Long-Range Dependence in Heart Rate Variability Data: ARFIMA Modelling vs Detrended Fluctuation Analysis

A Leite[1−3], AP Rocha[1,3], ME Silva[1,4], S Gouveia[1,3], J Carvalho[5], O Costa[6]

[1]Departamento de Matemática Aplicada, Universidade do Porto, Porto, Portugal
[2]Departamento de Matemática, Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
[3]Centro de Matemática da Universidade do Porto, Porto, Portugal
[4]Unidade de Investigação Matemática e Aplicações, Universidade de Aveiro, Aveiro, Portugal
[5]CIAFEL, Faculdade de Desporto, Universidade do Porto, Porto, Portugal
[6]Faculdade de Medicina, Universidade do Porto, Porto, Portugal

Abstract

Heart rate variability (HRV) data display non-stationary characteristics and exhibit long-range correlation (memory). Detrended fluctuation analysis (DFA) has become a widely-used technique for long memory estimation in non-stationary HRV data. Recently, we have proposed an alternative approach based on fractional integrated autoregressive moving average (ARFIMA) models. ARFIMA models, combined with selective adaptive segmentation may be used to capture and remove long-range correlation, leading to an improved description and interpretation of the components in 24 hour HRV recordings. In this work estimation of long memory by DFA and selective adaptive ARFIMA modelling is carried out in 24 hour HRV recordings of 17 healthy subjects of two age groups. The two methods give similar information on long-range global characteristics. However, ARFIMA modelling is advantageous, allowing the description of long-range correlation in reduced length segments.

1. Introduction

Cardiovascular variables such as heart rate, arterial blood pressure and the shape of the QRS complexes in the electrocardiogram, are almost "periodical" showing some variability on a beat to beat basis. This variability reflects the interaction between perturbations to the cardiovascular variables and the corresponding response of the cardiovascular regulatory systems. Therefore, both time and frequency analysis of such variability can provide a quantitative and noninvasive method to assess the integrity of the cardiovascular system. The discrete series of successive RR intervals (the tachogram) is the simplest signal that can be used to characterize heart rate variability (HRV) and has been applied in various clinical situations [1,2].

Ambulatory long-term HRV series typically correspond to 100 000 beats in a 24-h recording and display non-stationary characteristics, exhibiting long-range correlations [3,4]. An analysis of these correlations provides a means of distinguishing between sleep and wake states [5], healthy and diseased states [3] and monitoring the effect of ageing [6]. In recent years, detrended fluctuation analysis (DFA) has become a widely-used technique for the detection of long-range correlations in non-stationary data, where conventional fluctuation analyses such as power spectra and Hurst analysis cannot be reliably used [3]. An alternative approach to long-range correlations description in HRV data, proposed by Leite *et al* [4], is to use fractional integrated autoregressive moving average (ARFIMA) models, which are an extension of the well-known autoregressive moving average (ARMA) models. ARFIMA models, combined with selective adaptive segmentation may be used to capture and remove long-range correlation, leading to an improved description and interpretation of the components in 24 hour HRV recordings [4].

In this work, DFA combined with segmentation and selective adaptive ARFIMA modelling are used in the description of the long-term correlation structure in 24 hour HRV recordings of 17 healthy subjects of two age groups.

2. Long-range dependence

A stationary process $x(t)_{t\in Z}$ is said to have long-range correlations if there exists a real number $\gamma \in]0,1[$ and a constant $c_\rho > 0$ such that

$$\rho(k) \sim c_\rho |k|^{-\gamma},\ k \to \infty,$$

where $\rho(k) = \frac{cov[x(t),x(t+k)]}{var(x(t))}$ is the autocorrelation function. Alternatively, a stationary process $x(t)_{t\in Z}$ is said to have long-range correlations if there exists a real number $\beta \in]0,1[$ and a constant $c_f > 0$ such that

$$f(\omega) \sim c_f|\omega|^{-\beta},\ \omega \to 0,$$

where $f(.)$ is the spectral density function.

