Are "Scaling Patterns" Useful Tools for Exploring Fractality in Heart Rate Variability Data?

ER Bojorges-Valdez, JC Echeverría, R Valdés-Cristerna, MA Peña

Electric Engineering Department. Universidad Autónoma Metropolitana, Mexico City, Mexico

Abstract

Detrended fluctuation analysis (DFA) is becoming a widely used technique for exploring the structure of correlations in heart rate variability (HRV) data. This method provides a scaling or fractal exponent α derived from the behaviour of the root-mean-square fluctuations along different time scales n. Rather than just finding a single exponent, covering either short or long range, we recently suggested to track the local evolution of α as in this way scaling patterns (SP), which seem to provide more detailed characterisations of HRV data, are revealed. Here, we evaluate such potential advantage by classifying long-term data from 50 subjects in normal sinus rhythm and 29 congestive heart failure patients. Using the SP we achieved a significantly better classification of these data than using α, thereby confirming that the SP provide a useful assessment of the correlation structure in HRV data.

1. Introduction

The heartbeat fluctuations involve power-law correlations over different time scales that break down under pathological conditions [1, 2, 3]. Thus, among time and frequency methods, detrended fluctuation analysis (DFA) is becoming a widely used technique for exploring the structure of correlations in heart rate variability (HRV) data [4]. Essentially, with DFA it is explored the behaviour of the root-mean-square fluctuations *F* along different time scales *n*. With this aim, a scaling exponent, covering either the short (α_1) or long range (α_2), is generally derived from the log-log plot of such relationship [1]. A resulting exponent of 0.5 indicates white noise and the absence of long-term correlations, a value of 1 represents the behaviour of a $1/f$ process having persistent fractal correlations, and exponents of 1.5 or above reflect a Brownian motion and the existence of non-stochastic long-range correlations. Rather than just finding a single exponent, we recently suggested to track the local evolution of α along *n* by using a recursive least squares method [5]. This reveals a structure of correlations, or scaling behaviour, along scales that we define as *scaling patterns* (SP). Interestingly, the SP seem to provide more detail than α about the differences in the dynamics of HRV data from healthy subjects and patients with heart failure [5]. For instance, we have found patterns of HRV data not always presenting a uniform power-law behaviour, and the existence of dominant characteristic scales in abnormal physiological conditions [5]. In this study, we assess such potential advantage by classifying data from presumably healthy subjects and congestive heart failure patients using the conventional α_1 and α_2 exponents, as well as the SP.

2. Methods

Long-term 24 hours HRV data of 50 adults in normal sinus rhythm (NSR) and 29 congestive heart failure (CHF) patients were gathered from the RR interval databases of PhysioBank [6], which contain information derived from ECG recordings digitised at 128 Hz. NSR subjects aged 28 to 76, and CHF patients aged 34 to 79 (NYHA classes I, II, III).

DFA was applied to obtain α_1 and α_2 exponents as follows [1]. The original RR interval or HRV series is summed by:

$$Y(k) = \sum_{i=1}^{k} [RR(i) - RR_{ave}] \qquad (1)$$

where $Y(k)$ is the k-th value of the resulting series (k=1,2,...,N), $RR(i)$ is the i-th RR interval, and RR_{ave} is the mean RR value of the entire original series of length N. Next, $Y(k)$ is divided into windows, or boxes, having equal numbers of n beats or RR intervals. The local trend Y_n is obtained in each window by a least-squared line fit and is locally subtracted from $Y(k)$ to reduce the non-stationary artifacts. The average root-mean-square fluctuation, $F(n)$, is then calculated:

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [Y(k) - Y_n(k)]^2} \qquad (2)$$

