Fractals, Vol. 15, No. 4 (2007) 399–404

World Scientific
www.worldscientific.com

MULTISCALE ENTROPY ANALYSIS OF EEG FROM PATIENTS UNDER DIFFERENT PATHOLOGICAL CONDITIONS

JEONG-HYEON PARK, SOOYONG KIM and CHEOL-HYUN KIM
Department of Physics
Korea Advanced Institute of Science and Technology
Daejeon 305-701, Korea

ANDRZEJ CICHOCKI
Laboratory for Advanced Brain Signal Processing
RIKEN Brain Science Institute
2-1 Hirosawa, Wako-Shi, Saitama 351-0198, Japan

KYUNGSIK KIM*
Department of Physics, Pukyong National University
599-1 Daeyeon-Dong, Busan 608-737, Korea
kskim@pknu.ac.kr

Received December 14, 2006
Accepted February 21, 2007

Abstract

Multiscale sample entropy (MSE) of human electroencephalogram (EEG) data from patients under different pathological conditions of Alzheimer's disease (AD) was evaluated to measure the complexity of the signal. Quantifying the complexity level with respect to various temporal scales, MSE analysis provides a dynamical description of AD development. When compared to EEG data from normal subjects, EEG data from subjects with mild cognitive impairment (MCI) showed nearly the same complexity profile, but a scale discrepancy which may occur from a spectral abnormality. EEG data from severe AD patients showed a loss of complexity over the wide range of time scales, indicating a destruction of nonlinear structures in brain dynamics. We

*Corresponding author.

compare the MSE method and spectral analysis to propose that nonlinear dynamical approach combining a multiscale method is crucial for revealing AD mechanisms.

Keywords: Complexity; MSE; Alzheimer's disease; EEG.

1. INTRODUCTION

AD is one of the most common neurodegenerative disorders among the elderly population.[1–3] The conventional spectral analysis of EEG, a time series that contains ongoing brain activities, has mainly concerned showing spectral features in several frequency bands.[4–6] Although the spectral analysis has been successful in AD studies, nonlinear dynamical analysis is crucial if trying to capture higher order dynamical properties of brain.[7–9]

One of the statistical concepts of the nonlinear dynamics on complex systems is information entropy.[10] As a complexity measure, the entropy can be generalized to characterize the amount of information stored in the system. It is natural to assume that a highly complex system has a flexible structure for information processing in accordance with various temporal scales. Thus, it is necessary to investigate a system's complexity with respect to different temporal scales. Combining a coarse-graining method and the sample entropy (S_E), multiscale sample entropy (MSE) analysis[11,12] has provided a useful complexity measure for finite, noisy time series such as human gait signal,[13] heart rate variability,[14] and financial data.[15]

The aim of the present study is to describe how statistical properties of a signal change as a pathological condition proceeds. In order to represent the intermediate stage of AD, EEG data was collected from MCI subjects, who showed no clinical signs of AD at the time it was recorded but later developed AD.[16] Preclinical discrimination between normal subjects and MCI subjects is an important challenge for early treatment of the disease; once a subject with no clinical sign of AD is diagnosed as MCI, the onset of the manifestations of AD could be successfully prevented or at least slowed down.[17,18] Data was also collected from severe AD patients for comparison. Comparing these three kinds of subjects, we can track out dynamical transitions from a healthy state to a pathological state of brain. A linear discriminant analysis (LDA) was performed upon the sample entropy in order to clarify the clinical usefulness of the method.

2. METHODS AND DATA

2.1. Multiscale Sample Entropy Analysis and Linear Discriminant Analysis

We briefly describe the MSE algorithm for $m = 2$. For a time series of N points, $\{x_1, \ldots, x_i, \ldots, x_N\}$, consider the two-length templates: $u_2(i) = (x_i, x_{i+1})$, $1 \le i \le N-1$. The distance between two templates is defined by $d[u_2(i), u_2(j)] = \max\{|x(i) - x(j)|, |x(i+1) - x(j+1)|\}$ and they are said to be matched when the distance is less than a given tolerance r. Estimating total matched templates among all $(N-1)(N-2)/2$ pairs, the ratio $P^{(2)} = \frac{\#\text{ of matched templates}}{(N-1)(N-2)/2}$ defines the probability that two templates of length 2 are closed within the tolerance r. An identical procedure is carried out for all three-length templates, $u_3(i) = (x_i, x_{i+1}, x_{i+2})$, $1 \le i \le N-2$, to get $P^{(3)}$. Finally, S_E is calculated as $S_E = -\log(P^{(3)}/P^{(2)})$. S_E is able to detect the irregularity of patterns in a time series by estimating the conditional probability that sequences that are closed remain closed for an additional point.

