INSTITUTE OF PHYSICS PUBLISHING PHYSIOLOGICAL MEASUREMENT

Physiol. Meas. **26** (2005) 363–372 doi:10.1088/0967-3334/26/4/003

Leave-one-out prediction error of systolic arterial pressure time series under paced breathing

N Ancona[1], R Maestri[2], D Marinazzo[3,4,5], L Nitti[3,5,6], M Pellicoro[3,4,5], G D Pinna[2] and S Stramaglia[3,4,5]

[1] Istituto di Studi sui Sistemi Intelligenti per l'Automazione, CNR, Bari, Italy
[2] Dipartimento di Bioingegneria e Tecnologie Biomediche, Fondazione Salvatore Maugeri, IRCCS Istituto Scientifico di Montescano (PV), Italy
[3] TIRES-Center of Innovative Technologies for Signal Detection and Processing, University of Bari, Italy
[4] Dipartimento Interateneo di Fisica, Bari, Italy
[5] Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Italy
[6] Dipartimento di Biochimica Medica, Biologia Medica e Fisica Medica, University of Bari, Italy

E-mail: sebastiano.stramaglia@ba.infn.it

Received 19 November 2004, accepted for publication 18 February 2005
Published 4 April 2005
Online at stacks.iop.org/PM/26/363

Abstract

In this paper, we consider systolic arterial pressure time series from healthy subjects and chronic heart failure patients, undergoing paced respiration, and show that different physiological states and pathological conditions may be characterized in terms of predictability of time series signals from the underlying biological system. We model time series by the regularized least-squares approach and quantify predictability by the leave-one-out error. We find that the entrainment mechanism connected to paced breath, that renders the arterial blood pressure signal more regular and thus more predictable, is less effective in patients, and this effect correlates with the seriousness of the heart failure. Using a Gaussian kernel, so that all orders of nonlinearity are taken into account, the leave-one-out error separates controls from patients (probability less than 10^{-7}), and alive patients from patients for whom cardiac death occurred (probability less than 0.01).

Keywords: systolic blood pressure, paced breathing, predictability

1. Introduction

Physiological signals derived from humans are extraordinarily complex, as they reflect ongoing processes involving very complicated regulation mechanisms (Glass 2001), and can be used to diagnose incipient pathophysiological conditions. Many approaches to characterization and analysis of physiological signals have been introduced in recent years, including, for

example, studies of Fourier spectra (Akselrod *et al* 1981, Pinna *et al* 2002), chaotic dynamics (Babloyantz *et al* 1985, Poon and Merrill 1997), wavelet analysis (Thurner *et al* 1998, Marrone *et al* 1999), scaling properties (Nunes Amaral *et al* 1998, Ashkenazy *et al* 2001, Ivanov and Lo 2002), multifractal properties (Ivanov *et al* 1999, Nunes Amaral *et al* 2001), correlation integrals (Lehnertz and Elger 1998), $1/f$ spectra (Peng *et al* 1993, Ivanov *et al* 2001) and synchronization properties (Schafer *et al* 1998, Tass *et al* 1998, Angelini *et al* 2004). Less attention has been paid to the degree of determinism (Kantz and Schreiber 1997) of a physiological time series. It is the purpose of the present work to show that different physiological states, or pathological conditions, may be characterized in terms of the *predictability* of time series. In particular, we consider here the *predictability* of systolic blood pressure (SBP) time series under *paced respiration*, and show that a suitable index separates healthy subjects from chronic heart failure (CHF) patients. SBP is the maximal pressure within the cardiovascular system as the heart pumps blood into the arteries. Paced respiration (breathing is synchronized with some external signal) is a well-established experimental procedure to regularize and standardize respiratory activity during autonomic laboratory investigations (Cooke *et al* 1998), and is a useful tool for relaxation and for the treatment of chronic pain and insomnia, dental and facial pain, etc (Clark and Hirschman 1980, 1990, Freedman and Woodward 1992). Entrainment between heart and respiration rate (cardiorespiratory synchronization) has been detected in subjects undergoing paced respiration (Schiek *et al* 1997, Pomortsev *et al* 1998). Paced breathing can prevent vasovagal syncope during head-up tilt testing (Jauregui-Renaud *et al* 2003); in healthy subjects under paced respiration the synchronization between the main processes governing the cardiovascular system is stronger than the synchronization in the case of spontaneous respiration (Prokhorov *et al* 2003). However, a number of important questions remain open about paced breathing, including the dependence on the frequency of respiration and whether it affects the autonomic balance. In a healthy cardiorespiratory system, the regime of paced respiration induces *regularization* of related physiological signals (Brown *et al* 1993, Pinna *et al* 2003), in particular blood pressure time series smoothen and become more deterministic. To quantify this phenomenon, we face two problems at this point: (i) how may we model the SBP time series? (ii) what measure of predictability is the most suitable? In the present paper, we model time series by the regularized least-squares (RLS) approach (Mukherjee *et al* 2002). The choice of this class of model is motivated by the fact that it has several interesting properties. The most important is that such models have high generalization capacity. This means that they are able to predict complex signals when a finite and small number of observations of the signal itself are available. Moreover the degree of nonlinearity present in the modelling, introduced by a kernel method, may be easily controlled. Finally they allow an easy calculation of the leave-one-out (LOO) error (Vapnik 1998), the quantity that we use to quantify predictability. Our approach generalizes the classical autoregressive (AR) approach to time series analysis (Kantz and Schreiber 1997). It is worth mentioning that recently (Shalizi *et al* 2004) a measure of self-organization, rooted on optimal predictors, has been proposed. In the same spirit, LOO prediction error is related to the degree of organization of the underlying physiological system.

