35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Assessment of Spectral Bands of Interest in Airflow Signal to Assist in Sleep Apnea-Hypopnea Syndrome Diagnosis

Gonzalo C. Gutiérrez-Tobal*, *Student Member*, *IEEE*, Daniel Álvarez, *Member, IEEE*, Javier Gómez, Félix del Campo, and Roberto Hornero, *Senior Member*, *IEEE*

Abstract— **In this work, we analyze power spectral density (PSD) from single-channel airflow (AF) in the context of sleep apnea-hypopnea syndrome (SAHS) diagnosis. PSDs from SAHS-positive and SAHS-negative subjects were compared through Mann-Whitney test to find bands of interest. Thereby, we characterized three spectral bands (BW_1-BW_3) by their relative power (P_{R1}-P_{R3}) and established relationships with apneas and hypopneas. Then, the single and joint diagnostic ability of P_{R1}-P_{R3} was assessed by means of *K*-nearest neighbours (KNN), Fisher's linear discriminant (FLD), and logistic regression (LR) classifiers. The KNN and LR models, obtained from P_{R1}-P_{R3}, showed the best diagnostic ability after a leave-one-out cross-validation procedure. 87.7%-84.2% accuracy and 0.799-0.853 area under receiver operating characteristics curve (AROC) were achieved, respectively. Our results suggest that the bands of interest we defined are related to apneas and hypopneas and, therefore, can be useful in SAHS diagnosis.**

I. Introduction

The sleep apnea-hypopnea syndrome (SAHS) is a chronic disease characterized by recurrent episodes of total absence (apneas) or significant reduction (hypopnoeas) of airflow (AF) during sleep. SAHS is highly prevalent in western countries since up to 5% men and 2% women are affected [1]. It has been usually related to major cardiovascular illnesses [2], occupational accidents [3], and motor-vehicle collisions [4]. Recently, it has been also associated with cancer incidence [5].

Overnight polysomnography (PSG) is the standard test to diagnose SAHS. PSG involves monitoring and recording multiple physiological signals from patients, such as electroencephalogram (EEG), electrocardiogram (ECG), oxygen saturation (SpO_2) or AF. The origin of the signals can be electrical or mechanical and each one can involve one or several channels. The specialists perform an offline inspection of the signals to derive apnea-hypopnea index (AHI), which is used to establish SAHS and its severity. Hence, PSG is complex, costly, and time-consuming [6]. New diagnostic alternatives are needed in order to overcome these drawbacks.

Reduced sets of signals from PSG have been commonly investigated to decrease the complexity of SAHS diagnosis. In this preliminary work, we analyze single-channel AF, acquired with a nasal prong pressure (NPP) sensor. Since AF signal is directly modified by apneas and hypopneas [7], its study is a natural way of dealing with SAHS detection. Moreover, the recurrence of apneic events leads to a frequency analysis of AF. Previous studies successfully performed spectral analysis to investigate SAHS [8], [9], [10]. Thus, the study of spectral bands of interest is proposed.

The main objective is to evaluate how apneas and hypopneas modify the spectral information contained in AF of SAHS patients. Hence, we automatically compare the power spectral density (PSD) of AF recordings from SAHS-positive and SAHS-negative subjects in order to find differences along frequencies. Spectral bands are defined by the observation of these differences. The relative power of the bands is used to characterize them. The relationship between relative power, apneas and hypopneas is evaluated. Finally, we assess the single and joint diagnostic ability of data from the bands through *K*-nearest neighbours (KNN), Fisher's linear discriminant (FLD) and logistic regression (LR) classifiers [11]. We hypothesize that apneas and hypopneas modify the AF PSDs of SAHS patients in certain bands of interest. Hence, the information contained in these bands could be useful to assist in SAHS diagnosis.

