

Diagnosis of pediatric obstructive sleep apnea: preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients' home

Author names and affiliations: Gonzalo C. Gutiérrez-Tobal*¹, M. Luz Alonso-Álvarez², Daniel Álvarez¹, Félix del Campo^{3,4}, Joaquín Terán-Santos², and Roberto Hornero¹.

¹Biomedical Engineering Group, E.T.S.I. de Telecomunicación, Universidad de Valladolid, Valladolid, Spain.

²Unidad Multidisciplinar de Sueño. CIBER respiratorio. Hospital Universitario de Burgos, Burgos, Spain

³Facultad de Medicina, Universidad de Valladolid, Valladolid, Spain,

⁴Hospital Universitario Río Hortega, Valladolid, Spain.

***Corresponding Author:** Gonzalo C. Gutiérrez-Tobal, MSc, Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática, Universidad de Valladolid, Paseo Belén, 15, 47011, Valladolid, Spain; e-mail: gonzalo.gutierrez@gib.tel.uva.es; Phone number: +34 983423000 ext: 4716.

Abstract

The Obstructive Sleep Apnea Syndrome (OSAS) greatly affects both the health and the quality of life of children. Therefore, an early diagnosis is crucial to avoid their severe consequences. However, the standard diagnostic test (polysomnography, PSG) is time-demanding, complex, and costly. We aim at assessing a new methodology for the pediatric OSAS diagnosis to reduce these drawbacks. Airflow (AF) and oxygen saturation (SpO₂) at-home recordings from 50 children were automatically processed. Information from the spectrum of AF was evaluated, as well as combined with 3% oxygen desaturation index (ODI3) through a logistic regression model. A bootstrap methodology was conducted to validate the results. OSAS significantly increased the spectral content of AF at two abnormal frequency bands below (BW1) and above (BW2) the normal respiratory range. These novel bands are consistent with the occurrence of apneic events and the posterior respiratory overexertion, respectively. The spectral information from BW1 and BW2 showed complementarity both between them and with ODI3. A logistic regression model built with 3 AF spectral features (2 from BW1 and 1 from BW2) and ODI3 achieved (mean and 95% confidence interval): 85.9% sensitivity [64.5-98.7]; 87.4% specificity [70.2-98.6]; 86.3% accuracy [74.9-95.4]; 0.947 area under the receiver-operating characteristics curve [0.826-1]; 88.4% positive predictive value [72.3-98.5]; and 85.8% negative predictive value [65.8-98.5]. The combination of the spectral information from two novel AF bands with the ODI3 from SpO₂ is useful for the diagnosis of OSAS in children.

Keywords

Pediatric obstructive sleep apnea, airflow, oximetry, spectral analysis, at-home assessment

Abbreviations

Acc = accuracy; AF = airflow; AHI = apnea-hypopnea index; AROC = area under the receiver-operating characteristics curve; ECG = electrocardiogram; IQR = interquartile range; LR = logistic regression; *MA* = maximum amplitude of the power spectral density; *mA* = minimum amplitude of the power spectral density; M_{f1} - M_{f4} = first to fourth statistical moments of the power spectral density; NPV = negative predictive value; ODI = oxygen desaturation index; OSAS = obstructive sleep apnea syndrome; PPV = positive predictive value; PSD = power spectral density; PSG = polysomnography; RP = respiratory polygraphy; Se = sensitivity; SLR = stepwise logistic regression; Sp = specificity; SpO₂ = oxygen saturation of the blood.

1. Introduction

Obstructive Sleep Apnea Syndrome (OSAS) is a disorder characterized by recurrent episodes of apnea (complete absence of airflow) and hypopnea (significant reduction of airflow) during sleep [1]. Apneic events lead to oxygen desaturations and arousals which prevent patients from resting while sleeping, disrupting both their health and quality of life. OSAS can affect both adults and children. Common symptoms in children include overnight snoring and sleep difficulties [2], which may derive in other daytime symptoms and illnesses such as cognitive and behavioral irregularities, abnormal growth, and cardiovascular risks [3], [4]. Moreover, pediatric OSAS is known to be underdiagnosed [5], and the scientific literature reports up to 6 % of children affected [3]. This indicates the high prevalence of the disease which, in turn, leads to an intensive use of the healthcare services [6].

OSAS in children is diagnosed by means of nocturnal polysomnography (PSG) test, which acts as the “gold standard” [2]. PSG requires recording a wide range of physiological signals from patients overnight, including electroencephalogram (EEG), electrocardiogram (ECG), electromyogram (EMG), electrooculogram (EOG), thoracic and abdominal respiration movements, oxygen saturation (SpO_2), and airflow (AF) [1]. Hence, the necessary acquisition equipment is both complex and costly [6]. OSAS diagnosis is established according to the apnea-hypopnea index (AHI), which estimates the number of apneic events per hour of sleep time. To derive AHI, the physiological recordings need to be examined. Consequently, PSG is also time-consuming [7].

Furthermore, the equipment involved in PSG is often not well tolerated by children [8], interfering with their sleep routine.

To overcome these drawbacks a number of alternatives have been studied. One common approach is the use of a reduced set of signals from PSG to compute different estimations of AHI. In this regard, the respiratory disturbance index obtained from respiratory polygraphy (RP) was successfully assessed in an in-lab study with children involving 6 signals [9]: SpO₂, AF, heart rate, chest movements, body position, and snoring. The oxygen desaturation index (ODI), in combination with common symptoms, has been also recently evaluated as an alternative to PSG in pediatric patients [10]. On the other hand, the automatic analysis of physiological signals has been also proposed. In this sense, features from photoplethysmography time series have shown their usefulness in OSAS detection in children [11]. Moreover, studies conducting an automatic processing of the SpO₂ and ECG signals have been successfully performed in the context of adult and pediatric OSAS [12]-[16].

