Nonlinear measure of synchrony between blood oxygen saturation and heart rate from nocturnal pulse oximetry in obstructive sleep apnoea syndrome

D Álvarez1, R Hornero1, D Abásolo1, F del Campo2, C Zamarrón3, M López1
1 Biomedical Engineering Group, E.T.S.I. de Telecomunicación, University of Valladolid, Camino del Cementerio s/n, 47011, Valladolid, Spain
2 Hospital Universitario del Río Hortega, Servicio de Neumología, c/ Cardenal Torquemada s/n, 47010, Valladolid, Spain
3 Hospital Clínico Universitario, Servicio de Neumología, Travesía de la Choupana s/n, 15706, Santiago de Compostela, Spain

E-mail: dalvgon@ribera.tel.uva.es, robhor@tel.uva.es, danaba@tel.uva.es, fsas@telefonica.net, carlos.zamarron.sanz@sergas.es, miglop@tel.uva.es

Abstract
This study focuses on the analysis of the relationship between changes in blood oxygen saturation (\(\text{SaO}_2\)) and heart rate (HR) recordings from nocturnal pulse oximetry (NPO) in patients suspected of suffering from obstructive sleep apnoea (OSA) syndrome. Two different analyses were developed: a classical frequency analysis based on the magnitude squared coherence (MSC) and a nonlinear analysis by means of a recently developed measure of synchrony, the cross approximate entropy (cross-ApEn). A data set of 187 subjects was studied. We found significantly higher correlation and synchrony between oximetry signals from OSA Positive patients compared with OSA Negative subjects. We assessed the diagnostic ability to detect OSA syndrome of both the classical and nonlinear approaches by means of receiver operating characteristic (ROC) analyses with ten-fold cross-validation. The nonlinear measure of synchrony significantly improved the results obtained with classical MSC: 69.2% sensitivity, 90.9% specificity and 78.1% accuracy were reached with MSC, whereas 83.7% sensitivity, 84.3% specificity and 84.0% accuracy were obtained with cross-ApEn. Our results suggest that the use of nonlinear measures of synchrony could provide essential information from oximetry signals, which cannot be obtained with classical spectral analysis.

Keywords: Obstructive sleep apnoea, nocturnal pulse oximetry, blood oxygen saturation, heart rate, synchrony, magnitude squared coherence, nonlinear analysis, cross approximate entropy.

1. Introduction
Obstructive sleep apnoea (OSA) syndrome is a major sleep-related breathing disorder, affecting 1 to 5% of adult men and 2% of women in western countries (Young et al 2002). The gold standard diagnostic method for a definitive diagnosis of OSA is overnight polysomnography
(PSG). However, its relative high cost and complexity, along with the large number of people suspected of having this disease, limit its capacity as a diagnostic test for OSA (Whitelaw et al 2005). Due to its simplicity and low cost, nocturnal pulse oximetry (NPO) has been proposed as a substitute for PSG in the diagnostic assessment of patients with suspected sleep apnoea (Netzer et al 2001). NPO is a standard technique for monitoring the arterial blood oxygen saturation (SaO$_2$) in operating rooms, intensive care units and other critical care settings (Kirby et al 1997). Additionally, pulse oximeters accurately estimate heart rate (HR) at rest and during submaximal exercise (Iyriboz et al 1991) and are widely used for screening for the presence of congenital heart disease and cardiovascular malformations in newborns (Granelli et al 2008).

OSA is characterised by repetitive reduction or cessation of airflow due to partial or complete airway obstruction, leading to hypoxemia, bradycardia, arousals and fragmented sleep (Gilleminault et al 1978). The cardiovascular and pulmonary systems deliver oxygen to the tissues working in an integrated fashion: HR increases with inspiration and decreases with expiration. Moreover, under certain physiological conditions, HR is clearly a dependent variable of ventilation (Finley et al 1985). Apnoeic episodes during sleep are generally associated with profound bradycardia with cyclical variation of the HR during apnoea and resumption of breathing (Krahn et al 2006). Nightly episodes of hypoxia, arousals and swings in intrathoracic pressure due to OSA may lead to sustained elevation of blood pressure via pathophysiologic mechanisms that include chronically elevated sympathetic tone, alterations in baroreceptor function and cardiovascular remodelling (Young et al 2002). Additionally, recent studies relate OSA with secondary systemic hypertension (Peppard et al 2000).

The cardiovascular morbidity associated with OSA affects patients in a wide severity spectrum and single SaO$_2$ recordings do not reflect all the physiological aspects of this disease. NPO-based studies have been mainly focused on SaO$_2$ and few researchers have assessed the usefulness of HR and its relation with SaO$_2$ to characterise OSA syndrome. The synchronization between the increase in the very low frequency components of Holter ECG and episodes of sleep-related breathing disorders was suggested in the studies by Shiomi et al (1996) and Keyl et al (1997). The study by Zamarrón et al (2003) showed that spectral analysis of HR from NPO provided additional information to that obtained from SaO$_2$. Recently, Zamarrón et al (2005) showed that OSA patients presented a dynamic coordination between SaO$_2$ and HR from NPO in specific frequencies. In the present study, we investigate changes in simultaneously recorded SaO$_2$ and HR signals from NPO to search for additional and useful information in order to better characterise OSA syndrome. Our research is aimed at estimating the overall synchrony of each SaO$_2$ and HR pair, in order to test its possible usefulness in OSA diagnosis.

Spectral analysis has been commonly applied to investigate periodicities due to apnoea events in cardiorespiratory recordings (Penzel et al 2002, Zamarrón et al 2003). Additionally, spectral coherence has also been traditionally used to estimate similarities between physiological signals (Stam and van Dijk 2002, Zamarrón et al 2005). However, standard spectral estimators can be inconsistent and biased, especially in the presence of outliers and nonstationarities (Pincus 2000). Particularly, coherence estimates are not suitable to characterize nonstationary signals and only capture linear relations (Stam and van Dijk 2002).

