Investigating the Level of Significance of the Coherence Function in Cardiovascular Variability Analysis

L Faes, G Nollo, R Antolini

INFM & Department of Physics, University of Trento, Italy

Abstract

Although the presence of significant coupling between cardiovascular variability series is usually verified according to the threshold value of 0.5 in the coherence function (CF), parameters of spectral estimator should be considered. In this study the surrogate data technique was introduced to define the level of significance of CF. The proposed method determined a frequency-dependent threshold over which the hypothesis of zero coherence was rejected. The weighted covariance method and the autoregressive method were used to estimate CF on simulated series with different degrees of linear coupling and on real cardiovascular data. The threshold was dependent on the type and parameters of the adopted coherence estimator and on the record length, demonstrating the applicability of the method for all types of spectral estimator. This should favor its introduction for assessing the reliability of transfer function estimates in cardiovascular variability analysis.

1. Introduction

The coherence function (CF) is widely used as a frequency-domain measure of similarity between pairs of time series. Even though the coherence is able to detect only linear interactions, it is considered an important tool for system identification and/or source identification studies in different biological systems [1]. CF may be indicated as a correlation coefficient in the frequency domain and its value -ranging from 0 to 1- measures the amount of linear coupling between the two investigated processes at every frequency bin.

In real systems, an unresolved problem is the definition of the minimal CF level that can be considered significantly different from zero. Moreover, this threshold assesses what extent of linear coupling between two time series is needed to guarantee the reliability of transfer function estimates. De Boer et al. [2] proposed CF=0.5 as a gauge of reliability for gain and phase estimation between two cardiovascular variables. Even though it is empirical, this criterion has been widely used later on [3,4]. However, it should be remarked that different estimation procedures determine different thresholds necessary to accept the hypothesis that CF significantly exceeds zero. Indeed, Koopmans [5] demonstrated that for the weighted covariance (WC) coherence estimator this threshold was dependent on the degrees of freedom (deriving from the relations between sample size and window shape) and this theoretical approach has been followed by some recent works [6,7]. Being derived from asymptotic distributions of spectral estimators, the theoretical relationship holds only approximately in practical analysis. Moreover, an analogous approach is not pursuable for the class of autoregressive (AR) spectral estimators.

The technique of surrogate data has been introduced for the definition of the nonlinearity extent contained in various measures of coupling strength [8]. The method computed a discriminating parameter sensitive to nonlinearity on a set of surrogate data generated from the original time series according to the null hypothesis of gaussian linear stochastic process. Finding a statistical difference in the discriminating parameter between original and surrogate data leaded to reject the null hypothesis and thus to detect nonlinearity.

In this study, we propose to generate surrogate data consistent with the null hypothesis of linearly independent stochastic processes to establish a significance level for CF. In this way, a specific threshold can be derived on data for each estimation procedure as a function of frequency.

2. Methods

The theoretical CF between two time series $x(n)$ and $y(n)$ of N values (n=0,...,N-1), considered as realizations of two stochastic stationary processes, is given by [9]:

$$K_{xy}^2(f)=\frac{\left|P_{xy}(f)\right|^2}{P_x(f)P_y(f)} \tag{1}$$

where $P_x(f)$ and $P_y(f)$ are the autospectra and $P_{xy}(f)$ is the cross-spectrum of the two series. We utilized CF as discriminating parameter for the surrogate data analysis. Each pair of surrogate data ($x_k(n)$ and $y_k(n)$, k=1,...K) was generated by randomly shuffling the temporal order of the original time series. In this way, in the surrogate data the mean, the variance, and the histogram

0276-6547/01 $17.00 © 2001 IEEE

distribution were preserved, but any mutual correlation between the two series was destroyed. At each frequency bin, we set the threshold over which coherence values were assumed as statistically significant by evaluating their mean and standard deviation on the surrogate data distribution:

$$T(f)=mean_k\left\{K^2_{x_k y_k}(f)\right\}+\zeta(\alpha)sd_k\left\{K^2_{x_k y_k}(f)\right\} \qquad (2)$$

Under the hypothesis of gaussian distribution of the discriminating statistic, $\alpha=\mathrm{erfc}(\zeta(\alpha)/2^{1/2})$ is the probability to find values larger than the threshold if the null hypothesis is true. This definition of threshold considers that only one value of the discriminating statistic can be evaluated from the original data. If multiple realizations are available for the observational data, the null hypothesis of incoherent signals can be more properly rejected by tests that compare the two distributions of original and surrogate data (Kolmogorov-Smirnov or Mann-Whitney test). We performed the statistical test of Equation 2 generating K=20 pairs of surrogate data. The significance α was set to 0.05.

