Autonomic Neuroscience: Basic and Clinical xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Autonomic Neuroscience: Basic and Clinical

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

Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods

G.M. Ramírez Ávila [a,b,c], A. Gapelyuk [a], N. Marwan [b], H. Stepan [d], J. Kurths [a,b,e], Th. Walther [f,g], N. Wessel [a,*]

[a] Department of Physics, Humboldt-Universität zu Berlin, Berlin, Germany
[b] Potsdam Institute for Climate Impact Research, Potsdam, Germany
[c] Instituto de Investigaciones Físicas, Universidad Mayor de San Andrés, La Paz, Bolivia
[d] Department of Obstetrics and Gynecology, University of Leipzig, Leipzig, Germany
[e] Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen, United Kingdom
[f] Department of Pediatric Surgery and Department of Obstetrics, University of Leipzig, Leipzig, Germany
[g] Institute for Experimental and Clinical Pharmacology and Toxicology, Medical Faculty Mannheim, University of Heidelberg, Heidelberg, Germany

ARTICLE INFO

Article history:
Received 12 November 2012
Received in revised form 24 April 2013
Accepted 2 May 2013
Available online xxxx

Keywords:
Heart rate
Blood pressure
Cardiac dynamics
Heart rate
Preeclampsia
Recurrences
Networks
Time series analysis

ABSTRACT

It is urgently aimed in prenatal medicine to identify pregnancies, which develop life-threatening preeclampsia prior to the manifestation of the disease. Here, we use recurrence-based methods to distinguish such pregnancies already in the second trimester, using the following cardiovascular time series: the variability of heart rate and systolic and diastolic blood pressures. We perform recurrence quantification analysis (RQA), in addition to a novel approach, ε-recurrence networks, applied to a phase space constructed by means of these time series. We examine all possible coupling structures in a phase space constructed with the above-mentioned biosignals. Several measures including recurrence rate, determinism, laminarity, trapping time, and longest diagonal and vertical lines for the recurrence quantification analysis and average path length, mean coreness, global clustering coefficient, assortativity, and scale local transitivity dimension for the network measures are considered as parameters for our analysis. With these quantities, we perform a quadratic discriminant analysis that allows us to classify healthy pregnancies and upcoming preeclamptic patients with a sensitivity of 91.7% and a specificity of 45.8% in the case of RQA and 91.7% and 68% when using ε-recurrence networks, respectively.

1. Introduction

Nowadays, a severe pathology called preeclampsia (PE) affects healthy nulliparous women in a range between 2% and 7% worldwide (Sibai et al., 2005). The main features of PE are severe hypertension and proteinuria for which the pathophysiology is not well understood at present. Several strategies are used in order to predict PE, among which we can mention biochemical markers, such as fms-like tyrosine kinase 1 (sFlt-1), placental growth factor (PlGF), soluble endoglin (Ohkuchi et al., 2011; Rana et al., 2007), maternal autoantibody, angiotensin II type I receptor agonistic autoantibody (AT1-AA) (Siddiqui et al., 2010), urinary biomarkers (Carty et al., 2011), noninvasive cardiovascular (CV) indicators (Malberg et al., 2007; Walther et al., 2006), or a combination of the above (Stepan et al., 2008).

In recent years, recurrence methods based on recurrence plots (RP) have been successfully used in different fields of natural sciences as physics (Ngamga et al., 2012) and biology (Angus et al., 2012), but also to answer economic (Hirata and Aihara, 2012) or medical questions (Wessel et al., 2009). Recurrence quantification analysis (RQA), in particular, constitutes a very useful tool for the description and analysis of a systems diversity (Marwan, 2008; Marwan et al., 2007). More recently, the recurrence concept has been extended to networks and applied in novel time series analysis methods (Marwan et al., 2009), finding several applications such as in paleoclimate modeling (Donges et al., 2009).

The detection of cardiovascular disorders has been considerably improved due to both technological advances and new methods of time series analysis. Nevertheless, there are still unclear mechanisms that cannot be explained by standard data analysis. Nonlinear data analysis and modeling methods of CV physics allow to improve clinical diagnostics and also a better understanding of CV regulation. One of the most important aspects of these methods is that they focus on noninvasive measured biosignals. Among the biosignals that CV physics deals with are the heart rate variability (HRV) and the variabilities of systolic blood pressure (SBPV) and diastolic blood pressure (DBPV).

* Corresponding author.
E-mail address: wessel@physik.hu-berlin.de (N. Wessel).

1566-0702/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.autneu.2013.05.003

In this work, we apply the approach of RQA and ε-recurrence networks to analyze CV biosignals, obtained by noninvasive techniques, with the aim of developing a classification method to identify patients who develop PE in a pool of pregnancies within the second trimester.

2. Methods

2.1. Clinical aspects

We considered for this study 96 pregnancies with abnormal uterine perfusion (AUP), followed by means of Doppler sonography in the second trimester, between the 18th and the 26th week of gestation (WOG) of pregnancy, at the Department of Obstetrics and Gynecology of the University of Leipzig, Germany. Immediately after the Doppler examination, the blood pressure was measured noninvasively via finger cuff for 30 min (sampling rate: 100 Hz, Portapres device model 2, BMI-TNO, Amsterdam, The Netherlands). The continuous blood pressure curves were used to extract the time series of beat-to-beat intervals and systolic and diastolic blood pressures, allowing us to obtain the CV values (HRV, SBPV, and DBPV). The length of the dataset per variable is roughly of 1600 samples (heart beats). At the time of examination, the women were healthy, normotensive, without clinical signs of cervical incompetence, and on no medication. After the 30th WOG, 24 patients developed PE. Further details on the methodology can be found in Malberg et al. (2007). We point out that the root mean square errors of heart beats calculated from blood pressure curves (compared to ECG slope detection) is about 5–6 ms (Suhrbier et al., 2006). Therefore, the computation of the beat-to-beat-intervals from the distal pulse wave measurement as it has been performed in this paper is an acceptable alternative; however, this has to be confirmed in another comparative study.

2.2. Recurrence methods

The concept of recurrence applied to a single trajectory of the dynamical system allows us to obtain the recurrence matrix whose elements are given by $R_{i,j} = \Theta(\varepsilon - ||\mathbf{x}_i - \mathbf{x}_j||)$, where $\Theta(\cdot)$ represents the Heaviside function, $||\cdot||$ is a suitable norm, and ε is a threshold distance that should be chosen adequately according to the characteristics of the embedded attractor into the phase space. We use RQA and ε-recurrence networks with the aim of distinguishing between healthy individuals and patients with PE.

