Non-linear characteristics of blood oxygen saturation from nocturnal oximetry for obstructive sleep apnoea detection

D Álvarez¹, R Hornero¹, D Abásolo¹, F del Campo² and C Zamarrón³

¹ E.T.S. Ingenieros de Telecomunicación, Campus Miguel Delibes, Camino del Cementerio s/n, 47011 Valladolid, España.
² Hospital del Río Hortega, c/ Cardenal Torquemada s/n, 47010 Valladolid, España
³ Hospital Clínico Universitario, c/ Galeras s/n, 15705 Santiago de Compostela, España

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

Abstract
Nocturnal oximetry is an attractive option for the diagnosis of obstructive sleep apnoea (OSA) syndrome because of its simplicity and low cost compared to polysomnography (PSG). The present study assesses non-linear analysis of blood oxygen saturation (SaO₂) from nocturnal oximetry as a diagnostic test to discriminate between OSA positive and OSA negative patients. A sample of 187 referred outpatients, clinically suspected of having OSA, were studied using nocturnal oximetry performed simultaneously with complete PSG. A positive OSA diagnosis was found for 111 cases, while the remaining 76 cases were classified as OSA negative. The following oximetric indexes were obtained: cumulative time spent below a saturation of 90% (CT90), oxygen desaturation indexes of 4% (ODI4), 3% (ODI3) and 2% (ODI2), and the delta index (∆ index). SaO₂ records were subsequently processed applying two non-linear methods: central tendency measure (CTM) and Lempel-Ziv (LZ) complexity. Significant differences (p < 0.01) were found between OSA positive and OSA negative patients. Using CTM we obtained a sensitivity of 90.1% and a specificity of 82.9%, while with LZ the sensitivity was 86.5% and the specificity was 77.6%. CTM and LZ accuracy was higher than that provided by ODI4, ODI3, ODI2 and CT90. The results suggest that non-linear analysis of SaO₂ signals from nocturnal oximetry could yield useful information in OSA diagnosis.

Keywords: obstructive sleep apnoea, oximetry, blood oxygen saturation, non-linear analysis, classical oximetry analysis central tendency measure, Lempel-Ziv complexity

1. Introduction
Obstructive sleep apnoea (OSA) syndrome is characterized by recurrent episodes of upper airway obstruction during sleep, leading to total absence or markedly reduced airflow (Deegan and MacNicholas 1995). This disease is usually associated with hypoxemia, bradycardia, arousals and fragmented sleep (Gilleminault et al 1978). During apnoeas and the subsequent hyperventilation periods and arousals, OSA patients with intact autonomic nervous system experience a succession of bradycardia and tachycardia episodes (Zwillich et al 1982, Guilleminault et al 1984, Andreas et al 1992). However, treatment with continuous positive airway pressure (CPAP) decreases daytime sympathetic traffic and alleviates those changes in
heart rate (Narkiewicz et al 1999). The prevalence of OSA is estimated at 1 - 5% of adult men in western countries (Young et al 2002). This disease is associated with excessive daytime sleepiness (EDS), neurocognitive deficits, and psychological problems (Day et al 1999). Furthermore, this disease is a risk factor for traffic accidents (Findley et al 1995, Barbé et al 1998) and cardiovascular disease morbidity and mortality (Peter et al 1995, Leung and Bradley 2001).

The standard diagnostic test for OSA syndrome is overnight polysomnography (PSG) (AHCPR 1999), consisting in the recording of neurophysiological and cardiorespiratory signals subsequently used to analyze sleep and breathing. Given the wide prevalence of OSA, many health systems have been overwhelmed by the demand for PSG, which is expensive and time consuming (Lévy et al 1996, Herer et al 1999, Gagnadoux et al 2002, Kirk et al 2003). Nocturnal pulse oximetry is widely used in many medicine areas to determine patient’s blood oxygen saturation (SaO2) and heart rate because of its non-invasive nature and simplicity. The most widely used oximetry tools are the oxygen desaturation profile throughout the night, oxygen desaturation indexes (ODIs) measuring dips in the SaO2 signal (Sano et al 1998, Chiner et al 1999, Nuber et al 2000), and the cumulative time spent below an oxygen saturation of 90% (Chaudhary et al 1998, Golpe et al 1999).

Computer-applied mathematical techniques automatically plot out oximetry signals. In OSA, classical spectral analysis based on Fourier Transform has been performed to analyze the periodic nature of desaturation dips to extract features from power spectrum (Zamarrón et al 1999, 2003). Other studies try to provide additional information not based on periodicity but on the changes of physiological series. SaO2 signals tend to remain constant round 97% in normal subjects, varying slightly with age and independently of ethnicity, gender or weight (Netzer et al 2001). On the other hand, significant changes can be found in patients affected by OSA because of the recurrent episodes of apnoea. Some studies have tried to take advantage of this special behaviour of SaO2 making different measures of variability via the delta index (Lévy et al 1996, Olson et al 1999, Magalang et al 2003).

