Biomed Tech 2006; 51:215–219 © 2006 by Walter de Gruyter • Berlin • New York. DOI 10.1515/BMT.2006.040

# Modelling long-term heart rate variability: an ARFIMA approach

**Argentina S. Leite[1–3,*], Ana Paula Rocha[1,3], M. Eduarda Silva[1,4] and Ovídio Costa[5]**

[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] Faculdade de Medicina, Universidade do Porto, Porto, Portugal

## Abstract

Long-term heart rate variability (HRV) series can be described by time-variant autoregressive modelling. HRV recordings show dependence between distant observations that is not negligible, suggesting the existence of long-range correlations. In this work, selective adaptive segmentation combined with fractionally integrated autoregressive moving-average models is used to capture long memory in HRV recordings. This approach leads to an improved description of the low- and high-frequency components in HRV spectral analysis. Moreover, it is found that in the 24-h recording of a case report, the long-memory parameter presents a circadian variation, with different regimes for day and night periods.

**Keywords:** long-range correlations; selective adaptive segmentation; spectral analysis.

## Introduction

Cardiovascular variables such as heart rate, arterial blood pressure and the shape of the QRS complexes in the electrocardiogram are almost 'periodic', 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 cardiovascular regulatory systems. Therefore, both time and frequency analysis of such variability can provide a quantitative and non-invasive 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 characterise heart rate variability (HRV) and has been applied in various clinical situations [1, 10].

Ambulatory long-term HRV series typically correspond to 100 000 beats in a 24-h recording and exhibit non-stationary characteristics. These long-term HRV series can be described by time-variant autoregressive (AR) modelling. The AR models are called short-memory models, since their autocorrelations (ACFs) decay to zero exponentially, meaning that observations far apart are not correlated. However, the sample autocorrelations (SACFs) of HRV series show a different type of behaviour: the decay is very slow, indicating that the dependence between distant observations is not negligible. This is illustrated in Figure 1 by a tachogram for a normal subject, its SACF and the ACF of the fitted AR model. Correlations exhibiting this type of behaviour are called long-range correlations and the processes are denoted as long-memory or persistent processes [4].

This long-memory property is revealed in the frequency domain by a spectral density function that follows a power law (1/*f*) in the very low frequencies and has been the subject of several studies using detrended fluctuation analysis and coarse graining spectral analysis ([7] and references therein).

An alternative approach to long-memory description in HRV data is to use fractionally integrated autoregressive moving average (ARFIMA) models, which are an extension of the well-known autoregressive moving average (ARMA) models. ARFIMA models allow description of both the short- and long-term correlation structure and have been used in numerous application areas.

In this work, selective adaptive segmentation combined with ARFIMA models is used in the description of long-term HRV recordings.

## Materials and methods

### ARFIMA modelling of HRV

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

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

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

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

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

*Corresponding author: Argentina Soeima Leite, Departamento de Matemática Aplicada, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal
Phone: +351-220-100 869
Fax: +351-220-100 809
E-mail: amsleite@fc.up.pt

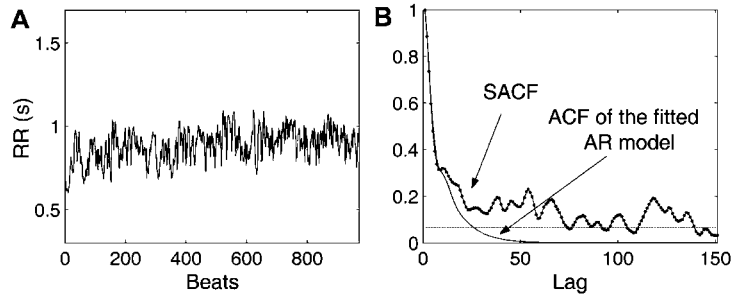


**Figure 1** (A) Tachogram of a normal subject. (B) SACF and ACF of the fitted AR model, AR(8) with order indicated by AIC criterion.

