MATHEMATICAL SIMULATION OF THE TIME SERIES OF MONITORING DATA ON NATURAL PROCESSES

V.M. Loginov and Yu.A. Kalush

Tuva Institute of Complex Development of Natural Resources,
Siberian Branch of the Russian Academy of Sciences, Kyzyl
Received July 6, 1995

Time series (TS) of data on natural processes are considered. The universal approach to mathematical simulation of time series based on the idea of dynamical switches is developed. The strict mathematical statement of the problem is given in addition to qualitative reasons. The ways of constructing the mathematical models of time series are shown. Advantages of the approach are illustrated by the examples connected with global changes in the Earth climate. The prediction of further behavior of the studied time series is discussed.

1. INTRODUCTION

When monitoring natural processes one, as a rule, deals with time series (TS) describing the temporal behavior of the object under investigation. Time series determine the behavior of the object and reflect the variety of all forces and fields affecting it. Since the whole complex of physical processes determines every time series, in order to understand the mechanisms of their influence on the object under study, it is important to create adequate mathematical models of the time series. That approaches are most valuable which allow one to approach time series corresponding to different natural processes from the common positions. For example, it may be time series of long-term observations of temperature, pressure, humidity of atmospheric air, temporal variations of one or another population, or temporal variations of solar or other radiation.

In this paper we present the universal approach to mathematical simulation of time series. The approach is based on some simple facts that are common for all time series arising at processing the data on natural processes. In particular, the characteristic feature of such time series is the fact that there are the time intervals, on which the parameter under investigation increases, and the intervals where it decreases. The law by which the parts of increase and decrease alternate can be random, determined, and also their mixture. The approach is based on the idea of dynamic and stochastic "switches" of the regimes at the points of the derivative sign alternation. The first results of such a direction were published in Refs. 1 and 2. The possibilities of the approach are illustrated (in Sec. 3) by examples of time series describing the global warming of climate in the current century, especially, the time series are considered of the "global change" of temperature and carbon dioxide concentration in the atmosphere (data from Refs. 3 and 4). Simple four-parameter mathematical models that approximate the real time series with the error not greater than 6% are also presented in these papers. The prediction capabilities of the time series models obtained are briefly discussed, as well as the possibilities of developing systems of automated mathematical simulation of the time series created on their basis.

2. DESCRIPTION OF THE METHOD

2.1. Essence of the method

It is convenient to illustrate the idea of the proposed method for processing and modeling the time series on the basis of the schematic time series describing the behavior of some object or process (Fig. 1). As was mentioned above, the time series are characterized by increase and decrease parts, with their duration and alternation following some law.

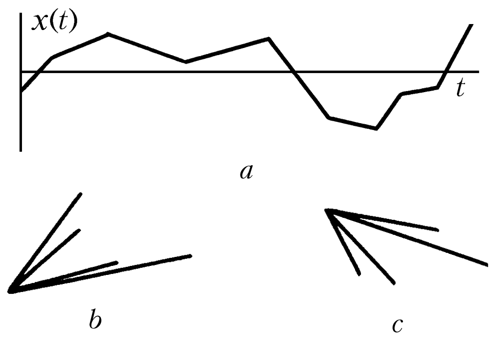

FIG. 1. Schematic time series (a), rays reflecting the increase and decrease periods of the time series (b and c).

Let us fix the points of switch from the regime of increase to the regime of decrease and consider the parts of the time series increase separately. Let us assume that the parts corresponding to the periods of increase has the same origin (Fig. 2*a*). The result is the "bundle of rays," where each ray is characterized by its angle and sets the increase of the function $x(t)$ on the corresponding time interval.

As a rule, the angle of the bundle opening is small for real time series, so that one always can select some effective (average) angle of the ray in one or another way. Then let us assume that the function $x(t)$ changes on the parts of the time series increase according to "the law of effective ray." Let us designate the partially smooth function obtained as $F^{(+)}(t)$. Let us accept that this function describes the time series $x(t)$ on the parts of its increase.