2.2.1. Recurrence quantification analysis

The RQA is a method of nonlinear data analysis that quantifies the number and duration of recurrences of a dynamical system presented by its state space trajectory. This method was developed by Zbilut and Webber (1992) and extended by Marwan et al. (2002). Several measures might be used to quantify the time series of a system when using RQA, such as the following: recurrence rate (RR), the percentage of recurrence points in an RP, corresponding to the correlation sum; determinism (DET), the percentage of recurrence points forming diagonal lines; laminarity (LAM), the percentage of recurrence points forming vertical lines; trapping time (TT), the average length of the vertical lines; and some other self-explanatory measures such as longest diagonal line (L_{MAX}) and longest vertical line (V_{MAX}). A more detailed description of these measures can be found in Marwan et al. (2007).

2.2.2. Recurrence networks

The basic idea of time series analysis based on complex network techniques relies on the fact that a time series may be transformed into a complex network from which we can extract the adjacency matrix, allowing us to obtain local and global network properties (Donner et al., 2011). We interpret the recurrence matrix **R** as the adjacency matrix of an unweighted and undirected complex network, commonly called the ε-recurrence network, which is associated with a given time series. Possible self-loops must be avoided in this network; thus, a Kronecker delta must be subtracted from the recurrence matrix. The elements of the adjacency matrix for an ε-recurrence network are thus

$$A_{i,j}(\varepsilon) = R_{i,j}(\varepsilon) - \delta_{i,j}, \tag{1}$$

where the ε-dependence is considered explicitly as in the case of RQA. There is no universal criterion for choosing ε, but the choice must be made avoiding too small values, which lead to a situation in which there are not enough recurrence points, or too large values, implying that every vertex is connected with many other vertices irrespective of their actual mutual proximity in phase space (Donner et al., 2010b). Having reconstructed the adjacency matrix **A** from a time series, we can apply appropriate network characteristics to analyze and obtain information on the underlying system (Donges et al., 2012). In Appendix A, there is an explanation of how to obtain the adjacency matrix, the associated network, and the 4-element motifs. In this work, we focus our interest on five global network measures: the average path length ($\mathcal{L}$), which is the mean value of the shortest geodetic path lengths $l_{i,j}$ considering all pair of vertices (i,j); the mean coreness ($\mathcal{C}\prime$), which is the average of the coreness (significance of a node and its "popularity" in the network) of all the vertices (Batagelj and Zaveršnik, 2002); the global clustering coefficient ($\mathcal{C}$), which is the average of the clustering coefficient of each vertex (ratio of triangles including vertex i and the number of triples centered on vertex i, where triple refers to a pair (j,k) of vertices that are both linked with i, but not necessarily mutually linked); the assortativity ($\mathcal{A}$), the tendency for vertices in networks to be connected to other vertices that are like (or unlike) them in some way (Newman, 2003); and the scale local transitivity dimension ($D_{\mathcal{T}}$), defined as $D_{\mathcal{T}} = \frac{\log \mathcal{T}}{\log(3/4)}$, where $\mathcal{T}$ is the transitivity (ratio of the number of triangles in the network times three and the number of linked triples of vertices). These four measures depend on ε and have a global character. A detailed description of networks and their properties can be found in Boccaletti et al. (2006).

3. Data processing and statistics

We use an algorithm that avoids artifacts such as extrasystolic beats. The original time series from consecutive R waves were filtered using a preprocessing algorithm that first removes obvious recognition errors, then applies an adaptive percent filter, and finally an adaptive controlling filter (Wessel et al., 2007). With the aim of using a recurrence approach, we consider the three CV indicators and several possible embeddings. An estimation of the coupling structure of CV indicators has been performed using nonlinear additive autoregressive models with external input, following the idea of Granger causality (Riedl et al., 2010). This coupling analysis shows that HRV, DBPV, and SBPV respond to respiration; SBPV respond to DBPV and the latter to HRV. In our case, we do not consider respiration; thus, the coupling structure may be represented as in Fig. 1(a), where, according to the coupling scheme, there is a delay between the HRV, the DBPV, and the SBPV. For simplicity, we write down the coupling structure as (HRV(t), DBPV ($t+1$), SBPV ($t+2$)), or simply $H(t)D(t+1)S(t+2) \equiv 012$.

We sought to predict whether or not a patient develops PE using the CV indicators embedded in a phase space determined by the structure of coupling. We consider a minimalist assumption in which the structure of coupling between HRV, DBPV, and SBPV is identical in each subject of a group and that this structure does not change during the measurement. In this study, we set out to test all the possible structures of coupling shown in Fig. 1 and a wide range of the threshold ε going from 0.01σ to 0.99σ, where σ is the standard deviation of the underlying process in the embedded phase space. From a simple CV time series corresponding to each patient, we construct a complex network for each possible structure of coupling and each value of ε. Then we compute the four network measures: $(\mathcal{C}, \mathcal{L}, \mathcal{C}\prime, D_{\mathcal{T}})$, and with these new measures, we perform an analysis to classify the groups of individuals: healthy and preeclamptic patients. For that purpose, we firstly verify

Please cite this article as: Ramírez Ávila, G.M., et al., Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. (2013), http://dx.doi.org/10.1016/j.autneu.2013.05.003