Recent studies (Álvarez et al 2006, del Campo et al 2006, Zamarrón et al 2006, Hornero et al 2007, Álvarez et al 2007) have shown the usefulness of nonlinear analysis of SaO$_2$ and HR from NPO to help in the diagnosis of OSA syndrome. Other researchers (Aguirre et al 1999) have previously estimated different nonlinear models of the dynamics of sleep apnoea based on SaO$_2$, HR and respiration, suggesting the nonlinearity of such data. A recently developed nonlinear measure of asynchrony between time series, the cross approximate entropy (cross-ApEn), has demonstrated to be complementary and often superior to classical spectral and correlation based methods (Pincus 2001). It has been mainly applied to study hormone time series dynamics (Veldhuis et al 2001, Liu et al 2005).

In the present study we apply both classical and nonlinear measures of synchrony to SaO$_2$ and HR recordings from NPO. Our research is aimed to assess whether cross-ApEn could
Nonlinear measure of synchrony from pulse oximetry in sleep apnoea

improve the classical approach based on the magnitude squared coherence (MSC). Finally, we evaluate their advantages and limitations in the diagnostic assessment of OSA syndrome.

2. Data set and sleep studies

A total of 187 subjects (147 males and 40 females) with ages between 21 and 81 years old and body mass indexes (BMI) ranging between 19.53 and 46.51 kg/m² suspected of having OSA were included in the study. All subjects presented, at least, one of the following symptoms reported by the subject or a bed mate: daytime hypersomnolence, loud snoring, nocturnal choking and awakenings, or apnoic events. Sleep studies were carried out usually from midnight to 08:00 AM in the Sleep Unit of Hospital Clínico Universitario in Santiago de Compostela, Spain. The Review Board on Human Studies approved the protocol, and all subjects gave their informed consent to participate in the study. Sleep studies consisted on a conventional PSG with a simultaneously acquired NPO. All patients were clinically stable during the whole recording time. None of the patients involved in the study suffered from chronic heart failure or insulin-dependent diabetes mellitus. Additionally, none of the patients showed clinical evidence of autonomic dysfunction or neuropathy. Individuals presenting permanent or paroxysmal atrial fibrillation or a pacemaker were excluded from the study. Patients were free from any medication which could influence sleep studies. Table 1 summarizes the demographic and clinical data of the patients included in this study.

All patients underwent overnight PSG (Ultrasom Network, Nicolet, Madison, WI, U.S.A.). Each PSG register included electroencephalogram (EEG), electrocardiogram (ECG), electrooculogram (EOG), chin electromyogram (EMG), measurement of chest wall movement and airflow measurement (three-port thermistor). In the airflow measurement procedure, the thermistor may not be sensitive for detecting hypopneas, leading to an underestimation of the AHI (Series and Marc 1999). However, the thermistor is the most common method for defining breathing events based on a flow measurement (Flemons et al 2003). The PSG register was analyzed over periods of 30 s according to the Rechtschaffen and Kales rules (Rechtschaffen and Kales 1968). Apnoea was defined as the cessation of airflow for more than 10 seconds and hypopnoea as the reduction of respiratory flow for at least 10 seconds accompanied by a 4% or more decrease in the saturation of haemoglobin. The average apnoea-hypopnoea index (AHI) was calculated for hourly periods of sleep. Subjects with an AHI greater than or equal to 10 (AHI ≥ 10) were diagnosed as suffering from OSA. If the subject had less than 3 hours of total sleep, the sleep study was repeated. The initial patient population was divided into two groups using the AHI derived from the PSG. 111 patients (59.4%) diagnosed as suffering from OSA syndrome composed the OSA Positive group, with an average AHI of 40.07 ± 19.64 events/hour, an age of 58.30 ± 12.88 years and a BMI of 30.45 ± 4.92 kg/m². The remaining 76 patients (40.6%) composed the OSA Negative group, with an AHI of 2.04 ± 2.36 events/hour, a mean age of 57.57 ± 12.87 years and a BMI of 28.42 ± 6.02 kg/m².

Pulse oximeters quantify SaO₂ by measuring the absorption of two frequencies of light by perfused tissues: two LEDs of differing wavelengths are placed on one side of the finger, whereas a photodetector detect the transmitted light at the opposite side (Coetzee et al 2000). The absorption of light passing through the body varies as blood pulses into the tissue with each beat of the heart. The HR can be accurately estimated from the variations in the amount of
transmitted light. Oximetry signals were saved to separate files and processed off-line by means of MSC and cross-ApEn. Each pair of oximetric recordings was automatically scanned to remove artefacts and drops to zero due to poor contact from the finger probe, patient movements or bad regional circulation.

3. Methods

3.1. Coherence function analysis

The coherence function between two wide-sense stationary random processes \( x \) and \( y \) is equal to the cross power spectrum \( G_{xy}(f) \) divided by the squared root of the product of their auto power spectra (Carter 1987). Mathematically, the coherence function is defined by

\[
\gamma_{xy}(f) = \frac{G_{xy}(f)}{\sqrt{G_{xx}(f)G_{yy}(f)}}.
\]

The magnitude squared coherence (MSC) \( C_{xy}(f) \) is a normalized cross-spectral density function given by

\[
C_{xy}(f) = \left| \gamma_{xy}(f) \right|^2, \quad 0 \leq C_{xy}(f) \leq 1,
\]

which provides a measure of the degree of correlation between each of the individual frequency components of two signals (Hampson and Mallen 2006).

