

Quantifying the Complexity of Short-Term Heart Period Variability through K Nearest Neighbor Local Linear Prediction

L Faes, S Erla, G Nollo

University of Trento, Trento, Italy

Abstract

The complexity of short-term heart period (HP) variability was quantified exploiting the paradigm that associates the degree of unpredictability of a time series to its dynamical complexity. Complexity was assessed through k -nearest neighbor local linear prediction. A proper selection of the parameter k allowed us to perform either linear or nonlinear prediction, and the comparison of the two approaches to infer the presence of nonlinear dynamics. The method was validated on simulations reproducing linear and nonlinear time series with varying levels of predictability. It was then applied to HP variability series measured from healthy subjects during head-up tilt test, showing that short-term HP complexity increases significantly from the supine to the upright position, and that nonlinearities are involved in the generation of HP dynamics in both positions.

1. Introduction

Characterization of the complexity of heart period (HP) variability is an important problem in the study of the cardiovascular control. Previous studies indicate that the complexity of HP time series, measured as the absence of repetitive patterns into the series, decreases with the reduction of the competition among regulatory subsystems, e.g. induced by cardiovascular diseases [1] or alterations of the experimental conditions [2]. Another significant issue is the evaluation of the impact of nonlinear dynamics in cardiovascular variability, as different physiological states have been associated with different degrees of nonlinearity in HP variability [3]. In short-term (up to few minutes) HP variability, complexity is usually estimated through entropy or predictability measures [2,3], while nonlinearity evaluation relies on the use of surrogate data approaches [3].

In this study we introduce a method to evaluate together complexity and nonlinearity of short HP time series. Exploiting the paradigm of k -nearest neighbor local linear prediction [4], the method quantifies complexity and regularity of the series, and allows to

infer the contribution of nonlinear dynamics through a comparison between linear and nonlinear predictability degrees. It is first tested on simulations reproducing linear and nonlinear dynamics with different complexity levels, and then tested on HP variability measured from healthy subjects during head-up tilt testing.

2. Methods

Let us consider a stationary time series $X(n)$, $n=1, \dots, N$. The series is first normalized by subtracting the mean and dividing to the standard deviation, to obtain the dimensionless series $x(n)$, $n=1, \dots, N$. To predict the current value $x(n)$, we consider the pattern formed by the L previous samples of the series, $x_L(n)=[x(n-1), \dots, x(n-L)]$, and identify its k nearest neighbors as the k patterns $x_L(n_1), \dots, x_L(n_k)$ having the lowest Euclidean distance to $x_L(n)$. To prevent overfitting, the neighbor patterns $x_L(n_j)$, $j=1, \dots, k$, are kept separated from $x_L(n)$ using a Theiler window of $W=N/10$ points [5]. A system of k linear equations (one for each neighbor pattern) was then set as

$$x(n_j) = \sum_{i=1}^L c(i)x(n_j - i), \quad j=1, \dots, k. \quad (1)$$

The system is solved through standard least squares optimization, yielding estimates of the unknown coefficients, $\hat{c}(1), \dots, \hat{c}(L)$, that are then used to predict the current value of the series as

$$\hat{x}(n) = \sum_{i=1}^L \hat{c}(i)x(n-i). \quad (2)$$

After predicting $N-L$ samples, the unpredictability of the series x can be evaluated as mean squared prediction error (MSPE) [3,6]

$$MSPE(L, k) = \frac{1}{N-L} \sum_{n=L+1}^N (x(n) - \hat{x}(n))^2, \quad (3)$$

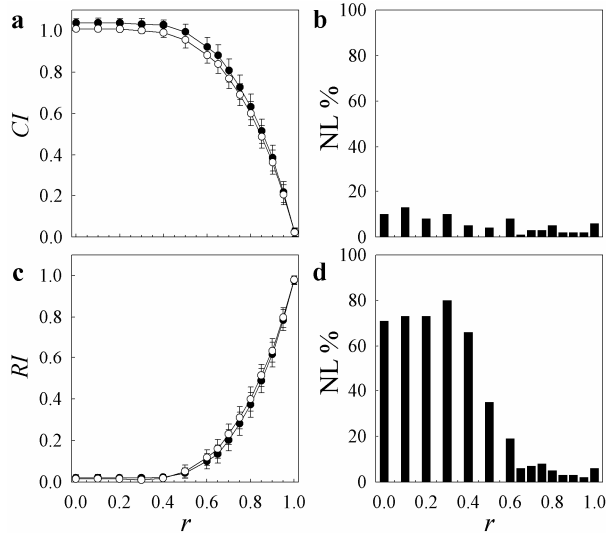


Figure 1. Linear simulation results. Distributions of the complexity (CI) and regularity (RI) indices evaluated over 100 simulation runs as at varying the pole modulus of the linear process using local prediction (black) and global prediction (white) are in **a** and **c**. Percentages of nonlinear dynamics detected by CI and RI are in **b** and **d**.

