Applied problems of nonlinear oscillation and wave theory

Izv. VUZ «AND», vol.11, № 3, 2003

THREE SUBPROBLEMS OF GLOBAL MODEL RECONSTRUCTION FROM TIME SERIES AND THEIR PECULIARITIES

T. Dikanev, D. Smirnov, V. Ponomarenko, and B. Bezruchko

We consider three main subproblems of global reconstruction of dynamical models from time series: selection of dynamical variables, selection of model function, and estimation of model parameters. Special techniques for their solution are presented. Their applications and prospects of the further development of empiric modeling methods are discussed. The approaches are illustrated in numerical and acoustic experiments.

1. Introduction

A traditional way of obtaining a mathematical model of a complex phenomenon from the first principles cannot often be realized in practice. Then, experimental data may become the main source of information about a system under investigation and problem of an *empiric* model construction may arise. Since observations of real-world processes are presented, as a rule, in the form of time series (discrete ordered sequences of observable values), the problem is called *modeling from time series*. It is important in physics, meteorology, medicine and physiology, etc. Since 1980s various methods for constructing deterministic low-dimensional models in the form of difference equations (maps) [1-3] and ordinary differential equations (ODEs) [4-16] have appeared in the framework of nonlinear dynamics. In particular, significant contribution to this field has been made by V.S. Anishchenko and his team [8-11].

In general, the problem of modeling from time series can be formulated as follows.

• There is a system of our interest («an object»).

• One picks out some quantities $\eta_1,\ldots,\eta_k$, which characterize the processes occurring in the system and which can be measured experimentally (they are called observables).

• A time series of these quantities (i.e. the finite sequence $\{\vec{\eta}(t_i)\}_{i=1}^{N}$, where $\vec{\eta}(t_i)=(\eta_1(t_i),\eta_2(t_i),\ldots,\eta_k(t_i))$, $t_i=i\Delta t$, Δt is a sampling interval) is measured.

• It is known that the object possesses a set of properties $\{P_1,\ldots,P_M\}$.

Based on the time series, **it is necessary** to construct a dynamical model capable of reproducing this time series and as many of the properties $\{P_1,\ldots,P_M\}$ as possible. Models are constructed in the form of differential equations (1) or discrete maps (2):

$$d\mathbf{x}/dt = \mathbf{F}(\mathbf{x}(t),\mathbf{c}), \qquad (1)$$

$$\mathbf{x}_{n+1} = \mathbf{F}(\mathbf{x}_n, \mathbf{c}), \tag{2}$$

where $\mathbf{x}=(x_1,\ldots,x_D)\in R^D$ is a state vector of the model, $\mathbf{F}$ is a smooth function, $\mathbf{c}\in R^P$ is a parameter vector, t is continuous time, n is discrete time. If the function $\mathbf{F}$ is expressed analytically in terms of elementary functions (in a closed form) for the entire phase space, then the model is called global. Further we consider only *global* models.

The success of modeling depends on several factors. It is exceedingly important to select properly *a model structure* that involves the choice of dynamical variables x_k (that is the relations between dynamical variables x_k and observables η_j) and choice of the form of the function $\mathbf{F}$.

So, the first subproblem of time series modeling is selection of dynamical variables and reconstruction of their time courses from the observed time series data. If there are too many observables then one should specify a subset of them to be used as model variables. If the number of observables is not sufficient for model construction or they cannot be used directly, then different combinations of available data are employed. Very popular methods are sequential derivatives and time delays ([5] and [6], respectively), both of them rest upon the celebrated Takens' results and their generalizations [17]. However, different ways of obtaining dynamical variables realizations, which are based on a priori information about the system under investigation or some peculiarities of its dynamics [8], may prove to be more appropriate for modeling. In Section 2 we present a technique for the selection of the best set of dynamical variables for modeling, which allows simultaneous convenient testing for nonlinearity.

The second subproblem is to specify the form of function $\mathbf{F}$. Algebraic polynomial is a standard recommendation [6], even though often inefficient [13]. To make polynomial more feasible different methods for spurious terms detection and exclusion were suggested which work well for a special situations [2,14,18-20]. In Section 3 we present a new method for spurious terms detection.

Third subproblem is technological: to estimate model parameters $\mathbf{c}$ (usually the least-squares routine is used). Finally, an obtained model should be validated. But even if the model is not sufficiently adequate, model coefficients may have their own value and serve for the characterization of the system. In Section 4 we consider such a situation where coefficients of a model map describing phase dynamics are used to solve an important problem of coupling characterization [21] and suggest extension of the known technique for the case of short and noisy time series.

2. Selection of dynamical variables

As it has been already mentioned, to construct model equations in the form $\mathbf{y}(t)=\mathbf{F}(\mathbf{x}(t))$ from a time series $\{\vec{\eta}(t_i)\}$, one forms, first of all, the series of state vectors $\{\mathbf{x}(t_i)\}$. Then, the time series of quantities to enter the left-hand side of model equations $\{\mathbf{y}(t_i)\}$ is obtained from the time series $\{\mathbf{x}(t_i)\}$ according to the chosen model type:

- via numerical differentiation of $\{\mathbf{x}(t_i)\}$ for ODEs, since $\mathbf{y}(t)=d\mathbf{x}(t)/dt$;
- via the shift of $\{\mathbf{x}(t_i)\}$ along the time axis for discrete maps, since $\mathbf{y}(t_i)=\mathbf{x}(t_{i+1})$.

Finally, the form of the function $\mathbf{F}$ is specified and its parameters are estimated.

Voluntary dynamical variables selection can make approximation of the dependencies $y_j(\mathbf{x})$ with a smooth function extremely problematic [22,23] or even make these dependencies many-valued. Here, we describe the method for assessing suitability and convenience of the selected variables $x_1,\ldots,x_D$ for constructing a global dynamical model. It is based on testing the time series $\{\mathbf{y}(t_i)\}$ and $\{\mathbf{x}(t_i)\}$ for single-valuedness and continuity of each dependency $y_j(\mathbf{x})$ in the entire region of an observed motion. It is crucial here that we use local characteristics rather than the averaged ones as in [24].

2.1. Description of technique. If a dependency $y(\mathbf{x})$ is single-valued and continuous in a domain V, then the difference $|y(\mathbf{x})-y(\mathbf{x}_0)|$ tends to zero when $\|\mathbf{x}-\mathbf{x}_0\|\to 0$ for each $\mathbf{x}_0 \in V$. In practice, violation of this condition may be viewed as a sign of many-valuedness or discontinuity of the dependency $y(\mathbf{x})$. Since the length of an observable time series is finite, the above-mentioned limit cannot, strictly speaking, be found. However, it is possible to trace a tendency in variations of the quantity $|y(t_i)-y(t_j)|$ when the vectors $\mathbf{x}(t_i)$ and $\mathbf{x}(t_j)$ are made closer and closer, down to a certain *finite* distance. Given sufficiently large amount of data N, high accuracy of measurements, and low noise level, the distance $\|\mathbf{x}(t_i)-\mathbf{x}(t_j)\|$ can be made sufficiently small for each local region of observed motion.

