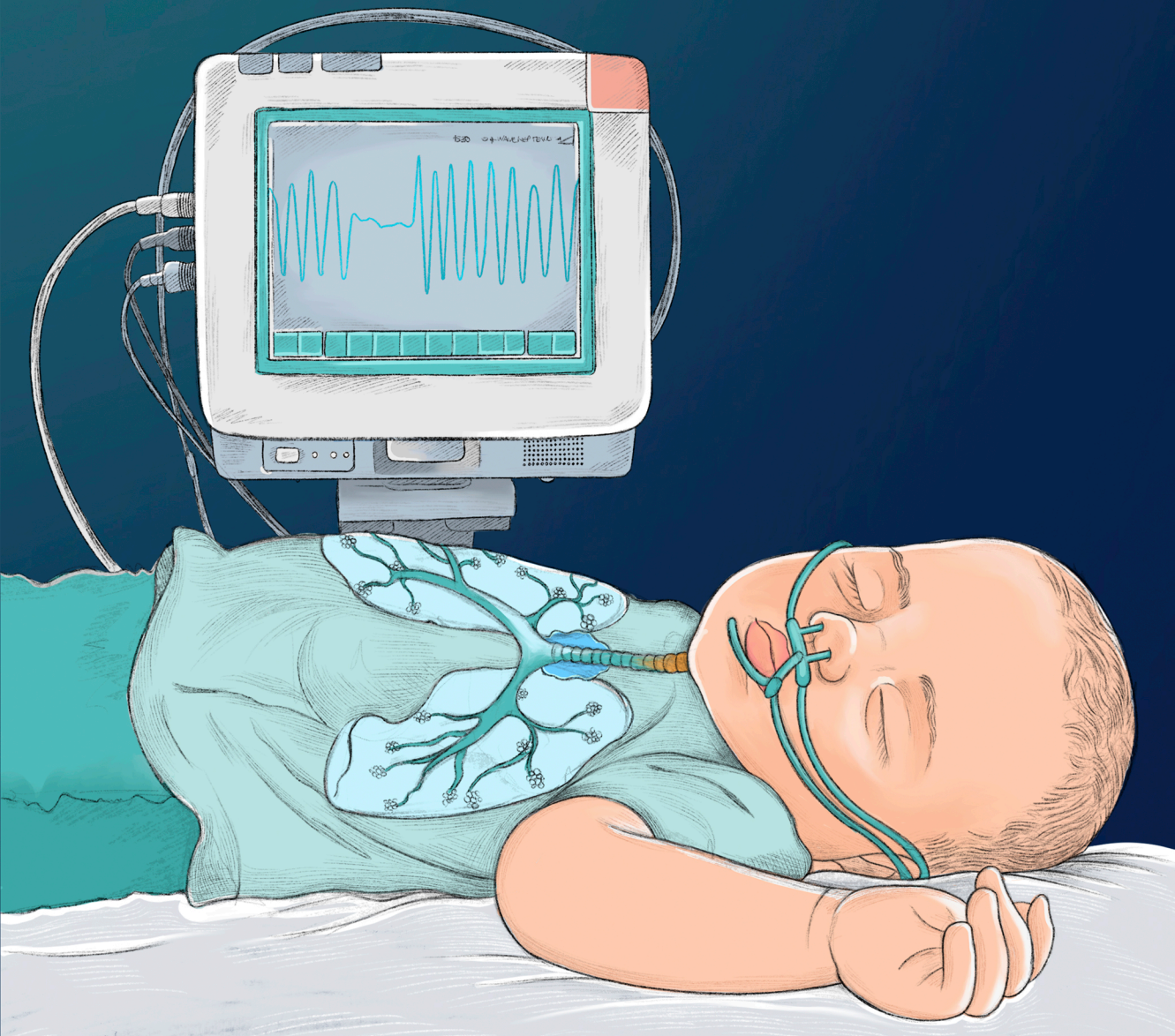
Automatic Analysis of Overnight Airflow to Help in the Diagnosis of Pediatric Obstructive Sleep Apnea



Author: Verónica Barroso-García

Advisors: Roberto Hornero and Gonzalo C. Gutiérrez

DOCTORAL THESIS

Universidad de Valladolid

Doctoral Program of
Information and Telecommunication Technologies

Doctoral Thesis

Automatic Analysis of Overnight Airflow to Help in the Diagnosis of Pediatric Obstructive Sleep Apnea

Thesis presented by **Verónica Barroso García**
to apply for the *Ph.D. degree*
from the *University of Valladolid*

Directed by:
Dr. Roberto Hornero Sánchez and Dr. Gonzalo C. Gutiérrez Tobal

2022
Valladolid, Spain

ACCEPTED: MARCH 2022
ISBN: 978-84-09-39400-5

To Gilda,
my friend, my life, my light.

Defensa

Título	Automatic Analysis of Overnight Airflow to Help in the Diagnosis of Pediatric Obstructive Sleep Apnea
Autor	Verónica Barroso-García
Directores	Dr. Roberto Hornero Sánchez y Dr. Gonzalo C. Gutiérrez-Tobal
Departamento	Teoría de la Señal y Comunicaciones e Ingeniería Telemática

---

Tribunal

Presidente	Dr. D.
Vocal	Dr. D.
Secretario	Dr. D.

acuerda otorgarle la calificación de

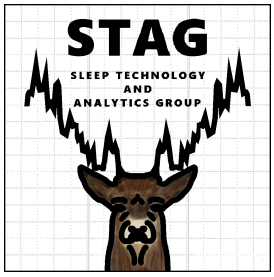
En Valladolid, a de del

Universidad de Valladolid

Escuela Técnica Superior de Ingenieros de Telecomunicación
Dpto. de Teoría de la Señal y Comunicaciones e Ingeniería Telemática

Research Stay for the International Mention

City:	Kuopio (Finland)
Faculty:	University of Eastern Finland
Department:	Applied Physics
Research group:	Sleep Technology and Analytics Research Group
Dates:	01/09/2021–03/12/2021
Duration:	94 days (3 months)
Supervisor:	Prof. Dr. Juha Töyräs



Agradecimientos

En esta etapa he recibido la ayuda, el apoyo y la confianza de muchas personas a las que quiero expresar mi mas sincero agradecimiento.

En primer lugar, al Dr. Roberto Hornero Sánchez y al Dr. Gonzalo C. Gutiérrez Tobal, mis directores de Tesis. Recuerdo esas primeras reuniones en las que me avisasteis reiteradamente de lo arduo que podía llegar a ser realizar un doctorado y dedicarse a la investigación. Hoy, seis años más tarde, puedo corroborar que es un camino difícil y sacrificado, pero que día a día he recorrido con las mismas ganas e ilusión con las que lo inicié. Y no sólo no me arrepiento de haber emprendido este camino, sino que volvería a hacerlo sin dudarlo, ya que vosotros me habéis dado la oportunidad de descubrir y formar parte de un mundo que realmente me apasiona. Gracias por guiarme y apoyarme en mi trayectoria investigadora. Vuestra paciencia, dedicación y experiencia han sido fundamentales para la realización de esta Tesis Doctoral.

