Beyond long memory in heart rate variability: An approach based on fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity

Argentina Leite, Ana Paula Rocha, and Maria Eduarda Silva

Citation: Chaos **23**, 023103 (2013); doi: 10.1063/1.4802035
View online: http://dx.doi.org/10.1063/1.4802035
View Table of Contents: http://chaos.aip.org/resource/1/CHAOEH/v23/i2
Published by the American Institute of Physics.

Beyond long memory in heart rate variability: An approach based on fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity

Argentina Leite,[1] Ana Paula Rocha,[2] and Maria Eduarda Silva[3]
[1]*Departamento de Matemática, Escola de Ciências e Tecnologia, Universidade de Trás-os-Montes e Alto Douro and CM-UTAD, Portugal*
[2]*Departamento de Matemática, Faculdade de Ciências, Universidade do Porto and CMUP, Portugal*
[3]*Faculdade de Economia, Universidade do Porto and CIDMA, Portugal*

(Received 23 April 2012; accepted 3 April 2013; published online 19 April 2013)

Heart Rate Variability (HRV) series exhibit long memory and time-varying conditional variance. This work considers the Fractionally Integrated AutoRegressive Moving Average (ARFIMA) models with Generalized AutoRegressive Conditional Heteroscedastic (GARCH) errors. ARFIMA-GARCH models may be used to capture and remove long memory and estimate the conditional volatility in 24 h HRV recordings. The ARFIMA-GARCH approach is applied to fifteen long term HRV series available at Physionet, leading to the discrimination among normal individuals, heart failure patients, and patients with atrial fibrillation. © *2013 AIP Publishing LLC* [http://dx.doi.org/10.1063/1.4802035]

Heart Rate Variability (HRV) data display non-stationary characteristics and exhibit long-range dependence in the mean. Non-parametric methodologies such as detrended fluctuation analysis have been widely used for the detection and estimation of long memory in HRV. Recently, Fractionally Integrated AutoRegressive Moving Average (ARFIMA) models have been proposed as a parametric alternative in this context. Another characteristic of HRV recordings is conditional heteroscedasticity (time-varying conditional variance), traditionally estimated by recursive least squares. In this work, an alternative approach based on ARFIMA models with Generalized AutoRegressive Conditionally Heteroscedastic (GARCH) innovations is proposed and applied to 15 long term HRV series available at Physionet.

With respect to the long range dependence, our study agrees with previous studies. However, this study also assesses persistent conditional volatility in HRV records, via the parametric ARFIMA-GARCH modeling. Thus, further characterization of the data is accomplished, indicating important differences among the volatility characteristics of the three groups: normal subjects (N) and patients with heart failure (C) present volatility whereas for patients with atrial fibrillation (A) it is reduced. Moreover, the long memory parameters in mean and volatility exhibit circadian variation, with higher values of the memory in the mean and lowest values of the memory in the conditional variance, during the day periods. The results abide the test of the surrogate data employed.

In summary, HRV series of N, C, and A subjects may be characterized by a time-varying fractal stochastic process in both the mean and the volatility.

I. INTRODUCTION

The characterization of the dynamics of a system has become an important and interdisciplinary problem, namely, in Biomedical applications. Cardiovascular variables such as heart rate, arterial blood pressure, and the shape of the QRS complexes in the electrocardiogram show variability on a beat to beat basis, as a response of the cardiovascular regulatory systems. Therefore, the 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 in the electrocardiogram (the tachogram) is the simplest signal that can be used to characterize HRV and has been applied in various clinical situations.[1] The analysis of ambulatory long term HRV series has become important for clinical diagnosis and risk assessment. These series correspond typically to 100 000 beats in 24 h recordings and exhibit non stationary characteristics with outliers, missing values, change points, and variability within as well as among individuals.[2]

A first approach currently used to describe non stationary long HRV series is time-variant AutoRegressive (AR) analysis using exponentially smoothed recursive least squares estimation, with fixed and varying forgetting factors, leading to the estimation of both the conditional mean and the conditional variance.[3–5] Another approach is based on the segmentation of the long record into short, approximately stationary records.[6] The segments are then usually modeled with short memory AR models.[2,4] This procedure leads to the achievement of parametric models for the conditional mean. However, it is well known that HRV recordings present long memory or persistence characteristics common in data arising from natural phenomena. In fact, Kobayashi and Musha[7] using 10 h recordings of healthy subjects found that the spectral density function obeys a power law ($1/f^{\alpha}$, $\alpha \approx 1$) in the very low frequencies. This preliminary observation was confirmed by Saul *et al.*,[8] using a large set of 24 h records. Goldberger *et al.*[9] and Goldberger and West[10] also reported long term variations in HRV records similar to those observed in long memory stochastic

processes such as fractional Gaussian noise or fractional Brownian motion. Later, non parametric approaches such as Detrended Fluctuation Analysis[11] and coarse-graining spectral analysis[12] have been applied to study the persistence in HRV. For a revision, see Cerutti *et al.*[13]

An alternative approach to long memory description in HRV data relying in time series analysis techniques was proposed by Leite *et al.*[14] using ARFIMA models. ARFIMA models, introduced by Hosking,[15] are an extension of the well-known AutoRegressive Moving Average (ARMA) models. They have special interest for applications because of their capability of modeling both short and long term behaviours of a time series. The authors Leite *et al.*[14] use segmentation of 24-h recordings of HRV to find that the long memory parameter (or scaling exponent) changes with time. These findings were corroborated later by Baillie *et al.*[16] In this work, this characteristic is designated by time-varying fractal in the sense that the memory parameter changes over time and between regimes (day and night).

ARFIMA and in particular AR, models are models for the conditional mean since they describe the changes in mean. However, HRV exhibits also changes in variance over time, with periods of large variability followed by periods of stability, suggesting heteroscedastic conditional variance. The conditional standard deviation is usually designated by volatility. These volatility clustering phenomena may be well described by conditional volatility models such as the GARCH models proposed by Bollerslev.[17] The GARCH models were originally proposed to model returns in financial time series but have lately found application in other areas. In fact, in 2006, Hu and Tsoukalas[18] used GARCH models to develop a HRV based apnea screening tool.

This work considers the possibility that HRV series is a long-memory process with time dependent conditional heteroscedasticity. To model data with such characteristics fractionally integrated ARFIMA models with GARCH innovations, ARFIMA-GARCH models, are appropriate. These models are an extension of the ARFIMA models and have been applied in economic and financial series, namely, by Baillie and Chung,[19] Ling and Li,[20] Ling,[21] and Vougas.[22] Their main advantage is that ARFIMA-GARCH models may provide a useful way of analysing the relationship between the conditional mean and variance of a process exhibiting long memory and time varying volatility.