The technique of testing consists in the following (Fig. 1, *a*). The domain V containing the set of vectors $\{\mathbf{x}(t_i)\}_{i=1}^{N_0}$ is partitioned into identical hypercubic boxes of the size δ. All boxes containing at least two vectors are selected. Let us denote them $s_1, s_2, \ldots, s_M$. The difference between the largest and the smallest values of y inside a box s_k is called *local variation* $\varepsilon_k = \max_{\mathbf{x}\in s_k} y(\mathbf{x}) - \min_{\mathbf{x}\in s_k} y(\mathbf{x})$. The largest local variation $\varepsilon_{max} = \max_{1\le k\le M} \varepsilon_k$ and its graph $\varepsilon_{max}(\delta)$ are used as the main characteristics of the investigated dependency. Suitability of the considered quantities $\mathbf{x}$ and y for global modeling is assessed using the following considerations [25].

• If a dependency $y(\mathbf{x})$ is single-valued and continuous, ε_{max} is sufficiently small for small δ and tends to zero for $\delta\to 0$ (Fig. 1, *b*, filled circles). The following statement is often correct: the less the slope of the graph $\varepsilon_{max}(\delta)$, the better are the dynamical variables for modeling.

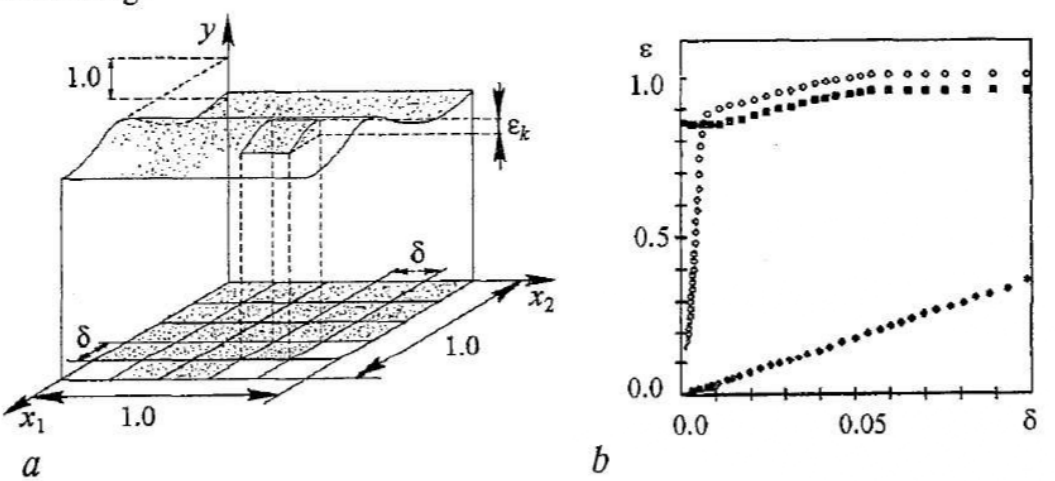

Fig. 1. (*a*) Illustration for a technique of testing a dependency $y(\mathbf{x})$ for single-valuedness and continuity in the case D=2. (*b*) Possible appearance of plots $\varepsilon_{max}(\delta)$ for different variants of dynamical variables

• If a single-valued and continuous dependency has a region of very steep slope (a «jump»), then ε_{max} remains rather big even for sufficiently small δ. However, further decrease in δ leads to decrease in ε_{max} and the graph $\varepsilon_{max}(\delta)$ exhibits a «kink» at the value of δ roughly equal to the size of the steep slope region (e.g., Fig.1,b, white circles). In such a case, the dependency $y(\mathbf{x})$ is difficult to approximate with a smooth function.

• If ε_{max} remains large and does not diminish for $\delta\to 0$ (Fig. 1, *b*, filled squares) then the considered variables are not appropriate for global modeling. Such situation may be related both to nonuniqueness of the dependency and high noise level.

2.2. Numerical example, refinement of technique and testing for nonlinearity. The above technique was already published and sufficiently illustrated previously [25]. Here, we describe briefly an approach to refinement of the technique and its use for assessment of nonlinearity of a dependence $y(\mathbf{x})$. Besides, we present the application of the refined technique to the analysis of a biological time series.

The procedure described in Section 2.1 is a technique with a *fixed-location* set of nonoverlapping boxes (independent of the distribution of data points). It has the disadvantage that a vector **x** lying near a box boundary is *never* compared to the close vectors from the neighboring boxes, but it *may* be compared to more distant vectors from its own box. It can lead to intensive oscillations in the $\varepsilon_{max}(\delta)$ for small δ in the presence of noise. The nonmonotony makes the assessment of the considered dynamical variables more difficult. An example of such situation is illustrated in Fig. 2, *c*, where the results of testing are presented for chaotic regime of the logistic map $x_{n+1}=\lambda-x_n^2$ at $\lambda=2.0$. The observable is $\eta_n=x_n+\xi_n$, where ξ_n is a sequence of independent identically uniformly distributed random values. We test the dependencies corresponding to the first iterate $\eta_{n+1}(\eta_n)$, to the second one $\eta_{n+2}(\eta_n)$, and to the third one $\eta_{n+3}(\eta_n)$ using the time series containing 1000 data points (see Figs 2, *a*, *b* for noise-free data).

The disadvantage of the technique may be obviated by using the set of overlapping boxes centered at the vectors of the time series (data-dependent location) instead of the fixed-location set of boxes. In other words, for each vector $\mathbf{x}(t_i)$ one should consider all its δ-neighbors, i.e. to calculate local variation of y in the box with the side 2δ centered at $\mathbf{x}(t_i)$. The number of considered boxes is then equal to the number of vectors N_0. The largest value of local variation obtained in such a way (let us denote it ε_{max}') monotonically decreases with decrease in δ. This advantage of the modified procedure is illustrated in Fig. 2, *c*, *d* for the above mentioned case of the logistic map.

