NONLINEAR DYNAMICS METHOD IN THE APPLICATION TO THE STUDY OF TIME SERIES
(selected from CEMA'22 Conference)

N. Ampilova*

* St. Petersburg State University, Comp. Sci. Dept.
Email : n.ampilova@spbu.ru

Abstract

Time series are widely used for representation data of different types. Along the traditional methods the approach of nonlinear dynamics – reconstruction of the attractor of the system generating the series – is successfully applied. It allows us to calculate correlation dimension of the attractor of the system under study (if it exists) or to establish that the system does not have any attractor.

In this work we apply this method to solve a practical problem to analyze EEG records for revealing the patients with epileptic activity. Additionally, we calculate entropy of a signal on amplitude coverage. This approach resulted in separation of 15 records into 2 classes – epileptic activity and other pathologies, and it is in accordance with the expert conclusions.

The implemented program system may be used both for investigations and educational purpose, and the method may be applied to time series of other types.

1. INTRODUCTION

The notion of time series naturally appears in practice of data processing and statistical analysis. Time series is an ordered sequence of pairs of measured values, one if which is time and other may have a different nature and dimension. Time series are the results of experiments, both real and computational. In particular, the records of various signals are time series. Main problems for time series are   Identification problem – for given observation data to find parameter of a system which generated this series.

Prognosis problem – for given observation data to predict future values of measured characteristics. The union of traditional methods of investigation of time series with the theory of dynamical systems lead to a new approach – the application of nonlinear dynamics methods to the investigation of time series of different nature, namely a reconstruction of the attractor of the system generating this series. [7]. The theoretical substantiation of the reconstruction idea was given by F.Takens in [14]. It is based on the reconstruction of an attractor in a space of a suitable dimension where the attractor does not have self-intersection, i.e is embedded. According to Whitney's theorem if an attractor of a system lies in a n-dimensional space then it may be embedded in a space with dimension 2n+1. Namely this dimension (embedding dimension) is calculated by well-known Grossberg-Procaccia algorithm. This algorithm calculates the correlation dimension of the attractor, which is the same for the attractor in initial space and the space of embedding. The Grossberg-Procaccia algorithm is a time delay method [2], in which from a given scalar time series one form state vectors with a given time delay. Another method of this class is the method of false neighbors.

Seemingly, nonlinear dynamics methods for the first time were used in medical applications to analyze EEG records [1], [4], [12]. Later on they were applied in geophysics, astrophysics [3], physics, economics for the analysis of financial markets 15].

The application of reconstruction algorithms for EEG analysis meets many problems, the main of which is non-stationarity of a signal – the state of a patient during recording procedure may change. In this case the record should be divided on several periods in accordance with these states. Besides that, the record length may be insufficient for correct estimation of correlation dimension. The existence of stochastic noise is one more problem.

Research experience in this area shows that the choice of the parameters of reconstruction – time delay and the value of proximity (ε) between state vectors depends on the type of a record and its length. It was noticed in [8] that sometimes the length of time series does not allow the correct choice of ε, and a modified algorithm for the calculation of correlation dimension was proposed. In [9], [10] the author considered an optimized algorithm for calculation of correlation dimension.

Different algorithms and their implementations, and various variants of the choice of parameters naturally lead to different estimations for correlation dimension. However, researchers note that when solving many practical problems, it is the changing of correlation dimension (for different types of records) but not its value is important. In this situation an error in calculation is not essential, and the results of calculations retain their significance.

It was shown in [9] that correlation dimension EEG records for children 4-6 years (recorded in the state of rest) are essentially less than for adults. In [11] the author compared dimensions for 16 channels and revealed the synchronization of α-rythms in different parts of brain. The authors of [13] obtained the estimation of correlation dimension and separated EEG records of patients with two types of disease.

One of important characteristics of time series is entropy. It may be calculated on amplitude or time coverage. For this purpose, Shannon entropy and the class of Renyi entropies are widely used. In [3], [6] the authors applied so called permutation entropy to determine the degree of noise for a time series. The union of nonlinear dynamics methods and entropy characteristics gives a more detailed description of a system under study.