The ARFIMA-GARCH approach is applied to fifteen 24-h HRV recordings provided by PhysioNet:[23] five from normal subjects, five from heart failure, and five from atrial fibrillation patients (http://www.physionet.org/challenge/chaos/). The small number of cases analysed and a reduced number of recordings in some segments, only 512 points in some cases, deems necessary testing the results for spurious long memory estimation through a comparison with surrogate data. The method of surrogate data is well described in the statistical hypothesis testing framework. A so called null hypothesis is put forward containing a statement about the population: in this case that HVR is short-memory. Then a discriminating statistic which quantifies the characteristic under study in the time series is chosen: here will be the ARFIMA-GARCH model parameter d. If this statistic obtained from the data is different from that expected under the null hypothesis, this hypothesis is rejected for the data. In the method of surrogate data, the distribution of the statistic is obtained by direct Monte Carlo simulation: an ensemble of surrogate data set is generated which share all properties of the observed time series except long memory (the one under the null hypothesis).[24]

The remaining of the paper is organized as follows. The next section describes the ARFIMA-GARCH model and some of its properties, estimation procedures, and test statistics for checking model adequacy. Section III illustrates the modeling of a short HRV segment and introduces the methodology to be used for 24-h HRV data. Section IV describes the data base and presents the results of its analysis enhanced with a discussion of the incorporation of a segmentation procedure for the long records. Finally, Sec. V concludes the paper with a discussion of the obtained results.

II. ARFIMA-GARCH MODEL

The most general model considered in this work is the ARFIMA(p, d, q)-GARCH(P, Q) process defined by the following equations:

$$\phi(B)(1-B)^d x_t = \theta(B)\epsilon_t, \tag{1}$$

$$\epsilon_t = \sigma_t z_t, \quad \sigma_t^2 = u_0 + \sum_{i=1}^{P} v_i\, \sigma_{t-i}^2 + \sum_{i=1}^{Q} u_i\, \epsilon_{t-i}^2, \tag{2}$$

where B is the backward-shift operator, $(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k}(-1)^k B^k$ is the fractional difference operator,[25] d is a real number, $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, and $\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$ are polynomials in B, $u_0 > 0, v_1, \ldots, v_P$, $u_1, \ldots, u_Q \geq 0$, $p, q, P, Q \in \mathbb{N}_0$ and z_t are independent and identically distributed random variables with zero mean and unit variance.

Equation (1) describes the conditional mean of the process with serially uncorrelated residuals, whereas Eq. (2) describes the conditional variance of the process which varies over time instead of being constant, as in traditional time series models.

In Eq. (1), the parameter d determines the long-term behaviour in the mean, whereas p, q and the coefficients in $\phi(B)$ and $\theta(B)$ allow for the modeling of short-range properties in the mean. In Eq. (2), the conditional variance σ_t^2 is modelled as dependent on its own lagged values and on the squared residuals of the mean equation. The parameters u_i characterize the short-range properties in the volatility and the parameters v_j characterize the persistence in the volatility.

For $P = Q = 0$, ARFIMA(p, d, q)-GARCH(P, Q) reduces to the ARFIMA(p, d, q) with constant variance, $\sigma_\epsilon^2 = u_0$. Additionally for $d = q = 0$, the model reduces to the classic short-memory AR(p) model. The population characteristics of ARFIMA processes have been extensively studied by Beran[25] and Baillie.[26] For $-0.5 < d < 0.5$ and all roots of $\phi(B)$ and $\theta(B)$ lying outside the unit circle, the

process in Eq. (1) is covariance stationary but the autocorrelation decays at a slow hyperbolic rate compared with exponential rate of decay of the stationary and invertible ARMA process. In the range $-0.5 < d < 0.5$, the long memory parameter is related to the Hurst coefficient,[16] H, to the fractal dimension,[13] D, and to the slope of the (generalized) spectral density in the low frequency range,[15] α, by $d = H - 0.5$, $H = 2 - D$ and $\alpha = 2d$, respectively. Moreover, for $0.5 \leq d < 1$, the process is non-stationary and mean reverting.

The GARCH(P, Q) model (2) is second order stationary if $\sum_{i=1}^{P} v_i + \sum_{j=1}^{Q} u_j < 1$, see Bollerslev.[17] Therefore, the ARFIMA(p, d, q)-GARCH(P, Q) is stationary if $-0.5 < d < 0.5$, all the roots of $\phi(B)$ and $\theta(B)$ lie outside the unit circle and $\sum_{i=1}^{P} v_i + \sum_{j=1}^{Q} u_j < 1$. Furthermore,

$$\sigma_\epsilon^2 = \mathrm{var}(\epsilon_t) = \frac{u_0}{1 - \sum_{i=1}^{P} v_i - \sum_{j=1}^{Q} u_j}. \tag{3}$$

In this work, we consider ARFIMA(p, d, 0)-GARCH(1, 1) models, since they are a natural extension of the classic AR(p) models usual in the analysis of HRV and give special attention to the parameters: d which characterizes the long memory in the mean and u_1 and v_1 which characterize the short-range and long-range properties in the conditional variance.

The spectral density function of a stationary ARFIMA (p, d, 0) process[15] is given by

$$f_\omega = f_\omega^* |1 - e^{-i\omega}|^{-2d}, \quad -\pi \leq \omega \leq \pi, \tag{4}$$

$$f_\omega^* = \frac{\sigma_\epsilon^2}{|\phi(e^{-i\omega})|^2} \tag{5}$$

with $\sigma_\epsilon^2 = u_0$, where f_ω^* is the spectral density of the AR(p) process. Since the autocorrelation function (ACF) of a GARCH(1,1) process is the same as that of white noise,[17] the spectral density function of a stationary ARFIMA($p, d, 0$)-GARCH($1, 1$) process is given by Eqs. (4) and (5) with σ_ϵ^2 defined by Eq. (3), where f_ω^* is the spectral density of the AR(p)-GARCH($1, 1$) process. For $0.5 \leq d < 1$, the process is non-stationary and Eq. (4) corresponds to a pseudo-spectral density.[27,28]

Given a time series $x_1, \ldots, x_n$, to estimate the parameters of an ARFIMA($p, d, 0$)-GARCH($1, 1$) model proceed as follows:[15,19,29]

1. estimate d using the semi-parametric local Whittle estimator;
2. define the filtered data $y_t = (1 - B)^d x_t$;
3. estimate the AR(p)-GARCH($1, 1$) parameters in the filtered data y_t.

The local Whittle estimate of the parameter d, $\hat{d}$, minimizes the objective function

$$\log\left\{\frac{1}{m}\sum_{j=1}^{m} \omega_j^{2d} I_{\omega_j}\right\} - \frac{2d}{m}\sum_{j=1}^{m} \log \omega_j,$$

where $I_{\omega_j} = \frac{1}{n}\left|\sum_{t=1}^{n} x_t e^{-it\omega_j}\right|^2$ with $\omega_j = \frac{2\pi j}{n}$ for $j = 1, \ldots, m$ are the first m harmonics of the periodogram. Robinson[30] has shown that the local Whittle estimator is consistent and asymptotically normal for $-0.5 < d < 0.5$,

$$m^{1/2}(\hat{d} - d_0) \xrightarrow{d} N\left(0, \frac{1}{4}\right), \tag{6}$$

where d_0 is the true value of d. Velasco[28] extended Robinson's results to show that the estimator is consistent for $-0.5 < d < 1$ and asymptotically normally distributed for $-0.5 < d < 0.75$. Using simulations, Nielsen & Frederiksen[31] verified that the local Whittle estimator is fairly robust to conditionally heteroscedastic errors. The local Whittle estimator depends on the choice of bandwidth m and is generally chosen in the range of $n^{0.5} \leq m \leq n^{0.65}$, where n is the sample size. In this work, we consider $m = n^{0.5}$ because it is less sensitive to the existence of short-memory components. In step 2, to approximate the filtered data $y_t = (1 - B)^d x_t$, a procedure in the frequency domain, proposed by Geweke and Porter-Hudak,[32] is used. This approach consists in calculating the Fourier transform of the observed series $x_1, \ldots, x_n$ which is then multiplied by the Fourier transform of the fractional difference operator based on $\hat{d}$ and, finally, calculating the inverse Fourier transform. In step 3, AR(p)-GARCH(1,1) parameters are estimated by maximum likelihood[17,20] using the GARCH Toolbox of MATLAB.[33] Initial parameters are obtained by least squares and the order p of the AR component determined by the Akaike Information Criterion (AIC).

