Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Nuclear Instruments and Methods in Physics Research A 534 (2004) 170–174
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

The genetic algorithm for a signal enhancement

L. Karimova*[,1], E. Kuadykov[a], N. Makarenko[a]

Laboratory of Computer Modelling, Institute of Mathematics, Pushkin Street 125, 480100 Almaty, Kazakhstan

Available online 30 July 2004

Abstract

The paper is devoted to the problem of time series enhancement, which is based on the analysis of local regularity. The model construction using this analysis does not require any a priori assumption on the structure of the noise and the functional relationship between original signal and noise. The signal itself may be nowhere differentiable with rapidly varying local regularity, what is overcome with the help of the new technique of increasing the local Hölder regularity of the signal under research. A new signal with prescribed regularity is constructed using the genetic algorithm. This approach is applied to enhancement of time series in the paleoclimatology, solar physics, dendrochronology, meteorology and hydrology.

PACS: 05.45.Df; 07.05.Kf; 07.05.Mh

Keywords: Noise reduction; Hölder regularity; Neutral networks

1. Introduction

Nowadays many noise reduction methods are known [1–4]. Let the signal Y be presented in the form of composition $Y = F(X, B)$ of a clean signal X and a noise term B. As a rule, the noise term B is assumed to be some, for example, Gaussian stochastic process, which does not depend on X and X in its turn is considered as a piecewise-smooth function of class $C^n, n \geqslant 0$. Finally, if F is a linear or quasi-linear functional dependence and if the B and X have different frequency characteristics, one can reduce the noise by a digital filter. However, if the X has a broad band power spectrum, then filtering by frequency cannot be applied. If the signal is produced by a certain deterministic chaotic system or described approximately by a manifold with a low dimension in the phase space then well-developed embedding technique of noise reduction can be employed [3,4]. Unfortunately, all these assumptions are too strict for many time series, such as proxy data, historical ones and paleodata, which are dealt with, when investigating paleoclimate and activity of the Sun [5,6]. Data of this kind are usually the result of

*Corresponding author.
E-mail addresses: karimova@math.kz, chaos@math.kz (L. Karimova).
[1]Participation in ACAT'2003 is support by ACAT03-Grant.

0168-9002/$ - see front matter
doi:10.1016/j.nima.2004.07.042

interference of large number of uncontrollable processes, for which correct models do not exist at all. Neither the functional relation $F(X, B)$ between the signal and noise nor the noise nature are known for such data. In this case nonlinear noise reduction methods can give erroneous results, because any verification whether the denoised signal is true is in fact impossible.

In this work we apply a more general approach to such data, which has been offered in the paper [7], for enhancement of paleodata. The method idea is based on properties of time series smoothness (Hölder regularity). Measured regularity is increased knowingly by a controllable constant value. After that a new signal with the obtained (prescribed) regularity is constructed. This signal is considered as a clean one. Thus, the noise reduction problem comes to the procedure of reconstructing a new signal on the basis of enhanced (prescribed) regularity [8].

2. Background

In our case we omit the requirement of explicit determination of the type of noise B and the clarification of what is the kind of interrelation $Y = F(X, B)$ between the noise and clean signal X. So, our primary task is the estimation of a signal local Hölder exponent [9]. More precisely we must associate each signal Y with its regularity function α_Y. Afterwards, for each point of the Y domain, the procedure of increasing local Hölder exponent is performed, so that the reconstructed signal $\tilde{X}$ regularity function is determined in the following way: $\alpha_{\tilde{X}} = \alpha_Y + \delta, \delta > 0$. Appropriateness of this procedure is confirmed by the fact that for a given signal its local Hölder exponent nearly everywhere is lower than it is in the case of the clean signal. It is due to typical noise non-regularity and is provided by functional analysis theorem about the regularity properties of a function composition.

