Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis

Verónica Barroso-García[a,*], Gonzalo C. Gutiérrez-Tobal[a], Leila Kheirandish-Gozal[b], Daniel Álvarez[a,c], Fernando Vaquerizo-Villar[a], Pablo Núñez[a], Félix del Campo[a,c], David Gozal[b], and Roberto Hornero[a,d]

[a] Biomedical Engineering Group, Universidad de Valladolid, Valladolid, Spain.

[b] Department of Child Health, The University of Missouri School of Medicine, Columbia, Missouri, USA.

[c] Sleep-Ventilation Unit, Pneumology Service, Río Hortega University Hospital, Valladolid, Spain.

[d] IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, Spain.

* Corresponding author: **Verónica Barroso-García**

Biomedical Engineering Group, E.T.S. Ingenieros de Telecomunicación, Universidad de Valladolid, Campus Miguel Delibes, Paseo Belén 15, 47011 – Valladolid, Spain.

Tel. +34 983 423000 ext. 4713

E-mail address: veronica.barroso@gib.tel.uva.es

URL: www.gib.tel.uva.es

Abstract- *Background and objective:* In-laboratory overnight polysomnography (PSG) is the gold standard method to diagnose the sleep apnoea-hypopnoea syndrome (SAHS). PSG is a complex, expensive, labour-intensive and time-consuming test. Consequently, simplified diagnostic methods are desirable. We propose the analysis of the airflow (AF) signal by means of recurrence plots (RP) features. The main goal of our study was to evaluate the utility of the information from RPs of the AF signals to detect paediatric SAHS at different levels of severity. In addition, we also evaluated the complementarity with the 3% oxygen desaturation index (ODI_3). *Methods:* 946 AF and blood oxygen saturation (SpO_2) recordings from children ages 0-13 years were used. The population under study was randomly split into training (60%) and test (40%) sets. RP was computed and 9 RP features were extracted from each AF recording. ODI_3 was also calculated from each SpO_2 recording. A feature selection stage was conducted in the training group by means of the fast correlation-based filter (FCBF) methodology to obtain a relevant and non-redundant optimum feature subset. A multi-layer perceptron neural network with Bayesian approach (BY-MLP), trained with these optimum features, was used to estimate the apnoea–hypopnoea index (AHI). *Results:* 8 of the RP features showed statistically significant differences (p-value <0.01) among the SAHS severity groups. FCBF selected the maximum length of the diagonal lines from RP, as well as the ODI_3. Using these optimum features, the BY-MLP model achieved 83.2%, 78.5%, and 91.0% accuracy in the test group for the AHI thresholds 1, 5, and 10 events/h, respectively. Moreover, this model reached a negative likelihood ratio of 0.1 for 1 e/h and a positive likelihood ratio of 13.7 for 10 e/h. *Conclusions:* RP analysis enables extraction of useful SAHS-related information from overnight AF paediatric recordings. Moreover, it provides complementary information to the widely-used clinical variable ODI_3. Thus, RP applied

to AF signals can be used along with ODI_3 to help in paediatric SAHS diagnosis, particularly to either confirm the absence of SAHS or the presence of severe SAHS.

Keywords- Airflow (AF); Children; Recurrence Plots (RP); Sleep Apnoea-Hypopnoea Syndrome (SAHS).

1. INTRODUCTION

Childhood Sleep Apnoea-Hypopnoea Syndrome (SAHS) is a breathing disorder characterized by recurrent airflow cessation (apnoeas) and/or significant airflow-reduction (hypopnoeas) episodes during sleep [1,2]. In spite of its high prevalence (1-4% of all children) [3], paediatric SAHS is an underdiagnosed disease whose adverse consequences include cognitive, behavioural, metabolic, and cardiovascular functions [3]. Early detection and treatment of the affected children is therefore of paramount importance.

In-laboratory overnight polysomnography (PSG) is the gold standard for paediatric SAHS diagnosis [4]. Paediatric PSG is performed in a sleep laboratory, suitable for children, and consists in recording a wide range of biomedical signals. Qualified medical personnel identifies and quantifies the severity of SAHS by means of these recordings. However, PSG is a complex, labour-intensive, and expensive test [5]. It is also time-consuming and both the facilities and the expertise needed to carry it out are not always available, which entails prolonged access delays [5]. Moreover, the multifarious sensors placed on child's body make PSG a particularly ill-at-ease test for some, leading to undesirable percentages of failed testing opportunities [6].

In order to deal with the inherent limitations of PSG, alternative simplified methods have been explored. The analysis of a reduced signal set, such as electrocardiogram (ECG), photoplethysmography (PPG), blood oxygen saturation (SpO_2), and airflow (AF), is a commonly used approach [7–14]. Several studies have analysed these signals through temporal and/or spectral analysis techniques, to assist in paediatric SAHS diagnosis [8–14].

In this study, we propose and have analysed the AF signal to help in paediatric SAHS diagnosis. In addition, 3% oxygen desaturation index (ODI_3), a conventional

oximetry index commonly used in the SAHS context [11,15,16], has been also obtained. Despite its widespread use, the extant literature indicates that ODI_3, which is obtained from SpO_2, underestimates both SAHS presence and severity [15,17]. Therefore, only using this index is insufficient to accurately simplify diagnosis. In this regard, the analysis of AF is a natural way of simplifying PSG, since the apnoeic events modify the amplitude of this signal [18]. Moreover, recent studies have already shown the usefulness of AF in diagnosing paediatric SAHS [13,14].

Based on the aforementioned considerations, our proposal is based on the use of non-linear recurrence plots (RPs) analysis to obtain useful features from AF [19]. The respiratory system is dynamic, non-linear, and non-stationary, which may lead to the presence of recurrences within a given state space [20,21]. Recurrence is a property of dynamic systems which refers to a point or state that occurs repeatedly throughout a given time series [19,20]. Therefore, RP analysis allows visualization of the recurrences of the phase-space states of a signal [19]. The occurrence of apnoeic events produces changes in the dynamics of the system, altering the amount and distribution of the recurrences in the RP [20]. Hence, RP analysis can provide information about these changes, even though the signals are non-stationary [20]. However, we are unaware of any studies characterizing paediatric SAHS by means of RP obtained from PSG signals. The properties of RP may further help to characterize the presence and severity of paediatric SAHS in AF recordings while overcoming some limitations of traditional Fourier-based analyses. The previous success of RP in characterizing other biomedical signals also supports its application to AF in SAHS context. It has been successfully used to automatically identify epileptic EEG signals [22], to monitor anaesthesia by EEG recordings [23], as well as to improve the diagnostic ability of ECG to detect SAHS in adults [20,24].

Hence, we hypothesized that the analysis of the RP applied to AF signals may be useful to obtain paediatric SAHS-related information. Accordingly, our main objective is to evaluate the utility of this information to detect SAHS at different severity degrees. Furthermore, its complementarity with ODI_3 is also addressed in our study.

2. SUBJECTS AND SIGNALS UNDER STUDY

In this study, AF and SpO_2 signals were recorded from 946 children referred to the Paediatric Sleep Unit at the Comer Children's Hospital of the University of Chicago, due to clinical SAHS suspicion. The Ethics Committee of the Hospital approved the study protocol. An informed consent was obtained from the legal caretakers of all children.

The subjects were diagnosed by specialised physicians according to the rules of the American Academy of Sleep Medicine (AASM) [18]. The apnoea-hypopnoea index (AHI), computed as the number of apnoea and hypopnoea events per hour (e/h) of sleep, was used to establish SAHS and its severity [18,25]. Common AHI thresholds of 1, 5, and 10 e/h were used to classify paediatric subjects into four SAHS-severity degrees [10,14,26–29]: no-SAHS (AHI < 1 e/h), mild (1 e/h ≤ AHI < 5 e/h), moderate (5 e/h ≤ AHI < 10 e/h), and severe SAHS (AHI ≥ 10 e/h). Hence, our database was divided according to these thresholds.

The subjects were randomly split into a training set (60%) and a test set (40%). Table 1 shows the clinical and demographic data of the population under study. No statistically significant differences (p-value >0.01) were found in age, gender, body mass index (BMI), and AHI between the training and test sets, after applying the Chi-square and the Mann-Whitney tests.

Table 1. Demographic and clinical data from the paediatric subjects under study. Data are presented as median [interquartile range] or n (%); BMI: body mass index; AHI: apnoea–hypopnoea index.

	All	Training group	Test group	*p*-value
Subjects (n)	946	570	376	-
Age (years)	6 [6]	6 [5]	6 [6]	0.9063
Males (n)	584 (61.73%)	339 (59.47%)	245 (65.16%)	0.0875
BMI (kg/m^2)	17.92 [6.17]	17.72 [6.74]	18.07 [6.01]	0.9610
AHI (e/h)	3.82 [7.80]	4.17 [8.34]	3.33 [6.44]	0.0340
AHI ≥ 1 (e/h)	783 (82.77%)	479 (84.04%)	304 (80.85%)	0.2185
AHI ≥ 5 (e/h)	397 (41.97%)	256 (44.91%)	141 (37.5%)	0.0263
AHI ≥ 10 (e/h)	225 (23.78%)	145 (25.44%)	80 (21.28%)	0.1601

PSG was conducted using a digital polysomnography system (Polysmith, Nihon Kohden America Inc., Irvine, CA, USA). Single-channel AF and SpO_2 signals were acquired during the PSG. Recordings lasting less than 3 hours were discarded [11]. The AF recordings, obtained with a thermistor, were sampled at 100 Hz [18]. These were normalised according to Varady et al. [30], to minimize possible differences in AF caused by age. AF artefacts were removed by comparing statistical measures of 30-second epochs [31]. The SpO_2 recordings, sampled at 25 Hz, were used to obtain ODI_3. Their artefacts were removed by discarding SpO_2 values <50% and changes with a slope ≥4%/second [16]. Figure 1(a) shows an example of the AF signal and figure 1(b) shows the corresponding SpO_2 signal. We can see 7 apnoea events (absence of AF, i.e., near cero amplitude) in 0.4, 0.9, 1.7, 2.6, 4.9, 5.4, and 6.3 minutes, and the corresponding desaturations.