Me gustaría mostrar mi agradecimiento a todos los miembros del Grupo de Ingeniería Biomédica de la Universidad de Valladolid, a Jesús P., Carlos G., María G., Javier G., Víctor M., Pablo N., Saúl R., Roberto R., Eduardo S., Víctor R., Aarón M., Víctor G., Marcos R., Sergio P., Selene M., Diego M., y especialmente a aquellos que forman parte de la línea de investigación de 'Apnea', a Daniel Á., Fernando V., Adrián M. y Jorge J. Pasar estos años con vosotros me ha hecho crecer personal y profesionalmente. Cada día he podido aprender algo nuevo, y no sólo de los aspectos técnicos, sino también de vuestra forma de ver la vida. Sin duda, esta Tesis también tiene un trocito de cada uno de vosotros.

También quiero dar las gracias a la Dra. Leila Kheirandish-Gozal y al Dr. David Gozal de la *University of Missouri School of Medicine* (Columbia, USA), así como a la Dra. Andrea Crespo y al Dr. Félix del Campo del Hospital Universitario Río Hortega (Valladolid, España). Sin su estrecha colaboración y su asesoramiento médico no habría sido posible llevar a cabo esta investigación.

I would also like to thank Prof. Dr. Juha Töyräs and the members from the Sleep Technology and Analytics Research Group from the University of Eastern Finland (Kuopio, Finland) for embracing me in one of the most prestigious research groups in the field of automatic obstructive sleep apnea diagnosis. My research stay in Kuopio was an wonderful experience, which has allowed me to grow personal and professionally. I would especially like to thank Timo Leppänen for his advice and help with my study, as well as for making me feel at home.

Finalmente, quisiera expresar mi gratitud a Gilda, por su cariño incondicional y por haber estado tantos años a mi lado. Mi fiel amiga, no hay ni un sólo día en el que no te eche de menos. Y por su puesto, mi más sincero agradecimiento a mi pareja, familia, amigos, y a todas aquellas personas que de una forma u otra me ayudaron a alcanzar esta importante meta.

Gracias por confiar en mí

Abstract

Obstructive sleep apnea (OSA) is a sleep-related breathing disorder characterized by presenting recurrent oronasal airflow cessations. These breathing cessations can be complete (apnea events) or partial (hypopnea events). The presence of these respiratory events can adversely affect the physiological and cognitive functions of children. In this regard, OSA can cause serious long-term neurocognitive deficiencies, behavioral disorders, as well as cardiovascular, metabolic and endocrine dysfunctions, drastically reducing their health and quality of life. Consequently, it is of the utmost importance that children are timely diagnosed and treated to prevent the negative consequences associated to OSA.

Pediatric OSA has a high prevalence, since it affects 5.7% of children between 2 and 18 years of age. According to the primary care clinical data requested from the *Subdirección General de Información Sanitaria Española* in March 2021, it is estimated that this disease affects 9.56‰ of children under 15 years in Spain. Despite its high prevalence, OSA is an underdiagnosed disease, estimating that 90% of affected children still do not have a medical diagnosis. In order to diagnose it, the subjects are referred to a specialized pediatric sleep unit where they undergo nocturnal polysomnography (PSG). This sleep study is based on simultaneously recording several neurophysiological and cardiorespiratory signals while the child sleeps. After, these recordings are visually inspected by sleep medical specialists for manual scoring of respiratory events and computing the apnea-hypopnea index (AHI). In children, this index is used to determine the presence and severity of OSA according to the thresholds 1, 5, and 10 events/h (e/h). PSG is effective, but also uncomfortable to children, complex, time-consuming, and relatively unavailable, which lead to long waiting lists and diagnostic delays. Then, great efforts have been made to search and develop simpler alternative methods that help diagnose pediatric OSA.

In this regard, several studies have focused their research on automatically analyzing a minimum set of cardiorespiratory signals involved in PSG. In this Doctoral Thesis, we propose to exhaustively characterize the behavior of nocturnal airflow (AF) in children to obtain relevant and useful information that helps to simplify the pediatric OSA diagnosis. This signal reflects the respiratory activity during sleep time, including the breathing pauses associated with OSA. In addition, AF can be easily acquired at the patient's home using a portable monitoring device with built-in thermistor. Thereby, AF analysis is a promising way to simplify the diagnosis of childhood OSA. Thus, we hypothesize that *the characterization of overnight AF by means of novel approaches can help and simplify pediatric OSA diagnosis*. Accordingly, the main objective of this Doctoral Thesis is to design, implement, and assess novel automatic signal processing methods that allow exhaustively characterizing the overnight AF from children and helping in the pediatric OSA diagnosis.

In order to achieve this goal, a four-stage methodology is proposed. Firstly, the recordings were subjected to a pre-processing stage to resample them and automatically remove noise and artifacts. Moreover, AF signals were standardized to minimize the effects caused by particular features unrelated to OSA. It would improve the quality of the AF recordings and would increase the effectiveness of subsequent analysis. Afterwards, a feature extraction stage was performed to comprehensively characterize the behavior of pediatric overnight AF by means of different techniques. In this regard, cardiorespiratory signals, and therefore AF, are dynamic, non-linear, and non-stationary. Consequently, non-linear, spectral, bispectral, recurrence plot (RP), and wavelet analyses have been conducted for adapting to the intrinsic properties of overnight AF and getting useful OSA-related information from it. The features derived from each of these methodological approaches could provide redundant information about the AF behavior. Thus, a feature selection stage has been applied to identify those features that provide relevant and complementary information, maximizing the diagnostic ability of AF. In this regard, forward stepwise logistic regression (FSLR) wrapper method and fast correlation-based filter (FCBF) method were used for this purpose. Finally, supervised machine-learning techniques have been applied to recognize patterns in AF features, infer behaviors from them, and use this information to automatically detect the presence and severity of OSA in children. This stage was conducted from three different approaches: discrimination between OSA-negative and OSA-positive pediatric subjects (binary classification task), classification of children according to their OSA severity degree (multiclass classification task), and AHI

estimation of each child (regression task). The binary and multiclass classification tasks were performed by means of logistic regression (LR) and adaptive boosting (AdaBoost.M2) algorithms, respectively. Regarding the regression task, it was performed through a multi-layer perceptron neural network (MLP) and a MLP with Bayesian approach (BY-MLP). In addition, the 3% blood oxygen desaturation index ($ODI3$), a clinical parameter used as a suboptimal alternative to PSG, was incorporated to the study. This allowed us to evaluate its complementarity with the information obtained from AF through the different methodological approaches.

Each of proposed characterization approaches enabled us to uncover behaviors of pediatric nocturnal AF that were previously unknown in OSA context. In this regard, the central tendency measure and spectral entropies showed that this disease reduces the variability and increases the irregularity of pediatric AF. The characterization conducted by means of RP-derived features revealed that OSA modifies the underlying dynamics and the phase-space of AF. Concretely, the occurrence of apneic events decreases the variability, the stationarity, and the complexity of AF signal, as well as the exponential divergence of its phase-space. Moreover, it also increases the dwell time at a certain phase state of AF (i.e., it does not change, or changes very slowly), its average prediction time, and its irregularity. In the case of the bispectral features, they showed that OSA reduces the non-gaussianity of AF, as well as the non-linear interaction of its harmonic components. Childhood OSA also decreases the phase coupling in the normal breathing band, shifting the coupling focus towards low frequency components related to the occurrence apneic events. In addition, the irregularity of AF signal increases in terms of amplitude and phase when the OSA severity is higher. Regarding the wavelet features, they revealed that OSA disturbs the energy distribution and the frequency components of AF signal. Concretely, apneic events reduce the AF detail signal amplitude and the energy produced in the normal breathing band. The frequency components of AF decrease, while its irregularity increases in terms of energy as the AHI is higher.

