Entropy analysis of dynamic properties of regional stock markets

H. Danylchuk[1], V. Derbentsev[2*], V. Soloviev[1], A. Sharapov[2]

[1]Bohdan Khmelnitsky National University of Cherkassy, Ukraine
[2]Kyiv National Economic University by Vadym Hetman, Ukraine
*Corresponding author. E-mail: derbv@mail.ru

Paper received 2016.01.30; Accepted for publication 2016.02.15.

Abstract. This paper examines entropy analysis of regional stock markets. We propose and empirically demonstrate the effectiveness of using such entropy as sample entropy, wavelet and Tsallis entropy as a measure of uncertainty and instability of such complex systems as regional stock markets. Our results show that these entropy measures can be effectively used as crisis prediction indicators.

Keywords: *complex systems, time series, regional stock market crisis, non-regular dynamics, entropy analysis, sample entropy, Tsallis entropy, wavelet entropy.*

Introduction. It is well known that the majority of complex systems cannot be accurately described with formal models. Therefore, one of the most promising approaches to the analysis of such systems, including financial, stock and foreign-currency exchange markets, is to investigate the dynamics of their time series.

Based on the results of numerous empirical studies [1-4] it is possible to reconstruct many of the dynamic and structural properties of a studied system by observing (studying) a single variable, e.g., scalar time series. As such, it is possible to determine the system's measures of complexity and randomness, to estimate the stability, time period of behaviour predictability, time of relaxation, etc.

This paper reports the results of the analysis of regional stock markets' dynamic properties using entropy time series analysis of relevant stock indices.

Brief review of pertinent literature. The concept of thermodynamic entropy as the measures of uncertainty of the system has been widely used in physics since the late 19th century through the work of Clausius and Boltzmann [5]. The concept of entropy as a measure of uncertainty or chaos of systems of different physical nature was further developed in the 20th century. Consider for example, Shannon entropy in information theory, Kolmogorov entropy in the theory of system dynamics, von Neumann entropy in quantum mechanics [5]. In recent years, along with other current methods of time series analysis, entropy analysis has been shown to be an effective method for the study of complex socio-economic [4-8] and biological systems [9-10]. Entropy indicators proved to be sufficiently representative and informative quantitative measures, which characterize the state of the system, its degree of complexity, randomness, and structural organization.

The aim of this work is to analyse the efficiency of such entropy measures as sample entropy [9-10], Tsallis entropy [11-12] and wavelet entropy [13-14] as measures of volatility and as indicators of regional stock market crises onset for the period from January, 2004 to January, 2016.

Methodology. Many different methods of entropy calculation are used in economic applications, such as Shannon entropy, Renyi entropy, Kolmogorov *K*-entropy, approximate entropy, sample entropy , multiscale entropy, wavelet entropy [5], etc. Next, we discuss in more detail the methods for obtaining sample entropy, Tsallis entropy and wavelet entropy, that we used to analyse the degree of instability of regional stock markets.

Approximate Entropy (ApEn) and *Sample Entropy (SampEn)*. The method for computing approximate entropy is described in [4, 9-10]. Approximate entropy is "regularity statistics" that determines the possibility of predicting fluctuations in time series. Intuitively, this means that the presence of repetitive fluctuation patterns in time series (subsequences of certain length that are built on successive time series values) results in higher degree of predictability than in the series without such repetitive patterns.

In calculating *ApEn* for a given time series consisting on *N* values $t(1)$, $t(2)$, $t(3)$, …, $t(N)$, two arguments *m* and *r* need to be selected. The first parameter *(m)* indicates the length of the sample, and the second parameter *(r)* indicates similarity criterion. We study the subsequences of the time series S_N, which consist of *m* values beginning with the value *i,* that are denoted with vectors $p_m(i)$.