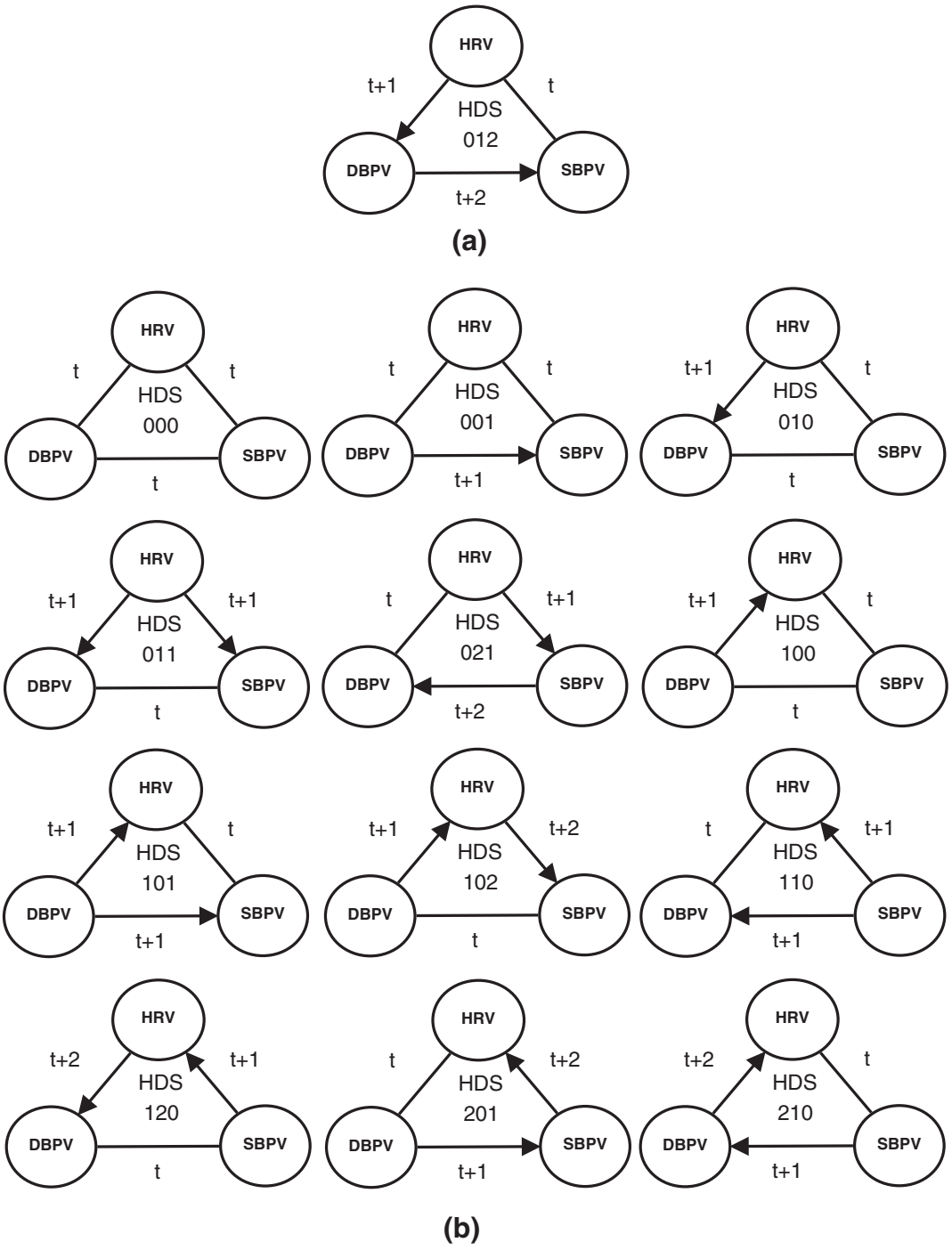

Fig. 1. (a) Coupling structure considering that HRV drives the DBPV and this in turn the SBPV (directed arrows from HRV to DBPV and from DBPV to SBPV). Note that when the variables are linked only by a line, it means that these are coupled but without any delay. This might be written schematically as $H(t)D(t+1)S(t+2) \equiv 012$; the latter number can change according to the delay among the sequential variables HRV, DBPV, and SBPV, represented as *HDS*. (b) All the other possibilities of coupling structures.

whether or not these new parameters are significant by means of a Mann–Whitney *U*-test and considering a significance level of 5%; here, the null hypothesis is that the data corresponding to control and pre-eclamptic patients are independent samples from identical continuous distributions with equal medians, against the alternative that they do not have equal medians.

3. Results

Firstly, for values of the threshold ε in the range of 0.01σ to 0.99σ, we obtain the matrices **R** and **A** that enable us to perform RQA and obtain the RP measures mentioned in Section 1 and the ε-recurrence networks measures described in Section 2. As an example of the networks obtained, Fig. 2 shows a visualization of the associated networks, obtained using the medians of the time series for patients exhibiting PE and control individuals. These representations are constructed using the coordinates of the nodes. An inspection of these networks (PE and control) allows us to perceive some differences between them, as, for example, the existence of more free nodes (more outliers from a statistical point of view) in the case of the control group network compared to the PE group network and the apparent node degree that seems to be higher in the control group network.

Please cite this article as: Ramírez Ávila, G.M., et al., Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. (2013), http://dx.doi.org/10.1016/j.autneu.2013.05.003

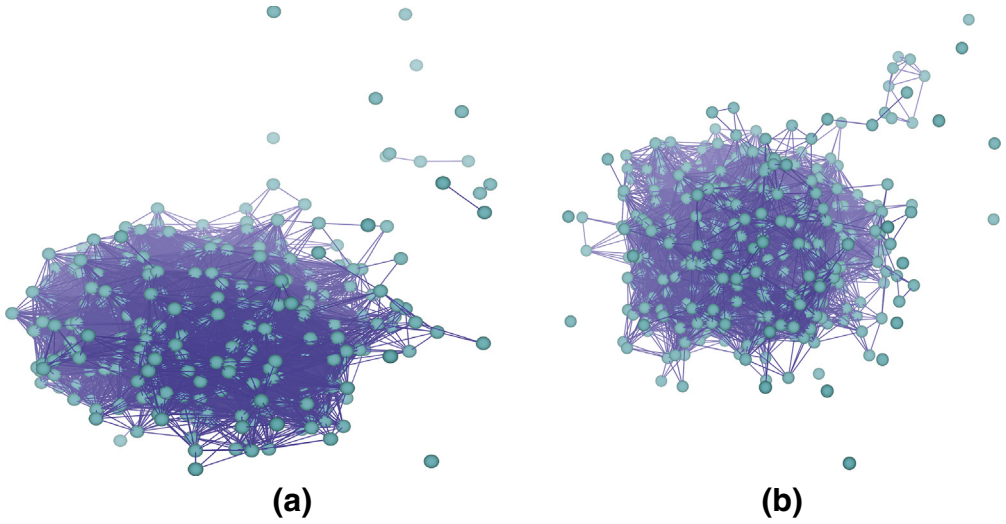

Fig. 2. (Color online) Visualization of the networks obtained using the time series of the medians for both groups of individuals, (a) PE and (b) control. The visualization has been obtained by means of the software Pajek (Pajek, 2011), with a 3-dimensional perspective and using all the nodes and their corresponding coordinates in the phase space HDS(t).

Nonetheless, this visual inspection is just a first check that cannot replace the quantification of the network measures.

The results for each RQA and network measure are represented in the phase plane, embedding (structure of the coupling) vs. ε, as shown in Figs. 3 and 4, respectively. The color code indicates the p-values of the statistical test when the null hypothesis H_0 of equal medians at 5% significance level is rejected. The white pixels denote that there is no difference between both groups ($p \geq 0.05$), and pink ones denotes the impossibility to compute tp. On the contrary, the black pixels represent the minimum p-value among all the possibilities on the phase plane.

