Theoretical and Mathematical Physics, **142**(1): 128–137 (2005)

THE PROBLEM OF PROCESSING TIME SERIES: EXTENDING POSSIBILITIES OF THE LOCAL APPROXIMATION METHOD USING SINGULAR SPECTRUM ANALYSIS

I. A. Istomin,* O. L. Kotlyarov,* and A. Yu. Loskutov*

We present algorithms for singular spectrum analysis and local approximation methods used to extrapolate time series. We analyze the advantages and disadvantages of these methods and consider the peculiarities of applying them to various systems. Based on this analysis, we propose a generalization of the local approximation method that makes it suitable for forecasting very noisy time series. We present the results of numerical simulations illustrating the possibilities of the proposed method.

Keywords: time series, forecast, chaos, local approximation

1. Introduction

Analyzing the results of an experiment is based on processing the data obtained. In many cases, these data are time series, that is, chronologically ordered sequences of values of one or several measured quantities. Processing time series in order to extract useful information from them about system properties is a very important and interesting problem. But the main attention in many cases is given not to studying system properties but to forecasting the dynamics of the time series generated by the system. In meteorology for example, just the weather forecast for the nearest future is most important, whereas a more global problem of studying climate peculiarities has less importance in the short-term perspective. Therefore, in addition to studying the properties of the system generating a time series, predicting its subsequent trajectory is often an important problem.

We encounter prognosis problems not exclusively in meteorology: they are also relevant in geophysics where the earthquake forecast is among the main directions of investigation, in astrophysics when studying solar activity, in financial analysis when forecasting stock prices and indices, and so on. In these and other areas of investigation, methods for forecasting (or extrapolating) time series belonging to the ARMA (AutoRegressive Moving Average) [1], [2] class have long been used. Their main idea is to express subsequent values of a series through previous values. This is the most widespread approach, which is often used in situations where we have no other information about a system except that contained in previous values of the series.

Special forecast methods based on papers by Takens [3] have also been elaborated in the framework of dynamical system theory. They are primarily aimed at forecasting irregular (chaotic and quasiperiodic) stationary time series generated by complex nonlinear systems. But progress in nonlinear dynamics methods has demonstrated that the forecast problem is much more complicated and often falls outside any theoretical scheme. In particular, the processed series must be sufficiently long, and the noise component must be small.

*Moscow State University, Moscow, Russia, e-mail: Loskutov@moldyn.phys.msu.ru.

Translated from Teoreticheskaya i Matematicheskaya Fizika, Vol. 142, No. 1, pp. 148–159, January, 2005. Original article submitted November 11, 2003; revised June 3, 2004.

In this paper, we propose one effective tool for studying relatively short, noisy series. It is based on the local approximation (LA) method first described in [4] in relation to forecasting chaotic time series. In its original formulation, this method had several advantages over the traditional method of autoregression when forecasting irregular time series, but it has not yet been widely applied, mostly because of difficulties appearing when analyzing short, noisy time series.

In the present paper, we consider the possibility of increasing the effectiveness of the LA method for forecasting noisy time series by implementing a preliminary data filtration using singular spectrum analysis (SSA) [5]. The SSA method was developed in the framework of nonlinear dynamics for processing time series, but it is mainly used currently to determine their basic constituents [6] and to damp the noise [7], although there are original forecast algorithms based on this method [8].

There currently exist many methods for processing time series. Among them are wavelet analysis, flicker-noise spectroscopy, etc. (see, e.g., [9]). All these methods are sufficiently effective and allow processing and, to some extent, predicting system dynamics. The generalization proposed in this paper makes it possible to forecast chaotic and quasiperiodic series in the presence of random noise components.

2. Methods for processing time series

2.1. Method of delays. A milestone in the majority of approaches for processing a time series $\{x_1, \ldots, x_N\}$ is the construction of the set of delayed vectors $\mathbf{z}_i = (x_i, x_{i+1}, \ldots, x_{i+m})^{\mathrm{T}}$ [10]. In the methods for time series analysis developed in the framework of nonlinear dynamics, this is also the first and necessary step.

