Variability analysis of the respiratory volume based on non-linear prediction methods

P. Caminal[1] L. Domingo[1] B.F. Giraldo[1] M. Vallverdú[1]
S. Benito[2] G. Vázquez[2] D. Kaplan[3]

[1]Biomedical Engineering Research Centre, Departament ESAII, Technical University of Catalonia, Spain
[2]Department of Intensive Care Medicine, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
[3]Department of Mathematics & Computer Science, Macalester College, Minnesota, USA

Abstract—*This work proposed and studied a method of automatically classifying respiratory volume signals as high or low variability by means of non-linear analysis of the respiratory volume. The analysis used volume signals generated by the respiratory system to construct a model of its dynamics and to estimate the quality of the predictions made with the model. Different methods of prediction evaluation, prediction horizons and embedding dimensions were also analysed. Assessment of the method was made using a database that contained 40 respiratory volume signals classified using clinical criteria into two classes: low or high variability. The results obtained using the method of surrogate data provided evidence of non-linear determinism in the respiratory volume signals. A discriminant analysis carried out using non-linear prediction variables classified the respiratory volume signals with an accuracy of 95%.*

Keywords—*Respiratory pattern variability, Non-linear prediction methods, Pressure support ventilation*

Med. Biol. Eng. Comput., 2004, **42**, 86–91

1 Introduction

THE POSSIBLE causes of breath-to-breath variability in the pattern of breathing have been discussed (BRUCE, 1996*a*; *b*; KHOO, 2000). Stochastic processes or dynamic behaviours of the autonomic nervous system (ANS) can produce this variability (MODARRESZADEH *et al.*, 1990; JUBRAN *et al.*, 1997; BRACK *et al.*, 2002). Analysis of respiratory variability provides a new tool to study the action of chemoreflexes without the application of external stimuli (VAN DEN AARDWEG and KAREMAKER, 2002). Determination of the variability of the respiratory volume also enables us to ascertain the ability of patients to control the mean tidal volume in response to alterations in respiratory demand (WRIGGE *et al.*, 1999).

Recently, it has been described that respiratory variability was reduced in patients with restrictive lung disease, compared with that of healthy subjects (BRACK *et al.*, 2002). One of the most challenging problems in intensive care (TOBIN, 2001) is the process of discontinuing mechanical ventilation, termed weaning. It has been hypothesised that the variability of the respiratory volume could be a convenient weaning criteria to reduce the number of patients not successfully weaned (DEL ROSARIO *et al.*, 1997).

Correspondence should be addressed to Prof. Pere Caminal; email: caminal@creb.upc.es

Paper received 9 June 2003 and in final form 7 October 2003
MBEC online number: 20043846

The traditional techniques of data analysis in the time and frequency domains are often not sufficient to characterise the complex dynamics of respiration. Various attempts have been reported to apply the concept of non-linear dynamics to the analysis of complex physiological systems (RIGNEY *et al.*, 1992; TURCOTT and TEICH, 1996; ACHERMANN, 1994) and to distinguish between variations that are random and those that are deterministic. The non-linear behaviour and time delays of the respiratory mechanisms of the ANS, together with muscle activity and the lungs, can introduce non-stochastic variability into the respiratory system. In this way, several studies have evidenced the non-linear, dynamic behaviour of the respiratory system. Several methods describing the non-linear deterministic variability of physiological time series have been proposed: correlation dimension, Lyapunov exponents, Kolmogorov–Sinai entropy etc. (BRUCE and DAUBENSPECK, 1995; SMALL *et al.*, 1999; AKAY *et al.*, 2002). SCHREIBER and SCHMITZ (1997) showed that non-linear prediction is an excellent method for detecting non-linearity in signals where determinism has not been established previously. Other approaches can present limitations according to the fractal nature of the time series (SAMMON *et al.*, 1993; WESSEL *et al.*, 1998; TAPANAINEN *et al.*, 1999) or even can lead to misinterpretations of the data (SMALL *et al.*, 1999). Cardiorespiratory synchronisation in humans and non-linear analysis of heart rate and respiratory dynamics have also been analysed using a prediction framework (HOYER *et al.*, 1998; 2002; CENSI *et al.*, 2000).

In this work, non-linear prediction methods were applied to find a set of indices that effectively characterise the variability of

the respiratory volume. As respiratory volume can be measured non-invasively, these indices may be advantageous in future automatic diagnosis of patients.

2 Materials and methods

2.1 *Analysed data*

A group of 20 patients on weaning trials from mechanical ventilation were studied in the Department of Intensive Care Medicine at Santa Creu i Sant Pau Hospital. According to a protocol approved by the local ethics committee and with informed consent obtained, the patients were each placed under two different levels of pressure support ventilation (PSV), classified as low PSV (5 ± 2 cm H_2O) and high PSV (12 ± 2 cm H_2O). The database therefore contains respiratory volume signals with different variabilities, mainly owing to the fact that changes in pressure support are often associated with changes in variability. The respiratory volume signals were obtained by means of a respiratory inductive plethysmograph. Respiratory volume at each PSV level was recorded for 30 min, with a sampling frequency of 250 Hz, and resampled at 10 Hz for this study. The 40 recordings of 30 min were classified by medical doctors into two classes, low (CLV) or high (CHV) variability, using clinical criteria based on respiratory rate, minute ventilation and a rapid shallow breathing index (CAPDEVILA, 1998).