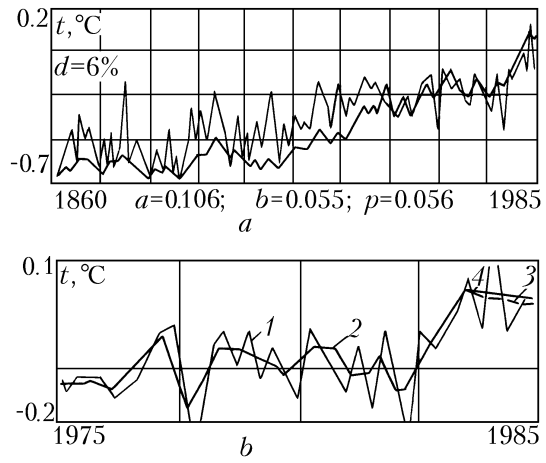

FIG. 2. Temporal series corresponding to the global change of temperature and its model (a); results of prediction of the time series of global change of temperature (b). Curve 1 shows actual data, curve 2 is the model; curve 3 is the result of prediction, and curve 4 is the result of averaging real data over the period of prediction.

Repeating the reasoning for the time intervals where the time series decreases (Fig. 2*b*), we obtain the function $F^{(-)}(t)$ that characterizes the time series $x(t)$ on the parts of its decrease to a certain accuracy. As a result, the time series $x(t)$ at the time of observation is approximated by some effective time series $F^{(+)}(t) + F^{(-)}(t)$. As to the construction, the approximation accuracy depends on the way of selecting the "effective rays".

The effective time series constructed in such a way will allow one to predict the behavior of the experimental time series for some period of time, if the law of the alternation of the regimes of increase and decrease is known (see below). When constructing a mathematical model of a real time series, the main problem of the mathematical statement is the choice of functions $F^{(+)}(t)$ and $F^{(-)}(t)$.

2.2. Mathematical statement of the problem

From the formal point of view, the majority of time series under study are the one-dimensional functions of time. However, if take the variable $x(t)$ as a variable characterizing the state of the natural medium at some point, then $x(t)$ becomes a multidimensional vector (let us call it the vector of the natural medium state). Its components are, for example, field of the atmospheric air velocities, concentration of one or another admixture, temperature, pressure, humidity, radiation intensity, etc. There are many time intervals $T_i^{(+)}$ for each of the components $x_i(t)$ of the vector of state $\mathbf{x}(t)$, where the variable $x_i(t)$ increases. Let the dynamics of the component $x_i(t)$ at the times $t \in T_i^{(+)}$ be described by the equation

$$\dot{x}_i = F_i^{(+)}(\mathbf{x}(t),\, t). \qquad (1)$$

The time series $x(t)$ within time intervals $T_i^{(-)} = T \setminus T_i^{(+)}$ is described by the partially decreasing function that is the solution of some different equation:

$$\dot{x}_i = F_i^{(-)}\,(\mathbf{x}(t),\, t), \qquad (2)$$

where T sets the total time of observation of the characteristic under study, $x_i(t)$. The back slash «$\setminus$» means the operation of subtraction of sets; the functions $F_i^{(+)}$ and $F_i^{(-)}$ are some scalar functions of the vector argument $\mathbf{x}(t)$ and the time t. Let us set the initial conditions for the component of state $x_i(t)$ at $t = 0$; as $x_i(t = 0) = x_{i0}$. As is seen from Eqs. (1) and (2), the mechanisms of interaction of physical subsystems are involved into the scheme of construction through the influence of the components of the vector of state $\mathbf{x}(t)$ on each other.