In order to evaluate multiscale S_E, we construct a coarse-grained time series corresponding to the scale factor, τ. First, the original time series of length N is divided into consecutive N/τ segments where each segment has length τ, second, each segment is taken an average, resulting the coarse-grained time series of the scale factor, τ. S_E is then evaluated for the coarse-grained time series as a function of scale factors.

LDA is a classification algorithm used to classify individual subjects into several classes based on a training data. It is implemented in order to reduce a high-dimensional feature set to a lower-dimensional feature set by a proper linear projection so that classes are most separable in the projected space. LDA was utilized here in order to demonstrate how

the MSE was able to classify all 86 subjects into their original groups.

2.2. Subjects and EEG Recording

Patients who only complained of memory impairment but had no apparent loss in general cognitive, behavioral, or functional status were recruited. Among them, 22 patients met the following criteria for MCI: a mini-mental state examination (MMSE) score of 24 or higher, a clinical dementia rating (CDR) scale score of 0.5 with memory performance of less than one standard deviation below the normal reference (Wechsler Logical Memory Scale and Paired Associates Learning Subtests, IV and VII, ≤ 9,[19] and/or ≤ 95 on the 30-minute delayed recall of the Rey-Osterreith figure test.[20] They developed, within one and a half years, probable or possible AD according to the NINDS-ADRDA criteria.[21] thirty-eight healthy individuals, as a control group, were recruited from among the family members of the patients participating in the study. An additional 26 severe AD patients were recruited from the same clinic.

The EEG data was recorded within one month after entering the study for all subjects. Ag/AgCl electrodes (disks of a diameter of 8 mm) were placed on 21 sites according to the 10–20 international system, with the reference electrode on the right earlobe. EEG data was recorded with a Biotop 6R12 (NEC San-ei, Tokyo, Japan) using an analog filtering bandpass at 0.5–250 Hz and a sampling rate of 200 Hz. Each EEG record was judged by inspection to be free from electrooculographic and movement artifacts and judged to contain minimal electromyographic activity.

3. RESULTS

3.1. MSE Analysis of EEG

The MSE analysis was subject to a clean and long (2×10^4 data points) EEG signal from a normal subject for the first inspection of EEG complexity. Figure 1 shows representative MSE curves with different values of tolerance r. Symbols of the curves indicate S_E values for their corresponding scale factors. The MSE curve for the EEG data has common features: it has a local maximum entropy value at a scale factor of 5 to 7; after reaching the maximum entropy, it gradually decreases and stabilizes to a base-line entropy value at a scale factor near 20. We observed that the tolerance, r, can be treated as

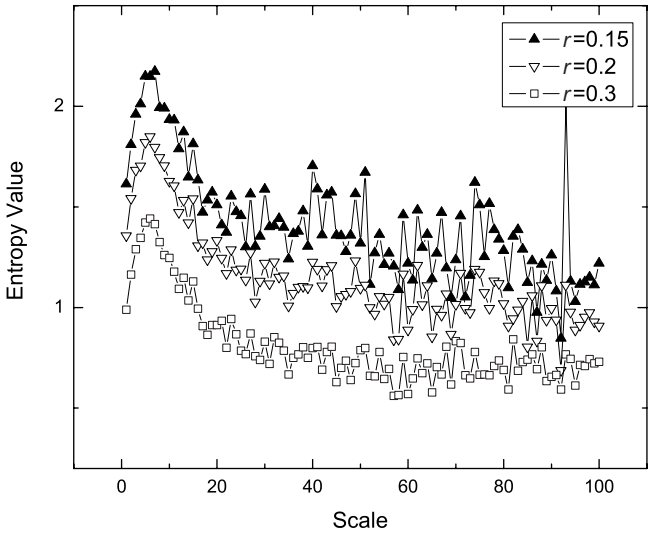

Fig. 1 Representative MSE curves derived from EEG data from a normal subject (100 seconds) with different tolerance level r. Sample entropy $S_E(m = 2, N = 2 \times 10^4)$ was evaluated for 100 scale factors.

a "filter" that discards uncorrelated noise. A large tolerance level (□), a coarse filter, makes templates pair easily matched, then induces a relatively low entropy value. On the other hand, a small tolerance level (▲), a fine filter, induces a relatively large entropy value, but the values are unstable at large scale. This is because the coarse-graining procedure reduces the sample size of templates by N/τ.

