*Manuscript - marked version

Regularity analysis of nocturnal oximetry recordings to assist in the diagnosis of sleep apnoea syndrome

J. Víctor Marcos, Roberto Hornero, Ian T. Nabney, Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, Félix del Campo

Number of words: 4850

Corresponding author: J. Víctor Marcos

- Address: Grupo de Ingeniería Biomédica, E.T.S. Ingenieros de Telecomunicación, Universidad de Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain
- Phone: 0034 983 184713
- E-mail: jvmarcos@gmail.com

Regularity analysis of nocturnal oximetry recordings to assist in the diagnosis of sleep apnoea syndrome

J. Víctor Marcos[a,*], Roberto Hornero[a], Ian T. Nabney[b], Daniel Álvarez[a], Gonzalo C. Gutiérrez-Tobal[a], Félix del Campo[c]

[a] *Grupo de Ingeniería Biomédica, E.T.S. Ingenieros de Telecomunicación, Universidad de Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain*

[b] *Non-linearity and Complexity Research Group, Aston University, Aston Triangle, B4 7ET, Birmingham, United Kingdom*

[c] *Hospital Universitario Pío del Río Hortega de Valladolid, Dulzaina 2, 47013, Valladolid, Spain*

Abstract

The relationship between sleep apnoea-hypopnoea syndrome (SAHS) severity and the regularity of nocturnal oxygen saturation (SaO_2) recordings was analysed. Three different methods were proposed to quantify regularity: approximate entropy (AEn), sample entropy (SEn) and kernel entropy (KEn). A total of 240 subjects suspected of suffering from SAHS took part in the study. They were randomly divided into a training set (96 subjects) and a test set (144 subjects) for the adjustment and assessment of the proposed methods, respectively. According to the measurements provided by AEn, SEn and KEn, higher irregularity of oximetry signals is associated with SAHS-positive patients. Receiver operating characteristic (ROC) and Pearson correlation

*Corresponding author. Phone: 0034 983 184713

Email addresses: jvmarcos@gmail.com (J. Víctor Marcos), robhor@tel.uva.es (Roberto Hornero), i.t.nabney@aston.ac.uk (Ian T. Nabney), dalvgon@gmail.com (Daniel Álvarez), gguttob@gmail.com (Gonzalo C. Gutiérrez-Tobal), fsas@telefonica.net (Félix del Campo)

analyses showed that KEn was the most reliable predictor of SAHS. It provided an area under the ROC curve of 0.91 in two-class classification of subjects as SAHS-negative or SAHS-positive. Moreover, KEn measurements from oximetry data exhibited a linear dependence on the apnoea-hypopnoea index, as shown by a correlation coefficient of 0.87. Therefore, these measurements could be used for the development of simplified diagnostic techniques in order to reduce the demand for polysomnographies. Furthermore, KEn represents a convincing alternative to AEn and SEn for the diagnostic analysis of noisy biomedical signals.
Keywords: Oxygen saturation, Entropy rate, Approximate entropy, Sample entropy, Kernel entropy, Density estimation

1. Introduction

Regularity is defined as the consistency of subpattern recurrence in a time series [1]. Regularity has shown to be a useful property of biomedical signals to discriminate those either generated by pathological systems or by the same system under different conditions [2]. Regular signals are characterised by a predictable behaviour, with recognizable patterns that repeat. Regularity is associated with the amount of information in a series, which, in a probabilistic sense, is a measure of the unexpectedness in the data [3]. Shannon [4] proposed the concept of entropy to evaluate the information (or uncertainty) in a message, which is modelled as a finite collection of random variables. In the context of infinite sequences or series, the entropy rate has been employed for the quantification of the amount of information [2]. Several metrics have been proposed to estimate the entropy rate of a series, with approximate entropy (AEn) [5] and sample entropy (SEn) [6] being the most common ones. A generalised entropy measure is given by the family of Renyi entropies (R_q), where q denotes the entropy order [7]. Lake [8] analysed the incorporation of the Renyi entropy into the entropy rate framework, showing that AEn and SEn approximate the differential Renyi entropy rate for $q = 1$ and $q = 2$, respectively.

AEn and SEn are based on the computation of probabilities by counting matches between signal subsequences of length m and $m + 1$. A match is found when the distance between two subsequences is lower or equal than a tolerance parameter r [6]. A different procedure to obtain the Renyi entropy rate of a series consists of substituting probability terms in AEn and SEn algorithms by the corresponding probability density functions [8, 3]. Several advantages are found in this approach. It suppresses the need of predefined rules for the choice of the tolerance parameter r, which can be freely varied in order to obtain confident estimates of the density functions. In addition, entropy estimates made with different values of r measure the same inherent quantity and can be compared directly [8, 9].

This approach requires the approximation of the (unknown) probability density function of the data, for which a finite set of samples extracted from the underlying series is initially available. Non-parameteric kernel density estimation based on the Parzen window method has been suggested for this purpose [8, 3]. Specifically, Gaussian kernels are of special interest since they result in a smooth and continuous profile of the approximated density [10]. Additionally, in the case of the quadratic entropy (R_2), i.e., the Renyi entropy of order $q = 2$, Gaussian kernels lead to the exact evaluation of the integral found in its definition [3]. In a preceding study, a kernel-based estimation of R_2 was adopted to assess the quadratic entropy rate of a time series [11]. The resulting measure, termed as kernel entropy (KEn), was proposed as an indicator of the irregularity of the series [11, 12].

Entropy analysis has yield successful results in several applications involving time series processing such as earthquake forecasting [13], exchange rating [14] or fault detection [15]. Furthermore, entropy measures of biomedical signals have been widely used to assess physiological differences between subjects [16, 17]. The present study focuses on this scenario. We explored the utility of entropy rate measurements of nocturnal oxygen saturation signals (SaO_2) in the context of sleep apnoea-hypopnoea syndrome (SAHS) diagnosis. Nowadays, a definitive diagnosis about SAHS is obtained from in-hospital evaluation of the patient's sleep through nocturnal polysomnography (PSG). This test enables the assessment of SAHS severity by means of the apnoea-hypopnoea index (AHI), which quantifies the number of apnoea and hypopnoea events per hour of sleep. To obtain the AHI of a patient, the sleep specialist must evaluate a large amount of clinical and physiological data that, in addition to SaO_2 series, include other signals such as the electrocardiogram (ECG), the electroencephalogram (EEG) or the respiratory airflow (AF) [18]. Therefore, PSG is a highly complex and time-consuming procedure.

Reliable indicators of SAHS severity automatically extracted from these data would enable an objective and simplified interpretation. Nocturnal oximetry recordings are of special interest as they reflect respiratory dynamics during sleep. Apnoeas and hypopnoeas are usually accompanied by hypoxaemia due to airflow reduction, which is reflected by a marked drop in the saturation value [19]. The diagnostic utility of oximetry signals has been previously evaluated through different methods. A straightforward approach is the use of oximetry parameters based on the computation of desaturation events or the time spent below a certain level of saturation [20, 21]. In addition, complex signal processing and pattern recognition techniques like neural networks or genetic algorithms have been employed for the extraction of useful descriptors from SaO_2 data [22, 23, 24]. According to the reported results, a higher diagnostic accuracy can be obtained through the combination of different features including statistical, spectral and non-linear ones. Correct diagnostic rates close to 90% have been reported for screening algorithms based on this approach [25, 26, 22].

Among other features, SaO_2 irregularity measured by the entropy rate has been employed as a descriptor of the influence of SAHS severity on its dynamic behaviour [27, 25]. The non-deterministic occurrence of apnoeic episodes tends to increase the uncertainty in the SaO_2 signal and, equivalently, its amount of information. As a result, signals from subjects suffering from SAHS are expected to have a higher entropy rate than those from control subjects. Previously, AEn has been employed to measure SaO_2 irregularity [28, 27]. These preceding studies showed the relationshp between higher irregularity of oximetry signals and SAHS severity, estimating that a correct diagnosis based on regularity analysis can be obtained for approximately 85% of the patients. However, AEn has proven to be a biased entropy estimator [6] and, thus, further analysis is required to extract robust conclusions on the relationship between SAHS severity and SaO_2 irregularity.