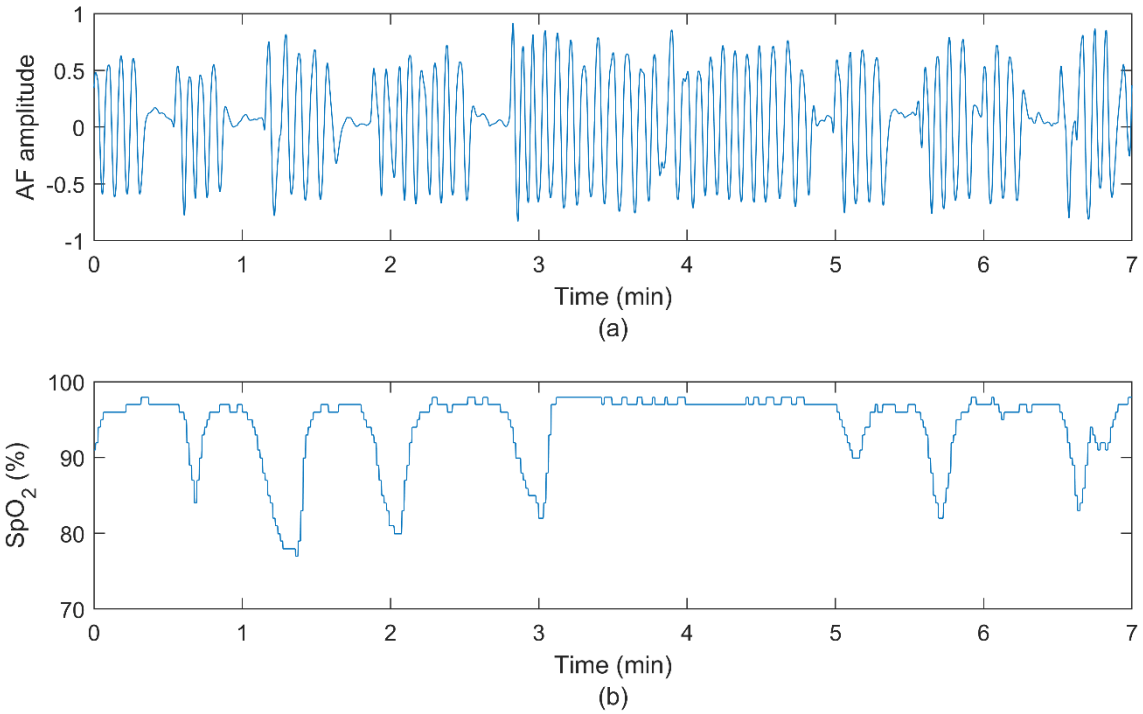

Figure 1. Normal breathing pattern followed by apnoeic events in (a) airflow signal (AF) and (b) corresponding blood oxygen saturation signal (SpO_2).

3. METHODS

ODI_3 was obtained from SpO_2 recordings according to Taha et al. [32]. Therefore, oxygen desaturation events were scored as a decrease ≥3% of SpO_2 at a rate of 0.1-4%/second, during 10-60 seconds. The number of desaturation events was divided by the number of recording hours to obtain *ODI_3*.

Afterwards, a four-stage study was carried out: (i) computation of RP from AF, (ii) RP feature extraction, (iii) feature selection through the fast correlation-based filter (FCBF) method, and (iv) AHI estimation using a multi-layer perceptron neural network with Bayesian training approach (BY-MLP). Figure 2 shows the block diagram of the method proposed in our study.

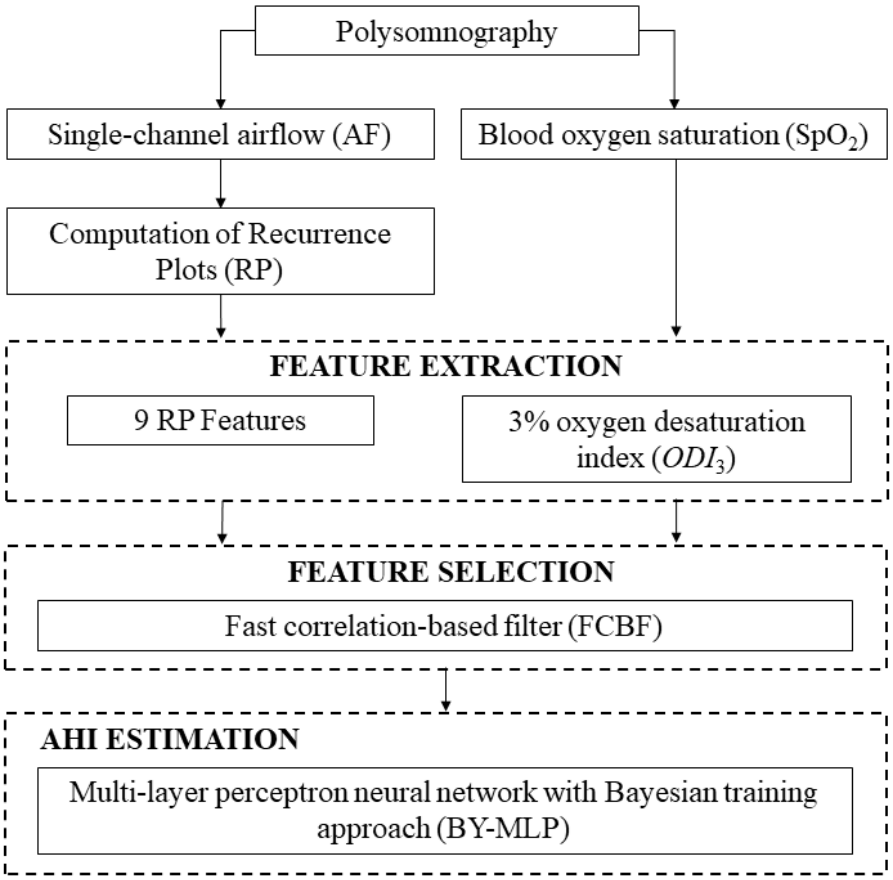

Figure 2. Block diagram of the proposed method. AHI: apnoea-hypopnoea index.

3.1. Recurrence plot computation

In the present context, recurrences are points or states that occur repeatedly. RPs are analytical tools for visualizing such recurrences and discovering hidden periodicities of dynamic systems, i.e., systems that evolve over time, like a physiological time series [19]. Since a dynamic system is defined by vectors representing trajectories in the *m*-dimensional phase-space, a RP is the graphic representation of the binary and symmetric recurrence matrix $R_{i,j}$, whose values are 1 if two trajectories are roughly equal (there is a recurrence) and 0 otherwise [19,20].

Dealing with discrete measurements, the first step towards obtaining an RP is to reconstruct the phase-space using the well-known Taken's time-delay method [19,20,33]:

$$\vec{x}_i = [u(i), u(i+\tau), ..., u(i+(m-1)\cdot\tau)], \quad (1)$$

where $u(i)$ is the value of the time series at time i, m is the embedding dimension, and τ is the time delay. The embedded dimension (m) and the time delay (τ) are parameters to be optimized in each context. The time delay has to be adjusted to not set up auto-correlated state vectors [20,24], with the auto-mutual information (AMI) function being commonly used for this purpose [19,20,34]. The embedding dimension must also be carefully selected, since an undue increase of m removes the isolated recurrences from the RP and increases the occurrence of spurious diagonal structures [19]. In this regard, the false nearest neighbour's (FNN) method has been successfully used to optimize m [19,34,35].

Once the phase-space has been reconstructed, the distance matrix, $D_{i,j}$, is calculated by the commonly used Euclidean distance norm [19,20]:

$$D_{i,j} = \left\| \vec{x}_i - \vec{x}_j \right\|, \tag{2}$$

where $i, j = 1, \ldots, L-(m-1)\cdot\tau$ and L is the time series length.

Afterwards, $R_{i,j}$ is calculated through the Heaviside function, being $R_{i,j}=1$ (i.e. there is a recurrence) if the distance is less than a certain threshold, ε, and $R_{i,j}=0$ otherwise [19]:

$$R_{i,j} = \begin{cases} 1 : D_{i,j} \leq \varepsilon \\ 0 : D_{i,j} > \varepsilon \end{cases}. \tag{3}$$

A proper selection of ε is essential too. If ε is too small, no recurrences will be shown in the RP and no information could be derived about the dynamic of the system [19]. By contrast, if ε is too large, almost all the points will be considered recurrences, leading to the appearance of many artefacts [19]. The fixed distance method, which fixes the threshold according to the standard deviation (σ) of each time series, has been widely used to select ε [20,34].

Finally, the RP is obtained by plotting the recurrence matrix.