In this paper, a new method for OSAS diagnosis in children is assessed. Our methodology is based on the only use of spectral data from single-channel AF and the 3% ODI (ODI3), both of them obtained at patient's home. The main objective is to evaluate the diagnostic usefulness of eventual differences in the AF spectrum of OSAS patients (OSAS-positive) and no-OSAS subjects (OSAS-negative) in combination with ODI3. As stated above, ODI3 is a commonly used parameter in OSAS studies. Moreover, the study of AF is a straightforward choice since apneas and hypopneas are defined on the basis of

its amplitude variations [17]. Additionally, the recurrence of apneic events naturally leads to the study of AF in the frequency domain. Recent works have shown that OSAS modifies the spectral content of AF recordings from adults at certain frequencies, and that the information contained in such frequencies is useful in OSAS detection [18], [19]. However, no studies have been found applying a similar analysis to AF recordings from children. According to the above mentioned, we pose the following research questions:

- i. How does OSAS modify the spectral information of airflow recordings from children?*
- ii. Are these changes useful to distinguish OSAS in children from at-home recordings?*
- iii. Is the airflow spectral information complementary to the classic oxygen desaturation index in pediatric OSAS detection?*

To answer them, we conduct an exploratory analysis of the power spectral density (PSD) of the AF recordings. We look for spectral bands of interest showing differences in OSAS-positive and OSAS-negative subjects, as well as their characterization. The single diagnostic performance of both the AF spectral features and the ODI3 are assessed. We also evaluate their usefulness and complementarity through logistic regression models. Our hypothesis is that the joint use of spectral information contained in single-channel AF and ODI3 could be useful to diagnose OSAS in children.

2. Methods and Materials

2.1. Subjects and Signals under Study

This study involved AF and SpO₂ recordings from 50 children ranging 3 to 13 years old (24 OSAS-negative and 26 OSAS-positive). All of them were referred to the unit of respiratory sleep disorders of the University Hospital of Burgos (Spain), due to clinical suspicion of OSAS (snoring and/or witnessed breathing pauses). Those children suffering from serious chronic medical or psychiatric co-morbidities, those who required urgent treatment, and those with symptoms suggestive of sleep disorders other than OSAS (e.g., parasomnias, narcolepsy, or periodic leg movements), were excluded. AF and SpO₂ were acquired during a polygraphy test performed at patients' home through an eXim Apnea polygraph (Bitmed®, Sibel S.A., Barcelona, Spain). The sensor used to obtain AF was a thermistor and the sample rate was 100 Hz. SpO₂ was recorded through an oximeter at the same sample rate. The physicians used the AHI derived from PSG to establish OSAS. For the overnight PSG, the Deltamed Coherence ® 3NT Polysomnograph, version 3.0 system (Diagniscan, S.A. ACH – Werfen Company; Paris, France) was used, recording EEG, right and left EOG, tibial and submental (leg and chin) EMG, ECG, AF by thermistor and nasal cannula, chest-abdomen movements with bands, body position, SpO₂ (Nellcor Puritan Bennett – NPB- 290®), snoring, and a continuous transcutaneous recording of carbon dioxide (PtcCO₂). The American Academy of Sleep Medicine (AASM) criteria were used to evaluate sleep states and respiratory events [17]. The median time between PSG and RP was 14 days ([6, 25], interquartile range, IQR). Apneas were scored after complete cessation of AF, as defined by the American Academy of Sleep Medicine [17].

Hypopneas were defined after a 50% reduction of AF accompanied by a 3% decrease in SpO₂ [17]. Amplitude cessations and reductions of AF required lasting 2 missed cycles in order to be considered as apneas or hypopneas, respectively [17]. An obstructive AHI threshold of 3 events/hour was used to distinguish OSAS-positive from OSAS-negative subjects [20]. ODI3 was estimated as the number of desaturations (at least 3%) per hour of recording. The interruption of the oronasal flow secondary to movements was not accounted for either the PSG or the RP. An uninterpretable AF signal was defined as no AF during 30 seconds of normal respiration, while respiratory motion signals and SpO₂ remained unchanged. Data were excluded from analysis if >60% of the AF was uninterpretable. The Ethics Committee of the University Hospital of Burgos accepted the protocol (approval #CEIC 936) and an informed consent was obtained for each subject. Table I summarizes clinical and demographical data from the subjects under study. No statistical significant differences in body mass index or age were found between groups (p -value $\gg 0.01$).

TABLE I. DEMOGRAPHIC AND CLINICAL DATA

Features	All	OSAS-positive	OSAS-negative
# Subjects	50	26	24
Age* (years)	5.3 ± 2.5	5.4 ± 2.7	5.2 ± 2.4
Male (%)	54.0	61.5	45.8
BMI ⁺ (kg/m ²)	16.5 ± 2.5	16.9 ± 3.0	16.1 ± 1.7
Recording Time (h)	8.9 ± 0.8	8.8 ± 1.0	9.0 ± 0.5
AHI (e/h)	9.9 ± 13.8	17.9 ± 15.4	1.3 ± 0.8

BMI: Body Mass Index; AHI: Apnea Hypopnea Index; * p -value = 0.76; + p -value = 0.94

2.2. Power spectral density of airflow

We computed the PSD of each AF recording to explore eventual differences between the spectral information of OSAS-positive and OSAS-negative groups. The estimation of the

PSDs was carried out by the non-parametric Welch method, which is suitable for non-stationary signals [21]. A Hamming window of 2^{15} samples (50% overlap) along with a discrete Fourier transform of 2^{16} samples were used. To avoid the influence of non-physiological factors, each PSD was normalized (PSDn) by dividing all their spectral components by their corresponding total power [22]. Thus, the amplitude values of the PSDns, as measured in 1/Hz, reflect the occurrence of AF events at each frequency.