It has been shown that non-linear methods are useful in the analysis of biomedical signals. Diverse techniques have been applied over a wide range of diseases extracting different non-linear characteristics. For instance, complexity and regularity measures of EEG signals have been taken to study the depth of anaesthesia (Zhang et al 2001) and Alzheimer disease (Abásolo et al 2005). Variability measures of ECG allow to distinguish between normal and chronic heart failure subjects (Cohen et al 1996). Some authors have applied non-linear methods over different respiratory patterns to study sleep stages and the coordination between brain and lungs (Burioka et al 2003) or panic disorder (Caldirola et al 2004).

In this study we applied two non-linear methods to analyze SaO2 signals: central tendency measure (CTM) and Lempel-Ziv (LZ) complexity. Common oximetric indexes were also computed. The aim of this study was to assess whether non-linear characteristics of SaO2 from nocturnal oximetry could discriminate between patient groups, providing useful information on the diagnosis of OSA syndrome.

2. Subjects and Signals

The population under study consisted of 187 patients (147 males and 40 females) with a mean ± standard deviation (SD) age of 57.97 ± 12.84 years and a body mass index (BMI) of 29.54 ± 5.51 kg/m². All subjects were suspected of having OSA because of daytime hypsomnolence, loud snoring, nocturnal choking and awakenings, or apnoeic events (or all four symptoms) reported by the subject or a bed mate.

Sleep studies consisted of 8 hours recordings. They 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 at this institution approved the protocol, and all subjects gave their informed consent to participate in the study. Patients were evaluated by an overnight pulse oximetry test in conjunction with a simultaneous conventional PSG study. PSG was carried out using a polygraph (Ultrasom Network, Nicolet, Madison, WI,
Non-linear analysis of oximetric recordings for apnoea detection

U.S.A.) and included EEG, ECG, electrooculogram, chin electromyogram, measurement of chest wall movement and air flow measurement (three-port thermistor). Thermistor is highly reliable in detecting static respiratory events (apnoeas). However, it is less effective with dynamic respiratory events (hypopnoeas) (Farré et al 1998), which may lead to diagnostic errors. The use of nasal cannula identifies changes in nasal pressure during breathing and can improve detection of dynamic respiratory events (Montserrat et al 1997). However, there may be a detection problem with nasal cannula when the patient breathes mainly through the mouth. One of the limitations of nasal pressure is false positive detection of apnoeas/hypopnoeas due to nasal obstruction or mouth breathing. Moreover, its reliance on flow limitation leads to ambiguous results. Although we are currently using both techniques for clinical application, it is important to remember that some of the data in our study were gathered before nasal cannula was validated by Heitman et al (2002). The American Academy of Sleep Medicine (AASM) Task Force suggested that differentiation of apnoeas from hypopnoeas was not necessary in clinical practice because both event types share a common pathophysiology and clinical consequences (AASM 1999). The PSG register was analyzed over periods of 30 s and during sleep phases I, II, III, IV and rapid eye movement, according to R&K rules (Rechtschaffen and Kales 1968). Apnoea was defined as the absence 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 (Zamarrón et al 1999, 2003, 2005, Meoli et al 2001). The average apnoea-hypopnoea index (AHI) was calculated for hourly periods of sleep. In this study, an AHI of 10 or more was considered as diagnosis of OSA. If the subject had less than 3 hours of total sleep, the sleep study was repeated.

After the PSG, a spirometry study was carried out using the conventional Collins spirometer. This study showed that 42 patients had chronic obstructive pulmonary disease (COPD). COPD was defined as a disease state characterized by airflow limitation that is not fully reversible. The airflow limitation is usually both progressive and associated with an abnormal inflammatory response of the lungs to noxious particles or gases (GOLD 2001). Moreover, 9 of the 42 patients with COPD (21.8%) present respiratory failure. According to the GOLD consensus (GOLD 2001), 22 of these patients (52.4%) could be classified as mild, 14 (33.3%) as moderate and 6 (14.3%) as severe. COPD subjects had a mean age of 62.26 ± 13.65 years and a BMI of 29.66 ± 17.31 kg/m².

