Nikolai G.Makarenko[1*], Lyailya M.Karimova[2], Olga A.Kruglun[3]
[1*] Leading scientist , Pulkovo Astronomical Observatory, Saint-Petersburg, Russia; Institute of Mathematics, Pushkin str. 200010, Almaty, Kazakhstan ng-makar@mail.ru, **Corresponding author**
[2] Principal scientist, Institute of Mathematics, Pushkin str. 200010, Almaty, Kazakhstan, klyailya@mail.ru,
[3] Senior scientist, Institute of Mathematics, Pushkin str. 200010, Almaty, Kazakhstan, okruglun@mail.ru

CASPIAN SEA LEVEL PREDICTION USING ARTIFICIAL NEURAL NETWORK AND EMPIRICAL MODE DECOMPOSITION

ABSTRACT

This paper demonstrates the possibility of using nonlinear modeling for prediction of the Caspian Sea level. Phase space geometry of the of a model can be reconstructed by the embedology methods. Dynamical invariants, such as the Lyapunov exponents, the Kaplan-Yorke dimension, and the prediction horizon were estimated from reconstruction. Fractal and multifractal analyses were carried out for various time intervals of the Caspian Sea level and multifractal spectra were calculated. Then, historical data resolution was improved with the help of fractal approximation. The EMD method was used to reduce noise of the time series. Global nonlinear predictions were made with the help of Artificial Neural Network for combinations of different empirical modes.

KEY WORDS: Caspian Sea level, Fractal Approximation, Multifractal Analysis, Empirical Mode Decomposition, Embedology, Nonlinear prediction

INTRODUCTION

The Caspian Sea is the largest intercontinental reservoir without water outflow which exhibits a unique global evolution over an extremely long interval of time. On geological scale, the history of the Caspian Sea is represented by alternations of transgressive and regressive phases reflected clearly by paleodata, reflected in scanty historical records, and shown by instrumental measurements during recent and a relatively short monitoring period. In Holocene, for example, the fluctuations of the Caspian Sea Level (CSL) were caused by climate changes. It is possible that the CSL was influenced not only by water balance, but also by tectonic changes of the seabed, and that these factors did not necessarily coincide in time [Golitsyn, 1995]. Up to date, the CSL has being raising for already 18 years. By the beginning of 1996 year, the CSL has reached –26,6 meters and the area of the Caspian Sea has increased by 40 000 square kilometers. The rising sea level and strong winds resulted in multiple problems for economic development. Economics of the regions near the Caspian Sea depended on modern sea fluctuations. Modeling the dynamics of a closed reservoir without outflows may appear straightforward on the first sight. However, simple balance models may not adequately describe the situation. Only nonlinear models may be applied to describe a chaotic dynamics of the CSL [Makarenko et al., 2004; Kozhevnikova & Shveikina, 2008]. Thus, in this paper, we applied nonlinear modeling to predict the CSL. This approach was based on reconstruction of phase dynamics from observed time series by means of topological embedding methods. A nonlinear global forecast has been made with the help of Artificial Neural Network (ANN). The CSL time series were constructed by fractal approximation.

EMBEDOLOGY AND NONLINEAR PREDICTION

The nonlinear approach for modeling and prediction of dynamic regimes of sea level can be based on chaotic dynamics [Makarenko, et al., 2004]. According to general assumptions about properties of an unknown dynamic model of sea level we can reconstruct the diffeomorphic copy of its attractor in an n-dimensional space. This technique, called embedology, is widely known [Sauer et al., 1991]. We used it to create the nonlinear scheme of sea level prediction. Embedology methods assume that an observed variable is typical and contains all characteristic elements of dynamics. However, the CSL data have being measured only since 1830 and they do not contain information about global variations of sea level. Thus, one can hope only for a short-term forecast. In fact, the prediction horizon is determined by the maximum positive value of the Lyapunov exponent [Schaw, 1981] according to $T \gg \log_2 N/kl^+_{max}$, where N is time series length, k = [1÷3]. Applying the embedding dimension with m = 3, τ = 2, and a number of nearest neighbors of 20, we obtained Lyapunov exponents: l_1^+ = 0.302797, l_2^- = –0.14458, l_3^- = –0.838176 for the instrumental time series. The Kaplan-Yorke dimension is 2.18876 that is close to the embedding dimension. In our case, the length is N = 1955 and the prediction horizon is $T \approx$ 12–36 months [Makarenko, et al., 2004].