At the next step we face the problem of function reconstruction based on its prescribed regularity obtained from enhancement (increase by δ) of local Hölder exponent. In offered method [7] this task is accomplished by means of some functional minimization. The latter guarantees the closeness of the reconstructed signal ($\tilde{X}$) and observational one (Y), and preserves the similarity of their regularity functions. Such a strategy allows us to process signals with essential non-regularity and even with singularities, simultaneously preserving information about a signal presented in its regularity function.

Now it is necessary to give the definition of pointwise regularity and the Hölder exponent. In general [9], for $x_0 \in R$ and a function $f(x): R \to R$, a positive real number α is called the regularity index of f at x_0, if there are a constant C and a polynomial $P(x)$ of order smaller than α so that, for all x in a neighborhood of x_0:

$$|f(x) - P(x - x_0)| \leqslant C|x - x_0|^{\alpha}. \quad (1)$$

The Hölder exponent $\alpha_f(x_0)$ is the supremum of all α such that (1) holds. If f is n times continuously differentiable at the point x_0, then one can use for the $P(x - x_0)$ the n-order Taylor expansion of f at x_0 and thus prove that $\alpha_f(x_0) > n$. Thus, the Hölder exponent $\alpha_f(x_0)$ measures how irregular f is at the point x_0: the higher the exponent, the more regular the f. Since α_f is defined at each point x, we may associate to $f(x)$ the function $x \to \alpha_f(x)$, which measures the evolution of its regularity. So, increasing α_f by a positive constant may be interpreted as a "smoothing" procedure. The numerical estimation of the Hölder exponent is rather simple [9]. There can be a chosen function, for example mother wavelet, $\psi(x)$, which is well localized and orthogonal to all the polynomials $P(x)$ up to some order, so that integration of both sides of (1) against $\psi(x - x_0/s)$ gives:

$$W(s, x_0)[f] = \frac{1}{s} \int \psi\left(\frac{x - x_0}{s}\right) f(x)\, \mathrm{d}x \sim s^{\alpha_f(x_0)} \quad (2)$$

at $s \to 0$. The function $W(s, b)[f]$ is called continuous *wavelet transformation* of f, where b and s are translation and scale parameters, respectively [10]. Thus, the Hölder exponent α_f can be obtained by estimating the power-law behavior of the wavelet transform at the position x_0, when the scale s varies.

Often discrete wavelet transformation is applied and b and s need to be discrete. For $s = 2^j$ and $b = 2^j k$ one can obtain orthonormal basis function for certain choices of ψ: $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$, where j indicates the scale, and k indicates a shift along the time axis. Then one gets $X(t) = \sum_{j,k} x_{jk}\psi_{jk}(t)$, where x_{jk} are wavelet decomposition coefficients of function $X(t)$. Supposing ψ has a sufficient regularity and sufficiently large number of vanishing moments, one can get [9] discrete analog of Eq. (2) in the form of

$$|x_{j,k}| \leqslant C2^{-j(\alpha_X + \frac{1}{2})}, \tag{3}$$

where C is a constant. Expression (3) is inequality, however, to estimate $\alpha_X(t)$, the application of a linear regression $\log_2|x_{j,k}|$ relative to such scales j is possible, as indices (j,k) correspond to the functions $\psi_{j,k}$, whose supports contain the point t.

3. Signal enhancement algorithm

Algorithmically, the task of signal enhancement can be accomplished in the following way [7]. Let X be an original (true) signal, while Y is an observational time series corrupted by noise. It is necessary to find a signal $\tilde{X}$ of smoothed regularity structure so as to meet the following conditions: (1) $\tilde{X}$ and Y must be close in L^2 metric space, (2) the function of local regularity $\alpha_{\tilde{X}}$ must be prescribed. If the function α_X is known, then we suppose $\alpha_{\tilde{X}} = \alpha_X$. If we do not know α_X, we estimate α_Y. Next we determine $\alpha_{\tilde{X}} = \alpha_Y + \delta$, where a value of $\delta > 0$ is chosen according to practical expediency.