2. Method

2.1. Regularized least-squares linear models for regression

Let us consider a set of ℓ independent, identically distributed data $S = \{(\mathbf{x}_i, y_i)\}_{i=1}^{\ell}$, where $\mathbf{x}_i$ is the n-dimensional vector of input variables and y_i is the scalar output variable. Data are

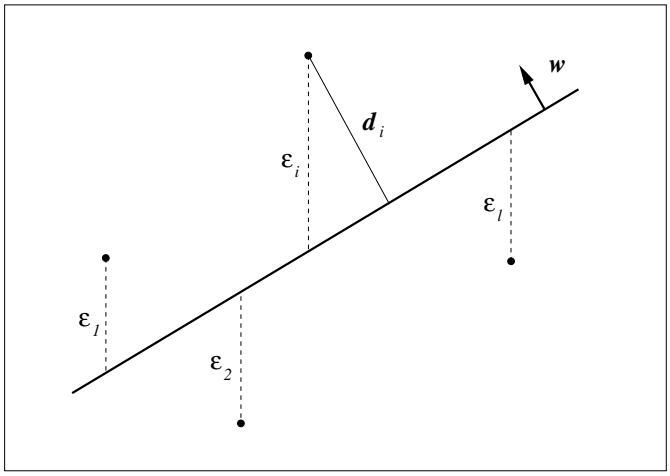

Figure 1. Geometrical interpretation of regularization.

drawn from an unknown probability distribution $p(\mathbf{x}, y)$. The problem of learning consists in providing an estimator $f_{\mathbf{w}} : \mathbf{x} \to y$, out of a class of functions $F(\mathbf{w})$, called *hypothesis space*, parametrized by a vector $\mathbf{w}$. Let us first consider the class of linear functions $y = \mathbf{w} \cdot \mathbf{x}$, where $\mathbf{w}$ is the n-dimensional vector of parameters. To provide a bias term in the linear function (to be included if $\mathbf{x}$ or y has a non-vanishing mean), a supplementary input variable (constant and equal to 1) is to be included in the input vector. In the regularized least-squares approach, $\mathbf{w}$ is chosen so as to minimize the following functional:

$$L(\mathbf{w}) = \frac{1}{\ell}\left[\sum_{i=1}^{\ell}(y_i - \mathbf{w} \cdot \mathbf{x}_i)^2 + \lambda\|\mathbf{w}\|^2\right], \tag{1}$$

where $\|\mathbf{w}\| = \sqrt{\mathbf{w} \cdot \mathbf{w}}$ is the Euclidean norm induced by the scalar product. The first term in functional L is called *empirical risk*, the mean square prediction error evaluated on the training data; the second term (*regularization term*) can be motivated geometrically by the following considerations. Let us view data $(\mathbf{x}_i, y_i)$ as points in a $(n + 1)$-dimensional space. Each function $y = \mathbf{w} \cdot \mathbf{x}$ determines an hyperplane in this space, approximating data points. The prediction square error on point i is $\epsilon_i = (y_i - \mathbf{w} \cdot \mathbf{x}_i)^2$; let d_i be the square distance between the point and the approximating hyperplane. It is easy to see that (see figure 1)