To estimate the dependence of the proposed method on the parameters of the specific spectral estimator, we fed a linear system having flat transfer function equal to 1 in the band 0÷0.5 Hz with a zero-mean white noise (x series). An independent white noise was added at the output of the linear block, yielding the output y of the overall system. The simulation was repeated for four different record lengths (N=180, 300, 420, 540), adjusting the noise level to obtain a theoretical coherence equal to 0.9. The threshold was estimated both by the WC method (Parzen window, spectral bandwidths 0.015 Hz and 0.03 Hz) and the AR method (model order: 8,16) [9].

In the second simulation, we performed the surrogate technique on data generated by the linear model:

$$\begin{aligned} x(n) &= \sqrt{2}\cdot 0.9\cdot x(n-1)+0.9^2\cdot x(n-2)+u(n) \\ y(n) &= x(n)+\beta\cdot w(n) \end{aligned} \qquad (3)$$

with $u(n)$ and $w(n)$ IID uncorrelated gaussian white noises. The series x is described as an autoregressive process with two complex conjugate poles having modulus 0.9 and phase $\pm\pi/4$. The level of coupling between x and y was modulated by the parameter β. Varying the value of β from 0 to 5 the two series were desynchronized as the value of the coherence peak at f=1/8 Hz decreased progressively towards zero. CF estimates were obtained on N=1024 samples by the AR technique, with model order set to 10, and the WC method, by averaging CF estimated on eight non overlapped segments smoothed with a Parzen window (Welch method) [9].

Eventually, an application of the surrogate technique to cardiovascular variability series was considered. The significance of CF between 300 consecutive values of heart period (RR interval of ECG) and systolic arterial pressure was assessed by the surrogate technique on a subject hospitalized after acute myocardial infarction. According to the usual approach followed for the cross-spectral analysis of cardiovascular signals [10], the CF was evaluated by the AR method. The sampling frequency was assumed equal to the reciprocal of the mean heart period.

3. Results

An example of three representative cases with different degrees of linear correlation was considered to illustrate how the surrogate technique works in determining the significance of CF (Fig. 1).

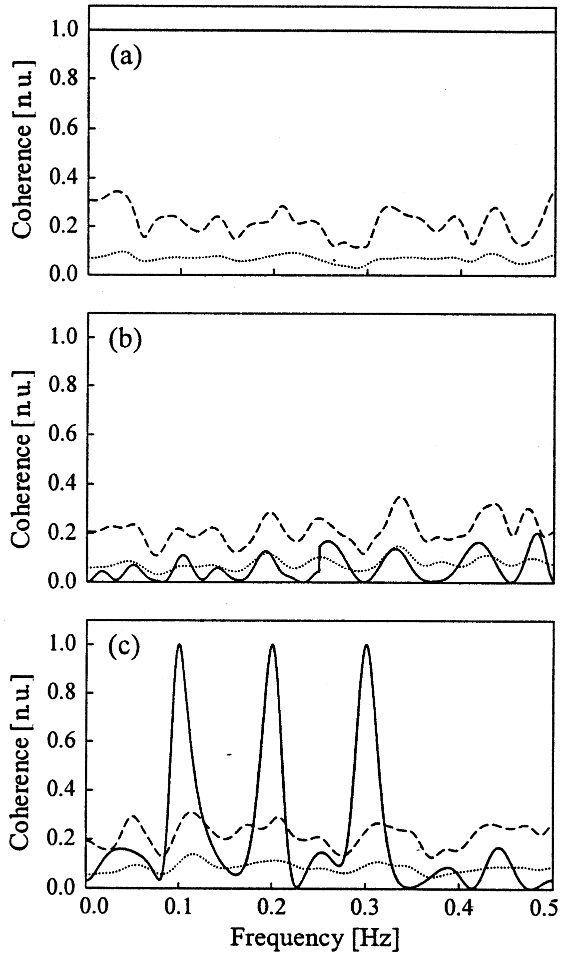

Figure 1. Coherence function (continuous line), average coherence on surrogate data (dotted line), and corresponding threshold (dashed line) evaluated for completely correlated (a), uncorrelated (b), and partially correlated (c) pairs of simulated series.

