Physics Letters A 373 (2009) 2245–2250

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

# Confidence bounds of recurrence-based complexity measures

Stefan Schinkel [a,*], N. Marwan [a,b], O. Dimigen [c], J. Kurths [b,d]

[a] Interdisciplinary Centre for Dynamics of Complex Systems, University of Potsdam, Germany
[b] Potsdam Institute for Climate Impact Research (PIK), Germany
[c] Department of Psychology, University of Potsdam, Germany
[d] Department of Physics, Humboldt University at Berlin, Germany

ARTICLE INFO

*Article history:*
Received 23 January 2009
Received in revised form 16 April 2009
Accepted 18 April 2009
Available online 24 April 2009
Communicated by C.R. Doering

*Keywords:*
Recurrence Plots
Complexity measures
Confidence bounds
Single trial ERP

ABSTRACT

In the recent past, recurrence quantification analysis (RQA) has gained an increasing interest in various research areas. The complexity measures the RQA provides have been useful in describing and analysing a broad range of data. It is known to be rather robust to noise and nonstationarities. Yet, one key question in empirical research concerns the confidence bounds of measured data. In the present Letter we suggest a method for estimating the confidence bounds of recurrence-based complexity measures. We study the applicability of the suggested method with model and real-life data.



## 1. Introduction

Recurrence Plots (RP) and their quantification (*recurrence quantification analysis*, RQA) [11] have become rather popular in various fields of science. The complexity measures based on RPs have helped to gain a deeper insight into diverse kinds of phenomena and experimental data. In this Letter we propose a straightforward extension to the existing RQA framework which allows us to not only compute these complexity measures, but also to estimate their confidence bounds. We do this by using a well-known resampling paradigm – the bootstrap. We show that the confidence bounds of RQA measures come with the regular analysis at virtually no extra costs and that the method can be useful for comparing univariate time series in a statistically sound fashion.

## 2. Recurrence Plots and their quantification

Recurrence is a fundamental property of dynamical systems. On this basis the data analysis tool called Recurrence Plot (RP) has been devised by Eckmann et al. [1] which visualises recurrences in the phase space of an $n$-dimensional state vector $\vec{x}_i$ $(i = 1, \ldots, N)$,

$$R_{i,j} = \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right), \tag{1}$$

where $\Theta$ is the Heaviside function, $\|\cdot\|$ is a norm and $\varepsilon$ is the recurrence threshold. The threshold $\varepsilon$ can be defined as an absolute value or in dependence on other criteria. For the examples in Fig. 1 we chose $\varepsilon$ so that the overall $RR$, Eq. (2), is 10%. The binary $N \times N$ matrix allows for a 2-dimensional visualisation of an $n$-dimensional attractor's recurrence properties (Fig. 1). From this matrix a number of well-defined complexity measures can be extracted (see [11] and references therein). If only univariate time-series are available the state vectors can be reconstructed using delay embedding with a given embedding dimension $m$ and a delay $\tau$ [14,19].

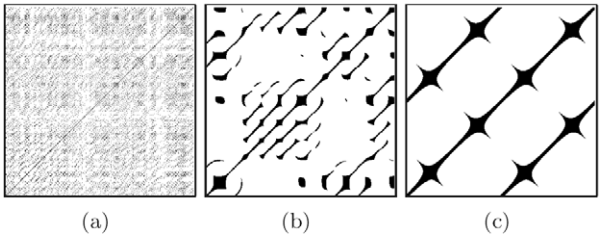


**Fig. 1.** Sample RPs for different dynamical systems. (a) Gaussian white noise, (b) the Lorenz attractor with $\rho = 28$, $\beta = 8/3$, $\sigma = 10$ and (c) a sine wave. Parameters were $m = 3$, $\tau = 3$, $\varepsilon$ adjusted to $RR = 0.1$.