3.2. Feature extraction: Recurrence plot analysis

A RP always has a main diagonal, i.e., line of identity (LOI), with respect to which the RP is symmetric. As shown in Figure 3, isolated recurrences, vertical (i.e. laminar segments), and diagonal (i.e. deterministic segments) structures can be found in a typical RP. There are several features that quantify these RP structures, known as recurrence quantification analysis (RQA) [19,20]:

- Features based on the recurrence density of the RP:
  - Recurrence rate (*REC*): it quantifies the percentage of recurrences of RP, including isolated recurrences and recurrences that form deterministic and laminar segments (see Figure 3), providing information about the variability of time series [19,36]. The more recurrences there are, the less variable the time series is [20]:

$$REC = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j} \,, \tag{4}$$

    where N is the number of rows in $R_{i,j}$.

- Features based on diagonal structures of the RP:
  - Determinism (*DET*): *DET* is the proportion of recurrences forming diagonal lines, i.e. forming deterministic segments (see Figure 3) [24,36]. This feature measures the determinism (predictability) of time series [19,20]. Thereby, longer diagonal lines and fewer isolated recurrences (high *DET*) imply more predictability of the time series [19]:

$$DET = \frac{\sum_{l=l_{\min}}^{N} l \cdot p(l)}{\sum_{l=1}^{N} l \cdot p(l)}, \tag{5}$$

    where l_{min} is the minimum length to consider a diagonal line and $p(l)$ is the histogram of diagonal lines of length l. In our work, we define the value l_{min} as 2,

which is the most commonly used [20,34].

- Average diagonal line length (*LEN*): *LEN* provides information about the average time that allows to predict the future trajectory of a dynamic system from the knowledge of its initial state (i.e. prediction time of a time series) [19,20,22]:

$$LEN = \frac{\sum_{l=l_{\min}}^{N} l \cdot p(l)}{\sum_{l=l_{\min}}^{N} p(l)}. \quad (6)$$

  This feature can be observed in Figure 3.

- Maximum length of the diagonal lines (*Lmax*): it measures the exponential divergence of the phase-space trajectory [19,20]. The faster the trajectory segments diverges, the shorter the diagonal lines [19]:

$$L\max = \max(l_i, i = 1,..., N). \quad (7)$$

  This feature can be observed in Figure 3.

- Shannon's entropy of the length distribution of the diagonal lines (*ENTR*): *ENTR* measures the complexity of the RP from its diagonal structures, i.e. regarding its deterministic segments [21,22]. Lower *ENTR* values indicate that the RP is less complex in respect of its diagonal lines [19].

$$ENTR = -\sum_{l=l_{\min}}^{N} p(l) \cdot \log(p(l)). \quad (8)$$

- Trend (*TREND*): it is the distribution of recurrences with respect to the LOI [19,36]. *TREND* reflects the non-stationarity of a signal [21,36]. Therefore, recurrences homogenously distributed (|*TREND*| close to zero) indicate a high level of stationarity in the signal [21]:

$$TREND = \frac{\sum_{i=1}^{\tilde{N}} (i - (\tilde{N}/2)) \cdot (REC_i - \langle REC_i \rangle)}{\sum_{i=1}^{\tilde{N}} (i - (\tilde{N}/2))^2}, \quad (9)$$

where $\tilde{N}$ is a number smaller than N to exclude the edges [19,36], REC_i is the number of recurrences in the diagonal lines with distance i to the LOI, and $\langle REC_i \rangle$ is the average of REC_i. We define $\tilde{N}$ as N-2, which is a commonly used value [19,36]. This feature can be observed in Figure 3.

- Features based on vertical structures of RP:
  - Laminarity (*LAM*): *LAM* is the proportion of recurrences forming vertical lines, i.e. forming laminar segments (see Figure 3) [22,24]. It represents the occurrence of laminar states in the RP, measuring the probability that a state does not change with time [19,34]. Higher *LAM* values indicate that states do not change or change slowly, resulting in less complexity [20]:

$$LAM = \frac{\sum_{v=v_{\min}}^{N} v \cdot p(v)}{\sum_{v=1}^{N} v \cdot p(v)}, \quad (10)$$

where v_{min} is the minimum length to consider a vertical line and $p(v)$ is the histogram of vertical lines of length v. We define the value v_{min} as 2, which is the most commonly used value [19,20,34].

  - Average vertical line length (trapping time, *TT*): *TT* estimates the average time that a system remains in a particular state [19,22,24]. The lower its value, the more complex the system is, as it stays briefly in a similar state [20]:

$$TT = \frac{\sum_{v=v_{\min}}^{N} v \cdot p(v)}{\sum_{v=v_{\min}}^{N} p(v)}. \quad (11)$$

This feature can be observed in Figure 3.

- Maximum length of vertical lines (*Vmax*): it gives information about the duration of the laminar states and the complexity of the signal [19,20]. The higher *Vmax* values are, the less complexity in the time series [20]:

$$V\max = \max(v_i, i = 1,...,N)\,. \tag{12}$$

This feature can be observed in Figure 3.

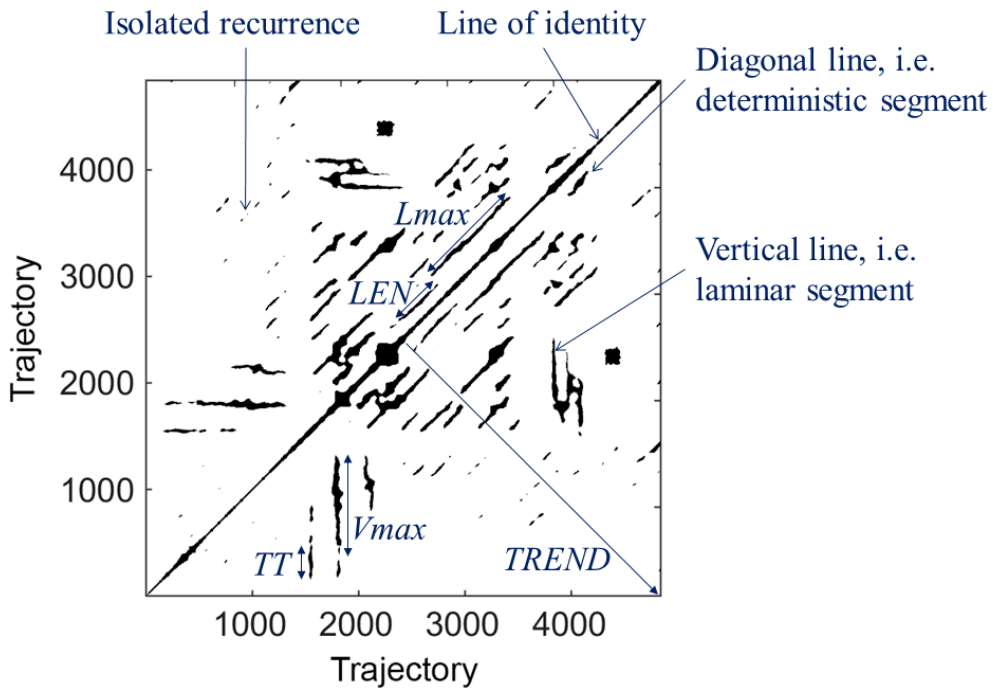

Figure 3. Example of typical structures and features of a RP.

3.3. Feature selection: Fast correlation-based filter

FCBF has proven its utility in a wide range of biomedical-related variable selection problems [37–39], including SAHS context [10,15,40]. Based on the symmetric uncertainty (*SU*), FCBF sorts the features in descending order of relevance (descending *SU*) and discards those of less relevance that are redundant [37]. Hence, an optimal subset of relevant and non-redundant features is obtained in order to maximise the diagnostic potential of the extracted information [37].

A bootstrapping methodology, with 1000 bootstrap replicates, was used during this stage to compose a more generalizable optimum feature subset [41]. FCBF was applied to each replicate and those features selected at least 50% of the times formed our optimum feature subset [10,15].

3.4. Apnoea–hypopnoea index estimation: Multi-layer perceptron neural network with Bayesian approach

MLP is an artificial neural network typically configured in 3 layers (input, hidden, and output layer) [10,15,42]. Each layer is composed of mathematical units called perceptrons, and each perceptron is connected to all perceptrons of the next layer [42]. There are as many inputs as variables in the feature space. The number of hidden layer perceptrons (N_H) is a hyper-parameter that must be optimized. In the present study, the output layer has one single perceptron to provide the estimated AHI, since this is a continuous variable. In order to optimize the weights and biases associated to the connections of the MLP, a Bayesian approach has been used in this study due to its previous success in the SAHS context [15,43].

3.5. Statistical Analysis

The RP features from AF did not pass the Lilliefors normality test. Therefore, the non-parametric Mann-Whitney and Kruskal-Wallis tests were used to evaluate statistically significant differences (p-value <0.01 after Bonferroni corrections for multiple comparisons) between the SAHS severity groups. Boxplots were used to show potential differences in RP features according to the degree of SAHS severity. Cohen's kappa (kappa) was used to measure the agreement between the actual diagnosis and the one derived from BY-MLP [44]. Regarding diagnostic performance of the BY-MLP, standard metrics were computed: sensitivity (Se: proportion of subjects with SAHS

rightly classified), specificity (Sp: proportion of subjects without the disease rightly classified), accuracy (Acc: proportion of overall subjects rightly classified), positive (PPV: proportion of positive test result which are true positives) and negative (NPV: proportion of negative test result which are true negatives) predictive values, positive (LR+: proportion of subjects with SAHS rightly classified with respect to the proportion of healthy subjects wrongly classified) and negative (LR-: proportion of subjects with SAHS wrongly classified with respect to the healthy subjects rightly classified) likelihood ratios [45,46]. Three optimum thresholds were used to evaluate the actual diagnostic ability of ODI_3 for AHI cut-off points 1, 5, and 10 e/h in the test group. Each optimum threshold was obtained as the closest point to [1,0] (100% Se and 100% Sp) of the receiver-operating characteristic (ROC) curve from the training group [46].

