

Finding Chaos in Finnish GDP

Radko Kříž

Faculty of Electrical Engineering, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

Abstract: The goal of this paper is to analyze the Finnish gross domestic product (GDP) and to find chaos in the Finnish GDP. We chose Finland where data has been available since 1975, because we needed the longest time series possible. At first we estimated the time delay and the embedding dimension, which is needed for the Lyapunov exponent estimation and for the phase space reconstruction. Subsequently, we computed the largest Lyapunov exponent, which is one of the important indicators of chaos. Then we calculated the 0-1 test for chaos. Finally we computed the Hurst exponent by rescaled range analysis and by dispersional analysis. The Hurst exponent is a numerical estimate of the predictability of a time series. In the end, we executed a recurrent analysis and displayed recurrence plots of detrended GDP time series. The results indicated that chaotic behaviors obviously exist in GDP.

Keywords: Chaos theory, gross domestic product (GDP), time series analysis, phase space reconstruction, Hurst exponent, largest Lyapunov exponent, recurrent analysis.

1 Introduction

Humanity has always been concerned with the question of whether the processes in the real world are deterministic in nature. Determinism can be understood variously. We know a lot of kinds of determinism, for example, philosophical, biological, economic, genetic, cultural, social, psychological, theological, technological, linguistic, environmental and mathematical. In this paper, we assume a mathematical sense of determinism, which is given by equations and initial conditions. Mathematical models that are not deterministic because they involve randomness are called stochastic.

Are the processes in the real world deterministic or stochastic in nature? The answer to this question is not clear. Both deterministic and stochastic dynamical systems have occupied an important place in the history of science. The first ideas were formulated in antiquity. In the beginning of the 18th century, there were seen many scientific progresses and discoveries. After Newton, most scientists worked on the assumption that the universe is controlled by strict natural laws. We are reminded of Laplace's demon^[1]. Laplace's demon is a universal scientist who is not limited by technical and mental restrictions. According to determinism, if someone knows the precise location and momentum of every atom in the universe, their past and future values for any given time are entailed; they can be calculated from the laws of classical mechanics. We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all the forces that set nature in motion, and all positions of all items of which nature is composed. If this intellect were also vast enough to submit these data for analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain

and the future, just like the past would be present before its eyes^[1]. It was a victory for determinism. In fact, classical determinism faces insurmountable limits even within the framework of classical mechanics itself. In particular, an accurate prediction of the future requires specifying a precise initial condition, but it is impossible. The second limitation lies in the difficulty of mathematical solutions of classical equations of motion.

Quantum mechanics with Heisenberg's Uncertainty principle appeared after that. The uncertainty principle says the paths of objects can only be predicted in a probabilistic way. This is the end of Laplace's demon with respect to classical determinism, but this is not the end of generalized determinism. Generalized deterministic theory must satisfy the following conditions: 1) In the framework of this theory, it is defined as a procedure by which we can clearly define the status of the studied system. 2) The equation that is formulated, uniquely determines future states of the specified initial state. Accordingly this definition is the quantum mechanics of deterministic theory.

The universe exists and develops over time. The "Big Bang" might have caused everything or it might just have been a coincidence. We can observe and describe the universe and things around us. In order to describe reality, we must simplify it through a relevant model. The only purely stochastic process is a mathematical model described by mathematical statistics. The statistical model often works and is the only possible description if we do not know the system. If we study some system, especially an economic system, we usually know a lot of information about this system. Of course, the only purely deterministic model is still just a model. We believe that for the description of real processes, it is better to use some deterministic model, when we know something more about this process. We think that to abandon determinism and to work only with classical methods is inefficient. Real processes in nature, according to the expectation of Mandelbrot^[2], lie somewhere between pure deterministic process and white noise. This is why we can describe reality either by a stochastic or deter-

Regular Paper
Special Issue on Recent Advances on Complex Systems Control,
Modelling and Prediction
Manuscript received August 13, 2013; revised October 28, 2013

ministic model. The Hurst coefficient can give us an answer to this.

An interesting case of determinism is deterministic chaos. The only purely stochastic process is a mathematical model described by mathematical statistics. The statistical model often works and is one of many possible descriptions if we do not know the system. This also applies to economic quantities, including forecasts for gross domestic product (GDP). The basic question is therefore the existence of chaotic behavior. If the system behaves chaotically, we are forced to accept only limited predictions. In this paper, we will try to show the chaotic behavior of GDP.

The methods used to analyze time series for detecting chaos can be classified according to Gilmore^[3] into metric, dynamical, and topological tools. Metric methods depend on the computation of distances in the system's attractor, and include the Grassberger-Procaccia correlation dimension^[4]. Dynamical methods deal with computing diverging orbits by estimating Lyapunov exponents. Topological methods are characterized by the study of the organization of the strange attractor, and they include close return plots and recurrence plots^[5].

Classical metric and dynamical methods are sometimes hardly applicable to economic time series. The main problem in analyzing the GDP time series is the lack of data and noisy data set. We do not have universal algorithms that work with small data sets and are robust against noise. Faggini^[6] claims that topological tools, like recurrence analysis, can solve this problem. Recurrence is a fundamental characteristic of many dynamical systems and was introduced by Poincaré in 1890^[7]. In this paper, we use metric, dynamical and topological methods. Recurrence analysis is a relatively new technique for the qualitative analysis of time series. Thanks to this technique, we can graphically detect hidden patterns and structural changes in time series.

2 Methods of analyzing

2.1 Phase space reconstruction

According to Henry et al.^[8], the main goal in nonlinear time series analysis is to determine whether or not a given time series is of a deterministic nature. If it is, then further questions of interest are: What is the dimension of the phase space supporting the data set? Is the data set chaotic?

The key to answering these questions is embodied in the method of phase space reconstruction, which has been rigorously proven by the embedding theorems of Takens^[9]. The Takens theorem was independently invented by Packard et al.^[10] among others. The Takens' theorem transforms the prediction problem from time extrapolation to phase space interpolation. The Takens' embedding theorem asserts that measured time series need not be components of the attractor, but only a sufficiently smooth transformation or maps of the component or components (so called measurement function) of the dynamical system under study^[11].

Let there be given a time series x_1, x_2, \dots, x_N which is embedded into the m -dimensional phase space by the time

delay vectors. A point in the phase space is given as

$$Y_n = x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}, n = 1, 2, \dots, N - (m-1)\tau \quad (1)$$

where τ is the time delay, and m is the embedding dimension. Different choices of τ and m yield different reconstructed trajectories. How can we determine optimal τ and m ?