This work proposes a method of automatically classifying the volume signals in high (HV) or low (LV) variability that does not necessarily match low and high PSV levels. For out-of-sample evaluation, the 40 volume recordings were organised into two sets: a training set and a testing set. A training set was selected that included patients presenting both CLV and CHV levels when the PSV was changed (nine patients and 18 volume recordings). A volume recording was considered correctly classified when the automatic classification coincided with the classification made by the medical doctor, considered as the gold standard.

2.2 *Non-linear prediction*

Fig. 1 shows CLV and CHV signals. The CHV signal in this case is at a lower frequency and, qualitatively, displays greater irregularity both in the waveform of a single cycle and in the spacing of cycles. The amplitude range of the signals is approximately the same. We sought to quantify this irregularity by measuring the autoregressive predictability of the signal. The time series is used to construct a model of the dynamics; the model is then used to predict other signal segments. The resulting prediction error quantifies irregularity.

There are different ways to construct dynamic models from data. As all the state variables of the systems were not directly measured or even known, we used the lag embedding technique to represent the system's state variables. By embedding the scalar time series D_t, the following vector sequence was created:

$$\boldsymbol{D}_t = (D_t, D_{t-1}, \ldots, D_{t-(m-1)})$$

where m is the embedding dimension. Each $\boldsymbol{D}_t$ is a point in the m-dimensional embedding space, and the embedded time series can be regarded as a sequence of points, one point at each time t. Each point represents the state of the system at that time.

A deterministic data set sampled at discrete times can be described by a discrete-time map

$$\boldsymbol{D}_{t+1} = \boldsymbol{F}(\boldsymbol{D}_t)$$

which is, however, immediately applicable only if the mapping $\boldsymbol{F}$ is known. With $\boldsymbol{F}$ unknown, some assumptions about its properties have to be made. With the minimum assumption that the mapping $\boldsymbol{F}$ is continuous, the following prediction scheme

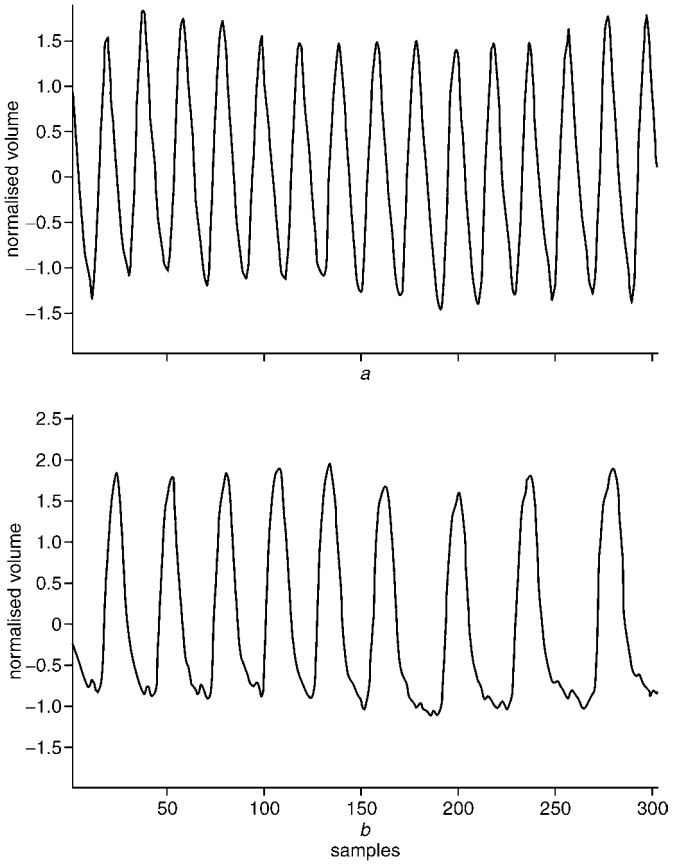

Fig. 1 *Respiratory volume recordings classified by medical doctors as (a) low variability (CLV) and (b) high variability (CHV)*

can be constructed (KANTZ and SCHREIBER, 2000). This method implements a non-linear regression model by stitching smoothly together a large number of locally linear models. The method works as follows: to predict the future state $\boldsymbol{D}_{t+1}$ given the present state $\boldsymbol{D}_t$, the state that is closest to $\boldsymbol{D}_t$ with respect to some norm is searched. Let us say that this closest point has time index a. The definition of determinism is that future events are set causally by past events. $\boldsymbol{D}_t$ describes the past events to D_{t+1}. Similarly $\boldsymbol{D}_a$ describes the past events to the measurement D_{a+1}. If $\boldsymbol{D}_t$ is close to $\boldsymbol{D}_a$, and if the system is deterministic, then it is expected that D_{a+1} will also be close to D_{t+1}. In the same way, D_{a+h} will be used as a predictor of D_{t+h} that will be called P_{t+h}.