Table 1 summarizes the demographic and clinical data of the subjects under study. After the PSG, subjects were classified into two groups. There was an OSA group consisting of 111 patients (59.4%) with an AHI of 40.07 ± 19.64 events/hour, an average age of 58.30 ± 12.88 years and a BMI of 30.45 ± 4.92 kg/m². There was also a non-OSA group including the remaining 76 subjects (40.6%), with an AHI of 2.04 ± 2.36 events/hour, an average age of 57.57 ± 12.87 years and a BMI of 28.42 ± 6.02 kg/m².

Recording of SaO₂ was carried out using a Criticare 504 oximeter (CSI, Waukesha, WI, U.S.A.) with a finger probe and sampling frequency of 0.2 Hz (one sample every 5 s). Although new oximeters work at higher frequencies, Warley et al (1987) showed that this sampling frequency provides reasonable resolution in SaO₂ variability. There is some underestimation of the peak SaO₂ in recovery post-apnoea, but the signal shape and variability were preserved (Warley et al 1987). The SaO₂ signals were saved to separate files and processed off-line by means of CTM and LZ. Figure 1 displays three common oximetric recordings. Figure 1(a) depicts a common OSA negative subject. Figure 1(b) shows a SaO₂ record with clearly marked desaturations, corresponding to an apparent OSA positive subject. Figure 1(c) illustrates more exactly the difficulty in the diagnosis of the disease. It shows the SaO₂ record for an uncertain OSA positive patient. In this case, dips in SaO₂ are not so extreme and the diagnosis by visual inspection is not evident.
### Table 1. Demographic and clinical characteristics of the subjects under study. Data are presented as mean ± SD unless otherwise indicated. OSA positive: patients with obstructive sleep apnoea. OSA negative: patients without obstructive sleep apnoea.

<table>
<thead>
<tr>
<th>Features</th>
<th>All subjects</th>
<th>OSA positive</th>
<th>OSA negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects (n)</td>
<td>187</td>
<td>111</td>
<td>76</td>
</tr>
<tr>
<td>Age (years)</td>
<td>57.97 ± 12.84</td>
<td>58.30 ± 12.88</td>
<td>57.57 ± 12.87</td>
</tr>
<tr>
<td>Males (%)</td>
<td>78.61</td>
<td>84.68</td>
<td>69.74</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>29.54 ± 5.51</td>
<td>30.45 ± 4.92</td>
<td>28.42 ± 6.02</td>
</tr>
<tr>
<td>Recording time (h)</td>
<td>8.19 ± 0.62</td>
<td>8.17 ± 0.75</td>
<td>8.22 ± 0.33</td>
</tr>
<tr>
<td>AHI (n/h)</td>
<td>40.07 ± 19.64</td>
<td>2.04 ± 2.36</td>
<td></td>
</tr>
<tr>
<td>COPD (n)</td>
<td>42</td>
<td>22</td>
<td>20</td>
</tr>
</tbody>
</table>

**Figure 1.** SaO₂ records from nocturnal oximetry for (a) a common OSA negative subject (b) an apparent OSA positive patient and (c) an uncertain OSA positive subject.

### 3. Methods

#### 3.1. Classical oximetry indexes

The following oximetry indexes were stored by our Criticare 504 oximeter: oxygen desaturation indexes of 4% (ODI4), 3% (ODI3), 2% (ODI2) and cumulative time spent below a saturation of 90% (CT90). The number of falls in SaO₂ of 4% or more, 3% or more and 2% or more were computed from baseline. Baseline was set initially as the mean level in the first 3 min of recording (Gyulay et al. 1993). ODI4 was computed for events with falls in SaO₂ greater than or equal to 4% per hour of recording, ODI3 was calculated for events with falls of SaO₂ greater than or equal to 3% per hour of recording and ODI2 was computed for events with falls of SaO₂ greater than or equal to 2% per hour of recording. CT90 were calculated as the percentage of time during which SaO₂ was below 90%. Furthermore, we have developed a computer program to compute the delta index (Δ index). This is a common measure of signal variability usually applied by researchers (Lévy et al. 1996, Olson et al. 1999, Magalang et al. 2003). The Δ index was computed as the sum of the absolute variations between two successive points, divided by the number of intervals (Lévy et al. 1996). Although the Δ index is usually computed for 12
Non-linear analysis of oximetric recordings for apnoea detection

seconds intervals (Lévy et al 1996, Olson et al 1999, Magalang et al 2003), we used 10 seconds intervals because of our sampling frequency.

3.2. Central Tendency Measure (CTM)

Second-order difference plots given by (1) are graphs centred round the origin used to assess the degree of chaos in a data set.

\[
[x(n + 2) - x(n + 1)] \text{ vs. } [x(n + 1) - x(n)]
\]

These kinds of scatter plots are very useful in modelling biological systems such as haemodynamics and heart rate variability (Cohen et al 1996). With this approach, rather than defining a time series as chaotic or not chaotic, the degree of variability or chaos is evaluated (Cohen et al 1996).