4. RESULTS

4.1. Training group

AMI was used in the training group to optimize τ in the phase-space reconstruction of AF signals. The τ was varied from 0.1 to 6 seconds. The optimum τ value (first relative minimum of AMI) was obtained for each subject of the training group. The median of these values determined the optimum τ for AF: τ=0.9 seconds. Regarding the optimum m value, FNN was used to obtain the minimum dimension where the number of false nearest neighbours is reduced to zero. The m was varied from 1 to 20 and FNN determined the adequate dimensional space for AF: m=3. Additionally, the threshold ε was varied from $(10^{-2}\cdot\sigma)$ to $(25\cdot10^{-2}\cdot\sigma)$ to optimise its value. The ε was fixed to $\varepsilon=10^{-2}\cdot\sigma$ since RP features obtained with this threshold had the highest average Spearman's correlation coefficient (RHO) with the AHI in the training group.

4.1.1. Exploratory Analysis

Since apnoeic events last at least 2 respiratory cycles [18], the AF signal of each subject was segmented into 30-second windows as a trade-off between ensuring that it is broad enough to include apnoeic events and these are a significant proportion of the information within each segment. RP of each subject was computed by averaging the recurrence points of RPs obtained for each window [19]. Moreover, Figure 4 illustrates the averaged RPs in each out of the four SAHS severity groups of the training group, with tones closer to red highlighting the presence of more recurrences. According to this figure, fading of recurrences towards the upper left and lower right corners is shown in RPs as darker blue tone regions. A slower fading and a higher occurrence of diagonal and vertical structures can be observed when AHI is higher. The combination of these structures formed clusters of recurrences close to the LOI, whose thickness is greater as SAHS severity increases.

Figure 5 shows the boxplots of the scaled RP features of the four SAHS severity groups in the training set. An increasing *LAM*, *LEN*, *Lmax*, *ENTR*, *REC*, *TT*, and *Vmax* tendency and a decreasing *TREND* tendency can be observed as AHI is higher. All RP features but *DET* showed significant differences (p-value <0.01) among the SAHS severity groups. The same features presented significant differences (p-value <0.01) between the severe SAHS group and the remaining ones. Additionally, *LAM* and *Lmax* showed significant differences between the no-SAHS and moderate SAHS groups, and between mild and moderate SAHS groups, respectively.

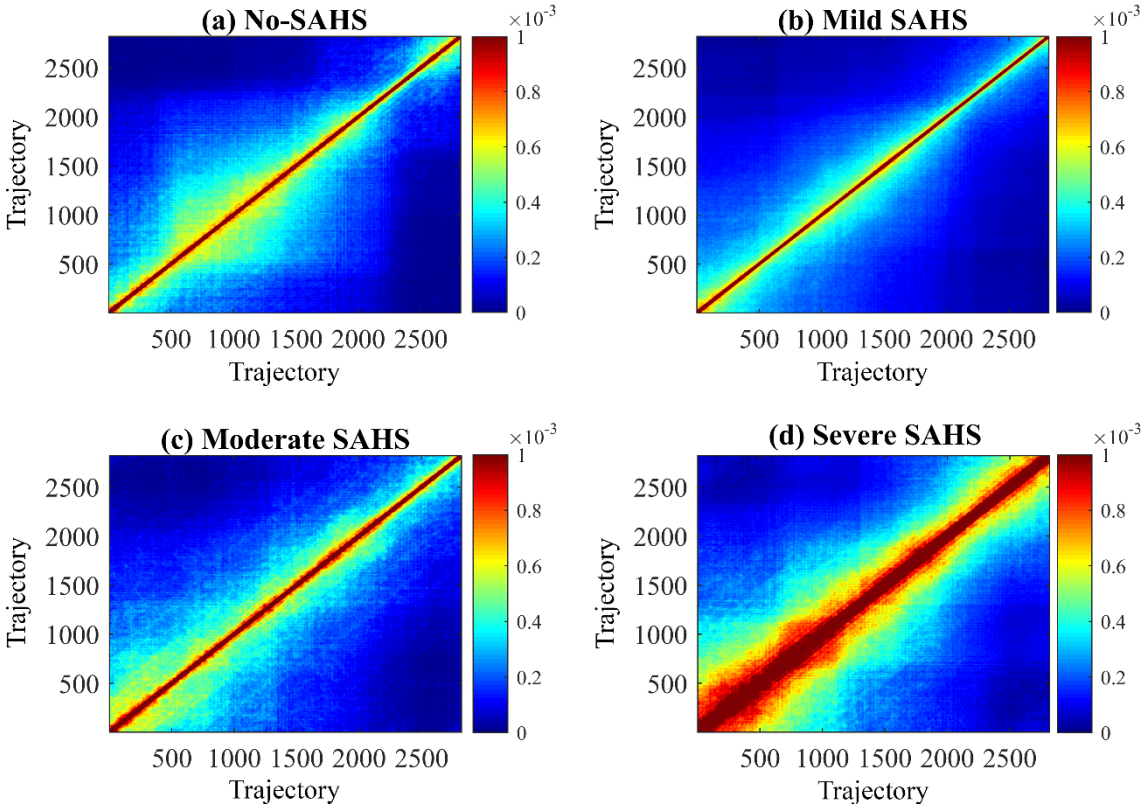

Figure 4. Averaged RP of the four SAHS severity groups in the training set: (a) no-SAHS, (b) mild, (c) moderate, and (d) severe SAHS. The trajectories of the m-dimensional phase-space are vectors that define the dynamic behaviour of AF. Having optimized the parameters dimension m = 3 and delay τ = 0.9 seconds, and being 100 Hz the sampling frequency and 30-seconds the window size, the equation 1 allow to define the phase-space trajectories of airflow recordings as $\vec{x}_1 = [u(1), u(91), u(181)]$ to $\vec{x}_{2820} = [u(2820), u(2910), u(3000)]$.

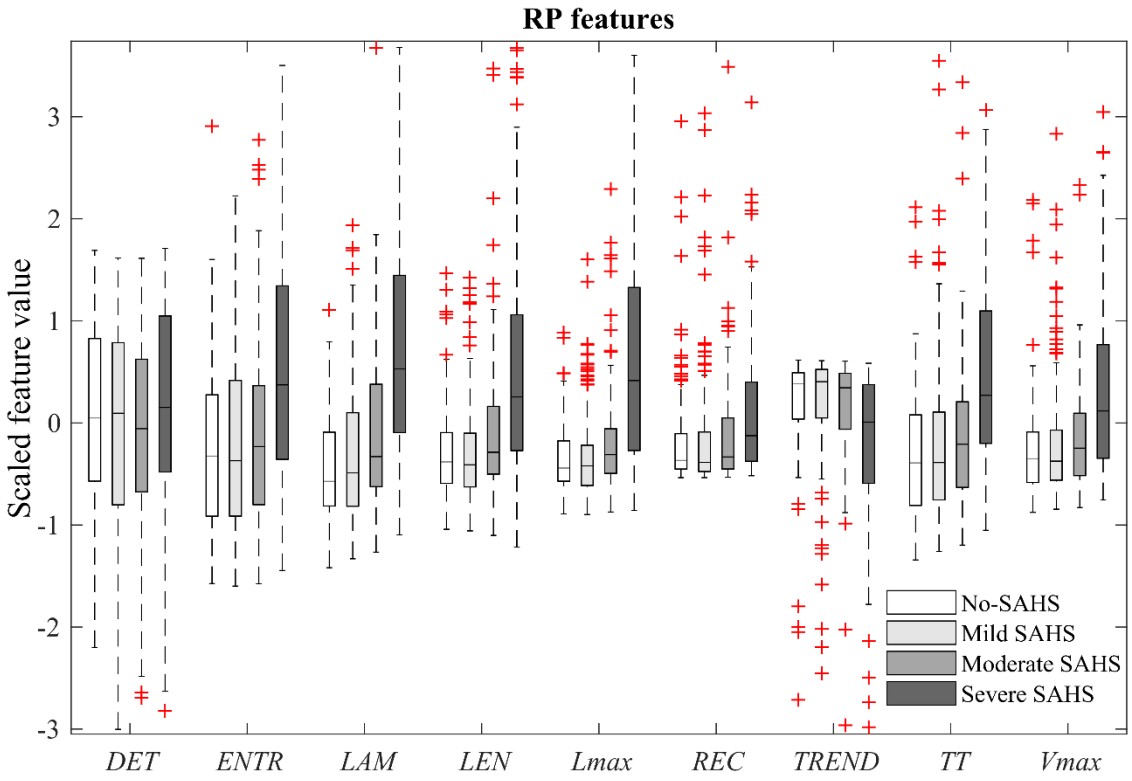

Figure 5. Boxplots of the scaled RP features extracted from AF signals from the training set: determinism (*DET*), Shannon's entropy of the length distribution of the diagonal lines (*ENTR*), laminarity (*LAM*), mean length of the diagonal lines (*LEN*), maximum length of the diagonal lines (*Lmax*), recurrence rate (*REC*), trend (*TREND*), mean length of the vertical lines (*TT*), and maximum length of the vertical lines (*Vmax*).