Next, the MSE method was used to compare the complexity of the EEG data from subjects with three different brain states: severe AD patients, MCI subjects who later progressed to AD, and normal subjects. The implementation was applied to 20-second (4×10^3 points) artifact-free segments of the signals. Figure 2 represents the MSE curves averaged over 21 channels and individuals for each group. Only $\tau \leq 16$ scale factors were included due to the lack of long artifact-free signals. The parameters of the MSE algorithm were chosen to be $m = 2$, $r = 0.15$. The MSE curves from normal subjects and MCI subjects have similar patterns: a local maximum at scale factor 6 and 7, respectively, followed by decreasing entropy values. The main difference between MSE curve from MCI subjects and that from normal subjects appears to be the shift of scale factors for their corresponding entropy values. The MSE curve from severe AD patients has an evidently lower level of entropy values than those of other two groups, particularly in the region where a maximum entropy value is assigned.

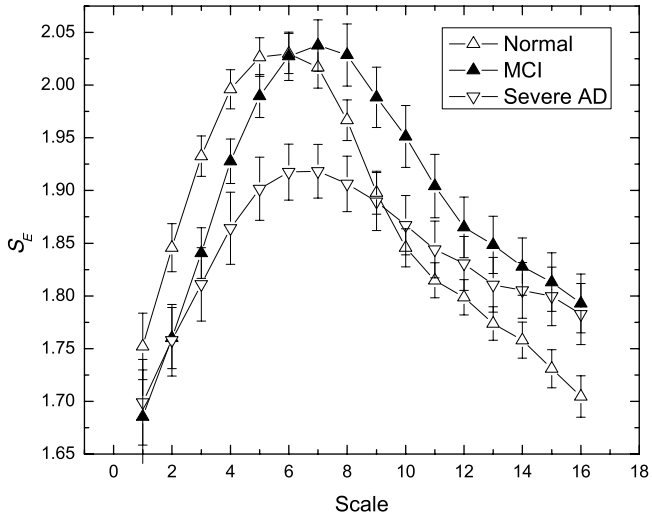

Fig. 2 MSE curves of EEG data derived from subjects with three different brain states: AD patients, MCI subjects who later progressed to AD, and normal subjects. $S_E(m = 2, r = 0.15, N = 4 \times 10^3)$ was evaluated at 16 scale factors. The symbols represent the mean values of S_E over each individual and channel, and the error bars represent the standard error, meaning the standard deviation divided by square root of the number of subject, $\sigma/\sqrt{n}$.

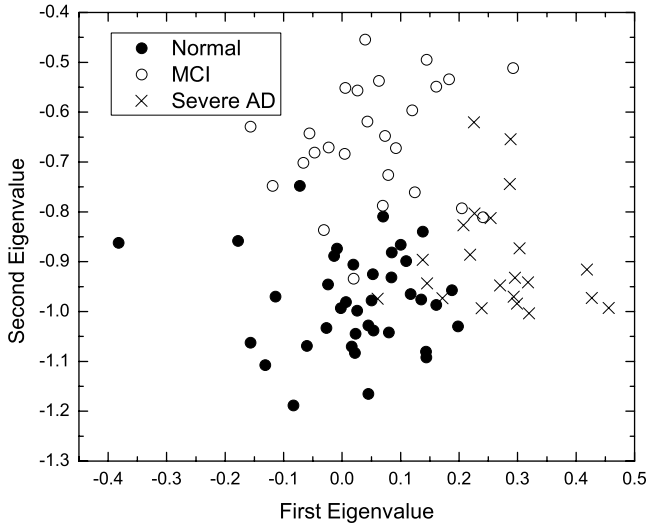

Fig. 3 Linear discriminant analysis of the three groups: the MCI subjects (○), the severe AD patients (×), and the normal control group (•). Eighty-six individuals are properly projected into a two-dimensional space due to their feature vectors which are constructed by 30 S_E values from EEG channel P7 and Fp1.