The method of delays passes from the initial one-dimensional (scalar) time series to a multidimensional (vector) representation similar to the one used in autoregression. Each multidimensional vector is there composed of a number m of successive values of the initial series. The result can be imagined as a set of "snapshots" of the series taken through a window sliding along the series such that only m consecutive terms of the series can be seen simultaneously through the window:

$$\mathbf{X} = \begin{pmatrix} \vdots & \vdots & & \\ f_{m+1} & f_{m+2} & & \\ \begin{bmatrix} f_m \\ \vdots \\ f_2 \\ f_1 \end{bmatrix} & \begin{bmatrix} f_{m+1} \\ \vdots \\ f_3 \\ f_2 \end{bmatrix} & \cdots & \begin{bmatrix} f_N \\ \vdots \\ f_{N-m+2} \\ f_{N-m+1} \end{bmatrix} \\ \downarrow & f_1 & & f_{N-m} \\ & \downarrow & & \vdots \end{pmatrix}, \tag{1}$$

where $f_1, f_2, \ldots, f_N$ are the values of the terms of the series at the instants $t = 1, 2, \ldots, N$. Each square bracket is a vector in the m-dimensional space of delays; the sequence of such vectors constitutes the matrix of observations $\mathbf{X}_{m\times(N-m+1)}$, where N is the number of terms of the initial series. This matrix, in each column of which we encounter parts of the same series shifted relative to each other, is the multidimensional representation of the initial scalar series in the space of delays.

The new result is that under special conditions, the space of delays can be considered the reconstruction of the phase space of a nonlinear dynamical system generating the time series (see the Takens theorem and its generalizations [3], [11]). We can therefore prove the possibility of describing a multidimensional system dynamics via the time series for the observable. Under special conditions, the possibility of describing and reconstructing the system dynamics in turn allows forecasting its future behavior [12].

Defining the reconstruction dimension. The passage to a multidimensional representation in the discrete case under consideration is described by the single parameter m, which is the dimension of the space of delays or the dimension of the embedding (reconstruction). The possibility of achieving an exact, reliable forecast relies strongly on choosing the value of this parameter properly. The restriction $m \geq 2d+1$, where d is the dimension of a system generating the series [3], [11], is rarely helpful for selecting the reconstruction dimension because d is often unknown a priori, whereas the problem of determining the system dimension based on the obtained data is very involved, especially in the case of short, very noisy series. Currently, the most popular algorithm for estimating the embedding dimension (and therefore the system dimension) is the Grassberger–Procaccia algorithm [13], [14], but even this algorithm turns out to be not very effective when processing short time series (less than 10^4 values).

The indicated difficulties in determining the reconstruction dimension can often be overcome in the framework of the forecast method under consideration. The given series is then split into two unequal parts, one following the other, and the shorter part is used to estimate the quality of the forecast made based on the other, longer part. We then take the dimension for which the best forecast is achieved as the optimal dimension for the given series.

2.2. Local approximation method. As mentioned above, autoregression methods belonging to the class of ARMA methods [1] are most frequently used nowadays.[1] The autoregression model of order p is

$$f_t = a_0 + a_1 f_{t-1} + a_2 f_{t-2} + \cdots + a_p f_{t-p} + \xi_t, \tag{2}$$

where f_t is the value of the series at the tth step and $\{\xi_t\}$ is a sequence of random quantities representing the white noise. In this case, to predict the subsequent trajectory of the series, we must first determine the autoregression order and then estimate the autoregression coefficients $\{a_0, \ldots, a_p\}$ based on the obtained data. The autoregression order is customarily chosen from the form of a particular autoregression function. The coefficients are estimated using all available data and are correspondingly assumed to be time independent. This approximation is therefore global.

Methods of global linear approximation provide a good approximation for a function so long as we have sufficient free parameters and the data array is sufficiently large to ensure stable estimates for these parameters. But when the function becomes sufficiently complex, we have no assurance that we can find a representation that can effectively approximate this function. We may then have a vicious circle: the more complex the function, the larger number of parameters we must estimate. This in turn means that we must increase the amount of data, and this increase may require introducing additional parameters into a model, and so on.

One way out of this circle is to use the LA method. Its basic idea is to split the whole domain of a function into several local domains and then construct approximating models and estimate the model parameters separately in each domain. These domains must be sufficeintly small to avoid "dramatic" changes of the function inside a domain. This allows applying simple (for instance, linear) models in each separate domain. The main condition for effectively applying the LA method is the proper choice of the local domain size or, which is practically the same, the number of neighbors. We must choose a number that suffices for the stability of the parameter estimate and such that adding a small number of new neighbors does not substantially affect the estimated parameter values.

The LA method [4] was the first method developed for forecasting time series that is based on the Takens theorem. This method also uses a representation of type (2) but has two basic differences from the autoregression method:

1. we estimate the coefficients in expression (2) for each local domain separately, and

[1]The recent rapid development of computing facilities has resulted in neural networks finding widening application in forecasting, but the ARMA method is still overwhelming in its range and intensity of application.