In order to define the spectral bands of interest, we looked for statistical significant differences between PSDns from OSAS-positive and OSAS-negative groups. Data were not normally distributed. Hence, we used a p -value based methodology consisting in applying the non-parametric Mann-Whitney U test to the amplitude values of the PSDns from both groups, at each frequency [19]. Fig. 1 shows the median values of the PSDns from OSAS-positive (black line) and OSAS-negative (grey line) samples. It also shows the p -values obtained in the comparison of both groups (light grey line). We found marked drops in the p -values around [0.06-0.2] Hz., [0.35-0.43] Hz., and [0.7-1] Hz. However, in order to avoid type I errors, we only defined as bands of interest those spectral bands in which the p -value were lower than 0.01. Thus, two bands were finally defined: 0.119-0.192 Hz (BW1); 0.784-0.890 Hz (BW2). At each band, we let 10% of components have a p -value above 0.01 to maintain coherence with the p -value tendency showed in Fig. 1. This avoids the disaggregation of one single homogeneous band into several due to spurious values.

We characterized these two bands by extracting six common spectral features from each of them:

- Maximum and minimum amplitude (MA , mA), computed as the highest and the lowest PSDn values in each band. These features measure the maximum and minimum occurrence of AF events at the bands.

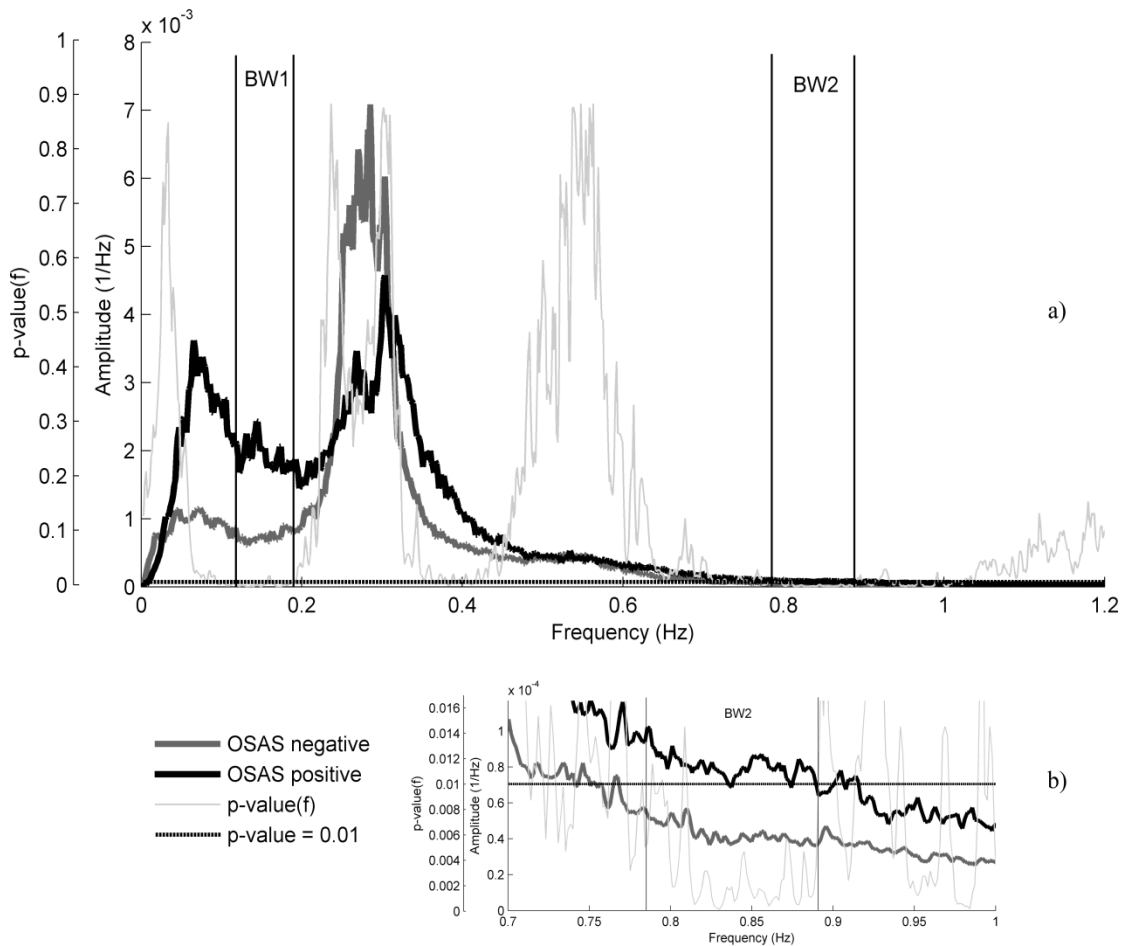


Figure 1. Median values of the PSDns from OSAS-positive (black line) and OSAS-negative (grey line) samples, and p -values at each frequency (light grey line). Significance level p -value = 0.01 (black dashed line). In Fig. 1 a, the spectral bands of interest BW1 and BW2 are delimited outside the normal respiratory rate. In Fig. 1 b, BW2 has been expanded for a better viewing.