According to Figs. 3 and 4, the significant values for each network measure occur only for some coupling structures and thresholds ε. Figs. 5 and 6 show the same plane as in Figs. 3 and 4 but considering the cases in which all the four RQA and network measures are simultaneously significant, i.e., $p < 0.05$ (black pixels). Fig. 5 shows the cases in which the set of RQA measures are (RR,DET,V_{MAX},TT) (a)–(b) and (RR,L_{MAX},LAM,V_{MAX}) (c)–(d), and Fig. 6 shows that the set of four network measures is $(\mathcal{C}, \mathcal{L}, \mathcal{C}l, \mathcal{D}_{\mathcal{T}})$. This case selection is an indirect multiple test correction: it may happen by chance with a probability of only $0.05^4 = 0.00000625$. The selected feature combinations are not necessarily the best classification sets, and the finding of better classification algorithms will be a task for further studies.

Inspection of Figs. 5(a and c) and 6(b) shows that there are only three occurrences in which the four RQA measures and 22 instances

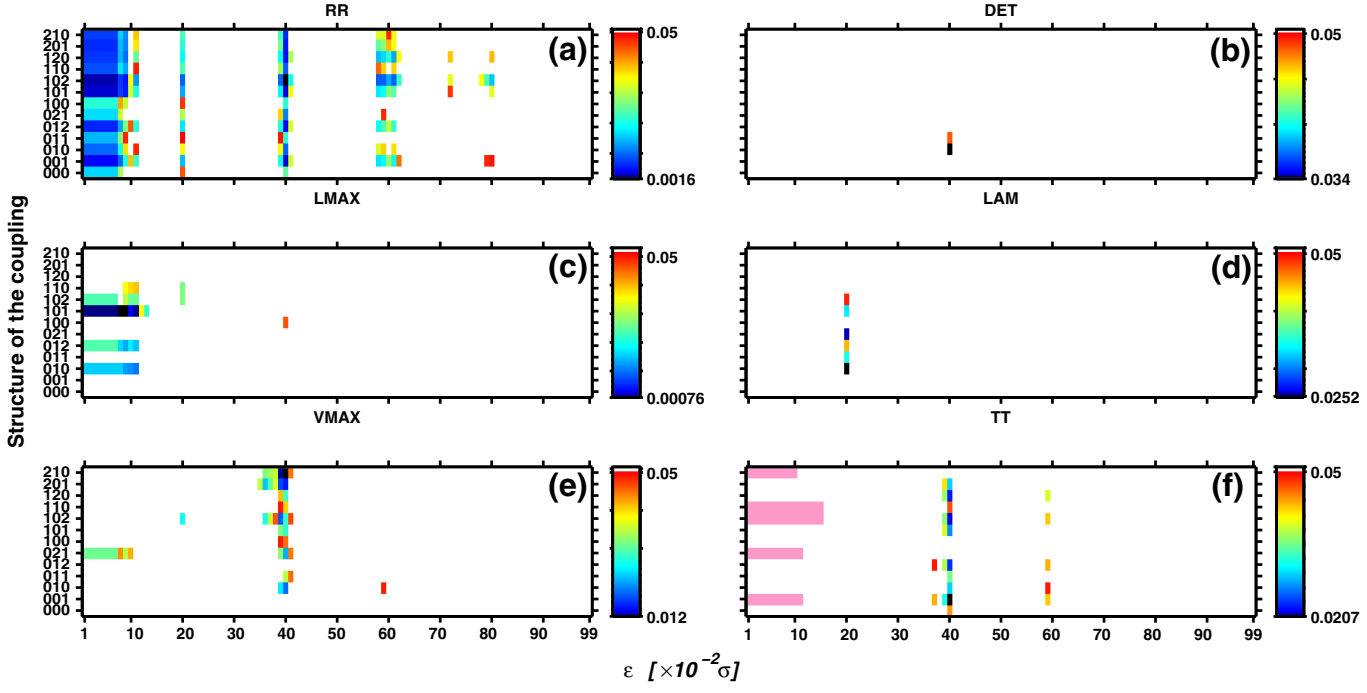

Fig. 3. (Color online) Phase plane structure of the coupling vs. **ε** showing the significance level p, computed by means of a Mann–Whitney U-test for establishing differences between the control and the PE groups and using the RP measures obtained by RQA. (a) RR, (b) DET, (c) L_{MAX}, (d) LAM, (e) V_{MAX}, and (f) TT. The color code indicates the p-values. Note that some special pixels are used such as white ($p \geq 0.05$; H_0 cannot be rejected), pink (it is not possible to compute the p-value; thus, the p-value is undetermined), and black (minimum p-value).

Please cite this article as: Ramírez Ávila, G.M., et al., Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. (2013), http://dx.doi.org/10.1016/j.autneu.2013.05.003