Two vectors (samples) $p_m(i)$ and $p_m(j)$ are similar if all the differences between pairs of their corresponding coordinates are smaller than *r*, or when

$$|t(i+k)-t(j+k)|<r \text{ для } 0\le k<m . \quad (1)$$

For set P_m of all vectors of length *m* in time series S_N we calculate the following value

$$C_{im}(r)=n_{im}(r)/(N-m+1), \quad (2)$$

where $n_{im}(r)$ is the number of vectors in P_m, that are similar to vector $p_m(i)$ (given the chosen similarity criterion *r*). The value $C_{im}(r)$ is part of a vector of length *m* that is similar to a vector of the same length, elements of which begin with *i*. For any given time series the value of $C_{im}(r)$ is calculated for each vector in P_m, and the average of $C_m(r)$ defines the distribution of similar vectors of length *m* in set S_N. Approximate entropy of time series S_N using vectors of length *m* and similarity criterion *r* is defined as follows:

$$ApEn(S_N, m, r) = \ln(C_m(r)/C_{m+1}(r)), \tag{3}$$

or as a natural logarithm of the relationship between the repetition of vectors of length m and the repetition of vectors of length m+1.

Thus, if there are such vectors in the time series, *ApEn* will assess the logarithmic probability that the intervals following each vector will be different. Lower values of *ApEn* correspond to higher probability that vectors are followed by similar ones. If the time series is very irregular then the existence of similar vectors cannot be predicted and the value of *ApEn* is relatively large.

Sample entropy $SampEn(S_N, m, r)$ is also calculated using equation (3), adding the following conditions: the length of vectors is not used in calculating conditional probabilities *SampEn*; similarity of a vector to itself is not included.

Sample entropy has the following advantages: the calculated value will have smaller error margin when vectors of small dimensionality are used; it maintains relative density (while approximate entropy loses this property); it is less dependent on the length of the time series.

Tsallis entropy. Constantino Tsallis proposed non-extensive (non-additive) entropy as a new entropy generalization [11-12]. Tsallis used the standard Shannon entropy expression introducing a new exponential function instead of the logarithmic one $\ln(x) \Rightarrow \ln_q(x) \Rightarrow (x^{1-q}-1)/(1-q)$ with some numeric parameter q. He proposed the following equation of q-entropy:

$$S_q = -\sum_i \left(p_i^q \ln_q(p_i)\right) = \left(1 - \sum p_i^q\right)/(q-1), \ \sum_{i=1}^{N} p_i = 1, \ q \in R. \tag{4}$$

For anomalous systems with long memory and / or long-term correlation, the q coefficient allows to determine the following behavioral characteristics of complex systems: unusual anomalous phenomena dominate in the system when $q < 1$; recurring phenomena dominate when $q > 1$; and when $q \to 1$, Tsallis entropy reduces to Shannon entropy. Large values of this coefficient could be treated as long memory parameter, because they correspond to long-term correlation of system's states.

The fundamental difference of Tsallis entropy from all others is in its non-additivity. In Tsallis' opinion, it is precisely this difference that allows for a more accurate description of systems with "long memory" and systems, in which every element interacts not only with its nearest neighbors, but with the entire system or some of its parts.

It is worth noting that the q indicator can be treated not just as a measure of system's complexity, but as a measure of its non-extensiveness. Empirical analysis of Tsallis entropy dynamics shows that the lower its value is for a fixed q over time, the lower is the complexity of the system dynamics [4].

Wavelet entropy. The algorithm for calculating wavelet entropy is given in [13-14]. We define normalized total wavelet entropy (*NTWE,* [4]) as

$$E_{WT} = -\sum_{j=1}^{N} p_j \cdot \ln p_j / X_{\max} \tag{5}$$

where $X_{\max} = \ln N$ is a normalising constant.

NTWE can be used as a measure of regularity (or chaos) of time series, thus providing useful information about hidden dynamic processes associated with the time series. An regularized process can be represented as a periodic mono-frequency signal (time series), that is a signal with a narrow frequency range. Wavelet representation of this time series uses only a single scale, meaning that all relative wevelet energy values will be approaching zero on all other scales except for the one that contains the representative frequency of the series. The relative energy on this scale will be approaching 1. Therefore, *NTWE* will take on a very low value.

The logical next step in the development of wavelet entropy calculation algorithms is to divide the time series into non-overlapping intervals/windows. To calculate new characteristics, we select windows of length L and create i intervals, $i = 1,...,N_T$, where $N_T = M / L$. On each interval, corresponding values of the time series are associated with the central point of the time window.