A one-to-one embedding can be obtained for any value of the time delay $\tau > 0$. However, very small time delays will result in near-linear reconstructions with high correlations between consecutive phase space points and very large delays might obscure the deterministic structure linking points along a single degree of freedom. If the time delay is commensurate with a characteristic time in the underlying dynamics, then this too may result in a distorted reconstruction.

In order to estimate τ , two criteria are important according to Kodba et al.^[12]. First, τ has to be large enough so that the information we get from measuring the value of x at time $n + \tau$ is significantly different from the information we already have by knowing the value of x at time n . Only then will it be possible to gather enough information about all other system variables that influence the value of x to reconstruct the whole attractor. Second, τ should not be larger than the typical time in which the system loses memory of its initial state. This is particularly important for chaotic systems, which are intrinsically unpredictable and hence lose memory of the initial state as time progresses.

Following this reasoning, Fraser and Swinney^[13] introduced the mutual information between x_n and $x_{n+\tau}$ as a suitable quantity for determining τ . The mutual information between x_n and $x_{n+\tau}$ quantifies the amount of information we have about the state $x_{n+\tau}$ presuming we know the state x_n . Now we can define mutual information function:

$$I(\tau) = - \sum_{h=1}^j \sum_{k=1}^j P_{h,k}(\tau) \ln \frac{P_{h,k}(\tau)}{P_h P_k} \quad (2)$$

where P_h and P_k denote the probabilities that the variable assumes a value inside the h -th and k -th bins, respectively, and $P_{h,k}(\tau)$ is the joint probability that x_n is in bin h and $x_{n+\tau}$ is in bin k . Hence, the first minimum of $I(\tau)$ marks the optimal choice for the time delay.

The embedding dimension m is conventionally chosen using the "false nearest neighbors" method. This method measures the percentage of close neighboring points in a given dimension that remain so in the next highest dimension. The minimum embedding dimension capable of containing the reconstructed attractor is that for which the percentage of false nearest neighbors drops to zero for a given tolerance level μ .

In order to calculate the fraction of false nearest neighbors the following algorithm is used according to Kennel^[14]. Given a point $p(i)$ in the m -dimensional embedding space, one first has to find a neighbor $p(j)$, so that

$$\|p(i) - p(j)\| \leq \mu. \quad (3)$$

We then calculate the normalized distance R_i between the

$(m + 1)$ -th embedding coordinate of points $p(i)$ and $p(j)$ according to the equation:

$$R_i = \frac{|x_{i+m\tau} - x_{j+m\tau}|}{\|p(i) - p(j)\|}. \quad (4)$$

If R_i is larger than a given threshold R_{tr} , then $p(i)$ is marked as having a false nearest neighbor. Equation (4) has to be applied for the whole time series and for various $m = 1, 2, \dots$ until the fraction of points for which $R_i > R_{tr}$ is negligible^[12].

2.2 The largest Lyapunov exponent

Lyapunov exponent λ of a dynamical system is a quantity that characterizes the rate of separation of infinitesimally close trajectories. Quantitatively, two trajectories in phase space with initial separation δZ_0 diverge.

$$\delta Z(t) \approx e^{\lambda t} |\delta Z_0|. \quad (5)$$

The largest Lyapunov exponent can be defined as

$$\lambda = \lim_{\delta Z_0 \rightarrow 0, t \rightarrow \infty} \frac{1}{t} \ln \frac{|\delta Z(t)|}{|\delta Z_0|}. \quad (6)$$

The limit $\delta Z_0 \rightarrow 0$ ensures the validity of the linear approximation at any time. Largest Lyapunov exponent determines a notion of predictability for a dynamical system. We can say that Lyapunov exponent measures local instability^[6]. A positive Lyapunov exponent is generally regarded as necessary but not sufficient to presence of chaos. A positive largest Lyapunov exponent is usually taken as an indication that the system is chaotic (provided some other conditions are met, e.g., phase space compactness)^[15].

We have used the Rosenstein algorithm^[16], which counts the largest Lyapunov exponent as

$$\lambda_1(i) = \frac{1}{i\Delta t} \cdot \frac{1}{(M-i)} \sum_{j=1}^{M-i} \ln \frac{d_j(i)}{d_j(0)} \quad (7)$$

where $d_j(i)$ is distance from the j point to its nearest neighbor after i time steps, and M is the number of reconstructed points. For more information see [16, 17]. Rosenstein algorithm is based on the following steps:

- 1) Estimate lag and mean period using the fast Fourier transform (FFT).
- 2) Reconstruct attractor dynamics using method of delays.
- 3) Find nearest neighbors. Constrain temporal separation.
- 4) Measure average separation of neighbors. Do not normalize.
- 5) Use least squares to fit a line to the data.

Generally, Lyapunov exponent estimates of economic data may not be so reliable, because the estimate requires a large number of observations. We chose Rosenstein algorithm, because it is easy to implement and fast because it uses a simple measure of exponential divergence that circumvents the need to approximate the tangent map. The algorithm is also attractive from a practical standpoint because it does not require large data sets. Furthermore, the method is accurate for small data sets because it takes advantage of all the available data^[16].

2.3 The 0-1 test for chaos

New test for the presence of deterministic chaos was developed by Gottwald and Melbourne^[18]. Their 0-1 test for chaos takes as input a time series of measurements, and returns a single scalar value usually in the range 0-1. In contrast the 0-1 test does not depend on phase space reconstruction but rather works directly with the time series given. The input is the time-series data and the output is 0 or 1, depending on whether the dynamics is non-chaotic or chaotic.

