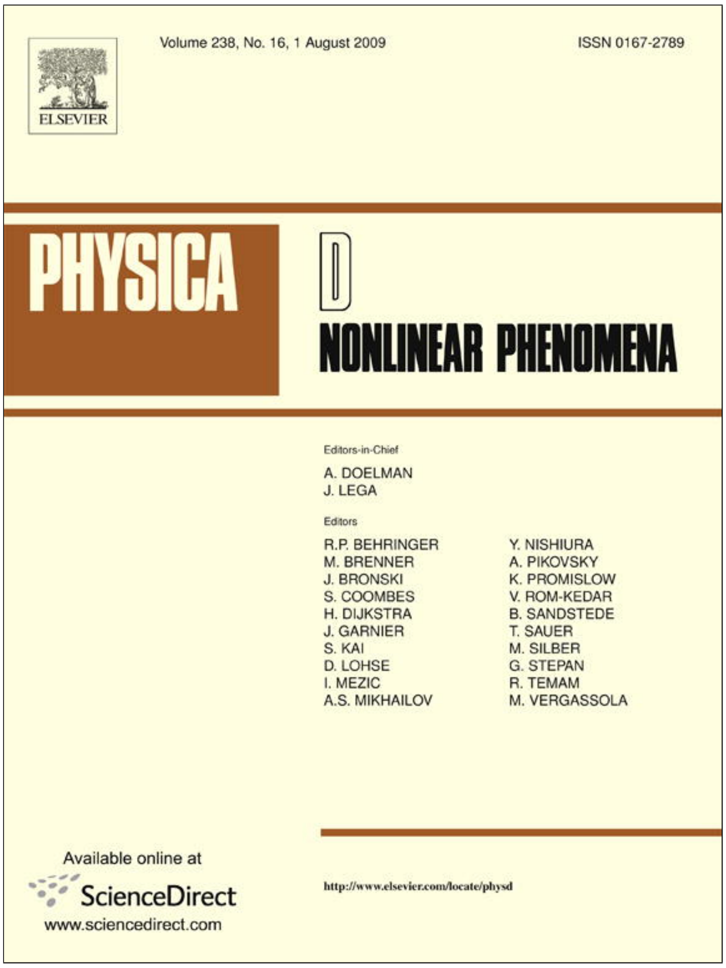
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Physica D 238 (2009) 1711–1715

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Comment on "Stochastic analysis of recurrence plots with applications to the detection of deterministic signals" by Rohde et al. [Physica D 237 (2008) 619–629]

Norbert Marwan [a,b,*], Jürgen Kurths [c,b]

[a] Interdisciplinary Centre for Dynamics of Complex Systems, University of Potsdam, Potsdam 14415, Germany
[b] Potsdam Institute for Climate Impact Research (PIK), P.O. Box 60 12 03, Potsdam 14412, Germany
[c] Department of Physics, Humboldt University Berlin, Newtonstr. 15, Berlin 12489, Germany

ARTICLE INFO

Article history:
Received 4 June 2008
Received in revised form
1 December 2008
Accepted 30 April 2009
Available online 18 May 2009
Communicated by R.P. Behringer

Keywords:
Recurrence plot
Stochastic processes
Detection
Data analysis

ABSTRACT

In the recent article "Stochastic analysis of recurrence plots with applications to the detection of deterministic signals" (Physica D 237 (2008) 619–629), Rohde et al. stated that the performance of RQA in order to detect deterministic signals would be below traditional and well-known detectors. However, we have concerns about such a general statement. Based on our own studies we cannot confirm their conclusions. Our findings suggest that the measures of complexity provided by RQA are useful detectors outperforming well-known traditional detectors, in particular for the detection of signals of complex systems, with phase differences or signals modified due to the measurement process.