$$d_i = \frac{\epsilon_i}{1 + \|\mathbf{w}\|^2}. \tag{2}$$

This equation shows that the smaller $\|\mathbf{w}\|^2$, the better the deviation ϵ_i approximates to the true distance d_i. Hence the role of the regularization term, whose relevance depends on the value of parameter λ and penalizes large values of $\|\mathbf{w}\|$, is to let the linear estimator be chosen as the hyperplane minimizing the mean square distance with the data points. It is easy to minimize functional L and get the optimal hyperplane:

$$\mathbf{w} = (\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{b}, \tag{3}$$

where $\mathbf{A}$ is the $n \times n$ matrix given by

$$\mathbf{A} = \sum_{i=1}^{\ell}\mathbf{x}_i\mathbf{x}_i^{\top}, \tag{4}$$

b is the n-dimensional vector given by

$$\mathbf{b} = \sum_{i=1}^{\ell} y_i \mathbf{x}_i, \tag{5}$$

while **I** stands for the identity matrix.

The empirical risk $E_e = 1/\ell \sum_{i=1}^{\ell} \epsilon_i$ is not a good measure of the quality of the estimator. What matters is the generalization ability, i.e. the prediction error on data points which have not been used to train the estimator. The following measure of the generalization performance, known as the LOO procedure, is both intuitive and statistically robust (one can show that the LOO error is almost unbiased, see Luntz and Brailovsky (1969)). For each i, data point i is removed from the data set. The approximating hyperplane is then determined on the basis of the residual set of $\ell - 1$ points; the square prediction error by this hyperplane on point i will be denoted as ϵ_i^{loo}. The LOO error is then defined as follows: $E_{\text{loo}} = 1/\ell \sum_{i=1}^{\ell} \epsilon_i^{\text{loo}}$. In principle, calculation of E_{loo} requires the estimation of ℓ hyperplanes, thus rendering this procedure unfeasible, or at least unpractical. However the class of model we are considering here allows calculating the LOO error after inversion of only one $\ell \times \ell$ matrix. It can be shown (Mukherjee *et al* 2002) that

$$E_{\text{loo}} = \frac{1}{\ell} \sum_{i=1}^{\ell} \left(\frac{y_i - \mathbf{w} \cdot \mathbf{x}_i}{1 - \mathbf{G}_{ii}} \right)^2, \tag{6}$$

where **w** is trained on the full data set, using (3), and **G** is an $\ell \times \ell$ matrix given by

$$\mathbf{G} = \mathbf{X}^{\top} (\mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{X}; \tag{7}$$

here we denote **X** as the $n \times \ell$ matrix whose columns are input data $\{\mathbf{x}_i\}$.

The value of the parameter λ is to be tuned to minimize the LOO error. In other words, this free parameter is to be tuned to enhance the generalization capability of the model. It is useful, for the nonlinear extension of these models, to express **w** as a linear combination of the vectors $\mathbf{x}_i$ for $i = 1, 2, \ldots, \ell$. Indeed, if $\ell > n$ one can suppose that vectors $\{\mathbf{x}_i\}$ span all the n-dimensional space, constituting an over-complete system of vectors. This means that there exist ℓ coefficients $\mathbf{c} = (c_1, c_2, \ldots, c_\ell)^{\top}$ such that

$$\mathbf{w} = \mathbf{X}\mathbf{c}. \tag{8}$$

Simple calculations yield

$$\mathbf{c} = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y}, \tag{9}$$

where $\mathbf{K} = \mathbf{X}^{\top}\mathbf{X}$ is an $\ell \times \ell$ matrix with generic element $K_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j$, while $\mathbf{y} = (y_1, y_2, \ldots, y_\ell)^{\top}$ is a vector formed by the ℓ values of the output variable. The prediction y, in correspondence to an input vector **x**, may then be written as a sum over the input data:

$$f : \mathbf{x} \to y = \sum_{i=1}^{\ell} c_i \mathbf{x_i} \cdot \mathbf{x}. \tag{10}$$

Equations (9) and (10) show that the evaluation of the linear predictor as well as the computation of the parameter vector **c** involves only scalar products of data in the input space. This property allows us to extend the regularized linear models to the nonlinear case, as we describe in the next subsection.

