Preparation of data for ecological studies with the use of REANALYSIS

V.V. Penenko and E.A. Tsvetova

Institute of Computational Mathematics and Mathematical Geophysics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk
Received March 3, 1999

The algorithms for transformation of the Reanalysis NCEP/NCAR data from their standard system to grid domains of variable structure are described. These data are used in basis models of hydrothermodynamics and admixture transfer in the atmosphere. To reconstruct the detailed spatiotemporal structure of fields, the algorithm for assimilation of data is developed. The algorithm provides for "hydrothermodynamic" interpolation with the basis models. Different versions of data reconstruction for the direct and inverse simulation are considered.

Introduction

In the problems on the environmental protection planning and evaluation of ecological future of industrial regions, the question always arises on suitable choice of the methods for description of background hydrometeorological situations. Actually, the problem of long-term forecasting of ecological situation under conditions of uncertain input data and external actions is stated.

It seems logical to use the scenario approach, which can provide a solution to the problem with a sufficient degree of certainty at the level of deterministic description provided that the set of scenarios is chosen properly.

There are several ways for realizing the scenario approach. One of these ways is to use the models of hydrothermodynamics for prognostic description of the behavior of the atmosphere. Naturally, in this case it is needed to answer the questions on the predictability of these models with the allowance made for the existing uncertainties. This very interesting scientific problem is still open.

Assuming relative stability of the climate, the retrospective hydrometeorological information can be used to form scenarios of the atmospheric circulation. This can be done with the use of Reanalysis NCEP/NCAR database,[1] which contains the systematized information on most important components of the function of state of the atmosphere for 40 years. Depending on the scales of phenomena under study, decomposition of the functions of state to the background and perturbations is used. Then the background processes are generated by the archived data, and the detailed spatiotemporal structure is reconstructed using the models, which take into account a variety of natural and anthropogenic factors and work in the assimilation mode.

The database of reanalysis is a well-structured information system, which is organized at the modern level and oriented toward supporting research and practical applications in the field of atmospheric physics and ecology. All the information about composition, structure of fields, coordinate systems, etc. can be found in Ref. 1.

In this paper we try to answer the question on how these data can be used in our interests? In other words, the aim is to develop a technique for reconstructing (from these data) the spatiotemporal structure of fields of the functions of state of the atmosphere with a preset resolution on the grid domains accepted in our basic models of hydrothermodynamics and admixture transfer.[2–4] The reconstruction must be carried out without the information loss in the initial data. This problem is an interesting example of applying the technique of data assimilation with the help of models and joint use of models and the observed data.[2]

Thus, we can formulate the concept of constructing the atmospheric model of the information type. The model should be based on the methods of data assimilation and interpolation with the help of the basic model of the processes under study. The former model uses the data of reanalysis together with other information available as an assimilated background, while the detailed structure of fields is calculated by use of the basic model playing the role of a spatiotemporal object for the interpolation.

Such models are convenient for studying natural objects under the atmospheric effect. It is worth using them for generation of the background in the mesoscale models of hydrothermodynamics and in simulating processes of transfer and transformation of pollutants provided that the direct contribution from the feedback of these processes to the atmospheric dynamics is insignificant. At the same time, their indirect influence is taken into account indirectly through the background data on real atmosphere. In the case of an insignificant feedback, the basic models must work in the forecasting mode without assimilation.

0235-6880/99/05 447-03 $02.00

System organization of models of the information type

Let us describe the idea and the general scheme of algorithms for constructing the model of an information type. Following Ref. 2, let us formulate the goal functional

$$\Phi_0^h(\varphi) \equiv \alpha_1\, [(S\varphi - \varphi_R)^{\mathrm{T}}\, M_1\, (S\varphi - \varphi_R)]^h_{D_t^R} +$$

$$+\, \alpha_2\, [(\varphi - \varphi_a)^{\mathrm{T}}\, M_2(\varphi - \varphi_a)]^h_{D_t^R} + [I^h(\varphi,\varphi^*,\mathbf{Y})]^h_{D_t^R}; \quad (1)$$