4.1.2. Optimum feature subset

Nine RP features from AF were obtained in the extraction stage. These features formed the FCBF algorithm input, which only selected *Lmax* more than 50% of the times (923 times). When ODI_3 was incorporated to the selection process, the FCBF chose *Lmax* (778 times) and ODI_3 (1000 times), highlighting its complementarity.

4.1.3. Multi-layer perceptron neural network with Bayesian approach: model optimisation and training

Two BY-MLP models were designed and trained using the corresponding selected features (BY-MLP^{AF}: *Lmax*; BY-MLP^{AF,ODI3}: *Lmax* and ODI_3). The N_H of these BY-MLP were varied from 1 to 30 to optimise its value. For each N_H, kappa was obtained through a leave-one-out cross-validation procedure in the training group and averaged for 10 runs to minimise the random initialization effect of the BY-MLP. The optimum N_H was 17 for BY-MLP^{AF} and 16 for BY-MLP^{AF,ODI3}, since they reached the highest kappa. The optimum N_H and feature subset were used to obtain the final models of BY-MLP^{AF} and BY-MLP^{AF,ODI3} using the whole training group.

4.2. Test group

Table 2 shows the diagnostic performance of BY-MLP^{AF} and BY-MLP^{AF,ODI3} models and single ODI_3, evaluated in the test group (AHI thresholds 1, 5, and 10 e/h). As expected, ODI_3 obtained lower Se in 1 and 5 e/h, underestimating the SAHS presence and agreeing with the literature [15,17]. The BY-MLP^{AF} achieved moderate diagnostic performance, outperforming single ODI_3 in several metrics. The combination of both approaches (BY-MLP^{AF,ODI3}) outperformed BY-MLP^{AF} and single ODI_3 in most of the performance metrics for the three common AHI thresholds, highlighting higher accuracies in 1 e/h and 10 e/h thresholds, as well as a LR- value of 0.1 for 1 e/h and a LR+ value of 13.7 for 10 e/h.

Table 2. Diagnostic performance of BY-MLP^{AF} and BY-MLP^{AF,ODI3} models, and ODI_3 in the test group for AHI thresholds 1, 5, and 10 e/h.

AHI threshold = 1 e/h							
	Se (%)	**Sp (%)**	**PPV (%)**	**NPV (%)**	**LR+**	**LR-**	**Acc (%)**
BY-MLP^{AF}	99,3	4,2	81,4	60,0	1,0	0,2	81,1
BY-MLP^{AF,ODI3}	97,7	22,2	84,1	69,6	1,3	0,1	83,2
ODI_3	59,9	86,1	94,8	33,7	4,3	0,5	64,9
AHI threshold = 5 e/h							
	Se (%)	**Sp (%)**	**PPV (%)**	**NPV (%)**	**LR+**	**LR-**	**Acc (%)**
BY-MLP^{AF}	80,9	48,9	48,7	81,0	1,6	0,4	60,9
BY-MLP^{AF,ODI3}	78,7	78,3	68,5	86,0	3,6	0,2	78,5
ODI_3	69,5	89,4	79,7	83,0	6,5	0,3	81,9
AHI threshold = 10 e/h							
	Se (%)	**Sp (%)**	**PPV (%)**	**NPV (%)**	**LR+**	**LR-**	**Acc (%)**
BY-MLP^{AF}	63,8	85,1	53,7	89,7	4,3	0,4	80,6
BY-MLP^{AF,ODI3}	78,8	94,3	78,8	94,3	13,7	0,2	91,0
ODI_3	81,3	88,5	65,7	94,6	7,1	0,2	87,0

5. DISCUSSION

In this study, we characterised overnight AF signals using common features extracted from RP analysis. We also assessed the utility of these features to detect paediatric SAHS and its severity, as well as its complementarity with ODI_3.

5.1. Airflow characterization in the paediatric sleep apnoea-hypopnoea syndrome context

The averaged RPs from the four SAHS severity groups showed a fade of recurrences, which is typical of non-stationary signals that vary slowly over time [19]. This fact revealed the non-stationarity of overnight AF regardless of the presence and severity of SAHS, reinforcing the need for a non-linear analytical tool to evaluate the signal. Moreover, the predictability of AF, measured by *DET*, presented neither visual nor statistical differences among the four severity groups. This suggests that the apnoeic AF signal is fundamentally predictable regardless the severity group it belongs to. Hence,

RPs would generally define nocturnal AF in children as non-stationary but predictable to some extent.

An increase in the density of recurrences in RPs (higher *REC*) was appreciated as SAHS severity increased, suggesting that apnoeic episodes decrease variability in the AF signal due to its amplitude being reduced to almost zero when these events occur. This results agree with those obtained in a previous work where we analysed the variability of the AF by means of the central tendency measure [14]. In addition, AF showed an increasing trend in the average prediction time (*LEN*) as AHI increases. This indicates that apnoeic events increase the time during which future phase-space trajectories of AF can be predicted from its initial state. Moreover, higher *ENTR* was appreciated as SAHS severity increased, revealing that the apnoeas and hypopnoeas incorporate a wider distribution of diagonal line lengths across the RP. Hence, apnoeic events could cause AF trajectory segments to behave similarly at different times, regardless of their duration. Regarding the degree of non-stationarity, higher absolute values of *TREND* were appreciated when the SAHS severity increased, highlighting that apnoeic events lead to less stationarity in the nocturnal airflow profile. Moreover, AF showed lower complexity (higher values of *TT* and *Vmax*) as AHI increased. This indicates that the occurrence of apnoeas and hypopnoeas could modify the AF dynamics by making it remain longer in a similar state. In this regard, several studies of SAHS in adults have analysed the complexity of AF by means of the Lempel-Ziv complexity [47,48]. However, they did not find potential differences between the SAHS severity groups. It suggests that these RP features could be more effective for this purpose. Despite the trends showed by *REC*, *LEN*, *ENTR*, *TREND*, *TT*, and *Vmax*, only the severe SAHS group presented significant differences (p-value <0.01) with the remaining groups. Therefore, these features may be depicting changes caused by SAHS, but only in certain circumstances such as those

produced in the most severely-affected children.

AF manifested a higher occurrence of laminar states (higher *LAM*) with higher AHI values, which agrees with the information provided by the averaged RPs. This fact suggests that the AF signal does not change state, or changes it very slowly, in the presence of apnoeas and hypopnoeas. This was supported by the significant differences found in *LAM* between the no-SAHS and moderate SAHS groups, and between the severe SAHS group and the remaining groups. Additionally, a decrease in exponential divergence (higher *Lmax*) of the AF phase-space was appreciated as SAHS worsened, indicating that apnoeic episodes could cause the trajectory segments of AF to separate more slowly. This was also observed in the averaged RPs, where the diagonal line clusters were thicker as SAHS severity increased. Moreover, only *Lmax* revealed significant differences between the mild and moderate SAHS group, and between the severe SAHS group and the remaining ones. Therefore, both laminar states (*LAM*) and the exponential divergence (*Lmax*) could be common signs of SAHS in AF and, consequently, be more useful for diagnosis purposes.

According to the aforementioned considerations, the AF characterization indicates that RPs can offer information related to SAHS, as well as the intrinsic nature of overnight AF in children.

5.2. Complementarity with the 3% oxygen desaturation index

In accordance with the relevance shown in the previous AF characterization, *Lmax* was automatically selected by FCBF to be included in the optimum SAHS-related feature subset. However, no other RP feature was selected, showing the redundancy of the remaining extracted features. In addition, ODI_3 was also selected by FCBF, which highlights the complementarity of the RP-derived *Lmax* with this widely-used clinical parameter.

5.3. Diagnostic usefulness and comparison with other studies

BY-MLP^{AF,ODI3} model achieved high diagnostic capability (83.24%, 78.46%, and 90.96% Acc for 1, 5, and 10 e/h, respectively), outperforming BY-MLP^{AF} and single *ODI₃* in most statistical metrics. BY-MLP^{AF,ODI3} obtained higher Se than single *ODI₃* in 1 and 5 e/h, as well as a similar one in 10 e/h, suggesting that SAHS underestimation from *ODI₃* can be minimised using the information extracted from RP of AF. In addition, BY-MLP^{AF,ODI3} reached a LR- of 0.1 for 1 e/h and a LR+ of 13.7 for 10 e/h. This fact is of the utmost importance since LR+ above 10 and LR- below 0.1 are considered to provide strong evidence to establish the presence or absence of a disease, respectively [49]. Accordingly, our proposed approach would be especially useful to confirm the absence of paediatric SAHS, as well as the presence of severe SAHS. Most severely-affected children have a high risk of suffering adverse health consequences and comorbidities [3,27]. Moreover, they can present residual SAHS, as well as persistent risk factors after surgical treatment [29]. Hence, early detection and treatment is required in these cases since a diagnostic delay can lead to serious and potentially irreversible sequelae [29]. In addition, the automatic detection of no-SAHS and severe SAHS cases would reduce the waiting times and the workload of qualified medical personnel, thus being able to focus on the less obvious cases.