Thus, nonlinear dynamics methods are applied for solving the problems of time series analysis. In this work we present the program system for solving identification problem – reconstruction of the attractor of a system and estimation of correlation dimension of the attractor by a time series generated by the system, and calculation entropy characteristics. The paper has the following structure. In the next section main notions, the description of the methods of calculation of correlation dimension and the entropy for amplitude coverage are given. Section 3 contains the results of experiments for discrete and continuous dynamical systems and EEG records.

2. MAIN NOTIONS

Scalar time series is an array on N numbers which are values of a variable x(t) at the moments $t_i = t_0 + \tau(i-1)$,where τ is called sampling period [7]. We should make a remark about the choice of parameter τ.

Time series may be obtained as trajectories of discrete dynamical systems or results of numerical integration of continuous ones. In the first case $\boldsymbol{\tau}$ =1, at that time in the last case parameter $\boldsymbol{\tau}$ is the step of the numerical method. When recording signal from encephalograph, this parameter depends on the recording device. It means that the time between two consecutive values of time series depends on the method of its forming. As the result, when showing restored attractor, we may obtain different representations.

2.1. Takens method

Let $\varphi_t(x)$ be a n-th order dynamical system defined on a compact N-dimensional manifold M, and let we obtained a time series as a result of observation of the system functioning on a coordinate j. Then $\varphi_\tau^j(x)$ is the value of j-th component of $\varphi_t(x)$ at the time τ. If the system has an attractor $A \subset M \subset R^n$, it may be restored in Euclidean space with dimension $2N+1$.

Define the map $F: M \to R^{2N+1}$ as the follows $F(x) = \left(\varphi_0^j(x), \varphi_\tau^j(x), \dots \varphi_{2N\tau}^j(x)\right)$, where τ is a period of the sample. In what follows we omit the denotation j for simplicity. Construct the vectors from the data of the time series $z_0 = (\varphi_0(x), \varphi_1(x), \dots, \varphi_{2N}), \dots z_i = (\varphi_i(x), \varphi_{i+1}(x), \dots, \varphi_{i+2N}), \dots, z_{K-2N} = (\varphi_{K-2N}(x), \varphi_{K-2N+1}(x), \dots, \varphi_K)$, where K is the length of the segment of the time series. In other words we construct z_i as a point in the space R^{2N+1}. By the Takens theorem [14] F is embedding M in R^{2N+1}, and it is the generic property. Hence, we have two systems: $\varphi: M \to M$, and $F: M \to R^{2N+1}$, which are connected by a nondegenerate change of variables $z = F(x)$. There is the characteristic that is invariant with respect to this change — correlation dimension, and we may obtain the properties of the attractor of the initial system as the properties of its copy in R^{2N+1}.

To determine the dimension of the embedding we follow the algorithm proposed by Grassberg and Procaccia [5]. It proposes to find such N for which there exists a functional dependence between values of the time series. If the system has an attractor then the points (trajectories) constructed by the time series are close. To estimate the closeness of points we use correlation integral and then calculate correlation dimension of the attractor.

The correlation integral estimates the number of pairs of points (constructed vectors z_i) which are ε-close:

$$C(\varepsilon) = \lim_{K\to\infty} \frac{1}{K^2} \sum_{n,n_1=1}^{K} \theta\left(\varepsilon - \rho\left(z_n, z_{n_1}\right)\right), \tag{1}$$

where K is the size the sample and θ is the Heaviside function. The correlation dimension of the attractor is defined as

$$D_c = \lim_{\varepsilon\to 0} \frac{\log C(\varepsilon)}{\log \varepsilon}, \tag{2}$$

and calculated approximately by the least square method as the angular coefficient of the line in coordinates $(\log \varepsilon, \log C(\varepsilon))$.

Thus, by changing the length of vectors z_i (denote it by k) we calculate D_c. This value may reach a stable value or not. In the first case we take the minimal value of k for the dimension of embedding, otherwise we believe that our series is a random noise, not a dynamical system.