To this end, the present study proposes a comparative analysis between different entropy metrics. In addition to AEn, we suggest entropy analysis of SaO_2 series based on SEn and KEn, which provide two different approaches to estimate the quadratic entropy rate of a signal. The present study aims to determine to which extent the irregularity of SaO_2 data is related to SAHS severity, as well as the most accurate method to quantify this relationship.

We hypothesise that a more confident assessment of the entropy of SaO_2 recordings can be obtained by means of kernel-based approximations to probability density functions as implemented by KEn. This method represents a novel approach for entropy estimation with respect to conventional procedures like AEn and SEn. The framework implemented by KEn suitably adapts to oximetry analysis since SaO_2 samples can be interpreted as observations of a continuous variable. Thus, probability density functions may provide a more reliable description of their statistical behaviour. This hypothesis is evaluated through an exhaustive regularity analysis of SaO_2 data using AEn, SEn and KEn.

2. Materials and methods

2.1. Subjects and signals

A total of 240 subjects suspected of suffering from SAHS took part in the study. They underwent PSG in the Sleep Unit of Hospital Universitario Pío del Río Hortega, Valladolid, Spain. The Review Board on Human Studies approved the protocol and each subject gave their consent to participate in the study. To draw useful conclusions on the effect of SAHS on SaO_2 dynamics, subjects affected by any other relevant respiratory disorder were excluded. The selected patients were continuously monitored using a polysomnograph (Alice 5, Respironics, Philips Healthcare, The Netherlands). A medical expert analysed the PSG recordings according to the rules proposed by Rechtschaffen and Kales [29]. Once apnoeas and hypopnoeas were identified, the AHI was obtained as the total number of events (i.e., the sum of apnoeas and hypopnoeas) divided by the total sleep time. The resulting value is expressed as the number of events per hour of sleep [30]. A threshold given by AHI = 10 h^{-1} was used to determine a positive diagnosis of SAHS [31].

A Nonin PureSAT pulse oximeter (Nonin Medical Inc., USA) was used to record oximetry signals at a sampling frequency of 1 Hz. These signals were subsequently saved to separate files to be processed offline. A preprocessing stage was initially applied to remove artefacts like marked drops or zero samples due to a bad contact of the probe during sleep. The criteria suggested by Magalang et al. [32] were taken into account to perform signal preprocessing. According to these criteria, all changes greater than 4%/s between consecutive sampling intervals and any sample lower than 20% were removed.

Figure 1 shows two oximetry recordings from our dataset once artefacts were removed. The signals correspond to a normal subject (AHI = 0.5 h^{-1}) and a subject with severe SAHS (AHI = 32.1 h^{-1}), respectively. In addition, a detailed view (12 minutes) of both recordings is provided. The differences between these signals reflect the influence of repeated apnoeas and hypopnoeas on SaO_2 dynamics. The signal from the normal subject is characterised by a near-constant saturation value along the night, with small fluctuations around the baseline level. This behaviour is confirmed when observed in detail, as it exhibits some variability without marked desaturation events. In contrast, the profile of the signal from the subject with severe SAHS reflects a significant instability as a consequence of repeated desaturations accompanying apnoeas and hypopnoeas. As observed in the zoomed segment of the signal, these desaturation events are more frequent when compared with the oximetry recording from the normal subject. Additionally, they are more pronounced and longer. Therefore, a distinct value of the entropy rate can be expected for these signals since they reflect different dynamics.

INSERT FIGURE 1 AROUND HERE

The hold-out method was used to prevent bias in the estimation of the performance of the three entropy metrics [10]. Therefore, the initial population was randomly divided into a training set with 96 subjects (40%) and a test set with 144 subjects (60%). The former was used to adjust user-dependent parameters in AEn, SEn and KEn algorithms. Signals in the test set were used to assess the diagnostic capability of these methods. Table 1 summarises the demographic and clinical data for subjects in training and test sets. Note that a higher proportion of SAHS-positive subjects was obtained due to the initial suspicion of SAHS in the population under study. In addition, the higher percentage of older males is motivated by the increased prevalence of SAHS in this group, as reported in previous studies. In the landmark study of the Wisconsin Sleep Cohort, including 602 men and women, 24% of men and 9% of women had AHI ¿ 5 h^{-1}, while 9% of men and 4% of women had AHI ¿ 15 h^{-1} [33]. In addition, it has been estimated that AHI increases with age for both men and women [34].

INSERT TABLE 1 AROUND HERE

2.2. Methods

We compared the utility of three different entropy metrics, namely AEn, SEn and KEn, to quantify the relationship between SAHS severity and the irregularity of SaO_2 data. The proposed methods represent distinct implementations of the Renyi entropy rate of a series. The mathematical definition of the latter is derived from the expression of the Renyi entropy, which is obtained as follows. Let $\mathbf{x} = [x_1, x_2, ..., x_N]$ be a continuous stochastic process composed of a sequence of N random variables x_i. The Renyi entropy of the process is given by [7]:

$$
\begin{aligned}
R_{q,N} &= \frac{1}{1-q}\log\left\{E\left[f^{q-1}\left(x_1,\ldots,x_N\right)\right]\right\} \\
&= \frac{1}{1-q}\log\int f^{q}\left(x_1,\ldots,x_N\right)dx_1\ldots dx_N
\end{aligned} \tag{1}
$$

where q determines the entropy order and $f(\mathbf{x}) = f(x_1, \ldots, x_N)$ is the joint probability density function of the set of variables x_i that compose the process. The Renyi entropy evaluates the amount of information (uncertainty) in $\mathbf{x}$.

Adding new variables to the process $\mathbf{x}$ will contribute to increase its information content, showing the dependence of the entropy on the process length specified by N [2]. Thus, a measure of the variation of the entropy due to the inclusion of a new variable x_i, i.e., the entropy rate, can be obtained. The differential Renyi entropy rate is defined by the following asymptotic limit [8]:

$$
\begin{aligned}
D_{q,N} &= \lim_{N\to\infty}\left(R_{q,N+1} - R_{q,N}\right) \\
&= \lim_{N\to\infty}\left\{-\log\left[\int f^{q}\left(\mathbf{x}^{(N+1)}\right)d\mathbf{x}^{(N+1)}\right] + \log\left[\int f^{q}\left(\mathbf{x}^{(N)}\right)d\mathbf{x}^{(N)}\right]\right\}
\end{aligned} \tag{2}
$$

where the superscripts (N) and $(N+1)$ denote the length of the process $\mathbf{x}$. In the following sections, a description of AEn, SEn and KEn is provided, showing the connection between each of these metrics and $D_{q,N}$ as expressed in (2).

2.2.1. Approximate entropy (AEn)

AEn is a family of metrics developed by Pincus [5] for the analysis of noisy data such as biomedical signals. Briefly, AEn estimates the entropy rate of a series from the computation of the probability of repetition for subsequences of length m. To this end, two subsequences are considered similar if the distance between them is lower than a threshold r. Mathematically, the algorithm to compute AEn is defined for a finite time series of length N given by $\mathbf{x} = [x_1, \ldots, x_N]$. From this series, extract overlapping m-length windows, from $\mathbf{x}_i^{(m)}$ to $\mathbf{x}_{N-m+1}^{(m)}$, defined as $\mathbf{x}_i^{(m)} = [x_i, x_{i+1}, \ldots, x_{i+m-1}]$. The distance $d\left[\mathbf{x}_i^{(m)}, \mathbf{x}_j^{(m)}\right]$ between two such vectors $\mathbf{x}_i^{(m)}$ and $\mathbf{x}_j^{(m)}$ is calculated as the maximum absolute difference between their respective scalar components. A tolerance value r is used to obtain the number of vectors $(j = 1, \ldots, N - m + 1)$ such that $d\left[\mathbf{x}_i^{(m)}, \mathbf{x}_j^{(m)}\right] \leq r$, which is denoted as $N^m(i)$. Then, the likelihood that a vector is within a distance r from vector $\mathbf{x}_i^{(m)}$ $(i = 1, \ldots, N - m + 1)$ is estimated as:

$$C_r^m(i) = \frac{N^m(i)}{N - m + 1} \tag{3}$$

The term $C_r^m(i)$ reflects the regularity of patterns of length m similar to $\mathbf{x}_i^{(m)}$ for a tolerance r. Equivalently, it estimates the probability of observing the m-length vector $\mathbf{x}_i^{(m)}$, implementing a discrete approximation to the probability density function $f\left(\mathbf{x}^{(m)}\right)$.