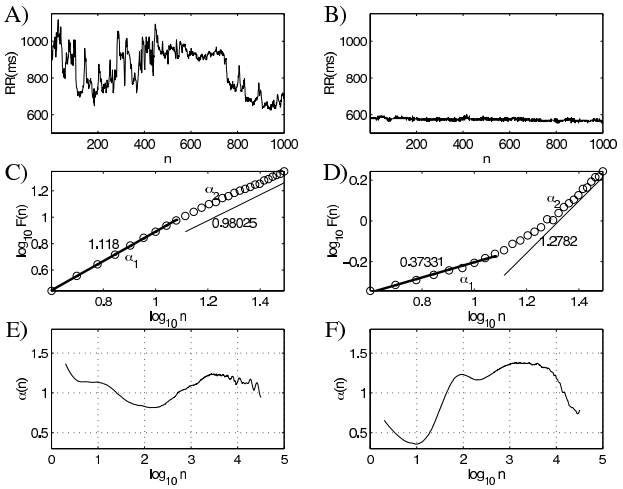

Figure 1. RR interval data from a NSR subject A) and a patient with CHF B). Whereas C) and D) show corresponding α_1 (bold) and α_2 (thin) slopes for both cases, E) and F) depict their SP.

The procedure is repeated for all boxes size or time scales. Finally, the relationship on a double-log graph between fluctuations $F(n)$ and time scales n can be approximately evaluated by a linear model $F(n) \sim n^{\alpha}$ that provides the scaling exponent α.

In this study α, was estimated by the slope of the double-log plot covering the short- (4 to 11 beats for α_1) or long-term range ($>$ 11 beats, for α_2) [1, 4].

In addition, the local evolution of α as a function of log time scales (or boxes size *n*) was also tracked by using a recursive least squares method, the $\alpha\beta$ filter. This provided gradients, $\alpha(n)$, which as specified above are referred to in this manuscript as SP, where the structure of correlations or scaling behaviour along scales can be identified distinctly [5].

Accordingly, the long-term RR data of NSR subjects and CHF patients were then classified by using α_1, α_2 and the SP ($\alpha(n)$).

Given that SP show several relevant features that are not revealed by scaling exponents (see Figure 1 and ref. [5]), different classification techniques were used. For α_x (α_1 or α_2) a threshold technique was utilised. In this scheme the values below the threshold belonged to one class and the rest to the other. The classification with SP was more complicated and Support Vector Machines (SVM) were used [7, 8]. As the classes (either NSR or CHF data) could not be separated in the original space, the values of SP were represented on a hyperspace via kernel transformations. Then, a training set was used to find a hyperplane that divided the hyperspace in two regions for the NSR and CHF classes (the separation hyperplane was determined by a set of points, called support vectors).

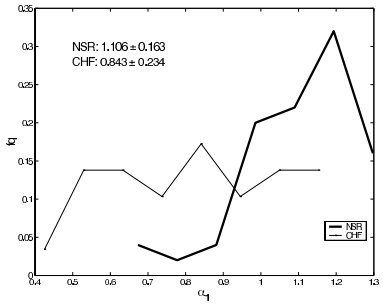

Figure 2. Histogram (fq expressed as %) of the α_1 exponent values for both classes.

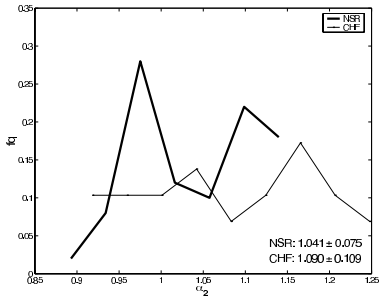

Figure 3. Histogram (fq expressed as %) of the α_2 exponent values for both classes.

These classification results (using α_x and SP) were compared via receiver operating characteristic (ROC) curves; but owing to the reduced number of samples, two resampling methods were used to obtain reliable ROC areas.

A bootstrap method [7] was applied to generate 200 sets of α_1 or α_2 exponents. Each set was classified, thereby obtaining accuracy indexes and ROC areas estimated by a Wilcoxon test. These values were finally averaged [9].

A 10-fold cross-validation method was applied to the SP classification. Thus, from the original SP two subsets were derived to train and test SVM. With this aim software gathered from internet was employed [10]. As above, since each fold provided a ROC area and accuracy index, results from the 10 folds were also averaged.

The different number of repetitions responded to the computational cost of each technique. For SVM four parameters were adjusted (kernel, cost, degree and γ), and around 300 classifications for each repetition were made to obtain the optimum local values. In this study each SP was formed with 473 equidistant points.