We use the CTM to obtain a quantitative variability measure via the second-order difference plots. The CTM is computed by selecting a circular region of radius \( \rho \) round the origin, counting the number of points that fall within the radius, and dividing by the total number of points. Given \( N \) data points from a time series, \( N-2 \) would be the total number of points in the scatter plot given by (1). Then, the CTM can be computed as (Jeong et al 2002),

\[
\text{CTM} = \frac{\sum_{i=1}^{N-2} \delta(d_i)}{N - 2}
\]

where

\[
\delta(d_i) = \begin{cases} 1 & \text{if } \left[ \left( x(i+2) - x(i+1) \right)^2 + \left( x(i+1) - x(i) \right)^2 \right]^{1/2} < \rho \\ 0 & \text{otherwise} \end{cases}
\]

The radius is chosen depending upon the character of the data. We used a new method (Hornero et al 1999) to select the radius. First, we compute the CTM with several radii for every SaO\(_2\) signal in both OSA positive and OSA negative groups. Then, we apply one way ANOVA to select the radii that achieve the most significant differences, i.e. the smallest \( p \)-values. If all the \( p \)-values are statistically significant (\( p < 0.01 \)), as it was the case, we choose the optimum radius in terms of the \( F \) statistic. \( F \) varies inversely and more slowly than the \( p \)-value, showing better little changes among radii. Table 2 contains the results of the one way ANOVA, showing the evolution of the \( F \) statistic with increasing radii. We can notice that higher \( F \) values were achieved when the radius tends to zero. Radii ranging from 0 to 1 provide the best results. Thus, we selected a radius \( \rho = 0.25 \) according to the values showed in table 2.

Previously to CTM analysis, every SaO\(_2\) record was first divided into epochs of 200 samples. We determined this epoch length according to the apnoea duration (between 25 s and 2 min) in order to include various disordered respiratory events in every epoch. The method was applied over every epoch, providing various values for each signal. Finally, these variability measures were averaged to obtain a single result per subject.

3.3. Lempel-Ziv (LZ) Complexity

The Lempel-Ziv (LZ) complexity measure for sequences of finite length was suggested by Lempel and Ziv (1976). It is a nonparametric, simple-to-calculate measure of complexity in a one-dimensional signal that does not require long data segments to compute (Zhang et al 1999). LZ complexity is related to the number of distinct substrings and the rate of their recurrence along the given sequence (Radhakrishnan and Gangadhar 1998). In the last years, LZ complexity has been widely applied in biomedical signal analysis. LZ has been used for characterization of DNA sequences (Gusev et al 1999) and to develop new methods for
Table 2. $F$ statistic for every radius. Each value was obtained applying one way ANOVA over CTM values corresponding to 111 OSA positive and 76 OSA negative patients and increasing the radius in 0.25 times from 0.25 to 5.

<table>
<thead>
<tr>
<th>Radius</th>
<th>$F$ statistic</th>
<th>Radius</th>
<th>$F$ statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>203.50</td>
<td>2.75</td>
<td>82.26</td>
</tr>
<tr>
<td>0.50</td>
<td>203.50</td>
<td>3.00</td>
<td>76.76</td>
</tr>
<tr>
<td>0.75</td>
<td>203.50</td>
<td>3.25</td>
<td>63.97</td>
</tr>
<tr>
<td>1.00</td>
<td>203.50</td>
<td>3.50</td>
<td>63.97</td>
</tr>
<tr>
<td>1.25</td>
<td>149.22</td>
<td>3.75</td>
<td>57.58</td>
</tr>
<tr>
<td>1.50</td>
<td>118.48</td>
<td>4.00</td>
<td>57.58</td>
</tr>
<tr>
<td>1.75</td>
<td>118.48</td>
<td>4.25</td>
<td>49.42</td>
</tr>
<tr>
<td>2.00</td>
<td>118.48</td>
<td>4.50</td>
<td>45.85</td>
</tr>
<tr>
<td>2.25</td>
<td>82.26</td>
<td>4.75</td>
<td>45.85</td>
</tr>
<tr>
<td>2.50</td>
<td>82.26</td>
<td>5.00</td>
<td>45.85</td>
</tr>
</tbody>
</table>