### 2.1. RQA measures

The information contained in an RP can be quantified by measures of complexity based on recurrence point density, *diagonal*

* Corresponding author.
*E-mail address:* schinkel@agnld.uni-potsdam.de (S. Schinkel).

0375-9601/$ – see front matter 
doi:10.1016/j.physleta.2009.04.045

and *vertical* line structures. The simplest measure is the *recurrence rate RR*,

$$RR = \frac{1}{N^2}\sum_{i,j} R_{i,j}, \tag{2}$$

which denotes the overall probability that a certain state recurs. A measure based on the distribution of *diagonal* structures $P(l)$ is the *determinism DET*

$$DET = \frac{\sum_{l\geqslant l_{\min}}^{l_{\max}} lP(l)}{\sum_{l=1}^{l_{\max}} lP(l)}, \tag{3}$$

the ratio of recurrence points in diagonal lines (of at least length $l_{\min}$) to all recurrence points. *DET* reflects how predictable a system is. The measures $L_{\max}$ and $\langle L\rangle$ are the maximum and average lengths of the diagonal lines in $P(l)$.

Further complexity measures quantify the *vertical* structures in an RP. The ratio of points forming vertical line structures of a minimal length $v_{\min}$ to all recurrence points is called *laminarity*

$$LAM = \frac{\sum_{v\geqslant v_{\min}}^{v_{\max}} vP(v)}{\sum_{v=1}^{v_{\max}} vP(v)}, \tag{4}$$

a measure sensitive to laminar states and regimes of intermittency. From the distribution of the vertical line structures $P(v)$ we can again compute the maximal vertical line length $V_{\max}$ and the average called *trapping time TT*.

These measures can be computed from the whole RP or in moving, possibly overlapping, windows of size $w$ shifted along the main diagonal of the RP by an increment of $s$. This approach is useful to reveal qualitative transitions in a system.

The RQA measures provide a qualitative description of a system in terms of complexity measures. It allows to detect transitions in the system's dynamics, e.g. transitions from period to chaos, from strange nonchaotic attractors (SNA) to chaos or even transitions from chaos to chaos [9,13].

In the next section we will focus on how to derive a quantitative judgement from these measures.

## 3. Confidence bounds of RQA measures

The RQA measures have been quite useful for the analysis of a variety of data. Yet, in order to not only detect qualitative changes in a system's dynamics but to be able to judge their significance or to compare two univariate time series, it is necessary to derive a quantitative judgement such as a confidence interval. For recurrence-based complexity measures those intervals can be estimated using a resampling paradigm.

### 3.1. Resampling statistics – the bootstrap

Statistical techniques based on resampling were among the first methods ever thought of. Sir R.A. Fisher himself introduced this idea when pondering over Gosset's $t$-distribution [3,4]. Due to lacking computational power, these ideas were not feasible at that time. With the advent of powerful, low-cost computers these methods have gained a broad interest and have been proven to be very reliable and powerful. In this Letter we focus on one particular resampling method – the bootstrap [2]. The bootstrap is a nonparametric method for estimating the variance of a statistic of interest. It relies on resampling of a given distribution *with* replacement and does not require any specific probability distribution. The bootstrap procedure works as follows:

Given a random sample $x_i$ $(i = 1, 2, \ldots, n)$ of size $n$, from an unspecified probability distribution we compute a statistic of interest, say, the mean $\langle x\rangle$. In order to estimate the variance of that statistic we draw at random and *with* replacement the same number $(n)$ of elements from $x_i$ to obtain the resampled distribution $x_i^*$. From $x_i^*$ we again compute the statistic of interest. *With* replacement means that we can draw the individual elements in $x_i$ more than once. Doing this a larger number of times[1] we obtain the empirical distribution of the statistic of interest, $\hat{P}_{\langle x\rangle}$. From the empirical distribution we can compute the percentiles $\alpha/2$ and $1-\alpha/2$ and define the $(100-\alpha)\%$ confidence interval (CI) as the range between those two percentiles.