To estimate the cross power spectral density (cross-PSD) and the auto power spectral density (PSD), we used the nonparametric Welch’s method (Welch 1967). Firstly, the method divides the time series into overlapping segments, applies a smooth time weighting to each segment and computes the discrete Fourier transform (DFT). Finally, all DFTs are averaged to obtain the PSD estimate. The Hanning smooth weighting function and an overlapping percentage of 50% are widely used to improve the performance of the power spectral estimation (Carter 1987). The MSC analysis was performed using Matlab®. In this study, a 300-sample Hanning window with 50% overlap was used. Thus, every record was divided at least into 38 segments, which ensured that each MSC estimate was asymptotically unbiased (Carter 1987). 512-points DFTs were used to estimate PSDs and cross-PSDs.

3.2. Cross approximate entropy (cross-ApEn)

Cross approximate entropy (cross-ApEn) is a two parameter family of statistics introduced as a measure of asynchrony between two paired time series (Pincus 2000). It was proposed as a method to compare correlated sequences, suggesting its application to physiological signals like HR and respiratory rate (Pincus and Singer 1996). Cross-ApEn is generally applied to compare sequences from two distinct yet intertwined variables. It allows the assessment of system

| Table 1. Demographic and clinical characteristics of the patient groups. Data are presented as mean ± SD or n (%). OSA Positive: patients with obstructive sleep apnea; OSA Negative: patients without obstructive sleep apnea; BMI: body mass index; AHI: apnea-hypopnea index. |
|---------------------------------|----------------|----------------|----------------|
| All Subjects                   | OSA Positive  | OSA Negative  |
| Subjects (n)                   | 187           | 111 (59.4%)    | 76 (40.6%)     |
| Age (years)                    | 58.0 ± 12.8   | 58.3 ± 12.9    | 57.6 ± 12.9    |
| Males (%)                      | 147 (78.6%)   | 94 (84.7%)     | 53 (69.7%)     |
| BMI (kg/m^2)                   | 29.5 ± 5.5    | 30.5 ± 4.9     | 28.4 ± 6.0     |
| Time recorded (h)              | 8.2 ± 0.6     | 8.2 ± 0.8      | 8.2 ± 0.3      |
| AHI (n/h)                      | 40.1 ± 19.6   | 2.0 ± 2.4      | 2.0 ± 0.3      |
evolution characteristics such as feedback and control, avoiding the requirement to model the underlying system (Pincus 2000). Cross-ApEn evaluates secondary as well as dominant patterns in data, quantifying changes in underlying episodic behaviour that do not reflect in peak occurrences and amplitudes (Veldhuis et al 2001). For two paired time series \( u(i) \) and \( v(i) \), cross-ApEn measures, within tolerance \( r \), the (conditional) regularity or frequency of \( v \)-patterns similar to a given \( u \)-pattern of window length \( m \). Greater asynchrony indicates fewer instances of (sub)pattern matches, quantified by larger cross-ApEn values (Pincus 2000).

Formally, given two equally sampled sequences of length \( N \), \( u = [u(1), u(2), \ldots, u(N)] \) and \( v = [v(1), v(2), \ldots, v(N)] \), cross-ApEn could be computed as follows (Pincus 2000):

1. Form the vector sequences \( x(i) = [u(i), u(i+1), \ldots, u(i+m-1)] \) and \( y(j) = [v(j), v(j+1), \ldots, v(j+m-1)] \). These vectors represent \( m \) consecutive \( u \) and \( v \) values, respectively, commencing with the \( i \)th point.

2. Define the distance between \( x(i) \) and \( y(j) \), \( d[x(i), y(j)] \), as the maximum absolute difference in their respective scalar components:

\[
d[x(i), y(j)] = \max_{k=1,2,\ldots,m} |u(i+k-1) - v(j+k-1)|.
\]

3. For each \( x(i) \), count the number of \( j \) (\( j = 1 \ldots N-m+1 \)) so that \( d[x(i), y(j)] \leq r \), denoted as \( N_i^m(r) \). Then, for \( i = 1 \ldots N-m+1 \), set

\[
C_i^m(r) = \frac{N_i^m(r)}{(N-m+1)}.
\]

The \( C_i^m(r) \) values measure within a tolerance \( r \) the regularity, or frequency, of \( v \) patterns similar to a given \( u \) pattern of window length \( m \).

4. Compute the natural logarithm of each \( C_i^m(r) \), and average it over \( i \),

\[
\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r).
\]

5. Cross-ApEn is defined by:

\[
\text{Cross-ApEn}(m, r, N)(v|u) = \phi^m(r) - \phi^{m+1}(r).
\]

Typically, cross-ApEn is applied to normalized \( \{u^*(i), v^*(i)\} \) time series, where \( u^*(i) = [u(i) - \text{mean}(u)] / \text{SD}(u) \) and \( v^*(i) = [v(i) - \text{mean}(v)] / \text{SD}(v) \).

Two important issues must be taken into account when computing cross-ApEn (Richman and Moorman 2000):

1. \( \text{Cross-ApEn}(m, r, N)(v|u) \) will not always be defined, because \( C_i^m(r)(v|u) \) in equation (4) will be equal to 0 in the absence of similar patterns between \( u \) and \( v \). To solve this, we used two correction strategies (Richman and Moorman 2000): \text{bias 0} and \text{bias max}. Both strategies assign non zero values to \( C_i^m(r)(v|u) \) and \( C_i^{m+1}(r)(v|u) \) in the absence of matches. The \text{bias 0} correction sets the bias towards zero by setting:

\[
C_i^m(r) = 0, \quad C_i^{m+1}(r) = 0.
\]

The \text{bias max} correction sets the bias towards the highest observable value of cross-ApEn by setting:
\[ C_i^m(r) = 1 \text{ if originally } C_i^m(r) = 0, \]
\[ C_i^{m+1}(r) = (N - m + 1)^{-1} \text{ if originally } C_i^{m+1}(r) = 0. \]  

In this research, both correction strategies were implemented, and their influence in the diagnostic ability of cross-ApEn were studied.