discovering patterns in DNA sequences by applying it to genomic sequences of Plasmodium falciparum (Stern et al 2001). It has also been applied to characterize the responses of neurons of the primary visual cortex to different kinds of stimulus (Szczepaski et al 2003) and to estimate the entropy of neural discharges (Amigó et al 2004). Other studies have computed LZ for recognition of structural regularities (Orlov and Potapov 2004) and to develop a new sequence distance measure for phylogenetic tree construction (Otu and Sayood 2003). It has also been applied to study the brain function (Wu and Xu 1991), brain information transmission (Xu et al 1997), ECG dynamics (Zhang et al 1997) and epileptic seizures (Radhakrishnan and Gangadhar 1998). Others authors have used LZ complexity to detect ventricular tachycardia and fibrillation (Zhang et al 1999), to predict movement in anaesthesia in animals (Zhang and Roy 1999), to estimate the depth of anaesthesia (Zhang and Roy 2001, Zhang et al 2001) and to quantify oscillations in uterine electromyography (Radhakrishnan 2002).

$LZ$ complexity analysis is based on a coarse-graining of the measurements, so before calculating the complexity measure $c(n)$, the signal must be transformed into a finite symbol sequence. In this study, the simplest way is selected to convert time series values ($x(i)$, $i = 1, 2, \ldots, N$) into a sequence of characters (zero and one). The median value is estimated as a threshold $T_d$. By comparison with the threshold, the signal data are converted into a 0-1 sequence $P$, given by (4) and (5) (Zhang et al 2001):

$$P = s(1), s(2), \ldots, s(n)$$

where

$$s(i) = \begin{cases} 0 & \text{if } x(i) < T_d \\ 1 & \text{otherwise} \end{cases}$$

Previous studies (Radhakrishnan and Gangadhar 1998, Wu and Xu 1991, Xu et al 1997, Zhang et al 1997, Zhang and Roy 1999) show that 0-1 conversion is enough to study the dynamic complexity of a system. In order to compute $LZ$ complexity, the sequence $P$ is scanned from left to right and the complexity counter $c(n)$ is increased by one unit every time a new subsequence of consecutive characters is encountered. The complexity measure can be estimated using the following algorithm (Zhang et al 1999, Zhang and Roy 2001, Zhang et al 2001):

1. Let $S$ and $Q$ denote two subsequences of $P$ and $SQ$ be the concatenation of $S$ and $Q$, while sequence $SQ\pi$ is derived from $SQ$ after its last character is deleted ($\pi$ means the operation to delete the last character in the sequence). Let $\nu(SQ\pi)$ denote the vocabulary of all different subsequences of $SQ\pi$. At the beginning, $c(n) = 1$, $S = s(1)$, $Q = s(2)$, therefore, $SQ\pi = s(1)$.

2. In general, $S = s(1), s(2), \ldots, s(r)$, $Q = s(r+1)$, then $SQ\pi = s(1), s(2), \ldots, s(r)$; if $Q$ belongs
to $\nu(SQ\pi)$, then $Q$ is a subsequence of $SQ\pi$, not a new sequence.

3. Renew $Q$ to be $s(r+1), s(r+2)$ and judge if $Q$ belongs to $\nu(SQ\pi)$ or not.

4. Repeat the previous steps until $Q$ does not belong to $\nu(SQ\pi)$. Now $Q = s(r+1), s(r+2), ...$, $s(r+i)$ is not a subsequence of $SQ\pi = s(1), s(2), ..., s(r+i-1)$, so increase $c(n)$ by one.

5. Thereafter, $S$ is renewed to be $S = s(1), s(2), ..., s(r+i)$, and $Q = s(r+i+1)$.

These procedures have to be repeated until $Q$ is the last character. At this time the number of different subsequences in $P$ – the measure of complexity – is $c(n)$.

In order to obtain a complexity measure which is independent of the sequence length, $c(n)$ should be normalized. If the length of the sequence is $n$ and the number of different symbols in the symbol set is $\alpha$, it has been prove (Lempel and Ziv 1976) that the upper bound of $c(n)$ is given by:

$$c(n) < \frac{n}{(1 - \epsilon_n) \log\alpha(n)}$$  \hspace{1cm} (6)

where $\epsilon_n$ is a small quantity and $\epsilon_n \to 0$ $(n \to \infty)$. In general, $n \log_d(n)$ is the upper bound of $c(n)$, where the base of the logarithm is $\alpha$, i.e.,

$$\lim_{n \to \infty} c(n) = b(n) = \frac{n}{\log\alpha(n)}$$  \hspace{1cm} (7)

For a 0-1 sequence, $\alpha = 2$, therefore

$$b(n) = \frac{n}{\log_2(n)}$$  \hspace{1cm} (8)

and $c(n)$ can be normalized via $b(n)$:

$$C(n) = \frac{c(n)}{b(n)}$$  \hspace{1cm} (9)

$C(n)$ reflects the arising rate of new patterns along with the sequence.