Once $\alpha_{\tilde{X}}$ is determined it is necessary to find a procedure, which would allow to estimate and modify simultaneously function α_Y of the observational signal Y as well as modify the given signal so as to obtain a clean signal $\tilde{X}$ with its prescribed regularity at the final stage. Such a procedure can be implemented only if relationship between the signal representation and its regularity structure is available and determined explicitly, where the *signal representation* means a certain mapping of the signal into some space with the help of a corresponding transformation. The wavelet-analysis of signals is a tool which can meet the requirements posed [9,10].

Consequently, the next step of the algorithm is to choose an appropriate family of basis wavelet functions. Then the signal Y is decomposed on this basis, and we obtain coefficients $\{y_{j,k}\}$. After that one should determine a procedure of the coefficients transformation in order to obtain coefficients $\{\tilde{x}_{j,k}\}$ which satisfy (3) for a given $\alpha_{\tilde{X}}$. As a result we reconstruct the required signal $\tilde{X}$ on the chosen basis.

Formally, for a given time series $Y = \{Y_1, Y_2, \ldots, Y_{2^n}\}$ and a given Hölder function $\alpha_{\tilde{X}}$, it is necessary to find a signal $\tilde{X}$, such as it provides a minimum to $\|\tilde{X} - Y\|_{L^2}$ and regression of logarithms of decomposition coefficients of $\tilde{X}$ at any point i $(i = 1, 2, \ldots, 2^n)$ equals $-(\alpha(i) + \frac{1}{2})$.

This can be expressed as [7] a search of minimum of

$$\sum_{j,k} (y_{j,k} - \tilde{x}_{j,k})^2 \to \min \tag{4}$$

subject to the restriction that

$$\sum_{j=1,n} s_j \log |\tilde{x}_{j,[(i-1)2^{i+1-n}]}| = -M_n\left(\alpha(i) + \frac{1}{2}\right), \tag{5}$$

where $[\cdot]$ denotes an integer part of a real number, and coefficients $s_j = j - \frac{n+1}{2}$, $M_n = \frac{n(n-1)(n+1)}{12}$ are determined uniquely from a given wavelet representation of the signal.

Searching the global minimum of (4)–(5) with standard optimization methods seems to be very difficult, in particular, there is no possibility to obtain an explicit formula of the $\{\tilde{x}_{j,k}\}$ coefficients estimation. Moreover, the problem is multidimensional (in the case of a transformation with scales $j = 1, 2, \ldots, 8$ the space dimension will be of 2^8 order), and the function has to be optimized as well as bound conditions are posed in nonlinear form.

To solve this problem genetic algorithm was used [11]. A population of genetic algorithm contained vectors of so-called modifiers $u_{j,k} \in [0, 1]$ of the wavelet coefficients $y_{j,k}$ of original data. The coefficients of the enhanced signal equaled the corresponding coefficients of Y multiplied by certain numbers, i.e. $\tilde{x}_{j,k} = u_{j,k} \cdot y_{j,k}$, $j = 1, 2, \ldots, n$, $k = 0, 1, \ldots, 2^j - 1$. The fitness function was represented as

$$F = \sum_{j,k} ((1 - u_{j,k})\tilde{x}_{j,k})^2 + \sum_i |\alpha_u(i) - \alpha_{\tilde{X}}(i)| \quad (6)$$

according to expressions (4)–(5). For numerical experiment there was used software FracLab [12], Wavelet Toolbox of MatLab for estimation of α_Y, and library Galib [13] for genetic algorithm.

4. Results

We used the approach described above for the enhancement of actual cosmogeneous isotopes ^{14}C and ^{10}Be time series. These time series are very important in climatology, paleontology, dendrochronology and others sciences, because they allow tracing of the solar activity variation over the large time intervals [6]. However, the time series contain the strong noise component, the nature and level of which are unknown.