Briefly, the 0-1 test takes as input a scalar time series of observations ϕ_1, \dots, ϕ_N . We have used the algorithm according to Dawes and Freeland^[19]. First, we must fix a real parameter c and construct the Fourier transformed series:

$$z_n = \sum_{j=1}^n \phi_j e^{ijc}, \quad n = 1, \dots, N. \quad (8)$$

Then we have computed the smoothed mean square displacement:

$$M_c(n) = \frac{1}{N-p} \sum_{j=1}^{N-p} |z_{j+n} - z_j|^2 - \left(\sum_{k=1}^N \frac{\phi_k}{N} \right)^2 \frac{1 - \cos nc}{1 - \cos c}. \quad (9)$$

Finally we have estimated correlation coefficient to evaluate the strength of the linear growth

$$r_c = \frac{\text{cov}(n, M_c(n))}{\sqrt{\text{cov}(n, n)\text{cov}(M_c(n), M_c(n))}}. \quad (10)$$

2.4 Correlation dimension

Euclidian dimension is given by the number of phase variables. However, for a deeper understanding of behavior of dynamical systems, we must define fractal dimension. The term “fractal” was first introduced by Mandelbrot^[2]. It comes from the Latin fractus, meaning an irregular surface like that of a broken stone. Fractals are non-regular geometric shapes that have the same degree of non-regularity on all scales. A fractal is a mathematical set that has a fractal dimension. There are many specific definitions of fractal dimension. Generally, the fractal dimension D , is a statistical quantity that gives an indication of how completely a fractal appears to fill space, as one zooms down to finer and finer scales. Let S be a set of points in a space of Euclidean dimension d . We now consider certain hypercubes of side ε , and calculate the minimum number of such cells, $N(\varepsilon)$, necessary to “cover” S . Let definite capacity dimension (Kolmogorov dimension), which is typical and common, be an example of fractal dimension:

$$D = \lim_{\varepsilon \rightarrow 0} \frac{\ln(N(\varepsilon))}{\ln \varepsilon^{-1}}. \quad (11)$$

Notice that fractal dimension is a real number. A non-integer dimension does not imply chaotic dynamic, but all strange attractors must have non-integer fractal dimensions^[20].

In practice, capacity dimension cannot be computed easily. A different approach has been designed by Grassberger and Procaccia^[4]. The method is based on the

concept of correlation dimension suggested by Grassberger and Procaccia^[4].

Correlation dimension describes the geometric or topological properties, but not the dynamics itself. Correlation dimension describes the dimensionality of the underlying process in relation to its geometrical reconstruction in phase space. The correlation dimension itself can be an indicator of the presence of chaos. Correlation dimension is calculated using the fundamental definition. The idea behind it is to construct a function $C(r)$ called correlation integral that is the probability that two arbitrary points on the orbit are closer together than r . Define the correlation integral $C(r)$ for data set of length M :

$$C(r) = \frac{1}{M(M-1)} \sum_{\substack{i, j = 1, \\ i \neq j}}^M \Theta(r - \|y_i - y_j\|) \quad (12)$$

where Θ is the Heaviside step function.

$$\Theta(y) = \begin{cases} 0, & \text{if } y < 0 \\ \frac{1}{2}, & \text{if } y = 0 \\ 1, & \text{if } y > 0. \end{cases} \quad (13)$$

Euclidean metric is used for all calculations in this paper. When a lower limit exists, the correlation dimension is then defined as

$$D_C = \lim_{r \rightarrow 0, M \rightarrow \infty} \frac{\ln(C(r))}{\ln(r)}. \quad (14)$$

The Hurst exponent H is directly related to fractal dimension D , because maximum fractal dimension for a planar tracing is 2:

$$D + H = 2. \quad (15)$$

2.5 Long memory in time series

The Hurst exponent (H) is widely used to characterize some processes. The Hurst exponent is used to evaluate the presence or absence of long-range dependence and its degree in a time-series. For more information see [5, 21]. The Hurst exponent is a measure that has been widely used to evaluate the self-similarity and correlation properties of fractional Brownian noise, the time-series produced by a fractional Gaussian process^[22]. We can describe self-similarity process with the following equation:

$$X(at) = a^H X(t) \quad (16)$$

where a is a positive constant, and H is the self-similarity parameter, for $0 < H < 1$.

We have used two methods for computing long memory in GDP time series. First, we have used a methodology known as rescaled range analysis or R/S analysis. This method was originally developed by Hurst^[23]. To calculate the Hurst exponent, one must estimate the dependence of the rescaled range on the time span n of observation. The Hurst exponent is defined in terms of the asymptotic behavior of the rescaled range as a function of the time span

of a time series as

$$E \left[\frac{R(n)}{S(n)} \right] = Cn^H \quad \text{as } n \rightarrow \infty \quad (17)$$

where $\left[\frac{R(n)}{S(n)} \right]$ is the rescaled range; $E[y]$ is expected value; n is number of data points in a time series, C is a constant^[20].

An algorithm for calculation is used from Qian and Khaled^[24]. To calculate the Hurst exponent, one must estimate the dependence of the rescaled range on the time span n of observation. The average rescaled range is then calculated for each value of n . For a (partial) time series of length n , $Y = Y_1, Y_2, \dots, Y_n$, the rescaled range is calculated as follows:

1) Create a mean-adjusted series:

$$U_t = Y_t - \frac{1}{n} \sum_{i=1}^n Y_i, \quad t = 1, 2, \dots, n. \quad (18)$$

2) Calculate the cumulative deviate series V :

$$V_t = \sum_{i=1}^n U_i, \quad t = 1, 2, \dots, n. \quad (19)$$

3) Compute the range R :

$$R(n) = \max(V_1, V_2, \dots, V_n) - \min(V_1, V_2, \dots, V_n). \quad (20)$$

4) Compute the standard deviation S

$$S(n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2}. \quad (21)$$

5) Calculate the rescaled range and average over all the partial time series of length n . The Hurst exponent is estimated by fitting the power law, according to the definition^[20].

Second method is dispersional analysis. This method was introduced by Bassingthwaite^[25]. In the original algorithm, the $x(t)$ series is divided into non-overlapping intervals of length n . The mean of each interval is computed, and then the standard deviation (SD) of these local means, for a given length n . These computations are repeated over all possible interval lengths. SD is related to n by a power law:

$$SD \propto n^{H-1}. \quad (22)$$

The quantity $(H - 1)$ is expressed as the slope of the double logarithmic plot of SD as a function of n . Obviously, as the number of means involved in the calculation depends on the number of available intervals, the SD's calculated from the highest values of n tend to fall below the regression line and bias the estimate according to Delignieres^[26].

Dispersional analysis can be regarded as a strong method for characterizing biological or natural time series, which generally show long-range positive correlation^[27].