II. Subjects and Signals

This study was conducted involving 57 subjects (45 SAHS-positive and 12 SAHS-negative). The AF recordings were acquired during overnight PSG through a NPP. The sample rate was 128 Hz. PSGs were carried out with a polygraph (E-Series, Compumedics) in the sleep unit of the Hospital Universitario Río Hortega (Valladolid, Spain). Physicians scored apneas and hypopneas according to the American Academy of Sleep Medicine (AASM) rules [7]. They established AHI = 10 events per hour (e/h) as the threshold for a positive diagnosis. All subjects were suspected of suffering from SAHS before undergoing PSG due to common symptoms like daytime sleepiness, loud snoring, nocturnal choking and awakenings, and referring apnoeic events. Table I shows some clinical and demographic data from the sample, including AHI, apnea index (AI), and hypopnea index (HI). The Review Board on Human Studies accepted the protocol. All the subjects gave their informed consent to participate in the study.

* This research was supported in part by the Proyecto Cero 2011 on Ageing from Fundación General CSIC, the "Consejería de Educación (Junta de Castilla y León)" under project VA111A11-2, and project TEC2011-22987 from Ministerio de Economía y Competitividad and FEDER. 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 (ESF).

G. C. Gutiérrez-Tobal, D. Álvarez, J. Gómez and R. Hornero, are with the Biomedical Engineering Group of the University of Valladolid, Valladolid, Spain (corresponding author phone: +34 983423000 ext.:4716 e-mail: gonzalo.gutierrez@gib.tel.uva.es)

F. del Campo is with the sleep unit of Hospital Universitario Rio Hortega in Valladolid, Spain (e-mail: fsas@telefonica.net).

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

TABLE I. DEMOGRAPHIC AND CLINICAL DATA

Features	All	SAHS+	SAHS-
# Subjects	57	45	12
Age (years)	51.3±16.3	54.0±15.8	41.5±14.8
Men (%)	71.9	75.5	58.3
BMI[a] (kg/m^2)	30.3±6.7	30.5±6.8	29.2±6.5
Rcording Time (h)	7.26±0.40	7.25±0.41	7.27±0.38
AHI[a] (e/h)	39.3±31.7	48.2±29.9	6.1±2.5
AI[a] (e/h)	21.2±27.9	26.4±29.3	1.9±1.5
HI[a] (e/h)	18.1±13.2	21.8±13.3	4.2±1.6

a. BMI: Body Mass Index;AHI: Apnea-Hypopnoea Index AI: Apnea Index; HI: Hypopnea Index

III. METHODS

A. Power Spectral Density Estimation

The power spectral density (PSD) of the recordings was estimated using the non-parametric Welch method, which is suitable for non-stationary signal analysis [12]. A Hamming window of 2^{15} samples, along with 50% overlap and 2^{16}-point discrete Fourier transform (DFT), was used. Then each PSD was normalized (PSDn) by dividing the amplitude values by its total power. Fig. 1 displays the average of the SAHS-positive and SAHS-negative normalized PSDs. Clear peaks are observed around the normal respiratory rate (0.25 Hz.) [13]. No studies showing a maximum respiratory rate in SAHS patients were found. However, up to 64 breathings per minute were reported involving other respiratory diseases [14]. Thus, we applied a low-pass filter using a cutoff frequency of 1.2 Hz.

B. Spectral Bands of Interest

We compared the PSDs from SAHS-positive and SAHS-negative groups to establish spectral bands of interest. The non-parametric Mann-Whitney test was applied to the PSD amplitude values at each frequency. Following this procedure, we observed 3 frequency ranges presenting statistical significance differences (p-value < 0.01):

- **BW$_1$**: 0.007-0.086 Hz.
- **BW$_2$**: 0.457-0.498 Hz.
- **BW$_3$**: 0.707-0.810 Hz.

Fig. 2 shows the p-value vs. frequency representation. Since respiratory rate at rest is around 0.25 Hz, BW$_1$, BW$_2$, and BW$_3$ correspond to abnormal frequencies. We characterized the three bands by obtaining their relative power (P_{R1}-P_{R3}), which can be computed as follows:

$$P_R = \sum_{f_i = f_1}^{f_N} PSDn(f_i), i = 1,2,\dots,N \qquad (1)$$

where N is the number of points in the band, and f_i each frequency component.

