Forecasting electricity prices using nonlinear method

Radko Kříž[1], Kateřina Seinerová[2]

Abstract

The goal of this paper is to analyze the electricity prices using chaos theory and to predict using nonlinear method. 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. The results indicated that chaotic behaviors obviously exist in electricity price series. If the system behaves chaotically, we are forced to accept limited predictions. Finally we computed predictions using a radial basis function to fit global nonlinear functions to the data. In this paper we analyze electricity price series of the biggest European energy markets EEX (Central European Energy Exchange).

Key words

Forecasting, Electricity prices, Chaos theory, Time series analysis, Phase space reconstruction, Gaussian radial basis function

JEL Classification: C53

1. Introduction

Electricity is one of those commodities with which price is highly volatile and with considerable number of jumps. That is especially caused by the fact that the electricity is not possible to store. Thus the prediction of electricity price is a really complicated problem. Forecasting electricity energy prices began in the 1990's, at the beginning of the liberalization process (Kříž et Kratochvíl, 2014). It is very important for electricity traders to forecast electricity prices with the highest possible precision as the precision of the forecast is connected to the trading strategy and thus with profit or lost. An electricity price depends on the demand and supply and on the costs of transportation in the grid. The first two features are strongly dependent on the weather, economic situation, government interventions etc. These dependencies result in a very complex fluctuation of electricity prices, which may seem to be slightly chaotic. Electricity is a flow commodity with unique characteristics that influence the way it is traded and thus the behavior of spot and futures prices in the market. The price is set by an interaction between supply and demand and in set in auction at a single time for the whole 24 hours of the following day. Demand for electricity is highly inelastic. In short run, it is absolutely inelastic so that the price is determined by the supply curve completely. The curve resembles upward sloping stairway, each step approximately represents a different type of a power plant and thus a different level of marginal costs. The price on the market rises until it reaches the marginal costs for a MWh of the power plant of the next level, after that

[1] University of Pardubice/Institute of Administrative and Social Sciences, Studentská 95, 532 10 Pardubice, Czech Republic, e-mail: Radko.Kriz@upce.cz.
[2] University of Pardubice/ Institute of Mathematics and Quantitative Methods, Studentská 95, 532 10 Pardubice, Czech Republic, e-mail: Katerina.Seinerova@upce.cz.

the supply rises (Kristoufek et Lunackova, 2013). The electricity spot prices can also be split up into deterministic and stochastic parts. The deterministic part describes the seasonality and all the periodic patterns (intra-day, intra-weak, peak, weekend patterns). This part is easily observable and we can describe it simply. The stochastic part (also called volatility), on the contrary, is unobservable and seems chaotic (Kříž et Kratochvíl, 2014). Real processes in nature, according to the expectation of Mandelbrot (1983), lie somewhere between pure deterministic process and white noise. This is why we can describe reality either by a stochastic or deterministic model. Long memory in time series are characteristic with values in past in past influencing the present and future values. Few works is focused to analyze a long memory in electricity price time series e.g. Kříž et Lešáková (2016b), Kristoufek et Lunackova (2013), Weron et Przybyłowicz (2002). According Weron et Przybyłowicz (2002) the whole complex process of electricity price formation results in behavior not observed in the financial or even other commodity markets. Forecasting of any commodity or derivative price is a common problem discussed in a lot of articles. Prediction on electricity prices have been discussed in several recent papers e.g. Kratochvíl et Starý (2013). Numerous new methods of electricity prices time series analysis have been developed for dealing with nonlinear data e.g. Conejo et al. (2005), Misiorek et al. (2006), Pao (2007).

To be able to use these methods and models we will have to prove first that electricity spot prices behave chaotically and that it is adequate to use a chaotic description. The basic question is therefore the existence of chaotic behavior. If the system behaves chaotically, we are forced to accept only limited predictions. Nevertheless chaotic behavior is much better than random processes (Kříž et Kratochvíl, 2014).

2. Methodology

2.1 Chaos theory

Chaos theory allows for the reconstruction of phase space from time series, which can be used for specifying the system states (Abardanel et al., 1973). This analysis is based on Takens (1980) embedding theorems. Takens' theorem transforms the prediction problem from time extrapolation to phase space interpolation (Kříž, 2013). For more information see Kříž et Lešáková (2016b).

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:

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

where τ is the time delay and m is the embedding dimension. Different choices of τ and m yield different reconstructed trajectories. Kodba et al. (2005) discuss how we can determine optimal τ and m. We use method developed by Fraser et Swinney (1986), which is based on 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. The mutual information function is:

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

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. The first minimum of $I(\tau)$ then marks the optimal choice for the time delay.