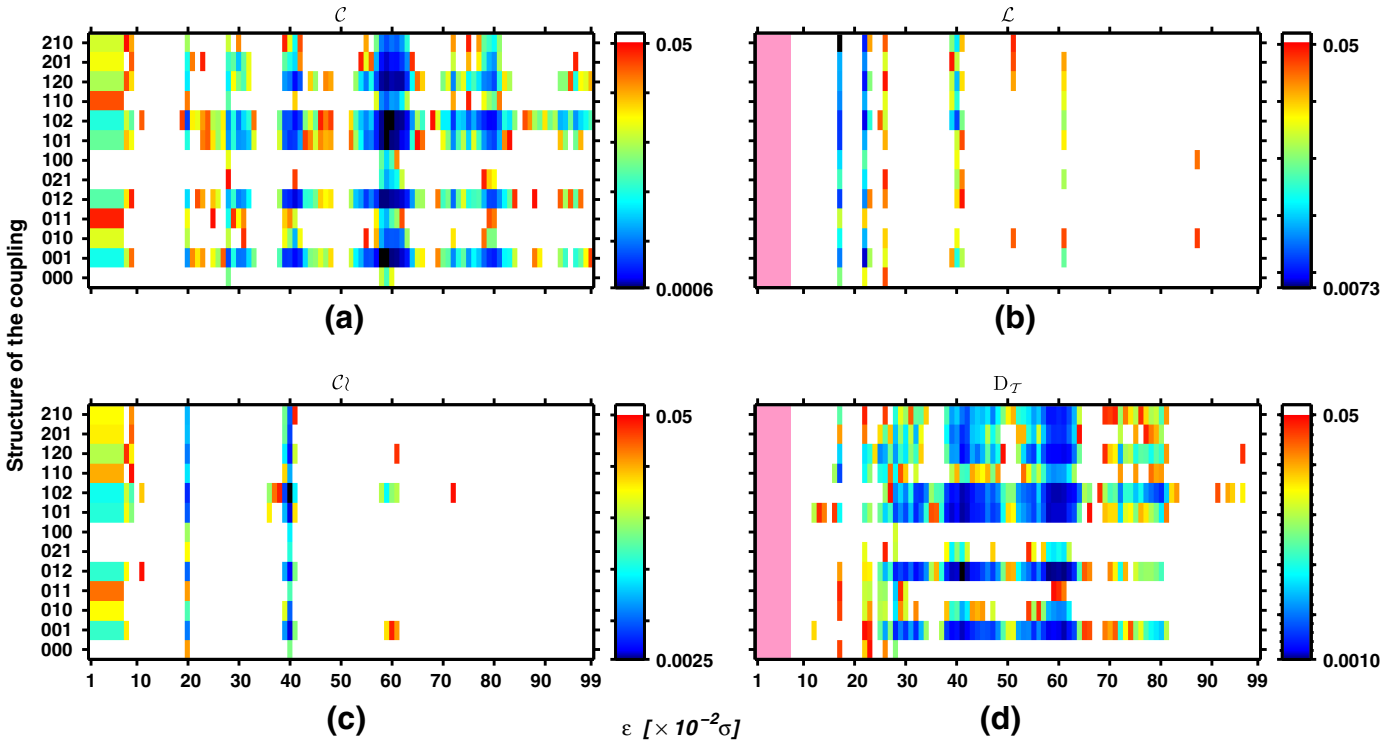

Fig. 4. (Color online) Same as Fig. 3 but using the network measures (a) $\mathcal{C}$, (b) $\mathcal{L}$, (c) $\mathcal{C}l$, and (d) $D_{\mathcal{T}}$.

in which the four network measures of each set satisfy simultaneously the statistical significance test, and we further restrict the analysis to these selected cases that do not necessarily correspond to the lower p-values. Now, considering these four measures as the parameters for the classification of control and PE groups, we perform a quadratic discriminant analysis for all the possible structures of the coupling and ε (Figs. 5(b and d) and 6(b)).

Table. 1 shows the statistical measures of the classification performance for the best selected feature sets of Figs. 5 and 6. Such measures are misclassification error rate (mer), which is the percentage of observations that are misclassified; sensitivity (se), which is the proportion of true positives that are correctly identified by the test; specificity (sp), which is the proportion of true negatives correctly identified by the test; positive predictive value (xtitPPV), which is the proportion of

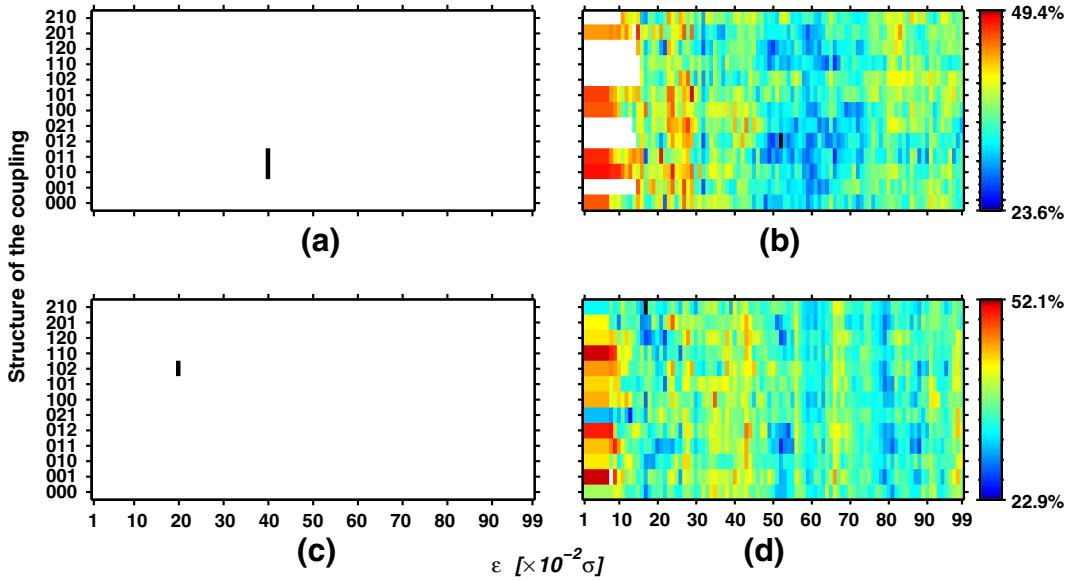

Fig. 5. (Color online) Same phase plane as in Figs. 3, showing the significant occurrences for the set (RR,DET,V_{MAX},TT) in which (a) the four considered RQA measures satisfy simultaneously the condition $p < 0.05$ (black pixels). (b) Misclassification errors (color code) in the classification of control and PE groups after a quadratic discriminant analysis for the four RQA measures. Panels (c) and (d) the same as panels (a) and (b) but for the set (RR,L_{MAX},LAM,V_{MAX}). In panels (b) and (d), the white pixels indicate that the discriminant analysis cannot be performed, and it is related to the fact that for these cases, at least one of the network measures has an undetermined p-value. The black pixel indicates the minimum value of the error.

Please cite this article as: Ramírez Ávila, G.M., et al., Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. (2013), http://dx.doi.org/10.1016/j.autneu.2013.05.003