In addition, it was observed that the information provided by AF through the different methodological approaches is complementary to the information from the classic $ODI3$. This complementarity was not only manifested in the selection stage, but also in the pattern recognition stage. In this regard, moderate-to-high accuracies (Acc) were achieved by the predictive models fed only with AF features: 60.0%–81.1% for 1 e/h, 57.1%–76.0% for 5 e/h, and 70.5%–80.6% for 10 e/h. However, significantly higher diagnostic accuracies were obtained when

AF features and $ODI3$ were combined: 78.0%–83.2% for 1 e/h, 78.5%–82.5% for 5 e/h, and 90.2%–91.0% for 10 e/h. Thereby, RP from AF signal and $ODI3$ achieved the highest Acc for 1 e/h (83.2%). Moreover, this approach obtained the lower negative likelihood ratio (LR– = 0.1), which is considered as a reliable clue to confirm the disease absence when it is ≤ 0.1. In the case of 5 e/h, the highest Acc was achieved by the bispectral analysis of AF and the $ODI3$ (82.5%). Regarding the AHI threshold 10 e/h, both RP and wavelet features from AF obtained 91.0% Acc in combination with the $ODI3$. However, AdaBoost model reached a remarkably higher positive likelihood ratio using wavelet features (LR+ = 19.0), which is considered as a strong inkling to confirm the disease presence when it is ≥ 10.

Based on the aforementioned considerations, the different methodological approaches proposed in this Doctoral Thesis allow adapting to the intrinsic properties of pediatric overnight AF, characterizing its behavior, and providing useful OSA-related information. These approaches enhance the ability of automatic AF analysis to determine the presence and severity of OSA in children. In this regard, the predictive models based on RP, bispectrum, and wavelet features obtained a high overall diagnostic performance along with the $ODI3$, outperforming other state-of-the-art studies and conventional approaches previously applied in adults. Thus, we can conclude that the characterization of overnight AF by means of these novel methods can help to simplify the OSA diagnosis in children. In addition, the high performance of the proposed models suggests that they could be incorporated into clinical practice as reliable automatic screening methods for pediatric OSA.

Acronyms

AASM	American Academy of Sleep Medicine
AB	Adaptive band
Acc	Accuracy
AdaBoost	Adaptive boosting
AF	Airflow
AHI	Apnea-hypopnea index
ANN	Artificial neural network
ANOVA	Analysis of variance
AMI	Auto-mutual information
AUC	Area under receiver-operating characteristics curve
BMI	Body mass index
BY-MLP	Multi-layer perceptron neural network with Bayesian approach
CNN	Convolutional neural networks
CSA	Central sleep apnea
CTM	Central tendency measure
CWT	Continuous wavelet transform
DAP	Decreases in amplitude fluctuations of the photoplethysmography
DFT	Discrete fourier transform
DWT	Discrete wavelet transform
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyograms
EOG	Electrooculogram
FCBF	Fast correlation-based filter
FN	False negatives
FNN	False nearest neighbors

FP	False positives
FSLR	Forward stepwise logistic regression
HRV	Heart rate variability
ICC	Intra-class correlation coefficient
IQR	Interquartile range
JCR	Journal citation reports
Kappa	Cohen's kappa
LASSO	least absolute shrinkage and selection operator
LDA	Linear discriminant analysis
Loo-cv	Leave-one-out cross-validation
LR	Logistic regression
LR+	Positive likelihood ratio
LR–	Negative likelihood ratio
LSTM	Long short-term memory
LZC	Lempel-Ziv complexity
MLP	Multi-layer perceptron neural network
NPV	Negative predictive value
ODI3	3% oxygen desaturation index
ODI4	4% oxygen desaturation index
OSA	Obstructive sleep apnea
PPG	Photoplethysmography
PPV	Positive predictive value
PRV	Pulse rate variability
PSD	Power spectral density
PSG	Polysomnography
PTTV	pulse transit time variability
QDA	Quadratic discriminant analysis
ROC	Receiver-operating characteristics
RP	Recurrence plot
RQA	Recurrence quantification analysis
RRV	Respiratory rate variability
SAHS	Sleep apnea-hypopnea syndrome
Se	Sensitivity
Sp	Specificity
SpO_2	Blood oxygen saturation
SU	Symmetrical uncertainty
SVM	Support vector machine

TN	True negatives
TP	True positives
XAI	Explainable artificial intelligence

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The study presented in this Doctoral Thesis focuses on characterizing overnight airflow (AF) to help determine the presence and severity of obstructive sleep apnea (OSA) in children. In order to carry out this characterization, AF was analyzed using different approaches that allowed adapting to the intrinsic properties of the signal, such as non-linearity and non-stationarity. Thus, each of these approaches enabled us to uncover behaviors of pediatric nocturnal AF that were previously unknown in OSA context. The results obtained in this research were published in four journals indexed in the Journal Citation Reports (JCR) from the Web of Science™, which has allowed presenting this Doctoral Thesis as a compendium of publications.

The thematic consistency of the articles included in this Doctoral Thesis is shown in section 1.1. The general context is briefly described in section 1.2, which presents biomedical engineering, physiological signal processing, and machine-learning study fields. Section 1.3 is devoted to pediatric OSA definition, prevalence, causes, and consequences. The standard method used to diagnose this disease, as well as its current limitations, are detailed in section 1.4. In this regard, simpler diagnostic methods are demanded to deal with these limitations, which motivates the research problem. Finally, section 1.5 provides a exhaustive state-of-the-art revision of previous studies focused on the simplification of pediatric OSA diagnosis.

1.1 Compendium of publications: thematic consistency

OSA is a sleep-related breathing disorder of high prevalence (Blechner and Williamson, 2016; Kaditis et al., 2016). This disease causes numerous adverse medical consequences in children, decreasing their health and quality of life (Blechner and Williamson, 2016; Kaditis et al., 2016). Therefore, it is important to diagnose and treat them in a timely fashion to prevent the worsening and irreversibility of these consequences (Alonso-Álvarez et al., 2011). In this regard, the standard method for pediatric OSA diagnosis, the polysomnography (PSG), is uncomfortable to children, complex, time-consuming, and relatively unavailable, which lead to long waiting lists and delays in diagnosis (Alonso-Álvarez et al., 2015; Tan et al., 2014). Consequently, novel and simpler diagnostic methods are required to deal with these limitations and accelerate OSA diagnosis in children.

In this context, the research carried out during the Doctoral Thesis has focused on exhaustively characterizing the behavior of nocturnal AF in children to obtain relevant and useful information that helps to simplify the diagnosis of pediatric OSA. As can be seen in Figure 1.1, the four papers included in the present compendium of publications share this research focus. Another connection among these articles is the used methodological framework. Thereby, the following four-stage methodology was conducted in each of them: pre-processing, feature extraction, feature selection, and machine-learning.