2. There is a direction dependence. Because of the logarithm in equation (5), \( \phi^m(r)(v|u) \) will not generally be equal to \( \phi^m(r)(u|v) \). Thus, cross-ApEn\((v|u)\) and its direction conjugate cross-ApEn\((u|v)\) are different in most cases. In the present study, both cross-ApEn\((v|u)\) and cross-ApEn\((u|v)\) were estimated, and the influence of the direction dependence in the diagnostic accuracy of cross-ApEn was tested.

In this study, we divided SaO\(_2\) and HR signals in 200-sample epochs before cross-ApEn was applied. This methodology is intended to reduce the computational time of cross-ApEn when applied to long-term signals, due to the multiple comparisons. Cross-ApEn data were subsequently averaged over the whole number of segments to obtain a single value per subject. The segment length (200 samples = 16.6 min) was also fixed taking into account the apnoea period (between 25 s and 2 min) (Álvarez et al 2006, del Campo et al 2006, Zamarrón et al 2006, Hornero et al 2007, Álvarez et al 2007), to include various apnoeic events in every epoch. The run length input parameter \( m \) was fixed to the widely established values \( m = 1 \) and \( 2 \) (Pincus 2000, 2001), whereas the tolerance window \( r \) varied from 0.1 to 1.0 in steps of 0.1, in order to assess the consistency of the method (Richman and Moorman 2000).

3.3. Statistical analysis

The linear relation between each measure and the AHI from PSG was studied by means of Pearson correlation. Student’s \( t \)-test was applied to verify if there were significant differences between the OSA Positive and the OSA Negative group. \( p \)-values < 0.01 were considered significant. A receiver operating characteristic (ROC) analysis was performed to assess the diagnostic ability of each method. The following statistics were derived from the ROC analysis: sensitivity (percentage of OSA Positive subjects correctly diagnosed), specificity (percentage of OSA Negative patients correctly diagnosed) and accuracy (proportion of total subjects under study rightly classified). To ensure statistical validity of results, a 10-fold cross-validation approach was applied. The population under study was divided into 10 subsets, so the ROC analysis was applied in 10 iterations. At each iteration, 9 subsets were used to compute an optimum threshold, which was applied to the remaining group to compute ROC derived parameters. A different subset was left out at each iteration. Finally, sensitivity, specificity and accuracy were averaged over the 10 iterations. The optimum decision threshold in the cross-validation procedure is selected looking for a compromise between sensitivity and specificity. This value is graphically determined from the ROC curve as the closest point to the left top corner (100% sensitivity and 100% specificity).

4. Results

4.1. MSC analysis

Figure 1 displays the averaged MSC curves for both the OSA Positive and the OSA Negative groups. The lower and upper limits of the apnoea interest frequency range (0.014–0.033 Hz) are represented by the dashed vertical lines. We fixed these limits within the frequency range (0.008–0.04 Hz), determined by the maximum (2 min) and the minimum (25 s) apnoea cycle length (Zamarrón et al 2003, 2005). The curve corresponding to the OSA Positive group (solid line) is above that corresponding to the OSA Negative group (dotted line). This is especially
Nonlinear measure of synchrony from pulse oximetry in sleep apnoea

Figure 1. Average MSC curves for the OSA Positive and the OSA Negative groups. Vertical dashed lines represent the lower and upper limits of the apnea interest region (0.014–0.033 Hz).

evident within the apnoea interest region, where a peak in the OSA Positive MSC curve is shown around 0.021 Hz.

The following spectral features were derived from the MSC curves for each patient (Zamarrón et al 2003, 2005): total area enclosed below the MSC curve ($A_{TOT}$), area enclosed below the MSC curve in the apnoea interest region ($A_{REG}$), the ratio of the area in the apnoea interest region to the total area enclosed below the MSC curve ($A_{REG}/A_{TOT}$) and the maximum of the MSC function within the apnoea interest region ($M_{REG}$). Table 2 summarises the results from the MSC analysis for the groups under study. All the spectral features values were larger in the OSA Positive group than in the OSA Negative, suggesting higher coherence between oximetric recordings from OSA patients. Significant differences between both groups ($p$-value $<< 0.01$) were obtained with all spectral features.

Table 3 shows the statistics from the 10-fold cross-validation analysis. We can see that all spectral features provide small sensitivity values, between 65% and 70%, whereas the specificity values range from 76.8% to 90.9%. The best results are obtained with $A_{REG}$, which provides 69.2% sensitivity, 90.9% specificity and 78.1% accuracy.