Every measurement of a continuous quantity is only valid up to some finite resolution, and this fact has to be taken into account. The finite resolution implies that looking for the single closest state is no longer the best that can be done, as interpoint distances are contaminated with an uncertainty. All points within a close region in phase space have to be considered to be equally good predictions *a priori*. Then the proposed prediction algorithm to be used forms a neighbourhood $U(\boldsymbol{D}_t)$ around the point $\boldsymbol{D}_t$. For all points $\boldsymbol{D}_{a_i} \in U(\boldsymbol{D}_t)$, that is, all points close to $\boldsymbol{D}_t$, look up the individual predictions $\boldsymbol{D}_{a_i+h}$. Then the matrix $\boldsymbol{H}$ of the application $\{\boldsymbol{D}_{a_i+h}\} = \boldsymbol{H}\{\boldsymbol{D}_{a_i}\}$ is obtained that transforms the points of the neighbourhood $U(\boldsymbol{D}_t)$ into their predictions. Finally, the prediction $\boldsymbol{P}_{t+h}$ is obtained by applying the matrix $\boldsymbol{H}$ to the vector $\boldsymbol{D}_t$. Two ways have been considered to define the neighbourhood

(i) the neighbours inside a hypersphere of radius ε around point $\boldsymbol{D}_t$
(ii) the K neighbours closest to point $\boldsymbol{D}_t$.

Given a method for making a prediction P_{t+h}, an actual measurement of D_{t+h} is needed to decide if the prediction is good or bad. The difference between P_{t+h} and D_{t+h} is the prediction error, which informs us about the quality of the

prediction. As a single prediction could be good or bad just by chance, to give a more meaningful indication of the determinism in the data, an average of many prediction errors should be taken.

Two different methods have been considered to define this indication of determinism: cross-prediction and leave-one-out auto-prediction. In the cross-prediction approach, the time series is broken into M segments. For each of the M segments, one at a time, the model is fitted, and then residuals are calculated on each of the other segments. The residuals are summarised by one number, the mean absolute value. The result is an $M \times M$ matrix of cross-predictabilities. In this study, the respiratory volume data set at each PSV level, which contains 18 000 samples, has been divided into $M = 3$ segments of 6000 samples. In this case, the 3×3 matrix has 6 entries (the diagonal elements that correspond to self-prediction are not computed), and their mean value is computed in each patient for each PSV level.

In the leave-one-out auto-prediction, the time series of length N is modelled N different times: for each model, a single data point is left out when the model is fitted, and the residual for the model is computed only for the left-out data point. The result is a set of residuals, one for each point, that provide an estimate of the prediction error of a model. In this study, the respiratory volume data set at each PSV level was divided into nine subsets of $N = 2000$ samples. In this way, the mean prediction error related to each patient for each PSV level corresponded to the mean absolute value of the prediction errors in the nine subsets.

A preprocessing step was applied to each respiratory volume data set to improve the analysis of the results. Each respiratory volume signal was normalised by subtraction by its mean value and division by its standard deviation. Figs 2*a* and *b* show the actual measurements and predictions for the respiratory volume of a patient with clinically labelled low and high variabilities (CLV and CHV), respectively. The different quality of the prediction is shown comparing CLV and CHV.

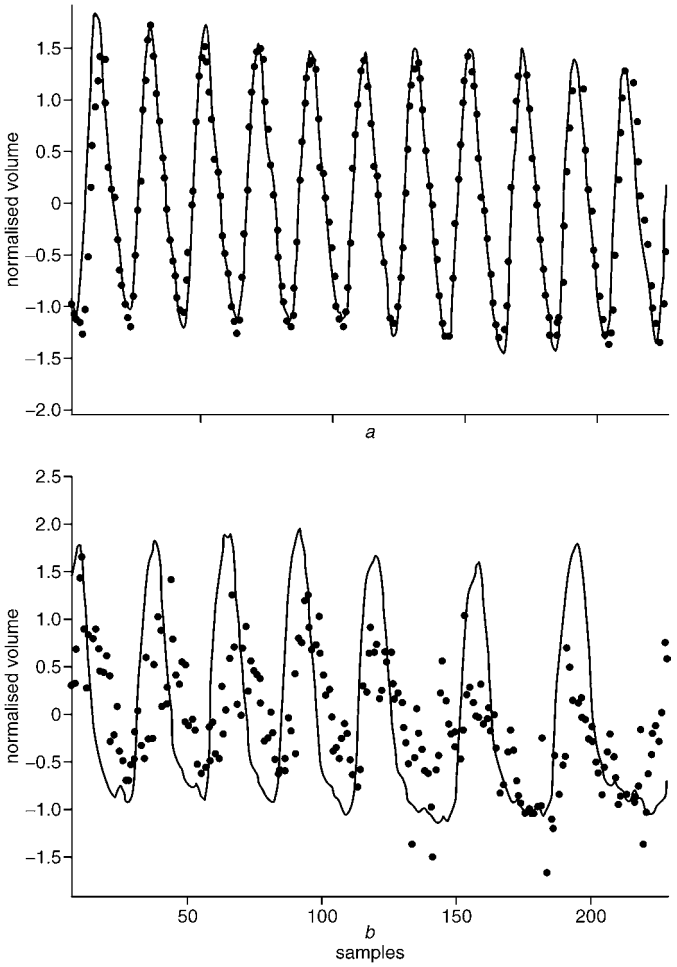

Fig. 2 *(—) Actual measurement and (•) one-step prediction of respiratory volume of patient with clinically labelled (a) low and (b) high variabilities. Mean respiratory period has been selected as prediction horizon*