In this work, two techniques are used to characterize the long-range correlations: DFA and ARFIMA modelling.

2.1. Detrended fluctuation analysis

DFA [3] has been established as an important tool for the detection of long-range correlations in non-stationary time series. The time series $x(t)$ of length N is first integrated to give

$$y(i) = \sum_{t=1}^{i}[x(t) - \bar{x}], i = 1, ..., N,$$

where $\bar{x}$ denotes the mean of the series. Next the integrated time series $y(i)$ is divided into segments of equal length k. In each segment, the local trend $y_k(i)$ is calculated by a least squares line fit. Next, the integrated time series $y(i)$ is detrended by subtracting the local trend $y_k(i)$ in each segment. The root-mean-square fluctuation of this integrated and detrended time series is given by

$$F(k) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[y(i) - y_k(i)]^2}.$$

The above computation is repeated for several segments of length k (different time scales). The relationship on a log-log graph between $F(k)$ and k can be approximately evaluated by a linear model $F(k) \sim k^{\alpha}$, where α is the scaling exponent. Values of $\alpha > 0.5$ for large values of k indicate long-range correlations in the data.

For stationary data with long-range correlations, $\gamma = 2 - 2\alpha$ and $\beta = 2\alpha - 1$.

2.2. ARFIMA approach

A class of processes with long-range correlations are the ARFIMA processes. These processes were introduced by Hosking [7] and have special interest for applications because of their capability of modelling both short- and long-term behaviour of a time series.

A stochastic process $x(t)_{t\in Z}$ is an ARFIMA(p,d,q), $p,q \in N \cup \{0\}$ and $d \in R$, if it satisfies the equation

$$\phi(B)\nabla^d x(t) = \theta(B)\epsilon(t),$$

where $\epsilon(t)_{t\in Z}$ is a Gaussian white noise WN(0,σ^2), $\phi(z) = 1-\phi_1 z-...-\phi_p z^p$ and $\theta(z) = 1-\theta_1 z-...-\theta_q z^q$ are polynomials such that $\phi(z) \neq 0$ and $\theta(z) \neq 0$ for $|z| \leq 1$, B is the backward-shift operator, $Bx(t) = x(t-1)$, ∇^d is the fractional difference operator defined by

$$\nabla^d = (1-B)^d = 1 + \sum_{j=1}^{\infty}\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}B^j,$$

and $\Gamma(.)$ is the gamma function. The parameter d determines the long-term behaviour, whereas p, q and the corresponding parameters in $\phi(B)$ and $\theta(B)$ allow the modelling of short-range properties. For $-0.5 < d < 0.5$, the ARFIMA(p,d,q) is stationary and invertible. Moreover, for $0 < d < 0.5$ the process has long-memory. ARFIMA models are adequate in HRV recordings and are used to capture and remove long-range correlations in these recordings [4].

Given a HRV series, $x(1), ..., x(N)$, the estimation of d can be obtained using the semi-parametric local Whittle estimator (LWE) [8]. Robinson [9] and Velasco [10] proved that LWE is consistent for $-0.5 < d < 1$.

For stationary data with long-range correlations, $\gamma = 1 - 2d$ and $\beta = 2d$ and d is related to α by $d = \alpha - 0.5$.

3. Results and discussion

The methodology presented above are applied to experimental 24 hour Holter HRV data of 17 healthy subjects, 7 aged 17-19 years (young) and 10 aged 65-77 years (old), obtained with a Mortara H-Scribe 12-lead ECG monitor. Sleeping and waking times were registered in the Holter diary.

To describe long-range correlations in the long-term HRV series (approximately 100 000 beats), ARFIMA modelling combined with selective adaptive segmentation is used [4]: the long record is decomposed into short records of variable length and the break points, which mark the end of consecutive short records, are determined using the AIC criterion for ARFIMA models. The short records thus obtained have a minimum length 512 and are subsequently modelled using ARFIMA models.