2.6 Recurrence analysis

Recurrence is a fundamental property of dynamical systems, which can be exploited to characterize the system's

behavior in phase space^[28]. Recurrence analysis, based on topological approach, was used to show recurring patterns and non-stationarity in time series^[29]. It was applied to study chaotic systems because recurrence is one of the most important features of chaotic systems. Thanks to recurrence analysis, it was possible to reveal a correlation in the data that was impossible to detect in the original time series. Recurrence analysis is also particularly suitable for economic time series that are characterized by noise and short data sets^[30].

A powerful tool for visualization and analysis is called recurrence plot (RP). Generally RP is an advanced technique of nonlinear data analysis. RP is a graphical method designed to locate hidden recurring patterns, nonstationarity and structural changes, introduced by Eckmann et al.^[31]. It is a visualization of a square matrix, in which the matrix elements correspond to those times at which a state of a dynamical system recurs (columns and rows correspond then to a certain pair of times).

We have a trajectory of a system in its phase space^[31]. The trajectory is vector and its components are points in phase space (1). The development of the systems is then described by a series of these vectors, representing a trajectory in an abstract mathematical space. Then, the corresponding RP is based on the following recurrence matrix:

$$\mathbf{R}_{i,j} = \begin{cases} 1 : \mathbf{y}_i \approx \mathbf{y}_j \\ 0 : \mathbf{y}_i \neq \mathbf{y}_j \end{cases}, \quad i, j = 1, \dots, N \quad (23)$$

where N is the number of considered states, and $\mathbf{y}_i \approx \mathbf{y}_j$ means equality up to an error (or distance) r . Note that this r is essential as systems often do not recur exactly to a formerly visited state but just approximately. Roughly speaking, the matrix compares the states of a system at times i and j . If the states are similar, this is indicated by a one in the matrix, i.e. $\mathbf{R}_{i,j} = 1$. If on the other hand the states are rather different, the corresponding entry in the matrix is $\mathbf{R}_{i,j} = 0$.

As our focus is on recurrences of states of a dynamical system, we define now the tool which measures recurrences of a trajectory in phase space: the recurrence plot, (23)^[31]. The RP efficiently visualizes recurrences and can be formally expressed by the matrix

$$\mathbf{R}_{i,j}(\mathbf{r}) = \Theta(r - \|\mathbf{y}_i - \mathbf{y}_j\|), \quad i, j = 1, \dots, N \quad (24)$$

where N is the number of measured points \mathbf{y}_i , r is a threshold distance, Θ is the Heaviside step function (13). For r -recurrent states, i.e., for states which are in an r -neighborhood, we introduce the following notion:

$$\mathbf{y}_i \approx \mathbf{y}_j \Leftrightarrow \mathbf{R}_{i,j} \equiv 1. \quad (25)$$

The RP is obtained by plotting the recurrence matrix (24), and using different colours for its binary entries, e.g., plotting a black dot at the coordinates (i, j) , if $\mathbf{R}_{i,j} \equiv 1$, and a white dot, if $\mathbf{R}_{i,j} \equiv 0$. Both axes of the RP are time axes and show rightwards and upwards (convention). Since $\mathbf{R}_{i,i} \equiv 1$ for all $i = 1, \dots, N$ by definition, the RP has always a black main diagonal line, the line of identity (LOI).

Furthermore, the RP is symmetric by definition with respect to the main diagonal, i.e. $\mathbf{R}_{i,j} \equiv \mathbf{R}_{j,i}$ ^[28]. For more information see Marwan^[28].

As already mentioned, RP visualizes trajectories in phase space. RP shows important information in the time evolution of these trajectories, because typical patterns in RPs are linked to a specific behavior of the system. RP is always symmetrical diagonally and contains at least one of the following structures called textures: single dots, diagonal lines as well as vertical and horizontal lines (the combination of vertical and horizontal lines obviously forms rectangular clusters of recurrence points); in addition, even bowed lines may occur^[28].

1) Single, isolated recurrence points can occur if states are rare, if they persist only for a very short time, or fluctuate strongly.

2) A diagonal line occurs when a segment of the trajectory runs almost in parallel to another segment.

3) A vertical (horizontal) line marks a time interval in which a state does not change or changes very slowly.

4) Bowed lines are lines with a non-constant slope. The shape of a bowed line depends on the local time relationship between the corresponding close trajectory segments.

Diagonal lines are the most interesting structures in this analysis. These lines indicate the existence of unstable periodic orbits and thus are characteristic of the presence of determinism. If there are only diagonal lines in the recurrent plot, then it is a periodic signal. Periodic and quasi-periodic systems have RPs with diagonal oriented, periodic or quasi-periodic recurrent structures (diagonal lines, checkerboard structures). Irrational frequency ratios cause more complex quasi-periodic recurrent structures (the distances between the diagonal lines are different).

3 Analysis of GDP time series

3.1 Input data

The Finnish GDP in current prices in millions of national currency is used in this paper. We have used data from the Eurostat between the years 1975–2012 (see Fig. 1). We have used data quarterly, seasonally adjusted and adjusted data by working days and data without seasonal adjustment. According to Eurostat^[32], seasonal adjustment is a treatment of infra-annual time series to remove the spurious effect of seasonal patterns from the series' trend and cycle. These patterns can be caused by weather, public holidays such as Christmas, the timing of school vacations or of dividend payments and a number of other reasons. Sometimes there can be a problem with the seasonal adjustment data. We have solved it by an independent analysis of time series with and without seasonal adjustment.

Generally, the main problem in analyzing the GDP time series is the lack of data. That is why we chose Finland where data is available since 1975. The analysis of such short time series in the context of nonlinear dynamics or in the presence of chaos can be questionable. We know, according to Horák or Galka^[21], that for this kind of method results are provable for at least 10^3 data-points. Analysis

of short time series (order of 10^1) may lead to a spurious estimation of the invariants, e.g., largest Lyapunov exponent. Despite the above, we have no choice but to analyze GDP time series in the context of nonlinear dynamics and try to find chaotic behavior of GDP growth rate time series. Therefore, all results are only estimates. The second problem can be the presence of trends in time series. Trended data are not suitable for future analysis to study chaos dynamics. There is no universal way to remove the trend from the data set. The results often depend on how the data are detrended. The trend is removed by subtracting the linear interpolation in this paper. Denote GDP without trend as $Y(t)$ (see Fig. 2).