The local maximum in MSE curve represents two aspects of underlying dynamics of EEG signal: (1) a scale factor where a maximum entropy value is assigned is closely related to spectral properties of the signal (we will deal with the relationship between EEG frequencies and MSE scale factors in the next section) and (2) the level of a maximum entropy value indicates the complexity level of the signal, i.e. in our case, a loss of complexity can be proposed as a generic feature of pathological dynamics, which is a wide consensus of nonlinear time series analysis.[22–24]

In a statistical point of view, MCI subjects, AD patients, and normal subjects are not distinguishable at scale factor 1, where traditional single-scale entropy approaches would fail to distinguish them. MCI and normal subjects start to be statistically separable at scale factor 2, and MCI subjects and severe AD patients at scale factor 3 (t-test, $p < 0.05$). The largest separation between severe AD patients and other two groups is obtained at scale factor 6 at which a significant pathological mechanism of AD could take place. However, the most separable scale factor in one case might not be valid for another case. When one tries to differentiate severe AD patients from other subjects, scale factor 6 provides valuable information, but scale factor 6 would yield poor results in case when differentiating MCI subjects and normal subjects. In this case, scale factor 6 can be considered to contain any additional information in early development of AD.

By using a classification algorithm, we examined how the MSE method could be used as a clinical tool for differentiating these groups. In Fig. 3 we present LDA results for three groups. Thirty S_E values, based on the significant scale factors in MSE results, were carefully selected from two EEG channels, P7 and Fp1, as a training set. The classification result reached at most 92% in success rate. Since only two channels of EEG data were utilized for the analysis, it is feasible to expect a higher classification rate if using an optimized feature selection.

3.2. MSE and EEG Spectral Analysis

We compare the MSE method and EEG spectral analysis in terms of our results. Figure 4 represents the power spectral density of EEG at the left occipital region (O1) from MCI subjects along with normal subjects. A decrease of mean frequency with an increase in delta (0.5–3.5 Hz) and theta (3.5–7.0 Hz) power was observed in almost every channel for MCI subjects (dotted curve), compared to that of normal subjects.

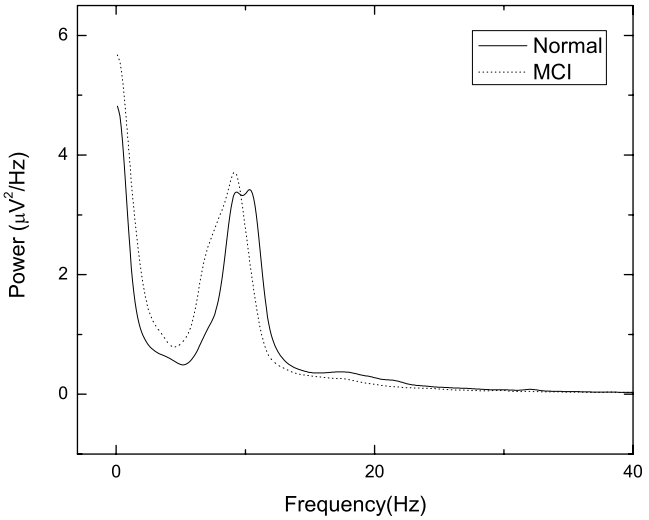

Fig. 4 Power spectral density (PSD) of the EEG channel O1 from MCI subjects and normal subjects (Welch method with Hanning windows).

To investigate how these spectral features are expressed in MSE curve, we evaluated an MSE curve for a synthetic EEG-like signal, a $1/f$ signal superimposed on a pure sinusoidal rhythm with various frequencies (Figs. 5a–5c). Figure 5d shows MSE curves for the composite signals with frequencies of $5, 10$ and $20\,\mathrm{Hz}$, as well as the original $1/f$ signal with its flat MSE curve. The sinusoidal component destroys the correlation of the $1/f$ signal at small time scale, inducing a local maximum entropy value to some extent before the component is being filtered out through the coarse-graining procedure. After the sinusoidal component is being filtered out completely, the curve stabilizes to a baseline entropy value. Numerically, we revealed that the scale factor spanning at least one period of a sinusoidal component is sufficient for S_E values to become stable. Thus, the scale factor where S_E values stabilize shifts to a large scale factor as the frequency of the pure sinusoidal component decreases, which is the reminiscent of the shifted MSE curve from MCI subjects compared to that from normal subjects (Fig. 2). In Fig. 1, the scale factor where the MSE curve stabilizes to a baseline entropy value ($\tau \approx 20$) corresponds to the alpha rhythm (7–13 Hz), the most characteristic frequency band for EEG. Thus, the shift of scale factors in MSE curves from MCI subjects or severe AD patients, as shown in Fig. 2, is closely associated with the slowing phenomena of EEG in pathological condition.