2.3 *Parameter setting*

The first analysis related to the non-linear prediction was performed to choose between auto-prediction or cross-prediction methodologies. Three patients (CRR, MMX and SAT), who clinically presented two different variability levels (CLV and CHV) when the PSV was changed, were randomly selected for the analysis. An embedding dimension $m = 2$ was considered. Two kinds of neighbourhood were analysed: the neighbours inside a hypersphere of radius $\varepsilon = 0.2$ and the $K = 20$ closest neighbours. Tables 1 and 2 present, as an example, the values obtained in patient CRR using the neighbours inside a hypersphere and the K closest neighbours, respectively. In the three patients analysed, the auto-prediction methodology presented the best statistical by significant differences (p-value) when CLV and CHV signals were compared. This methodology has therefore been selected for the following steps.

To decide the best kind of neighbourhood to discriminate the different irregularities of the respiratory volume, in low and high variabilities, the following neighbourhoods were considered: the neighbours inside hyperspheres of radius $\varepsilon = 0.1$, 0.2 and 0.3 and the $K = 20$ closest neighbours. The same three patients were analysed, and an embedding dimension $m = 2$ was considered. Table 3 presents, as an example, the values obtained in patient CRR. In the three patients analysed, the statistical significance (p-value) obtained when the CLV and CHV signals were compared was found not to be dependent on the different neighbourhood methodology. Then, as the radius of the hyperspheres could be dependent on the embedding dimension, the K closest neighbours methodology was selected for the following steps.

The following analysis was performed to select the best prediction horizon h. For each patient and for each PSV level, the mean respiratory period was calculated. This mean respiratory period translated to sample units is called h_{Tot}. Three prediction horizons were considered: 0.5 h_{Tot},h_{Tot} and 2 h_{Tot}. The three patients were analysed, and the embedding dimension $m = 2$ was considered. Table 4 presents, as an example, the values obtained in patient CRR using the different prediction horizons. In the three patients analysed, the statistical significance (p-value) obtained when CLV and CHV signals were compared was found not to be dependent on the considered h value. A prediction horizon of h_{Tot} was selected for the following steps.

2.4 *Non-linear determinism in the respiratory volume signal*

The typically slower frequency of the CHV signals suggests that a frequency-domain analysis using, for example, power spectrum analysis, could be effective at performing the discrimination. To assess to what extent our non-linear prediction method processes information not accessible to the linear method, we used the method of surrogate data (THEILER *et al.*, 1992; SCHREIBER and SCHMITZ, 1996). This method involves

Table 1 Mean ± standard deviation for mean prediction error of patient CRR with m = *2, ε = 0.20, when leave-one-out auto-prediction and cross-prediction are considered and statistical significance (p-value) when low and high variability levels are compared*

	CLV	CHV	p-Value
Leave-one-out auto-prediction	0.41 ± 0.06	0.82 ± 0.11	0.008
Cross-prediction	0.45 ± 0.04	0.90 ± 0.06	0.028

Table 2 Mean ± standard deviation for mean prediction error of patient CRR with m = 2, K = 20, when leave-one-out auto-prediction and cross-prediction are considered

	CLV	CHV	*p*-Value
Leave-one-out auto-prediction	0.36 ± 0.05	0.76 ± 0.09	0.008
Cross-prediction	0.43 ± 0.03	0.88 ± 0.06	0.027

Table 3 Mean ± standard deviation for mean prediction error of patient CRR with m = 2 when different radii ε of hyperspheres and K = 20 closest neighbours are considered

	CLV	CHV	*p*-Value
$\varepsilon = 0.1$	0.41 ± 0.05	0.81 ± 0.09	0.008
$\varepsilon = 0.2$	0.41 ± 0.06	0.82 ± 0.11	0.008
$\varepsilon = 0.3$	0.38 ± 0.06	0.81 ± 0.11	0.008
K neighbours	0.36 ± 0.05	0.76 ± 0.09	0.008

Table 4 Mean ± standard deviation for mean prediction error of patient CRR with m = 2 when different prediction horizons h *are considered*

	CLV	CHV	*p*-Value
0.5 h_{Ttot}	0.34 ± 0.07	0.67 ± 0.09	0.008
h_{Ttot}	0.36 ± 0.05	0.76 ± 0.09	0.008
2 h_{Ttot}	0.55 ± 0.07	0.81 ± 0.09	0.008

generating synthetic volume signals, called surrogate data, with the same Fourier spectra, mean, standard deviation and other percentiles as the original data. All the information that could be accessed by a linear power spectrum analysis, whatever form that analysis might take, is contained in the surrogate data. The algorithm to generate this surrogate data is based on the null hypothesis that the data come from a stationary linear process with Gaussian white noise inputs.