In order to address its characterization, AF was analyzed using different approaches that allowed adapting to the signal intrinsic properties. In this regard, the first article focused on the analysis of AF and respiratory rate variability (RRV) signals, which was directly obtained from AF, using the central tendency measure (CTM) and spectral entropies (Barroso-García et al., 2017). This approach allowed us to characterize the variability and the irregularity of these respiratory signals in the presence of apneic events. In the second article (Barroso-García et al., 2020), the characterization of AF was carried out by extracting nine features derived from recurrence plots (RP). These features provided novel information about the underlying dynamics and phase-space of AF signal. The third article focused on analyzing AF by means of thirteen bispectral features (Barroso-García et al., 2021a). Thereby, bispectral analysis revealed changes in the non-gaussianity degree and the phase coupling of AF signal, as well as in the non-linear interaction of its harmonic components. Lastly, we characterized AF using discrete wavelet transform (DWT) in the fourth paper (Barroso-García et al., 2021b). In this way,

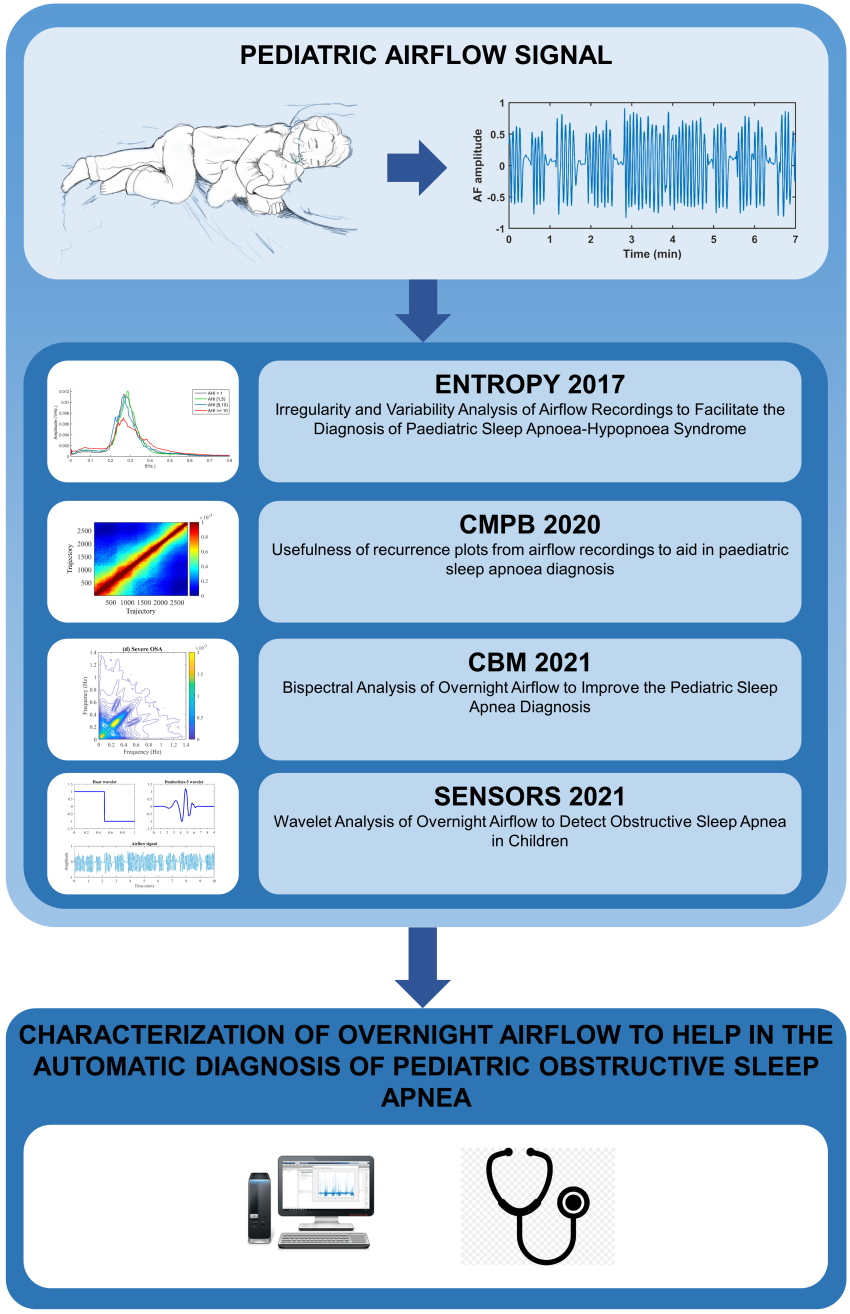


Figure 1.1: Thematic consistency between the publications of this Doctoral Thesis. CMPB:Computer Methods and Programs in Biomedicine, CBM: Computers in Biology and Medicine.

DWT-derived features provided useful information about the energy distribution and frequency components of AF signal. Thus, each of the approaches proposed in this research allowed us to uncover behaviors of pediatric nocturnal AF that were unknown in OSA context so far.

In addition to the characterization of AF signal conducted in these studies, the complementarity between the information obtained from AF and the 3% oxygen desaturation index ($ODI3$) was also evaluated. Therefore, the proposed approaches showed a high diagnostic ability to detect the presence and the severity degree of OSA in children, particularly when these are used along with the $ODI3$.

Titles, authors, and abstracts of the aforementioned articles, as well as the journals in which they were published are shown below. Due to this Doctoral Thesis is presented as a compendium of publications, the complete articles have been included in Appendix A for a suitable understanding of this document.

Irregularity and Variability Analysis of Airflow Recordings to Facilitate the Diagnosis of Paediatric Sleep Apnoea-Hypopnoea Syndrome (Barroso-García et al., 2017).

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, Félix del Campo, David Gozal, and Roberto Hornero. *Entropy*, vol. 19 (9), pp. 447, 2017. Impact factor in 2017: 2.305, Q2 in "PHYSICS, MULTIDISCIPLINARY" (JCR-WOS).

Abstract: The aim of this paper is to evaluate the evolution of irregularity and variability of airflow (AF) signals as sleep apnoea-hypopnoea syndrome (SAHS) severity increases in children. We analyzed 501 AF recordings from children 6.2 $\pm$ 3.4 years old. The respiratory rate variability (RRV) signal, which is obtained from AF, was also estimated. The proposed methodology consisted of three phases: (i) extraction of spectral entropy (SE_1), quadratic spectral entropy (SE_2), cubic spectral entropy (SE_3), and central tendency measure (CTM) to quantify irregularity and variability of AF and RRV, (ii) feature selection with forward stepwise logistic regression (FSLR), and (iii) classification of subjects using logistic regression (LR). SE_1, SE_2, SE_3, and CTM were used to conduct exploratory analyses that showed increasing irregularity and decreasing variability in AF, and increasing variability in RRV as apnoea-hypopnoea index (AHI) was higher. These tendencies were clearer in children with a higher severity degree (from AHI $\geq$ 5 events/hour). Binary LR models achieved 60%, 76%, and 80% accuracy for the AHI cutoff points 1, 5, and 10

e/h, respectively. These results suggest that irregularity and variability measures are able to characterize paediatric SAHS in AF recordings. Hence, the use of these approaches could be helpful in automatically detecting SAHS in children.

Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis (Barroso-García et al., 2020).

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Pablo Núñez, Félix del Campo, David Gozal, and Roberto Hornero. *Computer Methods and Programs in Biomedicine*, vol. 183, pp. 105083, 2020. Impact factor in 2020: 5.428, Q1 in "COMPUTER SCIENCE, THEORY & METHODS" (JCR-WOS).

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

Bispectral Analysis of Overnight Airflow to Improve the Pediatric Sleep Apnea Diagnosis (Barroso-García et al., 2021a).