- First to fourth statistical moments (M_{f1} - M_{f4}). Mean (M_{f1}), standard deviation (M_{f2}), skewness (M_{f3}), and kurtosis (M_{f4}), quantify central tendency, dispersion, asymmetry, and peakedness of the spectral data, respectively.

2.3. Logistic Regression: feature selection and classification

The logistic regression (LR) method is a supervised learning algorithm which estimates the posterior probability of a given instance \mathbf{x}_i (in our case, a vector containing the extracted features) belonging certain class C_k (in our case, $C_k = \text{OSAS-positive}$ or OSAS-negative). Hence, the posterior probability $p(C_k | \mathbf{x}_i)$, i.e. the probability of a subject belonging to OSAS-positive or OSAS-negative group, is computed through the logistic function:

$$p(C_k | \mathbf{x}_i) = \frac{e^{\beta_0 + \boldsymbol{\beta}^T \mathbf{x}_i}}{1 + e^{\beta_0 + \boldsymbol{\beta}^T \mathbf{x}_i}}, \quad (1)$$

where β_0 and $\boldsymbol{\beta}$ are obtained by the weighted least squares minimization procedure [23].

Thus, an instance \mathbf{x}_i is assigned to the class with larger posterior probability.

First, we used LR to automatically select those relevant and non-redundant features. This was performed through the stepwise LR method (SLR), proposed by Hosmer and Lemeshow [23]. Specifically, we applied the well-known forward-selection backward-elimination algorithm. Then, LR was also used to assess the joint potentiality of the selected features from BW1 and BW2 to predict OSAS in children.

2.4. Statistical analysis

Data did not pass the Lilliefors normality test. Hence, the non-parametric Mann-Whitney U test was used to evaluate statistical differences in the obtained features from OSAS-positive and OSAS-negative groups. Sensitivity (Se, percentage of OSAS-positive subjects rightly classified), specificity (Sp, percentage of OSAS-negative subjects rightly classified), accuracy (Acc, overall percentage of subjects rightly classified), positive predictive value (PPV, proportion of positive test results which are true positives), and negative predictive value (NPV, proportion of negative test results which are true negatives) were used to measure the diagnostic ability of both single features and LR models. In the assessment of single features, a receiver operating-characteristics (ROC) analysis was conducted. Thus, the area under the ROC (AROC) was also computed for each case.

2.5. Results validation: bootstrap 0.632

We used the bootstrap 0.632 algorithm to validate our results since it is particularly useful to estimate statistics in small-size samples [24]. Given a sample of N instances, this method proposes building B new samples (bootstrap samples) of size N by resampling with replacement from the original one [24]. A uniform probability is used to randomly select the instances for each B . Thus, the instances can be selected several times for a particular sample B_i , which acts as a training group and, most probably, will contain repeated instances [24]. Consequently, for each new sample, a number of instances from the original are not selected. These instances act as the test group.

The number of subjects in our database is $N=50$ and the number of bootstrap samples chosen was $B=1000$, since it ensures a proper estimation of the 95% confidence interval [25] (CI). Thus, 1000 new groups of size 50 were built, acting as training groups. As stated above, the instances not included in each case acted as the corresponding test groups. Following bootstrap 0.632, a statistic s obtained from a test set would be a downward estimation of the true one [25]. Hence, both the training and the test groups are used to compute s by weighting their corresponding estimations as follows [24]:

$$s = 0.632 \cdot s_{test} + 0.368 \cdot s_{training} . \quad (2)$$

Finally, the B estimations of s are averaged to show a global performance.

3. Results

3.1. Descriptive analysis and feature selection

Table II summarizes the values (median and IQR) of each spectral feature. Consistent with Fig. 1, the values of MA , mA , and Mf_1 in BW1 and BW2 were significantly higher in OSAS-positive than in OSAS-negative subjects (p -value < 0.01). Near to significant differences were found in Mf_2 from both bands, and there were no differences in Mf_3 and Mf_4 . As expected, ODI3 also showed statistical differences between groups (OSAS-negative: 0.87 e/h IQR [0.44, 1.9], OSAS-positive: 5.9 e/h IQR [1.8, 9.1], p -value < 0.01).

SLR was used twice to select relevant and non-redundant features. First, we applied SLR to the 12 spectral features previously obtained. Thus, mA from BW1, and Mf_3 and Mf_4 from BW2 were automatically selected by SLR to form the corresponding model

(SLR_{Spec}). Second, the selection process was repeated with the 13 features, i.e., including ODI3. In this case, ODI3, mA and Mf_4 from BW1, as well as Mf_3 from BW2 were selected for the model ($SLR_{Spec-ODI3}$).