The restored attractor is usually shown in projection to R^2 or R^3. For the plane one use coordinates (x(t), x(t+τ)), in R^3 – (x(t), x(t+τ), x(t+2τ)). In this work we use the projection on the plane.

2.2. Entropy on amplitude coverage

Consider the distribution of a signal by amplitude levels. Let x_{max}, x_{min} be maximal and minimal values of the signal respectively, and $\Delta = x_{max} - x_{min}$.. Divide Δ on N parts (levels) and define X_i as the number of $x(t)$ belonging to level i.

Define the normed distribution $\{p_i\}$ as $p_i = \frac{X_i}{\sum_i X_i}$ and calculate Shannon entropy $H(N) = -\sum_{i=1}^{N} p_i \ln p_i$.

3. NUMERICAL EXPERIMENTS

The most appropriate way to verify the Takens method is to use dynamical systems having attractors. The length of the obtained series may be taken arbitrary long. We consider examples for 3 types of data:

3.1. Henon map

The transformation is defined on R^2 and given by the formula

$$x_{n+1} = 1 - 1.4x_n^2 + y_n$$
$$y_{n+1} = 0.3x_n$$

It is well known that Henon map has attractor. The results of calculations: correlation dimension on x coordinate $D_c^x = 1.31$, correlation dimension on y coordinate $D_c^y = 1.25$.

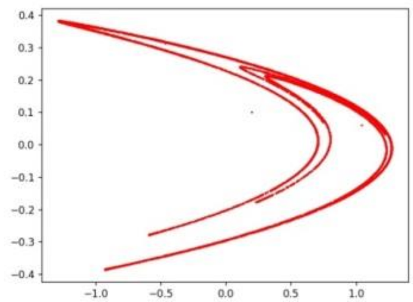

Figure1. Henon attractor. Initial point (0.2,0.1), 5000 iterations

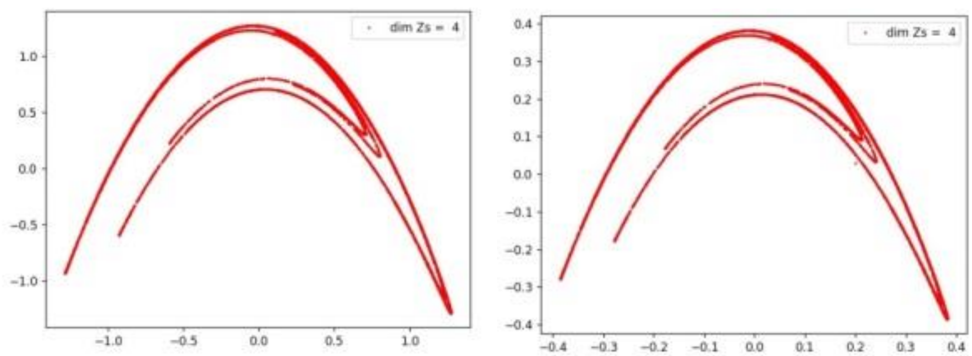

Figure 2. X- and Y- coordinate restored attractors for Henon map

3.2. Predator-prey system

The system is given by the system of differential equations and has the form

$\dot{x} = (0.7 - 0.65y)x,$

$\dot{y} = (-0.35 + 2.7x)y.$

For numerical integration we use 4th order Runge-Kutta method.

$$D_c^x = 1.05, D_c^y = 1.05.$$

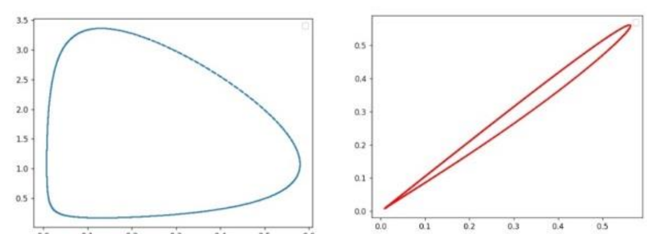

Figure 3. Attractor and x-coordinate restored attractor of predator-prey system. Initial point (0.5,0.55), 5000 iteration, $\tau = 0.1$