Table 3 summarizes previous studies aimed at automatically diagnosing paediatric SAHS by using a reduced set of biomedical signals. The diagnostic ability achieved in these studies is dependent on whether the AHI cut-off point is used to rule out or detect the paediatric SAHS (i.e., it is dependent on the AHI cut-off point selected in each study). The AHI cut-off point = 1 e/h is used to discard the presence of paediatric SAHS [7,26-29]. Therefore, diagnostic metrics like Acc and LR- are important for this purpose since they are related to the classification of healthy subjects. In this regard, only the studies

carried out by Shouldice et al. and Tan et al. [7,26] obtained higher Acc in this AHI threshold than the one obtained in our work. Moreover, it is of the utmost importance to assess the achieved LR- value, since a LR- below 0.1 is considered to provide strong evidence to establish the absence of a disease [49]. In this sense, our proposal obtained a LR- = 0.1 for 1 event/h, being this value lower than all those obtained in the state-of-the-art studies found in the context of paediatric SAHS.

In order to detect the presence of moderate SAHS, 5 e/h is used as threshold [8-12,14,26-29,40,50]. Therefore, Se, Acc, and LR+ are important diagnostic metrics for this purpose since they are related to the classification of SAHS subjects. In a previous study, Barroso-García et al. [14], reported 65.0% Se and 76.0% Acc for this AHI thresholds. However, our new proposed methodology obtained a higher Se and Acc (78.7% Se and 78.5% Acc), providing a better diagnostic capacity to detect moderate SAHS. On the other hand, the work carried out by Gil et al., Hornero et al., Álvarez et al., Vaquerizo-Villar et al., and Xu et al. [8,10,11,14,26,40,50], obtained a lower Se for this AHI threshold than the one in our work. Thus, the proportion of paediatric subjects correctly classified with AHI $\geq$ 5 e/h is considerably higher in our study, providing a higher diagnostic capacity to detect paediatric subjects with moderate SAHS, which is consistent with the purpose of using this AHI threshold. In addition, despite of the high Se and Acc values obtained by Lázaro et al. [9], the proportion of subjects with AHI < 5 e/h wrongly classified (LR+) is lower than in our study (3.5 versus 3.6), providing less evidence to establish moderate SAHS than our study.

Regarding the AHI cut-off point 10 e/h, this threshold is used to detect the presence of severe SAHS in children [10,14,26-29,50]. Thus, the most important diagnostic metrics for this purpose are again Se, Acc, and LR+. The Acc and LR+ obtained for this AHI cut-off point in our study are very high (91.0% Acc and 13.7 LR+) with respect to

other studies [10,14,28,40,50]. This fact is very important since, according to Deesks et al. [49], a LR+ above 10 is considered to provide strong evidence to establish the presence of a disease, (i.e., severe SAHS). These results far outperform those achieved in others state-of-the-art studies in the context of detection of severe paediatric SAHS, excluding the study carried out by Tan et al. [26]. However, Tan et al. [26] obtained a lower Se for the AHI threshold = 10 e/h. Thus, the proportion of subjects correctly classified with severe SAHS is considerably higher in our study, which is consistent with the purpose of using this AHI threshold.

Hence, our methodology obtained a high diagnostic Acc for 1 and 10 e/h, and provided stronger evidence to discard the SAHS presence (lower LR-) and to establish severe SAHS (higher LR+) than others state-of-the-art studies. Moreover, we have evaluated our hypothesis according to the AHI thresholds commonly used to determine the SAHS severity degrees (1, 5, and 10 e/h), and validated our methodology using a large cohort (946 subjects). This fact is also important to highlight, since this allowed us to achieve more generalizable results than others aforementioned studies [7-9,11,12,14,26,28,40,50] and, thus, reflect the actual underlying behaviour of the paediatric population.

Table 3. Summary of the state-of-the-art in the context of detection of paediatric SAHS.

Studies	Subjects (n)	Signal	AHI (e/h)	Methods (Analysis/Selection/Classifier)	Se (%)	Sp (%)	Acc (%)	PPV (%)	NPV (%)	LR+	LR-
Shouldice *et al.* (2004) [7]	50	ECG	1	Temporal and spectral analysis / – / QDA	85.7	81.8	84.0	85.7	81.8	4.7	0.2
Gil *et al.* (2010) [8]	21	PPG	5	Analysis of PTTV / Wrapper methodology / LDA	75.0	85.7	80.0	-	-	5.2*	0.3*
Tsai *et al.* (2013) [28]	148	SpO_2	1	Oxygen desaturation index of 4% (ODI_4) / – / – /	77.7	88.9	79.0*	-	-	7.0*	0.3*
			5		83.8	86.5	85.1*	-	-	6.2*	0.2*
			10		89.1	86.0	87.1*	-	-	6.4*	0.1*
Tan *et al.* (2014) [26]	100	ECG AF SpO_2 RIP	1	Comparison of the AHI obtained from PSG with the AHI directly estimated of respiratory polygraphic (RP) / – / – /	82.5	90.0	86.0*	97.1	56.3	8.3*	0.2*
			5		62.5	100	85.0*	100	80	Inf*	0.4*
			10		65.0*	100*	93.0*	100*	92.0*	Inf*	0.4*
Lázaro *et al.* (2014) [9]	21	PPG	5	Spectral analysis of PRV and DAP events detection / Wrapper methodology / LDA	100	71.4	86.7	-	-	3.5*	0*
Garde *et al.* (2014) [12]	146	SpO_2 PRV	5	Temporal and spectral analysis / Selection algorithm optimizing the AROC / LDA	88.4	83.6	84.9	76.9	92.6	5.4*	0.1*
Gutiérrez-Tobal *et al.* (2015) [13]	50	AF SpO_2	3	Spectral features and oxygen desaturation index of 3% (ODI_3) / FSLR / LR	85.9	87.4	86.3	88.4	85.8	6.8*	0.2*
Barroso-García *et al.* (2017) [14]	501	AF	1	Spectral entropy and Central tendency measure / FSLR / LR	60.5	58.6	60.0	81.2	25.0	1.1	0.9
			5		65.0	80.6	76.0	70.7	78.2	3.6	0.4
			10		83.3	79.0	80.0	52.8	93.5	4.0	0.2
Hornero *et al.* (2017) [10]	4191	SpO_2	1	Statistical, spectral, non-linear features, and ODI_3 / FCBF / MLP	84.0	53.2	75.2	81.6	53.7	1.8	0.3
			5		68.2	87.2	81.7	68.6	87.0	5.3	0.4
			10		68.7	94.1	90.2	67.7	94.3	11.6	0.3
Álvarez *et al.* (2018) [11]	142	SpO_2	5	Anthropometric, statistical moments, desaturation indices, symbolic dynamics / FSLR / LR	73.5	89.5	83.3	82.0	84.3	10.4	0.3
Vaquerizo-Villar *et al.* (2018) [40]	298	SpO_2	5	Anthropometric variables, ODI_3, spectral features from power spectral density and bispectrum / FCBF / MLP	61.8	97.6	81.3	95.5	75.5	25.3	0.4
			10		60.0	94.5	85.3	80.0	86.7	11.0	0.4
Xu *et al.* (2018) [50]	432	SpO_2	1	ODI_3 and third statistical moment of the spectral band of interest/ FCBF / MLP	95.3	19.1	79.6	82.0*	51.5*	1.2	0.2
			5		77.8	80.5	79.4	72.3*	84.7*	4.0	0.3
			10		73.5	92.7	88.2	75.8*	91.9*	10.1	0.3
Our proposal	946	AF SpO_2	1	Features of Recurrence Plots and ODI_3 / FCBF / BY-MLP	97.7	22.2	83.2	84.1	69.6	1.3	0.1
			5		78.7	78.3	78.5	68.5	86.0	3.6	0.3
			10		78.8	94.3	91.0	78.8	94.3	13.7	0.2

QDA: Quadratic discriminant analysis; PTTV: Pulse transit time variability; LDA: Linear discriminant analysis; RIP: Chest and abdominal movement by respiratory inductance plethysmography; PRV: Pulse rate variability; DAP: Decreases in amplitude fluctuations of the PPG signal; AROC: Area under the receiver operating characteristics curves; FSLR: Forward stepwise logistic regression; LR: Logistic regression model; FCBF: Fast correlation-based filter; MLP: Multi-Layer Perceptron neural network; BY-MLP: MLP with Bayesian approach.

* Computed from reported data.

5.4. Limitations

In spite of the usefulness illustrated by our proposed approach, this study has some limitations. Although our database is large, more subjects originating from other sleep laboratories would make our results more generalizable. Furthermore, this would be also convenient for the sake of the proportion of the subjects belonging to each of the SAHS severity groups. In addition, the selection of RP embedding parameters (m, τ, and ε) is a widely discussed topic for which there is no consensus as of yet [20]. The sensitivity of the RP analysis to these parameters, along with the lack of consensus on the method that should be used to optimise them, is another limitation. Consequently, although our study supports the use of AMI, FNN and the fixed distance method, the use, combination, and comparison of other methodologies could be the subject of future research. Moreover, it would also be interesting to evaluate other machine-learning algorithms with the ability to estimate AHI. In addition, conducting studies with AF signals recorded at home is another future goal. Finally, analysing the effects of our proposed method using other signals like ECG or PPG could be an interesting future research.