A set of surrogate data was generated for each volume signal tested. For all the signals (original data and surrogate data), a non-linear index was computed. Then, a statistical test was applied between the set of surrogate data and the original data.

If the null hypothesis was rejected, this suggested that the original data were due to a non-linear deterministic process and/or non-Gaussian inputs or non-stationarity. In the case of the signals analysed in this study, ten series of surrogate data were generated for each of the volume signals of the three patients CRR, MMX and SAT. The non-linear index selected was the mean prediction error.

2.5 *Discriminant analysis*

A discriminant analysis was applied to obtain a discriminant function that would enable the automatic classification of the volume signals as high (HV) on low (LV) variability. To know the best variables to be introduced in the discriminant analysis, a previous non-parametric analysis of variance test (Mann–Whitney) was used to analyse statistically the differences between the respiratory volume signals with CLV and CHV. Different variables from the classical time-domain analysis and from the described non-linear prediction analysis were considered.

In the classical time-domain analysis of the respiratory volume signal, for each patient and for each PSV level, the following time series were obtained: breath duration T_{tot}, inspiration time T_i and tidal volume V_t, related to the respiratory cycles of each 30 min recording. From these time series, the mean values of T_{tot}, T_i and V_t were obtained ($\overline{T_{tot}}$, $\overline{T_i}$ and $\overline{V_t}$).

From the respiratory volume signals training set, different discriminant functions were obtained and subsequently validated with the testing set. The validation was performed by comparing the results obtained from the discriminant functions with the classification made by medical doctors.

3 Results

Time-domain analysis of the respiratory volume signal was performed previously. Table 5 shows the results obtained with the mean values of T_{tot}, T_i and V_t when low and high variability levels, defined using clinical criteria, were compared in all 20 patients. $\overline{T_{tot}}$ and $\overline{T_i}$ variables present statistically significant differences ($p < 0.0005$ and $p = 0.03$, respectively). This change in $\overline{T_{tot}}$ reflects the slow frequency of the CHV signals and the differences between the populations of signals. However, as the populations overlap substantially, the classification of individual signals will not be very accurate.

Table 6 shows the results obtained when the surrogate data method was applied to the respiratory volume signals of CLV and CHV in the three selected patients CRR, MMX and SAT, who had both CLV and CHV recordings. The mean prediction error *mpe* of the original signal Q_D and the mean value ± standard deviation of the *mpe* of the surrogate data ($\mu_H \pm \sigma_H$) are presented. For both low and high variability recordings of the three patients, the respiratory volume signals of the patients analysed had significant differences with respect to the surrogate data generated, and so the null hypothesis could be rejected.

So that we can analyse the level of irregularity in the respiratory volume signals related to high variability in comparison with that related to low variability, Table 7 shows the mean prediction errors *mpe* obtained for $m = 2$ when all the patients are considered. The results show a statistically significant difference ($p < 0.0005$) between both groups (Mann–Whitney test).

The role of the embedding dimension m in the prediction errors was analysed in all the patients for each of the PSV levels. Fig. 3 shows, as an example, the relationship between the mean prediction error and the embedding dimension for the patient CRR. The line labelled CRR20 belongs to the CHV signal, and CRR06 belongs to the CLV signal.

Another way to characterise predictability involves finding the embedding dimension needed to model the dynamics of the patients with a low prediction error. For example, in patient CRR (Fig. 3), an embedding dimension $m = 8$ is needed to obtain a mean prediction error below 0.4 when analysing the CHV signal, whereas $m = 2$ is enough to obtain the same prediction error for the CLV signal. The values of the embedding dimension m needed to model the dynamics of the signals with a

Table 5 Mean ± standard deviation for classical time-domain analysis variables when low and high variability levels are compared in all 20 patients

	CLV	CHV	*p*-Value
$\overline{T_{tot}}$	2.48 ± 0.65	3.63 ± 1.04	<0.0005
$\overline{T_i}$	0.88 ± 0.12	1.04 ± 0.30	0.030
$\overline{V_t}$	466 ± 195	601 ± 265	ns

Table 6 Values of mean prediction error for volume signals and surrogate data with statistical significance

	Q_D	$\mu_H \pm \sigma_H$	*p*-Value
CRR–CLV	0.36	0.49 ± 0.01	<0.0005
CRR–CHV	0.72	0.75 ± 0.01	<0.0005
MMX–CLV	0.24	0.31 ± 0.01	<0.0005
MMX–CHV	0.33	0.39 ± 0.01	<0.0005
SAT–CLV	0.31	0.40 ± 0.01	<0.0005
SAT–CHV	0.70	0.79 ± 0.01	<0.0005

Table 7 Mean ± standard deviation of mean prediction errors mpe and embedding dimensions m needed to model dynamics of patients with reduced mean prediction error e of 0.35, 0.40 and 0.45 (me35, me40 and me45, respectively)