As it had been done with CTM, we estimated the LZ complexity for all 200 sample epochs of the SaO$_2$ signals. Then, we averaged these values to obtain a single complexity measure per subject.

3.4. Statistical analysis

To asse the diagnostic accuracy of these methods in order to discriminate subjects from OSA positive and OSA negative groups, we built receiver operating characteristic (ROC) curves. We used a computer program developed with Matlab® that calculated the sensitivity-specificity pair for different thresholds depending on the character of the values provided by each method. The sensitivity of a test is the percentage of patients in the OSA positive group correctly diagnosed, whereas the specificity is the percentage of subjects in the OSA negative group rightly classified by the test. The ROC curve is obtained representing sensitivity vs. 1-specificity. The area under the curve provides an idea of how good the test is. Accuracy is another parameter that evaluates globally the precision of a test, representing the portion of subjects under study correctly diagnosed. The optimum threshold is selected by our program 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).

One way ANOVA was applied to check for significant differences between OSA positive and OSA negative groups. $p$-values $< 0.01$ were considered significant.
4. Results

Classical oximetric indexes provided by the oximeter were processed. The $\Delta$ index was calculated using a computer program developed with Matlab®. CTM and LZ values were computed for each $\text{SaO}_2$ record. One way ANOVA and ROC curve analysis were applied to every classical oximetric index and to each non-linear method. ROC curves are displayed in Figure 2. Table 3 shows the statistical and test results for every method.

CT90 provided 71.4% sensitivity, 79.3% specificity and 75.0% accuracy for an optimum threshold of 1.6. With ODIs we obtained sensitivity values of 66.7%, 68.3% and 68.3%, specificity values of 84.9%, 83.0% and 81.1% and accuracy values of 75.0%, 75.0% and 74.1% for ODI2, ODI3 and ODI4 respectively. A sensitivity of 85.6%, a specificity of 81.6% and an accuracy of 84.0% were obtained for the $\Delta$ index with an optimum threshold of 0.17. All classical oximetry indexes provided significant differences ($p < 0.01$) between OSA positive and OSA negative subjects.

Figure 2. ROC curves for non-linear methods and classical oximetric indexes. The diagonal dashed line represents the worst case in which the test classifies correctly a subject with a probability of 0.5.

Table 3. Statistical and test results for classical oximetric indexes and non-linear methods.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>CT90</th>
<th>ODI2</th>
<th>ODI3</th>
<th>ODI4</th>
<th>$\Delta$ index</th>
<th>CTM</th>
<th>LZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSA negative</td>
<td>4.56 ± 17.05</td>
<td>6.07 ± 10.03</td>
<td>4.37 ± 7.91</td>
<td>3.12 ± 6.02</td>
<td>0.14 ± 0.14</td>
<td>0.71 ± 0.18</td>
<td>0.017 ± 0.007</td>
</tr>
<tr>
<td>OSA positive</td>
<td>23.93 ± 29.90</td>
<td>27.18 ± 24.54</td>
<td>25.69 ± 23.31</td>
<td>24.07 ± 22.46</td>
<td>0.44 ± 0.29</td>
<td>0.30 ± 0.20</td>
<td>0.029 ± 0.008</td>
</tr>
<tr>
<td>F statistic</td>
<td>17.46</td>
<td>34.35</td>
<td>40.36</td>
<td>43.45</td>
<td>68.73</td>
<td>203.50</td>
<td>119.05</td>
</tr>
<tr>
<td>Optimum threshold</td>
<td>1.6</td>
<td>10</td>
<td>6.7</td>
<td>4</td>
<td>0.17</td>
<td>0.58</td>
<td>0.022</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>71.4</td>
<td>66.7</td>
<td>68.3</td>
<td>68.3</td>
<td>85.6</td>
<td>90.1</td>
<td>86.5</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>79.3</td>
<td>84.9</td>
<td>83.0</td>
<td>81.1</td>
<td>81.6</td>
<td>82.9</td>
<td>77.6</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>75.0</td>
<td>75.0</td>
<td>75.0</td>
<td>74.1</td>
<td>84.0</td>
<td>87.2</td>
<td>82.9</td>
</tr>
<tr>
<td>ROC area</td>
<td>0.77</td>
<td>0.74</td>
<td>0.76</td>
<td>0.78</td>
<td>0.90</td>
<td>0.92</td>
<td>0.87</td>
</tr>
</tbody>
</table>
We used a computer program developed with Matlab® to represent the second order difference plots for all the recordings corresponding to both OSA positive and OSA negative groups. Figure 3(a) represents the second-order difference plot for a representative OSA positive subject, whereas figure 3(b) depicts the scatter plot for a common OSA negative subject. We can notice that points corresponding to OSA negative patients tend to be located close round the origin, while those points in OSA positive subjects are widely spaced in the diagram. As we previously said in the methodological section, we computed $CTM$ with various radii and we applied one way ANOVA to select the radius for which we obtain the most significant differences. We chose a radius equal to 0.25 to compute the $CTM$. With this radius the mean $CTM$ value for the OSA positive group was $0.30 \pm 0.20$ (mean $\pm$ SD). The $CTM$ values corresponding to the OSA negative group were significantly higher, according to their lower dispersion, with a mean $CTM$ value of $0.71 \pm 0.18$ (mean $\pm$ SD). The ROC curve corresponding to the $CTM$ analysis is displayed in figure 2. With this method we achieved an area under the ROC curve of 0.92. A sensitivity of 90.1%, specificity of 82.9% and accuracy of 87.2% was obtained for an optimum threshold of 0.58. Figure 4 represents the distribution of the $CTM$ values for each population. The dashed line depicts the optimum decision threshold provided by the ROC curve. We could distinguish two different but overlapped scattering areas for each group under study.