The average of the logarithmic likelihood over the complete set of samples is computed as:

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \log C_r^m(i) \tag{4}$$

From the previous analysis, $\phi^m(r)$ represents an estimation of the expected value of $\log\left[f\left(\mathbf{x}^{(m)}\right)\right]$, which in turn corresponds to the negative value of the Renyi entropy of order $q = 1$ [8]:

$$\phi^m(r) \approx E\left\{\log\left[f\left(\mathbf{x}^{(m)}\right)\right]\right\} = \int \log\left[f\left(\mathbf{x}^{(m)}\right)\right] f\left(\mathbf{x}^{(m)}\right) d\mathbf{x}^{(m)} = -R_{1,N} \tag{5}$$

Therefore, AEn approximates $D_{1,N}$, the differential Renyi entropy rate of order $q = 1$ [8], as expressed by its mathematical definition [5]:

$$AEn(m,r) = \lim_{N\to\infty}\left[\phi^m(r) - \phi^{m+1}(r)\right] \tag{6}$$

The following statisitc is adopted for the computation of AEn on finite time series:

$$AEn(m,r,N) = \phi^m(r) - \phi^{m+1}(r) \tag{7}$$

2.2.2. Sample entropy (SEn)

SEn adopts a similar approach to AEn for the estimation of the entropy rate, which is based on the estimation of the probability of repetition for a subsequence. As a substantial difference, self-matching is prevented in the SEn algorithm. It has been shown that self-matching, i.e., the comparison of a vector $\mathbf{x}_i^{(m)}$ with itself, introduces some bias in the computation of AEn. Richman and Moorman proposed SEn in order to avoid this problem [6]. For a time series $\mathbf{x} = [x_1, \ldots, x_N]$, only the first $N - m$ vectors of length m are considered for comparison in the SEn algorithm. Let $U_r^m(i)$ denote the probability that a m-length vector $\mathbf{x}_j^{(m)}$ is within a distance r from vector $\mathbf{x}_i^{(m)}$:

$$U_r^m(i) = \frac{N^m(i)}{N-m-1} \tag{8}$$

where $N^m(i)$ denotes the number of vectors $\mathbf{x}_j^{(m)}$ $(j \neq i, j = 1, \ldots, N-m)$ such that $d\left[\mathbf{x}_i^{(m)}, \mathbf{x}_j^{(m)}\right] \leq r$ and r is a tolerance value. Similarly to the term $C_r^m(i)$ in the AEn algorithm, $U_r^m(i)$ plays the role of the density function $f\left(\mathbf{x}^{(m)}\right)$, since it approximates the probability of observing $\mathbf{x}_i^{(m)}$. The average of the set of $U_r^m(i)$ values is given by the quantity $U^m(r)$:

$$U^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} U_r^m(i) \tag{9}$$

which provides an estimation of the expected value of the function $f\left(\mathrm{x}^{(m)}\right)$:

$$U^m(r) \approx E\left[f\left(\mathbf{x}^{(m)}\right)\right] = \int f^2\left(\mathbf{x}^{(m)}\right) d\mathbf{x}^{(m)} \tag{10}$$

It can be observed that the integral in the previous equation corresponds to the argument of the logarithm found in the definition of the Renyi entropy of order $q = 2$ or quadratic entropy. As a result, SEn can be interpreted as an estimation of the quadratic entropy rate as expressed by the following equation [6]:

$$SEn(m, r) = \lim_{N \to \infty} \left\{\log\left[U^m(r)\right] - \log\left[U^{m+1}(r)\right]\right\} \tag{11}$$

SEn is estimated by the statistics:

$$SEn(m, r, N) = \log\left[U^m(r)\right] - \log\left[U^{m+1}(r)\right] \tag{12}$$

2.2.3. Kernel entropy (KEn)

Unlike AEn and SEn algorithms, which involve the computation of probabilities, KEn is based on modelling the statistical properties of the generator of the data. Hence, the algorithm to obtain KEn estimates the density function of the signal subsequences. To this end, KEn approximates the quadratic entropy rate from the use of non-parameteric probability densitiy estimation techniques. The KEn assumes that the set of m-length vectors $\mathbf{x}_i^{(m)}$ from the original series has been generated according to the density function $f\left(\mathbf{x}^{(m)}\right)$ [8]. Hence, the Parzen window method with Gaussian kernels is proposed to estimate this function [11, 35, 3], resulting in the following expression:

$$\hat{f}\left(\mathbf{x}^{(m)}\right) = \frac{1}{N-1}\sum_{i=1}^{N-m+1} G\left(\mathbf{x}^{(m)} - \mathbf{x}_i^{(m)}, \mathbf{\Sigma}\right) \tag{13}$$

where $G\left(\mathbf{x}^{(m)}, \mathbf{\Sigma}\right)$ denotes the zero-mean Gaussian kernel with covariance matrix $\mathbf{\Sigma}$ evaluated at point $\mathbf{x}^{(m)}$. In our study, spherical Gaussians with a covariance matrix given by $\mathbf{\Sigma} = \sigma\mathbf{I}$ will be considered. The scalar σ is referred to as the kernel bandwidth and $\mathbf{I}$ denotes the identity matrix.

Using $\hat{f}\left(\mathbf{x}^{(m)}\right)$, the expected value of the density function $f\left(\mathrm{x}^{(m)}\right)$ is estimated by $J^m\left(\sigma\right)$, which is given by:

$$J^m\left(\sigma\right) = \int \hat{f}^2\left(\mathbf{x}^{(m)}\right) d\mathbf{x}^{(m)} \tag{14}$$

Note that the use of a Gaussian kernel estimator for $f\left(\mathbf{x}^{(m)}\right)$ enables the computation of the exact value of the integral:

$$\int \hat{f}^2\left(\mathbf{x}^{(m)}\right) d\mathbf{x}^{(m)} = \frac{1}{\left(N-m+1\right)^2} \sum_{i=1}^{N-m+1} \sum_{i=1}^{N-m+1} G\left[\mathbf{x}_j^{(m)} - \mathbf{x}_i^{(m)}, 2\sigma^2\mathbf{I}\right] \tag{15}$$

where $N-m+1$ is the total number of m-length vectors in the original time series. The negative logarithm of $J^m\left(\sigma\right)$ approximates the quadratic entropy $R_{2,N}$. Hence, KEn is an estimation of the quadratic entropy rate, which is obtained from the incorporation of this expression in the Renyi entropy rate framework [8]:

$$KEn\left(m,\sigma\right) = \lim_{N\to\infty} \left\{\log\left[J^m\left(\sigma\right)\right] - \log\left[J^{m+1}\left(\sigma\right)\right]\right\} \tag{16}$$

In practice, the following estimation is used for finite series:

$$KEn\left(m,\sigma,N\right) = \log\left[J^m\left(\sigma\right)\right] - \log\left[J^{m+1}\left(\sigma\right)\right] \tag{17}$$

It is worth noting that different techniques can be used to select the kernel bandwidth parameter σ of the Parzen density estimator from the original data. In this study, the Bayesian approach proposed by Zhang et al. [36] was applied for this purpose. According to this procedure, the elements of the covariance matrix $\mathbf{\Sigma} = \sigma\mathbf{I}$ are treated as parameters and the aim is to estimate their posterior distribution $f\left(\sigma \mid D\right)$, where D denotes the training set.

From Bayes theorem [10], the posterior density $f\left(\sigma \mid D\right)$ is proportional to the product of the prior density $f\left(\sigma\right)$ and the likelihood $f\left(D \mid \sigma\right)$:

$$f\left(\sigma \mid D\right) \propto f\left(D \mid \sigma\right) f\left(\sigma\right) \tag{18}$$

Note that the likelihood of observing the sample set D for a given value of σ can be approximated from the expression of $\hat{f}\left(\mathbf{x}^{(m)}\right)$. Therefore, the prior probability $f(\sigma)$ is the remaining factor to obtain $f(\sigma \mid D)$. Zhang et al. suggested the following functional form for the prior in order to make the sampling algorithm work properly [36]:

$$f(\sigma) \propto \prod_{k=1}^{m} \frac{\sigma_k}{\lambda + \sigma_k^2} \tag{19}$$

where λ controls the shape of the function. This prior density function aims to avoid large values of σ, for which the associated probability is small.