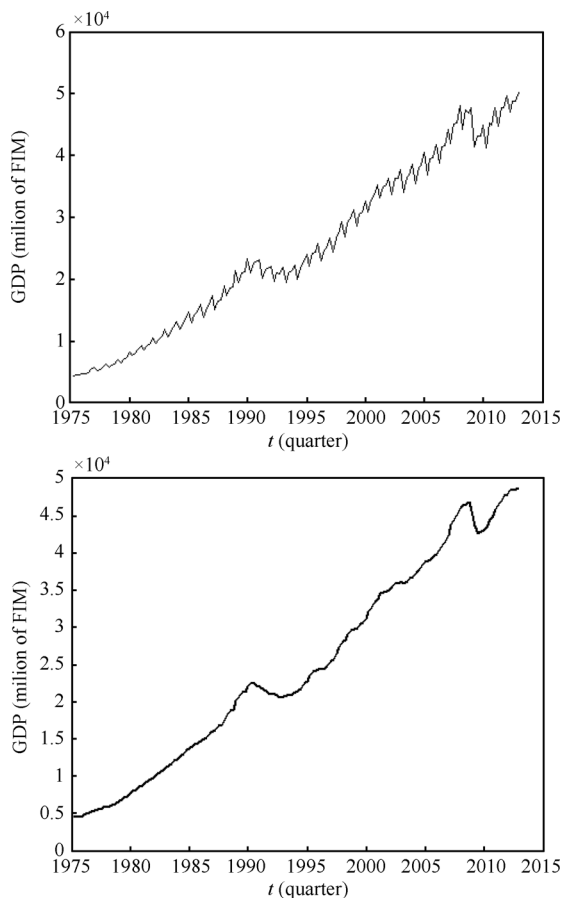


Fig. 1 The GDP of Finland. (a) Finland GDP without seasonal adjustment; (b) Finland GDP (seasonal adjustment)

3.2 Calculation of the time delay

In this chapter, we will use the mutual information approach to determine the time delay. This approach is described above. This variable is estimated from the graph (Fig. 3). The first minimum of the mutual information function $I(\tau)$ (2) marks the optimal choice for the time delay. Thus, the time delay τ is 2 for detrended GDP and 7 for seasonally adjusted detrended GDP. Values are different, because these are different time series.

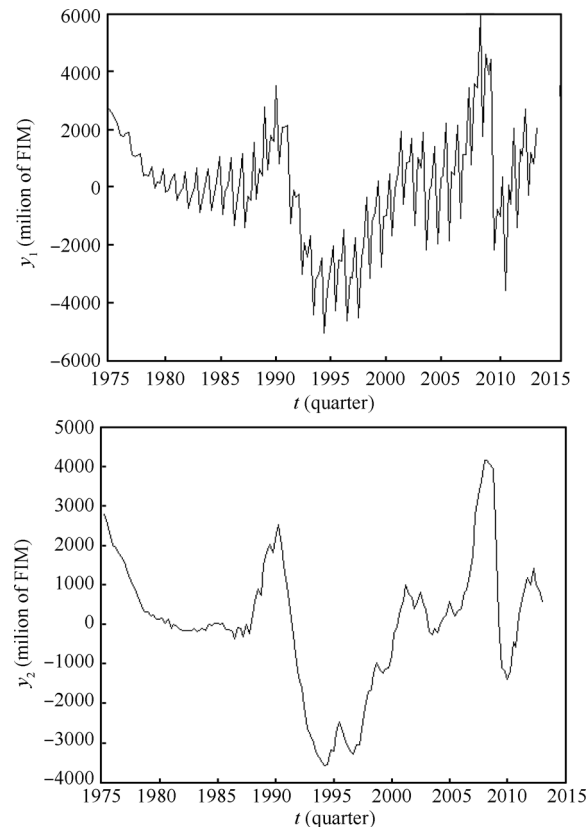


Fig. 2 Detrended GDP of Finland. (a) Detrended GDP of Finland GDP without seasonal adjustment; (b) Detrended GDP of Finland with seasonal adjustment

3.3 Calculation of the embedding dimension

In this chapter, we will use the false nearest neighbor method to determine the minimal sufficient embedding dimension. The embedding dimension m is chosen using the “false nearest neighbors” method. This variable is estimated from the graph (see Fig. 4). The minimum embedding dimension capable of containing the reconstructed attractor is that for which the percentage of false nearest neighbors drops to zero for a given tolerance level ε . Thus, the embedding dimension m is 3 for detrended GDP and 4 for seasonally adjusted detrended GDP.

3.4 Calculation of the largest Lyapunov exponents

In this chapter, we calculate the largest Lyapunov exponent as was shown above. We used the Rosenstein algorithm. The calculation of the largest Lyapunov exponent depends on the estimation of the embedding dimension. Importantly, for every relevant embedding dimension value 2-10, there is a positive largest Lyapunov exponent. A positive largest Lyapunov exponent is one of the necessary conditions for chaotic behavior. This shows that the GDP evolution is sensitive to the initial conditions. The value of the largest Lyapunov exponent was estimated at 0.013 for detrended GDP, and 0.014 for seasonally adjusted detrended GDP. Notice that both values from different time

series are very similar. It seems that seasonal adjustment of some time series is not important for the calculation of the largest Lyapunov exponent.

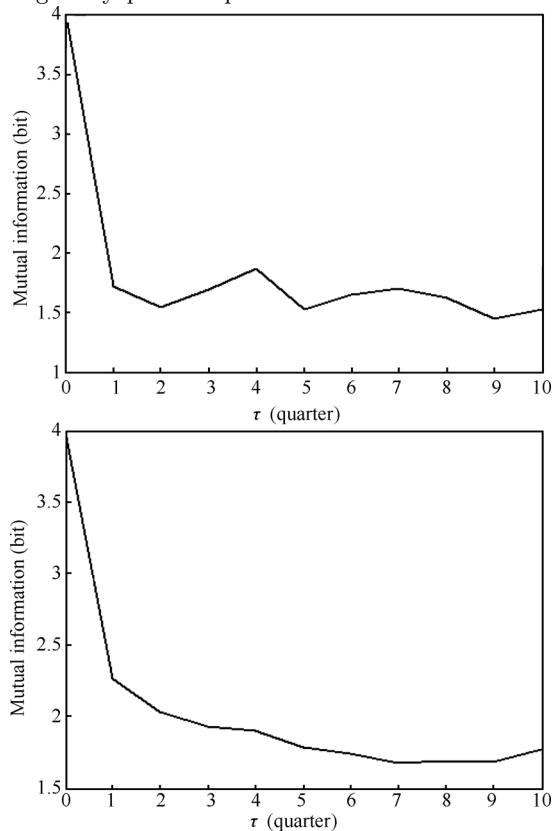


Fig. 3 Estimation of the time delay. (a) Estimation of the time delay GDP without seasonal adjustment; (b) Estimation of the time delay GDP with seasonal adjustment

3.5 Results of the 0-1 test for chaos

In this chapter, we calculate the correlation coefficient (10) as was shown above. The value of the correlation coefficient was computed at 0.95 for both time series. Again it seems that the seasonal adjustment of some time series does not affect the results of this test. The main advantage of this test is the easy interpretation of its result. The input is the time-series data, and the output is 0 or 1, depending on whether the dynamics is non-chaotic or chaotic. The correlation coefficient is near to 0 for non-chaotic data and near 1 for chaotic data. The value 0.95 is closer to 1. Hence we can assume that there is chaotic behavior in the Finnish GDP time series.