TABLE II. FEATURE VALUES FOR OSAS-POSITIVE AND OSAS-NEGATIVE GROUPS

	BW1			BW2		
	median [IQR]		p	median [IQR]		p
	OSAS-positive	OSAS-negative		OSAS-positive	OSAS-negative	
MA (10^{-3})	3.0 [1.3, 4.4]	1.3 [0.8, 2.4]	<0.01	0.11 [0.08, 0.28]	0.06 [0.05, 0.12]	<0.01
mA (10^{-4})	13.0 [7.0, 18.0]	5.1 [4.0, 6.7]	<<0.01	0.5 [0.4, 1]	0.3 [0.2, 0.5]	<<0.01
$Mf1$ (10^{-3})	1.9 [0.9, 2.7]	0.9 [0.5, 1.4]	<0.01	0.08 [0.06, 0.14]	0.04 [0.03, 0.07]	<0.01
$Mf2$ (10^{-4})	3.5 [1.4, 6.8]	1.5 [0.8, 3.9]	0.045	0.16 [0.11, 0.31]	0.09 [0.06, 0.19]	0.029
$Mf3$ (10^{-1})	5.9 [2.9, 8.7]	5.4 [3.5, 9.1]	0.993	3.9 [1.8, 6.6]	4.3 [1.9, 9.3]	0.541
$Mf4$ (10^0)	2.7 [2.3, 3.3]	2.7 [2.5, 3.3]	0.749	2.7 [2.4, 3.3]	2.5 [2.3, 3.1]	0.356

p : p -value of the Mann-Whitney U test; IQR: interquartile range.

3.2. Diagnostic performance

Table III shows the diagnostic performance of the spectral features and ODI3 after the bootstrap 0.632 procedure. Se, Sp, Acc, PPV, and NPV values (mean and 95% CI) were obtained by weighting the training and test estimations according to bootstrap 0.632, and averaging the results from the 1000 training-test group pairs. The best single feature in terms of Acc and AROC was the spectral mA from BW1 (76.3% [65.7-84.2]; 0.743 [0.584-0.871], respectively), outperforming ODI3 from oximeter (75.3% [67.0-83.4]; 0.676 [0.513-0.829]).

Table IV includes the diagnostic performance of SLR_{Spec} and $SLR_{Spec-ODI3}$. SLR_{Spec} , which only used spectral information from AF, outperformed all the single features in terms of Acc and AROC (79.1% [68.6-87.9]; 0.875 [0.723-1]). The $SLR_{Spec-ODI3}$ model, which combines spectral information from AF with ODI3 from SpO_2 , obtained the

highest results at each statistic (85.9% Se [64.5-98.7]; 87.4% Sp [70.2-98.6]; 86.3% Acc [74.9-95.4]; 0.947 AROC [0.826-1]; 88.4% PPV [72.3-98.5]; 85.8% NPV [65.8-98.5]).

TABLE III. DIAGNOSTIC PERFORMANCE OF THE SINGLE FEATURES

<i>BW1</i>	Se (%) [CI]	Sp (%) [CI]	Acc (%) [CI]	PPV (%) [CI]	NPV (%) [CI]	AROC [CI]
<i>MA</i>	60.2 [37.0,87.8]	71.7 [34.5,95.8]	65.6 [55.4,74.2]	71.7 [53.6,94.9]	62.9 [49.1,80.3]	0.651 [0.450,0.805]
<i>mA</i>	71.9 [47.1,87.4]	81.1 [65.3,100.0]	76.3 [65.7,84.2]	80.9 [62.3,100.0]	73.2 [58.7,87.3]	0.743 [0.584,0.871]
<i>Mf₁</i>	66.4 [40.5,85.9]	72.3 [50.2,95.1]	69.1 [54.7,77.3]	72.9 [57.9,91.6]	67.3 [53.2,83.6]	0.684 [0.515,0.825]
<i>Mf₂</i>	59.4 [32.0,85.3]	67.0 [31.9,91.6]	62.9 [52.4,72.7]	67.1 [51.4,88.7]	60.8 [46.0,77.2]	0.603 [0.416,0.769]
<i>Mf₃</i>	53.8 [29.2,78.7]	57.0 [30.8,79.2]	54.0 [43.4,67.7]	57.5 [40.1,75.5]	53.6 [38.5,72.1]	0.542 [0.422,0.675]
<i>Mf₄</i>	52.1 [27.3,77.9]	56.3 [29.3,80.8]	60.2 [42.2,67.9]	57.8 [40.5,77.9]	50.7 [25.8,71.6]	0.539 [0.372,0.698]
<i>BW2</i>						
<i>MA</i>	74.6 [53.1,95.6]	64.4 [43.8,82.7]	70.6 [60.5,79.2]	70.1 [57.6,83.2]	73.2 [55.2,93.9]	0.670 [0.498,0.809]
<i>mA</i>	83.4 [48.6,98.7]	65.6 [51.2,85.8]	74.8 [61.1,83.6]	72.3 [60.5,84.4]	80.8 [56.7,98.0]	0.730 [0.576,0.859]
<i>Mf₁</i>	79.7 [40.4,96.4]	65.2 [47.8,84.4]	72.7 [59.6,81.3]	71.2 [58.8,83.5]	76.6 [55.2,94.7]	0.698 [0.527,0.837]
<i>Mf₂</i>	74.6 [42.1,92.2]	64.2 [41.7,83.6]	69.5 [57.6,79.1]	69.3 [56.1,82.8]	71.1 [53.0,89.9]	0.627 [0.449,0.784]
<i>Mf₃</i>	52.4 [27.6,78.2]	52.1 [25.8,78.7]	52.4 [39.1,67.1]	53.4 [30.2,73.2]	50.9 [32.7,71.8]	0.557 [0.422,0.706]
<i>Mf₄</i>	60.3 [33.5,83.4]	58.9 [34.6,80.1]	59.6 [48.1,70.4]	61.5 [47.4,78.0]	58.1 [42.0,74.2]	0.574 [0.436,0.712]
<i>ODI3</i>	70.9 [49.5,94.6]	80.3 [46.5,100.0]	75.3 [67.0,83.4]	81.9 [62.9,100.0]	72.8 [37.0,87.8]	0.676 [0.513,0.829]

CI: 95% confidence interval.