C. Classifiers

1) K-nearest neighbours

KNN is a classifier based on the probability density estimation of the classes [11]. The probability of a class C_i given a pattern $\mathbf{x}_j$ (posterior probability, $p(C_i|\mathbf{x}_j)$) is estimated according to the nearest neighbours of $\mathbf{x}_j$ [11]:

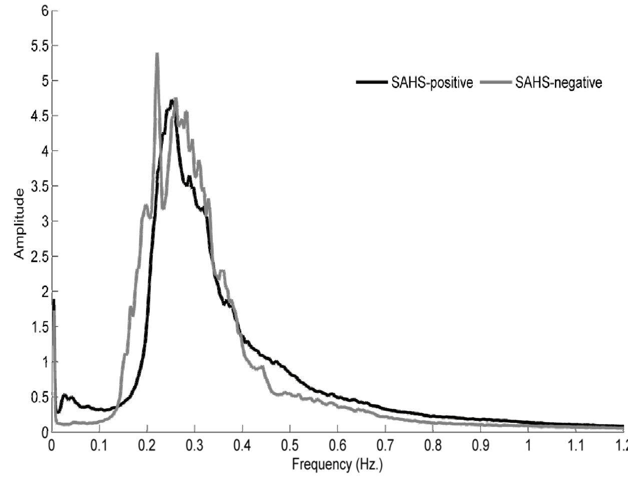

Figure 1. Average of the SAHS-positive and SAHS-negative normalized PSDs

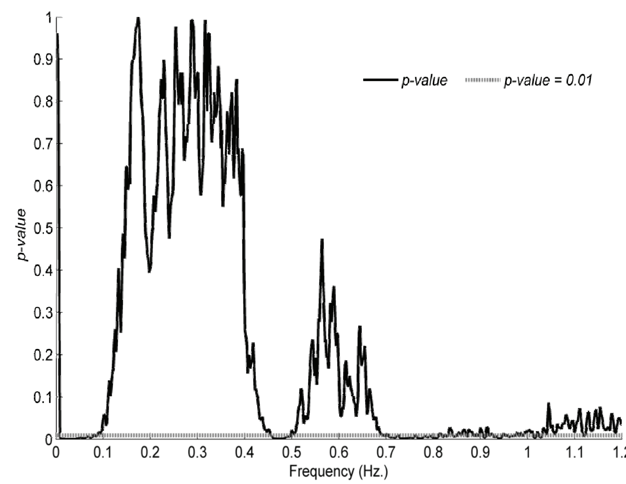

Figure 2. p-value vs. frequency representation

$$p(C_i \mid \mathbf{x}_j) = \frac{K_i}{K}, \qquad (2)$$

where K_i is the number of nearest neighbours belonging C_i and K is the total number of nearest neighbours. Thus $\mathbf{x}_j$ is assigned to the class with the largest posterior probability. K has to be adjusted by the user. Values of K too small could lead to noisy density estimation whereas too large values could equate posterior probabilities with prior probabilities [11]. A bootstrapping procedure was used to set up an appropriate K. Thus, 1000 new samples of size 57 were formed by resampling with replacement the original group [11]. Fig. 3 displays the evolution with K of the average accuracy of these samples. In order to keep the complexity of the model as low as possible, we selected K = 16 since no substantial improvement was observed from this value. Notice that choosing a K value above 23 has no sense since each new pattern would be always assigned to the SAHS-positive group.

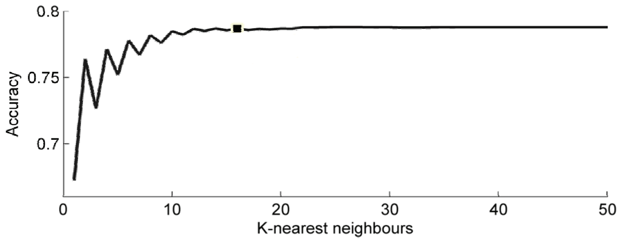

Figure 3. Evolution with K of the average accuracy.

2) *Fisher's linear discriminant*

In a two-class problem, FLD methodology projects data $\mathbf{x}$ into a one dimension space following:

$$y = \mathbf{w}^{\mathbf{T}}\mathbf{x}, \tag{3}$$

where $\mathbf{w}$ is the weight vector which performs the projection, and $\mathbf{x}$ a d-dimensional vector containing data from the classes. The objective is to optimize $\mathbf{w}$ such that the ratio of the between-class variance to the within-class variance is maximized [11]. Thereby, the separation of classes in the projected space is also maximized. Then, the two means of the projected data from classes (m_1, m_2) are computed. A new pattern is assigned to class C_i when its projection is closer to the corresponding m_i.