3.6 Calculation of correlation dimension

The correlation dimension is calculated using the Grassberger and Procaccia algorithm as was shown above. We have put the calculated data into a graph in logarithmic coordinates, and we have made a linear interpolation. On this basis, the correlation dimension for the small value of ε can be estimated. The estimate of the correlation dimension is 0.98 for detrended GDP and 0.91 for seasonally adjusted detrended GDP. As expected, the value of the correlation dimension is not an integer.

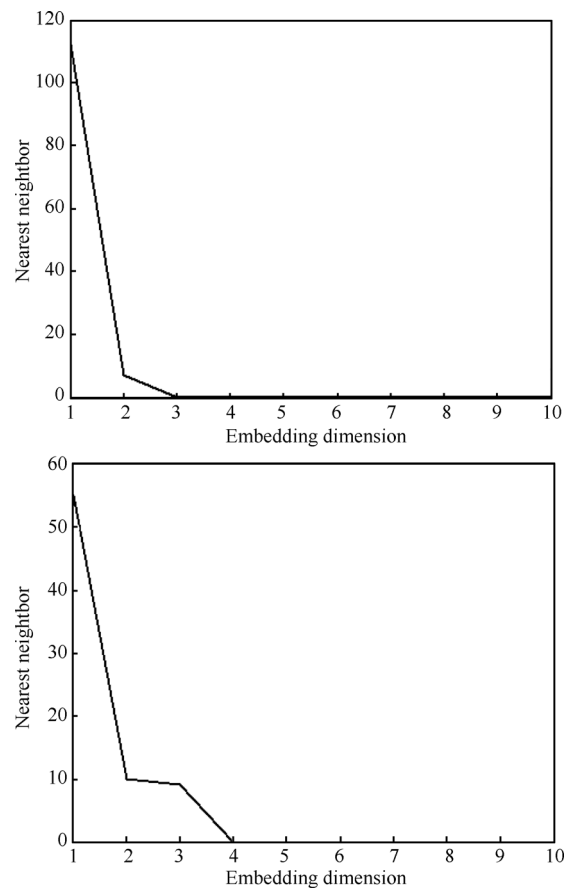


Fig. 4 Estimation of the embedding dimension. (a) Estimation of the embedding dimension GDP without seasonal adjustment; (b) Estimation of the embedding dimension GDP with seasonal adjustment

3.7 Calculation of the Hurst exponent

The rescaled range analysis gave us the value of the Hurst exponent 0.96 for detrended GDP and 0.92 for seasonally adjusted detrended GDP. The dispersional analysis gave us the value 0.87 for detrended GDP and 0.91 for seasonally adjusted detrended GDP.

All values indicate the presence of long memory in detrended GDP time series. Those values are between 0.5 and 1 which correspond with our expectations. We even predicted values somewhere between 0.75 and 1. Notice the following interesting points. First of all, the results of the seasonally adjusted detrended GDP show a higher degree of long-range dependence. Secondly, the rescaled range analysis and the dispersional analysis gave us very similar results for the seasonally adjusted detrended GDP.

We know that the value of H is between 0 and 1, whilst real time series are usually higher than 0.5. If the exponent value is close to 0 or 1, it means that the time-series has long-range dependence. We can assume that the true value lies somewhere between those values. We think that those values are sufficient for a credible prediction. Now we also know the fractal dimension $D_F = 2 - H$. We have estimated the value of the fractal dimension to lie between 1.04 and 1.13 for detrended GDP and 1.1 for seasonally adjusted

detrended GDP.

3.8 Recurrence plot

We have used in this research software called Visual Recurrence Analysis (VRA) developed by Eugene Kononov. VRA is a software package written in C++ using Borland C++ Builder 3.0 for topological analysis, qualitative and quantitative assessment, and non-parametric prediction of nonlinear and chaotic time series. Among all free software for recurrence analysis, VRA stands out as the easiest, fastest, and more user friendly program. It works under Windows in a menu-driven style, and it includes a wide range of recurrence analysis techniques^[8].

Recurrence plots of detrended GDP with and without seasonal adjustment are displayed in Fig. 5. In these recurrence plots, we can sense a certain structure and signs of diagonal lines. These recurrence plots are nothing like recurrent plots of white noise. We can thus exclude a purely random process. In both recurrence plots we can see a similar structure, but the seasonally adjusted data plot is fuzzier. In contrast, Fig. 5 (a) shows signs of areas with irregular diagonal structures. As already mentioned, diagonal lines are the most interesting structures in this analysis. We can thus exclude a purely periodic process, because purely periodic processes have only diagonal lines. These considerations confirm that it might be chaotic behavior.

3.9 Phase portrait of GDP time series

2D phase portrait of GDP is constructed so that each ordered pair of $\{Y_t; Y_{t-\tau}, t = \tau + 1, \dots, N\}$ is displayed in the plane where the x -axis represents the values of Y_t and y -axis value $Y_{t-\tau}$ (see Fig. 6). The individual points $\{Y_t; Y_{t-\tau}\}$ of phase space are connected by a line. The phase portrait of the detrended GDP without seasonal adjustment time series looks more chaotic.

4 Conclusions

Chaos theory has changed the thinking of scientists and the methodology of science. Making a theoretical prediction and then matching it to the experiment is not possible in chaotic processes. Long term forecasts are, in principle, also impossible according to chaos theory. The main problem is in the quantity and quality of data. Some improvement of measurement cannot help us adequately, because it is a fight against power of exponential rate. Nonlinear dynamics and chaos theory have also corrected the old reductionist tendency in science. Now it is known that real processes are nonlinear and a linear view can be wrong. The basic question is therefore—the existence of chaotic behavior. If the system behaves chaotically, we are forced to accept only limited predictions. But it is much better than random processes.