Long-range correlations in this long-term HRV series are also described by a global scaling exponent, obtained with DFA for $20 \leq k \leq 10000$ [3]. In order to obtain an adequate description of these correlations segmentation combined with DFA is used: the long record is decomposed into short records of constant length L. The short records are subsequently analysed by DFA for $L^{0.5} \leq k \leq \frac{L}{4}$ with $L = 4096$ beats. Using simulations of ARFIMA models, it was found that this value of L is the minimum allowable segment length for the estimation of the long-range correlations with DFA.

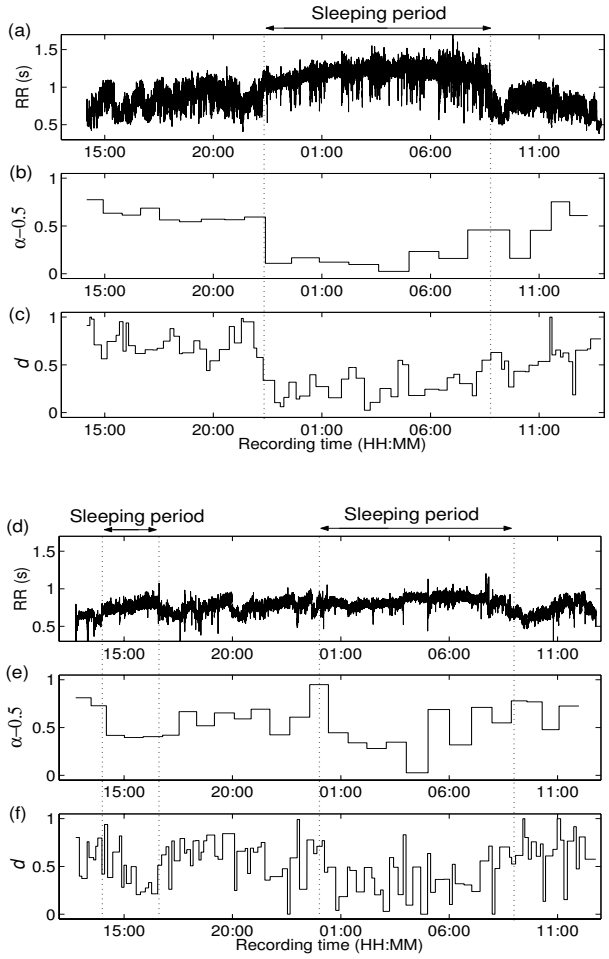

Figure 1. Tachograms of two healthy subjects, 24 hours Holter recordings, 19 years old (a) and 71 years old (d). Evolution over 24 hours of $\alpha - 0.5$ in (b) and (e) and $\hat{d}$ in (c) and (f). α is estimated using DFA combined with segmentation ($L = 4096$ beats) and $\hat{d}$ using selective adaptive segmentation combined with ARFIMA models (average of the short records is 1480 beats in (c) and 1222 beats in (f)). The dotted lines indicate the sleeping and waking times.

Figure 1 illustrates the typical results for a young (a) and an old (d) healthy subjects. The estimates $\alpha - 0.5$, in (b) and (e) and $\hat{d}$, in (c) and (f), evolve over time, presenting a circadian variation. Most of the values estimated during the sleeping period range from 0 to 0.5, whereas during the waking period they range from 0.5 to 1. For a global description of the long-range correlation structure in long-term HRV recordings, both methods contain similar information. However, selective adaptive ARFIMA estimates of long memory are based on short segments (average segments length 1480 beats in (c) and 1222 in (f) for ARFIMA vs $L = 4096$ in (b) and (e) for DFA). Therefore, the graphs indicate that ARFIMA modelling can be advantageous for a better description of long memory, namely during the transient periods, as the sleeping and waking times.