Due to this advantage, the plot $\varepsilon_{max}'(\delta)$ is more reliable and informative. Note also, that the plot $\varepsilon_{max}'(\delta)$ is a straight line if the system under investigation is linear. Therefore the plot $\varepsilon_{max}'(\delta)$ can serve as a test for linearity. Its concavity indicates nonlinearity of the system under investigation (Fig. 2, *d*). As an example of the proposed techniques application to a complex real-world system, let us briefly consider testing of an acoustic time series. This is a digitized recording of the human voice (in fact, air pressure variations), which was done when a man was pronouncing the sound [a:]. Sampling frequency is 44.1 kHz. The recording length is 10000 data points. A dependence

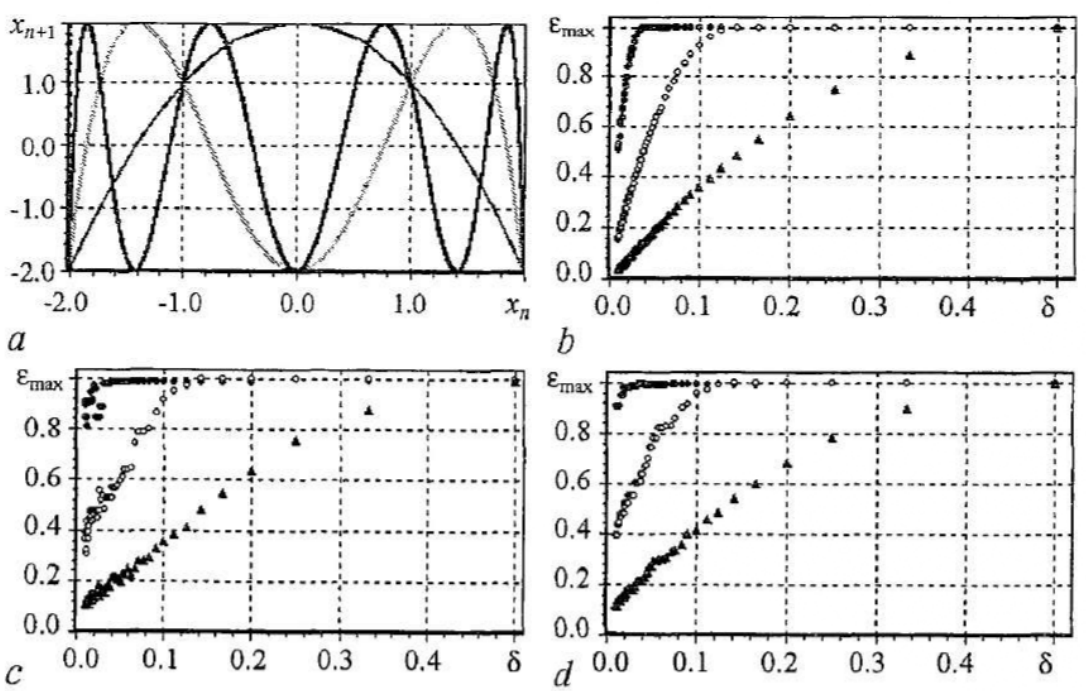

Fig. 2. Comparison of testing techniques in a numerical experiment. (*a*) The first, the second, and the third iterations of the chaotic logistic map. (*b*), (*c*) Results of testing with the fixed-set-of-boxes technique from noise-free and noisy data, respectively. (*d*) Results of testing with the modified technique from noisy data

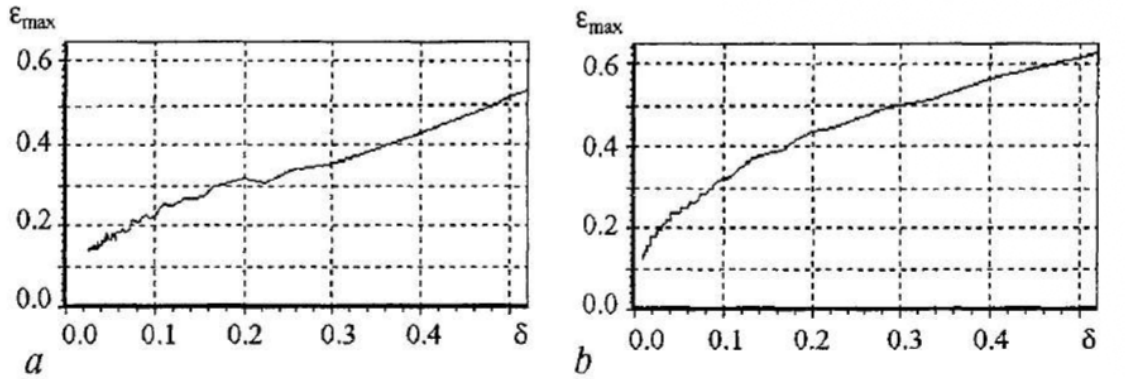

Fig. 3. Testing vocal time series described in Section 2.2. (*a*) The fixed-location set of boxes technique indicates many-valuedness. (*b*) The modified techniques indicates also nonlinearity

$\eta_{i+2}(\eta_i,\eta_{i-2},\eta_{i-4})$ is tested. We present the plot $\varepsilon_{max}(\delta)$ in Fig. 3, *a*; it does not indicate single-valuedness. Other conclusions can hardly be drawn from the figure because of the above-mentioned disadvantages of the fixed-location set of boxes technique. The modified technique leads to the monotone plot $\varepsilon_{max}'(\delta)$ (Fig. 3, *b*). It is easily seen that the dependence $\varepsilon_{max}'(\delta)$ is significantly concave that allow a conclusion about nonlinearity of the system under investigation.

3. Detection of spurious polynomial terms

Choice of the model function F is also very important problem, which comes into play after selection of dynamical variables. In the typical case of absence of detailed a priori information about proper function form, one usually uses algebraic polynomials relying upon many rigorous mathematical results (Weierstrass'theorem). But model with polynomials are often inefficient because of their very bad extrapolation properties that are determined by the presence of «spurious» terms (basis functions).

Theoretically, polynomial terms should be regarded as spurious if «true» values of their coefficients (coefficients of the «true» function expansion in a power series) are equal to zero. Detection with subsequent exclusion of several spurious terms from the model polynomial can lead to significant refinement of the model. But, different approaches to detection of spurious terms have been suggested: small absolute values of the corresponding coefficients [2], small values of the coefficients with respect to their standard deviation (Student's criterion), intensive variation of coefficients around zero when different parts of a transient time series are used for reconstruction [14], slight change of the approximation error when the term is excluded from the model [19]. Here we develop a new (and, in our opinion, more general) approach to detection of spurious terms.

Again, theoretically, rather typical situation is such that neither of terms is spurious. If the true function is, e.g., exponential, its expansion in a power series involves nonzero coefficients at each power of a variable. In such a situation, when time series is analyzed, adding of each term to a model structure would lead to decrease of the approximation error. Nonetheless, some of the terms are undesirable (practically spurious). We state that those terms are practically spurious which affect approximation errors only in a narrow domain of the phase space. We conjecture that such terms can be detected as those terms whose coefficients depend strongly on the distribution of the data points in the phase space. (If all coefficients slightly depend on the distribution of data points, one may reasonably guess that such model function describe an object not only for the domain explored by the observed training time series, but also in its neighborhood, that is the function has good extrapolation properties.)