Nevertheless, we also clearly assert that an uncritical application of RQA may lead to wrong conclusions.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Recently, Rohde et al. published an interesting article on signal detection and stochastic analysis [1]. Their article consists of mainly two parts. In the first part, the authors present a relationship between the distance matrix [2] and the variance of a time series (or signal). This relationship is important and has already motivated recent work on the application of the Wiener–Khinchin theorem on recurrence plots [3]. In the second part, Rohde et al. discuss and compare different methods for the detection of deterministic signals in noise, as recurrence based measures and traditional signal detectors. Since one of their final conclusions is that the performance of the recurrence quantification analysis (RQA) is weak, we tried to reproduce their results, but found several problems in the presented article (e.g. 2π is missing in Eq. (12), there is no normalisation of the data, or explanation on the calculation of the ROC has been left out). Moreover, the presented results are not derived from general cases. Therefore, the conclusion by Rohde et al. that the RQA would not perform well, cannot be generalised. However, the statements in the article connotes that this conclusion would be general. Although we find the presented comparison important and agree with it in general, we do not agree with such a general conclusion. Therefore, we comment on some points we think that are important to be discussed further. Our intention is not to rebut the findings of Rohde et al., but to justify and relativise their conclusions.

Recurrence plots (*RP*) and techniques related to RPs have become popular in the last decade. In order to apply these methods reliably, it is important to understand how the proposed measures of *recurrence quantification analysis* (*RQA*) are calculated and what they stand for. RQA provides measures of complexity quantifying structures in a RP [2]. RQA and RPs were introduced to study complex systems. For linear systems, traditional linear measures may be better in some cases.

A RP is a useful tool to visualise recurrences of phase space trajectories. In the most commonly used method, the RP is derived directly from the *distance matrix* $\mathbf{D} = D_{i,j}, i, j = 1, \dots, N$ (N is the length of the data series or trajectory):

$$D_{i,j} = \|\vec{x}_i - \vec{x}_j\| \tag{1}$$

by applying a threshold ε

$$R_{i,j} = \Theta(\varepsilon - D_{i,j}), \tag{2}$$

where Θ is the Heaviside function. Sometimes we find the appellation *recurrence plot* also referring to the distance matrix, Eq. (1), (as

* Corresponding address: Department of Transdisciplinary Concepts and Methods, Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, Potsdam 14412, Germany.

E-mail address: marwan@pik-potsdam.de (N. Marwan).

0167-2789/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2009.04.018

Author's personal copy

in [1]), which may cause confusion. The distance matrix should not be denoted as *recurrence plot* (or *unthresholded recurrence plot*), because it simply does not show recurrences, but only distances between all pairs of combinations of the phase space vectors $\vec{x}_i$ (the recurrence quantitative analysis (RQA) as well as the calculation of the dynamical invariants, such as C_2 or K_2, is based on the binary recurrence matrix). However, $\mathbf{D}$ is useful for the study of correlations within the data. We should note that the idea of a distance matrix is not new and can be found using different notations (e.g. *similarity plot and dot plot*) in several disciplines [4–7].

In the following we use the Euclidean norm in Eq. (1).

Although developed as quantifiers of structures in a RP (for interpretation of the line structures, cf. [8]), the RQA measures have a certain meaning in the sense of recurrences. For example, the density of points in a RP of size $N \times N$,

$$RR = \frac{1}{N^2} \sum_{i,j} \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|), \tag{3}$$

called *correlation sum*, *recurrence rate* or *percentage recurrences* [2,9], can be interpreted as the probability that any state of the system will recur in the future. In addition, the fraction of recurrence points forming diagonal line structures

$$DET = \frac{\sum_{l \geq l_{\min}} l\,P(l)}{\sum_{l} l\,P(l)}, \tag{4}$$

called *determinism* [2], can be interpreted as the probability that two closely evolving segments of the phase space trajectory will remain close for the next time step ($P(l)$ is the histogram of lengths of the diagonal line structures and $l_{\min}$ is the minimal length of a diagonal line necessary to be considered as a line; in the present work we use $l_{\min} = 2$). The notation *determinism* was not chosen to explain determinism in a mathematical sense, but to emphasise the observation that RPs of stochastic processes usually reveal fewer diagonal lines, whereas RPs of deterministic processes consist of many longer diagonal line structures. The occurrence of longer diagonal line structures is more related to the auto-correlation within the data (an auto-correlated stochastic process, as an iterated auto-regressive model (AR) does, can also have longer diagonal lines). Therefore, from a high value of this measure we cannot conclude that the process is deterministic.

Diagonal structures in a RP appear when two segments of the phase space trajectory run parallel within an ε-tube for some time, where ε is the minimal distance which is used to define recurrence. The length of a diagonal line corresponds to the time of such a parallel run. It is obvious that the states of stochastic processes (with zero or at least less auto-correlation) will not run closely for a longer time. Hence we will not find many or longer diagonal lines in a RP of such systems. Vertical recurrence structures appear if the state of the system changes very slowly, as is typical for intermittency. The vertical distance between points in a RP are related to recurrence times.