The empirical distribution could also be used to perform hypothesis testing. We opt for the estimation of confidence intervals only. The interpretation of hypothesis tests, especially $p$-values, the chosen indicator of significance, is currently under discussion and not agreed upon by the frequentist and Bayesian schools. Therefore we follow the suggestions of Hubbard and Lindsay [6] and only estimate the confidence intervals of the RQA measures. This allows us to not only detect transitions in the dynamics of one system or to differences between the dynamics of two systems but to provide a judgement whether those differences are statistically significant. While this is not statistical testing in the narrowest sense, we think this approach is more appropriate as the data is explicitly shown and the investigator may judge for him-/herself.

### 3.2. Structure preserving resampling

Since the bootstrap relies on resampling *with* replacement we cannot simply bootstrap the RP matrix as such for two reasons. First of all, we could draw one of the black points more than once. As the RP is a binary matrix by definition this is not possible. Secondly, randomly resampling an RP would necessarily result in a loss of most of the small-scale structures in it (i.e. diagonal and vertical lines). A loss of structures would result in an RP corresponding to noise. This is not desirable because we want to compare different systems against each other and not test against randomness/noise.

As stated above, RQA measures like *DET* or *LAM* rely on the distribution of line structures $P(l)$ and $P(v)$. Therefore we present a method that ensures that the structural elements are preserved during resampling. We only resample the distributions of diagonal and vertical lines, $P(l)$ and $P(v)$. It is important to note that we need to resample all lines in $P(l)$ and $P(v)$, even those of only length 1, thereby obtaining $P^*(l)$ and $P^*(l)$, respectively. The value of determinism is then computed according to:

$$DET^{*i} = \frac{\sum_{l\geqslant l_{\min}}^{l_{\max}} lP^{*i}(l)}{\sum_{l=1}^{l_{\max}} lP^{*i}(l)} \quad (i = 1, 2, \ldots, n_{bs}) \tag{5}$$

for each bootstrapped sample (see Fig. 2). The computation for a bootstrapped sample of $\langle L\rangle^*$, $LAM^*$ and $TT^*$ is done accordingly. Repeating this procedure $n_{bs}$ times we obtain the empirical distributions $\hat{P}_{DET}$, $\hat{P}_{\langle L\rangle}$, $\hat{P}_{LAM}$ and $\hat{P}_{TT}$. From the empirical distributions we can calculate the percentiles $\alpha/2$ and $1-\alpha/2$. The two-sided $(100-\alpha)\%$ confidence interval is then defined as the range between those two percentiles. The value $\alpha$ determines the spread of the interval, the smaller $\alpha$, the broader the interval. As we leave the structures in the RP intact, we refer to this procedure as *structure preserving resampling*.

For obvious reasons this approach is restricted to *DET*, $\langle L\rangle$, *LAM* and *TT*. $V_{\max}$ and $L_{\max}$ already represent maxima in the distribution and are therefore very unlikely to show variation and the upper bound cannot vary at all. Furthermore, we can apply this

[1] The number of resamplings is not generally agreed upon but common guidelines suggest values between 800 and 1500. Note also that the number of resampling decreases with the *randomness* of the samples drawn. In the present manuscript we use the MT19937 algorithm, which has a period of $2^{19937}-1$ [12].

procedure either to the whole RP or to moving windows in order to obtain CIs over time. For this purpose the resampling and CI estimation is applied to each window separately.

