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

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Goza, Daniel Álvarez, Fernando Vaquerizo-Villar, Pablo Núñez, Félix del Campo, David Goza, Roberto Hornero

ARTICLE INFO

Article history:
Received 26 March 2019
Revised 28 August 2019
Accepted 16 September 2019

Keywords:
Airflow (AF)
Children
Recurrence plots (RP)
Sleep Apnoea-Hypopnoea Syndrome (SAHS)

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 (ODI3).

Methods: 946 AF and blood oxygen saturation (SpO2) 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. ODI3 was also calculated from each SpO2 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 multilayer 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 ODI3. 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 event/h and a positive likelihood ratio of 13.7 for 10 events/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 ODI3. Thus, RP applied to AF signals can be used along with ODI3 to help in paediatric SAHS diagnosis, particularly to either confirm the absence of SAHS or the presence of severe SAHS.

© 2019 Elsevier B.V. All rights reserved.

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

https://doi.org/10.1016/j.cmpb.2019.105083
0169-2607/© 2019 Elsevier B.V. All rights reserved.
facilities and the expertise needed to carry it out are not always available, which entails prolonged access delays [5]. Moreover, the multivariable 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 (SpO2), 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 (ODI3), 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 ODI3, which is obtained from SpO2, 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 ODI3 is also addressed in our study.

2. Subjects and signals under study

In this study, AF and SpO2 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.

PSG was conducted using a digital polysomnography system (Polysmith, Nihon Kohden America Inc., Irvine, CA, USA). Single-channel AF and SpO2 signals were acquired during the PSG. Recordings lasting less than 3 h 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-s epochs [31]. The SpO2 recordings, sampled at 25Hz, were used to obtain ODI3. Their artefacts were removed by discarding SpO2 values <50% and changes with a slope ≥4% [16]. Fig. 1(a) shows an example of the AF signal and Fig. 1(b) shows the corresponding SpO2 signal. We can see 7 apnoea events (absence of AF, i.e., near zero amplitude) in 0.4, 0.9, 1.7, 2.6, 4.9, 5.4, and 6.3 min, and the corresponding desaturations.

3. Methods

ODI3 was obtained from SpO2 recordings according to Taha et al. [32]. Therefore, oxygen desaturation events were scored as a decrease ≥3% of SpO2 at a rate of 0.1–4%/s, during 10–60 s. The number of desaturation events was divided by the number of recording hours to obtain ODI3.

 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). Fig. 2 shows the block diagram of the method proposed in our study.

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 Rij,
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]:

\[ X_i = [u(i), u(i + \tau), \ldots, u(i + (m - 1) \cdot \tau)], \]

where \( u(i) \) is the value of the time series at time \( i \), \( m \) is the embedding dimension, and \( \tau \) is the time delay. The embedded dimension \( m \) and the time delay \( \tau \) are parameters to be optimized in each context. The time delay has to be adjusted to not set up autocorrelated 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, \( D_{ij} \), is calculated by the commonly used Euclidean distance norm [19,20]:

\[ D_{ij} = \| X_i - X_j \|. \]

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

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

\[ R_{ij} = \begin{cases} 1 : D_{ij} \leq \varepsilon \\ 0 : D_{ij} > \varepsilon \end{cases}. \]

A proper selection of \( \varepsilon \) is essential too. If \( \varepsilon \) 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 \( \varepsilon \) 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 \( \sigma \) of each time series, has been widely used to select \( \varepsilon \) [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 Fig. 3, isolated recurrences, vertical (i.e., lamar 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 Fig. 3), providing information about the variability of time series [19,36]. The more recurrences there are, the less variable the time series is [20]:

\[
REC = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{ij},
\]

where \( N \) is the number of rows in \( R_{ij} \).

• 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 Fig. 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]:

\[
DET = \frac{\sum_{l=\min}^{N} l \cdot p(l)}{\sum_{l=1}^{N} p(l)},
\]

where \( \min \) 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 \( \min \) 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]:

\[
LEN = \frac{\sum_{l=\min}^{N} l \cdot p(l)}{\sum_{l=\min}^{N} p(l)}.
\]

This feature can be observed in Fig. 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 diverge, the shorter the diagonal lines [19]:

\[
L_{\max} = \max(l, i = 1, ..., N).
\]

This feature can be observed in Fig. 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]:

\[
ENTR = -\sum_{l=\min}^{N} p(l) \cdot \log(p(l)).
\]

– 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]:

\[
TREND = \frac{\sum_{i=1}^{N} (i - \bar{N}/2) \cdot (REC_i - \langle REC \rangle)}{\sum_{i=1}^{N} (i - \bar{N}/2)^2},
\]

where \( \bar{N} \) 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 \( \langle REC \rangle \) is the average of \( REC_i \). We define \( \bar{N} \) as \( N - 2 \), which is a commonly used value [19,36]. This feature can be observed in Fig. 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 Fig. 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]:

\[
LAM = \frac{\sum_{v=\min}^{N} v \cdot p(v)}{\sum_{v=\min}^{N} p(v)}.
\]

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]:

\[
TT = \frac{\sum_{v=\min}^{N} v \cdot p(v)}{\sum_{v=\min}^{N} p(v)}.
\]

This feature can be observed in Fig. 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]:

\[
V_{\text{max}} = \max(v_i, i = 1, ..., N).
\]

This feature can be observed in Fig. 3.