The aforementioned system of equations is related to the case of concentrated dynamical systems, when the state of natural medium at fixed point is the single-parametric function of the temporal variable t. The vector of the state of natural medium in a more general statement depends both on time and spatial variables. In this case, Eqs. (1) and (2) should be interpreted as some operator equations for the vector of state. Then the dependence of the vector of state on the variables of the problem becomes multiparametric. Expansion of the approach to the case of the systems with distributed parameters will be stated in a separate paper.

One can write the system of equations (1) and (2) for the vector of state x in a more compact form. To do this let us introduce an indicator function $\alpha_i(t)$ so that $\alpha_i(t) = 1$ (or some constant) if $t \in T_i^{(+)}$, and $\alpha_i(t) = -1$ if $t \in T_i^{(-)}$. As a result, the system (1) and (2) is reduced to the equation of the form:

$$\dot{x}_i = \frac{1}{2}\,[F_i^{(+)}(x,\,t) + F_i^{(-)}(x,\,t)] +$$

$$+ \frac{1}{2}\,\alpha_i\,[F_i^{(+)}(x,\,t) - F_i^{(-)}(x,\,t)] \qquad (3)$$

with the initial condition $x_i(t = 0) = x_{i0}$. As is seen from Eq. (3), the resulting behavior of the system described by the vector of state $\mathbf{x}$ is the result of mixing the two dynamics. These dynamics are set by Eqs. (1) and (2) for each component. Depending on the properties of the sets $T_i^{(+)}$ and $T_i^{(-)}$, this mixing has determined the stochastic or mixed nature. Now using the qualitative scheme stated in the Section 2.1, let us take some functions $F^{(+)}$ and $F^{(-)}$ obtained by some averaging of the time series under study, $x_i(t)$, over the parts of increase and decrease as functions $F_i^{(+)}(\mathbf{x}(t), t)$ and $F_i^{(-)}(\mathbf{x}(t), t)$, respectively. The equation (3) for such a choice of the function $F^{(+)}$ and $F^{(-)}$ is the mathematical model of a real time series $x_i(t)$.

3. SIMULATION OF THE SPECIFIC TIME SERIES

As was mentioned in the previous section, the representation of a time series by a mathematical model constructed by mixing of stochastic and dynamical states depends on what functions $F^{(+)}(\mathbf{x}, t)$ and $F^{(-)}(\mathbf{x}, t)$ have been chosen on the interval of the study of the time series. We use the following ideas for choosing these functions. Numerous examples of the processes occurring in living and inanimate nature show the S–like temporal behavior, when the parameter under study increases, initially, according to a power or an exponential law, then the increase slows down, and the curve comes to the established regime and weakly depends on time during some interval. As a rule, physical reasons that restrict the increase of the characteristics are various nonlinearities in the dynamics of the system and its interaction with other subsystems. The time interval, on which the established regime of the system behavior is observed for real systems, is often finite. It can be followed by a sharp decrease of the parameter under study due to the development of non-linear effects also. The examples of such a behavior can be a number of kinetic processes in solid bodies, in particular, the processes of magnetizing and electroconductivity in magnetic and conductive materials,[5,6] the effects of electromagnetic radiation propagation in condensed media.[7,8] In living nature, the dynamics of any population, including the mankind, has analogous behavior.[9,10]

For the simplicity of simulation and to make results more descriptive, let us take the function $F^{(+)}(t)$ as a solution corresponding to the classic logistic model[9]

$$\dot{x} = ax - bx^2, \qquad (4)$$

where a and b are some positive numbers, i.e. the right-hand side of the logistic equation (4) is chosen as a function $F_i^{(+)}(\mathbf{x}(t),\,t)$. The solution of Eq. (4) at small t increases exponentially and reaches the stationary level $x_{\rm st} = a/b$ at a great time

$$x(t) = [a\exp(at)]/[c + b\exp(at)], \qquad (5)$$

where c is the constant determined by the initial condition $x(t_0 = 0) = x_0$. Let us simulate the regime of decrease by the exponentially decreasing function $x(t) = C\exp(-pt)$ that is the solution of the dynamic equation

$$\dot{x} = -px.$$

Obviously, the right–hand side of this equation sets the function $F_i^{(-)}(x,\,t) = -px$.