To determine how strongly a coefficient value depends on the distribution of training data points, consider its change under variation of the weight function $p(x_i)$ if coefficients values α_k are found by minimizing the weighted squared sum of errors

$$s^2 = \Sigma_{i=1}^{N} p(x_i)[F(x_i) - \Sigma_{k=1}^{M} \alpha_k g_k(x_i)]^2. \quad (3)$$

Here $F(x_i)$ are true (observed) values of the approximated function; $g_k(x_i)$ are basis functions (terms). Weight function $p(x_i)$ is normalized to unity $\Sigma_{i=1}^{N} p(x_i)=1$.

If the set of basis functions is orthonormal, then coefficient values read

$$\alpha_k = \Sigma_{i=1}^{N} p(x_i) F(x_i) g(x_i). \quad (4)$$

Approximated function can be expressed as

$$F(t) = \Sigma_{k=1}^{M} \alpha_k g_k(x) + \tilde{F}(x), \quad (5)$$

where $\tilde{F}(t)$ is approximation error.

Let us consider now how the values of coefficients will change under slight variations of the weight function $p \rightarrow p' = p + \hat{p}$. Because of weight function normalization the variation satisfies

$$\Sigma_{i=1}^{N} \hat{p}(x_i) = 0. \quad (6)$$

Impose also the condition of smallness of variation in the form

$$\Sigma_{i=1}^{N} \hat{p}^2(x_i) = \varepsilon^2. \quad (7)$$

The change of coefficient in the linear approximation is

$$\Delta\alpha_k = \Sigma_{i=1}^{M} \hat{p}(x_i) g_k(x_i) \tilde{F}(x_i). \quad (8)$$

To estimate the intensity of coefficient variation consider $\Delta\alpha_k$ as functional of $\hat{p}$. For its maximum with supplementary conditions (6) and (7) we have

$$\Delta\alpha_{k\max} = \varepsilon[\Sigma_{i=1}^{N} (g_k(x_i)\tilde{F}(x_i) - (1/N)\Sigma_{j=1}^{N} g_k(x_j)\tilde{F}(x_j))^2]^{1/2}. \quad (9)$$

If we suppose that at the beginning the weight function was uniform for all training time series points, that is $p(x_i)=1/N$, then as a consequence of orthogonality of basis functions and approximation error, we have

$$\Delta\alpha_{k\max} = \varepsilon[\Sigma_{i=1}^{N} (g_k(x_i)\tilde{F}(x_i))^2]^{1/2}. \quad (10)$$

As a criterion for exclusion (or inclusion) of basis function we can use the ratio between the maximum possible change of coefficient value (10) and the coefficient value itself

$$C_k = [\Sigma_{i=1}^{M} (g_k(x_i)\tilde{F}(x_i))^2]^{1/2} / [\Sigma_{i=1}^{M} g_k(x_i) F(x_i)]. \quad (11)$$

Formula (11) was derived for orthonormal basis functions. In practice this is usually not the case. However when we decide if the basis function is spurious or not we can consider only its projection orthogonal to all other basis functions $g_k'(x)$. After this we can freely use the method described above. We don't need to calculate this projection in

explicit form. Let $\tilde{F}_k(x)$ be an error of approximation in the absence of the k-th basis function, then the orthogonal projection of the k-th basis function is

$$g_k'(x) = (\tilde{F}_k(x) - \tilde{F}(x))/\alpha_k', \tag{12}$$

where α_k' is coefficient for this projection, calculated with least squares method.

After substitution of (12) into (11) for coefficient instability we have

$$C_k = [\Sigma_{i=1}^{M}(\tilde{F}_k(x_i) - \tilde{F}(x_i))^2(\tilde{F}(x_i))^2]^{1/2}/[\Sigma_{i=1}^{M}(\tilde{F}_k(x_i) - \tilde{F}(x_i))F(x_i)]. \tag{13}$$

Above we talked about exclusion of spurious terms from initially large basis, but adding the most suitable functions can also optimize the basis functions set. We can choose them using the same methodology, according to minimal value of criterion (13).

Let us illustrate proposed method on test example. The realization of x variable from Rössler system in chaotic regime is used as time series. Fig. 4, *b* shows phase trajectory of this system reconstructed with time delay method. Model is constructed in the form