The following equation is used to calculate wavelet energy on scale j for time window i

$$E_j^{(i)} = \sum_{k=(i-1)\cdot L+1}^{i\cdot L} \left|C_j(k)\right|^2, \ i = 1,...,N_T \tag{6}$$

General energy in this time window is equal to

$$E_{tot}^{(i)} = \sum_{j=-N}^{-1} E_j^{(i)} \tag{7}$$

The relative wavelet energy change over time and normalized general wavelet entropy can be obtained with the following equation:

$$E_{WT}^{(i)} = -\sum_{j=-N}^{-1} p_j^{(i)} \cdot \ln p_j^{(i)} / X_{\max}, \ p_j^{(i)} = E_j^{(i)} / E_{tot}^{(i)} \tag{8}$$

Results and discussion. To calculate entropy values we used the moving window procedure [4]. The concept of "moving window" reflects the essence of data processing: in the time series of interest we select a sequence of elements (window) of a given length (width of the window) and we calculate the entropy value for this sequence. Then, we move the window a given number of elements (window step), and, repeating the procedure for the new sequence, we obtain consecutive local entropy values. As a result, we obtain a time series of entropy values that helps determining characteristics of hidden patterns in the outcome series.

To study the dynamic properties of regional stock markets using entropy analysis we chose the following stock market indices PFTS (Ukraine), RTSI (Russia), FTSE (UK), SP500 (USA), SSEC (China), DAX (Germany) for the period from 01.01.2004 to 15.01.2016 [15].

Figure 1 shows the dynamics of the closing values of the stock market indices on a relative scale.

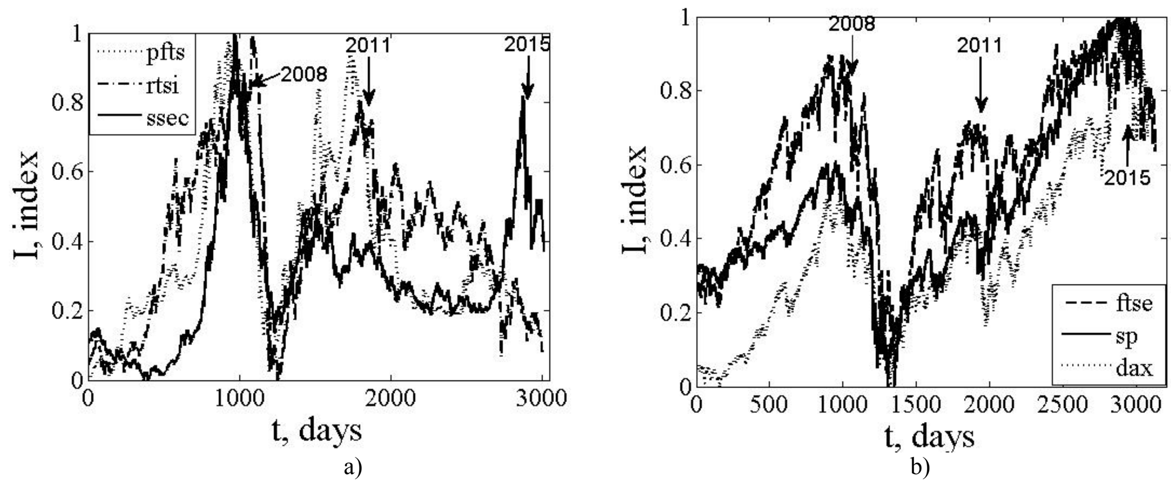


Figure 1. Dynamics of the stock market values for: a) Ukraine (pfts), Russia (rtsi), China (ssec); b) the UK (ftse), the USA (sp), Germany (dax) *Source: Calculated by the authors* [15].

Arrows in Figure 1 indicate crises of 2008, 2011 and 2015. We can see that the regional stock markets of the countries grouped in Figure 1 a) have similar dynamics amongst themselves, and those grouped in Figure 1 b) also have similar dynamics within that group. This observation leads to a conclusion that the developments of these markets have common trends.

Next we report the results of sample entropy calculation (SampEn) for stock markets. Figure 2 shows comparative dynamics of stock market indices, sample entropy and its average values.