2. expression (2) may contain nonlinear terms, i.e., higher degrees of series values at preceding instants (the maximum degree is called the approximation order).

We use the reconstruction dimension as the quantity having the sense of the autoregression order.

Hence, the LA method is better grounded in choosing the "autoregression order" and is more flexible in using initial data because different sets of coefficients are used on different parts of the trajectory. We can therefore approximate essentially nonlinear functions by a set of piecewise-linear approximations.

We now briefly describe the basic steps in the algorithm realizing the LA method.

1. Choosing the local representation. This step includes evaluating the reconstruction dimension and constructing a multidimensional series representation, i.e., the matrix of delays $\mathbf{X}$.

2. Determining the neighborhood or the number of neighbors. As a rule, the neighborhood can be determined by choosing the number of neighbors N^s. For this, we segregate the "closest" columns in the space of delays (among the columns of the matrix $\mathbf{X}$). The simplest criterion of closeness is the following. Given the metric $\|\cdot\|$, the set $\{\mathbf{y}^t\}$ is the set of nearest neighbors of the vector $\mathbf{x}$ for a *given* number of neighbors N^s if the sum $\sum_{t=1}^{N^s}\|\mathbf{y}^t-\mathbf{x}\|$ is minimum. We note that although we consider the dynamics of variable changes, the closeness of $\mathbf{y}^t$ to the values of $\mathbf{x}$ does not imply closeness in time.

3. Choosing the approximation model and its identification. At this step, for an already determined reconstruction dimension, we choose the approximation order (an approximation that is either linear or square in the series values at preceding instants is customarily used):

$$x_1^{t+T}=f^{t+T}(\mathbf{x}^t)=\mathbf{P}_q(x_1^t,x_2^t,\ldots,x_m^t), \tag{3}$$

where $T=1,2,\ldots,q$ is the polynomial degree, i.e., the approximation order. The LA schemes can therefore be classified according to the degree of the approximating polynomial. The model coefficients are usually estimated by the least squares method using the so-called singular expansion [15]. (See [16] for a detailed description and comparison of methods for calculating the model coefficients when using the LA method.)

In some cases of very short series, the zeroth-order approximation is used, i.e., it is assumed that subsequent values of the series are independent of preceding values inside each local domain. In this case, the "forecast" depends only on hitting the concrete local domain. Such a situation is well illustrated by an example pertaining to the weather forecast [17]: one possible way to predict tomorrow's temperature is to find the previous days when the temperature was maximally close to today's temperature and take the temperature on the day that followed the found day as the forecast. A higher-order approximation allows taking the influence on the forecast of the difference between today's temperature and the temperature on the found day into account.

4. Forecasting: At the last step, we forecast the next value of the series under the assumption that the evolution of the last term of the series is governed by the same law as for other vectors from its local neighborhood.

Extrapolation methods. Two main extrapolation methods [17] can be used to forecast a few steps forward. The first is the *iterative* forecast: the parameters of approximating model (3) are estimated for $T=1$. The further forecast for $T=2,3\ldots$ is the sequence of iterations, i.e., each time the predicted value is added to the series. It is then assumed that the obtained vector falls in the same domain in the space of delays (i.e., it evolves following the same law), and a forecast can be constructed for one more step forward, and so on. As shown in [18], the latter hypothesis may result in a systematic error in the forecast.

An alternative is the *direct* forecast: parameters are estimated separately at each value of T. This method allows organizing a more precise forecast because errors are not accumulated with steps in this case [19].

The noise influence. In the development of the LA method, it was assumed that the forecast accuracy is determined by the approximation quality, which in turn depends on the volume of available data. But

in many cases, the forecast accuracy is naturally limited by noise: even if we know the exact equations of motion, the noise narrows the predictability limits by causing errors in determining the initial conditions and by distorting deterministic trajectories. It was shown in [7] that in its consequences, the influence of noise on the forecast quality is very similar to the influence of approximation errors. The most dangerous error due to noise in the LA method is the erroneous choice of neighbors.

2.3. Singular spectrum analysis. The SSA method was primarily developed as a method for segregating periodic and quasiperiodic constituents from a series. It was also shown that this method can be used to improve the signal-to-noise ratio [7]. In recent variants of the SSA method, its capacities have been extended to predict the further dynamics of a series [8], [20]. We use this method here exclusively to damp the noise.