The reconstruction of the copy of the attractor into Euclidian space R^m gives possibility to obtain the following predictor [Sauer et al., 1991, Makarenko, 2003]:

$$y((k+l)\Delta t) =$$

$$= F(\mathbf{y}(k), \mathbf{y}(k-\tau), \mathbf{y}(k-2\tau), \ldots, \mathbf{y}(k-m\tau)).$$

We used $m = 27$ and $\tau = 37$ to construct delay vectors **y.** Unknown function F-predictor is nonlinear and continuous function of m vectors of the reconstruction. Their best approximation is found by ANN [Poggio & Girosi, 1989; Bishop, 2006]. ANN training was carried out using a set from available values of the CSL data. In the case when a lag of $\tau \neq 1$, one can obtain τ predicted points simultaneously, i.e. can construct a vector prediction. The lower estimate of the embedding dimension $m > 8$ was obtained with the help of the False Nearest Neighbors method. The prediction horizon is limited by the time series length and rate of divergence of close reconstructed trajectories. The prediction horizon of instrumental monthly CSL data was estimated at 12–36 months, as mentioned above.

FRACTAL APPROXIMATION OF THE CSL HISTORICAL DATA

In order to use a nonlinear context and obtain a nonlinear prediction for historical data from 600 BC, we applied fractal approximation [Barnsly, 2000; Karimova et al., 2003; Makarenko et al. 2004]. The latter could enhance the historical data that are poor in accuracy and have low time-resolution (a point per ten years). The usage of historical data in the prediction task is very important, because its accurate instrumental measurements reflect short time variations of the CSL and does not trace its global evolution. The main ideas of fractal approximation is as follows. The interpolation problem deals with a set of input pairs $\{(x_i, y_i)\}_{i=0}^{N}$ where the $0 = x_0 < x_1 < \ldots < x_N = 1$ are nodes and $y_i = F(x_i) \in R$ ordinates with some continuous function F:[0, 1] $\rightarrow R$. As a rule, in the case of smooth data, the input points are interpolated by a single-degree N polynomial, or by piecewise interpolations with a low-degree polynomial. Recent research has provided an alternative assumption that the interpolation function $F$is self-similar, and typically not smooth, but fractal. We note that a function F:[0, 1] $\rightarrow R$ is well defined by its graph, and use the same symbol to denote the set of points in its graph. Hence a point $(x, y) \in F$ if and only if. We also use the notation $F[x_1, x_2]$ to denote the graph of F over the interval $[x_1, x_2]$. Hence a point $(x, y) \in F[x_1, x_2]$ if $(x, y) \in F$ and $x \in [x_1, x_2]$. We construct an *Iterated Function System* (IFS)) [Barnsly, 2000] whose attractor is the graph of a function F:[0, 1] $\rightarrow R$. Such a function is called a *Fractal Interpolating Function* (FIF) [Cochran, W.O. et al. 1998]. For i = 1, 2, ..., N, let T_i:[0, 1] $\times R \longrightarrow$ [0, 1] $\times R$ has the form

$$T_i : \begin{bmatrix} t \\ x \end{bmatrix} \propto \begin{bmatrix} a & 0 \\ b_i & c_i \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} + \begin{bmatrix} d_i \\ e_i \end{bmatrix},$$

where $c_i < 1$ is given as a parameter controlling the roughness of the function, and a_i, b_i, d_i and e_i are determined either by the constraints

$$T_i(0, x_0) = (t_{i-1}, x_{i-1}),\ T_i(1, x_N) = (t_{i-1}, x_{i-1}),$$

or the "reflected" constraints

$$T_i(1, x_N) = (t_i, x_i),\ T_i(0, x_N) = (t_i, x_i).$$

Given a metric $d((t_1, x_1), (t_2, x_2)) = |t_1 - t_2| + \frac{(1-A)}{2B}|x_1 - x_2|$, where $A = \max_i|a_i|$ and $B = \max_i|b_i|$, it can be shown that each T_i has contractility $s = \max\{(1 + A)/2, C\}$, 1, where $C = \max_i|c_i|$. Hence, by the fixed point theorem, there exists one and only one function F satisfying the invariance $F = U_i T_i(F)$.