The embedding dimension m can be 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 et al. (1982). Given a point $p(i)$ in the m-dimensional embedding space, one first has to find a neighbor $p(j)$, so that

$$\|p(i)-p(j)\| \le \mu . \tag{3}$$

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:

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

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 $m = 1, 2, \ldots$ until the fraction of points for which $R_i > R_{tr}$ is negligible Kodba et al. (2005).

Lyapunov exponent λ of a dynamical system is a quantity that characterizes the rate of separation of infinitesimally close trajectories. The largest Lyapunov exponent can be defined as follows:

$$\lambda = \lim_{\substack{\delta Z_0 \to 0 \\ t \to \infty}} \frac{1}{t} \ln \frac{|\delta Z(t)|}{|\delta Z_0|} . \tag{5}$$

A positive largest Lyapunov exponent is usually taken as an indication that the system is chaotic. We have used the Rosenstein (1983) algorithm, which counts the largest Lyapunov exponent as follows:

$$\lambda_1(i) = \frac{1}{i\Delta t} . \frac{1}{(M-i)} \sum_{j=1}^{M-i} \ln \frac{d_j(i)}{d_j(0)} , \tag{6}$$

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.

2.2 Nonlinear prediction

Predictability is one way how correlations between data express themselves (Hegger et al., 1999). Most properties of chaotic systems are much more easily determined from the governing equations than from a time series. Unfortunately, the governing equations are usually not known, except for well controlled laboratory experiments. Analyzing an empirical model, and maybe synthetic time series data generated from it, can provide a valuable consistency test for the results of time series analysis. Chaotic dynamical systems generically show the phenomenon of structural instability (Kantz et Schreiber, 1997).

According Kantz et Schreiber (1997), using the reconstructed phase space for m and τ, a functional relationship f between the current state $X(t)$ and future state $X(t + T)$ can be given as

$$X(t+T) = f(X(t)) , \tag{7}$$

where T represents the number of time steps ahead that one wishes to perform the prediction. Function f represents the approximation to unknown dynamical system. It is shown that for sufficiently large values of the embedding dimension and if some additional conditions are satisfied, the reconstructed trajectory has the same topological and geometrical properties as the system's phase space trajectory (Takens, 1981). The predictive mapping can be expressed as

$$X(t+T) = f_p(X(t)) . \tag{8}$$

The aim is to find the predictor f_p, so that $x(t + T)$ can be predicted based on the reconstructed time series. If the time series is chaotic, then f_p is necessarily nonlinear. Several local and global approaches are available in the literature to find the function f_p (Farmer et Sidorowich, 1987).

The idea of locally linear predictions is following. If there is a good reason to assume that the relation

$$s_{n+1} = f(s_n) \qquad (9)$$

is fulfilled by the experimental data in good approximation (say, within 5%) for some unknown f and that f is smooth, predictions can be improved by fitting local linear models. They can be considered as the local Taylor expansion of the unknown f, and are easily determined by minimizing

$$\sigma^2 = \sum_{s_j \in U_n} (s_{n+1} - a_n s_j - b_n)^2 , \qquad (10)$$

with respect to a_n and b_n, where U_n is the ε-neighborhood of s_n, excluding s_n, as before. Then, the prediction is

$$\hat{s}_{n+1} = a_n s_n + b_n \qquad (11)$$

The minimization problem can be solved through a set of coupled linear equations, a standard linear algebra problem (Hegger et al., 1999).

The local linear fits are very flexible, but can go wrong on parts of the phase space where the points do not span the available space dimensions and where the inverse of the matrix involved in the solution of the minimization does not exist. Moreover, very often a large set of different linear maps is unsatisfying. Therefore many authors suggested fitting global nonlinear functions to the data, i.e. to solve

$$\sigma^2 = \sum_{n} (s_{n+1} - f_p(s_n))^2 , \qquad (12)$$

where f_p is now a nonlinear function in closed form with parameters p, with respect to which the minimization is done. The results depend on how far the chosen ansatz f_p is suited to model the unknown nonlinear function, and on how well the data are deterministic at all (Hegger et al., 1999). A radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin, or alternatively on the distance from some other point x_i, called a center, or alternatively on the distance from some other point c, called a center, so that

$$\Phi(x,c) = \Phi(\|x - c\|) \qquad (13)$$

Gaussian kernel is used in this analysis.