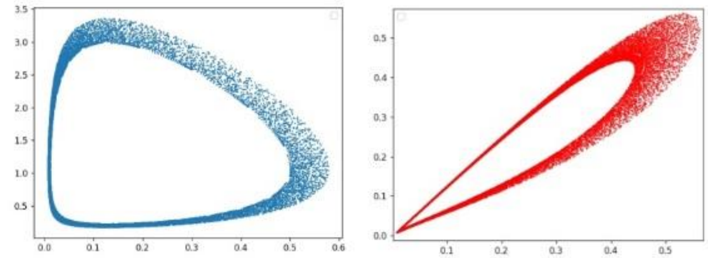

Figure 4. Attractor and x-coordinate restored attractor of predator-prey system Initial point (0.5,0.55), 5000 iteration, $\tau = 0.4$

3.3. EEG records

To calculate D_c we form vectors z_i by length k (starting from k=2) and calculate correlation integral for several values of ε by (2). Then we estimate D_c by the least square method. Increase k on 1 and repeat calculations. Compare obtained correlation dimensions. If they are close with a given accuracy δ we believe that the dimension reached a stable value. In this case we take the minimal value of k as the dimension of embedding. If dimensions are not close, we again increase k. In the situation when correlation dimension does not reach a limit value we consider the time series as a random noise, not a trace of a dynamical system.

It should be noted that one of main problems when calculating correlation integral is the choice of ε. Due to insufficient length of the time series it is difficult to take this parameter arbitrary, because the situation may occur when there are not pairs of points with such a distance between them. We use the following algorithm:

- take a sequence N_i of parts of a series by increasing length;
- for each N_i calculate distances between vectors z_i, z_j;
- take the minimal distance and one more as values of ε.

Note that to apply the least square method we need at least two values for ε.

We used 15 EEG records of patients with pathology and separated them into 2 classes – epileptic activity and other pathologies.

Example 1

Series length 514, frequency 80 Hz, time of recording 6.4s, the number of channels 16.

Results of calculation $D_c = 2.28, H = 1.37$ (using 0 channel)

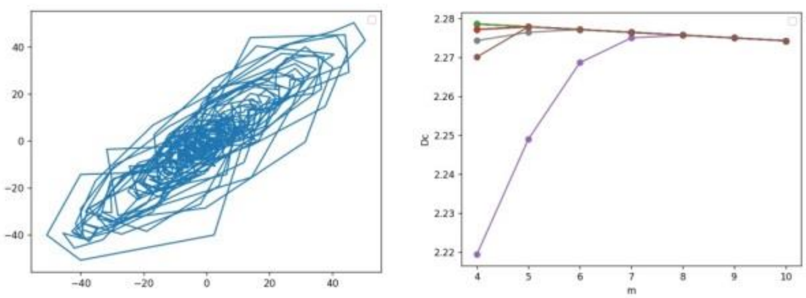

Figure 5. Restored attractor (0 channel) and the graph of stabilization of correlation dimension. Embedding dimension is 8

Example 2

The length of record 2139, frequency 80 Hz, time of recording: 26.73 s, 16 channels

$D_c \in [3.28 ; 3.44]$, H=2.59 (using 0 channel)

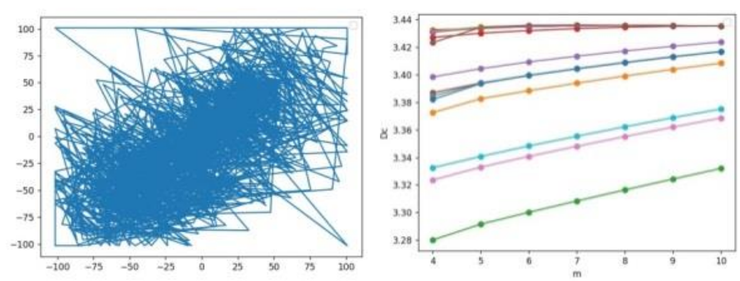

Figure 6. Restored attractor (0 channel) and graphs of correlation dimensions for 16 channels

Summarizing the results for 15 record one may conclude the following. 11 patients were preliminary considered by an expert as having epileptic activity, and 4 patients were diagnosed as having another pathology.