We have shown in this paper that the GDP time series is chaotic and contains long memory. First, we computed the values of the time delay τ and the embedding dimension m for time series with and without seasonal adjustment. These results were different, but the following results show compliance. The estimated largest Lyapunov

exponent is 0.013 for detrended GDP and 0.014 for seasonally adjusted detrended GDP. The value of the correlation coefficient in the “0-1 chaos test” was computed at 0.95 for both time series. Thus, according to the “0-1 test for chaos”, chaos is present in both time series. The estimate of the Grassberger-Procaccia correlation dimension is 0.98 for detrended GDP and 0.91 for seasonally adjusted detrended GDP. We have estimated the value of the fractal dimension to lie between 1.04 and 1.13 for detrended GDP and 1.1 for seasonally adjusted detrended GDP. Thus, the estimation of fractal dimension is close to the value 1. If the fractal dimension is low, the largest Lyapunov exponent is positive and the Kolmogorov entropy has a finite positive value, chaos is probably present. From these estimations, it can be concluded that the GDP time series is chaotic.

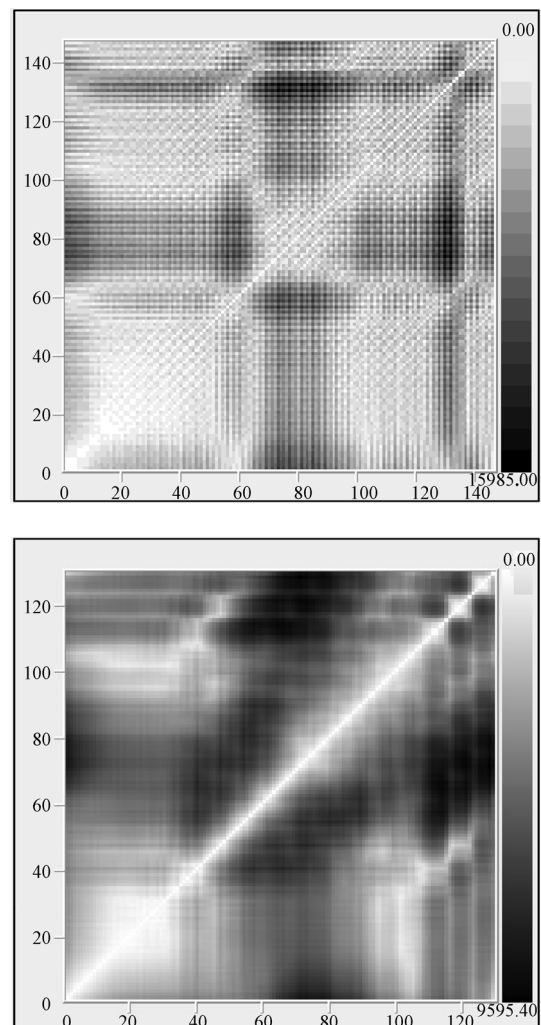


Fig. 5 Recurrence plot of detrended GDP. (a) Recurrence plot of detrended GDP with seasonal adjustment, where $\tau = 2$ and $m = 3$; (b) Recurrence plot of detrended GDP with seasonal adjustment, where $\tau = 7$ and $m = 4$

Long memory was deduced conclusively from the calculation of the values of the Hurst exponent. The values of the Hurst exponent lie between 0.75 and 1. We discovered two interesting points while analyzing long-range depen-

dence. Results for seasonally adjusted detrended GDP show a higher degree of long-range dependence. The rescaled range analysis and the dispersional analysis gave us similar results for seasonally adjusted detrended GDP.

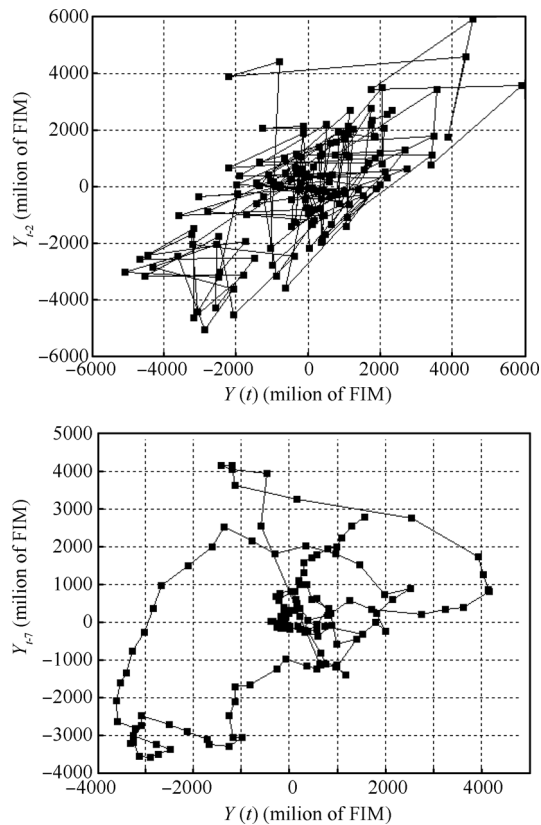


Fig. 6 Phase portrait of detrended GDP. (a) Phase portrait of detrended GDP without seasonal adjustment, where $\tau = 2$; (b) Phase portrait of detrended GDP with seasonal adjustment, where $\tau = 7$

In the end we have executed a recurrent analysis, more specifically we executed its visualization as represented by recurrence plots of detrended GDP time series with and without seasonal adjustment. In both recurrence plots we can see a similar structure, but the seasonally adjusted data plot is fuzzier. In contrast, Fig. 5 (a) shows signs of areas with irregular diagonal structures. Visually, we can exclude a purely random process as well as a purely periodic process. We can assume that this might be chaotic behavior.

We know that the main problem when analyzing GDP time series is the lack of data. In the future, we would like to focus on the proper statistical significance for nonlinearity and on predicting the GDP. In particular, the surrogate data approach (e.g., Theiler et al.^[33]) is a powerful tool for detecting actual nonlinear behavior, and distinguishing it from other phenomena.