Conditional heteroscedasticity in the series is assessed by formal hypothesis tests in which the null hypothesis that the series exhibits no conditional heteroscedasticity is tested against the alternative of a GARCH(P, Q) model. One such test is Engle's[34] which is based on the percentage of variation of x_t^2 that is explained in terms of linear dependence on its M lagged values, R^2. Thus, large values of R^2 are indicative of linear dependence of x_t^2 on $x_{t-1}^2, \ldots, x_{t-M}^2$. Under the hypothesis of Gaussianity nR^2 is asymptotically equivalent to the score test statistic and $nR^2 \sim \chi^2(M)$. Another test used in this work is the McLeod-Li[35] test based on the chi-squared statistic

$$Q = n(n+2)\sum_{k=1}^{L} \hat{r}^2(k)/(n-k) \sim \chi^2(L),$$

where L is the number of autocorrelations included in the statistic and $\hat{r}(k)$ is the sample autocorrelation of the squared data at lag k.

III. ARFIMA-GARCH MODELING OF HRV

To motivate the use of ARFIMA(p, d, 0)-GARCH(P, Q) models in HRV data consider Figs. 1–4. Fig. 1 represents in (a) the tachogram for a healthy subject with 1024 beats (segment of RR series for subject-N2, provided by PhysioNet[23]), in (b) the sample ACF of the data and in (c) the ACF of squared data. The ACF of the data shows a very slow decay indicating that the dependence between distant observations is not negligible and thus displaying long memory

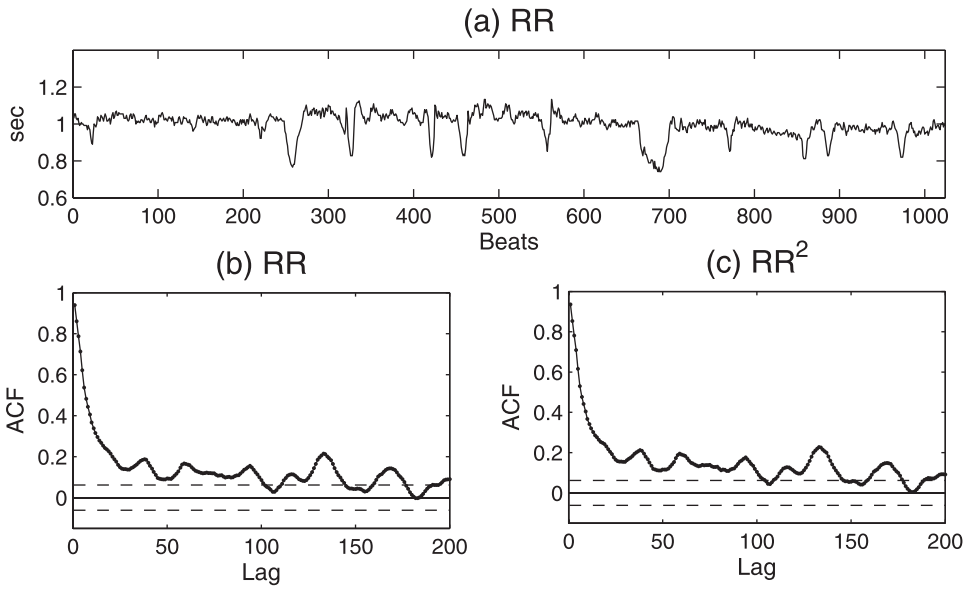

FIG. 1. Short HRV data: (a) tachogram of a normal subject (segment with 1024 beats of RR series for subject-N2, provided by PhysioNet), (b) ACF of the data, and (c) ACF of the squared data. The horizontal lines (- -) show the 95% confidence limits.

characteristics. The data are then modeled with an ARFIMA$(p,d,0)$ (Eq. (1) with constant variance $\hat{\sigma}_\epsilon^2 = \hat{u}_0 = 0.435 \times 10^{-3}$) with $p=8$ (selected by AIC criterion) and $\hat{d} = 0.313$, *Model I*. For $N=50$ surrogate series, the 95% tolerance interval (population coverage 95%; $K_\alpha = 2.065$) for d, $]-0.0756, 0.1106[$, does not contain the estimated valued in the observed series, $\hat{d} = 0.313$ indicating that the data has long memory in the mean. The residuals ($\hat{\epsilon}_t$) are displayed in Fig. 2(a) and the corresponding ACF in (b), exhibits little correlation indicating that the ARFIMA model is adequate to explain the dynamics and conditional mean of the data. However, the squared residuals exhibit significant autocorrelation in Fig. 2(c), indicating time-varying conditional variance. These results are confirmed by the p-value <0.001 of Engle and McLeod-Li tests applied to the residuals of ARFIMA. Now, to model this effect, a GARCH(0,1) model is entertained for the ARFIMA residuals, *Model II*: ARFIMA(8, d, 0)-GARCH(0, 1). The estimate for parameter u_1 is $\hat{u}_1 = 0.293$ but the p-values of Engle and McLeod-Li tests, 0.041 and 0.026, respectively, lead to the rejection the null hypothesis of no conditional heteroscedasticity.

The record is next modeled by an ARFIMA(8, d, 0)-GARCH(1, 1), *Model III*. The residuals ($\hat{z}_t = \hat{\epsilon}_t/\hat{\sigma}_t$) are displayed in Fig. 3(a) and the corresponding ACFs, in (b) and (c), indicate no significant autocorrelations in both the residuals and squared residuals. In fact, the p values for Engle and McLeod-Li tests do not lead to the rejection of the null hypothesis, indicating that there exists no significant conditional heteroscedasticity in the residual series. These results indicate that the ARFIMA(8, d, 0)-GARCH(1, 1) model is adequate explaining the conditional mean and conditional heteroscedastic variance of the data. The results for all models are summarized in Table I.