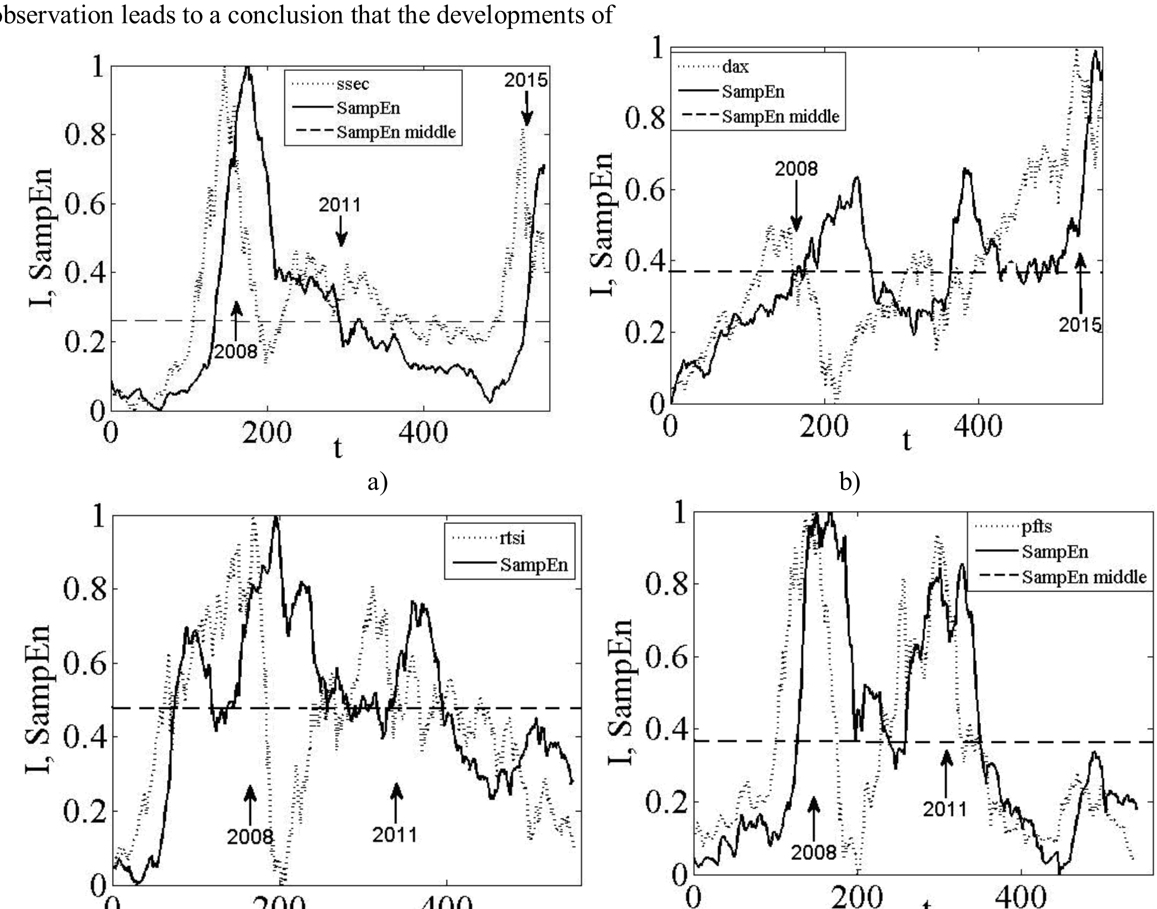
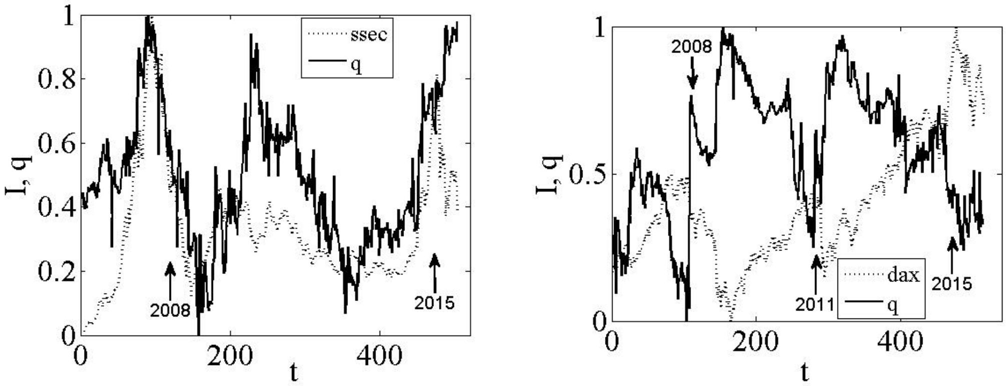