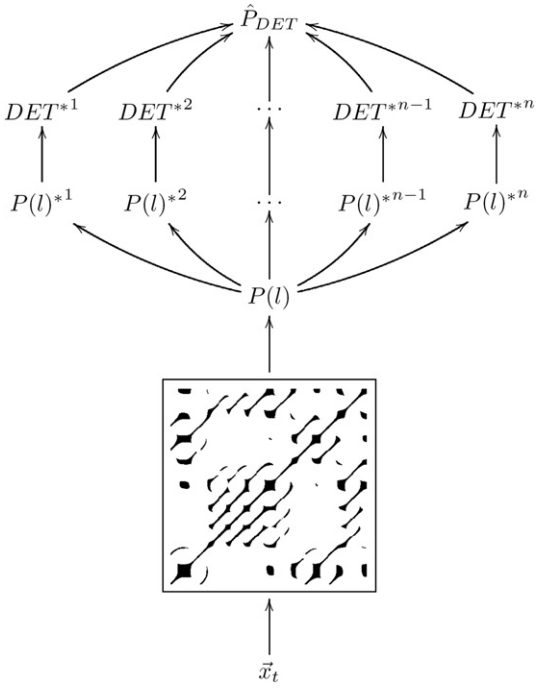


**Fig. 2.** The bootstrap procedure exemplified. It is important to note that only the line structures are resampled. In this procedure we estimate the variance of the RQA measures in question but do not construct surrogate data.

We can then plot the CIs as coloured bands ranging from the lower to the upper bound of the estimated CI. In Fig. 3 the CI of the complexity measures calculated from the logistic map in the chaotic regime is given as a red band, the CI of the logistic map with mutual transitions as a blue band.

## 4. Prototypical example

As a prototypical example we compare the logistic map in the chaotic regime (Fig. 3, upper panel):

$$x_{i+1} = a x_i (1 - x_i) \tag{6}$$

with $a = 3.92$ to the logistic map with mutual transitions (Fig. 3, middle panel) given as:

$$x_{i+1} = a_i x_i (1 - x_i). \tag{7}$$

In difference to the standard logistic map, in the latter the control parameter $a$ is changed with every iteration of the map ($a_{i+1} = a_i + \Delta a$).

For control parameter $a$ in the range of $[3.8; 3.88]$ ($\Delta a = 0.00001$) we find an island of stability starting at $a = 1 + \sqrt{8} \simeq 3.828$ with a period-3 window. For larger values of $a$ the system evolves into chaos by period-doubling. For $a \geqslant 3.856$ the system is again in a chaotic regime with a very short periodic window for $a = [3.858; 3.859]$. On the other hand, the standard logistic map is in the chaotic regime for the chosen control parameter ($a = 3.92$). In the chaotic regime the RP is mostly composed of short lines and single dots, thus the value of $\langle L \rangle$ should be low and the estimated confidence interval rather narrow as both systems are unpredictable. On the other hand, when the logistic map with transitions is in a periodic regime, the corresponding RP contains continuous, long diagonal lines close to the main diagonal and shorter, but still continuous lines, towards the edges of the plot.