3. Results

Figure 1 shows typical results of the RR data, scaling exponents and SP for both NSR and CHF cases. Clearly the CHF data involve a reduced RR variability, a low value of α_1, and different SP morphology than NSR data (the α_2

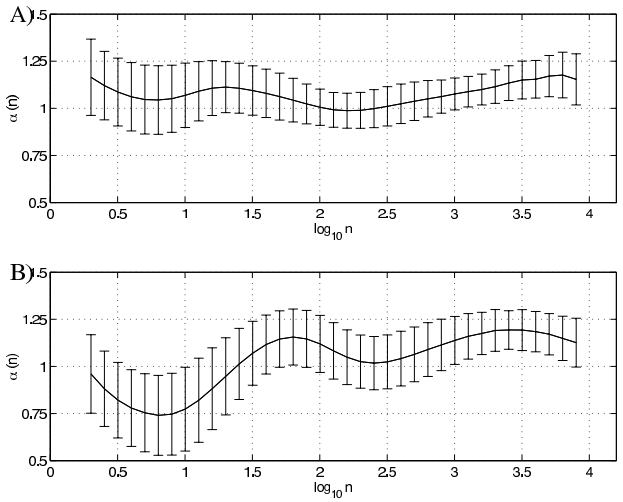

Figure 4. SP mean (± SD) from all NSR subjects A) and CHF patients B).

exponents are not perfectly adjusted along depicted points because these slopes were estimated from the whole long range ($n > 11$), and for clarity, the figure only presents a restricted range of n).

Figures 2 and 3, show the histogram of α_1 and α_2 values for all NSR and CHF cases. As the scaling exponents are overlapped in both figures, a perfect classification cannot be reached. Figure 4 shows the SP mean value (± SD) for data of both cases, which in the same way indicate that a classification with a 100% accuracy is unlikely without a transformation, because there exist scales where the SP of both conditions coincide.

ROC areas found by classifying RR data of NSR subjects and CHF patients with α_1, α_2 and SP were 0.814± 0.057, 0.650 ± 0.067, and 0.885 ± 0.049 (see Figure 5), and the accuracy values were 0.781, 0.630 and 0.854 respectively. According to these ROC areas, the SP provided significantly better classification than α_1 ($p < 10^{-2}$) and α_2 ($p < 10^{-6}$). The classification was also significantly improved by using α_1 as compared with α_2 ($p < 10^{-4}$). This result can also been appreciate in Figure 4, where the largest difference in $\alpha(n)$ between cases seems to lay in the short range. These p values were obtained by applying a Mann-Whitney test for unpaired samples, and indicate that SP combined with SVM become better classifiers than the α_x indexes. Yet the computational cost was higher because of the larger number of characteristics classified. By selecting SP formed with smaller number of points such cost can be reduced, but as this would also reduce the ROC area a further evaluation for finding a good compromise between performance and computational cost remains to be performed.

In the context of SVM, a polinomial kernel was determined as the best transformation for our data. The cost value was 1; γ value was approximately 0.002, and the transformation degree was 3. In this sense, the kernel was the most important parameter to adjust. It seems that the accuracy as a function of cost and γ had a smooth behaviour showing maximum values around 1 for both parameters.

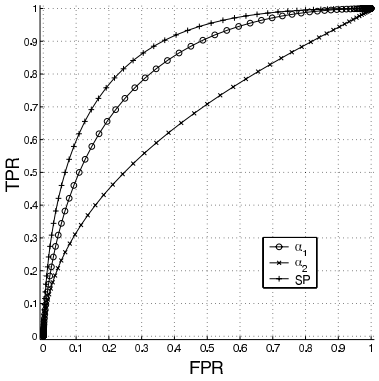

Figure 5. Figure shows the estimation of the averaged ROC curves, for all the cases a binormal model was assumed.

4. Discussion and conclusions

This study confirms that SP provide a more reliable assessment of the correlation structure in HRV data than scaling exponents, as it was revealed by achieving with them a significantly better classification of NSR and CHF data. Clearly, this was obtained by taking in consideration several features with the SP, but also by the fact that different classifiers were used.