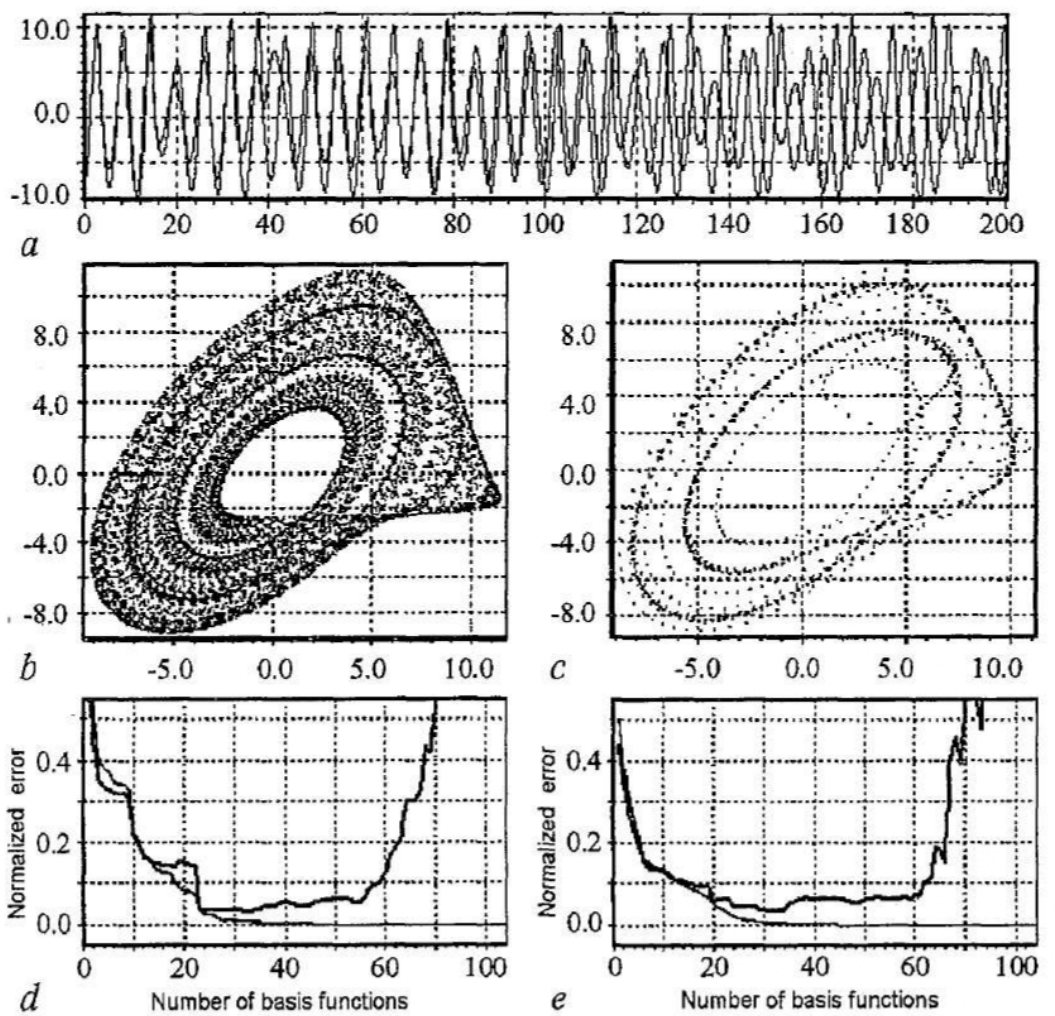

Fig. 4. (*a*) Comparison of test part of time series generated by Rössler system in chaotic regime with time series generated by optimized model. Good prediction for about 6 quasi-periods is observed. (*b*) Phase trajectory of Rössler system reconstructed with delay method from time series of x variable. (*c*) Phase trajectory generated by optimized model. (*d*) Error of approximation of training time series (thin curve) and error of test time series prediction (bold curve) as a function of basis functions number, added during process of basis functions set optimization with new method. (*e*) The same as (*d*) for old method

$$dx_1/dt = x_2,$$
$$dx_2/dt = x_3, \qquad (14)$$
$$dx_3/dt = f(x_1, x_2, x_3).$$

If we transform the Rössler equations into the form (14) we will have rational function f. We will try to construct the model in universal form using polynomial of the 6th order as function f. The short part of 200 data points (about 3 quasi-periods) is used as training time series. Let us add to initially empty basis such function for which the relative variation of coefficient (13) is minimal.

In Fig. 4, *d* the dependencies of training series approximation error (thin curve) and error of prediction of test time series (bold curve) on the number of added basis functions are shown. Errors are normalized by the standard deviation of the third derivative dx_3/dt of training time series. One can see that training series approximation error decreases monotonous, while the prediction error for test part of time series has minimum when the number of basis functions is 32 and it is equal to 0,032 (3.2% of standard deviation). At this moment the error on training series is 0.008. Before optimization (when we use full polynomial of the 6th order) the behavior of the model was divergent. After optimization the model generates stable trajectories and allows prediction of about 6th quasi-periods of test time series. In Fig. 4, *c* the phase trajectory generated by the model is shown. One can see that the trajectory is located in the same phase space domain as the trajectory of Rössler system, but the cycle of period 2 is eventually established.

In Fig. 4, *e* for comparison purposes the results of optimization of basis functions set with previously known method is shown. In this method we add to the basis such functions that allows maximal decrease of approximation error. One can see that the sequence of adding basis functions is different from proposed method. The best basis according to test series prediction error includes 33 functions. This best error is 0.034 (slightly worse than with new method), error of training time series approximation 0.006. The behavior of the optimal (from this point of view) model turns out to be divergent. So we can state that in this example new proposed method of optimization outperforms the well-known one.

4. Determining character of coupling between subsystems from time series

The problem of determining the presence and direction of interaction between two subsystems is very important in many fields, including physiology and medicine where interaction between human cardio-vascular and respiratory systems [26,27] and between different brain areas are of interest [28-30]. Thus, recently Rosenblum and Pikovsky suggested a very delicate and nice idea for characterization of weak coupling between subsystems from time series by estimating coefficients of a model map [21]. But their method works well for the case of very long time series (for «reasonable» noise level, time series should contain about $10^4 \ldots 10^5$ data points). In practice, nonstationarity of processes, impossibility to collect sufficient amount of data, and significant noise often require estimation of the interaction (coupling) characteristics under conditions of short observation interval. Here, we develop an extension of Rosenblum and Pikovsky approach to the case of short* and noisy time series. For a detailed consideration see [32].

* E.g., typical quasi-stationary segments of electroencephalogram (EEG) is about 5 seconds long [31]. EEGs are recorded at typical sampling rate of 200 Hz. Then, quasi-stationary segment contains about 10^3 data points. Roughly speaking, typical length of a short time series in practice is of the order of 10^3 data points.

4.1. Evolution map approach. The technique of Rosenblum and Pikovsky is based on construction of empiric model maps, describing *phase dynamics* of the subsystems, and is called *evolution map approach*. Having an original time series $\{\eta_{1,2}(t_i)\}_{i=1}^{N}$, where η_1, η_2 are observables, $t_i=i\Delta t$, $i=1,\ldots,N$, Δt is a sampling interval, one calculates the time realizations of phases $\{\phi_{1,2}(t_i)\}_{i=1}^{N_\phi}$ and construct a global model map, which characterizes the dependence of phase increments (over a time interval $\tau\Delta t$) on the phases of subsystems' oscillations, in the form

$$\Delta_{1,2}(t) \equiv \phi_{1,2}(t+\tau\Delta t) - \phi_{1,2}(t) = F_{1,2}(\phi_{1,2}(t),\phi_{2,1}(t),\mathbf{a}_{1,2}), \qquad (15)$$

where τ is positive integer; $\mathbf{a}_{1,2}$ are vectors of the coefficients of the functions $F_{1,2}$; $F_{1,2}$ are trigonometric polynomials:

$$F_1(\phi_1,\phi_2) = \Sigma_{i=1}^{L_1} a_{1,i} g_i(\phi_1,\phi_2), \qquad (16)$$

with $g_1=1$; $g_i=\cos(m_i\phi_1+n_i\phi_2)$ for even $i\geq 2$; $g_i=\sin(m_i\phi_1+n_i\phi_2)$ for odd $i\geq 3$; L_1 is the number of terms of the polynomial F_1. For $i\geq 2$, $m_{2i}=m_{2i+1}$ are nonnegative integer, $n_{2i}=n_{2i+1}$ are arbitrary integer, and by definition $m_1=1$, $n_1=0$.

Using the estimates of coefficients $\hat{\mathbf{a}}_{1,2}$, obtained from the time series via the least-squares routine, one computes intensities of influence of the second subsystem on the first one ($2\rightarrow 1$) $\hat{c}_1$ as

$$\hat{c}_1^2 = 1/(2\pi^2)\int_0^{2\pi}\int_0^{2\pi}(\partial F_1(\phi_1,\phi_2,\hat{\mathbf{a}}_1)/\partial\phi_2)^2 d\phi_1 d\phi_2 = \Sigma_{i=1}^{L_1} n_i^2 \hat{a}_{1,i}^2. \qquad (17)$$

Everything is similar for the influence of the first subsystem on the second one ($1\rightarrow 2$) $\hat{c}_2$. Directionality index is defined as $\hat{d}=(\hat{c}_2-\hat{c}_1)/(\hat{c}_2+\hat{c}_1)$. Since $\hat{c}_{1,2}\geq 0$, $\hat{d}$ takes the values only within the interval $[-1,1]$: $\hat{d}=1$ or $\hat{d}=-1$ corresponds to unidirectional coupling ($1\rightarrow 2$ or $2\rightarrow 1$, respectively), $\hat{d}=0$ for ideally symmetric coupling.