The most probable value of σ given the data in D is selected as the optimum. This value is obtained by sampling from the posterior probability $f(\sigma \mid D)$ using Markov Chain Monte Carlo (MCMC) techniques. The Metropolis-Hastings algorithm was used for this purpose [37, 38].

2.3. Statistical analysis

To assess the performance of the proposed entropy methods in the quantification of SaO_2 irregularity derived from SAHS, two different approaches were used. First, a two-class classification model was defined by dividing the initial population into SAHS-negative and SAHS-positive subjects. Receiver operating characteristic (ROC) analysis was used to assess the capability of AEn, SEn and KEn measurements for identifying SaO_2 signals from these two groups [39]. The area under the ROC curve (AUC), which represents the probability of correct classification for a randomly chosen pair of samples from the two possible categories [39], was used as a measure of classification performance.

The second model for SAHS diagnosis consists in estimating the AHI from the available oximetry signal. According to the value of this index, subjects are assigned to one of four severity groups [18]: no SAHS (AHI $< 5\text{h}^{-1}$), mild SAHS ($5\ \text{h}^{-1} \leq$ AHI $< 15\ \text{h}^{-1}$), moderate SAHS ($15\ \text{h}^{-1} \leq$ AHI $\leq 30\ \text{h}^{-1}$) and severe SAHS (AHI $> 30\ \text{h}^{-1}$). The utility of the three entropy measures to rank SAHS severity was assessed by means of the Pearson's correlation coefficient (ρ). It evaluates the linear relationship between AEn, SEn and KEn and the AHI. The value of ρ can be interpreted as the utility of a given method to predict the AHI.

3. Results

3.1. Selection of the input parameters

Entropy analysis based on the proposed methods require the prior selection of several parameters. These correspond to N, m, and r for AEn and SEn algorithms. In the case of KEn, only N and m are needed since a data-driven technique was used to optimise the value of σ. The parameter N denotes the length of the time series to be processed. In our study, the length of oximetry recordings was approximately 7 hours (i.e., more than 25000 samples), involving a large amount of data. However, as apnoeic events can take place at different moments during sleep and, in particular, during REM phases [18], the whole recording must be analysed for an objective assessment of SaO_2 dynamics. To reduce the computational load, we adopted the strategy suggested in preceding studies [40]. Hence, the original SaO_2 signal was divided into epochs of length $N = 512$ samples to estimate AEn, SEn and KEn. As the duration of apnoeas is typically between 10 seconds and 2 minutes, the chosen epoch length (approximately, 8.5 minutes) is large enough to include one or more complete events. The measurements obtained from all the epochs were then averaged to determine the final estimation for each method.

On the other hand, parameters m and r involve statistical considerations. Setting m large and r too small would result in inaccurate estimates of the probabilities in AEn and SEn algorithms. In contrast, a large value of r and a small m is generally too coarse to distinguish processes. Thus, m and r were set to the widely established values suggested by Pincus [5, 41] to obtain a statistically valid estimate of the proposed entropy measures: $m =$ 1 or 2, and $r = 0.1$, 0.15, 0.2 or 0.25 times the standard deviation (SD) of the original series.

It is worth noting that, while KEn enables the automatic selection of the parameter σ, the applied MCMC-based technique requires several design parameters to be specified. Figure 2 represents the shape of the prior $f(\sigma)$ for $\lambda = 0.1$, 1, 5 and 10. This function tends to be more concentrated near zero as the hyperparameter λ becomes smaller. Zhang et al. [36] did not find significant differences in the estimated density functions resulting from λ between 0.1 and 5. Thus, we set $\lambda = 5$ to avoid an excessive concentration of the probability density in a small range of σ values, favouring a smooth profile of $\hat{f}\left(\mathrm{x}^{(m)}\right)$. Furthermore, the variance of the proposal distribution for the Metropolis-Hastings algorithm must result in an acceptance rate between 20% and 30% of the total number of samples [36]. In our study, this requirement was satisfied by setting that variance to 0.015. Finally, in order to ensure the convergence of the sampling process, the number of samples to be omitted and the number of samples to be retained were set to 5000 while the starting σ value was set to 1% of the SD of the original series.

INSERT FIGURE 2 AROUND HERE

3.2. Training set

All the experiments in our study were conducted on Matlab 8.2.0. Initially, we evaluated different configurations of AEn, SEn and KEn on SaO_2 signals in the training set to find the optimum value of the input parameters m and r. For each method, the configuration with the highest performance was selected. The results achieved on the training set are summarised in Table 2. As can be observed, $m = 1$ and $r = 0.1$SD resulted in a more accurate characterisation of SAHS for AEn and SEn. For the KEn algorithm, the two evaluated configurations provided similar results. An AUC of 0.86 and a correlation coefficient of 0.82 were reached when m was set to 1. This configuration was slightly improved by setting $m = 2$ (AUC = 0.86 and $\rho = 0.83$). Therefore, $m = 2$ was finally selected as the optimum for KEn.

INSERT TABLE 2 AROUND HERE

3.3. Test set

AEn, SEn and KEn were computed on oximetry signals in the test set using the selected configurations. For each of the three methods, the Lilliefors test [42] was applied to assess the normality of the samples. The results showed a level of significance (p-value) higher than 0.1 for the distribution of AEn, SEn and KEn in both SAHS-negative and SAHS-positive groups, reflecting a valid assumption for normality. Subsequently, ROC and correlation analysis were conducted for each method. Table 3 summarises the obtained results. KEn showed to be the most accurate predictor of SAHS from the methods evaluated in our study. It achieved AUC $= 0.91$ and $\rho = 0.87$, which were substantially higher than the results provided by AEn (AUC $= 0.67$ and $\rho = 0.34$) and SEn (AUC $= 0.74$ and $\rho = 0.45$). The experiments reveal that a finer characterisation of SaO_2 irregularity was obtained by using a density estimation technique for entropy quantification as implemented by KEn.

INSERT TABLE 3 AROUND HERE

For each method, Figure 3 depicts the ROC curves and boxplots in SAHS-negative and SAHS-positive groups. The p-value in the figure corresponds to the level of significance for the difference between the means of each entropy metric in both groups, as obtained from the one-way ANOVA test [43]. It confirms that SEn and KEn provided statistically significant differences between both groups of subjects. In particular, a p-value close to 0 was obtained for KEn, reflecting notably more significant differences between SAHS-negative and SAHS-positive samples than AEn and SEn. Moreover, the results reflect the utility of the evaluated methods in two-class classification of patients. As can be observed, the distributions of AEn, SEn and KEn measurements reflect that higher irregularity, i.e., higher entropy rate, is associated with oximetry signals from SAHS-positive subjects.

INSERT FIGURE 3 AROUND HERE

On the other hand, correlation analysis indicates that SaO_2 regularity is directly related to SAHS severity, with larger entropy values obtained from signals corresponding to subjects with a higher AHI. Figure 4 depicts the AHI versus AEn, SEn and KEn measurements as well as the boxplots of these measurements in each severity group. The highest Pearson's correlation coefficient was obtained for KEn ($\rho = 0.87$), which showed a marked linear trend with respect to AHI. The boxplots obtained for the four severity groups show a higher dispersion of AEn and SEn measurements in each category as well as smaller differences between their median values. As a result, a significant overlap between different groups was observed for them. This overlapping was substantially smaller for KEn measurements, which provided a more accurate assessment of SAHS severity.

INSERT FIGURE 4 AROUND HERE

For a more rigorous evaluation of the differences achieved by each entropy metric between the four groups, we applied the one-way ANOVA test. The obtained results are summarised in Table 4. As can be observed, KEn provided statistically significant differences for any pair of severity groups under evaluation, but for the comparison between mild and moderate subjects. In the case of SEn, significant differences were found between normal subjects and patients with moderate SAHS, as well as between normal subjects and patients with a severe diagnosis of SAHS. Finally, AEn only achieved significant differences when subjects in normal and severe groups were compared. Thus, the analysis confirms the higher capability of KEn to capture differences in SaO_2 dynamics due to SAHS severity. In addition, the experiment also shows that signals corresponding to subjects with mild and moderate SAHS tend to reflect a similar behaviour, as no significant differences was observed for any of the evaluated entropy metrics.