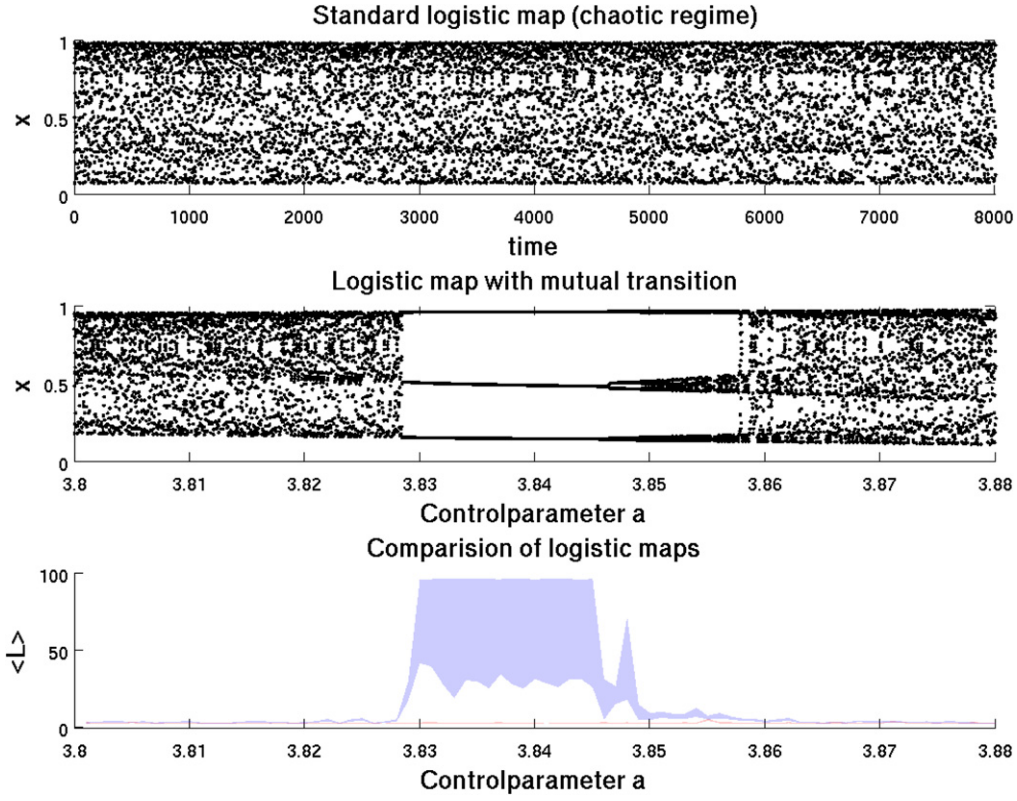


**Fig. 3.** Comparison of the standard logistic map and the logistic map with mutual transitions. The latter is in a periodic regime for $a = [3.828; 3.856]$ reflected in a significant increase in $\langle L \rangle$. The light blue band corresponds to the logistic map with transitions, the light red band to the logistic map in the chaotic regime. Parameters: $m = 2$, $\tau = 1$, $w = 100$, $s = 100$, fixed *RR* 10%. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this Letter.)

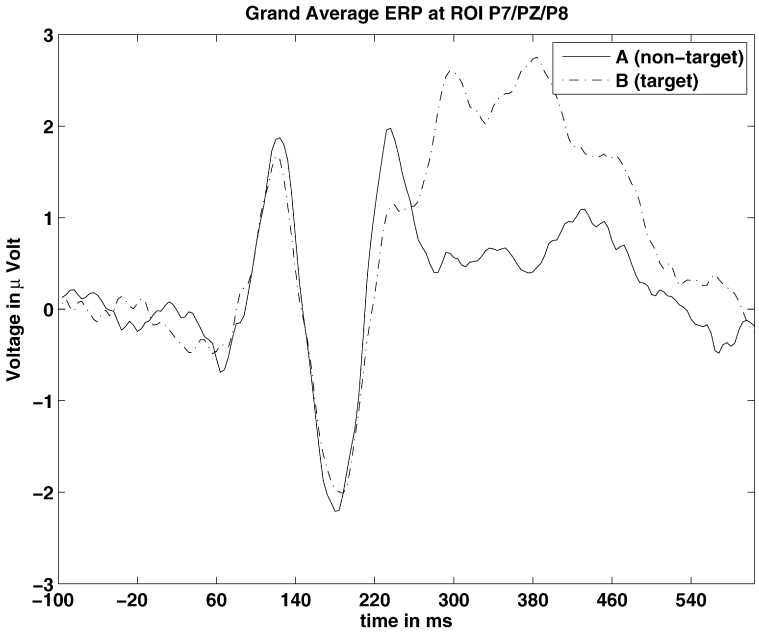


**Fig. 4.** Grand Average ERP for the visual oddball task. Spatial average at P7, PZ and P8 with 1800 measurements in the non-target (solid line) and target condition (dash-dotted line). The P300 in the target condition is clearly visible.