4.2. Short time series. For very long time series ($N\rightarrow\infty$) the estimates $\hat{c}_1$, $\hat{c}_2$, and $\hat{d}$ are unbiased and have practically no scattering, in other words, the method gives correct characterization of coupling. However, if the time series is short, the following important questions arise. Are the estimates $\hat{c}_{1,2}$, $\hat{d}$ biased or not? How can statistical significance of the results be estimated? To illustrate importance of the questions, let us consider a simple *demonstrative example*, when two subsystems are uncoupled and linear, that is a system of difference equations

$$\Delta_{1,2}(t) \equiv \phi_{1,2}(t+2\pi) - \phi_{1,2}(t) = 2\pi\omega_{1,2} + \varepsilon_{1,2}(t), \qquad (18)$$

where $\varepsilon_{1,2}$ are Gaussian random processes independent of each other with variances $2\pi D_{1,2}$. Obviously, correct values of coupling should be $c_1=c_2=d=0$ in this case.

We have carried out numerical experiment in the following way. Time realizations of original equations were simulated using the generator of pseudo-random numbers realized in the subroutine DRNNOR of the library IMSL. Initial conditions for each realization are random numbers $\phi_1(0)$, $\phi_2(0)$ distributed uniformly on the interval $[0,2\pi]$. We obtained 1000 short time realizations (1000 pairs of scalar time series) with the length $N_\phi=10^3$. The values of estimates $\hat{c}_{1,2}$ and $\hat{d}$ are computed for each of them. From the obtained sets of values we construct histograms.

The estimates $\hat{c}_{1,2}$ and $\hat{d}$ appear misleading. Their distributions are shown in Fig. 5. Thus, in the case of identical subsystems ($D_1=D_2$ and $\omega_1=\omega_2$) $\hat{c}_1$ is always positive and

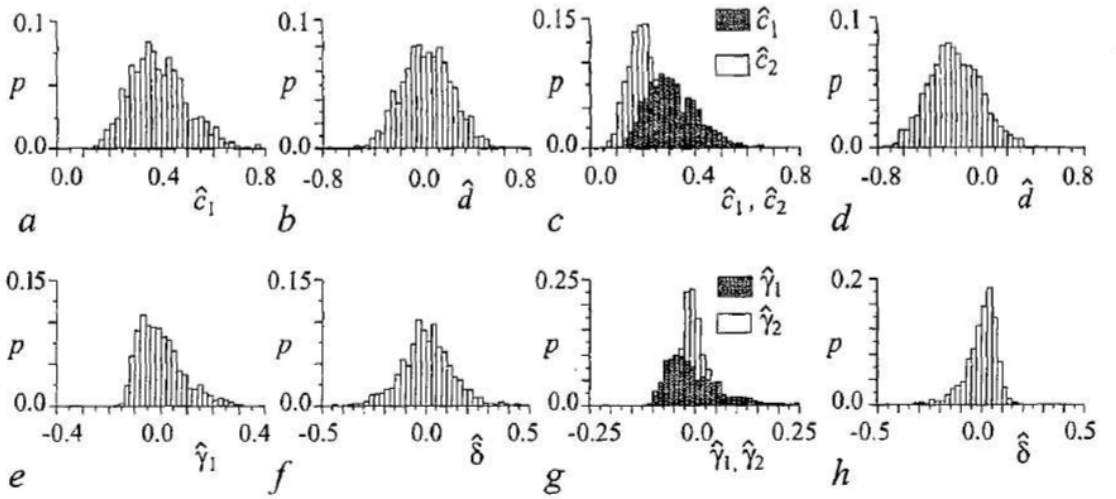

Fig. 5. Histograms for the estimates of coupling, constructed as a result of processing of 1000 time realizations of the equations (16) with $\omega_1=\omega_2=1.0$ (*a*), (*b*) $\hat{c}_1$ (a biased estimate) and $\hat{d}$ for identical subsystems with $D_1=D_2=0.4$. (*c*), (*d*) $\hat{c}_1$, $\hat{c}_2$ (exhibit different biases) and $\hat{d}$ (exhibits negative bias) for subsystems with different noise levels $D_1=0.4$, $D_2=0.1$. (*e*), (*f*), (*g*), and *h*) $\hat{\gamma}_1$, $\hat{\gamma}_2$, and $\hat{\delta}$ (unbiased estimates) for the situations corresponding to (*a*), (*b*), (*c*), and (*d*), respectively

takes sufficiently large values (Fig. 5, *a*), i.e. it is a biased estimate for $c_1=0$; $\hat{d}$ is unbiased, but exhibits quite a large scattering; even the values of $\hat{d}=\pm 0.4$ are encountered quite often (Fig. 5, *b*). Thus, it is very probable to get spurious indication of the presence of interaction from a single realization. The situation becomes even more complicated when subsystems are nonidentical. It is illustrated in Fig. 5, *c*, *d* for the case $D_1>D_2$, $\omega_1=\omega_2$. The estimates $\hat{c}_{1,2}$ are biased, bias in $\hat{c}_1$ being greater (Fig. 5, *c*). This leads to biasedness of $\hat{d}$ whose values are *systematically* less than zero (Fig. 5, *d*). Hence, predominant influence (*2→1*) is diagnosed, even though coupling is absent in reality.

4.3. Corrections to evolution map approach and novel unbiased estimates of coupling. By careful analytic consideration of the problem we found out the cause of biases and developed corrected estimates of coupling [32]. Novel unbiased estimate of c_1^2 is the quantity

$$\hat{\gamma}_1 = \hat{c}_1^2 - \Sigma_{i=1}^{L_1} n_i^2 \hat{\sigma}_{\hat{a}_{1,i}}^2, \qquad (19)$$

where $\hat{\sigma}_{\hat{a}_{1,i}}^2$ are unbiased estimates of variances $\sigma_{\hat{a}_{1,i}}^2$. Derivation of $\hat{\sigma}_{\hat{a}_{1,i}}^2$ is not trivial. Under some simplifying assumptions (noise $\varepsilon_{1,2}(t)$ are Gaussian, coupling between subsystems and their individual nonlinearities are very weak) we obtain the following analytic expression for $\hat{\sigma}_{\hat{a}_{1,i}}^2$:

$$\hat{\sigma}_{\hat{a}_{1,i}}^2 = (2\hat{\sigma}_{\varepsilon_1}^2/N)[1 + 2\Sigma_{j=1}^{\tau-1}(1-j/\tau)\cos((m_i\hat{a}_{1,1}+n_i\hat{a}_{2,1})j/\tau)e^{-j(m_i^2\hat{\sigma}_{\varepsilon_1}^2+n_i^2\hat{\sigma}_{\varepsilon_2}^2)/2\tau}], \; i>1, \qquad (20)$$

where $\hat{\sigma}_{\varepsilon_{1,2}}^2$ are estimates of noise variances, their derivation is straightforward. Normalized index d is replaced by nonnormalized quantity $\delta \equiv c_2^2 - c_1^2$, whose unbiased estimator is $\hat{\delta} = \hat{\gamma}_2 - \hat{\gamma}_1$.