INSERT TABLE 4 AROUND HERE

3.4. Comparison with conventional statistical features

To complete our study, we assessed the diagnostic utility of regularity analysis of SaO_2 data with respect to other statistical features commonly employed for the evaluation of biomedical signals. Conventionally, these features include the mean (f_{avg}), standard deviation (f_{sd}), coefficient of variation (f_{cv}), interquartile range (f_{iqr}) and dispersion indices (f_{sd1} and f_{sd2}) derived from the Poincare plot. The results achieved by these features are summarised in the Table 5.

INSERT TABLE 5 AROUND HERE

As can be observed, conventional features capture relevant diagnostic information about SAHS from nocturnal oximetry recordings. In particular, they achieved significantly high AUC values, showing a good capability to discriminate between SAHS-negative and SAHS-positive subjects. Our experiments reflect that all the evaluated features but f_{avg} achieved AUC higher than 0.95, which improves the AUC results provided by the three entropy metrics assessed in our research. However, correlation analysis reveals a lower ability of the conventional statistical features to detect small variations in the AHI. The obtained ρ values are significantly smaller than that achieved by KEn. The latter reached $\rho = 0.87$, whereas the highest correlation coefficient among the conventional features was $\rho = 0.77$, which was provided by

f_{sd1}. Therefore, correlation results show a stronger correspondence between SAHS severity and SaO_2 irregularity when compared to the signal properties evaluated by the proposed conventional methods.

4. Discussion and conclusions

Regularity analysis of SaO_2 recordings was performed using three different entropy algorithms: AEn, SEn and KEn. The obtained measurements show that more irregular signals are associated with SAHS-positive subjects, reflecting the influence of apnoea events on SaO_2 dynamics. Nevertheless, there were substantial differences between the diagnostic performances of AEn, SEn and KEn. This reveals a distinct reliability of the estimators implemented by these algorithms. The latter showed to be the most consistent entropy estimator, outperforming conventional entropy algorithms like AEn and SEn. Specifically, we found that KEn measurements from oximetry data could be used to estimate the AHI of a patient ($\rho = 0.87$).

The KEn method represents a novel strategy for entropy estimation. The main difference between KEn and the conventional AEn and SEn algorithms is the use of probability density functions for modelling the statistical distribution of the data. According to our results, this approach has shown to be a more suitable procedure when continuous variables like SaO_2 are analysed. In addition, the use of probability density functions involves other advantages for entropy estimation. First, a data-driven method as that proposed by Zhang et al. [36] can be applied to determine the value of the kernel bandwidth. As shown in our experiments, the variance σ of the Gaussian kernels is optimised for the series under evaluation instead of taking an arbitrary value fixed by the user. The automatic optimisation of σ suppresses one of the user parameters in AEn and SEn algorithms, reducing by four the number of KEn configurations to be evaluated. Second, it must be taken into account that kernel density estimation results in the exact computation of the integral in the definition of the Renyi entropy, as expressed in (15). Hence, the KEn algorithm avoids one of the approximations adopted in AEn and SEn, which corresponds to the expectation operator.

It is worth noting that the use of density functions overcomes the strong dependency of conventional entropy metrics like AEn or SEn on the tolerance parameter r. Small values of r lead to higher and less confident entropy estimates due to the reduced number of matches of length m and $m + 1$. In the context of density estimation, the value of r is chosen in order to obtain an accurate approximation of the target probability density function [8, 3, 9]. As a result, any value of r can be used for any time series, enabling the comparison between entropy results computed for distinct r [9]. As a matter of further study, the influence of the kernel chosen for density approximation (e.g., Gaussian, uniform, triangular or cosine kernels) on the resulting entropy estimates should be assessed.

From our experiments, we have found that SaO_2 irregularity is more closely related to SAHS severity than conventional statistical properties including mean, variance or dispersion indices extracted from Poincar plots. It is reflected by the Pearson's correlation coefficient, which was close to 0.90 in the case of KEn. This results shows the utility of the entropy rate of oximetry data to discriminate subjects with subtle differences between their AHI. Nevertheless, conventional statistical features should be taken into account for the implemention of accurate methods based on oximetry data. As reported in preceding studies, multivariate analysis including several uncorrelated features can yield higher diagnostic performance than univariate measures of a given signal property [22, 24].

Several limitations can be found in our study. We have demonstrated that KEn is a valuable approach to perform regularity analysis of SaO_2 data in the context of SAHS diagnosis. However, the information captured by KEn is not sufficient to provide a definitive diagnosis about SAHS as reflected by AUC < 1 and $\rho < 0.9$. KEn analysis should be then considered as a tool for the interpretation of SaO_2 data. In this vein, the role of nocturnal pulse oximetry in SAHS diagnosis must be analysed. As in our study, the results reported by other researchers reveal suboptimal diagnostic performance of oximetry-based methods, with sensitivity and specificity lower than 100% [26, 22, 24]. Hence, the number of false negative and false positive cases would prevent the use of these methods as an alternative for PSG. Instead, home unattended pulse oximetry could be adopted as a screening tool to reduce the number of required PSG tests, contributing to minimise the waiting time for a diagnosis about SAHS. On the other hand, KEn is computationally more expensive than AEn and SEn. We have estimated that the time required to compute KEn is approximately a hundred times that of AEn and SEn. Specifically, we estimated that the time required to compute AEn, SEn and KEn on a signal epoch (512 samples) was 0.72, 0.91 and 96.78 seconds, respectively, using Matlab 8.2.0 on an Intel i7 CPU at 3.4 GHz. Mainly, this difference is due to the MCMC-based procedure for automatic adjustment of σ. This procedure requires thousands of samples to converge, with several operations carried out for each of these samples. In order to reduce the computational load, the value of σ could be previously set by the user for the estimation of KEn. Nevertheless, this does not ensure the choice of the optimum σ for the underlying data.

In summary, our study confirms that SaO_2 signals from patients suffering from SAHS tend to be more irregular. This result suggests the unpredictable occurrence of apnoeas and hypopnoeas during sleep. Hence, regularity analysis could be used for the interpretation of nocturnal oximetry dynamics in the context of SAHS diagnosis. In particular, KEn showed to be a reliable predictor of SAHS. It could be considered to build new methods for automatic assessment of SAHS severity in order to reduce the demand for conventional PSG. Furthermore, it has been proved that KEn is a valuable metric for regularity analysis of biomedical data, representing an alternative to conventional methods such as AEn and SEn.

Acknowledgments

We would like to thank Dr. Dan Woodcock for his assistance with some of the experiments involved in this research work. Declaration statements: (i) competing interests: none; (ii) funding: this study has been partly funded by project VA059U13 from Junta de Castilla y León and project TEC 2011-22987 from Ministerio de Economía y Competitividad and FEDER grant; (iii) the study involves human subjects: the Review Board on Human Studies of the Hospital Universitario Pío del Río Hortega approved the protocol of this study and each subject gave their consent to participate.

References

[1] Pincus, S.M., Hartman, M.L., Roelfsema, F., Thorner, M.O., Veldhuis, J.D.. Hormone pulsatility discrimination via coarse and short time sampling. American Journal of Physiology-Endocrinology And Metabolism 1999;277(5):E948–E957.

[2] Costa, M., Goldberger, A.L., Peng, C.K.. Multiscale entropy analysis of biological signals. Phys Rev E 2005;71(2):021906.

[3] Principe, J.C.. Information theoretic learning: Rényi's entropy and kernel perspectives. New York: Springer; 2010.

[4] Shannon, C.E.. A mathematical theory of communication. Bell Syst Tech J 1948;5(1):3–55.

[5] Pincus, S.M.. Approximate entropy as a measure of system complexity. Proc Nat Acad Sci 1991;88(6):2297–2301.

[6] Richman, J.S., Moorman, J.R.. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000;278(6):H2039–H2049.