6. CONCLUSIONS

In summary, this is the first study in which RP features from AF have been used to help in the diagnosis of paediatric SAHS. We have shown that RP can offer useful information related to both paediatric SAHS and the intrinsic characteristics of overnight AF in children. Our results also revealed that the exponential divergence of the phase-space trajectories of AF provides complementary information to the oxygen desaturations events. Combining both approaches, a BY-MLP neural network reached a high diagnostic performance for 1, 5, and 10 e/h. Moreover, this model achieved strong evidence to rule out SAHS, as well as to predict severe cases. Hence, these findings suggest that the

analysis of RP applied to AF signal provides useful information that could be used along with ODI_3 to help in paediatric SAHS diagnosis.

ACKNOWLEDGEMENTS

This work was supported by 'Ministerio de Ciencia, Innovación y Universidades' and ‘European Regional Development Fund (FEDER)’ under projects DPI2017-84280-R and RTC-2017-6516-1, and by ‘European Commission’ and ‘FEDER’ under project ‘POCTEP 0378_AD_EEGWA_2_P’. V. Barroso-García was in receipt of a ‘Ayuda para financiar la contratación predoctoral de personal investigador’ grant from the Consejería de Educación de la Junta de Castilla y León and the European Social Fund. F. Vaquerizo-Villar was in receipt of a ‘Ayuda para contratos predoctorales para la Formación de Profesorado Universitario (FPU)’ grant from the Ministerio de Educación, Cultura y Deporte (FPU16/02938). P. Núñez was in receipt of a predoctoral scholarship ‘Ayuda para contratos predoctorales para la Formación de Profesorado Universitario (FPU)’ grant from the Ministerio de Educación, Cultura y Deporte (FPU17/00850). L. Kheirandish-Gozal and D. Gozal were supported by National Institutes of Health (NIH) grant HL130984.

CONFLICT OF INTEREST

There are no conflicts of interest that could inappropriately influence this research work.

ETHICAL APPROVAL

In all participants, the informed consent to be included in the research was obtained and the Ethical Committee of the Hospital approved the protocol.

AUTHORSHIP RESPONSIBILITY

- The material in this manuscript is original and contains no matter libellous or otherwise unlawful.

- The manuscript represents valid work and that neither this manuscript nor any other with substantially similar content under my authorship has been published or is being considered for publication elsewhere.

- I have participated sufficiently in the work to take public responsibility for all its content.

REFERENCES

[1] W.W. Flemons, M.R. Littner, J.A. Rowley, P. Gay, W.M.D. Anderson, D.W. Hudgel, R.D. McEvoy, D.I. Loube, Home Diagnosis of Sleep Apnea: A Systematic Review of the Literature - An Evidence Review Cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society, Chest. 124 (2003) 1543–79. doi:10.1378/chest.124.4.1543.

[2] C.L. Marcus, L.J. Brooks, K.A. Draper, D. Gozal, A.C. Halbower, J. Jones, M.S. Schechter, S.H. Sheldon, K. Spruyt, S.D. Ward, C. Lehmann, R.N. Shiffman, Diagnosis and management of childhood obstructive sleep apnea syndrome, Pediatrics. 130 (2012) 576–84. doi:10.1542/peds.2012-1671.

[3] A.G. Kaditis, M.L. Alonso Alvarez, A. Boudewyns, E.I. Alexopoulos, R. Ersu, K. Joosten, H. Larramona, S. Miano, I. Narang, H. Trang, M. Tsaoussoglou, N. Vandenbussche, M.P. Villa, D. Van Waardenburg, S. Weber, S. Verhulst, Obstructive sleep disordered breathing in 2- to 18-year-old children: diagnosis and management, Eur. Respir. J. 47 (2016) 69–94. doi:10.1183/13993003.00385-2015.

[4] C. Jon, Polysomnography in Children, in: R.B. Mitchell, K.D. Pereira (Eds.), Pediatr. Otolaryngol. Clin., Humana Press, Totowa, NJ, 2009: pp. 35–47. doi:10.1007/978-1-60327-127-1_5.

[5] P.J. Ryan, M.F. Hilton, D.A. Boldy, A. Evans, S. Bradbury, S. Sapiano, K. Prowse, R.M. Cayton, Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnoea/hypopnoea syndrome: can polysomnography be avoided?, Thorax. 50 (1995) 972–5. doi:10.1136/thx.50.9.972.

[6] E.S. Katz, R.B. Mitchell, C.M. D'Ambrosio, Obstructive Sleep Apnea in Infants, Am. J. Respir. Crit. Care Med. 185 (2012) 805–16. doi:10.1164/rccm.201108-1455CI.

[7] R.B. Shouldice, L.M. O'Brien, C. O'Brien, P. de Chazal, D. Gozal, C. Heneghan, Detection of Obstructive Sleep Apnea in Pediatric Subjects using Surface Lead Electrocardiogram Features, Sleep. 27 (2004) 784–92. doi:10.1093/sleep/27.4.784.

[8] E. Gil, R. Bailon, J.M. Vergara, P. Laguna, PTT Variability for Discrimination of Sleep Apnea Related Decreases in the Amplitude Fluctuations of PPG Signal in Children, IEEE Trans. Biomed. Eng. 57 (2010) 1079–88. doi:10.1109/tbme.2009.2037734.

[9] J. Lazaro, E. Gil, J.M. Vergara, P. Laguna, Pulse Rate Variability Analysis for Discrimination of Sleep-Apnea-Related Decreases in the Amplitude Fluctuations of Pulse Photoplethysmographic Signal in Children, IEEE J. Biomed. Heal. Informatics. 18 (2014) 240–6. doi:10.1109/JBHI.2013.2267096.

[10] R. Hornero, L. Kheirandish-Gozal, G.C. Gutiérrez-Tobal, M.F. Philby, M.L. Alonso-Álvarez, D. Álvarez, E.A. Dayyat, Z. Xu, Y.-S. Huang, M. Tamae Kakazu, A.M. Li, A. Van Eyck, P.E. Brockmann, Z. Ehsan, N. Simakajornboon, A.G. Kaditis, F. Vaquerizo-Villar, A. Crespo Sedano, O. Sans Capdevila, M. von Lukowicz, J. Terán-Santos, F. Del Campo, C.F. Poets, R. Ferreira, K. Bertran, Y. Zhang, J. Schuen, S. Verhulst, D. Gozal, Nocturnal Oximetry–based Evaluation of Habitually Snoring Children, Am. J. Respir. Crit. Care Med. 196 (2017) 1591–8. doi:10.1164/rccm.201705-0930OC.

[11] D. Álvarez, A. Crespo, F. Vaquerizo-Villar, G.C. Gutiérrez-Tobal, A. Cerezo-

Hernández, V. Barroso-García, J.M. Ansermino, G.A. Dumont, R. Hornero, F. del Campo, A. Garde, G.C. Gutierrez-Tobal, A. Cerezo-Hernández, V. Barroso-García, J.M. Ansermino, G.A. Dumont, R. Hornero, F. Del Campo, A. Garde, Symbolic dynamics to enhance diagnostic ability of portable oximetry from the phone oximeter in the detection of paediatric sleep apnoea, Physiol. Meas. 39 (2018) 104002. doi:10.1088/1361-6579/aae2a8.

[12] A. Garde, P. Dehkordi, W. Karlen, D. Wensley, J.M. Ansermino, G.A. Dumont, Development of a Screening Tool for Sleep Disordered Breathing in Children Using the Phone Oximeter™, PLoS One. 9 (2014) e112959. doi:10.1371/journal.pone.0112959.

[13] G.C. Gutiérrez-Tobal, M.L. Alonso-Álvarez, D. Álvarez, F. del Campo, J. Terán-Santos, R. Hornero, Diagnosis of pediatric obstructive sleep apnea: Preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients' home, Biomed. Signal Process. Control. 18 (2015) 401–7. doi:10.1016/j.bspc.2015.02.014.

[14] V. Barroso-García, G.C. Gutiérrez-Tobal, L. Kheirandish-Gozal, D. Álvarez, F. Vaquerizo-Villar, A. Crespo, F. del Campo, D. Gozal, R. Hornero, Irregularity and Variability Analysis of Airflow Recordings to Facilitate the Diagnosis of Paediatric Sleep Apnoea-Hypopnoea Syndrome, Entropy. 19 (2017) 447. doi:10.3390/e19090447.

[15] G.C. Gutiérrez-Tobal, D. Alvarez, A. Crespo, F. del Campo, R. Hornero, Evaluation of Machine-Learning Approaches to Estimate Sleep Apnea Severity from at-Home Oximetry Recordings, IEEE J. Biomed. Heal. Informatics. (2018). doi:10.1109/JBHI.2018.2823384.

[16] U.J. Magalang, J. Dmochowski, S. Veeramachaneni, A. Draw, M.J. Mador, A. El-Solh, B.J.B. Grant, Prediction of the Apnea-Hypopnea Index From Overnight Pulse Oximetry, Chest. 124 (2003) 1694–701. doi:10.1378/CHEST.124.5.1694.

[17] L. Chang, J. Wu, L. Cao, Combination of symptoms and oxygen desaturation index in predicting childhood obstructive sleep apnea, Int. J. Pediatr. Otorhinolaryngol. 77 (2013) 365–71. doi:10.1016/J.IJPORL.2012.11.028.

[18] R.B. Berry, R. Budhiraja, D.J. Gottlieb, D. Gozal, C. Iber, V.K. Kapur, C.L. Marcus, R. Mehra, S. Parthasarathy, S.F. Quan, S. Redline, K.P. Strohl, S.L.D. Ward, M.M. Tangredi, 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. 08 (2012) 597–619. doi:10.5664/jcsm.2172.