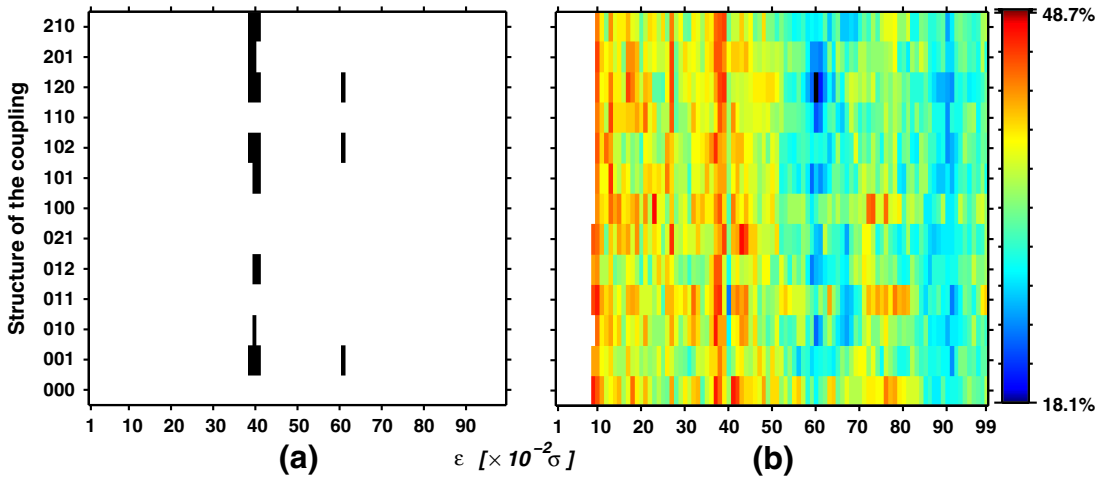

Fig. 6. (Color online) Same as in Fig. 5, but for the set of network measures $(\mathcal{C}, \mathcal{L}, \mathcal{C}l, \mathcal{D}_{\mathcal{T}})$.

patients with positive test results who are correctly diagnosed; and negative predictive value (NPV), which is the proportion of patients with negative test results who are correctly diagnosed.

From Table 1, we select the best situations in which the condition $p < 0.05$ is accomplished simultaneously by all the elements of the RQA and network sets. The best performance for the RQA analysis corresponds to the set (RR,L_{MAX},LAM,V_{MAX}), a coupling structure 102 and $\varepsilon = 0.20\sigma$, whose misclassification error is 31.2%, giving consequently the best values for the classification results, i.e., a sensitivity of 91.7%, a specificity of 45.8%, a PPV of 36.1%, and an NPV of 94.3%; whereas for the ε-recurrence networks, we have a coupling structure 120 and $\varepsilon = 0.61\sigma$, whose misclassification error is 20.1% and consequently a sensitivity of 91.7%, a specificity of 68.1%, a PPV of 48.9%, and an NPV of 96.1%.

Finally, other tools could be combined with the method used in this work such as motif distributions (Xu et al., 2008). As a glance of the latter, in Fig. 7 is shown the percentage of occurrence of subgraph size-4 motifs (computed using FANMOD, a software developed by Wernicke and Rasche, 2006) in the networks of both groups, PE and control, constructed using the medians. There is no significant difference among the percentage appearing in the networks of both groups.

4. Discussion

This work is based on recurrence methods: RQA and ε-recurrence networks, the latter is remarkable due to its novelty. The analysis of biosignals in their raw form does not give enough information to perform a suitable classification. On the contrary, when using RQA and ε-recurrence networks, the classification is possible with interesting results that allow us to validate a feasibility of these models. At a first sight, it seems that the ε-recurrence network is a tool more powerful than RQA for the classification. The latter could be related to the fact that the method based on ε-recurrence networks allows to distinguish between different dynamical regimes and also to detect corresponding dynamical transitions. In other words, the dynamical CV aspects should be better described by ε-recurrence networks. Nevertheless, further studies combining RQA, ε-recurrence networks, and detection of motifs could be important both to improve the classification results and for a better understanding of the underlying physiological phenomena involved in the CV indicators. Moreover, the combination with other methods such as those considered in Porta et al. (2009) could be also useful for achieving better classification.

The obtained results allow us to compare both recurrence methods and to realize that there are some important features that should be considered when combining both methods for classification analysis. Concerning the ε-values, the optimal values for RQA are less than those corresponding to recurrence network. In a first approach, considering the combined set $(TT, \mathcal{C}, \mathcal{L}, \mathcal{C}l, \mathcal{D}_{\mathcal{T}})$, we obtain the following statistical measures of the performance of a binary classification test when $\varepsilon = 0.60\sigma$ and the coupling 110: *mes* = 16.0 %, se = 91.7 %, sp = 76.4 %, PPV = 56.4 %, and NPV = 96.5 %. The problem with the latter results is that not all the measures of the set are significant ($p < 0.05$). Nevertheless, it is possible that by means of a simple relationship between the ε-values for RQA and the complex networks, we can improve the classification results.

The coupling among the CV indicators is an important point to study. The applied methods could give additional insights to understand the variability of the coupling between the CV indicators. In spite of the minimalist assumptions concerning the structure of the coupling, and just one value of ε in order to avoid the ambiguities stated in Donner et al. (2010a), our results give useful information for the classification and are similar to those obtained in Malberg et al. (2007).

The quantification made with the network measures is an indication of the differences between control individuals and preeclamptic patients, but of course, it is hard to estimate visually the differences in the networks structure. For that, it is mandatory to perform the

t1.1 **Table 1**
t1.2 Statistical measures of the performance of a binary classification test considering the best possible situations in which the four RQA and network measures satisfy or not simulta-
t1.3 neously the condition $p < 0.05$ (last column).

t1.4 Set	Coupling	ε [× σ]	mes [%]	se [%]	sp [%]	PPV [%]	NPV [%]	$p < 0.05$
t1.5 (RR,DET,V_{MAX},TT)	010	0.40	35.4	83.3	45.8	33.9	89.2	Yes
t1.6	012	0.52	23.6	95.8	56.9	42.6	97.6	No
t1.7 (RR,L_{MAX},LAM,V_{MAX})	102	0.20	31.2	91.7	45.8	36.1	94.3	Yes
t1.8	210	0.17	22.9	83.3	70.8	48.8	92.7	No
$(\mathcal{C}, \mathcal{L}, \mathcal{C}l, \mathcal{D}_{\mathcal{T}})$	120	0.61	20.1	91.7	68.1	48.9	96.1	Yes
t1.10	120	0.60	18.1	91.7	72.2	52.4	96.2	No

Please cite this article as: Ramírez Ávila, G.M., et al., Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. (2013), http://dx.doi.org/10.1016/j.autneu.2013.05.003

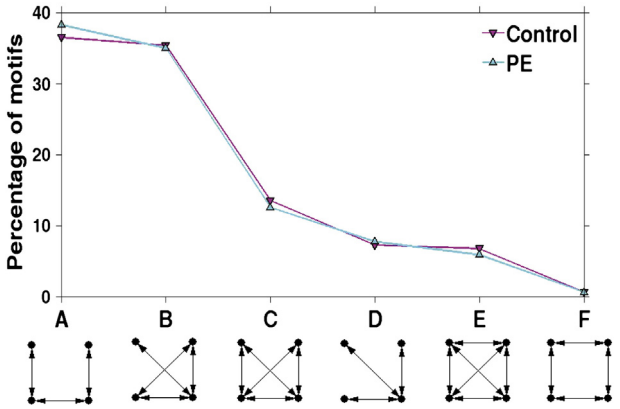

Fig. 7. (Color online) Subgraph size-4 motif distribution for both groups when using the medians to construct a network for each group.

recurrence network analysis finding the network measures which will be used later for the classification.