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 \( (p\text{-value} < 0.01 \) 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 correctly classified, specificity (Sp); proportion of subjects without the disease correctly classified), accuracy (Acc; proportion of overall subjects correctly 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 \( OD_{ij} \) 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 \( \tau \) in the phase-space reconstruction of AF signals. The \( \tau \) was varied from 0.1 to 6 s. The optimum \( \tau \) value (first relative minimum of AMI) was obtained for each subject of the training group. The median of these values determined the optimum \( \tau \) for AF: \( \tau = 0.9 \) s. 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 \( \epsilon \) was varied from \( (10^{-2} \cdot \sigma) \) to \( (25 \cdot 10^{-2} \cdot \sigma) \) to optimise its value. The \( \epsilon \) was fixed to \( \epsilon = 10^{-2} \cdot \sigma \) 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-s 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, Fig. 4 illustrates the averaged RPs in each out of the

![Fig. 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 \( \tau = 0.9 \) s, and being 100 Hz the sampling frequency and 30-s the window size, Eq. (1) allow to define the phase-space trajectories of airflow recordings as \( \tilde{x}_t = [u(1), u(91), u(181)] \) to \( x_{3200} = [u(2820), u(2910), u(3000)] \).](image-url)
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.

Fig. 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.

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 ODI1 was incorporated to the selection process, the FCBF chose Lmax (778 times) and ODI1 (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\textsuperscript{AF}: Lmax; BY-MLP\textsuperscript{REC,ODI}: Lmax and ODI1). The N\textsubscript{tr} of these BY-MLP were varied from 1 to 30 to optimise its value. For each N\textsubscript{tr}, 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\textsubscript{tr} was 17 for BY-MLP\textsuperscript{AF} and 16 for BY-MLP\textsuperscript{REC,ODI} since they reached the highest kappa. The optimum N\textsubscript{tr} and feature subset were used to obtain the final models of BY-MLP\textsuperscript{AF} and BY-MLP\textsuperscript{REC,ODI} using the whole training group.

4.2. Test group

Table 2 shows the diagnostic performance of BY-MLP\textsuperscript{AF} and BY-MLP\textsuperscript{REC,ODI} models and single ODI1, evaluated in the test group (AHI thresholds 1, 5, and 10 e/h). As expected, ODI1 obtained lower Se in 1 and 5 e/h, underestimating the SAHS presence and agreeing with the literature [15,17]. The BY-MLP\textsuperscript{AF} achieved moderate diagnostic performance, outperforming single ODI1 in several metrics. The combination of both approaches (BY-MLP\textsuperscript{REC,ODI}) outperformed BY-MLP\textsuperscript{AF} and single ODI1 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.

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 ODI1.

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, ODI3 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<sub>AF,ODI3</sub> model achieved high diagnostic capability (83.24%, 78.46%, and 90.96% Acc for 1, 5, and 10 e/h, respectively), outperforming BY-MLP<sub>AF</sub> and single ODI3 in most statistical metrics. BY-MLP<sub>AF,ODI3</sub> 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<sub>AF,ODI3</sub> 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 exclude
or detect the paediatric SAHS (i.e., it is dependent on the AHI threshold selected in each study). The AHI cut-off point = 1 e/h is more clinically useful to discard paediatric SAHS, rather than confirm its presence [7,26–29]. The global measurement of the diagnostic ability (Acc = 83.2%) obtained in our study for this threshold is high compared to other state-of-the-art studies. Our proposal outperformed most of the them and is close to the highest (Acc = 84.0% and 86.0%) [7,26]. Moreover, our LR− (−0.1) is the lowest among all the state-of-the-art studies found in the context of paediatric SAHS. This is of the utmost importance since it has been reported that a value ≤0.1 provides strong evidence to establish the absence of a disease [49]. Thus, our proposal is the most appropriate to discard the presence of SAHS.

AHI = 5 e/h is used as threshold to detect moderate SAHS [8–12,14,26–29,40,50]. The Acc reached in our study (78.5%) is slightly lower than the mean of the state-of-the-art studies (82.0%). However, most of them report an unbalanced Se/Sp pair. In contrast, our Se/Sp pair is almost perfectly balanced, which may be particularly desirable in order to provide similar importance to the classification of subjects above and below the threshold of this intermediate illness degree.

Finally, the AHI cut-off point 10 e/h is used to detect the presence of severe SAHS in children [10,14,26–29,50]. The Acc and LR+ obtained for this threshold in our study are very high (91.0% Acc and 13.7 LR+) compared to other studies [10,14,28,40,50]. This is another important result from our proposal since LR+ above 10 is considered to provide strong evidence to establish the presence of a pathological condition [49], (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, it is a clinical study that compares the conventional polysomnography (PSG) to the 4-channel respiratory polygraphy. The alternative diagnosis was carried out by medical specialists without any automated signal processing.

In conclusion, our study obtained a single classification model with high diagnostic abilities for all three thresholds. Thus, despite the limitations found for an AHI threshold of 5 e/h, our proposal is useful in a clinical context as a screening process for both discarding the presence of paediatric SAHS (LR= equal or less than 0.1) and detecting paediatric severe SAHS (LR+ higher than 10). Moreover, the statistical power of our results is significantly higher than the majority of the state-of-the-art studies due to the size of our cohort (946 subjects).

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 AML, FNN and the fixed distance method, the use, combination, and comparison of other methodologies could be the subject of future research. Moreover, it would be also 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 OD1 to help in paediatric SAHS diagnosis.

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

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

Declaration of Competing Interest

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

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_EGWA_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.

References


V. Barroso-García, G.C. Gutiérrez-Tobal and L. Kheirandish-Gozal et al./Computer Methods and Programs in Biomedicine 183 (2020) 105083 9