while its predictability can be calculated as squared correlation (SC) between original and predicted data [3]

$$SC(L, k) = \frac{\left(\sum_{n=L+1}^N (x(n)\hat{x}(n)) \right)^2}{\sum_{n=L+1}^N (x^2(n)) \sum_{n=L+1}^N (\hat{x}^2(n))}. \quad (4)$$

According to (3) and (4), $MSPE$ and SC are functions of the number k of neighbors involved into local linear prediction, and of the length L of the prediction patterns. The number of neighbors is a key parameter, as it determines the nature of the predictor (i.e. linear or nonlinear) [4]. Using a small neighborhood (low k) corresponds to perform a strict local prediction that allows the predictor to describe the possible nonlinear manifold of the data. On the contrary, a global linear prediction is performed with a large neighborhood (high k). In this study, we set $k_{min}=N/10$ to perform local prediction [5], and $k_{max}=N-L-W+1$ to perform global prediction [6]. The choice of L was optimized for each considered series, selecting the value L_{opt} that yielded the maximum predictability. This was possible since out-of-sample prediction avoids overfitting, and thus the predictability is not indefinitely increasing with L . Hence, application of local and global prediction led us to provide for the series X complexity indices (CI_l, CI_g) and regularity indices (RI_l, RI_g) as follows

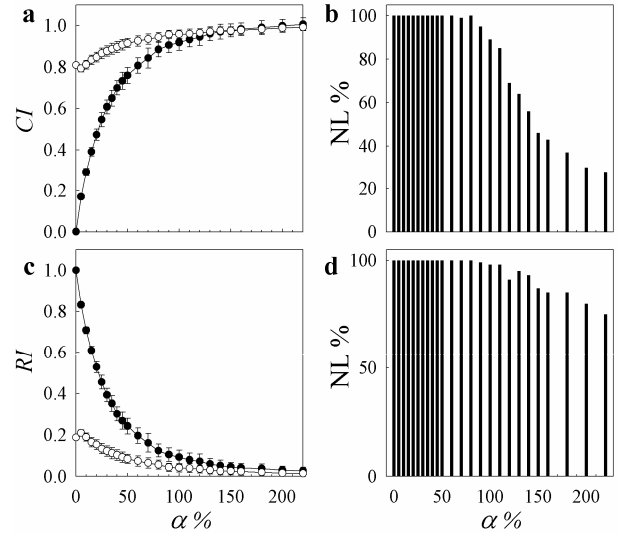


Figure 2. Nonlinear simulation results. Distributions of the complexity (CI) and regularity (RI) indices evaluated over 100 simulation runs as at varying the percentage of additive noise using local prediction (black) and global prediction (white) are in **a** and **c**. Percentages of nonlinear dynamics detected by CI and RI are in **b** and **d**.

$$CI_l = \min_{L=1, \dots, L_{max}} MSPE(L, k_{min}), CI_g = MSPE(L_{opt}, k_{max})$$

$$RI_l = \max_{L=1, \dots, L_{max}} SC(L, k_{min}), RI_g = SC(L_{opt}, k_{max}) \quad (5)$$

CI and RI measures range between 0 and 1, with high predictability corresponding to $CI \sim 0$ and $RI \sim 1$ and high unpredictability to $CI \sim 1$ and $RI \sim 0$. While RI is strictly bounded between 0 and 1, CI may slightly exceed 1 for very unpredictable series.

The comparison between local and global prediction was used to infer the presence of nonlinearities underlying the dynamics of the series X . Specifically, nonlinear dynamics were detected when local prediction yielded higher predictability than global prediction ($CI_l < CI_g$ or $RI_l > RI_g$).

3. Simulations

The proposed approach was tested on simulations reproducing both linear and nonlinear dynamics with different imposed degrees of predictability. Linear dynamics were realizations of a bivariate autoregressive process with two conjugate imaginary poles with modulus r , driven by white Gaussian noise w

$$x(n) = -r^2 x(n-2) + w(n). \quad (6)$$

Nonlinear dynamics were reproduced by the Tent map with control parameter set to produce chaotic dynamics