TABLE IV. DIAGNOSTIC PERFORMANCE OF THE LOGISTIC REGRESSION MODELS

	Se (%) [CI]	Sp (%) [CI]	Acc (%) [CI]	PPV (%) [CI]	NPV (%) [CI]	AROC [CI]
<i>SLR_{Spec}</i>	79.2 [59.1,96.6]	79.4 [59.3,95.8]	79.1 [69.6,87.9]	81.2 [65.2,94.5]	78.8 [60.4,95.4]	0.875 [0.723,1]
<i>SLR_{Spec-ODI3}</i>	85.9 [64.5,98.7]	87.4 [70.2,98.6]	86.3 [74.9,95.4]	88.4 [72.3,98.5]	85.8 [65.8,98.5]	0.947 [0.826,1]

4. Discussion

In this paper, an alternative diagnostic methodology for OSAS in children was developed by combining spectral information from AF with the classic ODI3 from SpO₂. Our proposal was assessed by answering three research questions.

- i. How does OSAS modify the spectral information of airflow recordings from children?*

We found that the spectral power of AF was significantly higher in OSAS-positive subjects at novel frequency bands below (BW1) and above (BW2) the typical respiratory range in children reported in previous studies (0.220-0.430 Hz) [3], [26], [27]. The relationship of BW1 with apneas and hypopneas can be explained on the basis of the definition of these apneic events in children. As stated in section 2.1, apneas and hypopneas require at least 2 missed breaths of length in order to be scored [17]. Missing 2 cycles means that the recurrence of these apneic events is every 2 normal breaths, at most. Therefore, their frequency has to be located below the half of the normal respiratory frequency range, modifying the spectrum of AF in such band. Since BW1 is located below the half of the normal respiratory band, it is consistent with the occurrence of apneas and hypopneas. On the other hand, differences in the high frequency band, BW2 (0.784-0.890 Hz.), may be explained as the typical respiratory overexertion after an apneic event, which increases the respiratory rate [9]. Moreover, the greater variability in the PSDn of OSAS-positive children in the range 0.35-0.43 Hz., which is shown in Fig. 1, is consistent with the decrease of the deep sleep stage time of these patients reported in other works [28]. During deep sleeping, respiration becomes more regular [29], which

leads to a condensed normal breathing band in the PSDn. OSAS interrupts the sleep cycle by the recurrence of arousals [1], causing respiratory instabilities [29] and, consequently, a more variable normal breathing rate.

ii. Are these changes useful to distinguish OSAS in children from at-home recordings?

Seven out of the 13 extracted features were significantly different in OSAS-positive than in OSAS-negative subjects, (6 out of 12 from AF, and ODI3). In the diagnostic ability assessment, mA from BW1 outperformed ODI3, whereas mA from BW2 performed similarly. Both SLR_{Spec} and $SLR_{Spec-ODI3}$ outperformed all the single features. Particularly high was the diagnostic ability of $SLR_{Spec-ODI3}$, which widely improved the performance of an in-lab 6-channel RP (74.2% Se, 81.8% Sp, 77.4% Acc, and 0.852 AROC) [9], only requiring information from 2 channels (thermistor and oximeter) recorded at patients' home. Additionally, SLR_{Spec} (information from single-channel AF only) also outperformed this 6-channel RP.

iii. Is the airflow spectral information complementary to the classic oxygen desaturation index in pediatric OSAS detection?

The study showed complementarity between features in two cases: first, between features from the two novel AF bands, since SLR automatically selected features from both of them to build the SLR_{Spec} and the $SLR_{Spec-ODI3}$ models; second, between features from the two spectral bands BW1-BW2 and the ODI3, since the latter was also selected for the $SLR_{Spec-ODI3}$ model.

Other recent studies also analyzed physiological signals to help in pediatric OSAS diagnosis. Shouldice *et al.* used 50 ECG recordings, and reached 85.7% Se, 81.8% Sp, and 84% Acc in a test set ($AHI \geq 1$), by applying a quadratic linear discriminant to 23 features [15]. Gil *et al.* investigated the diagnostic usefulness of the information contained in 21 PPG time series, reporting 75.0% Se, 85.7% Sp, and 80.0% Acc after a leave-one-out cross-validation procedure ($AHI \geq 5$) [11]. The relationship of high frequency inspiratory sounds (HFIS) to OSAS in children has been evaluated as well [30], [31]. Rembold and Suratt reported data to estimate that 10 HFIS events per hour can be useful to discriminate OSAS in children both for $AHI \geq 1$ (70% Se, 100% Sp, and 76.9% Acc) and $AHI \geq 3$ (61.5% Se, 100% Sp., and 80.8% Acc) [30]. Questionnaires and common symptoms have been also involved in screening tools for OSAS and sleep-disordered breathing. Spruyt and Gozal proposed a severity scale based on the answers of 1,133 children from general population to 6 sleep-related questions [32]. They used a predictive score which reached 59.0% Se, 82.9% Sp, 0.79 AROC, 35.4% PPV, and 92.7% NPV ($AHI \geq 3$). Kadmon *et al.* validated this 6-item questionnaire in a sample of 85 children referred to a pediatric sleep clinic [33], reaching 83.0% Se, 64.0% Sp, 0.65 AROC, 28.0% PPV, and 96% NPV ($AHI \geq 5$). Finally, Chang *et al.* combined symptoms (observable apnea, restless sleep, and mouth breathing) with ODI from 141 children to assess both a logistic regression model and a new discriminative score [10]. The former reached 76.6% of diagnostic accuracy whereas the latter achieved 60.0% Se, 86.0% Sp, 71.6% Acc, 84.0% PPV, and 64.0% NPV ($AHI \geq 5$). Our $SLR_{Spec-ODI3}$ outperformed the reported diagnostic ability in these studies, even though we used recordings obtained

from an unsupervised environment. However, Shouldice *et al* used a more restrictive AHI threshold to differentiate patients from control subjects and Gil *et al.*, as well as Rembold and Suratt, worked with one single channel.