It is worth noting that Model III leads to further characterization of the record. In fact, the estimate $\hat{v}_1 = 0.64$ indicates some persistence in the variance of the record. The

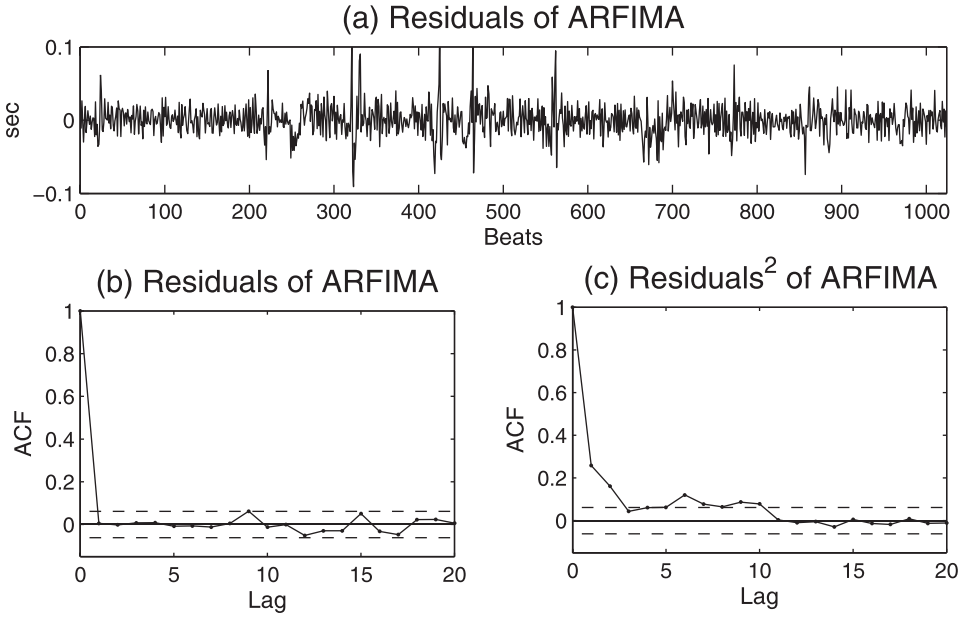

FIG. 2. Same data as Fig. 1: (a) residuals ($\hat{\epsilon}_t$) of the fitted ARFIMA(8, d, 0) model, (b) ACF of the residuals, and (c) ACF of the squared residuals. The horizontal lines (- -) show the 95% confidence limits.

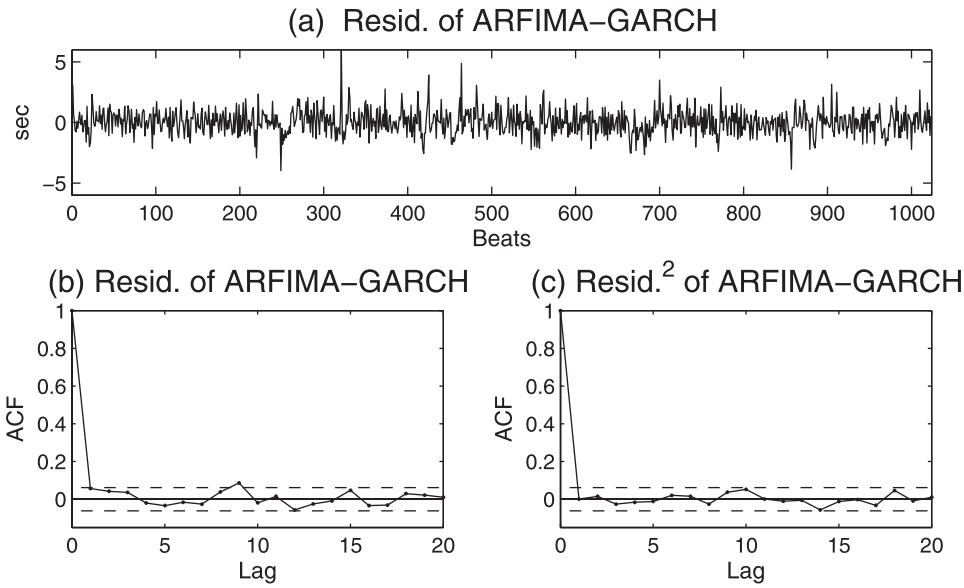

FIG. 3. Same data as Fig. 1: (a) residuals ($\hat{z}_t = \hat{\epsilon}_t/\hat{\sigma}_t$) of the fitted ARFIMA(8, d, 0)-GARCH(1, 1) model, (b) ACF of the residuals, and (c) ACF of the squared residuals. The horizontal lines (- -) show the 95% confidence limits.

conditional standard deviation estimate $\hat{\sigma}_t$ represented in Fig. 4(a), captures very well the heteroscedasticity in the original data series, plotted in Fig. 1(a). Furthermore, this estimate improves the identification of transient phenomena in comparison to AR analysis with recursive least squares estimation represented in (b). The unconditional standard deviation is shown as a horizontal line in Fig. 4(a).

These results indicate that ARFIMA($p, d, 0$)-GARCH(1, 1) models are adequate in HRV recordings, allowing more parsimonious modeling than AR(p) modeling (typically around $p = 12$ for AR, see Refs. 2 and 14). Similar results were obtained in other short HRV recordings.

Consider now ambulatory 24 h HRV recordings which are long (approximately 100 000 beats), exhibiting several non stationary characteristics with circadian variation in mean and variance, as well as many change points (Fig. 5). The usual approach in such situations is segmentation: the long record is decomposed into short records of variable length (S_i) which are homogeneous according to some criterion, e.g., AIC criterion, allowing the obtention of break points which mark the end of consecutive short records, see Neidźwiecki.[6] In this work, the segmentation approach is adopted and the break points are identified by a suitable AIC criterion for ARFIMA($p, d, 0$)-GARCH(1, 1) models:[20,36]

$$\mathrm{AIC} = n \ln \hat{\sigma}_\epsilon^2 + 2(p+4).$$

Thus the procedure is as follows:[14]

1. let S be the minimum allowable segment length and S_i initialized with S samples, $S_i \geq S$; each segment starts at the last break point determined and has length $S_i + S$;
2. a candidate break point is selected when $\mathrm{AIC}_0 - \mathrm{AIC}_1 \geq 0$, where AIC_0 is the value of the AIC in segment with $S_i + S$ samples and AIC_1 is the sum of AIC in the first S_i samples with AIC in the last S samples of the segment (the values of AIC are obtained considering a fixed value for p);
3. a new break point is obtained by selecting the most significant break points (maximum $\mathrm{AIC}_0 - \mathrm{AIC}_1$) from a set of candidates in the next S samples.

A minimum allowable length for the short records is set as 512 beats. This value is supported by a simulation study

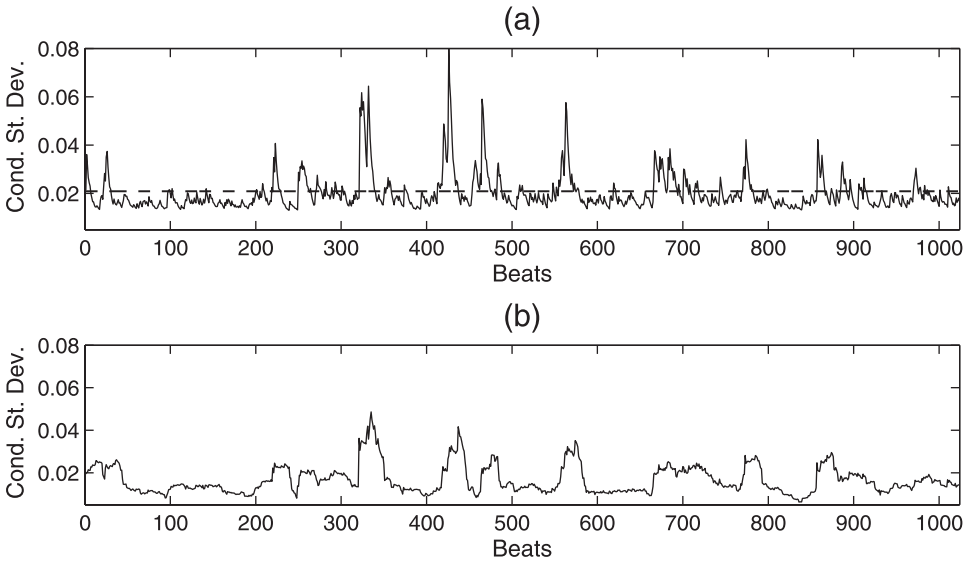

FIG. 4. Same data as Fig. 1: conditional standard deviations estimate $\hat{\sigma}_t$ from ARFIMA(8, d, 0)-GARCH(1, 1) modeling (a) and from time-variant AR analysis with recursive least squares estimation (b). The horizontal line, in (a), shows the unconditional standard deviation.