[7] Renyi, A.. On measures of entropy and information. In: Fourth Berkeley Symposium on Mathematical Statistics and Probability. 1961, p. 547–561.

[8] Lake, D.E.. Renyi entropy measures of heart rate gaussianity. IEEE Trans Biomed Eng 2006;53(1):21–27.

[9] Lake, D.E., Moorman, J.R.. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. Am J Physiol Heart Circ Physiol 2011;300(1):H319–H325.

[10] Bishop, C.M.. Neural networks for pattern recognition. Oxford: Oxford University Press; 1995.

[11] Woodcock, D., Nabney, I.. A new measure based on the renyi entropy rate using gaussian kernels. Tech. Rep.; Technical report, Aston University, Birmingham (UK); 2006.

[12] Marcos, J.V., Hornero, R., Nabney, I.T., Álvarez, D., Del Campo, F.. Analysis of nocturnal oxygen saturation recordings using kernel entropy to assist in sleep apnea-hypopnea diagnosis. In: Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2011, p. 1745–1748.

[13] Harte, D., Vere-Jones, D.. The entropy score and its uses in earthquake forecasting. Pure Appl Geophys 2005;162(6-7):1229–1253.

[14] Liu, L.Z., Qian, X.Y., Lu, H.Y.. Cross-sample entropy of foreign exchange time series. Physica A 2010;389(21):4785–4792.

[15] Yan, R., Gao, R.X.. Approximate entropy as a diagnostic tool for machine health monitoring. Mechanical Systems and Signal Processing 2007;21(2):824–839.

[16] Porta, A., Gnecchi-Ruscone, T., Tobaldini, E., Guzzetti, S., Furlan, R., Montano, N.. Progressive decrease of heart period variability

entropy-based complexity during graded head-up tilt. Journal of applied physiology 2007;103(4):1143–1149.

[17] Kapidžić, A., Platiša, M.M., Bojić, T., Kalauzi, A.. Nonlinear properties of cardiac rhythm and respiratory signal under paced breathing in young and middle-aged healthy subjects. Medical engineering & physics 2014;36(12):1577–1584.

[18] Qureshi, A., Ballard, R.D., Nelson, H.S.. Obstructive sleep apnea. J Allergy Clin Immunol 2003;112(4):643–651.

[19] Netzer, N., Eliasson, A.H., Netzer, C., Kristo, D.A.. Overnight pulse oximetry for sleep-disordered breathing in adultsa review. Chest 2001;120(2):625–633.

[20] Chiner, E., Signes-Costa, J., Arriero, J.M., Marco, J., Fuentes, I., Sergado, A.. Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies? Thorax 1999;54(11):968–971.

[21] Poupard, L., Philippe, C., Goldman, M.D., Sartène, R., Mathieu, M.. Novel mathematical processing method of nocturnal oximetry for screening patients with suspected sleep apnoea syndrome. Sleep and Breathing 2012;16(2):419–425.

[22] Marcos, J.V., Hornero, R., Alvarez, D., Aboy, M., Del Campo, F.. Automated prediction of the apnea-hypopnea index from nocturnal oximetry recordings. Biomedical Engineering, IEEE Transactions on 2012;59(1):141–149.

[23] Álvarez, D., Hornero, R., Marcos, J.V., del Campo, F.. Feature selection from nocturnal oximetry using genetic algorithms to assist in obstructive sleep apnoea diagnosis. Medical engineering & physics 2012;34(8):1049–1057.

[24] Morillo, D.S., Gross, N.. Probabilistic neural network approach for the detection of sahs from overnight pulse oximetry. Medical & biological engineering & computing 2013;51(3):305–315.

[25] Marcos, J.V., Hornero, R., Álvarez, D., del Campo, F., Zamarrón, C.. Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry. Medical engineering & physics 2009;31(8):971–978.

[26] Alvarez, D., Hornero, R., Marcos, J.V., del Campo, F.. Multivariate analysis of blood oxygen saturation recordings in obstructive sleep apnea diagnosis. IEEE Trans Biomed Eng 2010;57(12):2816–2824.

[27] Hornero, R., Álvarez, D., Abásolo, D., del Campo, F., Zamarrón, C.. Utility of approximate entropy from overnight pulse oximetry data in the diagnosis of the obstructive sleep apnea syndrome. IEEE Trans Biomed Eng 2007;54(1):107–113.

[28] Del Campo, F., Hornero, R., Zamarrón, C., Abasolo, D.E., Álvarez, D.. Oxygen saturation regularity analysis in the diagnosis of obstructive sleep apnea. Artif Intell Med 2006;37(2):111–118.

[29] Rechtschaffen, A., Kales, A.. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Los

Angeles: Brain Information Services, Brain Research Institute, University of California; 1968.

[30] Berry, R.B., Budhiraja, R., Gottlieb, D.J., Gozal, D., Iber, C., Kapur, V.K., et al. Rules for scoring respiratory events in sleep: update of the 2007 aasm manual for the scoring of sleep and associated events. J Clin Sleep Med 2012;8(5):597–619.

[31] Bloch, K.E.. Getting the most out of nocturnal pulse oximetry. CHEST Journal 2003;124(5):1628–1630.

[32] Magalang, U.J., Dmochowski, J., Veeramachaneni, S., Draw, A., Mador, M.J., El-Solh, A., et al. Prediction of the apnea-hypopnea index from overnight pulse oximetry. CHEST Journal 2003;124(5):1694–1701.

[33] Young, T.. Rationale, design, and findings from the wisconsin sleep cohort study: toward understanding the total societal burden of sleep-disordered breathing. Sleep medicine clinics 2009;4(1):37–46.

[34] Gabbay, I.E., Lavie, P.. Age-and gender-related characteristics of obstructive sleep apnea. Sleep and Breathing 2012;16(2):453–460.

[35] Lake, D.E.. Nonparametric entropy estimation using kernel densities. Meth Enzymol 2009;467:531–546.

[36] Zhang, X., King, M.L., Hyndman, R.J.. A bayesian approach to bandwidth selection for multivariate kernel density estimation. Comput Stat Data An 2006;50(11):3009–3031.

[37] Hastings, W.K.. Monte carlo sampling methods using markov chains and their applications. Biometrika 1970;57(1):97–109.

[38] Nabney, I.. NETLAB: algorithms for pattern recognition. London: Springer; 2002.

[39] Hanley, J.A., McNeil, B.J.. The meaning and use of the area under a receiver operating characteristic (roc) curve. Radiology 1982;143(1):29–36.

[40] Chen, Y., Chen, W.. Long-term tracking of a patients health condition based on pulse rate dynamics during sleep. Ann Biomed Eng 2011;39(12):2922–2934.

[41] Pincus, S.M., Goldberger, A.L.. Physiological time-series analysis: what does regularity quantify? Am J Physiol 1994;266:H1643–H1643.

[42] Lilliefors, H.W.. On the kolmogorov-smirnov test for normality with mean and variance unknown. Journal of the American Statistical Association 1967;62(318):399–402.

[43] Stoline, M.R.. The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way anova designs. The American Statistician 1981;35(3):134–141.

Figure captions

Figure 1: Two examples of the SaO_2 signals analysed in our study. (a) SaO_2 signal from a subject with AHI = 0.5 h^{-1}; (b) SaO_2 signal from a subject with AHI = 32.1 h^{-1}; (c) detailed view of the signal corresponding to the subject with AHI = 0.5 h^{-1}; (d) detailed view of the signal corresponding to the subject with AHI = 32.1 h^{-1}.

Figure 2: Prior probability distribution $f(\sigma)$ of the bandwidth parameter σ for $\lambda = 0.1$, 1, 5, 10.

Figure 3: Analysis of the results on two-class classification of subjects. ROC curves computed from the measurements of (a) AEn, (c) SEn and (e) KEn. Boxplots in SAHS-negative and SAHS-positive groups for (b) AEn, (d) SEn and (f) KEn.

Figure 4: Utility of the evaluated entropy metrics to rank SAHS severity. AHI versus SaO_2 regularity quantified by (a) AEn, (c) SEn and (e) KEn. Boxplots in the four SAHS severity groups for (b) AEn, (d) SEn and (f) KEn.