Further RQA measures quantify such diagonal line structures, and other measures quantify vertical structures in the RP (for definitions and meanings cf. [2,10]). In this present study we will focus only on the two measures *RR* and *DET*.

The measures of complexity provided by RQA can be used to distinguish different types of dynamical behaviour of systems. For example, they can be used to detect different types of transitions, such as those between period–chaos, chaos–chaos or strange non-chaotic behaviour [10–12]. The RQA can also distinguish between certain stochastic and deterministic processes [13,14].

Analysing several examples in noisy environments (linear additive noise) with a rather small signal-to-noise ratio (SNR), Rohde et al. claimed that "detectors based on certain statistics derived from recurrence plots are sub-optimal when compared to well-known detectors based on the likelihood ratio" and "their performance in classical signal detection problems does not compare well with traditional approaches" [1]. In our opinion this is a very strong statement. From our investigations we are able to show that the conclusions of Rohde et al. may hold for certain linear systems and under certain assumptions. In the following we will show that considering more general, complex and especially nonlinear processes, RP based measures are powerful tools for the detection of deterministic signals and outperform classical approaches.

2. Detection of deterministic signals in noise

Rohde et al. have compared the abilities of several RQA measures and traditional likelihood ratios in order to detect deterministic signals in an incoming signal (measurement). The problem can be formulated by considering the following two hypothesis:

H_0: signal absent, $x(t) = \xi(t)$
H_1: signal present, $x(t) = s(t) + \xi(t)$

where ξ is noise and s is a deterministic signal [1]. It is assumed that in the observed signal the original signal s is not changed (it comes in with linear additive noise). Based on a comparison of *receiver–operator characteristics* (*ROC*), they found that well-known detectors based on the likelihood ratio would outperform the measures of complexity derived by RQA. The idea for discriminating H_1 from H_0 is that by applying a threshold η on a certain discriminant measure λ; the test $\lambda \geq \eta$ would give us a decision in favour of hypothesis H_1.

The proposed discriminant in [1] was defined as the average over the entire squared distance matrix

$$\lambda = \frac{1}{N^2} \sum_{i,j} D^2_{i,j}, \tag{5}$$

which is related to the power detector.

2.1. Detection of unknown signals

For this study Rohde et al. [1] used a harmonic random process

$$s(t) = \cos(20t + \phi) \tag{6}$$

with a random phase ϕ (white uniformly distributed noise in the interval $[-\pi, \pi]$) and a sampling interval of $\Delta t = 1$.

The authors used an embedding procedure to reconstruct the phase space trajectory using an embedding dimension of $m = 4$ and a delay of $\tau = 12$, while for the RP they took a threshold of $\varepsilon = 1.5$. In order to calculate the ROC, the authors used 1000 realisations of a signal corresponding to H_0 and of a signal corresponding to H_1 (private communication with G.K. Rohde; this detail was not provided in [1]). Assuming a normal distribution for the detection measures, the resulting distributions of the measures were fitted by such a normal distribution, which is then used to derive the ROC. However, we found that, in general, normal distributions for the measures cannot be assumed, in particular for the average of the squared distance plot λ (Fig. 1). Thus, the ROC based on such a wrong assumption is biased, in particular for low embedding dimension and delay. Therefore, we used the empirically found distributions of the measures to derive the ROC.

In the calculation of the RP and the squared distance matrix, in [1] the data were not normalised. However, the normalisation is a crucial point. Adding noise to a signal, the variance of the noise corrupted signal can be different from the noise alone. Therefore, we can simply distinguish the signal from the noise using their variances. We would not need any other detector, neither the