The basic idea of the SSA method is to process the matrix $\mathbf{X}$ following an algorithm close to the principal component (PC) method. The use of the PC method is the most important ingredient of the SSA method, which makes it different from other nonlinear dynamics methods used to analyze and forecast time series. The essence of the PC method is reducing the dimension of the initial space of factors (here, the space of delays) by passing to more "informative" variables (coordinates). The new coordinates thus obtained are called PCs. We perform this passage using an orthogonal linear transformation.

The problem of passing to the PCs can actually be reduced to finding eigenvectors and eigenvalues of the matrix $\mathbf{X} \cdot \mathbf{X}^{\mathrm{T}}$ [21]. Let

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_M \end{pmatrix}$$

and $\mathbf{V}_{M\times M}$ be the respective matrices of eigenvalues and eigenvectors of the matrix $\mathbf{X} \cdot \mathbf{X}^{\mathrm{T}}$. Because the dimensions of reconstructions (and correspondingly of the matrices $\mathbf{X}$) used in LA and SSA may differ in the general case, we let M denote the reconstruction dimension for the SSA method. Then the PC matrix is $\mathbf{Z} = \mathbf{V}^{\mathrm{T}} \cdot \mathbf{X}$. Under such a transformation, the PCs are the sets of projections of points of the initial set on the eigenvectors. The eigenvalues then characterize the scattering of points on the axes of the new coordinate system. The first eigenvalue is maximum, and the others decrease monotonically.

The PCs have many useful properties. In the SSA method, we use the resulting decomposition to select the most significant constituents of a series and to eliminate random perturbations. The basic idea of this filtration is to use only the r most significant PCs, not all of them, to reconstruct the matrix $\mathbf{X}$.

In the general case, we obtain the approximation of the matrix $\mathbf{X}$

$$\widehat{\mathbf{X}} = \mathbf{V}_{M\times r} \cdot \mathbf{V}_{M\times r}^{\mathrm{T}} \cdot \mathbf{X}, \tag{4}$$

where $\mathbf{V}_{M\times r}$ is the part of the eigenvector matrix (the first r columns) corresponding to the first r PCs. After the matrix $\mathbf{X}$ is approximated, the next step is to reconstruct the initial series. When it is reconstructed over its first PC, the matrix $\mathbf{X}$ loses its initial diagonal form, and we must therefore average along all diagonals on which we initially had identical values:

$$\hat{f}_t = \begin{cases} \dfrac{1}{t}\sum_{i=1}^{t} \hat{x}_{M-t+i,t}, & t < M, \\ \dfrac{1}{M}\sum_{i=1}^{M} \hat{x}_{M-t+i,t}, & M \le t \le N-M+1, \\ \dfrac{1}{N-t+1}\sum_{i=1}^{N-t+1} \hat{x}_{i,t-M+i}, & N-M+1 \le t \le N. \end{cases} \tag{5}$$

The algorithm for reconstructing the time series using the SSA method therefore consists of three basic steps:

1. constructing the matrix $\mathbf{X}$,
2. calculating PCs and choosing the most significant among them, and
3. reconstructing the series by the chosen PCs.

Using this algorithm allows smoothing the initial series, decreasing the noise level, and increasing the signal-to-noise ratio. But forecast methods based on this algorithm [8] turn out to be insufficiently effective when processing nonperiodic series.

As does the LA method, the SSA admits several modifications, mainly related to a preliminary centering and/or normalization of the rows of the matrix $\mathbf{X}$. We use the variant with centered rows because it is closest to the standard PC method.

3. Generalizing the LA method

3.1. The SSA–LA algorithm. In the case of very noisy series, even increasing the number of observations cannot ensure effective application of the PC algorithm because the probability that false neighbors appear and true neighbors are expelled is high in this case. To improve the forecast quality, it therefore seems useful to combine these two methods (SSA and LA) into a unified method in which we use the SSA method only to filter the initial time series (to damp the noise) and use the LA method to produce a forecast.

The proposed method (called the SSA–LA method) can be treated as a generalization of the LA method to the case of very noisy time series. In the nearest future, we also want investigate the possibility of segregating the trend and periodic constituents in the SSA filtration process, which is especially important for applications to time series in economics.

The total SSA–LA algorithm contains two basic steps:

1. performing the SSA-filtration, i.e., damping the noise, and
2. forecasting the smoothed series using the LA method.

We have already described the algorithms of the SSA and LA methods. We note here that because of the special analysis [19], we choose SSA with centering and the direct variant of the LA method to construct the forecast.