CF was estimated on N=300 samples by the AR technique (model order=12). In Fig 1a the x series was constructed as an independent identically distributed (IID) gaussian white noise with zero mean and unitary variance and the y series was obtained as a linear transformation of x. The coherence estimate was equal to 1 at all frequencies and always far from the threshold given by Equation (2). In Fig. 1b x and y were IID uncorrelated gaussian white noises. CF estimated on original data never resulted significantly different from zero (i.e. $K^2_{xy}(f)$ was lower than $T(f)$ for each value of f). In Fig. 1c a complete linear coupling between x and y only at 0.1 Hz, 0.2 Hz, and 0.3 Hz was reproduced by corrupting by independent white noises the amplitude and the phase of three sine waves. CF resulted over the threshold only at the three frequencies exhibiting complete coupling.

As shown in Fig. 2, the threshold increased with decreasing the record length N. Moreover, it resulted related inversely to the bandwidth of the WC estimator and directly to the model order of the AR estimator.

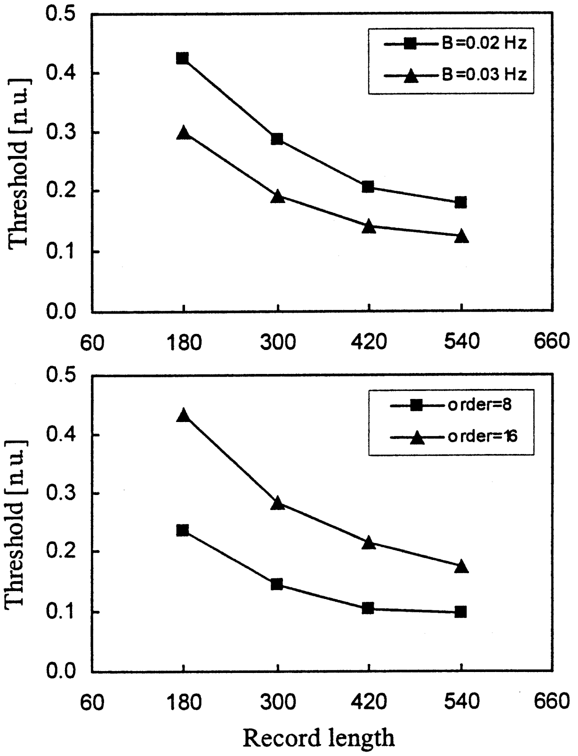

Figure 2. Dependence of the coherence threshold on the record length for different values of the bandwidth of the WC estimator (up) and of the model order of the AR estimator (down).

Fig. 3 shows the values of mean and threshold of CF evaluated on surrogate data as a function of β (second simulation). With these parameters the theoretical threshold was estimated as suggested in [5] by an F-test with 30 equivalent degrees of freedom, obtaining a value of 0.193. The threshold estimated by surrogate data, ranging from 0.183 to 0.202, was comparable with the theoretical one. Even though the autoregressive method provided a lower threshold than the weighted covariance method, both techniques consider $K^2_{.xy}(1/8)$ significantly different from zero when β=4 but not when β=5.

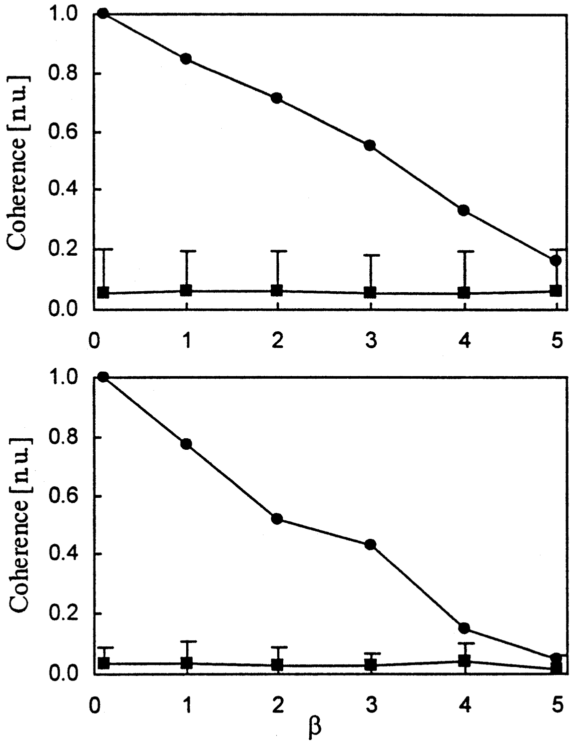

Figure 3. Estimates of CF of original data (circles), mean of CF of surrogate data (squares), and CF threshold (bar) for the simulation described in Eqn. 3 performed by the WC method (up) and the AR method (down).