Author's personal copy

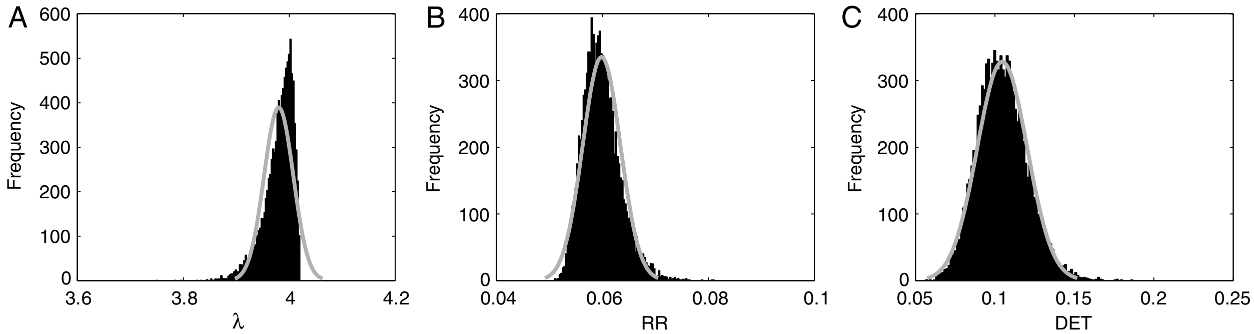


Fig. 1. Distributions of the measures (A) λ, (B) *RR* and (C) *DET* for 10,000 realisations of white Gaussian noise ($m = 2$, $\tau = 2$, $\varepsilon = 0.5$). Fits of corresponding Gaussian distributions are shown (in gray).

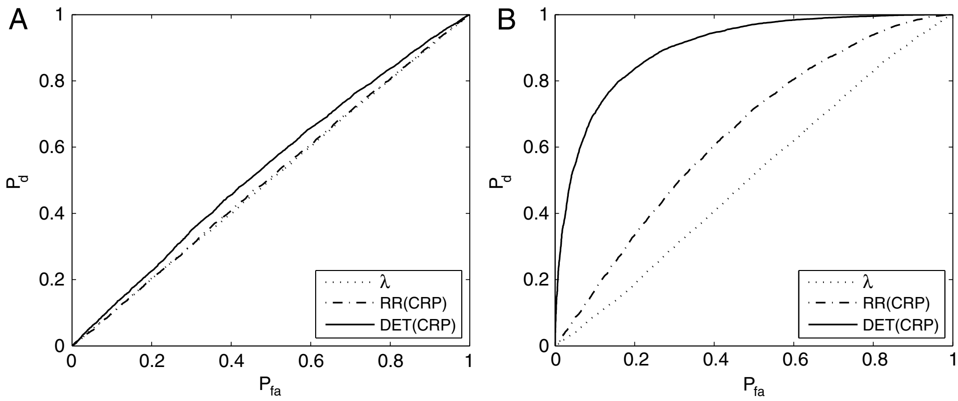


Fig. 2. ROC for the measures λ, *RR* and *DET* for the detection of a deterministic signal from white Gaussian noise. The signal used here is a harmonic signal $s(t) = \cos(20t+\phi)$, using a sampling interval of $\Delta t = 0.01$, data length $N = 1,000$ and 10,000 realisations (the same result can be achieved using $\Delta t = 1$). The SNR is (A) 0.1 and (B) 0.5; the embedding parameters are $m = 2$, $\tau = 15$ and the RP threshold $\varepsilon = 1$.

power detector, nor the RQA measures (and the variance may even outperform them). However, in practice there would not be such a link between the variances and the signals, and we would not know which signal is just noise and which one contains a signal. Therefore, for unknown signals it would be most likely that they will be normalised to some equal statistics. Here we will normalise the signals to the same standard deviation before analysis and apply the same recurrence threshold ε.

By applying normalisation, we are not able to detect a deterministic signal for a SNR of 0.1 with any of the discussed measures (Fig. 2A). SNR $= 0.1$ is indeed a rather high amount of noise and a big challenge for every method of signal detection (at least without prior knowledge about the signal).

Considering less noise, e.g. SNR $= 0.5$, we find that at least the RQA measure *DET* is able to detect the deterministic signal and even outperforms the detector λ (Fig. 2B; details of this analysis are added to the figure's caption). We have found similar results in favour of the RQA measure for the chirp signal and the Duffing system as used in [1]. The reason is that the RPs of such systems consist mainly of diagonal lines, which will remain as a significant amount if the signal is corrupted by some additional noise (Fig. 3B). Therefore, the values of RQA measures quantifying diagonal line structures, as *DET*, are high. In contrast, the RP of noise consists mainly of single recurrence points, causing very low values in the RQA measures quantifying diagonal line structures. For higher SNR (SNR > 1), the performance of the measure *RR* becomes better and *RR* is also able to detect the signal.

