FORECASTING THE CHARACTERISTICS OF THE ATMOSPHERIC OPTICAL CHANNEL ON THE BASIS OF A PHYSICAL-STATISTICAL APPROACH

V.S. Komarov, A.N. Kalinenko and S.A. Mikhailov

Institute of Atmospheric Optics,
Siberian Branch USSR Academy of Sciences, Tomsk

Received July 21, 1988

This paper considers a possible approach to the predictive evaluation of the optical state of the atmospheric channel. First, temporal changes in the optically active components of the atmosphere are calculated using physical and statistical methods. One then uses this information to predict optical characteristics of the atmospheric channel, which are employed as a database for remote sensing systems.

Research on the climate and environment of the Earth developing actively in a new direction based on remote optical sensing. It is difficult to imagine today how these problems might be successfully solved without using these data. But standing in the way of wider and more efficient use are significant obstacles, associated in particular with the problem of reliable interpretation of measurement data obtained by remote sounding, and with taking power losses of the optical radiation along a propagation path into account.

It is evident that to solve these problems it is necessary to have a priori information about the vertical distribution of such optically active components of the atmosphere as pressure, air temperature, humidity, ozone and others. The dimensionality n of the vector of optically active components $x_{\langle n \rangle}$ used for the solution of inverse problems and radiation correction of remote sensing results is determined at the design stage for the corresponding measurement system, and it depends on the methods used to handle the information obtained. It should be emphasized that not only the dimensionality of the vector X but also the type of a priori information plays an important role in solving the problems of remote optical sensing.

Modern approaches allow for the use of different types of information in the various stages of development and operation of optical measuring systems, i.e. climatic, predictive, and actual data. Thus, climatic information, usually represented by certain parameters of standard models of the atmosphere, can be used for modelling and long-term planning of system utilization. At the same time, predictive data (sometimes together with actual information, i.e., with observational results) can be used in the adoption of remote sensing results and short-term planning of experiments. Although the problems of providing optical measuring systems with the climatic information represented by the relevant models (including optical) were solved long ago[1–3], the problem of providing these systems with forecasting data still remains unsolved. This is related to the fact that the prediction of atmospheric optical channel characteristics, taking into account consideration regional peculiarities of altitude distributions of the optically active components of the atmosphere and different temporal scales for forecasting each of the physical components is one of the most complicated and difficult problems of atmospheric optics. By an atmospheric optical channel we mean an atmospheric volume in which optical radiation of a certain type (narrow- or broadband) propagates. The geometrical parameters of the channel are determined by the properties of the radiation source, the receiving aperture, and the geometry of the experiment. Its physical characteristics are determined by the interaction between the radiation and atmospheric components. The solution of the prediction problem is greatly complicated by the lack of systematic global observations of the optical characteristics of the atmosphere which are even more constable than the optically active components. It is also complicated by the poor state of our knowledge of physical relationships between meteorological quantities and atmospheric transmission in different spectral bands.

In spite of these problems, however, certain requirements have recently emerged which are necessary for a scientific formulation and solution of the problems involved in predicting the optical state of the atmosphere, though in doing so, one must choose appropriate approaches and methods. In the present paper, the authors have tried to give an account of some methodological aspects of solving this problem as applied to medium- and long-term prediction, i.e., 3–5 to 15 days or 2–3 weeks to some months, respectively. The problem of medium- and long-term prediction of the optical state of the atmospheric channel is solved via a physical statistical approach to the calculation of temporal changes in optically active components of the atmosphere, with subsequent evaluation of the required optical characteristics.

We emphasize here that the choice of a physical statistical approach is determined by the following:

– the theoretical forecasting limit (using hydrodynamic prediction scheme) is 1–2 weeks for synoptic scale processes, and at most 4 weeks for the phenomena on a planetary scale[4];

– existing hydrodynamical schemes do not enable us to take undersampled processes into account, and do not involve all of the optically active components;

– modern statistical methods enable one to identify periodic and quasiperiodic components fairly reliably, which in turn enables one to choice the most informative predictors and optimal forecast periods.