TABLE I. Estimates (standard deviation) for Models I, II, and III adjusted to the record represented in Fig. 1(a). Engle and McLeod-Li (ML) tests are applied to the residuals and the *p*-values are reported. Same data as Fig. 1.

	Model I	Model II	Model III
$\hat{d}$	0.313	0.313	0.313
$\hat{u}_0$	0.435×10^{-3}	0.304×10^{-3}	0.574×10^{-4}
	(0.125×10^{-4})	(0.122×10^{-4})	(0.102×10^{-4})
$\hat{u}_1$	…	0.293	0.233
	…	(0.036)	(0.032)
$\hat{v}_1$	…	…	0.640
	…	…	(0.044)
p-value$_{Engle}$	0.000	0.041	0.766
p-value$_{ML}$	0.000	0.026	0.794

undertaken by the authors which indicates that 512 beats ensures an adequate estimation of the long memory parameter d. These findings are corroborated by Nielsen and Frederiksen.[31] Furthermore, for $N = 50$ RR surrogate series of each of 150 recordings of a normal subject, with 512 beats, 79.3% of the 95% tolerance intervals (population coverage 95%) for d, do not contain the estimated valued $\hat{d}$ in the observed series. This indicates that the size 512 of the recording is adequate for long memory estimation in HRV. The segments thus obtained are modeled by ARFIMA-GARCH models allowing $0.5 < d < 1$, leading to ARFIMA-GARCH models which are non stationary but mean reverting; since the local Whittle estimator for d is still consistent in that interval, these segments can be modelled by ARFIMA-GARCH models.

This methodology leads to the possibility of describing the time evolution of the parameters of the model over the 24 h.

IV. RESULTS AND DISCUSSION

In this section, the above methodology is applied to long-term HRV series of 15 subjects provided by PhysioNet:[23] 5 series obtained from normal subjects (N, 21.8 ± 1.8 h; 91262 ± 8884 points; 29.8 ± 10.4 years), 5 series from congestive heart failure patients (C, 20.0 ± 0.1 h; 88701 ± 16085 points; 58.8 ± 9.3 years), and 5 series obtained while the patients were undergoing atrial fibrillation (A, 23.4 ± 1.9 h; 116677 ± 24041 points; age and starting time of the recording process are unavailable for this group). The results are first illustrated for a normal subject-N2, Fig. 6, a patient affected by congestive heart failure-C5, Fig. 7, and a patient in atrial fibrillation-A4, Fig. 8.

The long memory estimates $\hat{d}$ for N2 and C5, in Fig. 6(b) and Fig. 7(b), change over time showing circadian variation, with lowest values during the night periods, $0 < \hat{d} < 0.5$ in contrast with $0.5 < \hat{d} < 1$ for the day period. However, the long memory parameter estimates for patient A4, Fig. 8(b), does not exhibit circadian variation with most of the estimates ranging from 0 to 0.5. These findings are corroborated by surrogate data testing: $N = 50$ surrogate series for each of the segments of the long RR series are generated and d estimated. The corresponding 95% tolerance intervals (population coverage 95%; $K_\alpha = 2.065$) are represented by a grey region in Figs. 6(b), 7(b), and 8(b). It is clear from the figures that for subjects N2 and C5 the long memory estimates are statistically different from those obtained in the surrogate series (at 5% significance level). In fact, in 88% of the segments the tolerance interval does not contain $\hat{d}$ for subject N2, this percentage being of 90% for subject C5. For subject A4, the percentage of tolerance intervals that do not contain $\hat{d}$ drops to 64%.

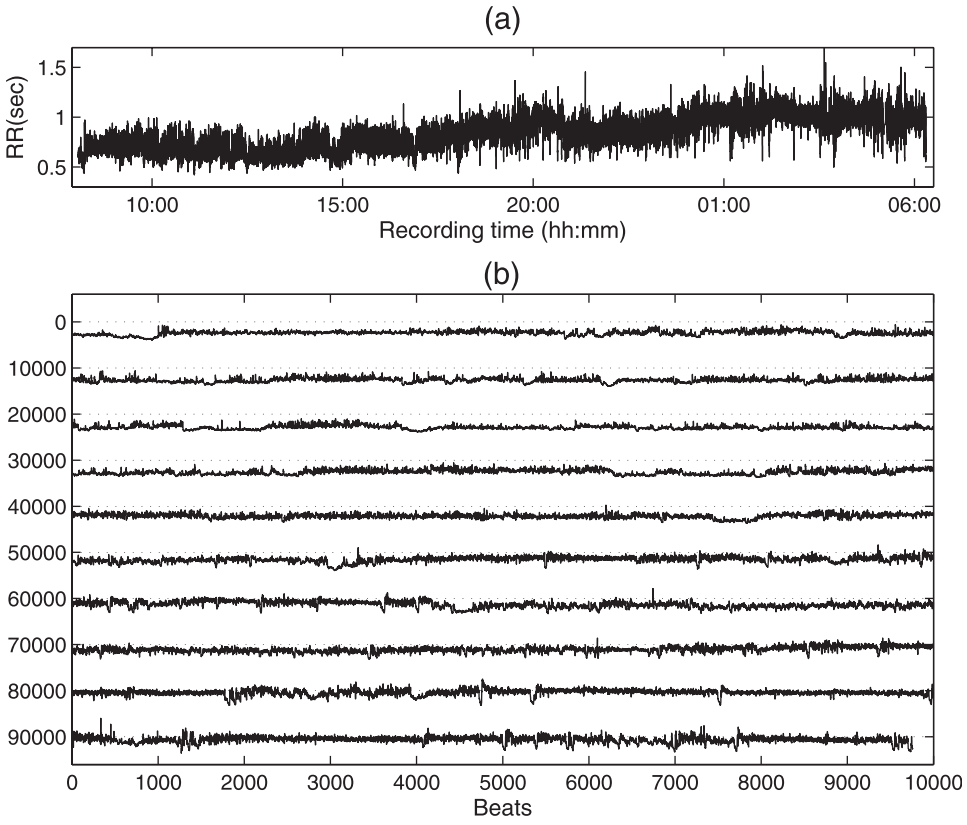

FIG. 5. Long HRV data: (a) tachogram of a normal subject, 24 h, and (b) the same tachogram, with 99 761 beats.