The authors of [1] also tested measures based on the vertical line structures in a RP (e.g., *laminarity*). Vertical line structures are a sign of laminar phases in the analysed system. Due to a lack

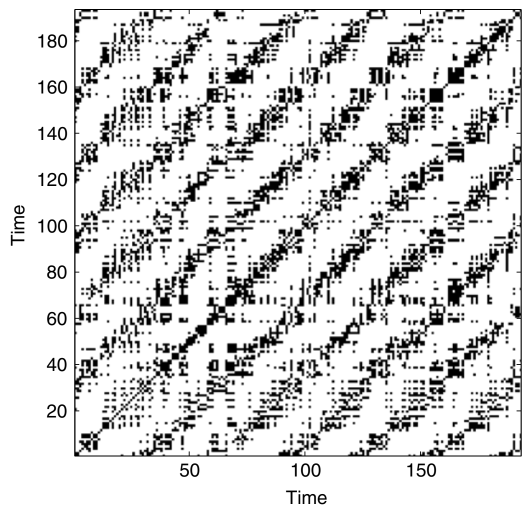


Fig. 3. Recurrence plot of the noise corrupted harmonic signal as used in Fig. 2 (SNR $= 2$). The embedding parameters are $m = 2$, $\tau = 8$ and the RP threshold $\varepsilon = 1$.

of laminar phases in the tested model, it is no surprise that such measures are not able to detect anything in this model.

2.2. Detection of known signals

In case the signal s is known, Rohde et al. proposed the application of a correlation detector and a cross recurrence plot (CRP) [13,15]. The measured signal is compared against the known

Author's personal copy

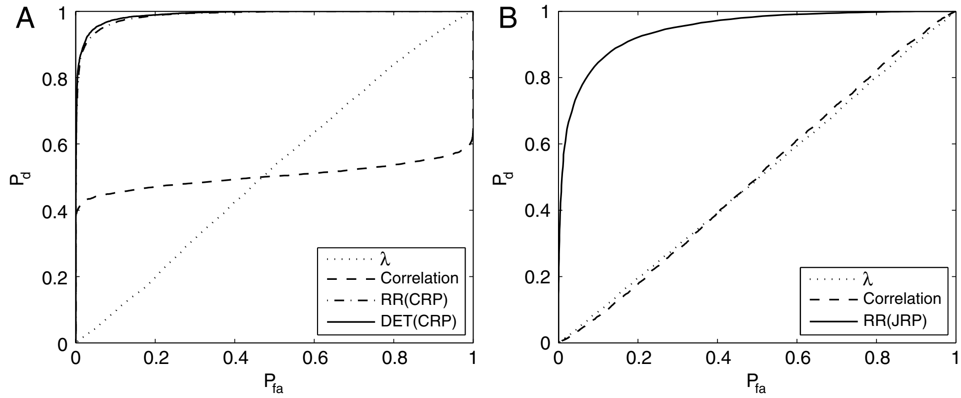


Fig. 4. (A) ROCs for the correlation receiver, the measure λ and the RQA measures *RR* and *DET* derived from a CRP for the detection of a noise corrupted signal from noise using a probe signal (SNR $= 0.5$). The incoming signal x consists of the original signal s plus noise ξ. (B) ROCs for the correlation receiver, the measure λ and the RQA measure *RR* derived from a JRP (SNR $= 0.3$). Here the signal s exists as a transformation of the measured signal x ($x = s^2 + \xi$). In both cases the signal is a stochastic harmonic process $s(t) = \cos(20t + \phi)$ with random phases ϕ, using a sampling interval of $\Delta t = 0.01$, data length $N = 1000$ and 10,000 realisations. The embedding parameters are $m = 2$, $\tau = 15$ and the RP threshold $\varepsilon = 1$ (for JRP, ε is chosen such that the *RR* of the RP is 0.1).

signal by the CRP. Rohde et al. compared the λ detector, Eq. (5), with the correlation detector. But they have not applied RQA measures (like *RR*) on such a CRP, as Zbilut et al. did [13]. Therefore, we argue that the statement that "the performance of the CRP detector proposed by Zbilut and colleagues [13] falls significantly below that of the traditional correlation receiver" [1] is actually not justified by the presented results.

For the signal and the probe, the authors used a cosine wave without phase randomisation. For this special case, the correlation receiver detector works very well, even for very low SNR as used by Rohde et al. (SNR $= 0.01$). For this low SNR, we have found that the CRP based RQA is indeed not able to detect the deterministic signals. However, it can detect deterministic signals for higher SNR.