References

- [1] P. S. Laplace. *A Philosophical Essay on Probabilities*, translated from the 6th French edition by Frederick Wilson Truscott and Frederick Lincoln Emory, New York: John Wiley & Sons, pp. 19. Dover Publications edition (New York, 1951) has same pagination, 1902.
- [2] B. B. Mandelbrot. *The Fractal Geometry of Nature*, San Francisco: W. H. Freeman and Co., 1983.
- [3] C. G. Gilmore. A new test for chaos. *Journal of Economic Behavior and Organization*, vol. 22, no. 2, pp. 209–237, 1993.
- [4] P. Grassberger, I. Procaccia. Characterization of strange attractors. *Physical Review Letters*, vol. 50, no. 5, pp. 346–349, 1983.
- [5] J. Belaire-Franch, D. C. Bayarri. Recurrence plots in nonlinear time series analysis: Free software, [Online], Available: <http://www.jstatsoft.org/v07/i09/paper>, 2001.
- [6] M. Faggini. Chaos detection in economics, Metric versus topological tools, *MPRA Paper*, University Library of Munich, Germany, 2010.
- [7] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta Mathematica*, vol. 13, no. 1, pp. 1–271. 1890. (in French)
- [8] B. Henry, N. Lovell, F. Camacho. Nonlinear dynamics time series analysis. *Nonlinear Biomedical Signal Processing, vol. II: Dynamic Analysis and Modeling*, Chap 1, M. Akay, Ed., New York: IEEE Press Series on Biomedical Engineering, pp. 1–39, 2001.
- [9] F. Takens. *Detecting Strange Attractors in Turbulence*, Berlin: Springer, vol. 98, 1981.
- [10] N. H. Packard, L. P. Crutchfield, J. D. Farmer, R. S. Shaw. Geometry from a time series. *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.
- [11] R. Kříž. Chaotic analysis of the GDP time series. In *Proceedings of Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems*, Springer, Switzerland, pp. 353–362, 2013.
- [12] S. Kodba, M. Perc, M. Marhl. Detecting chaos from a time series. *European Journal of Physics*, vol. 26, no. 1, pp. 205–215, 2005.
- [13] A. M. Fraser, H. L. Swinney. Independent coordinates for strange attractors from mutual information. *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.
- [14] M. B. Kennel, R. Brown, H. D. Abarbanel. Determining embedding dimension for phase space reconstruction using a geometrical construction. *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.
- [15] H. W. Lorenz. *Nonlinear Dynamical Economics and Chaotic Motion*, Berlin Heidelberg: Springer-Verlag, 1989.
- [16] M. T. Rosenstein, J. J. Collins, C. J. De Luca. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D*, vol. 65, no. 1–2, pp. 117–134, 1993.
- [17] T. Gotthans. Advanced Algorithms for the Analysis of Data Sequences in Matlab, Master dissertation, University of Technology Brno, 2010.
- [18] G. A. Gottwald, I. Melbourne. A new test for chaos in deterministic systems. In *Proceedings of the Royal Society of London A*, vol. 460, no. 2042, pp. 603–611, 2004.

- [19] J. H. P. Dawes, M. C. Freeland. The “0-1 test for chaos” and strange nonchaotic attractors, [Online], Available: <http://people.bath.ac.uk/jhpd20/publications/sna.pdf>, 2008.
- [20] R. Kríž. Chaos in GDP. *Acta Polytechnica*, vol. 51, no. 5, pp. 63–68, 2011.
- [21] A. Galka. *Topics in Nonlinear Time Series Analysis: With Implications for EEG Analysis*, World Scientific, 2000.
- [22] B. B. Mandelbrot, J. W. van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.
- [23] H. E. Hurst. Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, vol. 116, pp. 770–799, 1951.
- [24] B. Qian, K. Rasheed. Hurst exponent and financial market predictability. In *Proceedings of the 2nd IASTED International Conference on Financial Engineering and Applications*, IEEE, Cambridge, MA, USA, pp. 203–209, 2004.
- [25] J. B. Bassingthwaighte. Physiological heterogeneity: Fractals link determinism and randomness in structure and function. *News in Physiological Sciences*, vol. 3, no. 1, pp. 5–10, 1988.
- [26] D. Delignieres, S. Ramdani, L. Lemoine, K. Torre, M. Fortes, G. Ninot. Fractal analyses for “short” time series: A re-assessment of classical methods. *Journal of Mathematical Psychology*, vol. 50, no. 6, pp. 525–544, 2006.
- [27] J. B. Bassingthwaighte, G. M. Raymond. Evaluation of the dispersional analysis method for fractal time series. *Annals of Biomedical Engineering*, vol. 23, no. 4, pp. 491–505, 1995.
- [28] N. Marwan, M. C. Romano, M. Thiel, J. Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, vol. 438, no. 5–6, pp. 237–329, 2007.
- [29] J. P. Zbilut, A. Giuliani, C. L. Webber Jr. Recurrence quantification analysis as an empirical test to distinguish relatively short deterministic versus random number series. *Physics Letters A*, vol. 267, no. 2–3, pp. 174–178, 2000.
- [30] L. L. Trulla, A. Giuliani, J. P. Zbilut, C. L. Webber, Jr. Recurrence quantification analysis of the logistic equation with transients. *Physical Review Letters A*, vol. 223, pp. 255–260, 1996.
- [31] J. P. Eckmann, S. O. Kamphorst, D. Ruelle. Recurrence plots of dynamical systems. *Europhysics Letters*, vol. 4, no. 9, pp. 973–977, 1987.
- [32] ESS Handbook for Quality Reports, Luxembourg: Office for Official Publications of the European Communities, 2009.
- [33] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J. D. Farmer. Testing for nonlinearity in time series: The method of surrogate data. *Physica D: Nonlinear Phenomena*, vol. 58, no. 1–4, pp. 77–94, 1992.



Radko Kríž received his M. Sc. degree in economics and electrical engineering from the Czech Technical University in Prague, Czech Republic in 2001. From 2011 to 2012 he held a fellowship at Université de Montréal, Québec, Canada. He is currently a Ph.D. candidate at the Faculty of Electrical Engineering of the Czech Technical University in Prague. He works as a consultant and project manager. Since 2012, he has been assistant professor at University of Pardubice, Czech Republic. He has also worked at management level in manufacturing companies.

His research interests include chaos theory, nonlinear dynamics, economics, energetics and ecology.

E-mail: kriz@e-academia.eu