A class of processes with this property are the ARFIMA processes. These processes were introduced by Hosking [5] 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 ARFIMA($p$,$d$,$q$), $p,q\in N\cup\{0\}$ and $d\in\Re$, if it satisfies the equation

$$\phi(B)\nabla^d X_t=\theta(B)\varepsilon_t,$$

where $\{\varepsilon_t\}_{t\in Z}$ is Gaussian white noise WN(0,$\sigma^2$), $\phi(z)=1-\sum_{i=1}^{p}\phi_i z^i$ and $\theta(z)=1-\sum_{j=1}^{q}\theta_j z^j$ are polynomials such that $\phi(z)\neq 0$ and $\theta(z)\neq 0$ for $|z|\leq 1$, $B$ is the backward-shift operator, $BX_t=X_{t-1}$, $\nabla^d$ 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 $\Gamma(\cdot)$ 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 $-0.5<d<0.5$, ARFIMA($p$,$d$,$q$) is stationary and invertible. Moreover, for $0<d<0.5$, the process has long memory and for $d=0$ ARFIMA($p$,0,$q$) reduces to the usual short-memory ARMA($p$,$q$) model. In this work, we consider ARFIMA($p$,$d$,0) models, since they are a natural extension of the classic AR($p$) models. The spectral density function of $\{X_t\}$ is then given by:

$$f(\omega)=f_{SM}(\omega)\left|1-e^{-i\omega}\right|^{-2d},\quad -\pi\leq\omega\leq\pi \qquad (1)$$

where $f_{SM}(\omega)=\dfrac{\sigma^2}{2\pi\left|\phi(e^{-i\omega})\right|^2}$ is the spectral density of the corresponding short memory, ARFIMA($p$,0,0) [AR($p$)] process.

Given a HRV series, $X_1,\ldots,X_n$, estimation of the parameters of the ARFIMA($p$,$d$,0) models is as follows [3, 4]: estimate $d$ using the semi-parametric local Whittle estimator (LWE) and estimate the AR parameters in the filtered data $U_t=\nabla^d X_t$. Robinson [8] and Velasco [11] proved that the LWE is consistent for $-0.5<d<1$. The AR parameters are estimated from the Yule-Walker equa-

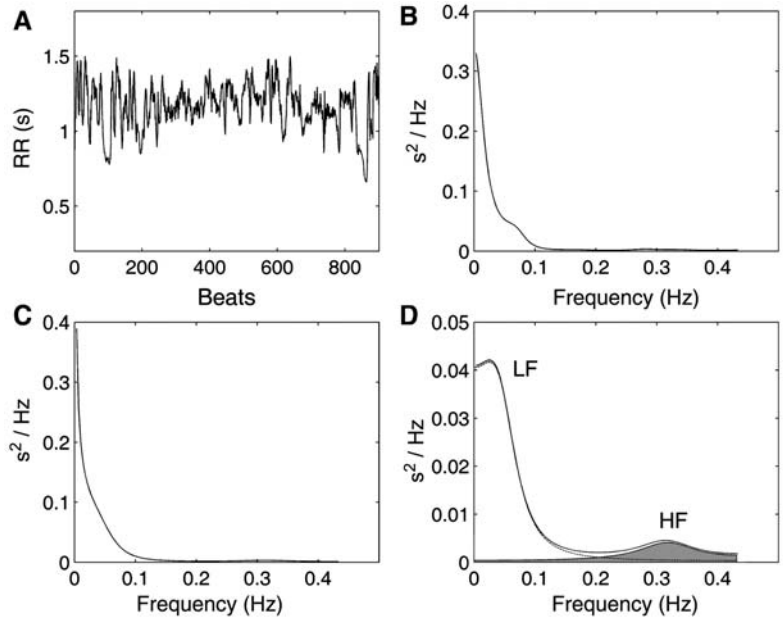


**Figure 2** (A) Tachogram of a normal subject during a sleeping period. (B) Spectrum using AR modelling. Spectrum using (C) ARFIMA modelling and (D) the corresponding $f_{SM}$ spectrum.
Panel (D) also shows the HF component resulting from the automatic decomposition of $f_{SM}$. The frequency axis is normalised by the mean RR.

**Table 1** Parameter estimates for AR($p$) and ARFIMA($p,d,0$) models of the tachogram in Figure 2A.

| | $\hat{d}$ | $p$ | $\hat{\phi}_1$ | $\hat{\phi}_2$ | $\hat{\phi}_3$ | $\hat{\phi}_4$ | $\hat{\phi}_5$ | $\hat{\phi}_6$ | $\hat{\phi}_7$ | $\hat{\phi}_8$ | $\hat{\phi}_9$ | $\hat{\phi}_{10}$ | $\hat{\phi}_{11}$ | $\hat{\sigma}^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AR | | 11 | -0.80 | -0.22 | -0.07 | 0.29 | 0.01 | -0.07 | 0.06 | -0.02 | -0.06 | -0.07 | 0.08 | 0.01 |
| ARFIMA | 0.30 | 4 | -0.51 | -0.25 | -0.13 | 0.24 | | | | | | | | 0.01 |