[19] N. Marwan, M. Carmen Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems, Phys. Rep. 438 (2007) 237–329. doi:10.1016/j.physrep.2006.11.001.

[20] S. Martín-González, J.L. Navarro-Mesa, G. Juliá-Serdá, G.M. Ramírez-Ávila, A.G. Ravelo-García, Improving the understanding of sleep apnea characterization using Recurrence Quantification Analysis by defining overall acceptable values for the dimensionality of the system, the delay, and the distance threshold, PLoS One. 13 (2018) e0194462. doi:10.1371/journal.pone.0194462.

[21] C.L. Webber, J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies, J. Appl. Physiol. 76 (1994) 965–73. doi:10.1152/jappl.1994.76.2.965.

[22] U.R. Acharya, S.V. Sree, S. Chattopadhyay, W. Yu, P.C.A. Ang, Application of recurrence quantification analysis for the automated identification of epileptic

EEG signals, Int. J. Neural Syst. 21 (2011) 199–211. doi:10.1142/S0129065711002808.

[23] K. Becker, G. Schneider, M. Eder, A. Ranft, E.F. Kochs, W. Zieglgänsberger, H.-U. Dodt, Anaesthesia Monitoring by Recurrence Quantification Analysis of EEG Data, PLoS One. 5 (2010) e8876. doi:10.1371/journal.pone.0008876.

[24] H.D. Nguyen, B.A. Wilkins, Q. Cheng, B.A. Benjamin, An Online Sleep Apnea Detection Method Based on Recurrence Quantification Analysis, IEEE J. Biomed. Heal. Informatics. 18 (2014) 1285–93. doi:10.1109/JBHI.2013.2292928.

[25] K. Spruyt, Pediatric Sleep-Disordered Breathing: Criteria and Spectrum of Disease, in: L. Kheirandish-Gozal, D. Gozal (Eds.), Sleep Disord. Breath. Child., Humana Press, New York, 2012: pp. 245–60. doi:10.1007/978-1-60761-725-9_18.

[26] H.-L. Tan, D. Gozal, H.M. Ramirez, H.P.R. Bandla, L. Kheirandish-Gozal, Overnight Polysomnography versus Respiratory Polygraphy in the Diagnosis of Pediatric Obstructive Sleep Apnea, Sleep. 37 (2014) 255–60. doi:10.5665/sleep.3392.

[27] G.D. Church, The Role of Polysomnography in Diagnosing and Treating Obstructive Sleep Apnea in Pediatric Patients, Curr. Probl. Pediatr. Adolesc. Health Care. 42 (2012) 2–25. doi:10.1016/J.CPPEDS.2011.10.001.

[28] C.-M. Tsai, C.-H. Kang, M.-C. Su, H.-C. Lin, E.-Y. Huang, C.-C. Chen, J.-C. Hung, C.-K. Niu, D.-L. Liao, H.-R. Yu, Usefulness of desaturation index for the assessment of obstructive sleep apnea syndrome in children, Int. J. Pediatr. Otorhinolaryngol. 77 (2013) 1286–90. doi:10.1016/J.IJPORL.2013.05.011.

[29] M. Luz Alonso-Álvarez, T. Canet, M. Cubell-Alarco, E. Estivill, E. Fernández-Julián, D. Gozal, M.J. Jurado-Luque, M.A. Lluch-Roselló, F. Martínez-Pérez, M. Merino-Andreu, G. Pin-Arboledas, N. Roure, F.X. Sanmartí, Ó. Sans-Capdevila, F. Segarra-Isern, M. Tomás-Vila, J. Terán-Santos, Consensus document on sleep apnea-hypopnea syndrome in children, Arch. Bronconeumol. 47 (2011) 2–18. doi:10.1016/S0300-2896(11)70026-6.

[30] P. Várady, T. Micsik, S. Benedek, Z. Benyó, A novel method for the detection of apnea and hypopnea events in respiration signals, IEEE Trans. Biomed. Eng. 49 (2002) 936–42. doi:10.1109/TBME.2002.802009.

[31] R.M. Rangayyan, Filtering for removal of artifacts, in: Biomed. Signal Anal., Second, John Wiley & Sons, 2015: pp. 91–231. doi:10.1002/9781119068129.ch3.

[32] B.H. Taha, J.A. Dempsey, S.M. Weber, M.S. Badr, J.B. Skatrud, T.B. Young, A.J. Jacques, K.C. Seow, Automated Detection and Classification of Sleep-Disordered Breathing From Conventional Polysomnography Data, Sleep. 20 (1997) 991–1001. doi:10.1093/sleep/20.11.991.

[33] F. Takens, Detecting strange attractors in turbulence, in: D. Rand, L.S. Young (Eds.), Dyn. Syst. Turbul. Warwick 1980. Lect. Notes Math. Vol 898, Springer, Berlin, Heidelberg, 1981: pp. 366–381. doi:10.1007/BFb0091924.

[34] S. Schinkel, O. Dimigen, N. Marwan, Selection of recurrence threshold for signal detection, Eur. Phys. J. Spec. Top. 164 (2008) 45–53. doi:10.1140/epjst/e2008-00833-5.

[35] M.B. Kennel, R. Brown, H.D.I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, Phys. Rev. A.

45 (1992) 3403–11. doi:10.1103/PhysRevA.45.3403.

[36] J.P. Zbilut, C.L. Webber, Recurrence Quantification Analysis, in: M. Akay (Ed.), Wiley Encycl. Biomed. Eng., John Wiley & Sons, Inc., Hoboken, NJ, USA, 2006. doi:10.1002/9780471740360.ebs1355.

[37] L. Yu, Liu, Efficient Feature Selection via Analysis of Relevance and Redundancy, J. Mach. Learn. Res. 5 (2004) 1205–24.

[38] E. Pereda, M. García-Torres, B. Melián-Batista, S. Mañas, L. Méndez, J.J. González, The blessing of Dimensionality: Feature Selection outperforms functional connectivity-based feature transformation to classify ADHD subjects from EEG patterns of phase synchronisation, PLoS One. 13 (2018) e0201660. doi:10.1371/journal.pone.0201660.

[39] A. Khaleghi, A. Sheikhani, M.R. Mohammadi, A.M. Nasrabadi, S.R. Vand, H. Zarafshan, M. Moeini, EEG classification of adolescents with type I and type II of bipolar disorder, Australas. Phys. Eng. Sci. Med. 38 (2015) 551–559. doi:10.1007/s13246-015-0375-0.

[40] F. Vaquerizo-Villar, D. Álvarez, L. Kheirandish-Gozal, G.C. Gutiérrez-Tobal, V. Barroso-García, A. Crespo, F. del Campo, D. Gozal, R. Hornero, Utility of bispectrum in the screening of pediatric sleep apnea-hypopnea syndrome using oximetry recordings, Comput. Methods Programs Biomed. 156 (2018) 141–149. doi:10.1016/J.CMPB.2017.12.020.

[41] I.H. Witten, E. Frank, M.A. Hall, Data mining: practical machine learning tools and techniques, Third, Morgan Kaufmann/Elsevier, Burlington, 2011. doi:10.1016/C2009-0-19715-5.

[42] C.M. Bishop, Neural networks for pattern recognition, Oxford university Press,

1996.

[43] J. V Marcos, R. Hornero, D. Álvarez, I.T. Nabney, F. del Campo, C. Zamarrón, The classification of oximetry signals using Bayesian neural networks to assist in the detection of obstructive sleep apnoea syndrome, Physiol. Meas. 31 (2010) 375–94. doi:10.1088/0967-3334/31/3/007.

[44] J. Cohen, A Coefficient of Agreement for Nominal Scales, Educ. Psychol. Meas. 20 (1960) 37–46. doi:10.1177/001316446002000104.

[45] W.W. Flemons, M.R. Littner, Measuring Agreement Between Diagnostic Devices, Chest. 124 (2003) 1535–42. doi:10.1378/CHEST.124.4.1535.

[46] M.H. Zweig, G. Campbell, Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine, Clin. Chem. 39 (1993) 561–77.

[47] G.C. Gutiérrez-Tobal, S. Member, D. Alvarez, F. del Campo, R. Hornero, S. Member, Utility of AdaBoost to Detect Sleep Apnea-Hypopnea Syndrome From Single-Channel Airflow, IEEE Trans. Biomed. Eng. 63 (2016). doi:10.1109/TBME.2015.2467188.

[48] G.C. Gutierrez-Tobal, R. Hornero, D. Álvarez, J.V. Marcos, F. del Campo, Linear and nonlinear analysis of airflow recordings to help in sleep apnoea–hypopnoea syndrome diagnosis, Physiol. Meas. 33 (2012) 1261–75. doi:10.1088/0967-3334/33/7/1261.

[49] J.J. Deeks, D.G. Altman, Diagnostic tests 4: likelihood ratios, BMJ. 329 (2004) 168–9. doi:10.1136/bmj.329.7458.168.

[50] Z. Xu, G.C. Gutiérrez-Tobal, Y. Wu, L. Kheirandish-Gozal, X. Ni, R. Hornero, D. Gozal, Cloud Algorithm-Driven Oximetry-Based Diagnosis of Obstructive Sleep Apnea in Symptomatic Habitually-Snoring Children., Eur. Respir. J.

(2018) 1801788. doi:10.1183/13993003.01788-2018.