That is why in the literature[5–7] necessity of developing simplified approaches based on the use of the physical statistical methods is often emphasized. This paper presents a discussion based on this approach.

METHODOLOGICAL BASIS

The approach discussed in Ref. 8 will be used as a methodological basis for a physical and statistical method of forecasting the optically active components. In according with the results obtained in Ref. 8, we shall assume that the formation of temporal structure of variations $X_l(t)$ of the local climatic system l is contributed to by many external (principally G, and partly g) and internal factors (W), which, on the one hand, are typical of this system, and, on the other hand, are components of the global climatic system L. Mathematically, this can be expressed by the following relation[8]:

$$X_1(t)=\sum^{L} f_1(G^2_{1,t})+\sum^{L} f_2(G_{1,t}) + \sum^{L} f_3(g_{m,t})+\sum^{L} f_4(W^2_{k,t}), \qquad (1)$$

which shows the contributions of the global L and local l components to the overall local variations $X_l(t)$ of the climatic system in the region under study. Thus, the first term on the right-hand side of Eq. (1) determines the contribution of the main fluctuations with annual periods, which appear in the "ocean-atmosphere" as a result of the temperature contrasts between equatorial and polar zones. The second term in this equation also determines the contribution of the main fluctuations, but with periods from weeks to months. These fluctuations are due to the displacements of the atmospheric centers of action, of the planetary frontal zones, motion of long waves, and so on. The third term in Eq. (1) takes into account the contribution of particular fluctuations with periods from several hours to some days, induced by cyclonic activity, that is, by the formation and movement of cyclones and anticyclones. And finally, the fourth term describes the contribution of fluctuations characteristic of the local climatic system l itself.

Thus the methodological basis consists of the following: the temporal structure of the variations of local climatic system is produced both by external and internal factors, with the main fluctuations (components of the first G^2 and of the second G terms) determining the fundamental picture of the variations of these systems, while the particular g and internal W^2 fluctuations only transform it, determining instability and geographical peculiarities in the main picture of variations.

Taking into account, as well as the fact that the combined influence of the internal and external factors on the $X_l(t)$ very often has the form of a law or trend[8], we held to the following main principles while solving the problems of forecasting:

1. the short-term forecast of the optically active components of the atmosphere must be coordinated with the forecast of the climatic background;
2. the forecasting techniques to be constructed must be based on the use of averaged characteristics (and not individual values of meteorological quantities) and of altitude fields presented in the form of some basic functions, allowing one to filter out high frequency "noise" in order to separate out the long period components;
3. the prognostic model must take into account the presence of physical geographical differences in the structure of local variations of the climatic system.

These principles were followed when constructing the technique for statistical analysis and the prognostic models of vertical distribution of optically active components of the atmosphere suggested by the authors, as applied to the solution of problems involved in forecasting the optical state of the atmospheric channel. This technique is discussed below.

THE MAIN STEPS IN THE STATISTICAL ANALYSIS AND CONSTRUCTION OF PREDICTIVE MODELS

The block diagram of a step-by-step treatment of the data on the altitude distribution of the optical conditions optically active components of the atmosphere is shown in Fig. 1. In accordance with the proposed technique, the statistical analysis and the construction of predictive models of the optically active components of the atmosphere can be separated into the following steps.

Step. 1. First of all, we single out the main physical parameters of the atmosphere determining changes in its optical characteristics in the given region, taking account the spectrum of optical radiation and the direction in which it propagates. Then, using multidimensional statistical analysis, one calculated the characteristics of the statistical structure of the vertical distribution of physical parameters of the atmosphere which include, first of all, thermodynamic quantities (temperature, air, pressure, and humidity) and the concentrations of minor gaseous components. The most important of these characteristics are vectors of the average values and standard deviations, covariance and correlation matrices, and empirical orthogonal functions (eigenvectors and eigenvalues of the covariance matrices).