3) *Logistic regression*

The LR classifier estimates posterior probabilities of a class C_i by the use of the logistic function. This function depends on linear combinations of inputs $\mathbf{x}_j$ [15]:

$$p(C_i \mid \mathbf{x}_j) = \frac{e^{w_0 + \mathbf{w}^{\mathbf{T}}\mathbf{x}_j}}{1 + e^{w_0 + \mathbf{w}^{\mathbf{T}}\mathbf{x}_j}} \tag{4}$$

where w_0 and $\mathbf{w}$ are computed following the maximum likelihood criterion by weighted least squares procedure [15]. As expected, a pattern $\mathbf{x}_j$ is assigned to the class with larger posterior probability.

D. Statistical Analysis

The Spearman's coefficient ρ was used to assess the correlation between P_{R1}-P_{R3} and the apneic events. Additionally, non-linear relationships with AHI, AI, and HI were evaluated by symmetrical uncertainty (*SU*). *SU* is a normalization of mutual information which ranges 0-1 [16]: 1 indicates that knowing one variable it is possible to completely predict the other, whereas 0 indicates that the two variables are independent [17]. We also evaluated the single and joint diagnostic ability of P_{R1}-P_{R3}. Thus, we computed the area under the receiver operating characteristics curve (AROC), sensitivity (Se, proportion of SAHS-positive patients correctly classified), specificity (Sp, proportion of SAHS-negative subjects correctly classified), accuracy (Acc, percentage of subjects correctly classified over the entire sample), 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). These values were obtained following a leave-one-out cross-validation (loo-cv) procedure.

IV. Results

A. Relationship between bands of interest and apneic events

Table II summarizes the results of the correlation (ρ) and *SU* analysis. We observed that the relative power from the very low band, BW_1, was significantly correlated with apneas but it was not with hypopneas. Conversely, the spectral power from BW_2 and BW_3 was significantly correlated with hypopneas but it was not with apneas. *SU* agrees with these results. It shows a higher value for apneas in the case of BW_1, and higher values for hypopneas in the case of BW_2 and BW_3. Hence, monotonic and non-linear relationships were stronger when considering BW_1 and apneas, and the same happened in the evaluation of BW_2, BW_3 and hypopneas.

B. Diagnostic ability assessment

Table III shows the diagnostic ability results for single P_{R1}, P_{R2}, and P_{R3}. It also displays the performance of the models obtained applying KNN, FLD and LR to a matrix containing the three relative powers. All results were achieved following loo-cv. P_{R2} reached the highest AROC (0.789) and P_{R3} the highest Acc (80.7%) when assessing single performance. However, KNN and LR models outperformed single features reaching 0.799-0.853 AROC and 87.7%-84.2% Acc, respectively. FLD achieved the highest values of Sp (83.3%) and PPV (94.3%), but both Acc and AROC were low (73.3% and 0.682).

V. Discussion and Conclusions

In this paper, the spectral information from single-channel AF was evaluated to help in SAHS diagnosis. We used Mann-Whitney test to find three spectral bands showing statistical significant differences between SAHS-positive and SAHS-negative groups. BW_1 was defined under the normal respiratory rate and its relative power was significantly correlated with apneas. BW_2 and BW_3 were above the normal respiratory rate and their relative power was significantly correlated with hypopneas. Since apneas are defined as the absence of AF during 10 seconds or more [7], these events must affect PSD under 0.1 Hz. BW_1, therefore, is consistent with the pathophysiology of apneas. Moreover, previous studies reported signs of an increase in respiratory rate in the presence of hypopneas [10]. This could support our findings in BW_2 and BW_3. However, more research is needed to confirm it. *SU* results agreed with the Spearman analysis of