Table captions

Training Set			
	All	SAHS-positive	SAHS-negative
Subjects	96	64	32
Age (years)	52.35 ± 13.76	54.88 ± 14.53	47.31 ± 10.59
Males (%)	77.08	84.38	62.50
BMI (kg/m^2)	29.83 ± 4.17	30.61 ± 3.86	28.27 ± 4.38
Recording Time (h)	7.25 ± 0.33	7.25 ± 0.35	7.25 ± 0.29
AHI (h^{-1})	24.75 ± 25.19	35.01 ± 25.16	4.23 ± 2.22
Test Set			
	All	SAHS-positive	SAHS-negative
Subjects	144	96	48
Age (years)	52.19 ± 13.73	54.71 ± 13.35	47.17 ± 13.20
Males (%)	77.78	83.33	66.67
BMI (kg/m^2)	29.83 ± 4.53	30.98 ± 4.65	27.54 ± 3.26
Recording Time (h)	7.24 ± 0.66	7.22 ± 0.78	7.30 ± 0.33
AHI (h^{-1})	26.39 ± 26.74	37.71 ± 26.17	3.75 ± 2.51

Table 1: Clinical and demographic features for subjects in training and test sets. Data are presented as mean $\pm$ standard deviation. BMI: body mass index; AHI: apnoea-hypopnoea index.

	AUC	ρ
AEn ($m = 1$, $r = 0.1$ SD)	0.68	0.32
AEn ($m = 1$, $r = 0.15$ SD)	0.57	0.19
AEn ($m = 1$, $r = 0.2$ SD)	0.50	0.18
AEn ($m = 1$, $r = 0.25$ SD)	0.48	0.15
AEn ($m = 2$, $r = 0.1$ SD)	0.56	0.18
AEn ($m = 2$, $r = 0.15$ SD)	0.55	0.13
AEn ($m = 2$, $r = 0.2$ SD)	0.51	0.13
AEn ($m = 2$, $r = 0.25$ SD)	0.52	0.14
SEn ($m = 1$, $r = 0.1$ SD)	0.74	0.40
SEn ($m = 1$, $r = 0.15$ SD)	0.68	0.31
SEn ($m = 1$, $r = 0.2$ SD)	0.63	0.33
SEn ($m = 1$, $r = 0.25$ SD)	0.63	0.34
SEn ($m = 2$, $r = 0.1$ SD)	0.67	0.22
SEn ($m = 2$, $r = 0.15$ SD)	0.61	0.13
SEn ($m = 2$, $r = 0.2$ SD)	0.58	0.17
SEn ($m = 2$, $r = 0.25$ SD)	0.60	0.19
KEn ($m = 1$)	0.86	0.82
KEn ($m = 2$)	0.86	0.83

Table 2: Selection of the input parameters using training data. Results achieved on the training set by the evaluated configurations of AEn, SEn and KEn. AUC: area under the ROC curve; ρ: Pearson's correlation coefficient; m: vector length parameter; r: tolerance; SD: standard deviation.

	AUC	ρ
AEn ($m = 1$, $r = 0.1$ SD)	0.67	0.34
SEn ($m = 1$, $r = 0.1$ SD)	0.74	0.45
KEn ($m = 2$)	0.91	0.87

Table 3: Asessment of the entropy metrics on test samples. Results achieved on the test set by the selected configurations of AEn, SEn and KEn. AUC: area under the ROC curve; ρ: Pearson's correlation coefficient; m: vector length parameter; r: tolerance; SD: standard deviation.

	Normal vs Mild	Mild vs Mod.	Mod. vs Sev.	Normal vs Mod.	Normal vs Sev.	Mild vs Sev.
AEn	> 0.001	> 0.5	> 0.05	> 0.005	< 0.0001	> 0.05
SEn	> 0.001	> 0.5	> 0.01	< 0.0001	< 0.0001	> 0.001
KEn	< 0.0001	> 0.001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

Table 4: One-way ANOVA test to evaluate the difference between the means of each entropy metric in the four SAHS severity groups: normal, mild, moderate and severe.

	AUC	ρ
f_{avg}	0.85	0.59
f_{sd}	0.96	0.76
f_{cv}	0.96	0.71
f_{iqr}	0.96	0.74
f_{sd1}	0.97	0.77
f_{sd2}	0.96	0.76

Table 5: Diagnostic results obtained for conventional statistical features of oximetry samples on the test set. AUC: area under the ROC curve; ρ: Pearson's correlation coefficient; f_{avg}: mean; f_{sd}: standard deviation; f_{cv}: coefficient of variation; f_{iqr}: interquartile range; f_{sd1}: first-order dispersion index from Poincare plot; f_{sd2}: second-order dispersion index from Poincare plot.