We estimated the $LZ$ complexity values for each subject using a computer program developed with Matlab®. We used these complexity values as thresholds to generate the ROC curve displayed in figure 2. The area under the curve was 0.87. The optimum threshold was found at 0.022. With this cut-off point, the method provided 86.5% sensitivity, 77.6% specificity and 82.9% accuracy. Figure 5 depicts these complexity values for every subject in both OSA positive and OSA negative groups. The dashed line represents the optimum threshold provided by the ROC curve. Patients in OSA positive group are generally characterized by higher complexity values (in the sense of number of new subsequences or distinct patterns contained in the time record) than OSA negative subjects. The mean $LZ$ complexity value for OSA positive patients was $0.029 \pm 0.008$ (mean $\pm$ SD). For subjects in the OSA negative group we obtained a mean complexity measure of $0.017 \pm 0.007$ (mean $\pm$ SD). Significant differences between both groups were found with one way ANOVA ($p < 0.01$).

![Second order difference plots](image)

*Figure 3.* Second order difference plots (a) for a representative OSA positive subject and (b) for a representative OSA negative subject.
Figure 4. Mean $CTM$ values of oxygen saturation ($\text{SaO}_2$) signals from nocturnal oximetry for patients in the OSA positive and OSA negative groups. The horizontal dashed line, set at 0.58, represents the optimum threshold between both populations.

Figure 5. Mean $LZ$ values of oxygen saturation ($\text{SaO}_2$) signals from nocturnal oximetry for patients in the OSA positive and OSA negative groups. The horizontal dashed line, set at 0.0218, represents the optimum threshold between both populations.

5. Discussion

The validity of nocturnal oximetry as a single screening method in the diagnosis of OSA has been widely studied in recent years (Lévy et al. 1996, Chiner et al. 1999, Netzer et al. 2001). This technique is a desirable alternative to PSG, because oximetry usually provides sufficient information about respiratory patterns as well as being simpler and less expensive (Netzer et al.
The most widely used oximetric indexes are the ODIs. However, consensus does not yet exist regarding desaturation thresholds for the diagnosis of OSA (Netzer et al 2001).

In this study, we applied two non-linear methods to quantify the effects in $SaO_2$ signals of recurrent apnoea events typical of OSA. To the best of our knowledge, there are no previous studies using non-linear analysis as a screening tool for OSA. Our findings indicate that CTM analysis provides greater variability values in $SaO_2$ signals corresponding to OSA positive patients, compared to OSA negative subjects. LZ complexity results are in line with this, providing higher complexity values in $SaO_2$ signals recorded from subjects affected by OSA syndrome. These results are similar to those obtained by authors using other methodology, such as the $\Delta$ index (Lévy et al 1996, Olson et al 1999, Magalang et al 2003). We also applied classical oximetric indexes. Non-linear methods improve the accuracy of CT90 and ODIs. CTM and LZ provide significantly higher sensitivity values than common oximetric indexes, maintaining similar specificity values. Looking at $\Delta$ index results, CTM enhances both sensitivity and specificity, leading to an improved accuracy. LZ provides higher sensitivity but lower specificity, resulting in a similar accuracy.

Regarding CTM results, specificity was 82.9% with 13 False Positive (FP) cases, 6 of which (46%) suffered from COPD, and another 2 subjects had a BMI $> 34$ kg/m$^2$. If we consider only the 7 non-COPD patients, specificity increases to 87.5%. Regarding the OSA positive group, sensitivity was 90.1% with 11 False Negative (FN) cases, 5 of which had an AHI $< 15$ events/hour. If we consider only patients with an AHI $\geq 15$ events/hour, sensitivity improves to 94%. With a similar approach, we achieved 76.8% specificity and 89% sensitivity values for the LZ complexity analysis.