2.2. Nonlinear models

The extension to the general case of nonlinear predictors is done by mapping the input vectors $\mathbf{x}$ into a higher dimensional space $\mathcal{H}$, called *feature space*, and looking for a linear predictor in this new space. Let $\Phi(\mathbf{x}) \in \mathcal{H}$ be the image of the point $\mathbf{x}$ in the feature space, with

$$\Phi(\mathbf{x}) = (\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \ldots, \phi_N(\mathbf{x}), \ldots),$$

where $\{\phi\}$ are real functions. Note that the number of components of the feature space can be finite, countable or even infinite uncountable. Moreover, suppose that one of the features be constant. This hypothesis allows us to write the linear predictor in the feature space $\mathcal{H}$ without making the bias term explicit. In the feature space induced by the mapping Φ, a linear predictor takes the form

$$y = f(\mathbf{x}) = \mathbf{w} \cdot \Phi(\mathbf{x}), \tag{11}$$

where now $\mathbf{w}$, according to the nature of the feature space, may have a finite or infinite number of components. Again, we hypothesize that $\mathbf{w}$ may be written as a linear combination of the vectors $\Phi(\mathbf{x}_i)$ with $i = 1, 2, \ldots, \ell$ (if this hypothesis would not be met, we thus determine a solution, constrained in the subspace, of the feature space, spanned by vectors $\{\Phi(\mathbf{x}_i)\}_{i=1,\ell}$). This means that there exist ℓ coefficients $(c_1, c_2, \ldots, c_\ell)^\top$ such that

$$\mathbf{w} = \sum_{i=1}^{\ell} c_i \Phi(\mathbf{x}_i). \tag{12}$$

In this hypothesis, the linear predictor in the feature space $\mathcal{H}$ takes the form

$$y = f(\mathbf{x}) = \sum_{i=1}^{\ell} c_i \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}), \tag{13}$$

and, therefore, will be nonlinear in the original input variables. The vector $\mathbf{c}$ is given by (9) with $\mathbf{K}$ being the $\ell \times \ell$ matrix with generic element $K_{ij} = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$. Note that the evaluation of the predictor on new data points and the definition of the matrix $\mathbf{K}$ involve the computation of scalar products between vectors in the feature space, which can be computationally prohibitive if the number of features is very large. A possible solution to these problems consists of making the following choice:

$$\Phi(\mathbf{x}) = (\sqrt{\alpha_1}\psi_1(\mathbf{x}), \sqrt{\alpha_2}\psi_2(\mathbf{x}), \ldots, \sqrt{\alpha_N}\psi_N(\mathbf{x}), \ldots),$$

where α_i and ψ_i are the eigenvalues and eigenfunctions of an integral operator whose kernel $K(\mathbf{x}, \mathbf{y})$ is a positive definite symmetric function. With this choice, the scalar product in the feature space becomes particularly simple because

$$\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) = \sum_{\gamma} \alpha_\gamma \psi_\gamma(\mathbf{x}_i)\psi_\gamma(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j), \tag{14}$$

where the last equality comes from the Mercer–Hilbert–Schmidt theorem for positive definite functions (Riesz and Nagy 1955). The predictor has, in this case, the form

$$y = f(\mathbf{x}) = \sum_{i=1}^{\ell} c_i K(\mathbf{x}_i, \mathbf{x}). \tag{15}$$

Analogously the LOO error can be calculated as follows:

$$E_{\text{loo}} = \frac{1}{\ell} \sum_{i=1}^{\ell} \left(\frac{y_i - \sum_{j=1}^{\ell} K_{ij} c_j}{1 - \mathbf{G}_{ii}} \right)^2, \tag{16}$$

where the matrix $\mathbf{G}$ can be shown to be equal to $\mathbf{K}(\mathbf{K} + \lambda\mathbf{I})^{-1}$. Many choices of the kernel function are possible, for example the polynomial kernel of degree p has the form