The results for the age groups of young and old are summarised in Figure 2 and Table 1. It is found that long-range correlation increases with age, both during sleeping and waking periods. This is consistent with previous results reported in literature concerning the value of global scaling exponent calculated with DFA [6]. However, selective adaptive ARFIMA estimates of long memory are based on short segments (average segments length 1243 beats for ARFIMA vs $L = 4096$ for DFA). Moreover, ARFIMA modelling allows long memory estimates from fixed length short segments of 512 beats, Table 1. The global results of this segmentation is similar to the results of the DFA combined with segmentation ($L = 4096$ beats) and selective adaptive ARFIMA modelling.

Table 1. $\alpha - 0.5$ and $\hat{d}$ for two groups of young (7 subjects) and old (10 subjects), during the 24 hour records and the sleeping and waking periods. For each case the average estimates $\pm$ standard deviations are presented.

Methods	Periods	Young	Old
DFA, $\alpha - 0.5$	24-h	0.44 ± 0.18	0.55 ± 0.19
segmentation with	Sleeping	0.31 ± 0.15	0.46 ± 0.20
$L = 4096$	Waking	0.50 ± 0.16	0.60 ± 0.17
ARFIMA , $\hat{d}$	24-h	0.45 ± 0.21	0.54 ± 0.24
select. adapt. segment.	Sleeping	0.35 ± 0.20	0.41 ± 0.22
mean(L)= 1243	Waking	0.49 ± 0.20	0.60 ± 0.22
ARFIMA, $\hat{d}$	24-h	0.47 ± 0.25	0.55 ± 0.26
segmentation with	Sleeping	0.38 ± 0.27	0.43 ± 0.27
$L = 512$	Waking	0.51 ± 0.23	0.60 ± 0.24

4. Conclusion

DFA combined with segmentation and selective adaptive ARFIMA modelling give similar information concerning the global characteristics of long-range correlations: circadian variation, with different regimes for sleeping and waking periods and increased values with age. However, selective adaptive ARFIMA estimates of long memory are based on shorter segments. Therefore, for a better description of these correlations, ARFIMA models can be advantageous, namely during the transient periods, as the sleeping and waking times. Moreover, ARFIMA modelling allows long memory estimates from fixed length short segments of 512 beats.

Acknowledgements

This work was partially supported by CMUP (financed by FCT Portugal through POCI2010/POCTI/POSI programmes, with national and CSF funds).