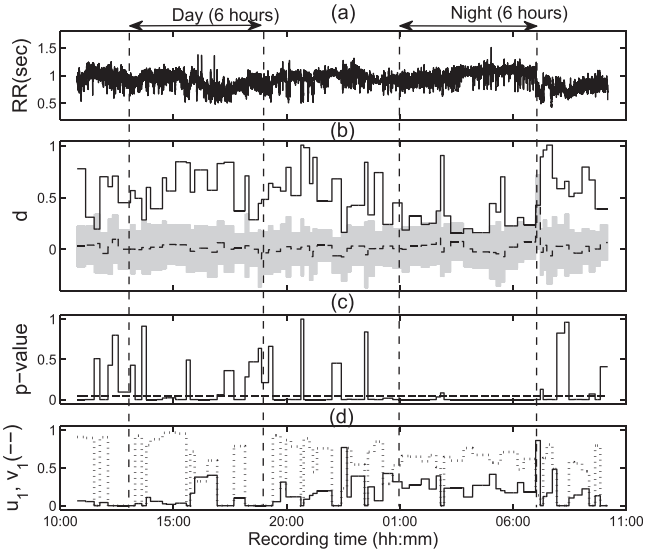

FIG. 6. (a) Tachogram of normal subject-N2, 24 h recordings provided by PhysioNet. Evolution over 24 h of $\hat{d}$ in (b) and $\hat{u}_1$(-) and $\hat{v}_1$(- -) in (d), estimated using ARFIMA-GARCH models and segmentation; a grey region shows the 95% tolerance intervals for d estimated from 50 surrogate data in (b); p-values of the McLeod-Li test for conditional heteroscedasticity in (c).

The residuals from the ARFIMA models are tested for GARCH effect. Plot (c) in Figs. 6–8 represent the p-values of the McLeod-Li test for conditional heteroscedasticity (the p values for Engle's test are similar). For subject-N2 and patient C5, the tests indicate that the data have volatility. Moreover, the volatility parameters estimates $\hat{u}_1$ and $\hat{v}_1$, Figs. 6(d) and 7(d), change over time with some circadian variation. Note that the estimated values for parameter v_1 are over 0.5 indicating some persistence in variance. Finally, the p-values for patient A4 indicate that the conditional variance of the data is constant over time, Fig. 8(c).

These ARFIMA-GARCH analyses are carried out for all the subjects of the three groups of patients. For groups N and

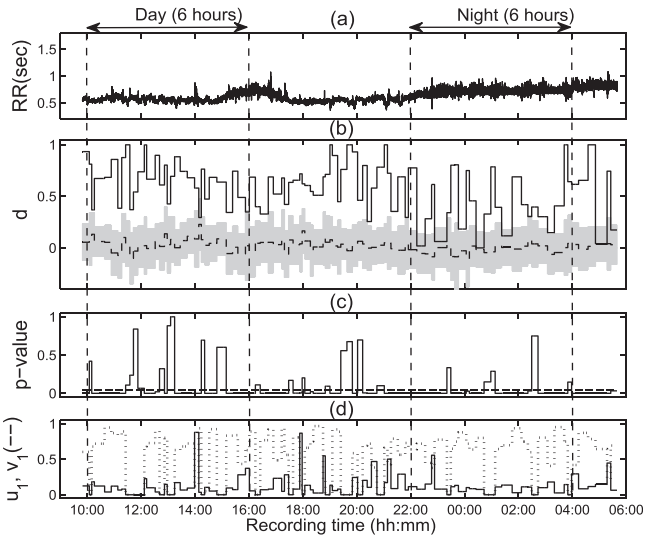

FIG. 7. (a) Tachogram of patient affected by congestive heart failure-C5, 24 h recordings provided by PhysioNet. Evolution over 24 h of $\hat{d}$ in (b) and $\hat{u}_1$(-) and $\hat{v}_1$(- -) in (d), estimated using ARFIMA-GARCH models and segmentation; a grey region shows the 95% tolerance intervals for d estimated from 50 surrogate data in (b); p-values of the McLeod-Li test for conditional heteroscedasticity in (c).

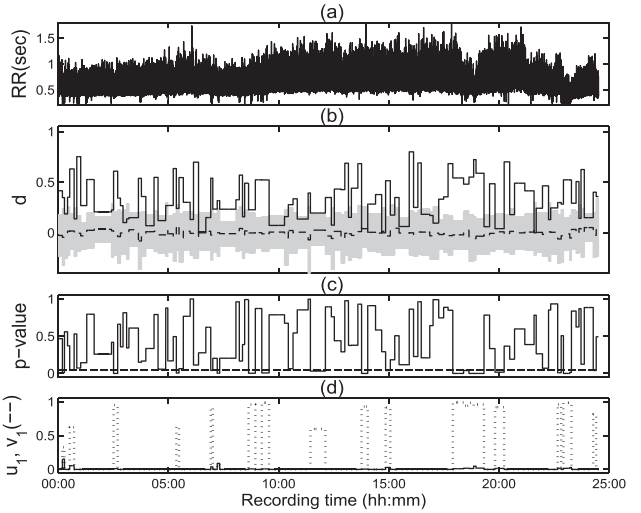

FIG. 8. (a) Tachograms of patient in atrial fibrillation-A4, 24 h recordings provided by PhysioNet. Evolution over 24 h of $\hat{d}$ in (b) and $\hat{u}_1$(-) and $\hat{v}_1$(- -) in (d), estimated using ARFIMA-GARCH models and segmentation; a grey region shows the 95% tolerance intervals for d estimated from 50 surrogate data in (b); p-values of the McLeod-Li test for conditional heteroscedasticity in (c).

C, the starting time of the Holter diary is available, enabling to distinguish between day and night periods. Therefore, for these two groups, the analyses are carried out for the 24 h period as well as for 6 h during day and night periods. For group A, the starting time of the Holter diary is unavailable and the analyses are performed only for the 24 h periods. Engle and McLeod-Li tests are applied to the residuals of ARFIMA model ($\hat{\epsilon}_t$) and to the residuals of ARFIMA-GARCH model ($\hat{z}_t = \hat{\epsilon}_t/\hat{\sigma}_t$) and the percentage of segments with conditional heteroscedasticity are reported. The results are summarized in Table II.