Fractal approximation can be used in the case of prior uncertainty, when available measurements are insufficient for determination of statistical characteristics of the concerned process. And it is the best tool for data approximation when the mean square of process augments depends on correlation interval according to the scaling law, i.e. $E[(x(t + \tau) - x(t))^2] \approx \tau^\gamma$, where γ is the Levy index and inverse value to the Hurst exponent.

In the case of the CSL, it has been shown that such time series satisfy the latter condition and we applied fractal approximation technique. The approximated data were historical decennial CSL time series measured from 600 BC to 2000 AD (261 records). The output of the fractal approximation procedure was the annual CSL time series, 2601 records in length. The time series of the annual CSL was obtained directly from the original historical time series.

MULTIFRACTAL CHARACTERISTICS

Multifractal formalism [Halsey et al., 1968; Riedi & Scheuring, 1997; Karimova et al., 2007] has proved to be a very useful technique in the study of both measures functions, deterministic as well as random. Multifractal analysis connects pointwise regularity of the function with a "size" of sets where regularity possesses some value. Function regularity may change abruptly from one point to the next. Pointwise regularity [Karimova et al., 2007] is a positive real number $\alpha(x)$, which describes a certain smoothness of the graph of a function at a point x. In general, let h be a nonnegative real number, $x_0 \in R$, a function $F(x)$:$R \rightarrow R$ is $C^h(x_0)$ if there exists $C > 0$, $\delta > 0$ and a polynomial $P(x)$ of the order smaller than h so that if

$$|x - x_0| \le \delta,\ |F(x) - P(x - x_0)| \le C|x - x_0|^h,$$

then the Hölder exponent of F at x_0 is a(x_0) = sup{h:$FisC^h(x_0)$}.

Let $E_\alpha = \{x \in R{:}\alpha(x) = \alpha\}$, then the fine (Hausdorff) multifractal spectrum [Riedi & Scheuring, 1997] is $f_H(\alpha) = \dim_H E_\alpha$, where $\dim_H E_\alpha$ is the Hausdorff dimension of the set E_α. Because $\dim_H$ of the set is never greater than its box dimension, one can estimate it by counting the boxes (or intervals) over F, the number of which increases roughly with the "right" Hölder exponent. In applications, however, one considers a *course grained version* f_G which is:

$$f_G(\alpha) = \lim_{\varepsilon\to\infty}\lim_{\delta\to\infty} \sup \frac{\log N_\delta(\alpha, \varepsilon)}{\log(1/\delta)}.$$

Here N_δ denotes the number of cubes of size δ with the coarse Hölder exponent roughly equal to α. Let $\mu(N_\delta)$ be a measure contained in a δ-cube, then a = $\log\mu(N_\delta)/\log\delta$.

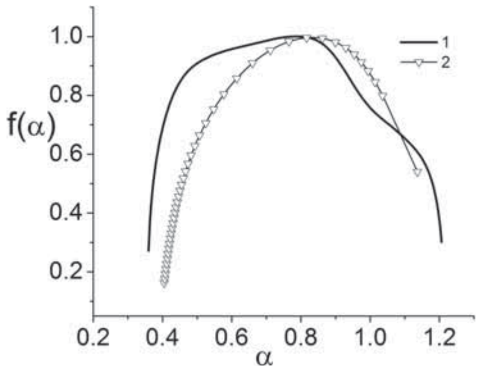

Fig. 1. The large deviation $f_G(\alpha)$ (1) and the Legendre $f_L(\alpha)$ (2) spectra for the fragment of the annual data constructed by fractal approximation using 1500-yr long decennial data

We computed the large deviation spectrum $f_G(\alpha)$ and the Legendre spectrum $f_L(\alpha)$ using FracLab software (http://fraclab.saclay.inria.fr/). Both Legendre and large deviation spectra of the time series are presented in Fig. 1. The multifractal spectra were computed for the fragment of the annual data constructed by fractal approximation using 1500-yr long decennial data. Fig. 1 shows that these spectra have similar maxima corresponding approximately to the Hölder exponent equal to 0.8.