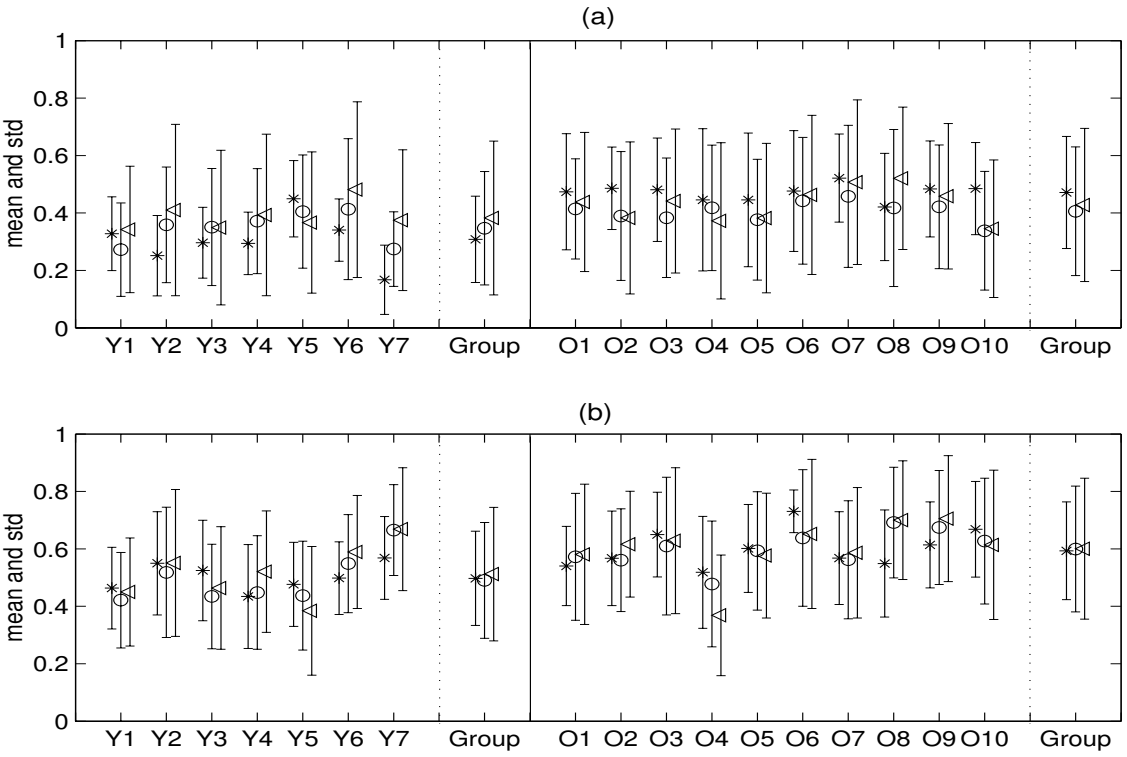

Figure 2. Average estimates and standard deviations of $\alpha - 0.5$ and $\hat{d}$ for each Holter recording during (a) sleeping period and (b) waking period. The estimated α is obtained using DFA combined with segmentation ($L = 4096$ beats, $*$) and $\hat{d}$ using selective adaptive ARFIMA modelling ($\circ$) and ARFIMA combined with segmentation ($L = 512$ beats, $\triangleleft$). The subjects are identified by Y (young) and O (old) and group estimates are presented on the right of each panel.

References

[1] Apple ML, Berger RD, Saul JP, Smith JM and Cohen RJ. Beat to beat variability in cardiovascular variables: noise or music? J Am Coll Cardiol 1989;14:1139–1148.

[2] Task Force of the European Society of Cardiology and North American Society of Pacing Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation 1996;93:1043–1065.

[3] Peng CK, Havlin S, Stanley HE and Golberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos 1995;5:82–7.

[4] Leite A, Rocha AP, Silva ME and Costa O. Modelling long-term heart rate variability: an ARFIMA approach. Biomedizinische Technik 2006;51:215–219.

[5] Ivanov PC, Bunde A, Amaral LAN, Havlin S, Fritsch-Yelle J, Baevsky RM, Stanley HE, Goldberger AL. Sleep-wake differences in scaling behavior of the human heartbeat: analysis of terrestrial and long-term space flight data. Europhys. Lett. 1999;48:594–600.

[6] Struzik ZR, Hayano J, Soma R, Kwak S, Yamamoto Y. Aging of complex heart rate dynamics. IEEE Transactions on Biomedical Engineering 2006;53:89–94.

[7] Hosking JRM. Fractional differencing. Biometrika 1981;68:165–176.

[8] Boukhan P, Oppenheim G and Taqqu MS. Theory and applications of long-range dependence. Boston: Birkhäuser 2003.

[9] Robinson PM. Gaussian semiparametric estimation of long range dependence. The Annals of Statistics 1995;23:1630–1661.

[10] Velasco C. Gaussian semiparametric estimation of nonstationary time series. Journal of Time Series Analysis 1999;20:87–127.

Address for correspondence:

Argentina Leite
Departamento de Matemática Aplicada
Faculdade de Ciências da Universidade do Porto
Rua do Campo Alegre 687, 4169-007 Porto, Portugal
E-mail address: amsleite@fc.up.pt.