The results indicate that HRV is a long memory process. However, for normal subjects and congestive heart failure patients the long memory $\hat{d}$, changes over time and presents circadian variation. Testing this hypothesis with surrogate data as before, it is found that among the normal subjects 85.7% of the segments during day time and 74.6% during night time have long memory. Similar numbers are observed for the congestive heart failure patients: 95.9% and 85.8%, respectively, see Table III. These results are in concordance with Baillie *et al.*[16] and Leite *et al.*[14,37,38] Moreover, patients suffering from congestive heart failure present increased values for $\hat{d}$, during night and day periods, while patients undergoing atrial fibrillation present lower values for $\hat{d}$, compared with normal subjects. In fact, only 54.2% of the segments in atrial fibrillation patients present statistically significance values for d as tested with surrogate data, Table III. These results are in agreement with the results obtained by Sassi *et al.*[39]

Regarding conditional heteroscedasticity, the results (Table II) indicate that it is observed in normal subjects and the patients suffering from congestive heart failure (in accordance to Baillie *et al.*[16]). However, most of the segments of patients undergoing atrial fibrillation do not present this feature (87.4% for Engle and 84.7% for Ljung-Box tests). The estimates for the volatility parameter u_1, $\hat{u}_1$, decrease for

TABLE II. ARFIMA$(p,d,0)$-GARCH(1,1) model analysis for the three groups of patients provided by PhysioNet: normal subjects N, congestive heart failure patients C, and patients were undergoing atrial fibrillation A, during 24 h, 6 h of night and 6 h of day periods (start times are unavailable for the group A). The Engle and McLeod-Li tests are applied to the residuals and percentage of segments with conditional heteroscedasticity are reported ($\hat{\epsilon}_t$ are the residuals of ARFIMA and $\hat{z}_t = \hat{\epsilon}_t/\hat{\sigma}_t$ are the residuals of ARFIMA-GARCH). For each case, the average estimates ± standard deviations are presented.

Parameter	Period	N	C	A
$\hat{d}$	24 h	0.443 ± 0.044	0.642 ± 0.142	0.258 ± 0.062
	Night-6 h	0.347 ± 0.047	0.540 ± 0.177	…
	Day-6 h	0.499 ± 0.080	0.691 ± 0.111	…
Seg. with heteroscedasticity in $\hat{\epsilon}_t$				
Engle test	24 h	77.9 ± 5.5	74.5 ± 7.4	12.6 ± 5.4
McLeod-Li test	24 h	78.0 ± 6.1	72.4 ± 11.0	15.3 ± 7.1
$\hat{u}_1$	24 h	0.132 ± 0.035	0.150 ± 0.039	0.004 ± 0.002
	Night-6 h	0.168 ± 0.063	0.170 ± 0.095	…
	Day-6 h	0.108 ± 0.037	0.132 ± 0.035	…
$\hat{v}_1$	24 h	0.561 ± 0.081	0.469 ± 0.087	0.127 ± 0.072
	Night-6 h	0.646 ± 0.051	0.520 ± 0.131	…
	Day-6 h	0.539 ± 0.180	0.420 ± 0.120	…
Seg. with heteroscedasticity in $\hat{z}_t$				
Engle test	24 h	11.0 ± 4.0	10.1 ± 6.9	4.3 ± 2.3
McLeod-Li test	24 h	8.9 ± 3.5	3.6 ± 3.4	3.5 ± 2.8

TABLE III. Percentage of segments (mean ± standard deviation) for which RR series $\hat{d}$ is outside the 95% tolerance interval estimated from surrogate data, for the three groups (normal subjects N, congestive heart failure patients C, and patients in undergoing atrial fibrillation A).

Period	N	C	A
24 h	80.8 ± 8.5	91.7 ± 8.9	54.2 ± 11.6
Night-6 h	74.6 ± 6.3	85.8 ± 14.3	…
Day-6 h	85.7 ± 13.7	95.9 ± 4.4	…

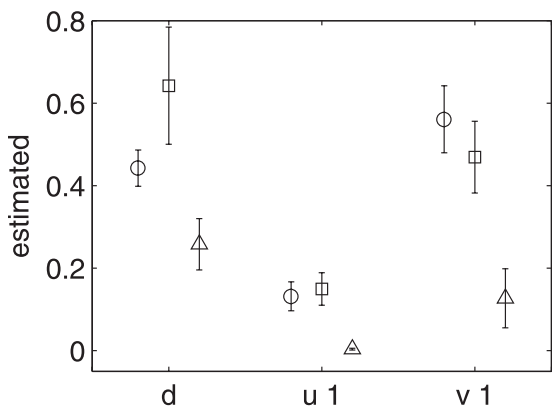

FIG. 9. Average estimates and standard deviations of $\hat{d}$, $\hat{u}_1$, and $\hat{v}_1$ for the three groups of patients: normal (N, ○), congestive heart failure patients (C, □) and patients in atrial fibrillation (A, △) during 24 h.

TABLE IV. p-value of Kruskal-Wallis rank sum test and multiple comparison results for the three groups (normal subjects N, congestive heart failure patients C, and patients in undergoing atrial fibrillation A) during 24 h. The symbol ✓✓ (✓) indicates significant differences at 5% (10%) level.

		Multiple comparison	
Parameter	Kruskal-Wallis test p-value	N versus A	C versus A
$\hat{d}$	0.004	…	✓✓
$\hat{u}_1$	0.007	✓	✓✓
$\hat{v}_1$	0.005	✓✓	✓✓

patients in atrial fibrillation. The estimates for the volatility parameter v_1, $\hat{v}_1$, decrease for patients suffering from congestive heart failure, during night and day periods, and for patients were undergoing atrial fibrillation, with the lowest values for the patients in atrial fibrillation. The results of the Engle and McLeod-Li tests applied to the residuals of ARFIMA-GARCH indicate that this model is adequate for the HRV data.

Finally, statistical differences among the three groups are studied for the 24 h period, relatively to d, u_1, and v_1, applying the Kruskal-Wallis rank sum test and multiple comparison procedures (5% level of significance). The results summarized in Table IV and illustrated in Fig. 9 indicate that the long range parameter in the mean $\hat{d}$, as well as the parameters governing the conditional variance dynamics, $\hat{u}_1$ and $\hat{v}_1$, differ between the groups C and A. Additionally, the parameter $\hat{v}_1$ differs between groups N and A.

V. CONCLUSION

In the last decade, research in HRV has introduced novel methods for assessing heart rate dynamics that may be used for risk stratification. Among those measures, the scaling exponents related to the fractal dimension for measuring long memory have a prominent role.[40] Observational studies have suggested that these scaling exponents may provide useful prognostic information in various clinical situations.

The present study also assesses the long memory properties of HRV records but using a parametric approach based in models for time series. However, it goes further in assessing also volatility in HRV records, via parametric ARFIMA-GARCH modeling. The methodology is applied to three groups of patients. The results regarding the long memory or fractal dimension of the HRV records are in accordance with other studies published in the literature: the long-memory parameter varies over time, showing circadian variation,

presenting lower values for patients undergoing atrial fibrillation (A) when compared with healthy subjects (N) and patients with congestive heart failure (C). Here, additional non-linear characterization of HRV is accomplished via two additional parameters that characterize HRV volatility. These parameters also present circadian variation, varying over the 24 h. This study indicates important differences among the volatility characteristics of the three groups: the healthy subjects as well as patients with congestive heart failure present volatility, whereas for patients with atrial fibrillation it is reduced. These volatility parameters turn out to contribute also to distinguish among the three groups. In fact, while the long memory parameter allows to statistically distinguish between the groups C and A, the volatility parameters allow to distinguish also between the patients N and A. A parametric approach such as the one proposed here has the advantage of allowing not only the obtention of new measures to describe the dynamics of HRV but also the assessment of traditional measures generally used in clinical practice such as low frequency and high frequency components. Even though it is acknowledged that the physiological background of novel methods of analyzing heart rate dynamics is poorly understood, well designed clinical studies and reproduction of results by independent researchers and multiple population samples may establish the predictive power of these measures for risk assessment.