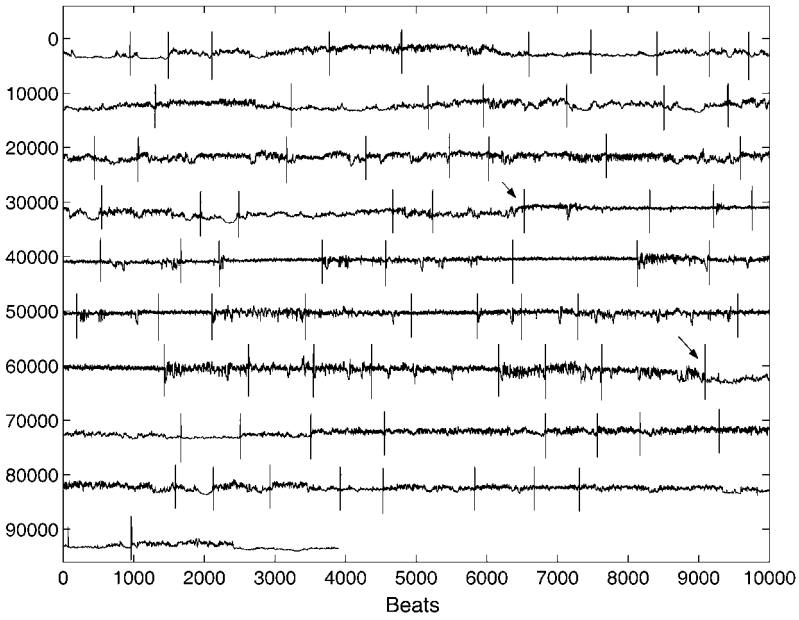


**Figure 3** Tachogram of a normal subject.
Holter recording (24 h), with localisation of the break points obtained using ARFIMA modelling combined with selective adaptive segmentation. The arrows indicate the sleeping and waking times.

tions and the Levinson algorithm, with the order $p$ determined by the Akaike information criterion (AIC) [3].

To illustrate the use of ARFIMA compared to AR models in HRV, the tachogram represented in Figure 2A was modelled using both methods. The results are summarised in Table 1. Note that the estimated value for $d$, $\hat{d}=0.30$ indicates that the record has long memory. Moreover, in ARFIMA(4,0.30,0), all the parameters are statistically significant, whereas in AR(11) there are only four statistically significant parameters, namely $\hat{\phi}_1$, $\hat{\phi}_2$, $\hat{\phi}_4$ and $\hat{\phi}_{11}$. The corresponding spectra, represented in Figure 2B,C, contain similar global information. However, the spectrum obtained after removing the long-range correlations, $f_{SM}$, shows more clearly the components associated with the low (LF, 0.04 to 0.15 Hz) and high (HF, 0.15 to 0.4 Hz) frequencies (Figure 2D). Furthermore, the usual automatic decomposition may be applied to the $f_{SM}$ spectrum, which is an AR spectrum, as illustrated in Figure 2D for the HF component.

These results indicate that ARFIMA models are adequate in HRV, allowing more parsimonious modelling and a better description of the spectral components. Similar results were obtained in other recordings.

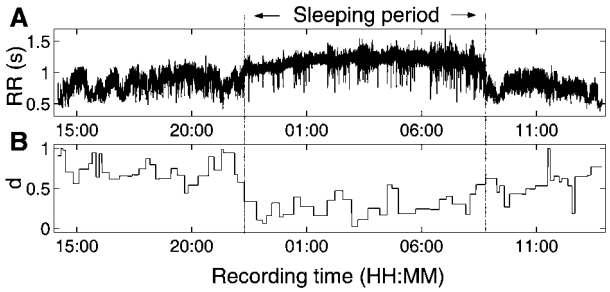


**Figure 4** (A) Tachogram of Figure 3 and (B) evolution over 24 h of $\hat{d}$.
The dotted lines indicate the sleeping (22:20 h) and waking (08:45 h) times.

### ARFIMA modelling of long-term data

Long-term HRV series can be described by AR modelling combined with adaptive segmentation [6]: the long record is decomposed into short records of variable length $L_i$ and the break points, which mark the end of consecutive shorts records, are determined using the AIC criterion for AR models.