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, David Gozal, and Roberto Hornero. *Computers in Biology and Medicine*, vol. 129, pp. 104167, 2021. Impact factor in 2020: 4.589, Q1 in "MATHEMATICAL & COMPUTATIONAL BIOLOGY" (JCR-WOS).

Abstract: Pediatric Obstructive Sleep Apnea (OSA) is a respiratory disease whose diagnosis is performed through overnight polysomnography (PSG). Since it is a complex, time-consuming, expensive, and labor-intensive test, simpler alternatives are being intensively sought. In this study, bispectral analysis of overnight airflow (AF) signal is proposed as a potential approach to replace PSG when indicated. Thus, our objective was to characterize AF through bispectrum, and assess its performance to diagnose pediatric OSA. This characterization was conducted using 13 bispectral features from 946 AF signals. The oxygen desaturation index $\geq$3% (*ODI*3), a common clinical measure of OSA severity, was also obtained to evaluate its complementarity to the AF bispectral analysis. The fast correlation-based filter (FCBF) and a multi-layer perceptron (MLP) were used for subsequent automatic feature selection and pattern recognition stages. FCBF selected 3 bispectral features and *ODI*3, which were used to train a MLP model with ability to estimate apnea-hypopnea index (AHI). The model reached 82.16%, 82.49%, and 90.15% accuracies for the common AHI cut-offs 1, 5, and 10 events/h, respectively. The different bispectral approaches used to characterize AF in children provided complementary information. Accordingly, bispectral analysis showed that the occurrence of apneic events decreases the non-gaussianity and non-linear interaction of the AF harmonic components, as well as the regularity of the respiratory patterns. Moreover, the bispectral information from AF also showed complementarity with *ODI*3. Our findings suggest that AF bispectral analysis may serve as a useful tool to simplify the diagnosis of pediatric OSA, particularly for children with moderate-to-severe OSA.

Wavelet Analysis of Overnight Airflow to Detect Obstructive Sleep Apnea in Children (Barroso-García et al., 2021b).

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, David Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, Leila Kheirandish-Gozal, and Roberto Hornero. *Sensors*, vol. 21 (4), pp. 1491, 2021. Impact factor in 2020: 3.576, Q1 in "INSTRUMENTS & INSTRUMENTATION" (JCR-WOS).

Abstract: This study focused on the automatic analysis of the airflow signal (AF) to aid in the diagnosis of pediatric obstructive sleep apnea (OSA). Thus, our aims were: (i) to characterize the overnight AF characteristics using discrete wavelet transform (DWT) approach, (ii) to evaluate its diagnostic utility, and (iii) to assess its complementarity with the 3% oxygen desaturation index ($ODI3$). In order to reach these goals, we analyzed 946 overnight pediatric AF recordings in three stages: (i) DWT-derived feature extraction, (ii) feature selection, and (iii) pattern recognition. AF recordings from OSA patients showed both lower detail coefficients and decreased activity associated with the normal breathing band. Wavelet analysis also revealed that OSA disturbed the frequency and energy distribution of the AF signal, increasing its irregularity. Moreover, the information obtained from the wavelet analysis was complementary to $ODI3$. In this regard, the combination of both wavelet information and $ODI3$ achieved high diagnostic accuracy using the common OSA-positive cutoffs: 77.97%, 81.91%, and 90.99% (AdaBoost.M2), and 81.96%, 82.14%, and 90.69% (Bayesian multi-layer perceptron) for 1, 5, and 10 apneic events/hour, respectively. Hence, these findings suggest that DWT properly characterizes OSA-related severity as embedded in nocturnal AF, and could simplify the diagnosis of pediatric OSA.

1.2 Context: biomedical engineering, physiological signal processing, and machine-learning

The current Doctoral Thesis is framed in the biomedical engineering field. This can be defined as an interdisciplinary branch of science dedicated to understand, modify or control biological systems by applying engineering knowledge and techniques (Bronzino and Peterson, 2014). Thus, one of the main benefits of biomedical engineering is that it can provide novel and effective technical solutions to issues and needs of our healthcare system, such as the necessary technology to monitor

physiological functions of a patient and assist in diagnosis and treatment of pathological conditions (Bronzino and Peterson, 2014). Biomedical engineering covers a wide range of research areas, such as physiological signal and image processing, physiological modelling, biomedical instrumentation, biotechnology, biomechanics, biomaterials or biosensors. Since we have analysed AF signal for the purpose of extracting relevant OSA-related information from it, this Doctoral Thesis is focused on the physiological signal processing area.

Physiological signals are produced by the different biological systems of human body (e.g., cardiovascular, nervous, and respiratory system) and provide information about its behavior (Bronzino and Peterson, 2014). Thus, the study of these signals would allow to detect changes in function of biological systems, as well as identify pathological conditions. However, the information contained in raw biomedical signals is hard to interpret and cannot always be evaluated by visual inspection. Consequently, a processing stage is required to extract and interpret relevant information. In this regard, signal processing enables to automatically extract features of physiological signals by means of mathematical and information theory techniques in order to characterize its behavior (Sörnmo and Laguna, 2005). This fact results in reduced subjectivity and increased reliability. Hence, physiological signal processing has become essential to understand the information obtained from biological systems and develop automatic diagnostic methods.

The present work is also encompassed in the machine-learning field. This is a branch of artificial intelligence focused on developing computer algorithms capable of automatically identifying regularities and complex patterns in data to infer behaviors and make predictions (Alpaydin, 2014; Bishop, 2006). There are different machine-learning approaches, such as supervised, unsupervised, and reinforcement learning. In this regard, supervised learning algorithms have been considered in the Doctoral Thesis since we had labeled data to perform classification and regression tasks (Alpaydin, 2014). Thereby, these methods build a mathematical model from labeled input data (i.e., training data), which is then applied to new inputs (i.e., test data) in order to automatically classify them into different categories (classification) or estimate a continuous variable (regression). Thus, machine-learning methods have been successfully used to develop novel automatic diagnostic methods in recent years.

This Doctoral Thesis aims at exhaustively characterizing the behavior of nocturnal AF in children to obtain relevant and useful information that helps to simplify the diagnosis of pediatric OSA. In order to reach this goal, novel signal processing and machine-learning methods have been studied and applied to

AF signal. Hence, the aforementioned study fields reflect the framework of this research.

1.3 Pediatric obstructive sleep apnea (OSA)

1.3.1 Definition and prevalence

Sleep apnea is a sleep-related respiratory disease characterized by the repeated occurrence of complete and/or partial oronasal airflow cessations (Alonso-Álvarez et al., 2011; Berry et al., 2012). These breathing cessations can be obstructive (obstructive sleep apnea, OSA) or central (central sleep apnea, CSA) (Berry et al., 2012). As shown in the Figure 1.2, OSA is originated by a total or partial occlusion of the upper airways, which blocks the airflow while the respiratory effort persists. Regarding CSA, it results when the brain stops sending stimulus for breathing impulse during certain time intervals, leading to cessation of both oronasal airflow and inspiratory effort movement. The combination of these apneic events results in mixed apneas, i.e., airflow cessations with an absent of respiratory effort (central) during one event portion and presence of inspiratory effort (obstructive) in another portion. Thus, information about the inspiratory effort is essential to be able to distinguish these types of events (Berry et al., 2012).