EMPIRICAL MODE DECOMPOSITION (EMD) FOR THE CSL

The main difficulty for construction of the global nonlinear prediction is an uncertainty associated with monthly instrumental time series. To avoid this difficulty we apply the well known Empirical Mode (EMD) decomposition [Huang et al., 1998] and approximate time series by the sum of smooth empirical modes. According to this method, it is possible to use a coarse-grained approximation of a signal, excluding high-frequency details, without breaking global structure of the signal (Fig. 2).

Thus, the method of decomposition of a signal by means of empirical modes [Huang et al., 1998; Flandrin et al., 2003] represents the signal as a set of functions corresponding to various oscillations in observed signal. A basic operation in EMD is the estimation of the upper and lower "envelopes" as interpolated curves between

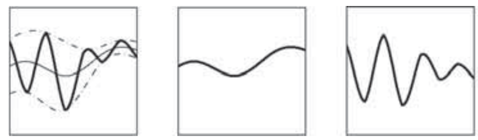

Fig. 2. Decomposition of the signal by means of empirical modes. From [Flandrin et al., 2004]

extremes. The nature of the chosen interpolation plays an important role, and our experiments tend to confirm (and are in agreement with) what is recommended in [Huang et al., 1998], specifically, that cubic splines are to be preferred. Other types of interpolation (linear or polynomial) tend to increase the required number of sifting iterations and to "over-decompose" signals by spreading out their components over adjacent modes.

Given a signal $x(t)$, the effective algorithm of EMD can be summarized as follows [Flandrin, et al., 2004]:

1. identify all extremes e_{min}, e_{max} of signal $x(t)$.

2. interpolate between minima and maxima, ending up with some envelope $e_{min}(t)$, $e_{max}(t)$.

3. compute the mean $m(t) = (\boldsymbol{e}_{min} + \boldsymbol{e}_{max})/2$.

4. extract the detail $d(t) = x(t) - m(t)$.

5. iterate on the residual $m(t)$.

In practice, the above procedure has to be refined by a *sifting* process which amounts to first iterating steps 1 to 4 upon the detail signal $d(t)$, until this latter can be considered as zero-mean according to some stopping criterion. Once this is achieved, the detail is referred to as an *Intrinsic Mode Function* (IMF), the corresponding residual is computed and step 5 is applied. By construction, the number of extremes is decreased when going from one residual to the next, and the whole decomposition is guaranteed to be completed with a finite number of modes. For calculation of empirical components software batch http://perso.ens-lyon.fr/patrick.flandrin/emd.html was applied.

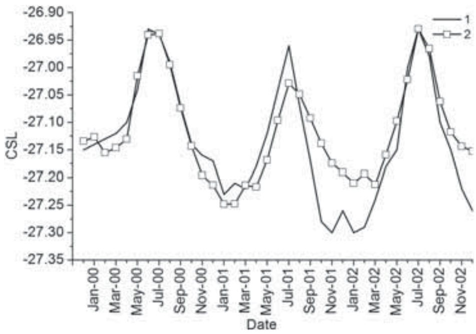

Fig. 3. The CSL real monthly data (1) and the prediction (2)

NUMERICAL RESULTS

During the experiment, the monthly data from January 1837 to December of 2002 were used. From 8 constructed empirical modes, the sum of modes 2, 4, 5, 6, 7, and 8 were taken, for which the correlation coefficient with the original data was 99,8%. For selected delay interval $\tau = 37$, the prediction interval was 37 months: December 1999 – December 2003.

All predictions of the CSL time series were carried out by ANN, namely *Statistica Neural Network v4.0E*. Fig. 3 demonstrates real CSL month data (1) and the prediction (2). Fig. 4 shows the CSL real monthly data, the sum of 2, 4, 5, 6, 7, and 8 modes of EMD decomposition and the mean values of three predictions made with the help of EMD. Deviation of predicted values from the real data did not exceed 1%.