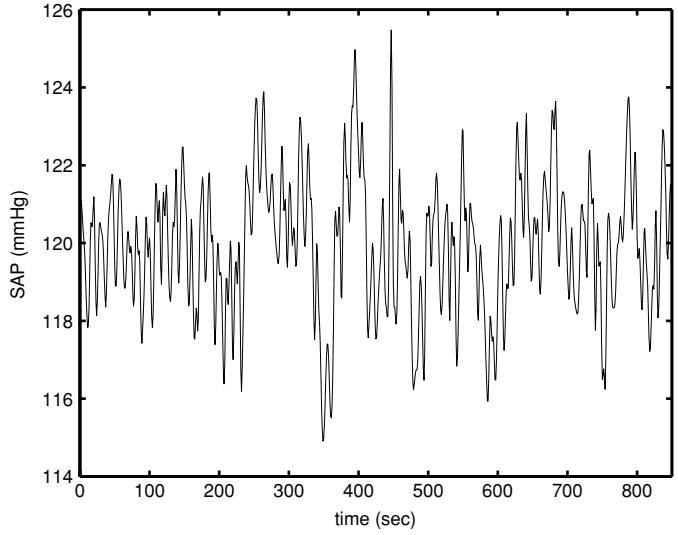

Figure 2. The time series of the systolic arterial pressure for one of the subjects examined.

$K(\mathbf{x}, \mathbf{y}) = (1 + \mathbf{x} \cdot \mathbf{y})^p$ (the corresponding features are made of all the powers of $\mathbf{x}$ up to the pth). The RBF Gaussian kernel is $K(\mathbf{x}, \mathbf{y}) = \exp -(\|\mathbf{x} - \mathbf{y}\|^2/2\sigma^2)$ and deals with all the degrees of nonlinearity of $\mathbf{x}$. Specifying the kernel function K, one determines the complexity of the function space within which we search the predictor, similarly to the effect of specifying the architecture of a neural network, that is the number of layers, the number of units for each layer and the type of activation functions which define the set of functions that the neural network implements. Note that, depending on the kernel function, we can have a countable or even an uncountable number of features. The last case applies, for example, to the Gaussian function. Use of kernel functions to implicitly perform projections, the *kernel trick*, is at the basis of support vector machines, a technique which has found application in several fields, including medicine (Bazzani *et al* 2001).

2.3. Physiological data

Our data are from 47 healthy volunteers (age: 53 ± 8 years, M/F: 40/7) and 275 patients with CHF (age: 52 ± 9 years, left ventricular ejection fraction: $28 \pm 8\%$, New York Heart Association class: 2.1 ± 0.7, M/F: 234/41), caused mainly by ischemic or idiopathic dilated cardiomyopathy (48% and 44% respectively), consecutively referred to the Heart Failure Unit of the Scientific Institute of Montescano, S Maugeri Foundation (Italy) for evaluation and treatment of advanced heart failure. Concerning the second group, cardiac death occurred in 54 (20%) of the patients during a 3-year follow-up, while the other 221 patients were still alive at the end of the follow-up period. All the subjects underwent a 10 min supine rest, recorded in the paced respiration regime (Cooke *et al* 1998, Rzeczinski *et al* 2002). To perform paced breathing, subjects were asked to follow a digitally recorded human voice inducing inspiratory and expiratory phases, at 0.25 Hz frequency. Non-invasive recording of arterial blood pressure at the finger (Finapres device) was performed. For each cardiac cycle, corresponding values of SBP were computed and re-sampled at a frequency of 2 Hz using a cubic spline interpolation. As an example, in figure 2 we report the SBP time series for one of the subjects.

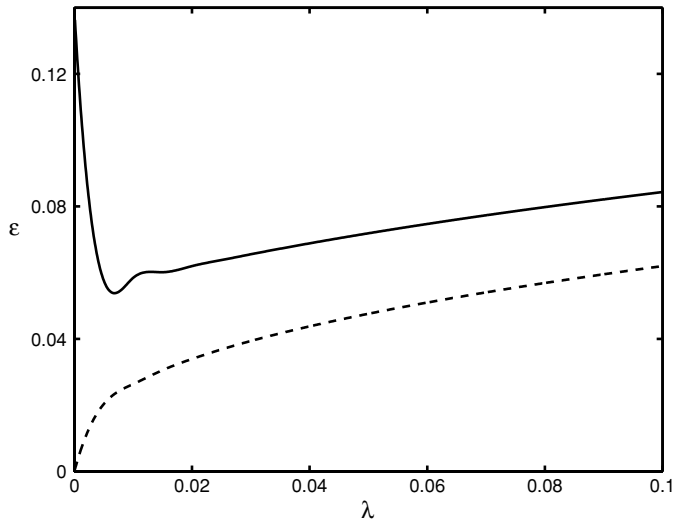

Figure 3. For a typical control subject, the LOO error (continuous line) and the empirical error (dashed line) are represented versus λ. A Gaussian kernel, with $\sigma = 8.5$, is used.