	CLV	CHV	*p*-Value
mpe	0.35 ± 0.09	0.63 ± 0.08	<0.0005
*me*35	3.3 ± 2.0	6.8 ± 1.7	<0.0005
*me*40	2.4 ± 0.7	5.9 ± 1.5	<0.0005
*me*45	2.1 ± 0.3	5.1 ± 1.4	<0.0005

Table 8 Validation, using test set, of discriminant functions of single variables obtained from training set

	Critical threshold	False HV	False LV	Accuracy (%)
$\overline{T_{tot}}$	2.86	4	1	77
$\overline{T_i}$	0.97	4	3	68
mpe	0.50	1	0	95
*me*35	5.3	0	2	91
*me*40	4.4	1	1	91
*me*45	3.9	3	1	82

prediction error *e* of 0.35, 0.40, 0.45 (*me*35, *me*40 and *me*45, respectively) have been calculated. Table 7 shows the values of *me*35, *me*40 and *me*45 when all the patients were analysed. The embedding dimension needed to model the dynamics of the patients with a low prediction error showed a statistically significant difference ($p < 0.0005$) between low and high variability signals (Mann–Whitney test).

The aim of the last part of this study was to obtain discriminant functions able to discriminate low and high respiratory pattern variability. From the respiratory volume signals of the training set, different discriminant functions were constructed using each single variable presented in Table 7 (*mpe*, *me*35, *me*40, *me*45), as well as $\overline{T_i}$ and $\overline{T_{tot}}$ variables. Table 8 shows the critical threshold of the discriminant functions, related to each one of the considered single variables, and the results achieved during the evaluation process with the 22 respiratory volume signals of the testing set. In this process, a signal is considered false HV when the discriminant function classifies it as high variability (HV) when it was considered by the medical doctor to be low variability (CLV), and a signal is considered as false LV when the discriminant function classifies it as low variability (LV), when it was considered by the medical doctor to be high variability (CHV). Accuracy is the percentage of volume signals correctly classified. The variables obtained with the non-linear prediction methodology present better discriminant results than the best variable proposed from the time-domain analysis. Table 9 shows the results obtained using discriminant functions of two variables. The mean prediction error *mpe* and the *mpe* combined with the embedding dimension needed to obtain an *mpe* of 0.40 or 0.45 achieve an accuracy of 95%.

4 Discussion and conclusions

To analyse respiratory pattern variability in respiratory volume signals, non-linear prediction methods were applied.

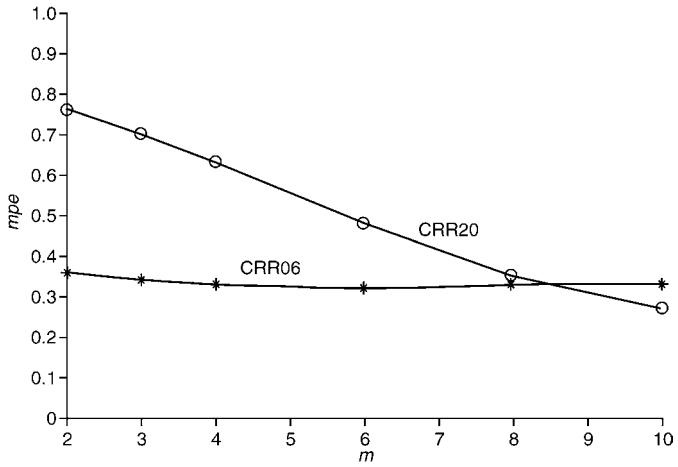

Fig. 3 *Prediction errors mpe obtained as function of embedding dimension m for patient CRR. Lines labelled CRR20 and CRR06 belong to CHV and CLV signals, respectively*

The volume time series were used to construct a model of the respiratory system dynamics, and the accuracy of the predictions made from the model were analysed. Two different ways were considered to define the indication of determinism: cross-prediction and leave-one-out auto-prediction. The auto-prediction methodology was selected because it presented the best statistically significant differences when CLV and CHV signals were compared. Two kinds of neighbourhood were analysed: the neighbours inside a hypersphere of radius ε and the K neighbours closest to a point in the phase space. The K closest neighbours methodology was selected because it produced the same statistical significance level as the neighbours inside a hypersphere, and this last method presented the inconvenience that the radii of the hyperspheres could be dependent on the embedding dimension. The incidence of different prediction horizons h was also considered. As the results were found not to be dependent on the considered h value, the mean respiratory period was selected as the prediction horizon.

Highly statistically significant differences were obtained when the mean prediction error *mpe* of the volume signals clinically classified as low variability (0.35 ± 0.09) were compared with high variability signals (0.63 ± 0.08): $p < 0.0005$. The embedding dimension needed to model the dynamics of the system with a low prediction error is also a good parameter to discriminate different respiratory patterns.