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

Here we take the following designations and definitions: φ is the function of state of the atmosphere; φ^* is an arbitrary, sufficiently smooth function, which plays the role of generalized Lagrange factor; this function is introduced to take into account the basic model as a set of restrictions imposed on the function of state; $\mathbf{Y}$ is the vector of model parameters; $Q(D_t)$ is the space of the functions of state satisfying the boundary conditions of the model; $Q^*(D_t)$ is the conjugate space; $R(D_t)$ is the set of allowed values of the model parameters; $D_t = D\times[0, t]$, D is the domain of variability of the spatial coordinates, $[0, t]$ is the time interval; φ_R is the vector of values of the function of state of the atmosphere; the values are taken from the Reanalysis database; φ_a is the *a priori* estimate of the function of state in D_t; M_1 and M_2 are positively defined matrices of weights; the superscript T denotes the transposition, while the superscript h is used for discrete analogs of the corresponding objects; D_t^R is the discrete grid domain, in which the Reanalysis data are presented; D_t^h is the grid domain for constructing discrete analogs of the basic model and for forming the discrete approximations of the sought function of state of the atmosphere; S is the operator of information transformation from the grid D_t^h to the grid D_t^R. Note that to prevent the loss in accuracy, the Reanalysis data are taken on their own grids.

The first term in Eq. (1) gives the measure of deviation of the sought values of the function from the Reanalysis data, while the second term gives the measure of deviation from *a priori* estimate. The parameters α_1 and α_2 set the relative weight of these functionals. The third term is a discrete analog of the integral functional in the variational formulation of the basic model. If there appears some additional information on the time interval $[0, t]$, then it is included as an additional term of the same type as the first two terms discussed above.

The model described in Ref. 3 is taken as a basic one. It, in particular, uses the hybrid system of coordinates (p-sigma): isobaric coordinates above the level of 500 mbar and coordinates following the Earth's terrain below this level. It provides a compromise between the method of presentation of the Reanalysis data and the need to describe processes in the models of hydrothermodynamics and pollutant transfer with the allowance for the terrain.

Functional (1) gives rise to the structure of algorithms for the information model of the atmospheric dynamics. It can be approximated with the use of the calculus of variations in combination with the methods of splitting and decomposition. Computational schemes and algorithms can be obtained from the conditions that the functional $I(\varphi, \varphi^*, \mathbf{Y})$ is stable relative to variations of the grid components of the functions φ and φ^* at the nodes of the grid D_t^h. The reference time interval for realization of one cycle of assimilation of the Reanalysis data is 24 hours.

This process involves two sets of data of the 00Z analysis and one more (intermediate) set of data of the 12Z analysis (00:00 and 12:00 EMT).[1] The number of diurnal cycles of assimilation is set as an input parameter of the model of reconstruction. Assignment of the working interval $[0, t]$ within the time interval of reanalysis and separation of the domain D inside the domain D_t^R are carried out parametrically too. In accordance with the research problem, two modes for reconstruction of fields are provided: direct and inverse simulation.

Quasistationary "hydrodynamic" objects for interpolation

To form the vector of *a priori* estimates of the function of state of the atmosphere, let us use the procedures for reconstructing fields over the vertical coordinate assuming the quasistationary behavior. These procedures are based on analytical solutions and discrete approximations of the equations of the quasistationary model of atmospheric dynamics:

$$u\,\frac{\partial u}{\partial x} + v\,\frac{\partial u}{\partial y} + w\,\frac{\partial u}{\partial p} = f\,(v - v_{\mathrm{g}}) + K\,\frac{\partial^2 u}{\partial p^2}\,, \quad (2)$$

$$u\,\frac{\partial v}{\partial x} + v\,\frac{\partial v}{\partial y} + w\,\frac{\partial v}{\partial p} = -\,f\,(u - u_{\mathrm{g}}) + K\,\frac{\partial^2 v}{\partial p^2}\,, \quad (3)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial p} = 0, \quad (4)$$

$$\partial H/(\partial\,\ln p) = -\,R\theta. \quad (5)$$

Here u, v, and w are the components of the velocity vector along the coordinates x, y, and p, respectively; u_{g} and v_{g} are the components of the geostrophic wind; H is the geopotential; θ is the temperature; p is the pressure; f is the Coriolis parameter; K is the coefficient of turbulence along the vertical direction; R is the universal gas constant. The model is closed with the corresponding boundary conditions. Operators of the model enter into equations of the basic model of the atmosphere.