To estimate reliability of numbers $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ obtained from a single realization, one needs the estimate of variance of $\hat{\gamma}_1$ (we denote it as $\hat{\sigma}_{\hat{\gamma}}^2$). After some algebra and

experimentation, we derive a semiempiric formula [32] for $\hat{\sigma}^2_{\hat{\gamma}_1}$ in terms of estimates $\hat{a}_{1,i}$ and $\hat{\sigma}^2_{\hat{a}_{1,i}}$ derived earlier, we do not present it here for the sake of brevity. Since $\hat{\gamma}_1$ has a skewed distribution for low order trigonometric polynomials $F_{1,2}$ typically used, we take asymmetric expression $[\hat{\gamma}_1-\alpha\hat{\sigma}_{\hat{\gamma}_1}, \hat{\gamma}_1+\beta\hat{\sigma}_{\hat{\gamma}_1}]$ as a confidence interval for c_1^2. We found constants α and β empirically to provide necessary significance level; 95% confidence interval is obtained if α=1.6, β=1.8. Conclusion about the presence of influence $(2\rightarrow 1)$ can be drawn with probability of error 0.025 provided

$$\hat{\gamma}_1 - \alpha\hat{\sigma}_{\hat{\gamma}_1} > 0. \qquad (21)$$

The degree of belief can be adjusted by changing α (and, hence, confidence interval width).

Conclusion of predominant direction of interaction can be drawn after estimation of the variance of $\hat{\delta}$. Its «good» estimate is $\hat{\sigma}^2_{\hat{\delta}}=\hat{\sigma}^2_{\hat{\gamma}_1}+\hat{\sigma}^2_{\hat{\gamma}_2}$. Our experiments show, that α=1.6 also provides approximately 95% confidence interval for δ in the form $\hat{\delta} \pm \alpha\hat{\sigma}_{\hat{\delta}}$. More accurately, the values

$$\hat{\gamma}_2 - \alpha\hat{\sigma}_{\hat{\gamma}_2} > 0 \quad \text{and} \quad \hat{\delta} - \alpha\hat{\sigma}_{\hat{\delta}} > 0 \qquad (22)$$

allow the statement about influence $(1\rightarrow 2)$ with probability of error 0.025 (similarly for $(2\rightarrow 1)$).

Results of application of the proposed estimates $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ to the above mentioned example (17) are presented in Fig. 5, *e-h*. Systematic errors in $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ are absent. Fig. 6 demonstrates usefulness of the interval estimates to ensure reliable conclusions of coupling direction.

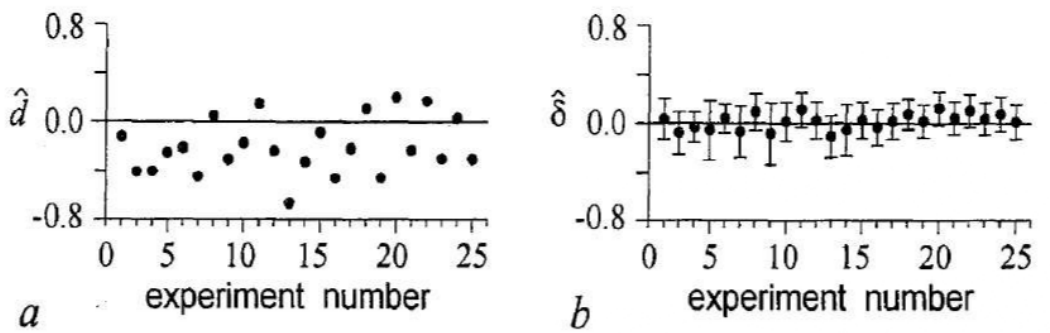

Fig. 6. Estimates of coupling for example (17), results obtained for the first 25 of the 1000 time realizations of the subsystems with different noise levels D_1=0.4, D_2=0.1. (*a*) $\hat{d}$ takes predominantly negative values. (*b*) $\hat{\delta}$ (circles) takes negative as well as positive values, estimated confidence intervals are shown as error bars and, as a rule, include zero

5. Summary

This paper illustrates some important details of the procedure of constructing mathematical model from a time series. Namely, three main subproblems are selected and their peculiarities are shown. Special techniques for better solutions of two of them are proposed:

- preliminary testing of time series of dynamical variables, which provides the variants which are the most suitable for modeling and allows convenient testing of experimental dependencies for nonlinearity;
- a procedure of model structure optimization, which allows elimination of

spurious terms from the polynomial model, whereby model performance can be significantly refined.

Finally, we have developed an approach to estimation of intensity and directionality of coupling between two subsystems in the case of short and noisy time series. Under certain assumptions (nonlinearity of subsystems and coupling between them are weak), *unbiased* estimates of intensity and directionality of interaction *provided with confidence* intervals are derived. In our opinion, suggested estimates are applicable for wide range of real-world processes, including signals of biologic origin when it is important to analyze short time series segments due to nonstationarity.

The work was supported by the RFBR (grant №. 02-02-17578), CRDF (Award REC-006), and the Russian Ministry of education.

References

1. *Crutchfield J.P. and McNamara B.S.* Equations of motion from a data series // Complex Systems. 1987. Vol. 1. P. 417-452.

2. *Judd K., Mees A.I.* On selecting models for nonlinear time series // Physica D. 1995. Vol. 82. P. 426-444.

3. *McSharry P.E. and Smith L.A.* Better nonlinear models from noisy data: Attractors with maximum likelihood // Phys. Rev. Lett. 1999. Vol. 83, №. 21. P. 4285-4288.

4. *Baake E., Baake M., Bock H.J., and Briggs K.M.* Fitting ordinary differential equations to chaotic data // Phys. Rev. A. 1992. Vol. 45, № 8. P. 5524-5529.

5. *Brown R., Rulkov N.F., and Tracy E.R.* Modeling and synchronizing chaotic systems from time-series data // Phys. Rev. E. 1994. Vol. 49, № 5. P. 3784-3800.

6. *Gouesbet G., Letellier C.* Global vector-field approximation by using a multivariate polynomial approximation on nets // Phys. Rev. E. 1994. Vol. 49. P. 4955-4972.