When the SSA and LA methods are applied together, it turns out that the first M and the last M terms of a time series are reconstructed with large deviations from the initial values, while the remaining part of the series is approximated much better. Apparently, the decrease in the approximation accuracy for boundary values is due to the shorter interval of averaging used to calculate in (5). We therefore eliminate the first M and the last M values of the averaged series and construct the forecast starting from the $(N-M+1)$th value.

3.2. Numerical results. We now present some numerical results demonstrating the capabilities of the SSA–LA method for forecasting very noisy irregular time series. Such time series were obtained from the differential–difference Mackey–Glass equation [22], often used as a test example (see [14], [17]),

$$\frac{dx}{dt} = \frac{ax(t-\Delta T)}{1+x^c(t-\Delta T)} - bx(t),$$

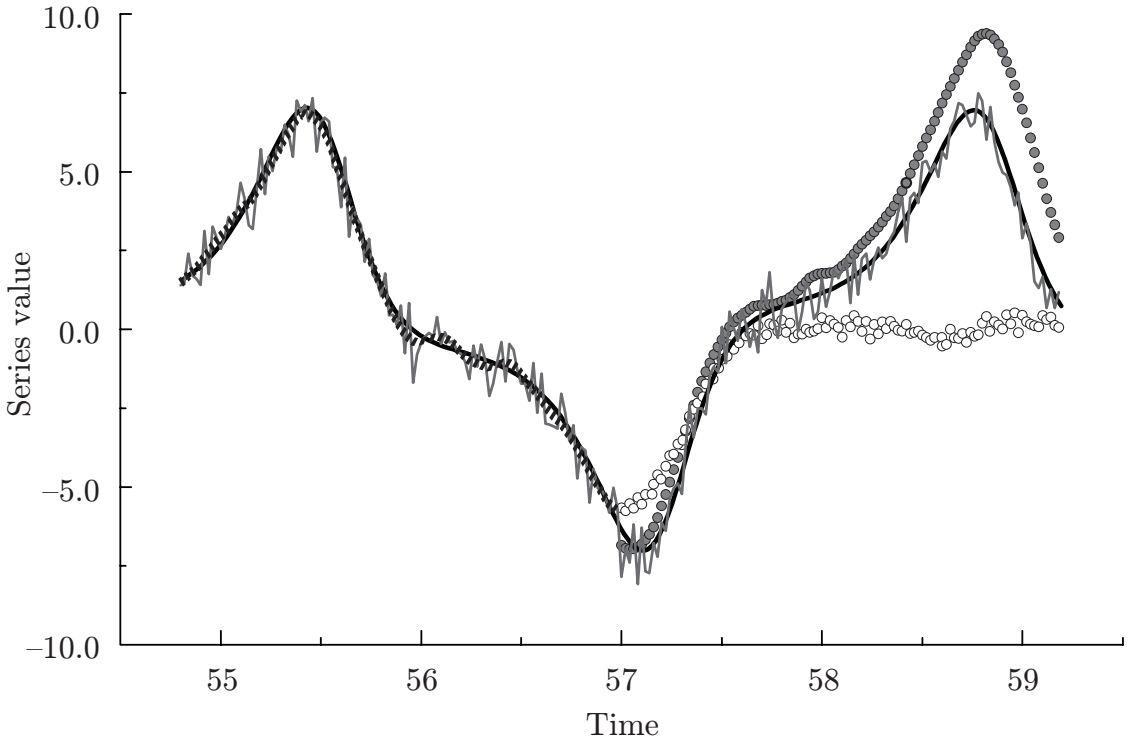

Fig. 1. The SSA filtration and the forecast for the x component of the Lorenz system with 5% Gaussian noise (from the maximum value). The forecast starts from the 57th point. The forecast comprises 110 points. The solid line shows the initial series; the broken line, the initial series with noise added; the dotted line, the result of smoothing the noised series; the empty circles, the forecast using the LA method; and the filled circles, the forecast using the SSA–LA method.

where $a = 0.2$, $b = 0.1$, $c = 10$, and $\Delta T = 17$. In addition, we used the well-known Lorenz system

$$\dot{x} = \sigma(y - x),$$

$$\dot{y} = (r - z)x - y,$$

$$\dot{z} = xy - bz,$$

where $\sigma = 5$, $r = 15$, and $b = 1$. The series were calculated once for all numerical experiments, while the noise component in the form of uncorrelated Gaussian noise with a zero mathematical expectation was generated separately for each experiment.