4.2. Cross-ApEn analysis

The usefulness of cross-ApEn in OSA diagnosis was assessed. Different combinations of the input parameters $m$ and $r$ were used, as well as the correction strategies bias 0 and bias max, and the two possible directions (SaO$_2$||HR) and (HR||SaO$_2$). Firstly, results were different for both directions of the method. No statistical differences were found between the groups under study when computing cross-ApEn(HR||SaO$_2$), whereas significant differences where found with (SaO$_2$||HR). This fits well with the nature of oximetry signals. SaO$_2$ are more regular than HR in both OSA Positive and OSA Negative subjects. The number of SaO$_2$ patterns similar to a fixed HR pattern when computing $C^m(r)(SaO_2||HR)$ will be greater than the number of similar patterns when comparing a HR record to a fixed SaO$_2$ pattern. Hence, cross-ApEn will be statistically better estimated in the first case (Pincus and Goldberger 1994), increasing the
Table 2. Average values (mean ± SD) for the features derived from the MSC function in the OSA Positive and the OSA Negative groups. $A_{TOT}$: total area enclosed below the MSC function; $A_{REG}$: area enclosed below the MSC curve in the apnea interest region (0.014–0.033 Hz); $A_{TOT}/A_{REG}$: ratio of the area in the apnea interest region to the total area enclosed below the MSC curve; $M_{REG}$: maximum of the MSC function within the apnea interest region (0.014–0.033 Hz).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>OSA Positive</th>
<th>OSA Negative</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{TOT}$</td>
<td>0.0118 ± 0.0059</td>
<td>0.0068 ± 0.0023</td>
<td>$p &lt;&lt; 0.01$</td>
</tr>
<tr>
<td>$A_{REG}$</td>
<td>0.0054 ± 0.0031</td>
<td>0.0024 ± 0.0012</td>
<td>$p &lt;&lt; 0.01$</td>
</tr>
<tr>
<td>$A_{TOT}/A_{REG}$</td>
<td>0.4340 ± 0.1270</td>
<td>0.3385 ± 0.0827</td>
<td>$p &lt;&lt; 0.01$</td>
</tr>
<tr>
<td>$M_{REG}$</td>
<td>0.5367 ± 0.2177</td>
<td>0.3147 ± 0.1117</td>
<td>$p &lt;&lt; 0.01$</td>
</tr>
</tbody>
</table>

Table 3. Diagnostic test results derived from the 10-fold cross validation analysis for the features derived from the MSC function.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{TOT}$</td>
<td>65.8</td>
<td>76.8</td>
<td>70.1</td>
</tr>
<tr>
<td>$A_{REG}$</td>
<td>69.2</td>
<td>90.9</td>
<td>78.1</td>
</tr>
<tr>
<td>$A_{TOT}/A_{REG}$</td>
<td>67.4</td>
<td>82.9</td>
<td>73.7</td>
</tr>
<tr>
<td>$M_{REG}$</td>
<td>67.7</td>
<td>80.5</td>
<td>72.7</td>
</tr>
</tbody>
</table>

differences between the OSA Positive and the OSA Negative groups. With respect to the correction strategies, no significant differences were found between bias 0 and bias max ($p$-value $>> 0.01$).

Table 4 summarises the average cross-ApEn($m,r,N$)(SaO$2$||HR) values with bias 0 correction for the OSA Positive and the OSA Negative groups varying the input parameters. Significant differences ($p$-value $<< 0.01$) were found with all combinations of the input parameters. Additionally, figure 2 shows the differences and the evolution of the average cross-ApEn values for the groups under study when the input parameters change. For a fixed $m$ and $r$ pair, cross-ApEn values are lower in the OSA Positive group than in the OSA Negative group. Hence, we could affirm that SaO$_2$ and HR paired records from OSA Positive patients are more synchronized than OSA Negative ones. Moreover, figure 2 demonstrates the relative consistency of cross-ApEn applied to oximetry signals with $r$ varying from 0.1 to 1.0 and $m$ equal to 1 or 2 (Richman and Moorman 2000).

Figures 3 (a) and (b) display the oximetric recordings for two common OSA Negative and OSA Positive subjects, respectively. Visually inspecting figure 3 (b), we can see how intervals with the greatest changes in HR coincide with the deepest SaO$_2$ desaturations. However, we cannot identify this behaviour in figure 3 (a). Figure 3 (c) depicts the evolution of cross-ApEn during the recording time for both subjects. Because of the synchrony between SaO$_2$ and HR recordings from the OSA patient, its cross-ApEn curve is always below that corresponding to the OSA Negative subject. Furthermore, cross-ApEn remains quite constant during the time of recording for the OSA patient, whereas the cross-ApEn curve for the OSA Negative subject shows marked changes.

Table 5 summarizes the statistics from the 10-fold cross-validation analysis of the cross-ApEn($m,r,N$)(SaO$2$||HR) with bias 0. The highest diagnostic accuracy were obtained with $m = 1$ and $r = 0.2$ (83.7% sensitivity, 84.3% specificity and 84.0% accuracy).

Figures 4 (a) and (b) display the ROC curves for the randomly selected iterations 3 and 7 from the 10-fold cross-validation procedure for the $A_{REG}$ feature. The average ROC curve is displayed in figure 4 (c). Classification statistics from $A_{REG}$ vary slightly when different diagnostic thresholds are applied. Additionally, figures 4 (d) – (f) display the same ROC curves for cross-ApEn(1,0.2). The consistence of cross-ApEn(1,0.2) among the iterations into the 10-
Table 4. Average cross-ApEn values (mean ± SD) in the OSA Positive and the OSA Negative groups. \( m \): cross-ApEn window length; \( r \): cross-ApEn tolerance.