ACKNOWLEDGMENTS

The authors are grateful to the anonymous referees whose comments helped to improve the paper. Research funded by *FEDER* through the programme COMPETE and by the Portuguese Government through the FCT - Fundação para a Ciência e a Tecnologia under the Projects PEst-C/MAT/UI0144/2011, PEst-C/MAT/UI4106/2011, and PEst-OE/MAT/UI4080/2011.

[1]M. L. Apple, R. D. Berger, J. P. Saul, J. M. Smith, and R. J. Cohen, "Beat to beat variability in cardiovascular variables: noise or music?," J. Am. Coll. Cardiol. **14**, 1139–1148 (1989).
[2]Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, "Heart rate variability, standards of measurement, physiological interpretation and clinical use," Circulation **93**, 1043–1065 (1996).
[3]A. M. Bianchi, L. T. Mainardi, and S. Cerutti, "Time-frequency analysis of biomedical signals," Trans. Inst. Meas. Control **22**, 215–230 (2000).
[4]A. P. Rocha, A. Leite, S. Gouveia, P. Lago, O. Costa, and A. F. Freitas, "Spectral characterization of long-term ambulatory heart rate variability signals," in *5th IMA Conference on Mathematics in Signal Processing, Warwick, UK*, 2000.
[5]L. T. Mainardi, A. M. Bianchi, and S. Cerutti, "Time-frequency and time-varying analysis for assessing the dynamic responses of cardiovascular control," Crit. Rev. Biomed. Eng. **30**, 175–217 (2002).
[6]M. Neidźwiecki, *Identification of Time-Varying Processes* (Wiley, New York, 2000).
[7]M. Kobayashi and T. Musha, "1/f fluctuation of heartbeat period," IEEE Trans. Biomed. Eng. **29**, 456–457 (1982).
[8]J. P. Saul, P. Albrecht, R. D. Berger, and R. J. Cohen, "Analysis of long term heart rate variability: Methods, 1/f scaling and implications," in *Computers in Cardiology 1987*, edited by IEEE Engineering in Medicine and Biology Society (IEEE Computer Society Press, New York, 1987), pp. 419–422.
[9]A. L. Goldberger, V. Bhargava, B. J. West, and A. J. Mandell, "On a mechanism of cardiac electrical stability: the fractal hypothesis," Biophys. J. **48**, 525–528 (1985).
[10]A. L. Goldberger and B. J. West, "Fractals in physiology and medicine," Yale J. Biol. Med. **60**, 421–435 (1987).
[11]C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Golberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," Chaos **5**, 82–87 (1995).
[12]Y. Yamamoto and R. L. Hughson, "Coarse-graining spectral analysis: new method for studying heart rate variability," J. Appl. Physiol. **71**, 1143–1150 (1991).
[13]S. Cerutti, F. Esposti, M. Ferrario, R. Sassi, and M. G. Signorini, "Long-term invariant parameters obtained from 24-h holter recordings: a comparison between different analysis techniques," Chaos **17**, 015108 (2007).
[14]A. Leite, A. P. Rocha, M. E. Silva, and O. Costa, "Modelling long-term heart rate variability: An ARFIMA approach," Biomed. Tech. **51**, 215–219 (2006).
[15]J. R. M. Hosking, "Fractional differencing," Biometrika **68**, 165–176 (1981).
[16]R. T. Baillie, A. A. Cecen, and C. Erkal, "Normal heartbeat series are non-chaotic, nonlinear, and multifractal: New evidence from semiparametric and parametric tests," Chaos **19**, 028503 (2009).
[17]T. Bollerslev, "Generalized autoregressive conditional heteroscedasticity," J. Econometr. **31**, 307–327 (1986).
[18]M. Y. Hu and C. Tsoukalas, "Conditional volatility properties of sleep-disordered breathing," Comput. Biol. Med. **36**, 303–312 (2006).
[19]R. T. Baillie and C. F. Chung, "Analysing inflation by the fractionally integrated ARFIMA-GARCH model," J. Appl. Econometr. **11**, 23–40 (1996).
[20]S. Ling and W. K. Li, "On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity," J. Am. Stat. Assoc. **92**, 1184–1193 (1997).
[21]S. Ling, "Adaptive estimators and tests of stationary and nonstationary short- and long-memory ARFIMA-GARCH models," J. Am. Stat. Assoc. **98**, 955–967 (2003).
[22]D. V. Vougas, "Analysing long memory and volatility of returns in the Athens stock exchange," Appl. Financ. Econ. **14**, 457–460 (2004).
[23]A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit and PhysioNet: Components of a new research resource for complex physiologic signals," Circulation **101**, e215–e220 (2000).
[24]J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, "Testing for nonlinearity in time series: the method of surrogate data," Physica D **58**, 77–94 (1992).
[25]J. Beran, *Statistics for Long-Memory Processes* (Chapman & Hall, New York, 1994).
[26]R. T. Baillie, "Long memory processes and fractional integration in econometrics," J. Econometr. **73**, 5–59 (1996).
[27]C. M. Hurvich and B. K. Ray, "Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes," J. Time Ser. Anal. **16**, 17–41 (1995).
[28]C. Velasco, "Gaussian semiparametric estimation of nonstationary time series," J. Time Ser. Anal. **20**, 87–127 (1999).
[29]P. Boukhan, G. Oppenheim, and M. S. Taqqu, *Theory and Applications of Long-Range Dependence* (Birkhäuser, Boston, 2003).
[30]P. M. Robinson, "Gaussian semiparametric estimation of long range dependence," Ann. Stat. **23**, 1630–1661 (1995).
[31]M. O. Nielsen and P. H. Frederiksen, "Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration," Econometric Rev. **24**, 405–443 (2005).
[32]J. Geweke and S. Porter-Hudak, "The estimation and application of long memory time series models," J. Time Ser. Anal. **4**, 221–238 (1983).
[33]MATLAB, *version 7.10.0 (R2010a)*, The MathWorks Inc., Natick, Massachusetts, 2010.
[34]R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," Econometrica **50**, 987–1007 (1982).
[35]G. M. Ljung and G. E. P. Box, "On a measure of lack of fit in time series models," Biometrika **65**, 297–303 (1978).
[36]J. Beran, R. J. Bhansali, and D. Ocker, "On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes," Biometrika **85**, 921–934 (1998).

[37]A. Leite, A. P. Rocha, M. E. Silva, S. Gouveia, J. Carvalho, and O. Costa, "Long-range dependence in heart rate variability data: ARFIMA modelling vs detrended fluctuation analysis," Comput. Cardiol. **34**, 21–24 (2007).
[38]A. Leite, A. P. Rocha, and M. E. Silva, "Long memory and volatility in HRV: An ARFIMA-GARCH approach," Comput. Cardiol. **36**, 165–168 (2009).
[39]R. Sassi, M. G. Signorini, and S. Cerutti, "Multifractality and heart rate variability," Chaos **19**, 028507 (2009).
[40]H. V. Huikuri, J. S. Perkiömäki, R. Maestri, and G. D. Pinna, "Clinical impact of evaluation of cardiovascular control by novel methods of heart rate dynamics," Philos. Trans. R. Soc. London, Ser. A **367**, 1223–1238 (2009).