The results obtained using the surrogate data method mean that the non-linear prediction method detects signs of non-linearity, non-stationarity or non-Gaussianity in the signals. However, note that the prediction errors for the surrogate data in the different classes of CHV and CLV signals follow roughly the same pattern of variability as for the original data. That is, there is a lower non-linear prediction error for surrogates from CLV signals than for surrogates from CHV signals. As the surrogate data have, by construction, no statistically identifiable non-linear, non-stationary, or non-Gaussian components, this suggests that it may be possible to find some linear analysis method that can perform a discrimination between CLV and CHV similar to the one using non-linear prediction. This does not necessarily mean, however, that the physiological mechanisms generating the linear structures are themselves linear. The hypotheses on the physiological mechanisms governing respiratory volume variability are based on the non-linear dynamic interactions between various components of the respiratory control system, such as the lung vagal afferents and the respiratory pattern generator, or through the propagation of stochastic disturbances around the chemoreflex loops (Bruce, 1996*a*; Khoo, 2000).

Table 9 Validation, using test set, of discriminant functions of two variables obtained from training set

	False HV	False LV	Accuracy (%)
mpe and *me*35	2	1	86
mpe and *me*40	0	1	95
mpe and *me*45	0	1	95

The discriminant analysis carried out with the training set, when the mean prediction error was used, obtained discriminant functions able to classify, with an accuracy of 95%, the test respiratory volume signals, whereas the discriminant analysis using classical time-domain variables presented lower accuracy (77%). These results indicate that non-linear prediction is a promising methodology to study respiratory pattern variability. It should be validated by a larger number of patients, especially to investigate further the discriminant functions.

The clinical relevance of such a method to discriminate respiratory volume variability is related to the study of the action of chemoreflexes without application of external stimuli and the analysis of the ability of patients to control mean tidal volume in response to alterations in respiratory demand. Furthermore, this method could be a convenient weaning criterion to reduce the number of patients not successfully weaned.

Acknowledgments—This work was partially supported by CICYT grant TIC2001-2167-C02-01 and FIS grant 99/0226.

References

ABARBANEL, H. D., BROWN, R., and KADTKE, J. B. (1989): 'Prediction and chaotic nonlinear systems: Methods for time series with broadband Fourier spectra', *Phys. Rev. A*, **41**, pp. 1782–1807

ACHERMANN, P., HARTMANN, R., GUNZIGER, A., GUGGENBUHL, W., and BROBELY, A. A. (1994): 'All-night sleep EEG and artificial stochastic control signals have similar correlation dimensions', *Electroencephalogr. Clin. Neurophisiol.* , **90**, pp. 384–387

AKAY, M., LIPPING, T., MOODIE, K., and HOOPES, P. J. (2002): 'Effects of hypoxia on the complexity of respiratory patterns during maturation', *Early Hum. Devel.*, **70**, pp. 55–71

BRACK, T., JUBRAN, A., and TOBIN, M. J. (2002): 'Dyspnea and decreased variability of breathing in patients with restrictive lung disease', *Am. J. Respir. Crit. Care Med.*, **165**, pp. 1260–1264

BRUCE, E. N., and DAUBENSPECK, J. A. (1995): 'Mechanisms and analysis of respiratory variability', in 'Control of breathing' (Marcel Dekker, 1995), pp. 285–314

BRUCE, E. N. (1996*a*): 'Temporal variations in the pattern of breathing', *J. Appl. Physiol.*, **80**, pp. 1079–1087

BRUCE, E. N. (1996*b*): 'Measures of respiratory pattern variability', in 'Bioengineering approaches to pulmonary physiology and medicine' (Plenum Press, 1996), pp. 149–160

CAPDEVILA, X., PERRIGAULT, P. F., RAMONATXO, M., ROUSTAN, J. P., PERAY, P., FRANCOISE, A., and PREFAUT, C. (1998): 'Changes in breathing pattern and respiratory muscle performance parameters during difficult weaning', *Crit. Care Med.*, **26**, pp. 79–87

CENSI, F., CALCAGNINI, G., LINO, S., SEYDNEJAD, S. R., KITNEY, R. I., and CERUTTI, S. (2000): 'Transient phase locking patterns among respiration, heart rate and blood pressure during cardiorespiratory synchronisation in humans', *Med. Biol. Eng. Comput.*, **38**, pp. 416–426

DEL ROSARIO, N., SASSOON, C. S., CHETTY, K. G., GRUER, S. E., and MAHUTTE, C. K. (1997): 'Breathing pattern during acute respiratory failure and recovery', *Eur. Respir. J.*, **10**, pp. 2560–2565

HOYER, D., KAPLAN, D. T., SHAAF, F., and EISELT, M. (1998): 'Determinism in bivariate cardiorespiratory phase space sets. How to detect nonlinear coordinations', *IEEE Eng. Med. Biol.*, **17**, pp. 26–31

HOYER, D., LEDER, U., HOYER, H., POMPE, B., SOMMER, M., and ZWIENER, U. (2002): 'Mutual information and phase dependencies: Measures of reduced non-linear cardio-respiratory interactions after myocardial infarction', *Med. Eng. Phys.*, **24**, pp. 33–43

JUBRAN, A., GRANT, B. J. B., and TOBIN, M. J. (1997): 'Effect of hyperoxic hypercapnia on variational activity of breathing', *Am. J. Respir. Crit. Care Med.*, **156**, pp. 1129–1139