But using a stochastic harmonic process, Eq. (6), i.e. considering phase randomisation, the correlation receiver also fails (even for high SNR), because of the random phase mismatch between the signal and the probe. In contrast, the CRP based RQA is then able to detect the deterministic signal where the well-known detector fails (Fig. 4A).

To understand the reason why the CRP works where the correlation receiver fails let us assume that we have two identical signals $\vec{x} = \vec{y}$. The corresponding CRP would be the same as the RP of one of those signals. We focus on the main diagonal line in the RP, which is also known as *line of identity* (*LOI*) [16]. Now we add small distortions to one of these signals, say a small amount of additive noise. The *LOI* is now interrupted (because we do not have identical states anymore; this line is now called *line of synchronisation* (*LOS*) [16]). The same happens to other diagonal line structures in the CRP. For two different systems, such diagonal lines correspond to those times, when the states of the two systems are similar and experience a similar evolution [2]. If the two systems are rather similar but have a difference in their phase, then simply the *LOS* departs from the main diagonal [2,16]. The lengths of the diagonal lines remain unchanged. Therefore, the RQA of the CRP would detect the deterministic signal even if there is a phase difference between signal and probe.

It is very important to emphasise that the correlation receiver detector works only well for linear problems. If we generalise the problem by allowing some functional change of the signal due to observation, i.e. by modifying the alternative hypothesis to

$$H_1: \text{ signal present, } x(t) = f(s(t)) + \xi(t),$$

the correlation receiver will fail (even for high SNR and no phase difference between signal and probe; Fig. 4B).

Instead of using a CRP, we suggest the calculation of a joint recurrence plot (JRP) for the detection of known signals [2,17]. The difference between these bi-variate methods is that a CRP tests for simultaneously occurring similar states, whereas a JRP tests for simultaneously occurring recurrences of states [2]. This is qualitatively different. The advantage of the JRP is that it is not as sensitive to the strong change of the amplitude of a state due to the noise as a CRP would be.

A JRP is the Hadamard product of the RPs of both systems. For simplicity, the recurrence thresholds for the two RPs should be chosen in such a way that the RPs contain the same number of recurrence points (i.e. *RR* should be equal). If both systems have the same recurrence structure, their JRP equals their RPs. Considering an increasing difference in their recurrence structures, the JRP loses more and more points (*RR* decreases). Therefore, the recurrence rate *RR* of the JRP (relative to the *RR* of the original RPs) can be used to compare two systems regarding their coinciding recurrence structure.

The measure *RR* derived from the JRP is able detect a deterministic signal using a probe signal, even if the signal was modified during the measurement process, as simulated by the applied function f. For example, in the case of $f(s) = s^2$ and SNR $= 0.3$, the correlation receiver is not able to detect the signal, whereas the *RR* derived from the JRP detects it clearly (Fig. 4B).

3. Conclusion

We conclude that the RQA can indeed be a powerful tool for the detection of deterministic signals and, hence, confirm previous studies [13,14,18,19]. However, a large amount of noise (as often occurs in real data, e.g. from EEG analysis, cardiology or geology) reduces the ability to detect deterministic signals of each method. The advantage of the RQA based approach is its ability to detect deterministic signals in the case of phase differences and nonlinear transformations of the input signal.

Finally, we should clearly state that all applications of RQA in order to detect determinism are pure heuristic approaches. High values of the RQA measure *determinism* do not imply a deterministic system in an exact mathematical sense. Using RQA we cannot conclude that a system is deterministic or not, but we can distinguish between systems of certain recurrence behaviour which may be characteristic for typical processes, such as white noise, correlated noise, chaotic maps, (quasi-)periodic processes etc. (in fact, using the discussed RQA measures and even the λ measure, we are able to distinguish an AR process from white noise — note that an AR process is not a deterministic process). Moreover, the application of measures based on recurrence line structures

Author's personal copy

needs justification in terms of the purpose of the intended analysis. For example, measures quantifying vertical recurrence structures are not appropriate if we are not interested in the detection of laminar phases. Therefore, it is important to understand the idea behind the measures of complexity provided by RQA before uncritical application.

Acknowledgements

We thank G.K. Rohde, J.M. Nichols, B.M. Dissinger and F. Bucholtz for the study on this important problem of signal detection for very low SNR, and J.M. Nichols, U. Schwarz and N. Wessel for constructive discussions. This study was made possible by grants from project MAP AO-99-030 of the Microgravity Application Program/ Biotechnology from the Human Spaceflight Program of the European Space Agency (ESA).

