novosibirsk, 1985), 254 pp.
11. *LINAL Library for Solution of Linear Algebra Problems. On-line Material* (Computer Center SB RAS, Novosibirsk, 1985), 144 pp.