Some limitations have to be addressed in this study. The sample size should be larger to empower the generalization ability of our results. Although the bootstrap 0.632 algorithm is known to provide good estimates from small datasets [24], the assessment of our methodology in a larger sample is a very interesting future target. Additionally, a larger sample would let us define the AF bands of interest through an independent set of subjects. Nonetheless, our bands were consistent with the pathophysiology of the apneic events. A wide sample of subjects would be also useful to optimize the set of selected features. Moreover, since our methodology relies on a classification problem, it only provides information about the presence of OSAS and not about its severity. In this sense, future work focused on estimating the AHI or assessing different AHI thresholds could complement our findings. The only use of a thermistor to record AF is another limitation of the study since the AASM recommends the use of a thermistor to score apneas and a nasal pressure transducer to score hypopneas [17]. However, our approach does not rely on event scoring and, in spite of using thermistor alone, results showed high diagnostic ability. Recent studies have shown high performance when using automatic analysis of single-channel AF from thermistor in adults [18], [19]. Finally, the application of different spectral or non-linear measures, as well as the training of more complex classification models, may be also useful to enhance our methodology.

5. Conclusion

To the best of our knowledge, this is the first time that the spectral information of AF recordings from children is analyzed in the context of OSAS. We showed that OSAS in children significantly modifies the PSDn of AF at two abnormal respiratory bands. Diagnostic ability of single features from these novel bands is similar to that of classic ODI3. Additionally, the information contained in the two bands showed complementarity both between them and with ODI3. Our optimum LR model, built with information from thermistor and oximeter at-home recordings, outperformed the diagnostic ability reported in previous in-lab studies focused on finding new alternatives to standard PSG. These results suggest that the spectral information contained in AF recordings is useful to help in pediatric OSAS and that its combination with ODI3 could be beneficial to diagnose OSAS in children at home.

Acknowledgments

This study was supported by the Proyecto Cero 2011 on Ageing from Fundación General CSIC, the project TEC2011-22987 from Ministerio de Economía y Competitividad, the project VA059U13 from the Consejería de Educación de la Junta de Castilla y León, FEDER and SEPAR. G. C. Gutiérrez-Tobal was in receipt of a PIRTU grant from the Consejería de Educación de la Junta de Castilla y León and the European Social Fund.

References

- 1 Patil SP, Schneider H, Schwartz AR, Smith PL. Adult Obstructive Apnea: Pathophysiology and Diagnosis. *Chest*. 2007;132(1):325-337, doi:10.1378/chest.07-0040.
- 2 Marcus CL, Brooks LJ, Davidson S et al. Diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics*. 2012;130(3): e714-e755, doi: 10.1542/peds.2012-1672.
- 3 Guilleminault C, Lee JH, Chan A. Pediatric obstructive sleep apnea syndrome. *Arch Pediatr Adolesc Med*. 2005;159:775-785, doi: 10.1001/archpedi.159.8.775.
- 4 Kim J, Bhattacharjee R, Kheirandish-Gozal L, Spruyt K, Gozal D. Circulating Microparticles in Children with Sleep Disordered Breathing. *Chest*. 2011;140(2):408-417, doi:10.1378/chest.10-2161.
- 5 Lipton AJ, Gozal D. Treatment of obstructive sleep apnea in children: do we really know how? *Sleep Med Rev*. 2003;7(1):61-80, doi: 10.1053/smr.2001.0256.
- 6 Reuveni H, Simon T, Tal A, Elhayany A, Tarasiuk A. Health care services utilization in children with obstructive sleep apnea syndrome. *Pediatrics*. 2002;110(1):68-72.
- 7 Bennet JA, Kinnear WJM. Sleep on the cheap: the role of overnight oximetry in the diagnosis of sleep apnoea hypopnoea syndrome. *Thorax*. 1999;54:958-959, doi:10.1136/thx.54.11.958.
- 8 Katz ES, Mitchell RB, D'Ambrosio CM. Obstructive sleep apnea in infants. *Am J Respir Crit Care Med*. 2012;185(8):805-816, doi: 10.1164/rccm.201108-1455CI.
- 9 Alonso-Álvarez ML, Terán-Santos J, Cordero-Guevara JA et al. Reliability of respiratory polygraphy for the diagnosis of sleep apnea-hypopnea syndrome in children. *Arch Bronconeumol*. 2008;44:318-323, doi:10.1016/S1579-2129(08)60052-X.
- 10 Chang L, Wu J, Cao L. Combination of symptoms and oxygen desaturation index in predicting childhood obstructive sleep apnea. *Int J Pediatr Otorhi*. 2013;77(3):365-371, doi: 10.1016/j.ijporl.2012.11.028.