The results of our analysis are presented in Figs. 1–3. For the forecast, we processed series of each type comprising 3600 points each, from which we used only 110 points to estimate the forecast quality. We took $M = 18$, $r = 6$ and $m = 6$, $N^s = 21$ as the respective parameters of the SSA filtration and the LA. The problem of consistency of the LA and SSA parameters was solved empirically at this stage.

We first consider the example of forecasting the x component of the Lorenz system (Fig. 1). In the left part of the plot (before the instant 57.0), we present the result of the SSA filtration of the noisy signal (the last 110 points), which provides an apparently accurate reconstruction of the initial series without noise. But the phase is reconstructed with a small error, seen most explicitly in the domain of the maximum. We start the forecast from the instant 57.0. This forecast is indicated by circles in the figure. We see that preliminary filtration there resulted in a substantial increase of the forecast accuracy.

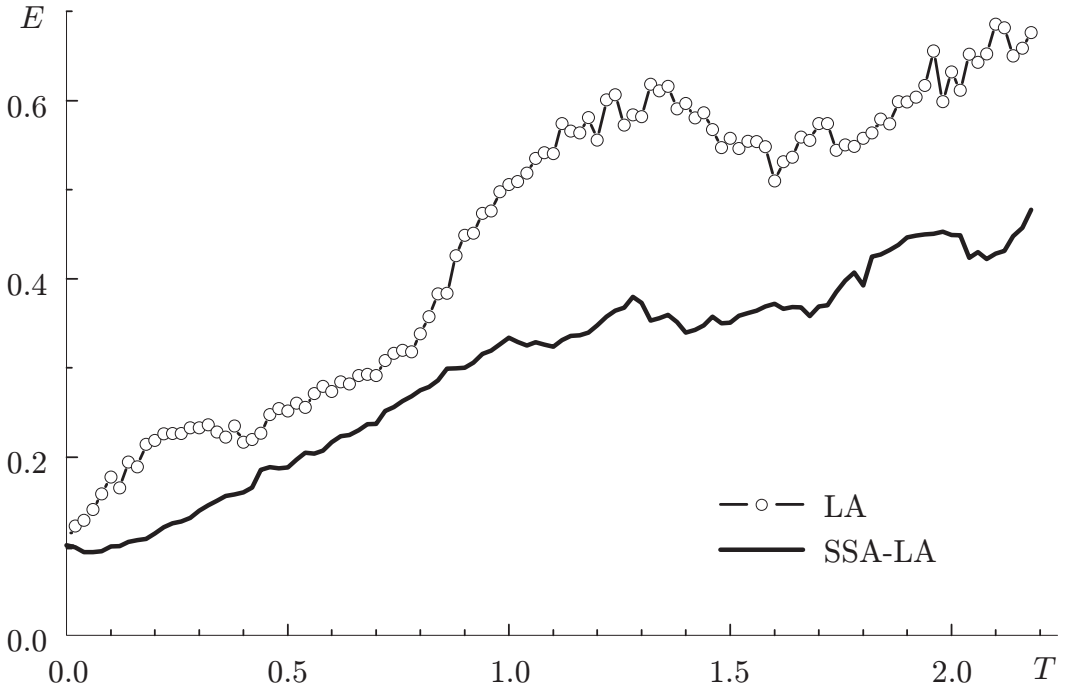

Fig. 2. The forecast error for the x component of the Lorenz system (for data with 5% noise applied). We used median averaging over 500 initial instants.

The forecast accuracy, generally speaking, can strongly depend on the initial instant of the forecast. To obtain a more objective estimate of the quality of forecasts obtained with the standard LA method and with the SSA–LA method, we must use a characteristic that is less sensitive to the initial instant. We take the normalized mean squared error (the forecast error) as such a characteristic for comparing the accuracies of forecasts obtained by different methods (see [17])

$$E = \sqrt{\frac{\langle (f_{t+T} - \hat{f}_{t+T})^2 \rangle_t}{\langle (f_t - \langle f_t \rangle_t)^2 \rangle_t}}. \tag{6}$$

Here we average over the initial instants of the forecasts. Of course, the less the forecast error, the better the forecast is. If the quantity E becomes greater than unity, then the forecast fails, and it is better to take just the mean of the series instead. To achieve better reliability of the obtained error, we use the median, not the arithmetic mean, when averaging over the initial instants of the forecasts in relation (6).