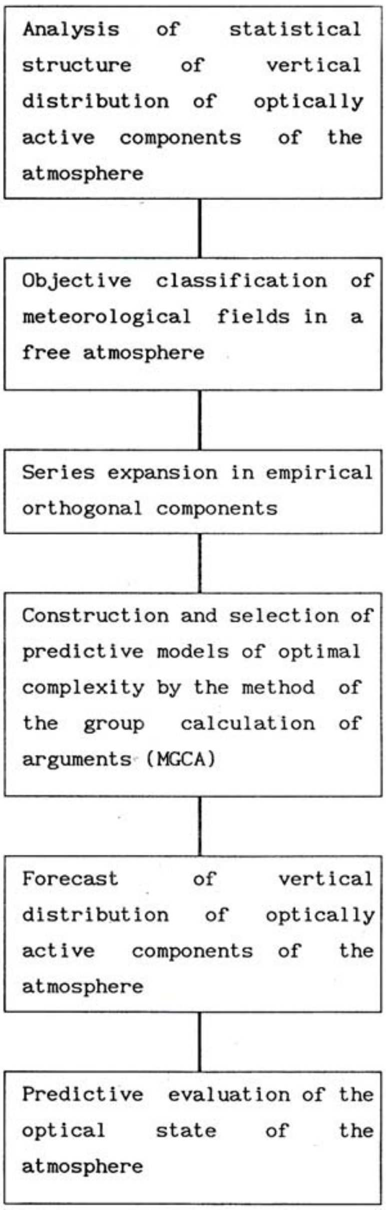

FIG. 1. Block diagram of statistical analysis of the optical conditions of the atmospheric channel.

Step 2. In the second step, using empirical orthogonal functions, one classifies the fields of optically active components of the atmosphere affecting the propagation of optical radiation in the given spectral band. The method of objective classification is based on the evaluation of similarity between the fields of variations of one of the optically active components in different regions of the globe which can occur due to the atmospheric processes of planetary and synoptic scales. A detailed description of the methods is given in Ref. 1. An overview of the results and methods of objective classification of the optically active components of the atmosphere, along with some results of the objective zoning of the northern hemisphere relative to the "pressure-temperature-humidity-ozone" complex is given in Ref. 3.

The use of objective classification of the fields optically active components in solving the problems of forecasting the state of an atmospheric optical channel enables one to divide the entire globe into a number of quasihomogeneous areas (from the point of view of variations). The use of quasihomogeneous areas in turn enables one to develop predictive models that take into account regional peculiarities of the atmospheric processes of different scales, and to take into consideration thereby the physical and geographical differences in the structure of local variations of the climatic system.

Step 3. As we have already noted, the methods being constructed must use as predictors the models of altitude fields presented in the form of some basic functions enabling one to filter out high frequency "noise" and separate out the long-period components. Natural orthogonal components (NOC) were used for this purpose, which allowed us to restrict the description of the fields of pressure, temperature, humidity and ozone to the first m terms in the series expansion in NOC. The values of m are 2 or 3 for the pressure, 3 to 5 for the temperature, 3 or 4 for the humidity and 5 to 7 for the ozone (the order of the covariance matrices of humidity is $p = 15$, for the remaining quantities $p = 27$). It should be noted that errors in the series expansion representation of the fields of optically active components in terms of finite number of NOC do not exceed in each particular case the measurements errors.

Starting with this, in the third step one forms the time series of the first m coefficients in the series expansion in the NOC $\alpha_k(t_i)$ for quasi homogeneous regions identified in the objective classification step. Here k is the number of NOC ($k = 1, \ldots, m$), i designates a particular observation ($I = 1, \ldots, N$) and N is the number of samples. In the succeeding two steps, while constructing and choosing optimal predictive models, the coefficients $\alpha_k(t_i)$ are used as predictors. This enables us to keep the representation error of the fields of optically active components at the level of the measurement errors, and to filter out more changeable processes on meso- and microscales. If one takes the latter into consideration, then instability of the evaluation of the parameters of the predictive models will occur.