- 11 Gil E, Bailón R, Vergara JM, Laguna P. PTT variability for discrimination of sleep apnea related decreases in the amplitude fluctuations of PPG signal in children. *IEEE Trans Biomed Eng.* 2010;57(5):1079-1088, doi: 10.1109/TBME.2009.2037734 .
- 12 Marcos JV, Hornero R, Álvarez D, del Campo F, Aboy M. Automated detection of obstructive sleep apnea syndrome from oxygen saturation recordings using linear discriminant analysis. *Med Biol Eng Comput.* 2010;49(9):895-902, doi:10.1007/s11517-010-0646-6.
- 13 Álvarez D, Hornero R, Marcos JV, et al. Assessment of feature selection and classification approaches to enhance information from overnight oximetry in the context of apnea diagnosis. *Int J Neural Syst.* 2013;23(5):1-18, doi: 10.1142/S0129065713500202.
- 14 Penzel T, McNames J, De Chazal P, Raymond B, Murray A, Moody G. Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. *Med Biol Eng Comput.* 2002;40(4):402-407, doi: 10.1007/BF02345072.
- 15 Shouldice RE, O'Brien LM, O'Brien C, De Chazal P, Gozal D, Heneghan C. Detection of obstructive sleep apnea in pediatric subjects using surface lead electrocardiogram features. *Sleep.* 2004;27(4):784-92.
- 16 Vavrina J. Computer assisted pulse oximetry for detecting children with obstructive sleep apnea syndrome. *Int J Pediatr Otorhi.* 1995;33(3):239-248, doi:10.1016/0165-5876(95)01217-6.
- 17 Iber C, Ancoli-Israel S, Chesson AL, Quan SF. *The AASM Manual for the Scoring of Sleep and Associated Events.* Manual, American Academy of Sleep Medicine; 2007.
- 18 Gutiérrez-Tobal GC, Hornero R, Álvarez D, Marcos JV, del Campo F. Linear and nonlinear analysis of airflow recordings to help in sleep apnoea–hypopnoea syndrome diagnosis. *Physiol Meas.*, 2012;33(7):1261-1275, doi:10.1088/0967-3334/33/7/1261.
- 19 Gutiérrez-Tobal GC, Álvarez D, Marcos JV, del Campo F, Hornero R. Pattern recognition in airflow recordings to assist in the sleep apnoea–hypopnoea syndrome diagnosis. *Med Biol Eng Comput.* 2013;51:1367-1380, doi: 10.1007/s11517-013-1109-7.

- 20 Alonso-Álvarez ML, Canet T, Cubel-Alarco M, Estivill E, Fernandez-Julian E, Gozal D, Jurado-Luqué MJ, Lluch-Roselló A, Martínez-Pérez F, Merino-Andreu M, Pin-Arboledas G, Roure N, Sanmartí F, Sans-Capdevila O, Segarra-Isern F, Tomás-Vila T, Terán-Santos J. Consensus document on sleep apnea-hypopnea syndrome in children. *Arch Bronconeumol*. 2011;47(5): 1-18, doi:10.1016/S0300-2896(11)70026-6.
- 21 Welch PD. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on time Averaging Over Short, Modified Periodograms. *IEEE Trans on Audio and Electroacoustics*, 1967;AU-15:70-73.
- 22 Sörnmo L, Laguna P. *Bioelectrical signal processing in cardiac and neurological applications*. London, U.K./New York: Elsevier/Academic; 2005.
- 23 Hosmer DW, Lemeshow S. *Applied Logistic Regression*. New York, NY, USA: John Wiley and Sons; 2000.
- 24 Witten IH, Frank E, Hall MA. *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, MA, USA: Morgan Kaufmann/Elsevier; 2011.
- 25 Efron B, Tibshirani RJ. *An introduction to the bootstrap*. New York, NY, USA: Chapman and Hall; 1994.
- 26 Tabachnik E, Muller N, Toye B, Levison H. Measurement of ventilation in children using the respiratory inductive plethysmograph. *J Pediatr*. 1981;99(6):895-899, doi: 10.1016/S0022-3476(81)80012-1.
- 27 Fleming S, Thompson M, Stevens R, et al. (2011). Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. *The Lancet*. 2011;377(9770):1011-1018, doi: 10.1016/S0140-6736(10)62226-X.
- 28 Guilleminault C, Korobkin R, Winkle R. A review of 50 children with obstructive sleep apnea syndrome. *Lung*. 1981;159(1):275-287, doi: 10.1007/BF02713925.
- 29 Shea SA. Behavioural and arousal-related influences on breathing in humans. *Exp Physiol*. 1996;81(1):1-26.
- 30 Rembold CM, Suratt PM. Children with obstructive sleep-disordered breathing generates high-frequency inspiratory sounds during sleep. *Sleep*. 2004;27(6):1154-61.

- 31 Rembold CM, Suratt PM. An upper airway resonator model of high-frequency inspiratory sounds in children with sleep-disordered breathing. *J Appl Physiol.* 2004;98(5):1855-61. doi: 10.1152/jappphysiol.01231.2004.
- 32 Spruyt K, Gozal D. Screening of Pediatric Sleep-Disordered Breathing Pediatric Sleep-Disordered Breathing Complaints. A Proposed Unbiased Discriminative Set of Questions Using Clinical Severity Scales. *Chest.* 2012;142(6): 1508-1515, doi:10.1378/chest.11-3164.
- 33 Kadmon G, Shapiro CM, Chung SA, Gozal, D. Validation of a pediatric obstructive sleep apnea screening tool. *Int J Pediatr Otorhi.* 2013;77(9):1461-1464, doi: 10.1016/j.ijporl.2013.06.009.