$$\Phi(x) = e^{-c^2 x^2} \qquad (14)$$

Hence the prediction made is:

$$f_p = a_0 + \sum_{i=1}^{n} a_i \Phi(\|x - x_i\|) \qquad (15)$$

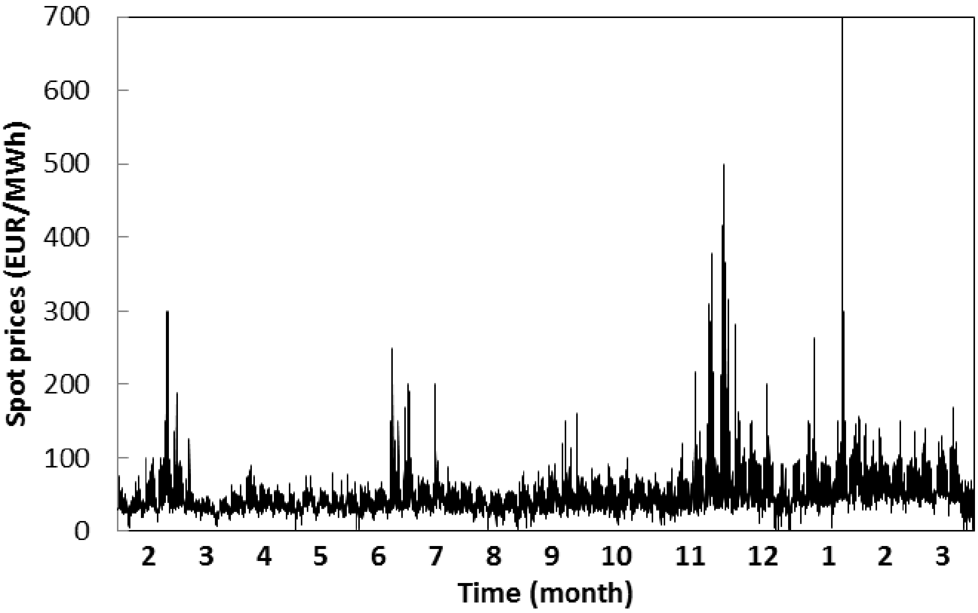
respectively

$$x_{t+1} = a_0 + \sum_{i=1}^{n} a_i e^{-(c\|x - x_i\|)^2} \qquad (16)$$

3. Empirical analysis

We used the data from one of the biggest European energy markets EEX (Central European Energy Exchange). This exchange has a lot of participants and good liquidity. We use PHELIX hourly spot prices from 8.2.2005 to 31.3.2006 (fig.1). This is 10000 samples, which is a sufficient amount of data for our study.

Figure 1: Electricity spot prices (PHELIX hourly spot prices from 8.2.2005 to 31.3.2006)

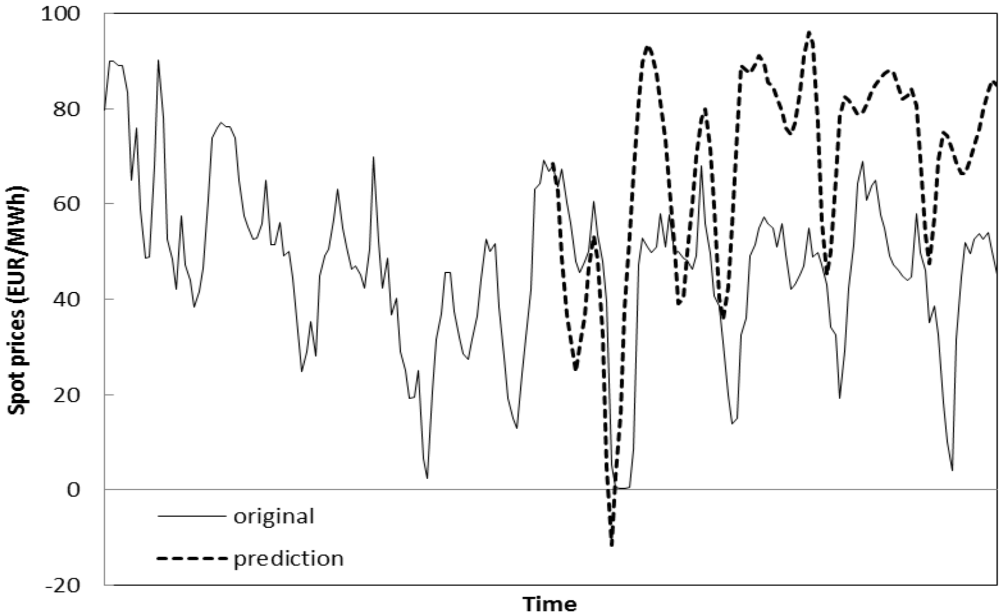


At first we used the mutual information approach to determine the time delay τ and the false nearest neighbor method to determine the minimal sufficient embedding dimension m. τ is estimated as the first minimum of the mutual information function $I(\tau)$ (2) marks the optimal choice for the time delay. Thus, the time delay τ is 7. The embedding dimension m is chosen using the "false nearest neighbors'" method, estimated from the graph. 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. Thus, the embedding dimension m is 7.