<table>
<thead>
<tr>
<th>( m )</th>
<th>( r )</th>
<th>OSA Positive</th>
<th>OSA Negative</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
<td>3.69 ± 0.79</td>
<td>4.73 ± 0.41</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td>2.83 ± 0.72</td>
<td>4.15 ± 0.59</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td>2.20 ± 0.62</td>
<td>3.51 ± 0.70</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td>1.70 ± 0.50</td>
<td>2.88 ± 0.73</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>1.34 ± 0.39</td>
<td>2.35 ± 0.71</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td>1.07 ± 0.29</td>
<td>1.87 ± 0.63</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.7</td>
<td></td>
<td>0.87 ± 0.21</td>
<td>1.46 ± 0.54</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td>0.71 ± 0.16</td>
<td>1.13 ± 0.42</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.9</td>
<td></td>
<td>0.59 ± 0.12</td>
<td>0.86 ± 0.33</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td>0.50 ± 0.09</td>
<td>0.68 ± 0.24</td>
<td>( p &lt; 0.01 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( m )</th>
<th>( r )</th>
<th>OSA Positive</th>
<th>OSA Negative</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
<td>4.17 ± 0.65</td>
<td>4.96 ± 0.31</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td>2.98 ± 0.91</td>
<td>4.48 ± 0.59</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td>2.28 ± 0.81</td>
<td>3.88 ± 0.75</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td>1.79 ± 0.63</td>
<td>3.25 ± 0.82</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>1.43 ± 0.48</td>
<td>2.70 ± 0.83</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td>1.15 ± 0.35</td>
<td>2.16 ± 0.75</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.7</td>
<td></td>
<td>0.94 ± 0.26</td>
<td>1.70 ± 0.65</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td>0.78 ± 0.19</td>
<td>1.33 ± 0.52</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>0.9</td>
<td></td>
<td>0.65 ± 0.14</td>
<td>1.01 ± 0.40</td>
<td>( p &lt; 0.01 )</td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td>0.55 ± 0.10</td>
<td>0.79 ± 0.30</td>
<td>( p &lt; 0.01 )</td>
</tr>
</tbody>
</table>

Figure 2. Mean cross-ApEn(SaO\(_2||HR\)) for both the OSA Positive and OSA Negative groups varying cross-ApEn input parameters.

A fold cross-validation procedure could be seen. Furthermore, the AROC for cross-ApEn(1,0.2) in figure 4 (f) is significantly higher than the AROC for the A\(_{\text{REG}}\) feature in the figure 4 (c). The optimum threshold selected at each iteration and the classification statistics obtained using the corresponding test sub-set for A\(_{\text{REG}}\) and cross-ApEn(1, 0.2) are shown in table 6.

Finally, table 7 shows the statistics from the Pearson correlation test between the AHI derived from PSG and the measures involved in the study. Additionally, correlation between spectral and nonlinear features is included. High correlation coefficients were obtained between
Figure 3. SaO$_2$ and HR records from NPO for a common (a) OSA Negative subject and (b) OSA Positive patient. (c) Evolution of the cross-ApEn(SaO$_2$||HR) for both the common OSA Negative and OSA Positive subjects with the recording time.

Table 5. Diagnostic test results derived from cross-validation for the cross-ApEn analysis. $m$: cross-ApEn window length; $r$: cross-ApEn tolerance.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$r$</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>81.9</td>
<td>84.3</td>
<td>82.9</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>83.7</td>
<td>84.3</td>
<td>84.0</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
<td>81.0</td>
<td>78.9</td>
<td>80.2</td>
</tr>
<tr>
<td>0.4</td>
<td>0.4</td>
<td>82.8</td>
<td>76.6</td>
<td>80.2</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>81.9</td>
<td>76.6</td>
<td>79.7</td>
</tr>
<tr>
<td>0.6</td>
<td>0.6</td>
<td>83.7</td>
<td>75.4</td>
<td>80.3</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7</td>
<td>82.0</td>
<td>75.2</td>
<td>79.2</td>
</tr>
<tr>
<td>0.8</td>
<td>0.8</td>
<td>80.2</td>
<td>76.6</td>
<td>78.6</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9</td>
<td>77.4</td>
<td>66.1</td>
<td>72.7</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>76.7</td>
<td>64.8</td>
<td>71.6</td>
</tr>
</tbody>
</table>

Spectral features from MSC, whereas lower correlation was observed between these indexes and cross-ApEn. $A_{REG}$ and cross-ApEn(1,0.2) were significantly correlated with AHI ($p$-value $<< 0.01$), with Pearson correlation coefficients $\rho = 0.67$ and $\rho = -0.68$, respectively. This suggests that, as OSA severity increases, both correlation and synchrony between oximetric recordings increases: coherence increases and cross-ApEn decreases.
5. Discussion and conclusion

The advantages and limitations of a nonlinear synchrony measure between \( \text{SaO}_2 \) and HR from NPO as an alternative to classical spectral-based methods was assessed. We analyzed the oximetric recordings of 187 patients (111 with a positive diagnosis of OSA and 76 with a negative diagnosis of OSA). A 10-fold cross-validation study was applied to evaluate the diagnostic performance of each method. Firstly, a classical spectral analysis of coherence by means of the MSC was performed. Four common spectral features were derived from the MSC curves, \( A_{\text{TOT}} \), \( A_{\text{REG}} \), \( A_{\text{REG}}/A_{\text{TOT}} \) and \( M_{\text{REG}} \) were significantly higher \((p\text{-value} \ll 0.01)\) in the OSA Positive group than in the OSA Negative one, resulting in higher linear correlation at the
Table 7. Pearson correlation coefficients between AHI and spectral and nonlinear features derived from NPO. \( \rho \): Pearson correlation coefficient; \( A_{TOT} \): total area enclosed below the MSC function; \( A_{REG} \): area enclosed below the MSC curve in the apnea interest region (0.014–0.033 Hz); \( A_{TOT}/A_{REG} \): ratio of the area in the apnea interest region to the total area enclosed below the MSC curve; \( M_{REG} \): maximum of the MSC function within the apnea interest region (0.014–0.033 Hz); \( \text{cross-ApEn}(1,0.2) \): cross approximate entropy with input parameters \( m = 1 \) and \( r = 0.2 \).