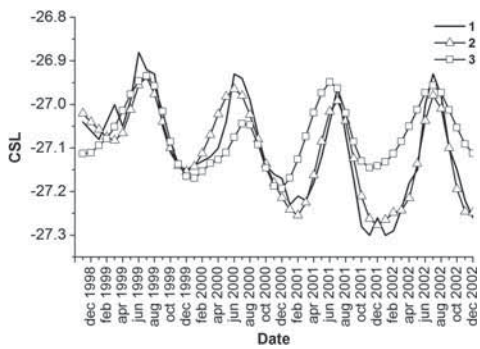

Fig. 4. Comparison of the real CSL monthly data (1), the sum of 2, 4, 5, 6, 7, and 8 modes of EMD decomposition (2) and the mean value of three predictions based on EMD decomposition (3)

CONCLUSION

We have discussed the method of the CSLs predictions based on the reconstruction of a dynamical system with the help of embedding time series in Euclidian space of an appropriate dimension. This approach makes it possible to construct vector prediction, i.e. to forecast a consecutive set at the same time. Therefore, one can avoid an exponential increase of errors inherent in a step-by-step prediction. The prediction is realized by means of the ANNs, that are optimal approximating tool for an unknown continuous and multivariable function, i.e. a nonlinear predictor.

The ANN is trained by transforming input examples into the outputs as the known answer of the training set constructed by the known records from the past. To use the decennial historical data together with the annual instrumental data, we constructed fractal approximation of the decennial data that allowed increasing the time series. The necessary property of statistical scale invariance was verified by multifractal spectra. Additional improvement of the data was implemented with the help of EMD technique that allowed delicate noise filtration without disturbance of correlation of the time series. The global nonlinear predictions were made with the help of ANN for combinations of different empirical modes.

Our experiments based on the approaches mentioned above have shown possibility to forecast the CSLs at 1–3 yrs intervals, which is extremely useful in practice. It is necessary to point out that a nonlinear prediction represents an ill-posed problem as it produces a great number of possible variants. Selection of the most probable variant from a set of predictions remains an open challenge.

ACKNOWLEDGEMENTS

This research was funded by the scientific and technical program of the Republic of Kazakhstan "Resource evaluation and prediction of water resources usage in Kazakhstan during the environment and anthropogenic climatic changes". ■

REFERENCES

1. Barnsley, M.F. (2000) Fractals Everywhere. Hawley Rising, 550 p.

2. Bishop, Ch.M., (2006) Pattern Recognition and Machine Learning. Springer, 738 p.

3. Cochran,W.O., Hart J.C., Patrick J. F. (1998) On Approximating Rough Curves With Fractal Functions. Proceedings of the Graphics Interface 1998 Conference, June 18–20, 1998, Vancouver, BC, Canada. pp. 65–72.

4. Flandrin, P., Gonçalves, P. (2004) Empirical mode decompositions as data-driven wavelet-like expansions. *Inter. Journ. of Wavelets, Multiresolution and Information Processing*. Vol. 2, pp. 1–20.

5. Golitsyn, G.S. (1995) Changes in the Caspian Sea Level as aProblem for Diagnosing and Predicting Regional Climatic Changes. *Izv. RAN, Fiz. Atm. Okeana* Vol. 31, pp. 385–391. (in Russian).

6. Halsey, T.C., Jensen, M.H., Kadanoff, L.P., Procaccia, I., Schraiman, B.I. (1968) Fractal measures and their singularities: the characterizations of strange sets. *Phys. R*ev. A., Vol. 33, pp. 1141–1151.

7. Huang, N.E., Shen, Zh., Long, St.R., and et al. (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis Proc. R. Soc. Lond. A., Vol. 454, pp. 903–995.

8. Karimova, L., Mukhamejanova, S., Makarenko, N. (2003) Fractal geometry methods and neurocomputing for Caspian Sea level forecasting. EGS – AGU – EUG Joint Assembly, Abstracts from the meeting held in Nice, France, 6–11 April 2003, abstract #186.