According to the aforementioned considerations, respiratory cessations can be complete (apneas) or partial (hypopneas) (Alonso-Álvarez et al., 2011; Berry et al., 2012). The American Academy of Sleep Medicine (AASM) defines apnea in children as a total absence or $\geq$90% reduction of airflow for at least 2 breaths, and hypopnea as a reduction between 30% and 90% of airflow for at least 2 breaths that

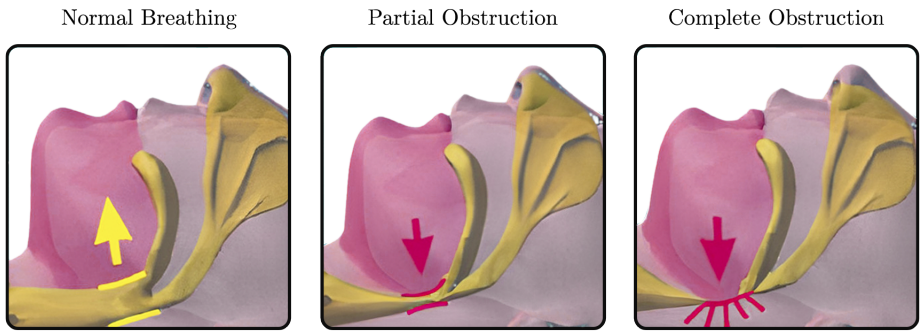


Figure 1.2: Representation of normal breathing, partial, and complete obstruction. Figure modified from Clínica Dental Ceballos (2017).

is accompanied by a ≥3% blood oxygen desaturation or an arousal (Berry et al., 2012). It should be noted that these criteria are different for children and adults, since the minimum duration established by the AASM to score an event as apneic is 2 respiratory cycles for children and 10 seconds for adults (Alonso-Álvarez et al., 2011; Berry et al., 2012).

Currently, pediatric OSA is a common condition that constitutes a major public health problem. Recent epidemiological studies have shown a high prevalence of OSA in children, estimating that this disorder affects about 5.7% of the pediatric population (Lumeng and Chervin, 2008; Marcus et al., 2012). A multicenter study, involving 4,191 pediatric subjects, revealed that approximately 43.3% of children whom are referred to a specialized sleep laboratory for clinical OSA suspicion suffer from mild OSA, 12.8% moderate, and 16.8% severe degree (Hornero et al., 2017). According to the primary care clinical data requested from the *Subdirección General de Información Sanitaria Española* in March 2021, the estimated prevalence of this pediatric disorder in Spain is 9.56‰, affecting 10.83‰ of boys and 8.20‰ of girls under the age of 15. However, despite its high prevalence, OSA is still an underdiagnosed disease (Kheirandish-Gozal, 2010).

1.3.2 Causes, clinical consequences and related pathologies

The soft tissue around the naso and oropharynx relaxes during sleep, which can occlude the airways and cause a sporadic episode of obstructive apnea in healthy children. However, certain anatomical and functional alterations of the upper airways are those that favor its obstruction (Arens and Muzumdar, 2010; Moffa et al., 2020). In this sense, the most common cause of OSA in children is adenotonsillar hypertrophy, i.e., enlarged adenoids and tonsils (Alonso-Álvarez et al., 2011; Moffa et al., 2020). Other risk factors that predispose to suffer from OSA are abnormalities in the maxillofacial or craniofacial anatomy, such as micrognathia, retrognathia, Pierre Robin syndrome or Treacher Collins syndrome (Bitners and Arens, 2020; Moffa et al., 2020). As an adipose tissue increase in the pharyngeal area can cause a narrowing of the airways, obesity can also increase the risk of suffering from this disease (Bitners and Arens, 2020; Moffa et al., 2020). However, unlike what occurs in adults, this narrowing is more associated with adenotonsillar hypertrophy than increased adipose tissue in children (Carroll and Loughlin, 1992; Marcus, 2000). In addition, neurological and neuromuscular disorders, such as cerebral palsy, myotonic dystrophy, or different types of myopathy predispose the appearance of OSA in childhood (Arens and Muzumdar, 2010).

The main clinical manifestations that characterize the presence of OSA are respiratory pauses during sleep, snoring, and daytime hypersomnolence (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). Unlike adults, snoring is usually continuous and excessive daytime sleepiness is often less frequent in children (Carroll and Loughlin, 1992). Pediatric subjects affected by OSA suffer from gas exchange abnormalities, hypoxia, arousals, and fragmented sleep, affecting their physiological and cognitive functions (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). In this regard, this disorder can cause serious consequences in their cardiovascular system, such as increased sympathetic activity, arterial hypertension, pulmonary hypertension, endothelial damage, right heart dysfunction/failure (cor pulmonale), or other alterations in ventricular function (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). Childhood OSA can also cause alterations in their central nervous system, leading to several neurocognitive and behavioral disorders such as inattention, hyperactivity, irritability, emotional lability, depression, neurcognitive deficits, learning difficulties, and behavioral problems (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). Other complications associated with OSA also include endocrine and metabolic changes, enuresis, and somatic growth delay (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). Thus, this disease drastically decreases health and quality of life of affected children. Consequently, a timely diagnosis and treatment is essential to avoid these clinical outcomes.

1.4 Pediatric OSA diagnosis

In order to diagnose OSA, children are referred to a specialized pediatric sleep unit where they undergo a Type 1 sleep study (overnight polysomnography, PSG), which is the gold standard diagnostic test (Jon, 2009). This sleep study is based on simultaneously recording several neurophysiological and cardiorespiratory signals by means of different body sensors while the child sleeps. These signals include body position, electroencephalogram (EEG), electrooculogram (EOG), chin and anterior tibialis electromyograms (EMG), electrocardiogram (ECG), oximetry (SpO_2), oral and nasal airflow (AF), thoracic and abdominal respiratory effort, and snoring sound (Collop et al., 2007; Jon, 2009). These recordings are then visually inspected by sleep medical specialists for manual scoring of apnea and hypopnea events. The number of apneas and hypopneas per hour of sleep (apnea-hypopnea index, AHI) is the most widely used indicator to determine the presence and severity of pediatric OSA (Spruyt, 2012). Regarding the AHI thresholds established to determine the OSA severity degree, 1, 5 and 10 events per hour (e/h) are com-

monly used to differentiate among no-OSA (AHI $<$ 1 e/h), mild OSA (1 e/h $\leq$ AHI $<$ 5 e/h), moderate OSA (5 e/h $\leq$ AHI $<$ 10 e/h), and severe OSA (AHI $\geq$ 10 e/h) in children (Alonso-Álvarez et al., 2011; Hornero et al., 2017; Tan et al., 2014). It should be noted that these criteria are more conservative than in adults, where the AHI thresholds 5, 15 and 30 e/h are used to establish these severity degree (Kapur et al., 2017).

Although PSG is the gold standard test to effectively diagnose childhood OSA, it has several limitations. A high number of biomedical recordings are monitored in PSG, which requires suitable and expensive equipment (Collop et al., 2007; Ryan et al., 1995). Specialized medical personnel should be present throughout the test and then visually inspect the acquired recordings, which makes it complex and intensively laborious (Collop et al., 2007; Ryan et al., 1995). Another limitation is that child should spend a night hospitalized in the pediatric sleep unit, i.e., in a different environment than usual, which can affect sleep development and characteristics (Jon, 2009). Moreover, pediatric subjects should be attached to multiple body sensors, which can be uncomfortable and even distressing for them (Jon, 2009). It should also be noted that not all hospitals have specialized pediatric sleep units or these are overwhelmed by increasing demand. This fact hinders access to PSG, generating long waiting lists and diagnosis delays (Alonso-Álvarez et al., 2015).