Step 4. In the fourth step, for the same quasi homogeneous regions one must construct and choose the predictive models of optimal complexity. In doing so, the time series of coefficients of the expansion in NOC $\alpha_k(t_i)$ are used as the initial information.

To construct and choice optimal predictive models we have used the methods of self-organization of the models. A description of these methods can be found in papers (9)–(10). We should only like to note here that for problems of such a level of complexity, these methods are fairly efficient, but they do not rule out the use of other methods of forecasting. The main peculiarities of the methods of self-organization of the models (particularly, the method of group calculation of arguments-MGCA) are as follows:

1. Self-organization of the models similar to regression analysis refers to experimental methods of modeling, since it is based on the analysis of tables of the observation data obtained in either a passive or active experiment. It should be noted, that self-organization of physical and predictive models is also possible even if the initial data are very noisy. At present, the MGCA algorithms enable one to reconstruct a physical model of the object even in the case when the noise in the measurement data exceeds the regular signal by several-fold[11].

2. In the case of self-organizing models, the structure of predictive models is not specified a priori, but is chosen from among a number of computer-generated models according to the selection criteria chosen by an expert. In this procedure a certain number of points (models) close to the coordinate origin are chosen in criterion space. The final model selection procedure consists of choosing the best model according to the main criterion.

3. Since any internal criterion used to compare models (calculated using all the points in the table of initial data, used as predictors) leads to an incorrect rule the more complicated the model, the more accurate it is the selection of self-organizing models is made using external criteria constructed according to the experimental data which have not been employed to generate the probable models. By the principle of self-organization while the structure of the model is gradually made more complicated, the values of the external criteria first decrease and then increase, which means that there exists a minimum ! of a criterion which indicates the model of I optimal complexity.

4. These methods give us an opportunity to forecast complicated processes with incomplete a priori (i.e., without measuring many important arguments). Modeling on the basis of incomplete information is antithetical to the idea of increasing the information basis before summing up the maximum number of all the influences.

5. The objective nature of self-organization methods. Experts (the authors of the model) can influence the result of the forecast only by changing the criteria of selection, but otherwise they do not interfere in the self-organization of the model. The self-organization is intended to reduce a priori information introduced into the computer modeling of the system as much as possible. An operator should only feed observational data into the computer and define the class of fundamental functions and the criteria in the most general form, and in some cases, he would take part in the specification of the model.

To construct and choose models of optimal complexity we have used the MGCA algorithms to identify a harmonic trend with incommensurate frequencies (in one row or in many rows[12]) in order to predict 3–15 steps ahead in the table of initial vertical profiles of temperature, pressure, humidity; we also used the MGCA algorithm with a linear autoregressive model to predict vertical ozone profiles.

Step 5. In the fifth step, on the basis of models of optimal, complexity each of

$$\delta X_p(t+\Delta t)=\sum_{k=1}^{m} a_k(t+\Delta t)F_{kp}, \qquad (2)$$

obtained as a result of the forecast are used to reconstruct the predicted profiles of the optically active components using the formulas

$$X_p(t+\Delta t)=\bar{X}_p+\delta X_p(t+\Delta t), \qquad (3)$$

where $\delta X_p(t+\Delta t)$ and $\bar{X}_p$ are the predicted variation and the average of the component X at the level p, F_{kp} is an element of the eigenvector F_k at level p, and m Is the number of terms in the expansion.

The prognostic profiles of the optically active components of the atmosphere obtained as a result of the reconstruction are used in the next step to predict the optical state of the atmospheric channel within the limits of the quasihomogeneous zone chosen.