3. Results

Let us denote $\{x_i\}_{i=1,\ldots,N}$ the time series of SBP values, which we assume to be stationary (this assumption is justified by the short length of the recording). The models previously introduced are used to make predictions of the time series. We fix the length of a window m, and for $k = 1$ to ℓ (where $\ell = N - m$), we denote $\mathbf{x}_k = (x_{k+m-1}, x_{k+m-2}, \ldots, x_k)$ and $y_k = x_{k+m}$; we treat these quantities as ℓ realizations of the stochastic variables $\mathbf{x}$ (input variables) and y (output variable). In the preprocessing stage, the time series are normalized to have zero mean and unit variance, but are not filtered. We use $m = 30$, so that the input pattern receives contributions from frequencies greater than 0.066 Hz, thus including part of LF (low frequency 0.04–0.15 Hz) and HF (high frequency 0.15–0.45 Hz) frequency bands, the major rhythms of heart rate and blood pressure variability. All the formalism previously described is applied to model the dependency of y from $\mathbf{x}$, i.e. to forecast the time series on the basis of m previous values: LOO error is a robust measure of its predictability. We use a Gaussian kernel and a polynomial of 1, 2 and 3°.

To show the role of the parameter λ, in figure 3 we depict, for a typical control subject, both the LOO error and the empirical error versus λ. As λ increases, the empirical risk monotonically increases, whilst the LOO error shows a minimum at a finite value of λ ensuring the best generalization capability. It is worth mentioning that one can recover the classical autoregressive approach from our method, setting $\lambda = 0$ in the linear kernel model. We fix the value of λ once for all subjects, by minimizing the average LOO error on a subset made of an equal number of control and CHF time series. This procedure yields $\lambda = 0.01$ for a Gaussian kernel and polynomial of 1, 2 degree, whilst for the third-order polynomial kernel the optimal value we find is $\lambda = 0.1$.[7]

We thus evaluate the LOO error for all the 322 subjects (table 1). In any case, healthy subjects are characterized by a smaller LOO error than patients. Moreover, dead CHF patients show greater LOO error than alive patients. Hence the seriousness of the heart disease appears

[7] For a Gaussian kernel, σ was also similarly tuned to minimize the LOO error, and fixed equal 8.5.

Table 1. Mean values of LOO error. In the parentheses the standard deviation is reported.

Kernel	Controls	CHF	CHF alive	CHF dead
Gaussian	0.039(0.025)	0.081(0.025)	0.077(0.025)	0.097(0.025)
1-poly	0.002(0.004)	0.016(0.03)	0.013(0.03)	0.027(0.04)
2-poly	0.002(0.004)	0.08(0.15)	0.07(0.15)	0.1(0.3)
3-poly	0.01(0.01)	0.15(0.3)	0.148(0.3)	0.153(0.3)

Table 2. p values. In the parentheses the area under the ROC curve is reported.

Kernel	Controls versus CHF	CHF alive versus CHF dead
Gaussian	1.03E-08(0.84)	0.0088(0.64)
1-poly	0.0011(0.70)	0.1825(0.57)
2-poly	0.0010(0.75)	0.1289(0.56)
3-poly	0.0121(0.61)	0.1429(0.53)

to be correlated to the LOO error. The regularized linear model seems to be the best model of SBP time series. For the Gaussian kernel we verify that LOO errors from controls and patients are Gaussianly distributed and check the homogeneity of the variances of the two groups; we apply the t-test to evaluate the probability that LOO error values, relative to controls and patients, were drawn from the same distribution (the null hypothesis); for polynomial kernels, the nonparametric Wilcoxon test is applied (table 2). For all kernels, the null hypothesis can be rejected, also after the Bonferroni correction (which lowers the threshold to $0.05/4 = 0.0125$). The Gaussian kernel shows the best separation between the two classes. We have also tested the separation between dead and alive patients, and the results are displayed in table 2. Only when the Gaussian kernel is used, the p value is lower than 0.0125: since all orders of nonlinearities contribute to the Gaussian modelling, this result suggests that the phenomenon here outlined is an effect with strong nonlinear contributions. We also evaluate the diagnostic power of the LOO error by measuring the area under the receiver-operating-characteristic (ROC) curve (Swets 1988), a well-established index of diagnostic accuracy; the maximum value of 1.0 corresponds to perfect assignment (unity sensitivity for all values of specificity) whereas a value of 0.5 arises from assignments to a class by pure chance. As one can see in table 2, the accuracy is good (i.e. between 0.8 and 0.9) only when the Gaussian kernel is used to discriminate controls from patients.