The advantages of the recurrence methods shown in this work lie in their easy applicability to the analysis of biosignals and offer new possibilities both in the understanding of PE pathogenesis and to envisage new therapeutic strategies.

Acknowledgment

This work has been financially supported by the German Academic Exchange Service (DAAD), the Deutsche Forschungsgemeinschaft (grant nos. KU 837/20-1 and KU-837/29-2), the Federal Ministry of Economics and Technology (grant no. FKZ KF2248001FR9), and the European projects EU NEST-pathfinder and BRACCIA.

Appendix A. Obtaining networks from time series

With the aim of illustrating the process in which the networks are obtained from the time series, we use a simple example using a normalized time series from a control individual using the coupling structure $(H(t+1),D(t+2),S(t))$ and the threshold value $\varepsilon = 0.61\sigma$ that lead to the best discrimination (see Table 1). From Eq. (1), we determine whether or not a link relies two specific points. For simplicity, we show in Fig. 8 the case in which only the first three points of the time series are considered. A sphere of diameter ε is associated to each point and according to Eq. (1), the value of $A_{i,j} = 1$ (existence of a link), only if the distance between the points i and j is less than ε, i.e., if their ε-spheres intersect. Fig. 8(a) shows that only points 1 and 2 satisfy the above-mentioned condition; thus, a link (thicker and less dark line) is established between them. On the contrary, point 3 does

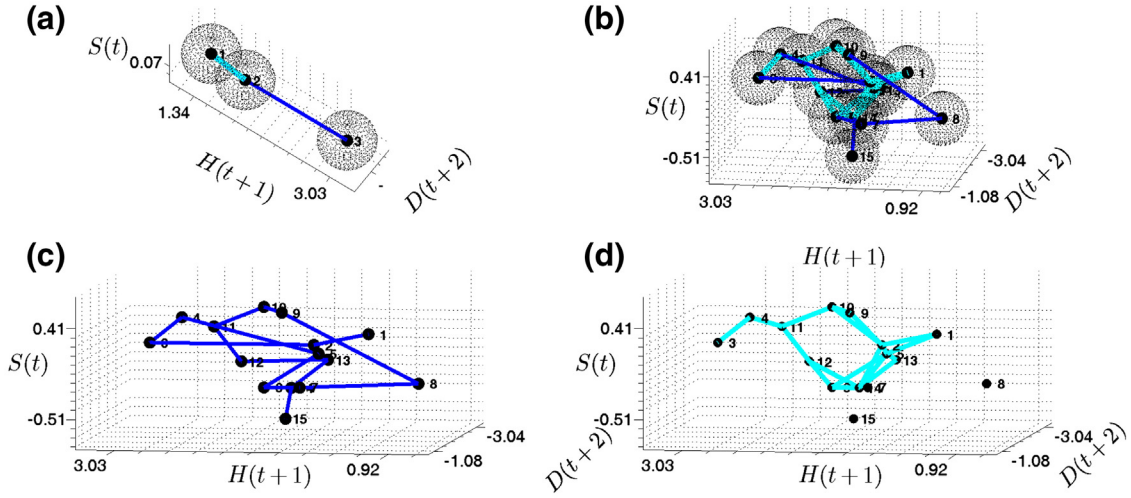

Fig. 8. (Color online) Detailed description of the **ε**-recurrence network method onto the embedding phase space $(H(t+1),D(t+2),S(t))$ using the **ε**-sphere. (a) The first and second points of the time series are intersecting their **ε**-spheres; thus, a link is associated to both points in the network structure. (b) The representation of the first 15 points, and proceeding in the same way as in panel (a), the network structure is obtained (links represented by thicker and less dark lines). (c) Time series. (d) Complex network represented into the phase space.

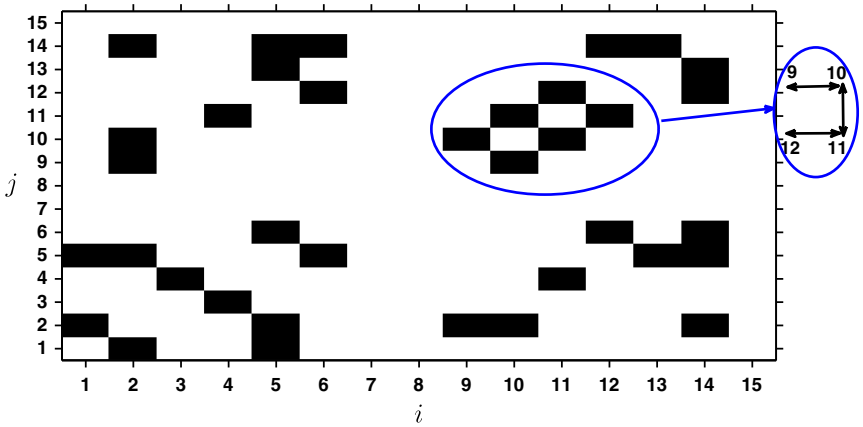

Fig. 9. (Color online) Representation of the adjacency matrix elements from which it is possible to extract the motifs. For instance, the elements 9, 10, 11, and 12 give rise to the motif shown besides the image.

Please cite this article as: Ramírez Ávila, G.M., et al., Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. (2013), http://dx.doi.org/10.1016/j.autneu.2013.05.003

not have any link with the other points, and only the time series sequence (1 → 2 → 3) is represented by the thinner and darker line. In Fig. 8(b), we extend the procedure for the first 15 points of the time series and then we extract the time series sequence 1 → 2 … → 15 and the network in Figs. 8(c) and 8(d), respectively. The representation of the adjacency matrix is possible using an image where the black and white pixels correspond to the values 1 and 0, respectively. The value $A_{i,j} = 1$ indicates the existence of a link between the nodes i and tj. Fig. 9 shows the adjacency matrix corresponding to the first 15 elements of the above-mentioned time series. As a glance of how a 4-element motif is obtained, the elements and the corresponding motif are circled in the figure. Given the consideration of only 15 elements of the time series, there are 63 subgraphs (4-element motifs).