These time series are sufficiently irregular. Fig. 1 demonstrates Hölder exponent α characterizing the irregularity of the original ^{14}C time series. For enhancement of the time series we have chosen $\delta = 1.5$. The enhancement has resulted in suppression of the high-frequent component up to 4 years, while the low-frequent part of the spectral density has been preserved (Fig. 2). Analogous enhancement procedure, applied to the ^{10}Be time series, has led to emphasizing of the 11-years Solar component, which cannot be observed by means of wavelet analysis of the ^{10}Be original data. Besides, the enhanced time series give possibility to estimate dynamical characteristics of these data with the help of deterministic chaos methods.

It should be noted that the method can be applied under the most general assumptions about signal and noise nature.

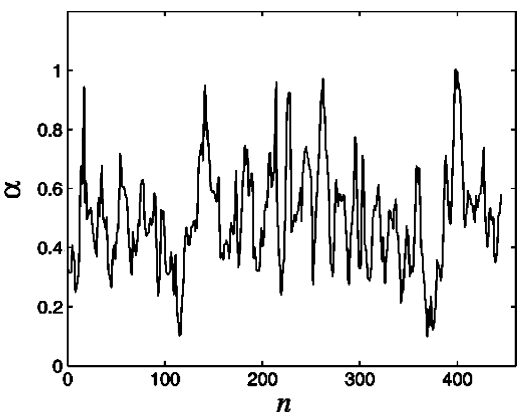

Fig. 1. The Hölder regularity function α of the ^{14}C time series.

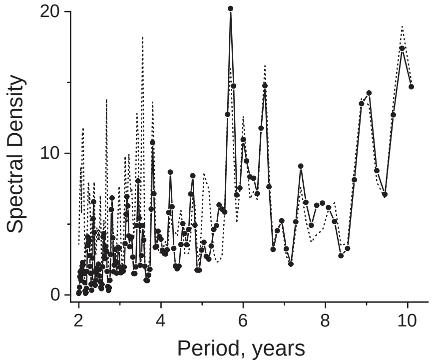

Fig. 2. Periodograms of the original (dotted line) and enhanced (line with circles) ^{14}C time series.

Acknowledgements

The support from INTAS Grant No 2001-0550 is gratefully acknowledged.

References

[1] E.J. Kostelich, J.A. Yorke, Phys. Rev. A. 38 (1988) 1649.
[2] E.J. Kostelich, T. Schreiber, Phys. Rev. E. 48 (1993) 1752.
[3] P. Grassberger, R. Hegger, H. Kantz, C. Schaffrath, T. Schreiber, Chaos 3 (1993) 127.
[4] J. Bröcker, U. Parlitz, Chaos 11 (2001) 319.
[5] V.A. Dergachev, V.S. Veksler, The application of radiocarbon method for investigation of environment of the past, Leningrad, 1991 (in Russian).
[6] V.A. Dergachev, N.G. Makarenko, L.M. Karimova, E.B. Danilkina, Geochronometria 20 (2001) 45.
[7] J. Lévy Véhel, E. Lutton, Evolutionary signal enhancement based on Hölder regularity analysis, EVOIASP2001, Lake Como, Italy, Springer Verlag, Berlin, Lecture Notes in Computer Science 2038, 2001.
[8] K. Daoudi, J. Lévy Véhel, Y. Meyer, Constructive Approximation 014 (03) (1998) 349.

[9] A. Arneodo, E. Bacry, S. Jaffard, J.F. Muzy, J. Stat. Phys. 87 (1997) 179.
[10] M. Holschneider, Wavelets. An Analysis Tool, Clarendon Press, Oxford, 1995.
[11] P. Collet, E. Lutton, M. Schoenauer, J. Louchet, Take it EASEA, PPSN VI, Paris, France, September 16–20, Lecture Notes in Computer Science 1917, Springer Verlag, Berlin, 2000.
[12] FracLab, Available ⟨http://www-rocq.inria.fr/fractales//⟩.
[13] M. Wall, Galib homepage: ⟨http://lancet.mit.edu/ga, MIT⟩.