9. Karimova, L.M., Kuandykov, Y.B., Makarenko, N.G., Novak, M.M., Helama, S., (2007) Fractal and topological dynamics for the analysis of paleoclimatic records. Physica A: Statistical Mechanics and its Applications, N 373, pp. 737–746.

10. Kozhevnikova, I.A., Shveikina, V.I. (2008) Nonlinear dynamics of level in the Caspian Sea. Water Resources, Vol. 35, pp. 297–304.

11. Makarenko, N.G., Karimova, L.M., Kyandykov, Y.B., Novak, M.M. (2004) Nonlinear Dynamics and Prediction of the Caspian Sea Level. Thinking in Patterns, M. M. Novak (ed.). World Scientific, pp. 91–102.

12. Makarenko, N.G. (2003) Embedology and Neural Networks, Moscow. Lectures on neuroinformatics. " Neuroinformatics-2003", part. 1, pp. 86–148, (in Russian).

13. Poggio, T., Girosi, F. A (1989) Theory of networks for approximation and learning. MIT AI Lab. Technical Report, Memo No. 1140, Paper No. 31.

14. Riedi, R. H., Scheuring, I., (1997) Conditional and Relative Multifractal Spectra. Fractals, N 5, pp. 153–168.

15. Sauer, T., Yorke, J.A., Casdagli, M. (1991) Embedology. J. Statist. Phys., Vol. 65, pp. 579–616.

16. Schaw, R. (1981) Strange attractors, chaotic behavior, and information flow. Z. Naturforsch, Vol. 36a, pp. 80–112.

Nikolai Makarenko was born in Troitzk, Russia, in 1945. He studied at the Ural State University, graduated in 1967 and obtained the Diploma on "Astronomy". He is a leading scientist, Doctor of Physical and Mathematical Sciences of GAO, Pulkovo, and leading scientist, also Doctor of Technical Science of the Institute of Mathematics, Almaty, Kazakhstan. The focus of his research lies in Nonlinear dynamics, Neuroinformatics, Morphological Image Analysis. Main publications: Makarenko N., Knyazeva I. Multifractal analysis of digital images. // Izv. VUZ. Applied nonlinear dynamics. 2009. V. 17. P. 84–97 (in Russian).
Makarenko N., Kruglun O., Makarenko I., Karimova L. Multifractal segmentation of remote sensing data. // Issled. Zemli iz Kosmosa, (2008), № 3. P. 1–9 (in Russian).
Makarenko N., Karimova L.Contemporary nonlinear methods of time series prediction // Econophysics. Modern physics in search of economy theory. V.V. Kharitonov, A.A. Ezov (Eds). Moscow. (2007). P. 316–352 (in Russian).

Lyailya Karimova was born in Almaty, Kazakhstan, in 1948. She studied at the Saint Petersburg State University, graduated in 1973 and obtained the Diploma on "Radiophysics". She is a leading scientist of Institute of Mathematics, PhD in Physical and Mathematical Science. The focus of her research lies on Nonlinear dynamics, Multifractal Analysis, Morphological Image Analysis. Main publications: L. Karimova, Y. Kuandykov, N. Makarenko, M.M. Novak and S. Helama. Fractal and topological dynamics for the analysis of paleoclimatic records. Physica A., (2007), V. 373, P. 737–746.
Makarenko N., Karimova L., Novak M.M. Investigation of global solar magnetic field by computational topology methods // Physica A. (2007). V. 380. P. 98–108.

Olga Kruglun was born in Almaty, Kazakhstan, in 1981. She studied at the Kazakh National University, graduated in 2004 and obtained the Diploma on "Mathematics". She is a senior scientist of Institute of Mathematics, PhD in Technical Science. The focus of her research lies on Multifractal Analysis, Morphological Image Analysis, Neural Networks. Main publications:
Helama, S., Makarenko, N.G., Karimova L.M., Kruglun, O.A., et al. Dendroclimatic transfer functions revisited: Little Ice Age and Medieval Warm Period summer temperatures reconstructed using artificial neural networks and linear algorithms // Ann. Geophys., (2009). V. 27, P. 1097–1111.