We use the value of interval were correlation dimensions and entropies on amplitude coverage lie. For records corresponding epileptic activity $D_c \in [2.27 ; 2.46]$, and $H \in [0.99 ; 1.37]$. In this case correlation dimension is practically the same for all channels, which means the synhronization process characterizing epilepsy.

For records of other pathology $D_c \in [0.6 ; 1.5] \cup [3.10 ; 3.44]$, $\mathrm{H} \in [2.02 ; 2.59]$, and there is no synchronization. Thus, these signs allow the differing epileptic activity from other type of pathology.

6. CONCLUSION

In this work we implemented the investigation of time series by nonlinear dynamics method – reconstruction of attractor and estimation of correlation dimension. The entropy of a time series on amplitude coverage is calculated as an additional characteristic. Continuous and discrete dynamical systems having attractor and EEG records were considered as test examples.

For dynamical systems the correlation dimension is in accordance with known results, being the calculation is easier than for capacity dimension. For continuous systems the dependence of graphical representation of attractor on the choice of the step of numerical method is illustrated. For EEG records the implemented method allowed the separation of data on 2 classes, which in agreement with expert diagnosis.

These algorithms may be modified and applied to analysis of more complex time series.

ACKNOWLEDGEMENTS

The author thanks V. Lavruhin for help in computer experiments.

REFERENCES

[1] A. Babloyantz, A. Destexhe, "Low dimensional chaos in an instance of epilepsy". Proc. Nat. Acad. Sci. (USA), 1986, a3: 3513-3517.

[2] B.P. Bezruchko, D.A. Smirnov, "Mathematical modeling and chaotic time series", 2005 ,izd. GosUNZ "College", Saratov.(in Russian)

[3] O.V. Chumak, "Entropies and fractals in data analysis", https://www.researchgate.net/publication/235247584,2022-04-05 (in Russian)

[4] G.W. Frank, T. Lookman, M.A.H. Nerengerg, C. Essex, J. Lemieux, and W. Blume, "Chaotic time series analyses of epileptic seizures". Physica D, 1990, 46: 427-438.

[5] P. Grassberger, I. Procaccia "Measuring the strangeness of strange attractors.", Physica D. №1, v.9, 1983. P. 189-208.

[6] S.A. Makarkin, A.V. Starodubov, Y.A. Kalinin, "Application of permutation entropy in the analysis of chaotic, noise and chaotic-noise series", Journal of technical physics, 2017, v. 87, N. 11, .1712-1717.

[7] G.G. Malinezkii, A.B. Potapov, "Modern problems of nonlinear dynamics", 2002, URSS, Moskva (in Russian).

[8] O.Y. Mayiorov, V.N. Fenchenko, "Calculation of correlation dimension and entropy EEG signals on cluster computer systems", Klinicheskaia informatika I telemedicine,2014,v.10,N.11,10-20 (in Russian).

[9] A.A. Mekler, "Application of nonlinear analysis of dynamical systems for EEG signal processing", Vestnik novih medicinskih teknologii, 2007,v. Xiv, № 1, 73-77(in Russian).

[10] A.A. Mekler, "Program system for the EEG analysis by the dynamical chaos methods", PhD Thesis, SpB, 2006.(in Russian).

[11] D.A. Nikolaeva, "Application of correlation dimension method for analysis of EEG of patients with epilepsia",. Differential equations and control processes, №2, 2009,43-51,https://diffjournal.spbu.ru/pdf/darina.pdf (in Russian).

[12] J. Roschke, J. Aldenhoff, "The dimensionality of human's electroencephalogram during sleep", Biol.Cybernetics,1991,64,p 307-313.

[13] M.I. Shpitonkov, "Calculation of correlation dimension for physiological time series", Trudi ISA RAN,2020, v.702,75-79 (in Russian).