4. CONCLUSION

The valuation of a nonlinear measure is subjected to how it fits the parameter that we are to determine. For most cases, the parameter may not be inherent in a trivial scale, and a significant scale is hard to be assigned in advance. Even though the significant scale is possibly found, intrinsic or extrinsic noise could lead to incorrect results when we count on a single-scale analysis. In our study, multiscale analysis is crucial for both early and severe development of AD. Considering MCI as an intermediate stage toward AD, dynamical changes owing to AD development can be effectively described by an MSE curve. When the brain is changed into mild dementia from the normal state, this transition is

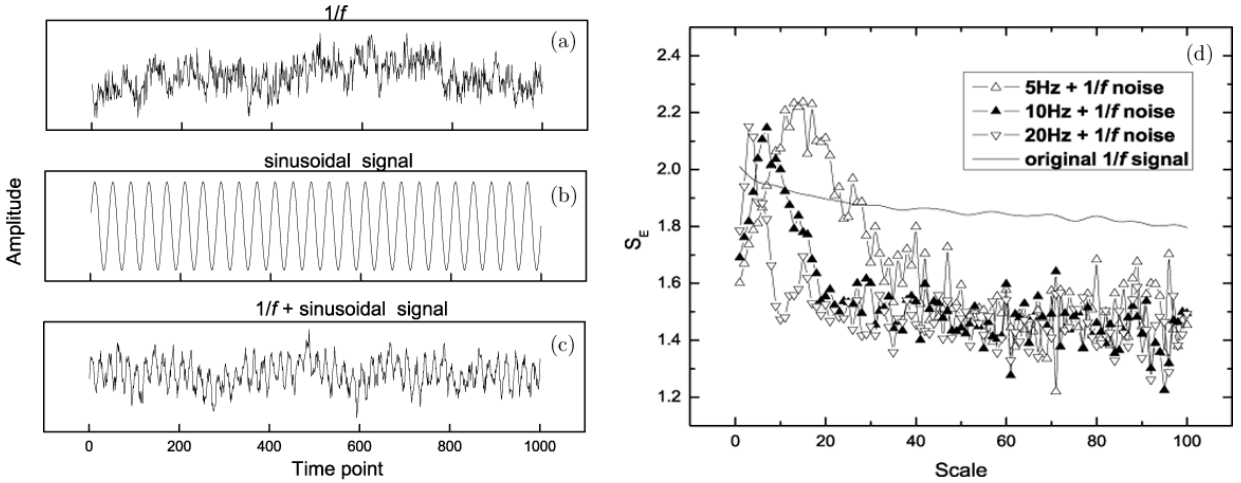

Fig. 5 Procedure to construct a synthetic EEG-like signal. **(a)** A $1/f$ signal is superimposed on **(b)** pure sinusoidal rhythm to make **(c)** a composite signal. **(d)** MSE curves for the composite signals with three different frequencies: $5, 10$ and $20\,\mathrm{Hz}$. Dotted curve indicates flat MSE curve for the original $1/f$ signal.

initially represented in the MSE curve as the shift of scale factors without any entropy losses. In this case, EEG complexity of MCI subjects can still be regarded as preserving nonlinear properties as that of normal subjects. The scale discrepancy in MCI subjects, possibly coming from a spectral abnormality, is an initial symptom of AD, which allows effective early treatment,[17,18] as far as average entropy value maintains the level of normal case. As the disease become severe, the complexity of the signal is remarkably reduced, and this nonlinear transition is represented in the MSE curve as an entropy loss over the wide range of scale factors, which is regarded as an incurable brain state. Clinically, maintaining an entropy level over significant scales is more relevant for the preclusion of a development toward severe AD than early detection of AD.