<table>
<thead>
<tr>
<th></th>
<th>AHI</th>
<th>cross-ApEn ((1,0.2))</th>
<th>( A_{TOT} )</th>
<th>( A_{REG} )</th>
<th>( A_{REG}/A_{TOT} )</th>
<th>( M_{REG} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHI</td>
<td>1</td>
<td>-0.68</td>
<td>0.61</td>
<td>0.67</td>
<td>0.49</td>
<td>0.64</td>
</tr>
<tr>
<td>cross-ApEn ((1,0.2))</td>
<td>-0.68</td>
<td>1</td>
<td>-0.60</td>
<td>-0.62</td>
<td>-0.42</td>
<td>-0.62</td>
</tr>
<tr>
<td>( A_{TOT} )</td>
<td>0.61</td>
<td>-0.60</td>
<td>1</td>
<td>0.92</td>
<td>0.46</td>
<td>0.84</td>
</tr>
<tr>
<td>( A_{REG} )</td>
<td>0.67</td>
<td>-0.62</td>
<td>0.92</td>
<td>1</td>
<td>0.73</td>
<td>0.92</td>
</tr>
<tr>
<td>( A_{REG}/A_{TOT} )</td>
<td>0.49</td>
<td>-0.42</td>
<td>0.46</td>
<td>0.73</td>
<td>1</td>
<td>0.76</td>
</tr>
<tr>
<td>( M_{REG} )</td>
<td>0.64</td>
<td>-0.62</td>
<td>0.84</td>
<td>0.92</td>
<td>0.76</td>
<td>1</td>
</tr>
</tbody>
</table>

frequency components in the apnoea interest band. The relationship between these oximetric recordings was subsequently investigated applying a nonlinear measure of synchrony: the cross-ApEn. We developed an exhaustive study of the method varying the input parameters \( m \) and \( r \), showing the relative consistence of cross-ApEn when applied to oximetry signals. For a fixed \( m \) and \( r \) pair, the OSA Positive group had an average cross-ApEn significantly lower than the OSA Negative group \((p\text{-value} \ll 0.01)\), which corresponds to a higher degree of synchrony between \( \text{SaO}_2 \) and HR. Results from the nonlinear analysis of oximetry signals by means of cross-ApEn agree with classical frequency analysis. We showed that the recurrent apnoea events typical of OSA cause coordinated fluctuations in \( \text{SaO}_2 \) and HR signals, leading to high \( A_{TOT} \), \( A_{REG} \), \( A_{REG}/A_{TOT} \) and \( M_{REG} \) values and, similarly, to low cross-ApEn values. Additionally, we found that the diagnostic accuracy obtained with cross-ApEn significantly improved that obtained with spectral features derived from the MSC curve. Classical coherence analysis provided low sensitivities and moderate to high specificities: \( A_{REG} \) reached the best results, with 69.2% sensitivity, 90.9% specificity and 78.1% accuracy. On the other hand, moderate to high sensitivities and specificities were obtained with cross-ApEn, leading to accuracies greater than 80% in ten combinations of the input parameters: \( \text{cross-ApEn}(1,0.2) \) reached the best results, with 83.7% sensitivity, 84.3% specificity and 84.0% accuracy. The mechanisms controlling the interrelation between the respiratory and the cardiovascular systems are not completely described applying the spectral methods commonly used. Our results suggest that the relationship between changes in \( \text{SaO}_2 \) and HR from NPO due to apnoea events is better characterized by nonlinear analysis.

To our knowledge, this is the first study assessing the synchrony between \( \text{SaO}_2 \) and HR signals from NPO by means of cross-ApEn. Furthermore, this is the first study that assesses the diagnostic ability of MSC and cross-ApEn analyses of oximetric recordings to detect OSA. Although previous studies had evaluated the relationship between \( \text{SaO}_2 \) and HR, our research extends these analyses providing new measures of synchrony between both signals and quantitative results of their diagnostic ability in OSA screening. An increase in the spectral frequency components of HR from Holter ECG was linked to synchronised episodes of hypoxemia in the study by Shiomi \textit{et al} (1996). In (Keyl \textit{et al} 1997), frequency analysis was applied to nocturnal Holter ECG epochs during episodes of periodic and normal breathing determined by oximetry. The authors suggest that spectral analysis of HR provides an accurate methodology to detect episodes of sleep-related breathing disorders. In the study by Noda \textit{et al} (1998), 24-h simultaneously recorded \( \text{SaO}_2 \) and Holter ECG signals were used to search for daytime sleep consequences in patients with OSA. However, all these studies did not analyze \( \text{SaO}_2 \) and HR signals jointly and no results about its diagnostic ability in OSA screening were provided. In (Zamarrón \textit{et al} 2003), independent frequency analyses of \( \text{SaO}_2 \) and HR from nocturnal oximetry were developed to screen for OSA syndrome. Different spectral features
Nonlinear measure of synchrony from pulse oximetry in sleep apnoea

were derived from the periodograms. The researchers showed that taking the periodograms together improved the results provided by the independent spectral analysis of SaO$_2$ or HR. This study reached high diagnostic accuracies by visual inspection but no automated assessment was made. Furthermore, none of the spectral features was derived from the joint analysis of the oximetric recordings. In a latter study by Zamarrón et al (2005), a simultaneous analysis of SaO$_2$ and HR from NPO was developed by means of the coherence function. Three different patterns were identified: positive, negative and undetermined. Both positive and negative patterns were mainly associated to OSA Positive patients, while a significant number of subjects had an undetermined pattern. No features were derived from the coherence function and no results about the ability of coherence to help in OSA diagnosis were provided.