It will be shown below that, even for such a choice of simplest dynamic regimes, their mixing and use as an approximation of the real time series provides quite satisfactory accuracy.

Let us consider the parts of increase of the time series. Two points are known for each part, the beginning of the period $(x_1,\,t_1)$ and its end $(x_2,\,t_2)$. Obviously, one can draw an infinite number of the curves of the form (5) through two points $(x_1,\,t_1)$ and $(x_2,\,t_2)$, so let us use the least square method (LSM) to select the coefficients a and b and the constant c. LSM supposes minimization of the rms deviation

$$\sigma^2 = 1/n\sum_{i=1}^{n}\,(x_i - f(t_i))^2, \qquad (6)$$

where $f(t_i)$ is the value of the approximation function at the point t_i. Since Eq. (5) is not convenient for taking derivatives, let us transform it by means of the substitution $X = 1/x$, $\theta = \exp(-at)$. As a result, we obtain the quasilinear dependence

$$X = k\theta + m, \qquad (7)$$

where $k = c/a$ and $m = b/a$, which is more acceptable for minimization. Using Eqs. (6) and (7), we obtain the expressions that allow us to determine a and b for each period of the increase in the time series. Since linearization of the initial dependence gives some error, the coefficients a and b were recalculated in the vicinity $(a \pm 0.2a)$, $(b \pm 0.2b)$ for better accuracy. Then the coefficients a and b were averaged as mean weighted over all parts of the increase of the time series. The constant c was determined from the condition of sewing together the solutions at the initial points $x(t = t_1) = x_1$. The coefficient p and the constant C are determined unambiguously for each part of the decrease of the time series by the points of the beginning and the end of a decrease period. Then the coefficient p is similarly averaged over all parts of the decrease in the time series. By substituting the results obtained into Eq. (3), we obtain the mathematical model of a real time series

$$\dot{x} = (1/2)\,((a - p)x - bx^2) + $$
$$+ (1/2)\,\alpha(t)((a + p)x - bx^2), \qquad (8)$$

where $\alpha(t) = \pm 1$ at the points of the derivative sign alternation. Obviously, the model obtained depends on three parameters a, b, and p in an explicit form and on the fourth parameter μ, the frequency of switches of the regimes of increase and decrease, in an implicit form. Let us present the numerical values of these parameters for the time series under consideration.

–	a	b	p
Global temperature change	0.106	0.055	0.056
Carbon dioxide concentration	0.090	0.003	0.04

Shown in Fig. 2*a* is the time series of the “global changeB of temperature from the data of Ref. 3 and its mathematical model. Relative deviation of the model from the real series in the meaning of average values is 6%, and the relative deviations for each of 126 points considered vary from 40 to 0.2%, that is 17% on average. The real series is shown by the thin line, and the model is shown by a bold line.

The results of prediction of the series by means of the model (8) are shown in Fig. 2*b*. Prediction of the behavior of the time series is shown by dashed line in the right corner of the figure, the solid line is averaging of the real series values over the same time interval.

Figures 3*a* and *b* show the temporal series of the carbon dioxide concentration from the data of Ref. 4 as well as its mathematical model and the result of prediction. It is easy to notice a good agreement of the calculated behavior of the time series with actual data. As is seen from Figs. 2 and 3, the predicted behavior of the time series model does not leave the limits of accuracy of approximation of the real series at the time intervals of about 1.5 years.