In Fig. 2, we show the dependence of error (6) on the length of the forecast for the x component of the Lorenz system. As in Fig. 1, the forecast with preliminary filtration is better in the case of averaging. The forecast error for the SSA–LA algorithm is always less than that for the standard LA method. Dependences similar to the one in Fig. 2 were constructed for various noise amplitudes and always, except for the case of zero noise, the SSA–LA method provided more accurate predictions.

Following the same criterion (the forecast error), we performed comparisons for series obtained using the Mackey–Glass equation. The results of this comparison are very close to those obtained for the Lorenz series. The corresponding dependences of the forecast error on the length of the forecasts for different values of noise are presented in Fig. 3. For this series, the SSA–LA algorithm is also more accurate than the standard LA method.

The numerical analysis results thus indicate that the preliminary SSA filtration allows increasing the accuracy of a forecast obtained with the LA method substantially and that unifying these two methods

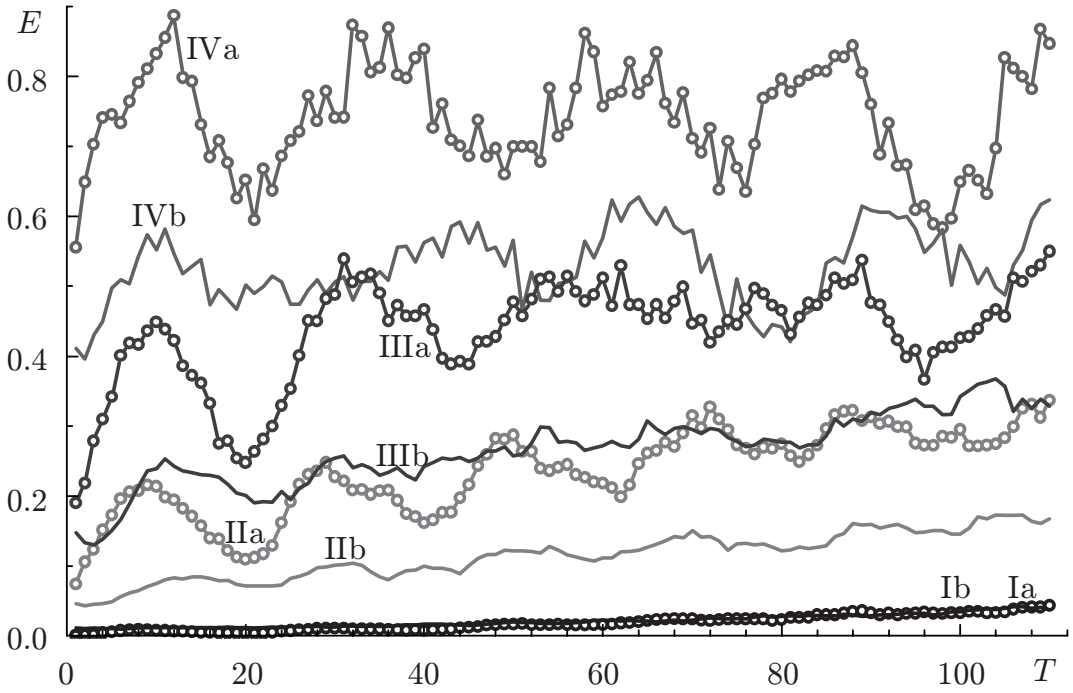

Fig. 3. The dependence of forecast error obtained with (a) the LA and (b) the SSA–LA methods for the Mackey–Glass equation on the forecast length in the absence of noise (I) and at the noise levels constituting 1.5% (II), 5% (III), and 20% (IV) of the maximum amplitude value. We used median averaging over 500 starting points.

into a hybrid SSA–LA algorithm can be especially useful when forecasting noisy series for which the LA method alone is insufficiently effective.

4. Conclusion

Forecasting time series using the LA method, in principle, is currently more an opportunity than an actual research tool. At the same time, the LA algorithm is certainly advantageous to the standard autoregression because using the fractional-linear, not global linear, approximation allows using LA to successfully forecast irregular time series for which the linear autoregression representation cannot be used. Segregating into domains inside which we use the linear approximation then occurs in the space of delays, not in the time scale. This results in true neighboring states falling into the same domain, which allows using a linear approximation inside each domain, whereas a global linear approximation corresponding to approximating the whole system phase space by a single hyperplane may be absolutely unrealistic.

The main obstacle restricting the use of the LA method is that its application becomes effective only when forecasting long time series. Then, of course, we must use much longer series for very noisy data, and series of sufficient lengths are seldom obtainable for actual systems. To relax this restriction, we propose using the preliminary series filtration with the SSA method, which has won a good reputation as an effective tool for damping noise, especially in irregular time series, i.e., in the case where Fourier filtration cannot be used.