7. *Hegger R., Kantz H., Schmuser F., Diestelhorst M., Kapsch R.-P., and Beige H.* Dynamical properties of a ferroelectric capacitors observed through nonlinear time series analysis // Chaos. 1998. Vol. 8, № 3. P. 727-754.

8. *Janson N.B., Pavlov A.N., and Anishchenko V.S.* One method for restoring inhomogeneous attractors // Int. J. of Bifurc. and Chaos. 1998. Vol. 8, № 4. P. 825-833.

9. *Anishchenko V.S. and Pavlov A.N.* Global reconstruction in application to multichannel communication // Phys.Rev. E. 1998. Vol. 57, № 2. P. 2455-2457.

10. *Anishchenko V.S., Pavlov A.N., and Janson N.B.* Global reconstruction in the presence of a priori information // Chaos, Solitons & Fractals. 1998. Vol. 8. P. 1267-1278.

11. *Pavlov A.N., Janson N.B., and Anishchenko V.S.* Reconstruction of dynamical systems // Radiotekh. i Electron. 1999. Vol. 44, № 9. P. 1075-1092.

12. *Anosov O.L., Butkovskii O.Ya., and Kravtsov Yu.A.* Reconstruction of dynamical systems from chaotic time series: brief review // Izv. VUZ. Applied Nonlinear Dynamics. 2000. Vol. 8, № 1. P. 29-51.

13. *Bezruchko B. and Smirnov D.* Constructing nonautonomous differential equations from a time series // Phys. Rev. E. 2001. Vol. 63. 016207.

14. *Bezruchko B., Dikanev T., and Smirnov D.* Role of transient processes for reconstruction of model equations from time series // Phys. Rev. E. 2001. Vol. 64. 036210.

15. *Timmer J., Rust H., Horbelt W., and Voss H.U.* Parametric, nonparametric and parametric modelling of a chaotic circuit time series // Phys. Lett. A. 2000. Vol. 274. P. 123-134.

16. *Horbelt W., Timmer J., Bunner M.J., Meucci R., and Ciofini M.* Identifying physical properties of a CO_2-laser by dynamical modeling of measured time series // Phys. Rev. E. 2001. Vol. 64. 016222.

17. *Sauer T., Yorke J.A., and Casdagli M.* Embedology // J. Stat. Phys. 1991. Vol. 65, № 3-4. P. 579-616.

18. *Le Sceller L., Letellier C., and Gouesbet G.* Structure selection for global vector field reconstruction by using the identification of fixed points // Phys. Rev. E. 1999. Vol. 60, № 2. P. 1600-1606.

19. *Aguirre L.A., Freitas U.S., Letellier C., Maquet J.* Structure-selection techniques applied to continuous-time nonlinear models // Physica D. 2001. Vol. 158. P. 1-18.

20. *Small M., Judd K. and Mees A.* Modeling continuous processes from data // Phys. Rev. E. 2001. Vol. 65. 046704.

21. *Rosenblum M.G. and Pikovsky A.S.* Detecting direction of coupling in interacting oscillators // Phys. Rev. E. 2001. Vol. 64. 045202(R).

22. *Letellier C., Maquet J., Le Sceller L., Gouesbet G., and Aguirre L.A.* On the non-equivalence of observables in phase space reconstructions from recorded time series // J. Phys. A: Math. Gen. 1998. Vol. 31. P. 7913-7927.

23. *Letellier C., Aguirre L.A.* Investigating nonlinear dynamics from time series: The influence of symmetries and the choice of observables // Chaos. 2002. Vol. 12, № 3. P. 549-558.

24. *Kaplan D.T.* Exceptional events as evidence for determinism // Physica D. 1994. Vol. 73. P. 738-748.

25. *Smirnov D., Bezruchko B. and Seleznev Ye.* Choice of dynamical variables for global reconstruction of model equations from time series // Phys. Rev. E. 2002. Vol. 65. 026205.

26. *Bracic-Lotric M. and Stefanovska A.* // Physica A. 2000. Vol. 283. P. 451.

27. *Janson N.B., Balanov A.G., Anishchenko V.S., McClintock P.V.E.* // Phys. Rev. E. 2002. Vol. 65. 036212.

28. *Arnhold J., Grassberger P., Lehnertz K., and Elger C.E.*// Chaos in Brain? / Eds K. Lehnertz, J. Arnhold, P. Grassberger, and C.E. Elger. World Scientific, 2000, p. 325.

29. *Rodriguez E. et al.* // Nature (London) 1999. Vol. 397. P. 430.

30. *Andrzejak R.G, Lehnertz K., Mormann F., Rieke C., David P., and Elger C.E.* // Phys. Rev. E. 2001. Vol. 64. 061907.

31. *Kaplan A.Ya.* // Uspekhi Fiziol. Nauk. 1998. Vol. 29(3). P. 35.

32. *Smirnov D. and Bezruchko B.* Estimation of interaction strength and direction from short and noisy time series // Phys. Rev. E. 2003, to be published.

Saratov State University
Institute of RadioEngineering
and Electronics of RAS, Saratov Branch

Received 04.09.03

УДК 517.9: 519.6

ТРИ ПОДЗАДАЧИ РЕКОНСТРУКЦИИ ГЛОБАЛЬНОЙ МОДЕЛИ ПО ВРЕМЕННЫМ РЯДАМ И ИХ ОСОБЕННОСТИ

Т. Диканев, Д. Смирнов, В. Пономаренко, Б. Безручко

Мы рассматриваем три основные подзадачи глобальной реконструкции динамических моделей по временным рядам: выбор динамических переменных, выбор функций модели и определение параметров модели. Представлены

специальные методы для их решения. Обсуждаются их приложения и перспективы дальнейшего развития методов эмпирического моделирования. Данные подходы иллюстрируются в численных и акустических экспериментах.

Smirnov Dmitry Alekseyevich was born in 1977. He studied physics in Department of Nonlinear Processes of Saratov State University (DNP SSU) and graduated in 1999. He performed his candidate dissertation work in DNP SSU and Saratov Department of Institute of Radio Engineering and Electronics of RAS (SD IRE RAS) and obtained the Ph.D. degree in 2001. He has received several international and Russian awards for young scientists. Now, he is a senior researcher in Laboratory of Nonlinear Dynamical Modeling of SD IRE RAS. His scientific interests include mathematical modeling of complex systems, time series analysis, and dynamical systems theory. He is an author of more than 45 publications including 14 papers in refereed journals.

Bezruchko Boris Petrovich was born in 1946. He graduated from Department of Physics, Saratov State University (1969). He is a doctor of Physics and Mathematics (he obtained this degree in 1995) and a professor in the Chair of Electronics, Oscillations, and Waves of Saratov State University. Field of his scientific activity involves experimental investigation of nonlinear phenomena in radiophysical and electronic systems, mathematical modeling from experimental observable time series. He published more than 100 papers in Russian and foreign journals.