Therefore the value for $\langle L \rangle$ should increase and the corresponding confidence interval should become broader because resampling lines closer to the edges of the plot or closer to the main diagonal is equally likely. Hence we should be able to distinguish both systems when one is in a periodic regime, i.e. the confidence intervals should not overlap. For the RQA we use an embedding dimension of $m = 2$ and a delay of $\tau = 1$, for the bootstrap procedure we chose 1000 resamplings.

If the complexity measures, respectively their confidence bounds, perform as expected, we should be able to distinguish the logistic map with mutual transition from its counterpart at least in the periodic range as such states are not present in the standard logistic map with the chosen control parameter. This is indeed the case. Starting when both systems are chaotic, the confidence bounds for the RQA measure $\langle L \rangle$ of the chaotic and the test system overlap. For the island of stability $\langle L \rangle$ (Fig. 3, lower panel in light blue) is significantly higher and the confidence intervals do not overlap, whereas in the chaotic regime the confidence intervals do overlap. Even the very small periodic window for $a = [3.858; 3.859]$ occurring as the system evolves into chaos is detected. The confidence bounds then start to overlap again as both systems are chaotic.

## 5. Application

We apply the suggested procedure to measurements of electric brain signal (electroencephalogram, EEG) in a study on event-related potentials (ERPs).

### 5.1. Material and methods

The paradigm used was a visual *oddball* featuring a prominent P300, a centro-parietal positivity peaking at about 300 ms after the presentation of a task relevant stimulus [18].

The stimuli were red and green disks presented in randomised, equiprobable order. Stimulation duration was 100 ms, the interval between successive stimuli 900 ms. The task was to count the items of one colour (green or red) thereby constituting the non-target (A) and the target condition (B) (items to be counted).

The EEG was recorded from 40 Ag/AgCl electrodes (impedances $\leqslant 5$ kΩ) at a sampling rate of 250 Hz using a BrainAmp DC amplifier (Brain Products GmbH, Munich, Germany). All electrodes were initially referenced to an electrode on the left mastoid bone (A1) and converted to average reference off-line. Details of data acquisition and pre-processing can be found in [20]. The EEG data shown here was spatially averaged over 3 centro-parietal electrodes (P7, PZ and P8) and lowpass-filtered at 20 Hz.

For the RQA we use *order patterns* recurrence plots (OPRP) as introduced by [5]. In this variant of RPs the data series is symbolised according to local rank relations and recurrence is defined in this symbol space (see [5] for details). OPRPs have been shown to be a suitable tool for ERP analysis [10,15] and further omit the need for selecting a threshold $\varepsilon$, which cannot be estimated easily [16]. The embedding parameters were estimated using the commonly accepted methods of false nearest neighbours and mutual information: $m = 3$, $\tau = 7$ [7]. We use a moving window of size $w = 60$ shifted by $s = 1$ datapoints (frames). The time axis was rescaled to adjust for pattern length and $w/2$ to align time to the middle of a moving window. For the bootstrap procedure we used 1200 resamplings.

### 5.2. Results

The P300 is clearly visible in the Grand Average ERP (Fig. 4) which is comprised of a total of 1800 trials. The challenge now is to detect the component in the individual measurements. Note that these measurements are extremely noisy, nonstationary and rather short (276 sample points in total).

For our analysis we focus on the RQA measure *LAM* for which we computed the 95% confidence interval. *LAM* is sensitive to slowly changing dynamics in a system and has been proven to be sensitive to ERP components [8,17]. From the single trial measurements it is not feasible to assume a P300 (Fig. 5, upper panels). Yet, the RQA does detect the component even in a single measurement, reflected by a higher level of *LAM* in the target condition (B) (Fig. 5, lower panels, red areas). Note that the method is only sensitive to the ERP effect as the CIs always overlap except for the relevant time window. Secondly, the P300 varies in its time course. This