Whilst in general the SP of NSR subjects involve a uniform power-law behaviour (*i.e.* $\alpha(n)$ closer to unity along different scales), the CHF data present lack of correlations at short-range scales, and two projections above unity at intermediate- and long-range scales (figures 1 and 4). As discussed before these CHF features are not fully reflected by the scaling exponents (see Figure 1). These features are originated from the impaired autonomic response, Chyne-Stokes type respiration, and probably from the accentuated hormonal activity that are typical of this condition [5]. Also important is to discuss that we found some of the NSR subjects presenting SP more similar to the CHF abnormal features, as well as some CHF patients not clearly presenting them. Hence, the SP could be useful to indicate early subclinical manifestations in presumably healthy subjects, or to complement evaluations, like the NYHA classification, which could be affected by the physician criterion.

The DFA technique is applied to detect structural or dynamical rather than quantitative changes in HRV data. Ac-

cordingly, the advantage of performing a detailed analysis using SP may be relevant as we have confirmed here that DFA provides potential information, which appears to be ignored by using single or coarse scaling exponents derived from linear estimations over predefined ranges (α_x). This idea has also been studied by Makowiec [11], who obtained similar conclusions using a multifractal technique.

Thus, the SP may become useful tools for exploring and evaluating the potential for HRV analysis to be used widely in clinical practice, a circumstance that, as indicated by Mahon *et al.* [12], still needs to be elucidated. Furthermore, the SP could help clarifying why spectral an scaling measures are correlated during strictly controlled condition, but their relationship is weak during free running recordings[4, 13].

Acknowledgements

First author acknowledges the financial support provided by CONACyT (Mexican Council of Science and Technology) for M Sc. studies, and John Goddard for his lectures and suggestions.

References

[1] Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos 1995;5:82–87.

[2] Peng CK, Hausdorff JM, Goldberger AL. Fractal mechanisms in neuronal control: human heartbeat and gait dynamics in health and disease. In Walleczek J (ed.), Self-organized biological dynamics and nonlinear control. Cambridge, UK: Cambridge University Press, 2000; 66–96.

[3] Goldberger AL, Peng CK, Lipsitz LA. What is physiologic complexity and how does it change with aging and disease? Neurobiology of Aging 2002;23:23–26.

[4] Pikkujamsa SM, Makikallio TH, Sourander LB, Raiha IJ, Puukka P, Skytta J, Peng CK, Goldberger AL, Huikuri HV. Cardiac interbeat interval dynamics from childhood to senescence. Comparison of conventional and new measures based on fractals and chaos theory. Circulation 1999; 100:393–399.

[5] Echeverría JC, Woolfson MS, Crowe JA, Hayes-Gill BR, Croaker GDH, Vyas H. Interpretation of heart rate variability via detrended fluctuation analysis and $\alpha\beta$ filter. Chaos 2003;13:467–475.

[6] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit and physionet. Circulation 2000;101:e215–e220.

[7] Duda RO, Hart PE, Stork DG. Pattern Classification. John Wiley and sons, 2001.

[8] Cristiani N, Shawe-Taylor. An introduction to Support Vector Machines and another kernel-based learning methods. Cambridge University Press, 200.

[9] Hanley JA, McNeil B. The meaning and use of the area under receiver operating characteristic (roc) curve. Radiology 1982;143:29–34.

[10] Chang CC, Lin CJ. LIBSVM: a library for support vector machines, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[11] Makowiec D, Galaska R, Dudkowska A, Andrzej R, Zwierz M. Long-range dependencies in heart rate signals-reviseted. Physica A statistical mechanics and its applications 2006; 369:632–644.

[12] Mahon NG, Hedman AE, Padula M, Gang Y, Savelieva I, Wakatre JEP, Malik MM, Huikuri HV, McKenna WJ. Fractal correlation properties of R-R interval dynamics in asymptomatic relatives of patients with dilated cardiomyopathy. The European Journal of Heart Failure 2002;4:151–158.

[13] Tulppo MP, Hughson RL, Makikallio TH, Airaksinen KEJ, Seppamen T, Huikuri HV. Effects of exercise and passive head-up tilt on fractal and complexity properties of heart rate dynamics. Am J Physiol Heart Circ Physiol 2001; 280:H1081–H1087.

Address for correspondence:

Name: Erik René Bojorges Valdez
Full postal address: Av. San Rafael Atlixco #186, Del. Iztapalapa, México D.F., C.P. 09340, México
E-mail address: bovedder@gmail.com