The evaluation of the significance of CF between the RR interval variability and the systolic pressure variability for the post-myocardial infarction patient is reported in Fig. 4. Whereas the surrogate data approach lead to consider the two main peaks of coherence as significant, the assumption of the 0.5 empirical threshold fails in detecting significance of the flow frequency CF peak.

4. Discussion

The reported examples and simulations suggest that with the surrogate approach different threshold values for CF can be obtained depending on the specific method adopted for the estimation of the coherence (parametric or non parametric estimation, estimator parameters, sample size). For this reason, the surrogate technique is entitled to establish the significance of coherence and consequently the reliability of the correspondent transfer function more properly than the empirical 0.5 level. This ability should avoid to discard from the analysis subjects with low but significant coupling. The need of accurately evaluate the degree of coupling is particularly important in post-infarction patients, often characterized by a depressed, but not certainly absent, link between the cardiovascular variables [10].

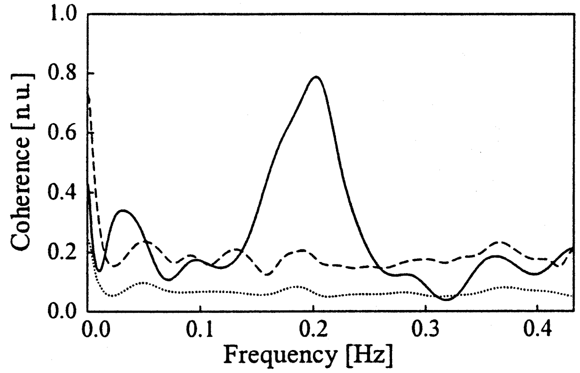

Figure 4. Estimation of coherence (continuous line), mean of coherence on surrogate data (dotted line), and zero-coherence threshold (dashed line) by the AR method for a post-myocardial infarction patient.

The comparable values between the theoretical and the surrogate threshold obtained by applying the WC estimator suggest the feasibility of the proposed statistical approach for the determination of the CF level necessary to reject the hypothesis of uncoupling between the investigated series.

It is known that the parameters of the adopted CF estimator (i.e., the record length and the spectral bandwidth or the model order) affect its variance and thus the reliability of its estimates [5,9]. The threshold obtained by the surrogate data was related to such parameters in the same way of the variance (see Fig. 2). Thus, with the proposed method the presence of coupling is tested by accounting for the parameters that affect the confidence of spectral and cross-spectral estimates and should be performed independently on the kind of CF estimator. Particularly, the surrogate technique could be appropriate for assessing the significance of CF also when it is estimated by the AR method, for which no theoretical criteria have been provided.

References

[1] Challis RE, Kitney RI. Biomedical signal processing (in four parts). Part 3. The power spectrum and coherence function. Med Biol Eng Comput 1991;29(3):225-41.
[2] De Boer RW, Karemaker JM, Strackee J. Relationships between short-term blood-pressure fluctuations and heart-rate variability in resting subjects: a spectral analysis approach. Med Biol Eng Comput 1985;23:352-8.
[3] Robbe HWJ, Mulder LJM, Ruddel H, Langewitz WA, Veldman JBP, Mulder G. Assessment of baroreceptor reflex sensitivity by means of spectral analysis. Hypertension 1987;10:538-43.
[4] Saul JP, Berger RD, Albrecht P, Stein SP, Hui Chen M, Cohen RJ. Transfer function analysis of the circulation: unique insights into cardiovascular regulation. Am J Physiol 1991;261:H1231-45.
[5] Koopmans LH. The spectral analysis of time series. New York: Academic Press, 1974.
[6] Taylor JA, Carr DL, Myers CW, Eckberg DL. Mechanisms underlying very-low-frequency RR-interval oscillations in humans. Circulation 1998;98:547-55.
[7] Pinna GD, Maestri R. Reliability of transfer function estimates in cardiovascular variability analysis. Med Biol Eng Comput 2001;39:338-47.
[8] Prichard D, Theiler J. Generating surrogate data for time series with several simultaneously measured variables. Phys Rev Lett 1994;73:951-4.
[9] Kay SM. Modern spectral estimation. Theory & application. New Jersey: Prentice Hall, Englewood Cliffs, 1988.
[10] Nollo G, Porta A, Faes L, Del Greco M, Disertori M, Ravelli F. Causal linear parametric model for baroreflex gain assessment in patients with recent myocardial infarction. Am J Physiol 2001;280:H1830-9.

Address for correspondence:

Faes Luca
Dipartimento di Fisica, Università di Trento
Via Sommarive 14, 38050 Povo (Trento), Italy
E-mail: faes@science.unitn.it