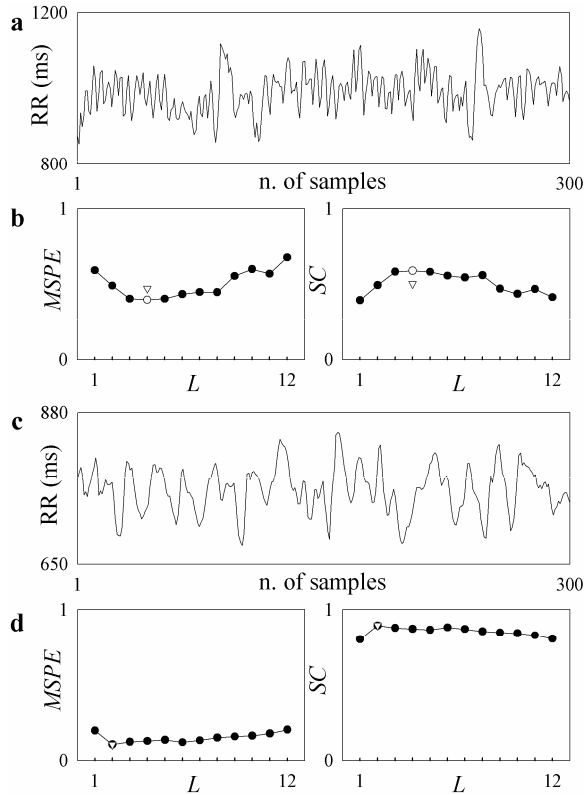


Figure 3. HP variability series measured in the supine and upright position for a representative subject are in **a** and **c**. The corresponding $MSPE$ and SC functions calculated as in (3) and (4) for local prediction ($k=k_{min}$) are plotted in **b** and **d** (black circles), along with the values of CI_l and RI_l (white circles), and of CI_g and RI_g (triangles).

$$x(n+1) = \begin{cases} 1.8x(n) & \text{if } 0 < x(n) \leq 0.5 \\ 1.8[1-x(n)] & \text{if } 0.5 < x(n) \leq 1 \end{cases} \quad (7)$$

To blur nonlinear dynamics, the Tent map was corrupted with additive Gaussian white noise with variance equal to an assigned percentage $\alpha\%$ of the variance of the map generated in (7).

One hundred realizations of the two simulations, each lasting 300 samples, were generated for each value assigned to the parameters r and $\alpha\%$. Results are depicted in Figs. 1 and 2, indicating that: (i) values of the complexity and regularity indices reflect the predictability of simulated series, as increasing the pole modulus r in linear series and decreasing the amount $\alpha\%$ of noise in nonlinear series determined a larger predictability measured by lower CI values and higher RI values; (ii) local and global prediction led to super impossible distributions of the indices for linear time series (Fig. 1a,c), while local prediction is more accurate than global prediction to determine the predictability of

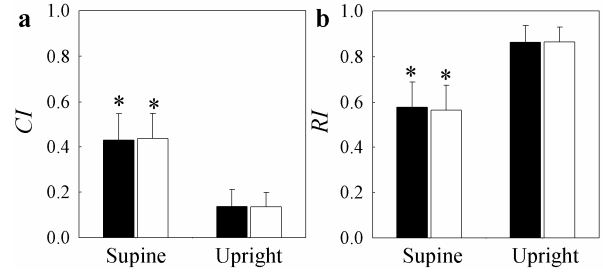


Figure 4. Real data results. Distributions over 10 subjects (mean+SD) of the complexity (CI) and regularity (RI) indices evaluated using local prediction (black) and global prediction (white) are in **a** and **b**. * $p < 0.001$ Supine vs. Upright, Student t-test for paired data.

nonlinear time series (Fig. 2a,c); (iii) in presence of non random time series, i.e. approximately $r > 0.5$ in Fig. 1 and $\alpha\% < 100$ in Fig. 2, both CI and RI indices are able to detect the nature of the simulated dynamics, as $NL\%$ is very low for linear series and very high for nonlinear series (note that for low r and high $\alpha\%$ the series are almost fully unpredictable, thus rendering meaningless the assessment of nonlinear dynamics).

4. Application to heart period variability

The experimental protocol included ten young healthy subjects (24.1 ± 2.1 yrs) in whom the surface ECG (lead II) was acquired in the supine position after subjects' stabilization, and in the upright position after passive 60° head-up tilt. After ECG digitalization (1KHz sampling rate, 12 bit resolution), QRS complexes were located by parabolic interpolation and the RR intervals measured as the temporal distance between consecutive R peaks. The series were then cleaned up from artifacts and two stationary sequences of 300 beats were selected in the supine and upright positions.

Fig. 3 reports an example of HP variability measured in the two positions for a representative subjects, as well as the results of the prediction analysis. In the supine position, minimum $MSPE$ and maximum SC for the local prediction were found at $L_{opt}=4$. In correspondence, global prediction yielded a lower predictability degree, suggesting the presence of nonlinear dynamics underlying the investigated series. In the upright position, RI and CI were evaluated for $L_{opt}=4$, and no differences between local and global prediction were noticed, suggesting that the HP dynamics is substantially linear.