Figure 2. Comparative dynamics of sample entropy (SampEn), average entropy values (SampEn middle) and stock market indices (I) for: a) China (ssec), b) Germany (dax), c) Russia (rtsi), d) Ukraine (pfts) *Source: Calculated by the authors*

We can see in Figure 2 that for all examined time series, sample entropy significantly increases and exceeds the average before a crisis. Such behavior suggests that the markets can approach or be in a state of uncertainty. In the post-crisis period the sample entropy values drop below average. Sample entropy values at this level indicate that the markets are stable or steadily growing. This fact indicates that sample entropy as an prognostic indicator is sensitive to pre-crisis states. It is also worth noting the possibility of determining the extent of a crisis, which demonstrates sensitivity of sample entropy to changes in market behavior.

Figure 3 presents comparative dynamics of stock market indices with corresponding values of *q* coefficient, which is considered to be an indicator of time series' complexity.

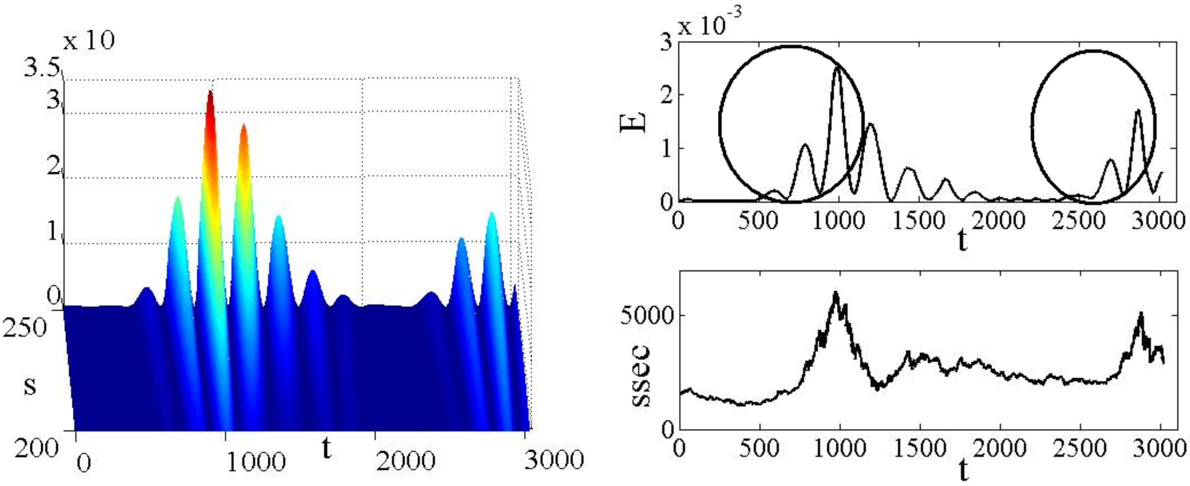


a) b)

Figure 3. Comparative dynamics of stock market index (I) with the corresponding values of *q* coefficient for: a) China (ssec); b) Germany (dax) *Source: Calculated by the authors*

The rapid growth rate of *q* at the time of crisis indicates the increased complexity of the (stock market) system at that time.

Figure 4 presents the energy of wavelet coefficients and wavelet entropy for the Chinese stock market.



a) b)

Figure 4. Energy of wavelet coefficients (a) and wavelet entropy (b) for the Chinese stock market (ssec) *Source: Calculated by the authors*

From Figure 4 we can see that both the energy surface (Fig. 4 a) and wavelet entropy (circled in Fig. 4 b) during the pre-crisis period form three characteristic increasing waves. It was demonstrated empirically that the point at which the third wave exceeds the second wave marks the development of a crisis event. Peak of the third wave is the starting point of a crisis. Other time series that we studied also showed similar dynamics.