4. Discussion

We consider here the SBP time series in healthy subjects undergoing paced breath and in patients with heart disease, and we show that the LOO prediction error of physiological time series may usefully be used as a measure of organization of the underlying regulation mechanisms, and can thus be used to detect changes of physiological state and pathological conditions. We propose the use of RLS models in time series prediction because they allow fast calculation of the LOO error and their degree of nonlinearity can be easily controlled. We find that the entrainment mechanism connected to paced breath, that renders the arterial blood pressure signal more deterministic and thus more predictable, is less effective in patients, and this effect correlates with the seriousness of the heart failure; paced breathing conditions seem suitable for diagnostics of a human state. Using a Gaussian kernel, so that all orders of nonlinearity are taken into account, the leave-one-out error separates controls from patients

(probability less than 10^{-7}), and alive patients from patients for whom cardiac death occurred (probability less than 0.01). In our opinion, the LOO error, as a measure of determinism and complexity, is a concept that has potential application to a wide variety of physiological and clinical time series data.

References

Akselrod S, Gordon D, Ubel F A, Shannon D C and Cohen R J 1981 Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control *Science* **213** 220–2

Angelini L *et al* 2004 Steady-state visual evoked potentials and phase synchronization in migraine patients *Phys. Rev. Lett.* **93** 38103–6

Ashkenazy Y, Ivanov P C, Havlin S, Peng C K, Goldberger A L and Stanley H E 2001 Magnitude and sign correlations in heartbeat fluctuations *Phys. Rev. Lett.* **86** 1900–3

Babloyantz G A, Salazar J M and Nicolis C 1985 Evidence of chaotic dynamics of brain activity during the sleep cycle *Phys. Lett.* A **111** 152–6

Bazzani A, Bevilacqua A, Bollini D, Brancaccio R, Campanini R, Lanconelli N, Riccardi A and Romani D 2001 An SVM classifier to separate false signals from microcalcifications in digital mammograms *Phys. Med. Biol.* **46** 1651–63

Brown T E, Beightol L A, Koh J and Eckberg L D 1993 Important influence of respiration on human R-R interval power spectra is largely ignored *J. Appl. Physiol.* **75** 2310–7

Clark M E and Hirschman R 1980 Effects of paced respiration on affective responses during dental stress *J. Dent. Res.* **59** 1533–7

Clark M E and Hirschman R 1990 Effects of paced respiration on anxiety reduction in a clinical population *Biofeedback Self Regul.* **15** 273–84

Cooke W K, Cox J F, Diedrich A M, Taylor J A, Beightol L A, Ames IV J E, Hoag J B, Seidel H and Eckberg L D 1998 Controlled breathing protocols probe human autonomic cardiovascular rhythms *Am. J. Physiol.* **274** H709–18

Freedman R R and Woodward S 1992 Behavioral treatment of menopausal hot flushes: evaluation by ambulatory monitoring *Am. J. Obstet. Gynecol.* **67** 436–9

Glass L 2001 Synchronization and rhythmic processes in physiology *Nature* **410** 277–84

Ivanov P C and Lo C C 2002 Stochastic approaches to modelling of physiological rhythms *Modelling Biomedical Signals* ed G Nardulli and S Stramaglia (London: World Scientific) pp 28–51

Ivanov P C, Nunes Amaral L A, Goldberger A L, Havlin S, Rosenblum M B, Struzik Z and Stanley H E 1999 Multifractality in healthy heartbeat dynamics *Nature* **399** 461–5

Ivanov P C, Nunes Amaral L A, Goldberger A L, Havlin S, Rosenblum M G, Struzik Z and Stanley H E 2001 From $1/f$ noise to multifractal cascades in heartbeat dynamics *Chaos* **11** 641–52