In view of the long-memory properties of the data, the adaptive segmentation here is based on ARFIMA modelling. The break points are obtained using the AIC criterion for ARFIMA($p,d,0$) models [2]:

$$\mathrm{AIC}(p)=n\log\hat{\sigma}^2+2(p+2).$$

Specifically, let $N$ be the minimum allowable segment length and $L_i$ initialised with $N$ samples, $L_i \geq N$. Each segment $S_i$ starts at the last break point determined and has length $L_i+N$. A new break point is selected when $D_{aic}=\mathrm{AIC}_0\text{-}\mathrm{AIC}_1 \geq 0$, where $\mathrm{AIC}_0$ is the value of the AIC in

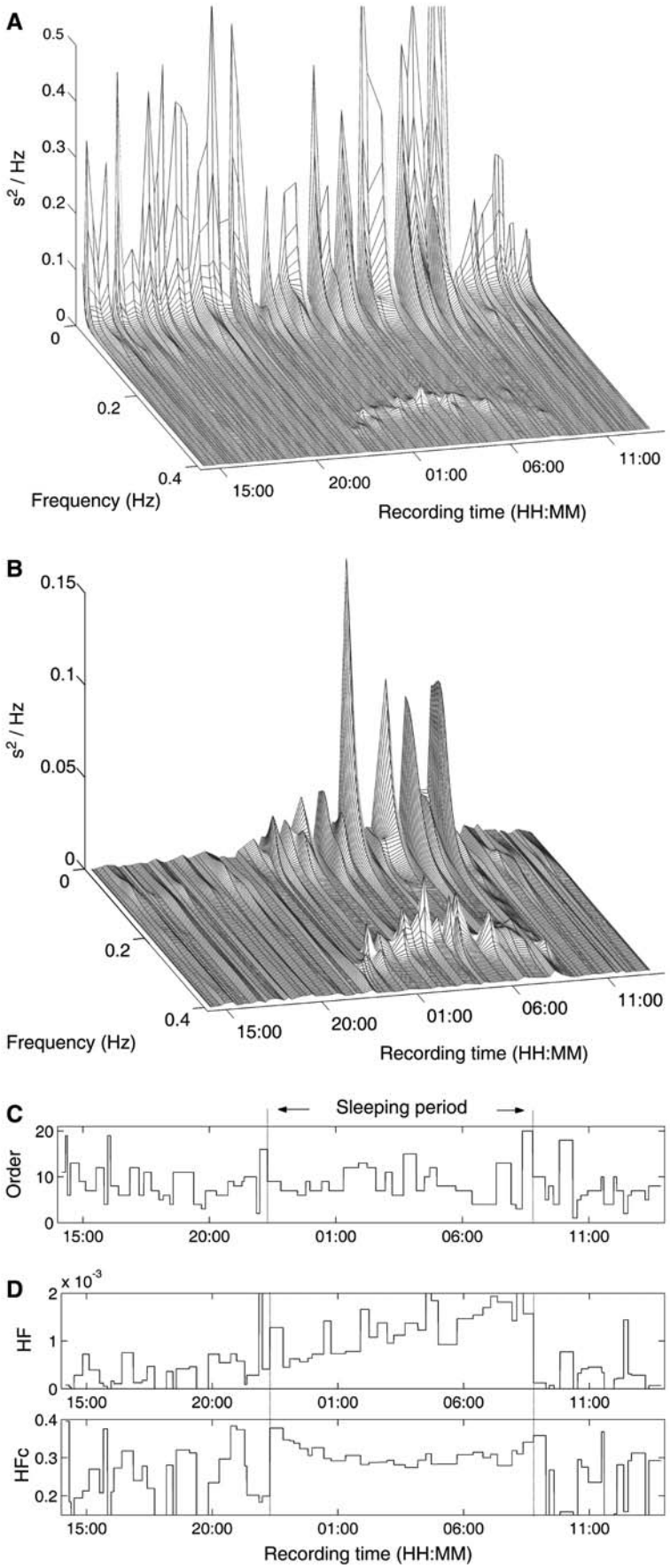


**Figure 5** Three-dimensional spectra (pseudo-spectra where appropriate) of the long-term HRV series in Figure 3: (A) $f(\cdot)$ and (B) $f_{SM}(\cdot)$ for the selected AIC model orders in (C). (D) Evolution over time of the HF and HFc parameters determined automatically from $f_{SM}(\cdot)$.
Spectra were obtained with interpolation and resampling of the individual spectra to ensure correct matching of the frequency axis.

segment $S_i$ and $AIC_1$ is the sum of AIC in the first $L_i$ samples with AIC in the last $N$ samples of the segment $S_i$. The values of the AIC are obtained considering a fixed value for $p$.