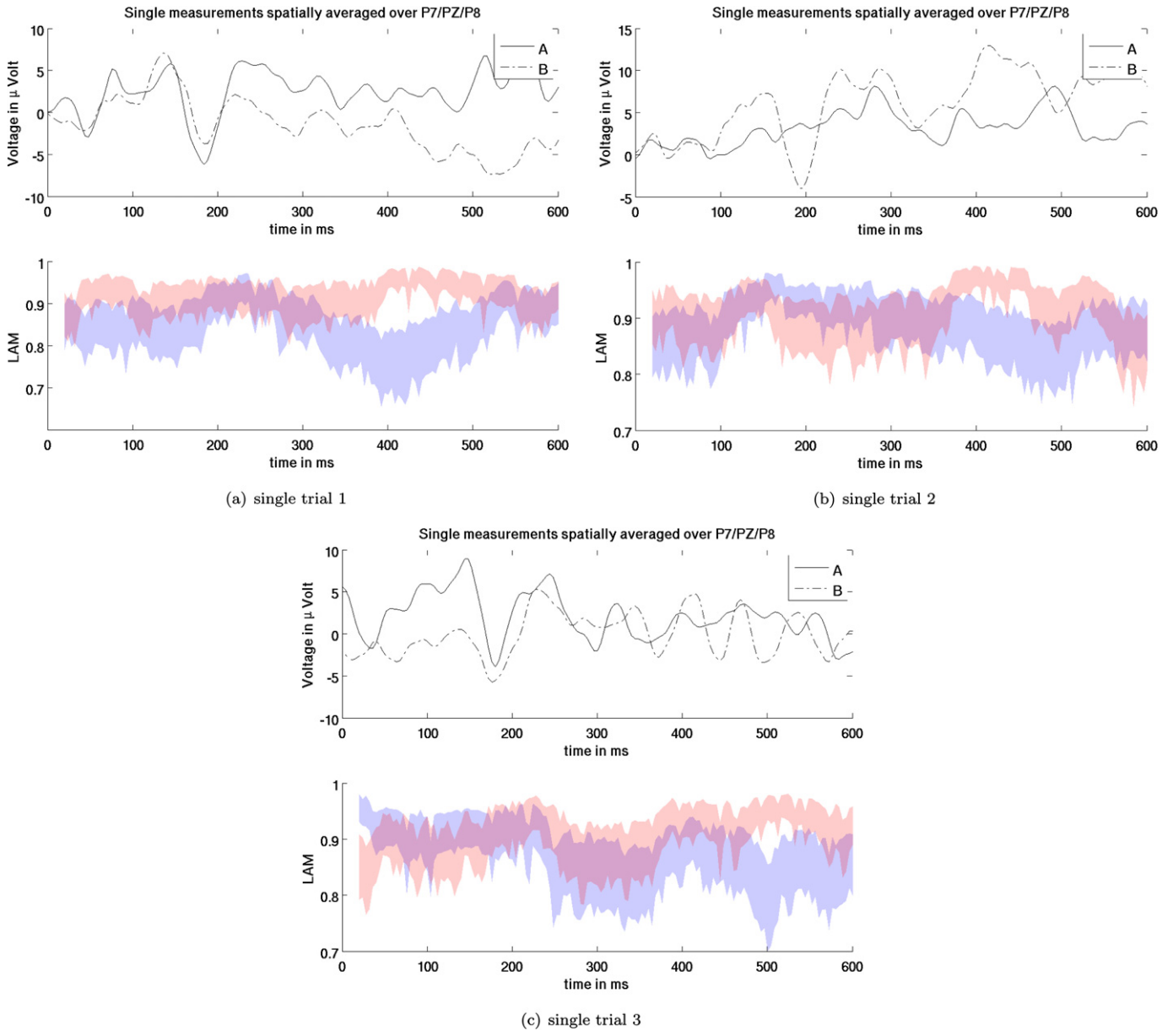


**Fig. 5.** Comparison of three single trials (a)–(c) in the visual oddball task. Spatial average at P7, PZ and P8. From the voltage measurements it is hard to assume a P300. The RQA does detect the component and is only sensitive to the effect in question, confidence bounds do not overlap only in the relevant time window. RQA parameters: $m = 3$, $\tau = 7$, $w = 60$ and $s = 1$, order patterns RP. Time axis rescaled to compensate for pattern length and window size.

phenomenon is well known and referred to as *latency jitter*. In the data at hand the P300 varies in onset (about 380 ms to 450 ms), and its duration (50 to 250 ms). These are again well-known facts but are not accessible with methods that rely on averaging only.