Equations (2)–(5) determine the structure of the objects of interpolation of the "hydrodynamic" type. They allow efficient reconstruction of field of meteorological elements (wind velocity components, temperature, geopotential) on the vertical p-sigma grid in D_t^h from the data of reanalysis on the grid D_t^R. Different modifications of numerical schemes can be applied here. An example of interpolation equations for the geopotential and temperature can be found in Ref. 5. These equations are closed up with the use of equation of hydrostatics (5). The procedure of interpolation for the horizontal components of the velocity vector with the use of the data on geopotential is based on analytical solutions of equations (2) and (3). The method for construction of such solutions and their concrete form are given in Ref. 6. Once u and v are reconstructed, the vertical velocity w is calculated with the use of the equation of discontinuity (4) in p-σ coordinates together with the data of reanalysis on the field of vertical velocities. This calculation is performed in the assimilation mode. Interpolation along the horizontal coordinates is carried out with the procedures of at least the second order of the accuracy of approximation.

The fields reconstructed in such a way are used to calculate, on the D_t^h grid, the background values for mesoregional models of atmospheric dynamics. They are also used to form the vector of *a priori* estimates of the function of state in the second term in Eq. (1). The procedures of vertical interpolation collectively take part in formation of the operator S for transition from D_t^h to D_t^R in the first term of Eq. (1).

Lake Baikal is one of the natural objects, for studying which the information model of the atmosphere of this type is used.[3] Comparative estimates of the scales of process formation in the regional atmosphere and in the lake indicate that characteristic time scales for the lake exceed the characteristic scales of the period of predictability of numerical models of the atmospheric circulation. This fact favors a preferable use of the information models wherever it is possible.

The model of an admixture transfer in the atmosphere in combination with the information model of the atmosphere has been used as a basis for scenario calculations of the influence areas and the functions of pollution danger of Lake Baikal and its region due to atmospheric transfer from pollution sources in the Northern Hemisphere. The calculated results confirm the conclusions on the value of characteristic scales in space and time in the atmosphere–lake system and on the influence of Sayano-Altaiskii cyclogenesis on the character of transfer in the southern regions of Siberia.[4]

Conclusion

The information model resolves the problem on predictability of the model describing the behavior of the atmosphere. It possesses the properties of a usual hydrodynamic model and allows reconstructing the structure of fields of the function of state of the atmosphere with a preset resolution for any region of the Earth and in any time interval within the time interval of reanalysis. The first version of the model of this type has been implemented.

Acknowledgments

This work was partially supported by the Russian Foundation for Basic Research (Grants Nos. 97–05–96511 and 98–05–65318) and the Program of Integrated Research of the Siberian Branch of the Russian Academy of Sciences (IG SO RAN–97 No. 30).

References

1. E. Kalnay, M. Kanamitsu, R. Kistler, et al., "ull. Amer. Meteor. Soc. **77**, 437–471 (1996).
2. V.V. Penenko, "ull. NCC. Num. Model. in Atmosph., No. 4, 32–51 (1996).
3. V.V. Penenko and E.A. Tsvetova, Atmos. Oceanic Opt. **11**, No. 6, 514–516 (1998).
4. V.V. Penenko and E.A. Tsvetova, Atmos. Oceanic Opt. **12**, No. 6 (1999), in press.
5. Yu.M. Liberman, Meteorol. Gidrol., No. 10, 106–108 (1986).
6. K. Kao, J. Appl. Meteor. **20**, 386–390 (1981).