In this methodology, when a break point is identified, $N$ samples are necessary to start the search for the next break point. Therefore, no break points are identified in segments of length $N$. This limitation may be overcome by selecting the most significant break points (maximum $D_{aic}$) from a set of candidates in the next $N$ samples.

This segmentation is referred to as selective adaptive segmentation. The short records thus obtained are subsequently modelled using ARFIMA models. The selective segmentation leads to segments of increased length and consequently to a better estimation of the memory parameter, $d$.

## Results and discussion

The methodology presented above is illustrated in a case report. The data are an experimental 24-h Holter recording of a 19-year-old normal subject, obtained with a Mortara H-Scribe 12-lead ECG monitor (Mortara Instrument Inc., Milwaukee, WI, USA). The sleeping and waking times were registered by the subject in the Holter diary. Figure 3 shows the break points thus obtained (marked with dotted vertical lines) with $N=512$ and $p=12$ (a value reported in the literature [9]). ARFIMA modelling combined with selective adaptive segmentation leads to adequate localisation of rhythm differences and transient periods, as the sleeping and waking times, marked with arrows.

As illustrated in Figure 4B, $\hat{d}$ evolves over time, presenting a circadian variation. Most of the values estimated during the night (sleeping period) range from 0 to 0.5, whereas during the day (waking period) they range from 0.5 to 1. Hence, during the day the majority the HRV recordings are non-stationary and Equation (1) is then said to be a pseudo-spectral density [11].

Figure 5A represents the three-dimensional parametric spectra (pseudo-spectra where appropriate), $f(.)$, of the long-term HRV series and Figure 5B represents the corresponding parametric $f_{SM}(\cdot)$, computed from AR models with orders selected by AIC and represented in Figure 5C. During the night, a significant HF component is observed, corresponding to dominance of the parasympathetic nervous system activity and to the regularity of respiratory activity. However, these features are emphasised in $f_{SM}(\cdot)$ spectra (Figure 5B). Furthermore, the use of $f_{SM}(\cdot)$ allows the application of current approaches to the analysis of HRV. Specifically, Figure 5D represents the evolution over time of the HF component and its typical central frequency, HFc, determined automatically from the parametric modelling. Similar results were obtained from the analysis of other Holter recordings of normal subjects.

## Conclusion

HRV recordings are long-memory or persistent processes, i.e., they show spectra that follow a power law in the very low frequencies, hindering the analysis of LF and HF frequencies. ARFIMA models are used to capture and remove long memory in HRV recordings, leading to an improved description of low- and high-frequency components in the spectra. ARFIMA modelling combined with selective adaptive segmentation of 24-h HRV data shows that the memory parameter has a circadian variation with two regimes: sleeping and waking periods. The ARFIMA approach described in this work allows the use of current techniques in the analysis of HRV.

## References

[1] Appel ML, Berger RD, Saul JP, Smith JM, Cohen RJ. Beat to beat variability in cardiovascular variables: noise or music? J Am Coll Cardiol 1989; 14: 1139–1148.
[2] Beran J, Bhansali RJ, Ocker D. On unified model selection for stationary and nonstationary short and long memory autoregressive processes. Biometrika 1998; 85: 921–934.
[3] Brockwell PJ, Davis RA. Time series: theory and methods. New York: Springer-Verlag 1991.
[4] Doukhan P, Oppenheim G, Taqqu MS. Theory and applications of long-range dependence. Boston: Birkhäuser 2003.
[5] Hosking JRM. Fractional differencing. Biometrika 1981; 68: 165-176.
[6] Niedzwiecki M. Identification of time-varying processes. Somerset, NJ: John Wiley 2000.
[7] Rienzo MD, Mancia G, Parati G, Pedotti A, Zanchetti A. Frontiers of blood pressure and heart rate analysis. Amsterdam: IOS press 1997.
[8] Robinson PM. Gaussian semiparametric estimation of long range dependence. The Annals of Statistics 1995; 23: 1630-1661.
[9] Rocha AP, Leite A, Gouveia S, Lago P, Costa O, Freitas F. Spectral characterization of long-term ambulatory heart rate variability signals. In: Digest of the 5th IMA Conference on Mathematics in Signal Processing, Warwick, UK, 2000. Essex: IMA ed.: 1–4.
[10] Task Force of the ESC and NASPE. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Eur Heart J 1996; 17: 354–381.
[11] Velasco C. Gaussian semiparametric estimation of non-stationary time series. Journal of Time Series Analysis 1999; 20: 87–127.