ACKNOWLEDGMENTS

This research was partially supported by the Korea Science and Engineering Foundation, and the KOSEF-JSPS cooperation program. This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST) (R01-2006-000-10470-0).

REFERENCES

1. J. Jeong, EEG dynamics in patients with Alzheimer's disease, *Clin. Neurophysiol.* **115** (2004) 1490–1505.
2. T. Musha, T. Asada and F. Yamashita, A new EEG method for estimating cortical neuronal impairment that is sensitive to early stage Alzheimer's disease, *Clin. Neurophysiol.* **113** (2002) 1052–1058.
3. A. Cichocki, S. L. Shishkin and T. Musha, EEG filtering based on blind source separation (BSS) for early detection of Alzheimer's disease, *Clin. Neurophysiol.* **116** (2005) 729–737.
4. K. Bennys, G. Rondouin, C. Vergnes and J. Touchon, Diagnostic value of quantitative EEG in Alzheimer's disease, *Neurophysiol. Clin./Clin. Neurophysiol.* **31** (2001) 153–160.
5. M. Penttila, J. V. Partanen, H. Soininen and P. J. Riekkinen, Quantitative analysis of occipital EEG in different stages of Alzheimer's disease, *Electroencephalogr. Clin. Neurophysiol.* **60** (1985) 1–6.
6. J. J. Claus, V. I. H. Kwa and S. Teunisse, Slowing on quantitative spectral EEG is a marker for rate of subsequent cognitive and functional decline in early Alzheimer disease, *Alzheimer Dis. Assoc. Disord.* **12** (1998) 167–174.
7. D. Ruelle, *Chaotic Evolution and Strange Attractors: The Statistical Analysis of Time Series for Deterministic Nonlinear Systems* (Cambridge University Press, Cambridge, 1989).
8. H. Kantz, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).
9. W. Klonowski, Chaotic dynamics applied to signal complexity in phase space and in time domain, *Chaos Solitons Fractals* **14** (2002) 1379–1387.
10. J.-P. Eckmann and D. Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.* **57** (1985) 617–656.
11. M. Costa, A. L. Goldberger and C. K. Peng, Multiscale entropy analysis of biological signals, *Phys. Rev. E* **71** (2005) 021906–18.
12. M. Costa, A. L. Goldberger and C. K. Peng, Multiscale entropy analysis of complex physiologic time series, *Phys. Rev. Lett.* **89** (2002) 068102–4.
13. M. Costa, C. K. Peng, A. L. Goldberger and J. M. Hausdorff, Multiscale entropy analysis of human gait dynamics, *Physica A* **330** (2003) 53–60.
14. M. Costa and J. Healey, Multiscale entropy analysis of complex heart rate dynamics, *Comput. Cardiol.* **30** (2003) 705–708.
15. S. Pincus and R. E. Kalman, Irregularity, volatility, risk, and financial market time series, *Proc. Natl. Acad. Sci. USA* **101** (2004) 13709–13714.
16. R. C. Petersen, G. E. Smith and S. C. Waring, Mild cognitive impairment: clinical characterization and outcome, *Arch. Neurol.* **56** (1999) 303–308.
17. K. Blennow and H. Hampel, CSF markers for incipient Alzheimer's disease, *Lancet Neurol.* **2** (2003) 605–613.
18. S. T. DeKosky and K. Mare, Looking backward to move forward: early detection of neurodegenerative disorders, *Science* **302** (2003) 830–834.
19. D. Wechsler, *Wechsler Memory Scale: Revised Manual* (San Antonio, Psychological Corp., 1987).
20. J. Hodges, *Cognitive Assessment for Clinicians* (Oxford Medical Publications, Oxford, 1993).
21. G. McKhann, D. Drachman and M. Folstein, Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease, *Neurology* **34** (1984) 939–944.
22. A. L. Goldberger, C. K. Peng and L. A. Lipsitz, What is physiologic complexity and how does it change with aging and disease?, *Neurobiol. Aging* **23** (2002) 23–26.
23. M. Palus, Coarse-grained entropy rates for characterization of complex time series, *Physica D* **93** (1996) 64–77.
24. A. Porta, S. Guzzetti and N. Montano, Entropy, entropy rate, and pattern classification as tools to typify complexity in short heart period variability series, *IEEE Trans. Biomed. Eng.* **48** (2001) 1282–1291.