We should take into account some limitations of our study. The sample size could be larger and OSA Positive patients were predominantly studied. Further work is required to apply our methodology to a new and larger data set with a wide spectrum of sleep-related breathing disorders, as well as to study groups of especial interest, such as patients with lung or cardiovascular diseases and young snorers. Our research is aimed at discriminating OSA Positive patients from OSA Negative subjects. However, it would be interesting to study additional groups, like a control population consisting of healthy subjects. This would increase our knowledge of the methods involved in our study (MSC and cross-ApEn) and would also confirm the consistency of both methods. An additional drawback could be linked to our methodology and it should be mentioned. Our study is based solely on oximetry data: SaO$_2$ and HR were simultaneously recorded with a single oximeter, eliminating the need for additional equipment. This a priori advantage could introduce a correlation between both signals. Although our results agree with previous studies assessing the relation between oximetric recordings and HR from Holter ECG, further analyses are necessary to test our methodology. Moreover, the sampling frequency used in this study (0.2 Hz) is lower than that used by current oximeters (1 Hz). The study by Wiltshire et al (2001) showed that low sampling rates provide SaO$_2$ recordings with a low number of artefacts. However, low sampling frequencies potentially reduce the sensitivity and increase the specificity of a diagnostic test (Flemons et al 2003). An additional limitation should be taken into account, related to the lack of repeated sleep studies in successive nights. Night-to-night variability of AHI from conventional PSG is widely known (Dean et al 1992, Chediak et al 1996, Carlile and Carlile 2008). However, a recent study emphasize that the high cost and complexity of PSG limits its repeatability (Levendowski et al 2009). This study also suggests that sleep studies at home by means of portable equipment would be less variable. Thus, our methodology, based on portable NPO and automated nonlinear analysis, could benefit from this issue. However, repeated sleep studies (both in PSG and NPO) among two consecutive nights would be necessary to validate the performance of our methodology. Finally, an additional issue regarding the continuous positive airway pressure (CPAP) therapy should be taken into account. Our study is aimed at developing a screening tool for OSA syndrome, in order to discriminate OSA Positive patients, who will benefit from respiratory therapies like continuous positive airway pressure (CPAP). CPAP is currently the most widely used treatment for severe OSA (Lindberg et al 2006, Marshall et al 2006). Nevertheless, whether CPAP reduces sleepiness in subjects with mild to moderate OSA syndrome (AHI between 5 and 30 e/h) is not proven (Marshall et al 2006, Gay et al 2006). Thus, further work is needed to assess the accuracy of our methodology for screening those patients who will benefit from CPAP.

In summary, the results reported in this study suggest that nonlinear analysis of joint SaO$_2$ and HR from NPO could provide additional and useful information to help in sleep apnoea detection. Our results agree with previous studies that assessed the relationship between desaturations and changes in HR due to periodic hypoxia (Shiomi et al 1996, Keyl et al 1997, Noda et al 1998, Zamarrón et al. 2003, 2005). We found that the recurrence of apnoea events in patients with OSA syndrome determined an important increase in the synchrony between SaO$_2$ and HR, which leads to a significant decrease of cross-ApEn. Additionally, our results suggest that nonlinear analysis by means of cross-ApEn could be better suited to help in OSA diagnosis.
than classical frequency analysis. However, we need to carry out this methodology with larger populations to verify prospectively whether cross-ApEn may be an easy tool to use in OSA diagnosis on an ambulatory basis.

Acknowledgements

This work has been partially supported by Ministerio de Ciencia e Innovación and FEDER grant TEC 2008–02241, the grant project from Consejería de Sanidad de la Junta de Castilla y León SAN/673/VA03/08 and Sociedad Castellano – Leonesa y Cántabra de Patología Respiratoria (SOCALPAR).

References

Carrière J and Carrière N 2008 repeat study of 149 patients suspected of having sleep apnea but with an AHI<5 Sleep 31 A153
Chediac AD, Acvedo-Crespo JC, Seiden DJ, Kim HH and Kiel MH 1996 Nightly variability in the indices of sleep-disordered breathing in men being evaluated for impotence with consecutive night polysomnograms Sleep 19 589–92
Dean RJ and Chaudhary BA 1992 Negative polysomnogram in patients with obstructive sleep apnea syndrome Chest 101 105–8
Findley L J, Farkas G A and Rochester D F 1985 Changes in heart rate during breathing interrupted by recurrent apneas in humans J. Appl. Physiol. 59 536–42
Gay P, Weaver T, Loube D and Iber C 2006 Evaluation of positive airway pressure treatment for sleep related breathing disorders in adults Sleep 29 381–401
Kirby R R, Taylor R W and J M Civetta 1997 Handbook of Critical Care (2nd Ed.) (Philadelphia, USA Lippincott-Raven)
Nonlinear measure of synchrony from pulse oximetry in sleep apnoea


Levendowski D, Steward D, Woodson BT, Olmstead R, Popovic D and Westbrook P 2009 The impact of obstructive sleep apnea variability measured in-lab versus in-home on sample size calculations International Archives of Medicine 2 1–8


Marshall NS, Barnes M, Travier N, Campbell AJ, Pierce RJ, Neill AM and Gander PH 2006 Continuous positive airway pressure reduces daytime sleepiness in mild to moderate sleep apnoea: a meta-analysis Thorax 61 430–4

Netzer N, Eliasson A H, Netzer C and Kristo DA 2001 Overnight Pulse Oximetry for Sleep-Disordered Breathing in Adults Chest 120 625–33


Peppard P E, Young T, Palta M and Skatrud J 2000 Prospective study of the association between sleep-disordered breathing and hypertension N. Engl. J. Med. 342 1378–84


Pincus S M 2000 Irregularity and asynchrony in biologic network signals Methods Enzymol. 321 149–82


Sériès F and Marc I 1999 Nasal pressure recording in the diagnosis of sleep apnoea hypopnea syndrome Thorax 54 506–10


Zamarrón C, Pichel F and Romero P V 2005 Coherence between oxygen saturation and heart rate obtained from pulse oximetric recordings in obstructive sleep apnoea *Physiol. Meas.* **26** 799–810
Zamarrón C, Hornero R, del Campo F, D Abásolo and Álvarez D 2006 Heart rate regularity analysis obtained from pulse oximetric recordings in the diagnosis of obstructive sleep apnea *Sleep Breath.* **10** 83–9