All these inherent drawbacks of PSG have led to search and development of alternative methods that simplify and accelerate pediatric OSA diagnosis before the consequences become more severe (Alonso-Álvarez et al., 2015). In this regard, sleep studies performed with portable monitoring equipment have been proposed as the main alternative to PSG (Type 1) (Collop et al., 2007). Among these studies, we can distinguish comprehensive portable polysomnography (Type 2), modified portable sleep apnea testing (Type 3), and continuous recording of one or two biomedical parameters (Type 4) (Flemons et al., 2003; Kapur et al., 2017):

- Type 1. It acquires up to 32 physiological recordings in a hospital environment.
- Type 2. This study consists in recording a minimum of 7 physiological signals: EEG, EOG, EMG, ECG, SpO_2, AF, and respiratory effort.
- Type 3. Between 4 and 7 signals are recorded in this study, including AF and respiratory movement (or alternatively 2 channels of respiratory movements), as well as ECG and SpO_2.

- Type 4. It is the simplest sleep study, which usually monitors AF and/or SpO_2. All sleep studies that do not meet Type 3 criteria are also included in this group.

One of the portable monitoring advantages is that this is less expensive than PSG (Collop et al., 2007). In addition, it incorporates fewer channels and can be performed in the patient's home, resulting less complex and uncomfortable for children (Chiner et al., 2020).

1.5 Simplification of pediatric OSA diagnosis: cardiorespiratory signals

Due to the high prevalence of OSA in children (5.7%) and its relationship with other severe pathologies (hypertension, heart failure, or neurocognitive deficit) (Kaditis et al., 2016; Marcus et al., 2012), the search for diagnostic alternatives to PSG has become a major concern and the motivation of several investigations in recent years. In this regard, multiple alternatives focused on the automatic analysis of a reduced set of PSG-derived signals have been evaluated (Garde et al., 2019; Gil et al., 2010; Gutiérrez-Tobal et al., 2015; Hornero et al., 2017; Jiménez-García et al., 2020; Lazaro et al., 2014), being single-channel approach one of the most used.

Since repeated bradycardia and tachycardia episodes are manifested in the presence of apneas and hypopneas, some studies have addressed the analysis of cardiac signals such as ECG (Pu et al., 2005; Shouldice et al., 2004) or heart rate variability (HRV) (Deng et al., 2006; Martín-Montero et al., 2020, 2021). In this context, Shouldice et al. (2004) followed a conventional approach to analyze the ECG signal, applying temporal and spectral techniques. The study conducted by Pu et al. (2005) based their research on the temporal feature extraction from the ECG that, together with several cardiorespiratory index, allowed them to effectively detect obstructive events. Non-linear and spectral methods were also used by Deng et al. (2006), who showed the usefulness of jointly applying both approaches to the HRV signal. Martín-Montero et al. (2020) used a spectral analysis approach to characterize the classic frequency bands of the HRV signal, as well as to define and characterize novel specific OSA-related bands in children. In a later study (Martín-Montero et al., 2021), the authors analyzed these bands using bispectral techniques, which allowed them to obtain additional information about the particularities that pediatric OSA causes in HRV.

Others, such as those conducted by Gil et al. (2009, 2010), Lazaro et al. (2014), and Dehkordi et al. (2016), used the photoplethysmography (PPG) to help in childhood OSA diagnosis. Based on event detection, Gil et al. (2009) obtained information about the spectral behavior of HRV during the amplitude fluctuation decreases of the PPG signal. In another study carried out by the authors (Gil et al., 2010), this approach was used to analyze the pulse transit time variability (PTTV) through the temporal evolution of spectral features, which was particularly useful for discriminating OSA-related events in children. Conventional spectral methods also used by Lazaro et al. (2014) to analyze the pulse rate variability (PRV) signal, which is directly derived from PPG. Another study based on characterizing the PRV signal is the one conducted by Dehkordi et al. (2016). They applied temporal, spectral, and detrended fluctuation methods, and showed the utility of combining these approaches to diagnose pediatric OSA from the PPG signal.

The usefulness of automatic analysis of SpO_2 recordings has also been widely assessed in recent years (Álvarez et al., 2018; Calderón et al., 2020; Garde et al., 2014a, 2019; Hornero et al., 2017; Vaquerizo-Villar et al., 2018a,b,c; Xu et al., 2018). This signal measures the blood oxygen saturation levels, allowing to quantify the number, duration, and depth of oxygen desaturations associated with the occurrence of apneic events (Berry et al., 2012). Thus, several studies have focused on extracting features from SpO_2 using different methodological approaches to characterize the particularities of pediatric OSA (Álvarez et al., 2018; Garde et al., 2014a, 2019; Hornero et al., 2017; Vaquerizo-Villar et al., 2018a,b,c; Xu et al., 2018). In this context, Garde et al. (2014a) designed a phone oximeter based on statistical, non-linear, and spectral features from SpO_2 to detect the disease. In a subsequent study (Garde et al., 2019), the authors aimed to diagnose the different OSA severity degrees and used time and frequency features from SpO_2 for it. Hornero et al. (2017) developed a multicenter study focused on estimating the children's AHI from oximetric indices, statistical moments, spectral and non-linear features, all of them derived from the SpO_2 signal. To the best of our knowledge, their work is the one that has used the largest number of pediatric oximetry recordings in the automatic OSA diagnosis context. Later, their predictive model was also successfully validated in the study conducted by Xu et al. (2018). In addition to conventional analysis, other novel methods have been applied to characterize the behavior of pediatric nocturnal oximetry. In this regard, Álvarez et al. (2018) analyzed the SpO_2 signal by means of symbolic dynamics techniques in order to parameterize the non-linear changes caused by OSA. Moreover, this signal has been characterized through bispectral, detrended fluctuation, and wavelet analy-

sis in Vaquerizo-Villar et al. (2018a,b,c), respectively, which has allowed to obtain novel information about the OSA-related particularities in SpO_2. Another common approach has been to extract clinical indexes from this signal, such as oxygen desaturation indexes that quantify the number of drops $\geq$3% ($ODI3$) and $\geq$4% ($ODI4$) from SpO_2 signal per hour of sleep (Calderón et al., 2020; Chang et al., 2013; Kirk et al., 2003; Tsai et al., 2013). These indexes are commonly used by medical specialists as a suboptimal diagnostic alternative when PSG is not available (Kaditis et al., 2015; Magalang et al., 2003). However, they underestimate the OSA severity (Calderón et al., 2020; Kirk et al., 2003; Oeverland et al., 2002), thus they are usually used along with additional automated features to maximize its diagnostic ability (Garde et al., 2019; Gutiérrez-Tobal et al., 2015; Hornero et al., 2017; Jiménez-García et al., 2020).