KANTZ, H., and SCHREIBER, T. (2000): 'Determinism and predictability', in 'Nonlinear time series analysis' (Cambridge University Press, 2000), pp. 42–57

KAPLAN, D., and GLASS, L. (1995): 'Characterizing chaos', in 'Understanding nonlinear dynamics' (Springer-Verlag, 1995), pp. 314–338

KHOO, M. C. K. (2000): 'Determinants of ventilatory instability and variability', *Respir. Physiol.*, **122**, pp. 167–182

MODARRESZADEH, M., BRUCE, E. N., and GOTHE, B. (1990): 'Non-random variability in respiratory cycle parameters of humans during stage 2 sleep', *J. Appl. Physiol.*, **69**, pp. 630–639

RIGNEY, D. R., OCASIO, W. C., CLARK, K. P., WEI, J. Y., and GOLDBERGER, A. L. (1992): 'Deterministic mechanism for chaos and oscillations in heart rate and blood pressure', *Circulation*, pp. 651–659

SAMMON, M., ROMANIUK, J. R., and BRUCE, E. (1993): 'Bifurcations of the respiratory pattern associated with reduced lung volume in the rat', *J. Appl. Physiol.*, **75**, pp. 887–901

SCHREIBER, T., and SCHMITZ, A. (1996): 'Improved surrogate data for nonlinearity tests', *Phys. Rev. Lett.*, **77**, pp. 635–638

SCHREIBER, T., and SCHMITZ, A. (1997): 'Discrimination power of measures for nonlinearity in a time series', *Phys. Rev. E*, **55**, pp. 5443–5447

SMALL, M., JUDD, K., LOWE, M., and STICK, S. (1999): 'Is breathing in infants chaotic? Dimension estimates for respiratory patterns during quiet sleep', *J. Appl. Physiol.*, **86**, pp. 359–376

TAPANAINEN, J. M., SEPPäNEN, M. D., LAUKKANEN, R., LOIMAALA, A., and HUIKURI, H. V. (1999): 'Significance of the accuracy of RR interval detection for the analysis of new dynamic measures of heart rate variability', *Ann. Noninvas. Electrocardiol.*, **4**, pp. 10–18

THEILER, J., EUBANK, S. E., LONGTIN, A., GALDRIKIAN, B., and FARMER, D. (1992): 'Testing for nonlinearity in time series: the method of surrogate data', *Phys. D*, **58**, pp. 77–94

TOBIN, M. J. (2001): 'Advances in mechanical ventilation', *New Engl. J. Med.*, **344**, pp. 1986–1996

TURCOTT, R. G., and TEICH, M. C. (1996): 'Fractal character of the electrocardiogram: distinguishing heart failure and normal patients', *Ann. Biomed. Eng.*, **24**, pp. 269–293

VAN DEN AARDWEG, J. G., and KAREMAKER, J. M. (2002): 'Influences of chemoreflexes on respiratory variability in healthy subjects', *Am. J. Respir. Crit. Care Med.*, **165**, pp. 1041–1047

WESSEL, N., MEYERFELDT, U., SCHIRDEWAN, A., KURTHS, J., and VOSS, A. (1998): 'Short-term forecasting of life-threatening arrhythmias with finite time Lyapunov exponents', *Ann. Conf. IEEE Eng. Med. Biol. Soc.*, **20**, pp. 326–329

WRIGGE, H., GOLISCH, W., ZINSERLING, J., SYDOW, M., ALMELING, G., and BURCHARDI, H. (1999): 'Proportional assist versus pressure support ventilation: effects on breathing pattern and respiratory work of patients with chronic obstructive pulmonary disease', *Intens. Care Med.*, **25**, pp. 790–798

Authors' biographies

PERE CAMINAL is a Professor in the Department of Control Engineering (DCE), at the Technical University of Catalonia (UPC), Spain. His research interests include modelling of biological systems and biomedical signal processing.

LUIS DOMINGO is a PhD candidate in Biomedical Engineering at the Research Centre on Biomedical Engineering (CREB-UPC).

BEATRIZ F. GIRALDO received the PhD in Biomedical Engineering from the UPC in 1996. She is Assistant Professor in the DCE at UPC. Her research interests include biomedical signal processing.

MONTSERRAT VALLVERDÚ received the PhD in Biomedical Engineering from the UPC in 1993. She is Associate Professor in the DCE at UPC. Her research interests include complexity analysis of the hidden information in cardiac and respiratory signals.

SALVADOR BENITO is a Professor in the Department of Medicine, at the University Autonoma of Barcelona (UAB), Spain. His research interests include lung function, mechanical ventilation and critical care.

GUILLERMO VÁZQUEZ is a Professor in the Department of Medicine, at the University Autonoma of Barcelona (UAB). His research interests include lung function, mechanical ventilation and critical care.

DANIEL KAPLAN is Associate Professor in the Department of Mathematics and Computer Science at Macalester College in St. Paul, Minnesota, USA. His research interests include techniques of nonlinear dynamics time-series analysis.