Figure 1
Click here to download high resolution image

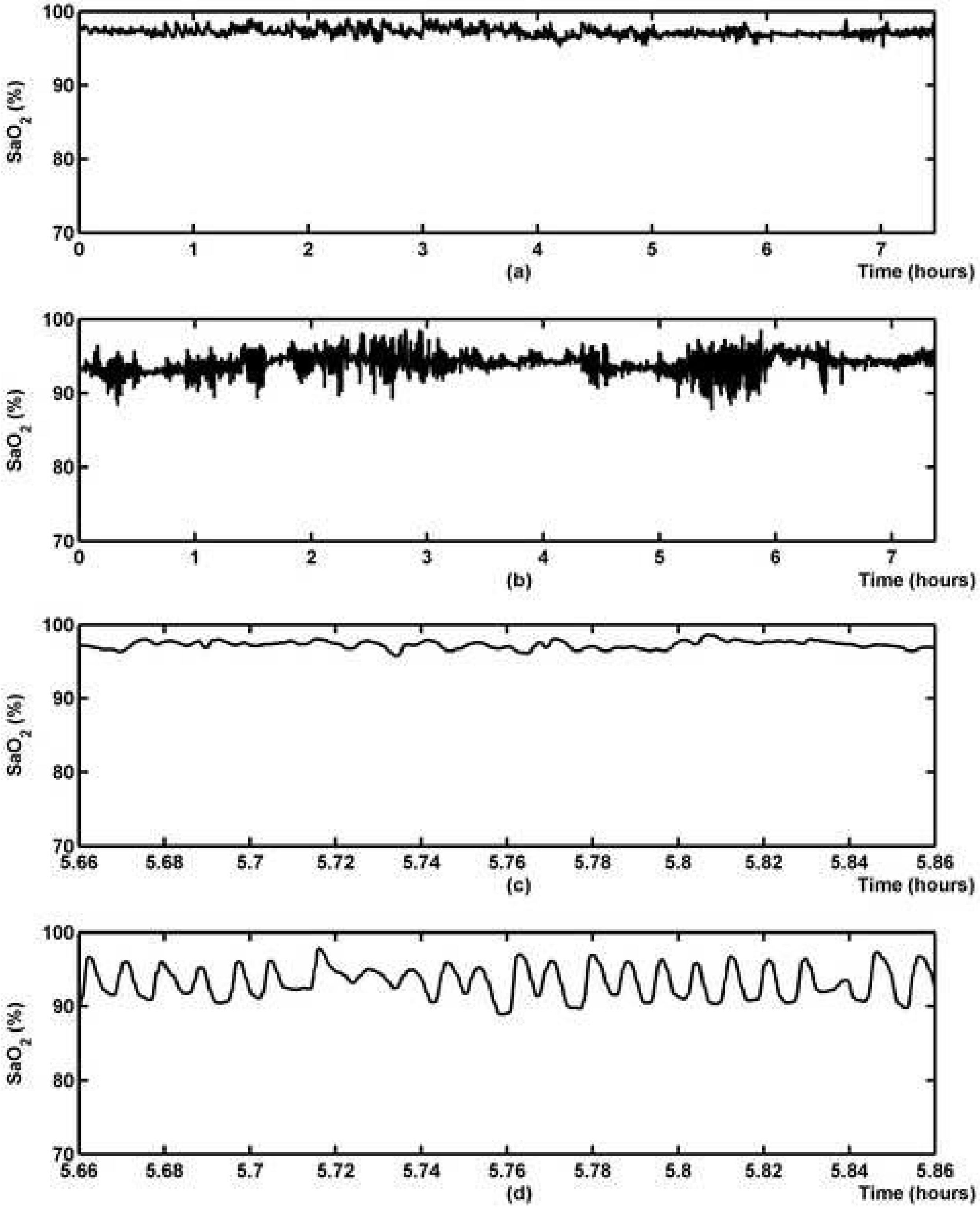

Figure 2
Click here to download high resolution image

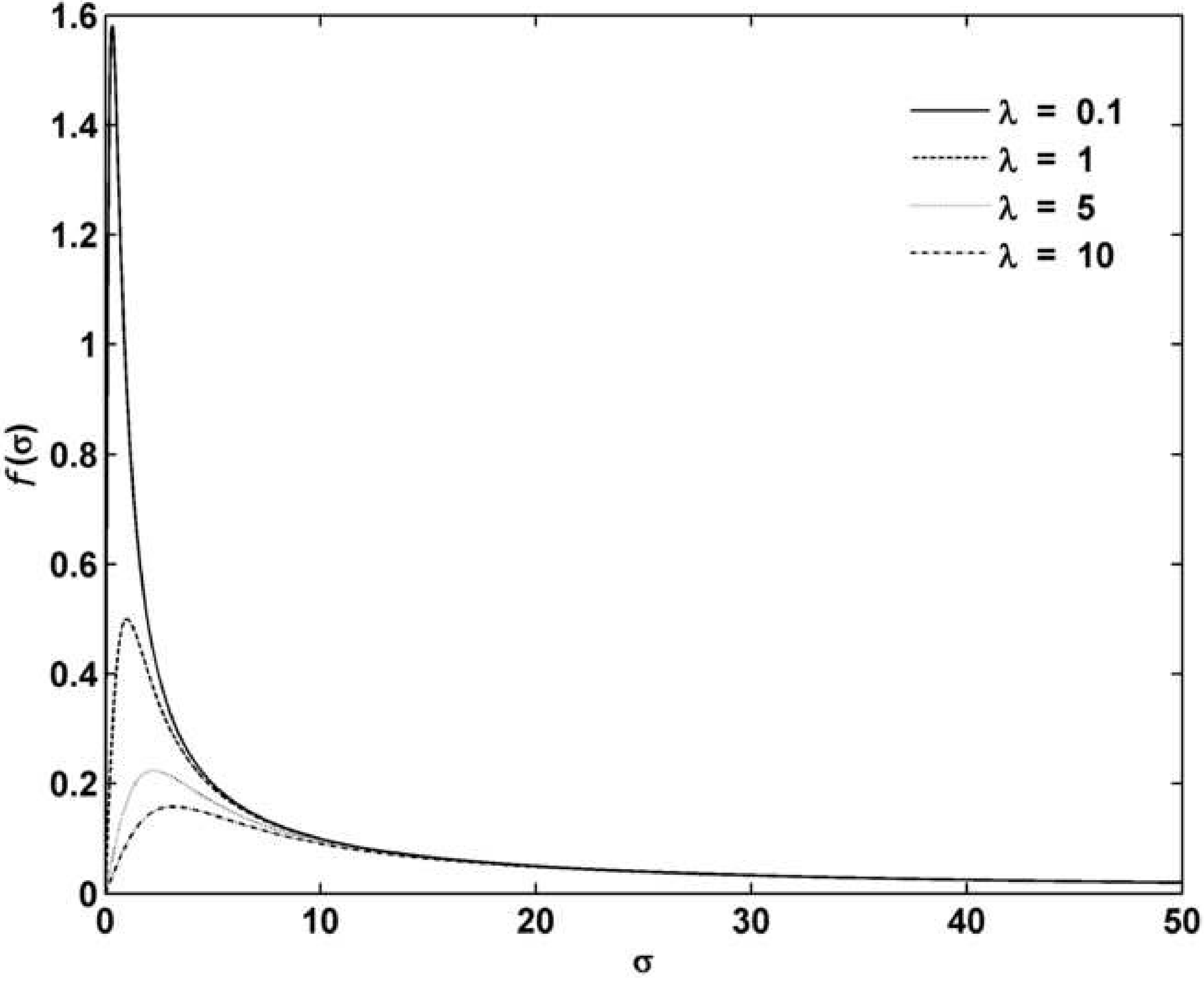

Figure 3
Click here to download high resolution image

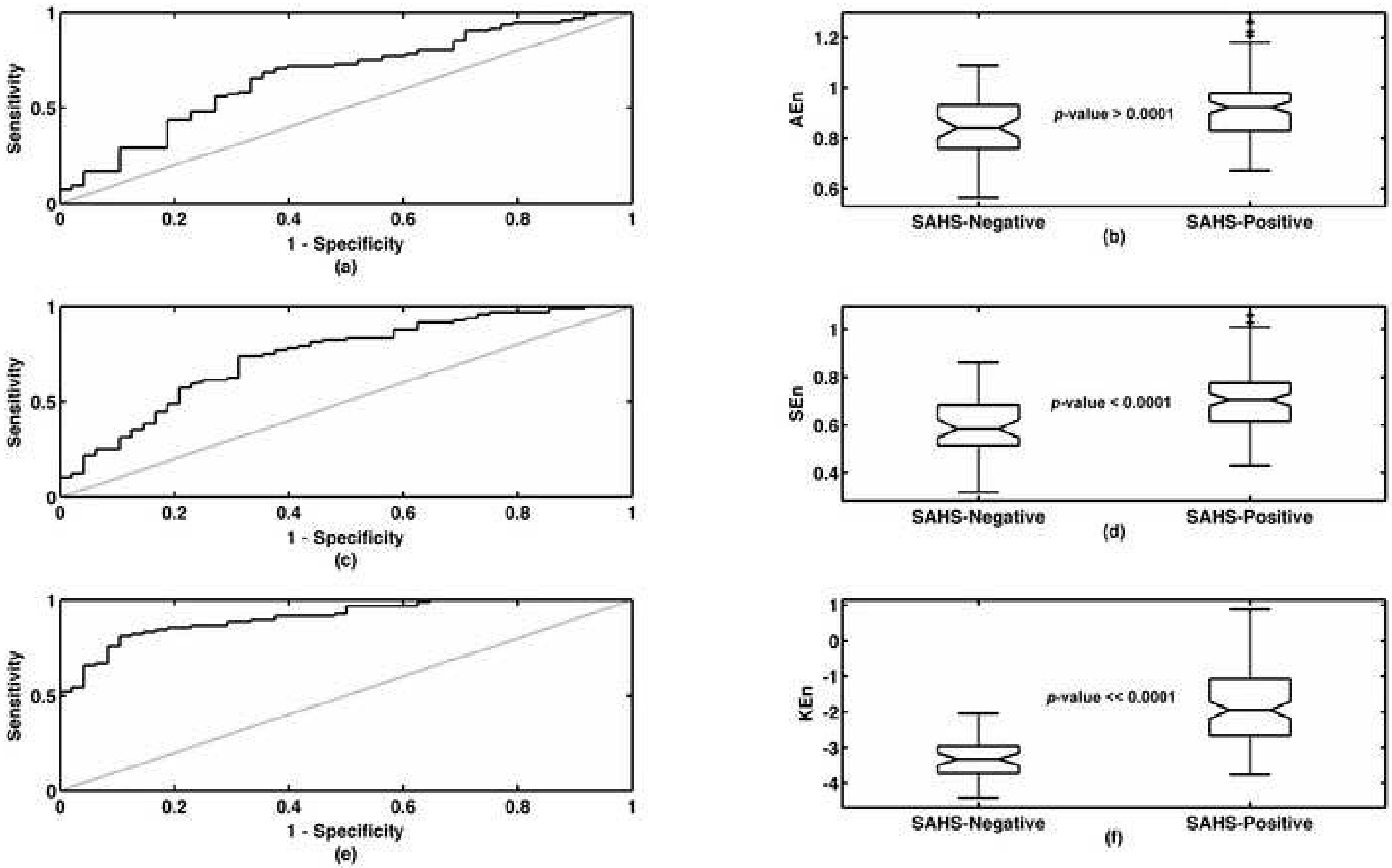

Figure 4
Click here to download high resolution image