Jauregui-Renaud K, Marquez M F, Hermosillo A G, Sobrino A, Lara J L, Kostine A and Cardenas M 2003 Paced breathing can prevent vasovagal syncope during head-up tilt testing *Can. J. Cardiol.* **19** 698–700

Kantz H and Schreiber T 1997 *Nonlinear Time Series Analysis* (Cambridge: Cambridge University Press)

Lehnertz K and Elger C E 1998 Can epileptic seizures be predicted? Evidence from nonlinear time series analyses of brain electrical activity *Phys. Rev. Lett.* **80** 5019–22

Luntz A and Brailovsky V 1969 On estimation of characters obtained in statistical procedure of recognition *Tecnicheskaya Kibernetica* **3** (in russian)

Marrone A, Polosa A D, Scioscia G, Stramaglia S and Zenzola A 1999 Multiscale analysis of blood pressure signals *Phys. Rev.* E **60** 1088–91

Mukherjee S, Rifkin R and Poggio T 2002 Regression and classification with regularization *Lectures Notes in Statistics: Nonlinear Estimation and Classification, Proc. MSRI Workshop* vol 171 ed D D Denison, M H Hansen, C C Holmes, B Mallick and B Yu (Berlin: Springer) pp 107–24

Nunes Amaral L A, Goldberger A L, Ivanov P C and Stanley H E 1998 Scale-independent measures and pathologic cardiac dynamics *Phys. Rev. Lett.* **81** 2388–91

Nunes Amaral L A, Ivanov P C, Aoyagi N, Hidaka I, Tomono S, Goldberger A L, Stanley H E and Yamamoto Y 2001 Behavioral-independent features of complex heartbeat dynamics *Phys. Rev. Lett.* **86** 6026–29

Peng C K, Mietus J, Hausdorff J M, Havlin S, Stanley H E and Goldberger A L 1993 Long-range anticorrelations and non-Gaussian behavior of the heartbeat *Phys. Rev. Lett.* **70** 1343–46

Pinna G D, Gobbi E, Maestri R, Robbi E, Fanfulla F and La Rovere M T 2003 Effect of paced breathing on cardiovascular variability parameters *IEEE-EMBS Asian-Pacifical Conf. on Biomedical Engineering*

Pinna G D, Maestri R, Raczak G and La Rovere M T 2002 Measuring baroreflex sensitivity from the gain function between arterial pressure and heart period *Clin. Sci. (Lond)* **103** 81–8

Pomortsev A V, Zubakhin A A, Abdushkevitch V G and Sedunova L F 1998 *Proc. XVII Congress of Physiologists of Russia* ed G A Kuraev (Rostov: Rostov State University) p 316

Poon C S and Merrill C K 1997 Decrease of cardiac chaos in congestive heart failure *Nature* **389** 492–5

Prokhorov M D, Ponomarenko V I, Gridnev V I, Bodrov M B and Bespyatov A B 2003 Synchronization between main rhythmic processes in the human cardiovascular system *Phys. Rev.* E **68** 041913–22

Riesz F and Nagy B S 1955 *Functional Analysis* (New York: Ungar)

Rzeczinski S, Janson N B, Balanov A G and McClintock P V E 2002 Regions of cardiorespiratory synchronization in humans under paced respiration *Phys. Rev.* E **66** 051909–17

Schafer C, Rosenblum M G and Abel H H 1998 Heartbeat synchronized with ventilation *Nature* **392** 239–40

Schiek M, Drepper F R, Engbert R, Abel H H and Suder K 1997 Transition between two different cardiorespiratory synchronization regimes during paced respiration *Jahreskongress der DPG, Rostock* **76**

Shalizi C R, Shalizi K L and Haslinger R 2004 Quantifying self-organization with optimal predictors *Phys. Rev. Lett.* **93** 118701–4

Swets J A 1988 Measuring the accuracy of diagnostic systems *Science* **240** 1285–93

Tass P, Rosenblum M G, Weule J, Kurths J, Pikovsky A, Volkmann J, Schnitzler A and Freund H-J 1998 Detection of n:m phase locking from noisy data: application to magnetoencephalography *Phys. Rev. Lett.* **81** 3291–4

Thurner S, Feurstein M C and Teich M C 1998 Multiresolution wavelet analysis of heartbeat intervals discriminates healthy patients from those with cardiac pathology *Phys. Rev. Lett.* **80** 1544–7

Vapnik V 1998 *Statistical Learning Theory* (New York: Wiley)