Methodological differences should be taken into account when comparing our results to previous studies. In the first place, not all studies are based on the same definition of AHI. For example, some previous studies conservatively employ values of AHI $\geq 15$ events/hour (Ryan et al 1995, Lévy et al 1996, Lacassagne et al 1997, Sano et al 1998, Olson et al 1999, Magalang et al 2003), whereas others use AHI $\geq 5$ events/hour to diagnose mild OSA (Chaudhary et al 1998, Littner 2000, Fernández et al 2004). We used a threshold of AHI $\geq 10$ events/hour.

Other studies based on classical oximetry analysis achieved better or similar values in either sensitivity or specificity, but not both. The studies by Rodríguez et al (1996), Lacassagne et al (1997), Zamarrón et al (1999), Olson et al (1999) and Magalang et al (2003) provided higher sensitivity but lower specificity, whereas Ryan et al (1995) and Brouillette et al (2000) achieved higher specificity but lower sensitivity. Our results demonstrate that both CTM and LZ methods provide significant sensitivity and specificity values. This may be as a result of determining the optimum threshold of the ROC curve as the closest point to the left top corner. We found three studies that presented similar results, although they have other limitations. One was based on variability measures (Lévy et al 1996) and two were based on ODIs (Sano et al 1998, Nuber et al 2000). The study carried out by Lévy et al (1996), which quantifies variability via the $\Delta$ index, does not improve the statistics obtained measuring variability with the CTM. The specificity achieved by Sano (1998) is greater than that obtained in our study with both CTM and LZ. However, the sensitivity decreases to 73.5%, leading to worse accuracy. Nuber et al (2000) achieve high sensitivity (91.8%) and good specificity (77.8%), but their study is based on a small sample (40 subjects).

We should take into account some limitations of our study. Although our sample size is an improvement over previous studies, it could be larger. Further research is required to prospectively assess the accuracy of our methodology, i.e. enlarging the data set and considering a wider spectrum of sleep-related breathing disorders. Furthermore, the proposed methods should be applied to healthy subjects to completely assess our methodology. In addition, the data collection process could be enhanced. Our oximeter takes one sample every five seconds. Sampling at higher frequencies we could record $SaO_2$ signals more accurately, improving subsequent analyses. The selection of the optimum radius for the CTM as the one that achieves the most significant differences in terms of the $p$-value is another limitation of our study. Further research is needed applying our methodology to a larger population to validate the results obtained with this radius.
In summary, this study showed that non-linear analysis of blood oxygen saturation from nocturnal oximetry provides useful information for the diagnosis of OSA syndrome. CTM and LZ significantly improved the accuracy of classical oximetric indexes in most cases. With CTM and LZ we obtained high sensitivity and we maintained good specificity values. The best results were achieved applying CTM, with 90.1% sensitivity, 82.9% specificity and 87.2% accuracy. This suggests that non-linear analysis provides suitable methods to study OSA and it seems to fit well with \( \text{SaO}_2 \) behaviour.

Acknowledgments

This work has been partially supported by a grant project from SOCALPAR (Sociedad Castellano-Leonesa y Cántabra de Patología Respiratoria) and Consejería de Sanidad de la Junta de Castilla y León under project VA15/04.

References


Agency for Health Care Policy and Research (AHCPR) 1999 Systematic Review of the Literature Regarding the Diagnosis of Sleep Apnea: Summary, Evidence Report/Technology Assessment Number 1 vol E002 (Department of Health and Human Services) (U. S. Public Health Service)

American Academy of Sleep Medicine Task Force (AASM) 1999 Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research Sleep 22 667–89


Gagnadoux F, Pelletier-Fleury N, Philippe C, Rakotonanahary D and Fleury B 2002 Home Unattended vs Hospital Telemonted Polysomnography in Suspected Obstructive Sleep Apnea Syndrome Chest 121 753–8
Non-linear analysis of oximetric recordings for apnoea detection


Golpe R, Jimenez A, Carpioz R and Cifrian JM 1999 Utility of home oximetry as a screening test for patients with moderate to severe symptoms of obstructive sleep apnea *Sleep* 22 932–7


Littner M 2000 Polysomnography in the Diagnosis of the Obstructive Sleep Apnea-Hypopnea Syndrome *Chest* 118 286–8


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


Otu H H and Sayood K 2003 A new sequence distance measure for phylogenetic tree construction *Bioinformatics* 19 2122–30


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