Step 6. For most problems concerned with the utility of remote sensing systems, it is quite enough to have predictive data on the attenuation coefficient or atmospheric transmission functions for a source of narrow band quasimonochromatic or wideband radiation. But, as we have seen, the problem on direct forecasting of these characteristics is very difficult at present. On the other hand, advances, the development of computers, and corresponding data systems have made it possible to carry out correct calculations of these characteristics for certain spectral bands using known altitude profiles of the optically active components of the atmosphere[13,14]. The development of fast, efficient computational[16] enables one to efficiently calculate beforehand the state of a specific atmospheric channel (in a cloudless atmosphere) using the predicted altitude profiles of temperature, pressure, humidity. and gas components. Extinction by aerosols is calculated with the help of models[2,16].

CONCLUSION

The method suggested in this paper enables one to obtain beforehand for different time intervals the optical state assessments of the atmospheric channels in a cloudless atmosphere provided that necessary information on the optically active components' vertical profiles is available. This method has been implemented at the Institute of Atmospheric Optics of the USSR Academy of Sciences in the form of corresponding data bases and specialized software for the ES1055M computer. Tests of the method have been carried out using 10 years' data obtained at aerological and ozonometrical stations in London, Buffalo, Resolute, and Hoenpasenberg, and have demonstrated good reliability and efficiency of this approach. Some results obtained using these methods will be presented in further publications.

REFERENCES

1. V.E. Zuev and V.S. Komarov, *Statistical Models of the Temperature and Gaseous Components of the Atmosphere*, (D. Reidel Publishing Company, Dordrecht, 1987).
2. V.E. Zuev and G.M. Krekov, *Optical Models of the Atmosphere*, (Gidrometeoizdat, Leningrad,, 1986).
3. V.S. Komarov and V.A. Remenson, Optika Atmosfery, **1**, No. 7, 3 (1988).
4. K.Ya. Kondrat'ev, Global Climate. Scientific and Technical Results. Series of Meteorology and Climatology, Moscow, **17**, 313 (1987).
5. *Long-Range Meteorological Forecasts*, (Gidrometeoizdat, Leningrad, 1985).
6. V.E. Prival'skii, *Climatic Variability, Stochastic Models, Forecasting, Spectra,* (Nauka, Moscow, 1985).
7. G.N. Chichasov, *Monthly Weather Forecasts, their Present Status and Prospects*, (All-Union Research Institute of Hydrometeorological Information-World Meteorological Center, Obninsk, 1984).
8. A.A. Isaev, Proc. All-Union Research Institute of Hydrometeorological Information – World Meteorological Center, **140**, 104, (1987).
9. A.G. Ivakhnenko and I.A. Myuller, *Self-Organization of Forecasting Models*, (Tekhnika, Kiev, 1985, VEB Verlag Technik, Berlin, 1984).
10. A.S. Ivakhnenko and V.S. Stepashko, *Noise Immunity of Simulation*, (Naukova Dumka, Kiev, 1985).
11. A.S. Ivakhnenko and V.S. Stepashko, *Numerical Simulation of the Noise Immunity of Multicriterion Model Selection,* Avtomatika, No. 4, **26** (1982).
12. *Reference Book of Standard Simulation Programs*, ed. by A.G. Ivakhnenko, (Tekhnika, Kiev, 1980).
13. O.K. Voitsekhovskaya, V.E. Zuev, and Vl.G. Tyuterev, *Atmospheric Molecular Spectroscopy Information Retrieval Systems*, Opt. Atmos., **1**, No. 3, **3** (1988).
14. V.S. Komarov, A.A. Mitsel', S.A. Mikhailov, et al., *Software and Data Support for Problems of Molecular Spectroscopy*, Opt. Atmos. **1**, No. 5, 84 (1988).
15. A.A. Mitsel', V.P. Budenko, and K.M. Firsov, *Appriximate Methods for Calculating the Absorption Functions of Overlapped Lines*, Opt. Atmos., **1**, No. 2, 45 (1988).
16. G.M. Krekov and R.F. Rakhimov, *Optical Models of Atmospheric Aerosol*, Publications of Tomsk Division, Siberian Branch of the Acad. Sci. USSR, Tomsk (1986).