TABLE II. Relationship between the Bands of Interest and Apneic Events

Features		AHI	AI	HI
P_{R1}	ρ	***0.561****	***0.593****	*0.200*
	SU	0.329	0.382	0.157
P_{R2}	ρ	***0.342****	*0.236*	***0.380****
	SU	0.170	0.134	0.160
P_{R3}	ρ	*0.318*	*0.166*	***0.441****
	SU	0.209	0.164	0.259

*Statistically significant correlation (p-value < 0.01)

TABLE III. DIAGNOSTIC ABILITY THROUGH LEAVE-ONE-OUT CROSS-VALIDATION

	Se(%)	Sp(%)	Acc(%)	PPV(%)	NPV(%)	AROC
P_{R1}	68.9	66.7	68.4	88.6	36.4	0.778
P_{R2}	77.7	66.7	75.4	89.7	44.4	0.789
P_{R3}	84.4	66.7	80.7	90.5	53.3	0.738
KNN	91.1	75.0	87.7	93.2	69.2	0.799
FLD	73.3	83.3	73.7	94.3	45.5	0.682
LR	88.9	66.7	84.2	90.9	61.5	0.853

correlation. The non-linear relationship with apneas was higher for BW_1 whereas the non-linear relationship with hypopneas was higher for BW_2 and BW_3. Nonetheless, the *SU* values were all low.

KNN, FLD and LR models were obtained using the three relative powers from BW_1-BW_3. KNN and LR widely outperformed P_{R1}, P_{R2}, and P_{R3} in terms of accuracy (87.7% and 84.2%). They also achieved higher AROC values (0.799 and 0.854). FLD reached low Acc and AROC, which could be explained by the underlying linearity assumption of the method. These results suggest that BW_1-BW_3 contain complementary and useful information about SAHS. Recent studies focused on automatic detection of SAHS through data from AF [18-21]. Sample sizes ranged between 30 and 200. Se, Sp, PPV and NPV ranged 82.1-97.0%, 76.0-90.0%, 50-82.1% and 83.9-100%, respectively (AHI threshold = 10 e/h). None of them performed a global analysis of AF signal. All were aimed at scoring respiratory events. Although the goal of this preliminary study is not to achieve the highest diagnostic performance but to show the utility of the defined spectral bands, our diagnostic ability results are close to these successful works.

Some limitations have to be addressed. The sample size should be larger to improve the generalizability of results and the number of SAHS-negative subjects should be higher. A large enough sample could allow the use of a training group from which independently deriving the spectral bands of interest and the *K* value for KNN. The use of a *p*-value = 0.01 to define the bands is also a limitation. However, although the limits of the bands could change, results reported in this study are consistent enough to suggest that SAHS modifies spectral information under and above the normal respiratory range. Another limitation is presented when using AHI = 10 e/h as the threshold for SAHS diagnosis. Despite it is a widely used cutoff, others are often considered to establish SAHS and its severity (5, 15, 20, 30, e/h) [18-21]. Our methodology could easily be adapted to these values, and the corresponding results would complete this study.

In summary, after analyzing single-channel AF recordings we defined three spectral bands of interest located into abnormal respiratory ranges. We found significant correlations between their relative powers, apneas, and hypopneas. KNN and LR classifiers applied to these relative powers showed high diagnostic performance. These results suggest that the information included in the defined bands are related to apneas and hypopneas, and can be useful to assist in SAHS diagnosis.

REFERENCES

[1] T. Young, P. E. Peppard and D. J. Gottlieb, "Epidemiology of Obstructive Sleep Apnea: A Population Health Perspective," *Am. J. Respir. Crit. Care. Med.*, vol. 165, pp. 1217-1239, 2002.

[2] F. Lopez-Jiménez, F. H. S. Kuniyoshi, A. Gami, and V. K. Somers, "Obstructive Sleep Apnea," *Chest*, vol. 133(1), pp 793-804, 2008.

[3] E. Lindberg, N. Carter, T. Gislason, and C. Janson, "Role of Snoring and Daytime Sleepiness in Occupational Accidents," *Am. J. Respir. Crit. Care Med.*, vol. 164, pp. 2031-2035, 2001.

[4] A. Sassani, L. J. Findley, M. Kryger, E. Goldlust, C. George, and T. M. Davidson, "Reducing Motor-Vehicle Colissions, Costs, and Fatalities by Treating Obstructive Sleep Apnea Syndrome" *Sleep*, vol. 27(3), pp. 453-458, 2003.