Numerical experiments showed that such a hybrid of SSA and LA methods always provided results more accurate than those from the standard LA method. Moreover, our results are practically independent of the noise level, the length of the forecast, and the system that generated the time series under investigation. This proves that using the SSA filtration method as a necessary constituent of the SSA–LA algorithm can

dramatically enlarge the applicability domain of the LA method and may demonstrate its advantages over traditional forecast methods when used in practice.

REFERENCES

1. G. E. P. Box and G. Jenkins, *Time Series Analysis, Forecasting, and Control*, Holden-Day, San Francisco, Calif. (1970).
2. A. S. Monin and L. I. Piterbarg, "Forecasting weather and climate [in Russian]," in: *Limits of Predictability* (Yu. A. Kravtsov, ed.), Nauka, Moscow (1992), pp. 28–53; English transl., Springer, Berlin (1993), pp. 7–44.
3. F. Takens, "Detecting strange attractors in turbulence," in: *Dynamical Systems and Turbulence* (Lect. Notes Math., Vol. 898, D. A. Rand and L. S. Young, eds.), Springer, Berlin (1981), pp. 336–381.
4. J. D. Farmer and J. J. Sidorowich, *Phys. Rev. Lett.*, **59**, 845–848 (1987).
5. R. Vautard, P. Yiou, and M. Ghil, *Phys. D*, **58**, 95–126 (1992).
6. A. Yu. Loskutov, I. A. Istomin, O. L. Kotlyarov, and K. M. Kuzanyan, *Astron. Lett.*, **27**, 745–753 (2001).
7. M. Ghil et al., *Rev. Geophys.*, **40**, No. 3, 1–41 (2002).
8. D. L. Danilov and A. A. Zhiglyavskii, eds., *Basic Ingredients of Time Series*: *The "Caterpillar" Method* [in Russian], St. Petersburg Univ. Publ., St. Petersburg (1997).
9. I. M. Dremin, O. V. Ivanov, and V. A. Nechitailo, *Phys. Usp.*, **44**, 447–478 (2001); A. V. Deshcherevskii, A. A. Lukk, A. Ya. Sidorin, G. V. Vstovsky, and S. F. Timashev, *Nat. Hazards and Earth Syst. Sci.*, **3**, No. 2, 159–164 (2003).
10. D. S. Broomhead and G. P. King, *Phys. D*, **20**, 217–236 (1986).
11. R. Mañe, "On the dimension of the compact invariant sets of certain nonlinear maps," in: *Dynamical Systems and Turbulence* (Lect. Notes Math., Vol. 898, D. A. Rand and L. S. Young, eds.), Springer, Berlin (1981), pp. 230–242.
12. G. G. Malinetskii and A. B. Potapov, *Contemporary Problems of Nonlinear Dynamics* [in Russian], URSS, Moscow (2000).
13. P. Grassberger and I. Procaccia, *Phys. Rev. Lett.*, **50**, 346–349 (1983).
14. P. Grassberger and I. Procaccia, *Phys. D*, **9**, 189–201 (1983).
15. D. Kahaner, C. Moler, and S. Nash, *Numerical Methods and Software*, Prentice-Hall, Upper Saddle River, N. J. (1989).
16. D. Kugiumtzis, O. C. Lingjoerde, and N. Christophersen, *Phys. D*, **112**, 344–360 (1998).
17. J. D. Farmer and J. J. Sidorowich, "Exploiting chaos to predict the future and reduce noise," in: *Evolution, Learning, and Cognition* (Y. C. Lee, ed.), World Scientific, Singapore (1988), pp. 277–330.
18. K. Judd and M. Small, *Phys. D*, **136**, 31–44 (2000).
19. A. Yu. Loskutov, I. A. Istomin, O. L. Kotlyarov, and D. I. Zhuravlev, *Moscow Univ. Phys. Bull.*, No. 6, 3–21 (2002).
20. A. Loskutov, I. A. Istomin, K. M. Kuzanyan, and O. L. Kotlyarov, *Nonlinear Phenomena in Complex Systems*, **4**, No. 1, 47–57 (2001).
21. A. M. Dubrov, V. S. Mkhitaryan, and L. I. Troshin, *Multidimensional Statistical Methods* [in Russian], Finansy i Statistika, Moscow (2000).
22. M. C. Mackey and L. Glass, *Science*, **197**, 287–289 (1977).