Results of the analysis extended to the 10 subjects are summarized in Fig. 4. According to both CI and RI measures, more than 50% of the HP variance was described by local and global predictors in the supine position, and the predictability increased significantly with the transition to the upright position. The

predictability degrees were not statistically different when evaluated by local and global prediction. However, the comparison of the two approaches in the single time series suggested that nonlinearities are involved in HP dynamics ($CI < CI_g$ and $RI > RI_g$) in 5 subjects at rest and 4 subjects after tilt.

5. Discussion

The results of the present study suggest that k-nearest neighbor local linear prediction is an efficient tool to quantify complexity in short time series. Simulations showed that both predictability and unpredictability measures are able to follow changes in complexity imposed in short linear and nonlinear time series. Application to real data evidenced that the indices detect the complexity of short-term HP dynamics, as well as its significant reduction in the transition from the supine to the upright position. These results agree with the expected physiological behavior: for instance, a simplification of the heart rate dynamics, leading to a lower complexity of HP series, has been previously reported in correspondence of the rise of a dominant low frequency oscillation induced by the sympathetic activation associated to head-up tilt [7]. Our results also indicate that in the considered experimental protocol local and global predictors provide comparable complexity degrees, and that complexity can be assessed equally well using predictability- or unpredictability-based indices.

Besides the quantification of complexity, this study introduces also a new approach to assess the presence of nonlinear dynamics in short-term HP variability. The traditional method to test for nonlinearity in time series is that making use of a set of surrogate data, that are copies of the original series in which the property under investigation (in this case nonlinear dynamics) is destroyed, while other properties (e.g., amplitude distribution and linear correlation) are maintained [8]. Here we propose an alternative approach, based on comparing the predictability (or unpredictability) degrees given by local prediction and global prediction. Our hypothesis is that if a local model yields a better prediction than a global model, indicated either by a lower CI or a higher RI , the improvement might be ascribed to a significant contribution of nonlinear dynamics which are detected by the local model but not by the global one. Simulation results showed that this approach works well for both linear time series, in which prediction is slightly better using a global model, and nonlinear time series, in which the predictability is much larger when a local model is adopted. The application on real cardiovascular data seems also effective, showing that –although global and local prediction yielded comparable average results– nonlinear dynamics might

be involved in about a half of the considered subjects, in either body position. Previous studies have also shown how nonlinear dynamics may contribute to short-term HP variability both in the supine and in the upright body positions [3,9]. Hence, although its agreement with the traditional surrogate approach has not been tested yet, this new method to look for nonlinearity in time series seems promising; it offers also the advantages of being independent on the structure of the analyzed time series (while surrogate approaches may fail to reproduce some features while destroying nonlinearity), and of demanding a reduced computational effort.

References

- [1] Bigger JT, Steinman RC, Rolnitzky LM, Fleiss JL, Albrecht P, Cohen RJ. Power law behavior of RR-interval variability in healthy middle-aged persons, patients with recent acute myocardial infarction, and patients with heart transplants. *Circulation* 1996;93:2142-51.
- [2] Porta A, Guzzetti S, Montano N, Pagani M, Somers V, Malliani A, Baselli G, Cerutti S. Information domain analysis of cardiovascular variability signals: evaluation of regularity, synchronisation and co-ordination. *Med Biol Eng Comput* 2000;38:180-8.
- [3] Porta A, Guzzetti S, Furlan R, Gneccchi-Ruscione T, Montano N, Malliani A. Complexity and nonlinearity in short-term heart period variability: comparison of methods based on local nonlinear prediction. *IEEE Trans Biomed Eng* 2007;54:94-106.
- [4] Casdagli M. Nonlinear prediction of chaotic time series. *Physica D* 1989;35:335-56.
- [5] Faes L, Porta A, Nollo G. Mutual nonlinear prediction as a tool to evaluate coupling strength and directionality in bivariate time series: Comparison among different strategies based on k nearest neighbors. *Phys Rev E* 2008;in press. doi:10.1103/PhysRevE.78.026201.
- [6] Faes L, Nollo G. Bivariate nonlinear prediction to quantify the strength of complex dynamical interactions in short-term cardiovascular variability. *Med Biol Eng Comput*. 2006;44:383-92.
- [7] Montano N, Gneccchi Ruscone G, Porta A, Lombardi F, Pagani M, Malliani A. Power spectrum analysis of heart rate variability to assess the changes in sympathovagal balance during graded orthostatic tilt. *Circulation* 1994;90:1826-31.
- [8] Schreiber T, Schmitz A. Surrogate time series. *Physica D* 2000;142:346-82.
- [9] Braun C, Kowallik P, Freking A, Haderl D, Kniffki KD, Meesmann M. Demonstration of nonlinear components in heart rate variability of healthy persons. *Am. J. Physiol Heart Circ. Physiol* 1998;275:H1577-84.

Address for correspondence

Luca Faes
 Laboratorio di Biofisica e Biosegnali
 Via Sommarie 14, 38050 Povo, Trento, Italy.
 luca.faes@unitn.it