Then we calculated the largest Lyapunov exponent as was shown above. We used the Rosenstein algorithm. The value of the largest Lyapunov exponent was estimated at 0,0006 for embedding dimension 7. A positive largest Lyapunov exponent is one of the necessary conditions for chaotic behavior. This result is consistent with the work of Kříž et Kratochvíl (2014), where was shown that the electricity price series is chaotic.

Finally we computed predictions using a Gaussian radial basis function to fit global non-linear functions to the data. This prediction using RBF for the first few hours is very good even though this is an area with a high volatility (fig.2). In many cases the prediction correctly estimated the peaks and trends of the electricity prices time series. Predictions using RBF can be used in long term.

Figure 2: Electricity spot prices and its prediction (last 200 values)



4. Conclusion

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. 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 shown in this paper that the electricity price time series is chaotic. First, we computed the values of the time delay $\tau = 7$ and the embedding dimension $m = 7$. The estimated largest Lyapunov exponent is 0,0006. If the correlation 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 electricity prices time series is chaotic. Finally we computed predictions using a Gaussian RBF to fit global nonlinear functions to the data. Considering all these findings, we recommend the Gaussian RBF to fit global non-linear functions as one of the methods used for prediction. As it may not be reliable under certain circumstances, it should be used in combination with other prediction methods.

Supported by the specific science grant SGS_2016 of the University of Pardubice.

References

[1] Abarbanel, H. D., Brown, R., Sidorowich, J. J. and Tsimring, L., S. (1993). The analysis of observed chaotic data in physical systems. *Reviews of modern physics*. 65(4), p. 1331.

[2] Conejo, A. J., Contreras, J., Espinola, R., and Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting*, 21(3), pp. 435-462.

[3] Farmer, D., J. and Sidorowich, J., J. (1987). Predicting chaotic time series, *Phys. Rev. Lett*. 59, pp. 85–848.

[4] Fraser, A. M. and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical review A*. 33(2), p.1134.

[5] Hegger, R., Kantz, H., Schreiber, T. (1999). Practical implementation of nonlinear time series methods: The TISEAN package Chaos, 9(2), pp. 413-435.

[6] Kantz, H. and Schreiber, T. (1997). *Nonlinear time series analysis*. Cambridge University Press, Cambridge.

[7] Kennel, M., B., Brown, R. and Abarbanel, H., D. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A*, 45(6), p. 3403.

[8] Kodba, S., Perc, M., Marhl, M. (2005). Detecting chaos from a time series. *European journal of physics*. 26(1), p. 205.

[9] Kratochvíl, Š. and Starý, O. (2013). Predicting the Prices of Electricity Derivatives on the Energy Exchange." *Acta Oeconomica Pragensia* 2013(6). pp. 65-81.

[10] Kristoufek, L. and Lunackova, P., (2013). Long-term Memory in Electricity Prices: Czech Market Evidence. *Czech Journal of Economics and Finance*. 63(5), pp 407-424.

[11] Kříž, R. and Kratochvíl, Š. (2014). Analyses of the chaotic behavior of the electricity price series. In *ISCS 2013: Interdisciplinary Symposium on Complex Systems*. Springer Berlin Heidelberg, Germany, pp. 215-226.

[12] Kříž, R. (2013). Chaotic Analysis of the GDP Time Series. In: *Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems*. Springer. pp. 353-362.

[13] Kříž, R., & Lešáková, P. (2016a). Prediction of NOX Concentration Time Series Using the Chaos Theory. In *Proceedings of the Second International Afro-European Conference for Industrial Advancement AECIA 2015* (pp. 465-475). Springer International Publishing.

[14] Kříž, R. and Lešáková, P. (2016b). Long memory in electricity price time series. In *SGEM2016 Conference Proceedings*.

[15] Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. W.H. Freeman and Co., New York, USA.

[16] Misiorek, A., Trueck, S., and Weron, R. (2006). Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models. *Studies in Nonlinear Dynamics & Econometrics*, 10(3).

[17] Pao, H. T. (2007). Forecasting electricity market pricing using artificial neural networks. *Energy Conversion and Management*, *48*(3), pp. 907-912.

[18] Rosenstein, M., T., Collins, J., J. and De Luca, C., J. (1993). A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65(1), pp. 117-134.

[19] Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980* (pp. 366-381). Springer Berlin Heidelberg.

[20] Weron, R. and Przybyłowicz, B. (2000). Hurst analysis of electricity price dynamics. *Physica A: Statistical Mechanics and its Applications*, 283(3), pp. 462-468.