[5] F. Campos-Rodriguez, M. A. Martínez-García, M. Martínez, J. Duran-Cantolla, M. de la Peña, M. J. Masdeu, M. Gonzalez, F. del Campo, I. Gallego, J. M. Marín, F. Barbe, J. M. Montserrat and R. Farre, "Association between obstructive sleep apnea and cancer incidence in a large multicenter spanish cohort," *Am. J. Respir. Crit. Care Med.*, vol. 187, pp. 99-105, 2013.

[6] J. A. Bennet, W. J. M. Kinnear WJM, "Sleep on the cheap: the role of overnight oximetry in the diagnosis of sleep apnoea hypopnoea syndrome," *Thorax* vol. 54, pp. 958-959, 1999.

[7] C. Iber, S. Ancoli-Israel, A. L. Chesson, and S. F. Quan. *The AASM Manual for the Scoring of Sleep and Associated Events*. Manual, American Academy of Sleep Medicine, 2007.

[8] D. Álvarez, R. Hornero, J. V. Marcos, and F. Del Campo, "Multivariate analysis of blood oxygen saturation recordings in obstructive sleep apnea diagnosis," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 12, pp. 2816–2824, 2010.

[9] J. V. Marcos, R. Hornero, D. Álvarez, M. Aboy, and F. Del Campo, "Automated prediction of the apnea-hypopnea index from nocturnal oximetry recordings," *IEEE Trans. Biomed. Eng.*, vol. 59, pp. 141–9, 2012.

[10] G. C. Gutiérrez-Tobal, R. Hornero, D. Álvarez, J. V. Marcos and F. del Campo, "Linear and nonlinear analysis of airflow recordings to help in sleep apnoea-hypopnoea syndrome diagnosis," *Physiol. Meas.*, vol. 33, pp. 1261-75, 2012.

[11] C. M. Bishop, *Pattern Recognition and Machine Learning.* New York, NY: Springer, 2006.

[12] P. D. Welch, "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*, vol. AU-15, pp. 70-73, 1967.

[13] R. Farré, J.M. Montserrat, M. Rotger, E. Ballester and D. Navajas, "Accuracy of thermistors and thermocouples as flow-measuring devices for detecting hypopnoeas," *Eur. Respir. J.*, vol. 11, pp. 179-182, 1998.

[14] K. C. Ong and Y. T. Wang, "Factors associated with improvement in breathing capacity during exercise in patients with chronic obstructive pulmonary disease," *Respirology*, vol. 8, pp. 332-338, 2003.

[15] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*. New York, NY: John Wiley and Sons, 2000.

[16] I. H. Witten, E. Frank and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, MA: Morgan Kaufmann/Elsevier, 2011.

[17] L. Yu and H. Liu, "Efficient feature selection via analysis of relevance and redundancy," *J. Mach. Learn. Res.*,vol. 5, pp. 1205-1224, 2004.

[18] F. R. de Almeida, N. T. Ayas, H. Ueda, P. Hamilton, F. C. Ryan and A. A. Lowe, "Nasal pressure recordings to detect obstructive sleep apnea," *Sleep Breath*, vol. 10, pp. 62-69, 2006.

[19] M. K. Erman, D. Stewart, D. Einhorn, N. Gordon and E. Casal, "Validation of the ApneaLinkTM for the screening of sleep apnea: a novel and simple single-channel recording device," *J. Clin. Sleep. Med.*, vol. 3, pp. 387-392, 2007.

[20] H. Nakano, T. Tanigawa, T. Furukawa and S. Nishima, "Automatic Detection of Sleep-disordered breathing from a single-channel airflow record," *Eur. Respir. J.*, vol. 29, pp. 728-736, 2007.

[21] H. Chen, A. A. Lowe, Y. Bai, P. Hamilton, J. A. Fleetham and F. R. de Almeida, "Evaluation of a portable recording device (ApneaLinkTM) for case selection of obstructive sleep apnea," *Sleep Breath*, vol. 13, pp. 213-219, 2009.