## 6. Discussion

The extension to the existing RQA framework presented here is straightforward and easily applicable. The RQA is known to be a useful tool for analysing noisy and nonstationary data, as is the case with EEG recordings. By providing confidence bounds to a number of recurrence-based complexity measures, we cannot only detect properties of the investigated data, but we can also extend the analysis to statistical comparisons, which is often called for in empirical and experimental research. As shown here, the presented method even enables us to perform single trial ERP analysis.

The main advantage of our approach is obvious. Not only are we able to discriminate the two conditions qualitatively, which would be possible with the RQA as such, but we are also able to give a quantitative judgement. As shown here, we can distinguish two experimental conditions with 95% confidence even on the scale of single measurements.

## Acknowledgements

This work supported by grants of the German Research Foundation (DFG) in the Research Group FOR 868 and the SFB 555 *Komplexe nichtlineare Systeme* and by the EU in the COST Action BM0601 and the *BioSim Network of Excellence*. The software used for this Letter, the CRPtoolbox & extensions, is available for download at http://tocsy.agnld.uni-potsdam.de.

## References

[1] J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Europhys. Lett. 5 (1987) 973.
[2] B. Efron, R.J. Tibshirani, An Introduction to the Bootstrap, Chapman and Hall, London, New York, 1993.

[3] R.A. Fisher, The Design of Experiments, Hafner, New York, 1935.
[4] W.S. Gosset, Biometrika 6 (1) (1908) 1.
[5] A. Groth, Phys. Rev. E 72 (4) (2005) 046220.
[6] R. Hubbard, R.M. Lindsay, Theory Psychol. 18 (1) (2008) 69.
[7] H. Kantz, T. Schreiber, Nonlinear Time Series Analysis, University Press, Cambridge, 1997.
[8] N. Marwan, A. Meinke, Int. J. Bifur. Chaos Cognition Complex Brain Dyn. 14 (2) (February 2004) 761.
[9] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Phys. Rev. E 66 (2) (2002) 026702.
[10] N. Marwan, A. Groth, J. Kurths, Chaos Complex. Lett. 2 (2/3) (2007) 301.
[11] N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Phys. Rep. 438 (2007) 237.
[12] M. Matsumoto, T. Nishimura, ACM Trans. Model. Comput. Simul. 8 (3) (1998).
[13] E.J. Ngamga, A. Nandi, R. Ramaswamy, M.C. Romano, M. Thiel, J. Kurths, Phys. Rev. E 75 (3) (2007) 036222.
[14] N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Phys. Rev. Lett. 45 (9) (1980) 712.
[15] S. Schinkel, N. Marwan, J. Kurths, Cognitive Neurodyn. 1 (4) (2007) 317.
[16] S. Schinkel, O. Dimigen, N. Marwan, Eur. Phys. J. Special Topics 164 (2008) 45.
[17] S. Schinkel, N. Marwan, J. Kurths, Brain Signal Anal. Based Recurrences, submitted for publication.
[18] S. Sutton, M. Braren, J. Zubin, E.R. John, Science 150 (1965) 1187.
[19] F. Takens, Lect. Notes Math. 898 (1) (1981) 366.
[20] M. Valsecchi, O. Dimigen, R. Kliegl, W. Sommer, M. Turratto, Microsaccadic inhibition and P300 enhancement in a visual oddball task, Psychophysiology, in press.