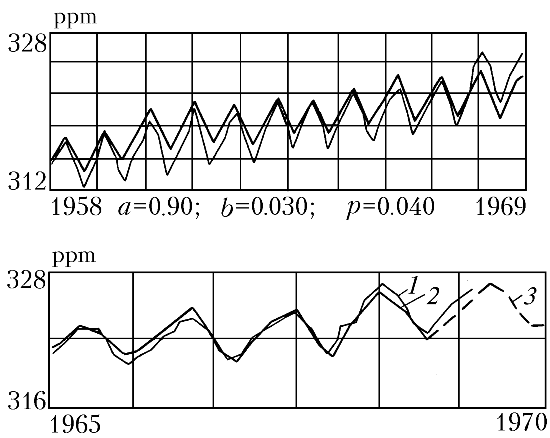

FIG. 3. Temporal series of the carbon dioxide concentration in the atmosphere (a); results of prediction of the time series of the carbon dioxide concentration in the atmosphere (b): curve 1 is actual data, curve 2 is the model, and curve 3 is the result of prediction.

4. DISCUSSION

The proposed approach to mathematical simulation of temporal series allows one to avoid many difficulties arising at the traditional approaches to simulation of complex time series. It is usually accepted in the traditional approach that one can describe the dynamics of the time series by systems of nonlinear dynamic equations, often supposing the explicit time dependence of coefficients of these equations.. When simulating the complex objects, the state of which is described by nonlinear equations, the additional time dependence of the coefficients of equations leads to the complex mathematical problems. The problems often arise already at the stage of the proof of theorems of the existence of solutions and their classification. The above approach to simulation of the time series uses very simple mathematical models.

We have selected the solutions of logistic equation and linear equation of the first order as simulating functions. It is clear that the class of functions, by means of which one can simulate the time series based on an idea of mixing dynamic and stochastic states, is not limited by this fact. In particular, such functions can be polynomials. Let us note that when simulating the temporal series we fix the points of the derivative sign alternation, thus obtaining some average characteristic times of increase and decrease of the time series. Such times, together with the characteristic time of the time series derivative sign alternation, are the fundamental parameters for describing the real time series. By means of them one can extrapolate the results of simulation to the times when the observations are absent.

As applied to the time series considered in this paper, the accuracy of predictions is within the limits of the accuracy of simulation (<6%) at the times not less than the characteristic time of the derivative sign alternation. The simplicity of simulation and possibility of extrapolating allow one to say about the possibility of creating an automated system of mathematical simulation of the natural processes described by the time series. Another approach to simulation of time series based on the stochastic distribution of the points of switches of the regimes and the choice of time series sample the most close to reality from the many possible ones is also possible. We will discuss the results of prediction and stochastic simulation in our future papers.

ACKNOWLEDGMENT

This work was supported by the Russian Foundation of Fundamental Research (Grant No. 93–01–12269).

REFERENCES

1. V.M. Loginov, Modeling, Measurement, and Control **53**, No. 3, 57–63 (1995).
2. V.M. Loginov and Yu.A. Kalush, Sibirskii Ekologicheskii Zh., No. 3 (1995).

3. J. Leggette, ed., *Global Warming: Greenpeace Report* (Russian translation) (State University Publishing House, Moscow, 1993), 272 pp.
4. B. Nebel, *Science of Environment: How Is the World Organized* [Russian translation] (Mir, Moscow, 1993), Part 1, 424 pp.
5. L.D. Landau and E.M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1974).
6. S. Chandrasekhar, *Liquid Crystals* [Russian translation] (Mir, Moscow, 1980), 344 pp.
7. A.S. Gurvich, A.I. Kon, V.L. Mironov, and S.S. Khmelevtsov, *Laser Radiation in the Turbulent Atmosphere* (Nauka, Moscow, 1976), 275 pp.
8. V.P. Lukin, *Atmospheric Adaptive Optics* (Nauka, Novosibirsk, 1986), 248 pp.
9. M. Bygon, J. Harper, and K. Townsand, *Ecology. Individuals, Populations, and Associations* (Mir, Moscow, 1989), Part 1, 667 pp.
10. M. Bygon, J. Harper, and K. Townsand, *Ecology. Individuals, Populations, and Associations* (Mir, Moscow, 1989), Part 2, 478 pp.