The calculation for the RQA was performed using the *CRP Toolbox*, available at tocsy.pik-potsdam.de/crp.php.

References

[1] G.K. Rohde, J.M. Nichols, B.M. Dissinger, F. Bucholtz, Stochastic analysis of recurrence plots with applications to the detection of deterministic signals, Physica D 237 (5) (2008) 619–629. doi:10.1016/j.physd.2007.10.008.

[2] N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems, Phys. Rep. 438 (5–6) (2007) 237–329. doi:10.1016/j.physrep.2006.11.001.

[3] J.P. Zbilut, N. Marwan, The Wiener–Khinchin theorem and recurrence quantification, Phys. Lett. A 372 (44) (2008) 6622–6626. doi:10.1016/j.physleta.2008.09.027.

[4] H. Sakoe, S. Chiba, Dynamic programming algorithm optimization for spoken word recognition, IEEE Trans. Acoust. Speech Signal Process. ASSP 26 (1) (1978) 43–49.

[5] J.V. Maizel, R.P. Lenk, Enhanced graphic matrix analysis of nucleic acid and protein sequences, Proc. Natl. Acad. Sci. 78 (12) (1981) 7665–7669.

[6] J. Foote, Visualizing music and audio using self-similarity, in: Proceedings of ACM Multimedia'99, Orlando, FL, 1999, pp. 77–80.

[7] R. Cutler, L. Davis, Robust periodic motion and motion symmetry detection, in: Proceedings of the Conference on Computer Vision and Pattern Recognition, South Carolina, USA, Vol. 2, 2000, p. 2615.

[8] N. Marwan, J. Kurths, Line structures in recurrence plots, Phys. Lett. A 336 (4–5) (2005) 349–357. doi:10.1016/j.physleta.2004.12.056.

[9] P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors, Physica D 9 (1–2) (1983) 189–208. doi:10.1016/0167-2789(83)90298-1.

[10] E.J. Ngamga, A. Nandi, R. Ramaswamy, M.C. Romano, M. Thiel, J. Kurths, Recurrence analysis of strange nonchaotic dynamics, Phys. Rev. E 75 (3) (2007) 036222. doi:10.1103/PhysRevE.75.036222.

[11] L.L. Trulla, A. Giuliani, J.P. Zbilut, C.L. Webber Jr., Recurrence quantification analysis of the logistic equation with transients, Phys. Lett. A 223 (4) (1996) 255–260. doi:10.1016/S0375-9601(96)00741-4.

[12] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Recurrence plot based measures of complexity and its application to heart rate variability data, Physical Review E 66 (2) (2002) 026702. doi:10.1103/PhysRevE.66.026702.

[13] J.P. Zbilut, A. Giuliani, C.L. Webber Jr., Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification, Phys. Lett. A 246 (1–2) (1998) 122–128. doi:10.1016/S0375-9601(98)00457-5.

[14] 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, Phys. Lett. A 267 (2–3) (2000) 174–178. doi:10.1016/S0375-9601(00)00098-0.

[15] N. Marwan, J. Kurths, Nonlinear analysis of bivariate data with cross recurrence plots, Phys. Lett. A 302 (5–6) (2002) 299–307. doi:10.1016/S0375-9601(02)01170-2.

[16] N. Marwan, M. Thiel, N.R. Nowaczyk, Cross recurrence plot based synchronization of time series, Nonlinear Process. Geophys. 9 (3/4) (2002) 325–331.

[17] M.C. Romano, M. Thiel, J. Kurths, W. von Bloh, Multivariate recurrence plots, Phys. Lett. A 330 (3–4) (2004) 214–223. doi:10.1016/j.physleta.2004.07.066.

[18] T. Aparicio, E.F. Pozo, D. Saura, Detecting determinism using recurrence quantification analysis: Three test procedures, J. Econ. Behav. Organ. 65 (3–4) (2008) 768–787. doi:10.1016/j.jebo.2006.03.005.

[19] S. Schinkel, O. Dimigen, N. Marwan, Selection of recurrence threshold for signal detection, Eur. Phys. J.—Special Topics 164 (2008) 45–53. doi:10.1140/epjst/e2008-00833-5.