References

Angus, D., Smith, A., Wiles, J., 2012. IEEE. Trans. Vis. Comput. Graph. 18, 988–997.

Batagelj, V., Zaveršnik, M., 2002. An O(m) Algorithm for Cores Decomposition of Networks. University of Ljubljana, Institute of Mathematics, Physics and Mechanics, Department of Mathematics.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U., 2006. Complex networks: structure and dynamics. Phys. Rep. 424 (4–5), 175–308.

Carty, D.M., Siwy, J., Brennand, J.E., Zürbig, P., Mullen, W., Franke, J., McCulloch, J.W., North, R.A., Chappell, L.C., Mischak, H., Poston, L., Dominiczak, A.F., Delles, C., 2011. Urinary proteomics for prediction of preeclampsia. Hypertension 57 (3), 561–569.

Donges, J.F., Donner, R.V., Trauth, M.H., Marwan, N., Schellnhuber, H.J., Kurths, J., 2009. Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. PNAS 108 (51), 20422–20427.

Donges, J.F., Heitzig, J., Donner, R.V., Kurths, J., 2012. Analytical framework for recurrence network analysis of time series. Phys. Rev. E 85 (4), 046105.

Donner, R.V., Zou, Y., Donges, J.F., Marwan, N., Kurths, J., 2010. Ambiguities in recurrence-based complex network representations of time series. Phys. Rev. E 81 (1), 015101.

Donner, R.V., Zou, Y., Donges, J.F., Marwan, N., Kurths, J., 2010. Recurrence networks—a novel paradigm for nonlinear time series analysis. New J. Phys. 12 (3), 033025.

Donner, R.V., Small, M., Donges, J.F., Marwan, N., Zou, Y., Xiang, R., Kurths, J., 2011. Recurrence-based time series analysis by means of complex network methods. Int. J. Bifurcation Chaos 21 (4), 1019–1046.

Hirata, Y., Aihara, K., 2012. Physica A 391, 760–766.

Malberg, H., Bauernschmitt, R., Voss, A., Walther, T., Faber, R., Stepan, H., Wessel, N., 2007. Analysis of cardiovascular oscillations: a new approach to the early prediction of pre-eclampsia. Chaos 17 (1), 015113.

Marwan, N., 2008. A historical review of recurrence plots. Eur. Phys. J. - Spec. Top. 164 (1), 3–12.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., Kurths, J., 2002. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. Phys. Rev. E 66 (2), 026702.

Marwan, N., Romano, M.C., Thiel, M., Kurths, J., 2007. Recurrence plots for the analysis of complex systems. Phys. Rep. 438 (5–6), 237–329.

Marwan, N., Donges, J.F., Zou, Y., Donner, R.V., Kurths, J., 2009. Complex network approach for recurrence analysis of time series. Phys. Lett. A 373 (46), 4246–4254.

Newman, M.E., 2003. Phys. Rev. E 67, 026126.

Ngamga, E.J., Senthilkumar, D.V., Prasad, A., Parmananda, P., Marwan, N., Kurths, J., 2012. Phys. Rev. E 85, 026217.

Ohkuchi, A., Hirashima, C., Matsubara, S., Takahashi, K., Matsuda, Y., Suzuki, M., 2011. Threshold of soluble fms-like tyrosine kinase 1/placental growth factor ratio for the imminent onset of preeclampsia. Hypertension 58 (5), 859–866.

http://pajek.imfm.si/doku.php?id=pajek [Pajek - program for large network analysis.].

Porta, A., Aletti, F., Vallais, F., Baselli, G., 2009. Multimodal signal processing for the analysis of cardiovascular variability. Philos. Trans. R. Soc. A 367 (1892), 391–409.

Rana, S., Karumanchi, S.A., Levine, R.J., Venkatesha, S., Rauh-Hain, J.A., Tamez, H., Thadhani, R., 2007. Sequential changes in antiangiogenic factors in early pregnancy and risk of developing preeclampsia. Hypertension 50 (1), 137–142.

Riedl, M., Suhrbier, A., Stepan, H., Kurths, J., Wessel, N., 2010. Short-term couplings of the cardiovascular system in pregnant women suffering from pre-eclampsia. Philos. T. Roy. Soc. A 368 (1918), 2237–2250.

Sibai, B., Dekker, G., Kupferminc, M., 2005. Pre-eclampsia. Lancet 365 (9461), 785–799.

Siddiqui, A.H., Irani, R.A., Blackwell, S.C., Ramin, S.M., Kellems, R.E., Xia, Y., 2010. Angiotensin receptor agonistic autoantibody is highly prevalent in preeclampsia. Hypertension 55 (2), 386–393.

Stepan, H., Geipel, A., Schwarz, F., Krämer, T., Wessel, N., Faber, R., 2008. Circulatory soluble endoglin and its predictive value for preeclampsia in second-trimester pregnancies with abnormal uterine perfusion. Am. J. Obstet. Gynecol. 198 (2), 175.e1–175.e6.

Suhrbier, A., Heringer, R., Walther, T., Malberg, H., Wessel, N., 2006. Comparison of three methods for beat-to-beat-interval extraction from continuous blood pressure and electrocardiogram with respect to heart rate variability analysis. Biomed. Tech. 51, 70–76.

Walther, T., Wessel, N., Malberg, H., Voss, A., Stepan, H., Faber, R., 2006. A combined technique for predicting pre-eclampsia: concurrent measurement of uterine perfusion and analysis of heart rate and blood pressure variability. J. Hypertens. 24 (4).

Wernicke, S., Rasche, F., 2006. Fanmod: a tool for fast network motif detection. Bioinformatics 22 (9), 1152–1153.

Wessel, N., Malberg, H., Bauernschmitt, R., Kurths, J., 2007. Nonlinear methods of cardiovascular physics and their clinical applicability. Int. J. Bifurcation Chaos 17 (10), 3325–3371.

Wessel, N., Suhrbier, A., Malberg, H., Bretthauer, G., Riedl, M., Marwan, N., Kurths, J., 2009. Detection of time-delayed interactions in biosignals using symbolic coupling traces. Europhys. Lett. 87, 10004.

Xu, X., Zhang, J., Small, M., 2008. Superfamily phenomena and motifs of networks induced from time series. PNAS 105 (50), 19601–19605.

Zbilut, J.P., Webber Jr., C.L., 1992. Phys. Lett. A 171, 199–203.

Please cite this article as: Ramírez Ávila, G.M., et al., Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. (2013), http://dx.doi.org/10.1016/j.autneu.2013.05.003