Note that examination of sub-intervals in the time series will allow for a more accurate analysis and monitoring of the markets than just entropy analysis over a single long period of time, because several crises of different magnitude developed over the time period under study. In addition, it is advisable to apply the sliding window method, which allows keeping track of these precursors in real time.

Conclusion. In this paper, we investigated the effectiveness of using entropy analysis of regional stock markets non-regular dynamics. The empirical results have shown that the sample entropy, Tsallis and wavelet entropy can be used as indicators of pre-crisis states. All these three entropy measures have an anticipatory pattern of change, and therefore they can serve as a crisis predictor. However, each event has its own individual lead time, which cannot be generalized.

Thus, we can conclude that the use of entropy analysis, along with other modern methods of time series analysis allows monitoring of financial markets in order to study their stability and to identify pre-crisis states.

REFERENCES

1. Sornette D. (2003) Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton University Press.
2. Peters, E. E. (1994). Fractal Market Analysis: Applying Chaos Theory to Investment and Economics. John Wiley and Sons.
3. Mantegna R. N. , Stanley H. E. (2000). An Introduction to Econophysics. – Cambridge: Cambridge University Press.
4. Derbentsev V., Serdyuk O., Soloviev V., Sharapov A. Synergetics and Econophysics Research Methods of Dynamic and Structural Properties of Economic Systems. Cherkasy: Brama-Ukraine, 2010. – 300 p.
5. Chyumak O. Entropy and Fractals in Data Analysis. Moscow: Regular and Chaotic Dynamics, 2011. – 164 p.
6. Weidlich W. Socio-Dynamics: A Systematic Approach to Mathematical Modelling in the Social Science. Taylor and Francis, London, 2002.
7. Soloviev V., Danilchuk H., Serdyuk O.A. Research of the crises phenomena on time series basis with help of entropy characteristics // Information Technologies, Management and Society. Theses of the International Conference. Information Technologies and Management. 2008 April 10-11, Information System Institute, Riga, Latvia. – P.41-42.
8. Soloviev V., Serdyuk O. Using of wavelet entropy for analysis of complex economic systems // Econophysics Colloquium and beyond, Alcona, Italy, September 27-29, 2007.
9. Joshua S., Richman J., Moorman R. Physiological time-series analysis using approximate entropy and sample entropy // Am. J Physiol. Heart Circ. Physiol. 278: H2039-H2049, 2000.
10. Pincus S.M. (1991) Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. Vol.88. – P. 2297-2301.
11. Tsallis C. Nonextensive Statistics: Theoretical, Experimental and Computational Evidences and Connections / C. Tsallis // Brazilian Journal of Physics. – Vol. 29. – №. 1. – March, 1999.
12. Tsallis C. Introduction to Nonextnsive Statistical Mechanics, Approaching a Complex Word / C. Tsallis // Springer, New York, 2009. – 382 p.
13. Rosson O.A, Blanco S, Yordanova J, et al. Wavelet entropy: a new tool for analysis of short duration brain electrical signals[J]. J Neurosci Meth, vol. 105, P. 65-75, Jan. 2001.
14. Quiroga, R.Q; Rosso, O.A; Basar E.; Schürmann M. Wavelet entropy in event-related potentials: A new method shows ordering of EEG oscillations. Biol. Cybern. 2001, 84, 291–299.
15. http://finance.yahoo.com.

Энтропийный анализ динамических свойств региональных фондовых рынков

А. Б. Данильчук, В. Д. Дербенцев, В. Н. Соловьев, А. Д. Шарапов

Аннотация. Статья посвящена энтропийному анализу региональных фондовых рынков. В работе предложено и эмпирически доказано эффективность использования в качестве меры неопределенности и нестабильности сложных систем, в частности региональных фондовых рынков, таких энтропийных показателей, как энтропия шаблонов, Тсаллиса и вейвлет-энтропия. Полученные результаты свидетельствуют о том, что эти показатели могут быть использованы в качестве индикаторов-предвестников кризисных явлений.

Ключевые слова: *сложные системы, региональные фондовые рынки, кризисные явления, нестабильная динамика, энтропийный анализ, энтропия шаблонов, Энтропия Тсаллиса, вейвлет-энтропия.*