Previous studies addressing OSA detection in children and adults have shown the high diagnostic potential of AF signal (Gutiérrez-Tobal et al., 2015, 2016; Jiménez-García et al., 2020; Koley and Dey, 2013). According to the criteria established by the AASM, apnea and hypopnea events are defined based on AF reductions (Berry et al., 2012). Therefore, OSA diagnosis simplification naturally leads to the analysis of this signal. As shown in the Figure 1.3, the presence of apneic and hypopneic events modifies the overnight AF behavior. Thus, when a partial respiratory cessation (hypopnea) occurs, the amount of inspired and expired air is limited, which is reflected in a notable reduction of AF amplitude. Regarding the total respiratory cessation (apnea), the airflow into the lungs is completely blocked, causing AF signal to reach values close to zero. Thus, several research have focused on the automatic analysis of this signal as a useful alternative to PSG (Gutiérrez-Tobal et al., 2015, 2016; Jiménez-García et al., 2020; Koley and Dey, 2013). In adults, this AF analysis is usually performed by means of conventional statistical, non-linear, and/or spectral methods. Some studies have successfully used these techniques to automatically detect apnea and hypopnea events, such as those conducted by Han et al. (2008), Nakano et al. (2007), and Koley and Dey (2013). However, an approach increasingly adopted in recent years is to apply these techniques to characterize OSA particularities in complete AF signal, thus taking advantage of the information from the whole overnight signal instead of only the specific apneic events. In this regard, the studies carried out by Álvarez et al. (2020) and Gutiérrez-Tobal et al. (2013, 2016) achieved accuracies that ranged between 86.5%–95.8% for 5 e/h, 86.5%–91.5% for 10 e/h, 81.0%–85.4% for 15 e/h, and 78.1%–83.3% for 30 e/h only with AF signal in adults. Like in OSA diagnosis simplification in adults, some studies have already evaluated

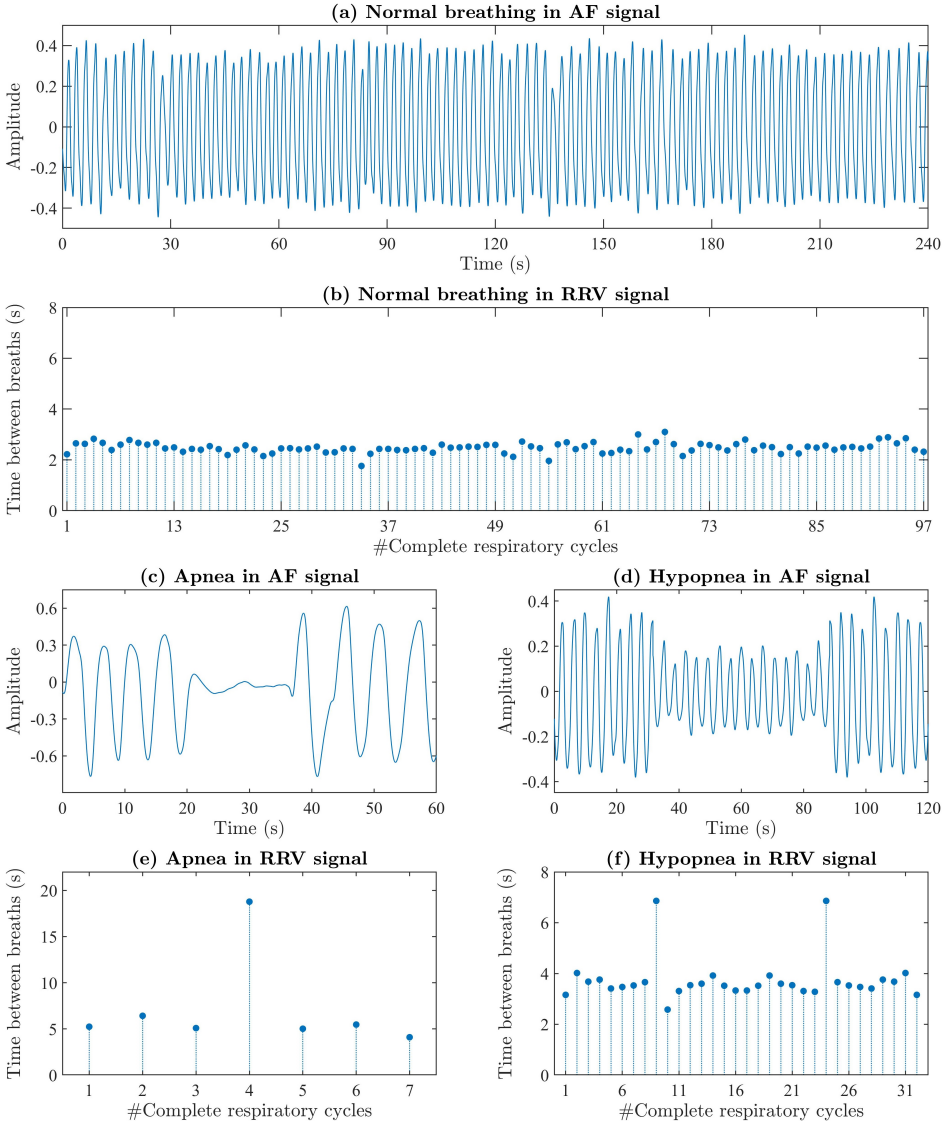


Figure 1.3: (a) Normal breathing, (c) Apnea, and (d) hypopnea events in airflow signal (AF); (b) normal breathing, (e) apnea, and (f) hypopnea events in its corresponding respitarory rate variability signal (RRV).

the usefulness of AF to characterize and diagnose pediatric OSA (Gutiérrez-Tobal et al., 2015; Jiménez-García et al., 2020). However, these studies were based on direct application of conventional methods previously used in adults (non-linear and spectral techniques), leading to a significant decrease in the diagnostic performance of the evaluated methods: accuracies of 79.1% for 3 e/h (Gutiérrez-

Tobal et al., 2015), 80.5% for 1 e/h, 62.8% for 5 e/h, and 79.0% for 10e/h (Jiménez-García et al., 2020), using only AF.

Another respiratory signal that has been used in the context of OSA in adults is respiratory rate variability (RRV) (Gutierrez-Tobal et al., 2012; Gutiérrez-Tobal et al., 2013). This signal measures the elapsed time between two consecutive respiratory cycles and is directly derived from AF (Gutierrez-Tobal et al., 2012; Gutiérrez-Tobal et al., 2013). To obtain this, it is necessary to apply an algorithm that allows for the location of the inspiratory peaks (Korten and Haddad, 1989). These peaks correspond to the relative maximums of AF, so that the algorithm is based on the analysis of first and second derivative of AF to locate them. Afterwards, the elapsed time between these peaks is calculated, thus obtaining the RRV signal (Cysarz et al., 2008). As can be seen in Figure 1.3, the amount of inspired air is suddenly reduced when a hypopnea occurs. This fact leads to an increase in elapsed time between the inspiratory peaks just before and after the event, i.e., when the respiratory flow changes from normal to reduced and from reduced to normal. In the case of apneas, the subject completely stops breathing, therefore the time between breaths increases during the collapse and then abruptly decreases to restore normal respiratory flow. Consequently, OSA also alters the behavior of RRV signal. Thus, this signal has been characterized by means of conventional methods and has shown great utility to help detect OSA in adults, achieving an accuracy of 75.7% for 10 e/h and up to 91.5% when RRV was combined with AF (Gutierrez-Tobal et al., 2012; Gutiérrez-Tobal et al., 2013). In pediatric OSA context, RRV signal was used for the first time in this Doctoral Thesis.

Additionally, the usefulness of automatic AF analysis in the OSA diagnosis context, and by extension of RRV, is supported by the easy acquisition of this single-channel signal. As shown in the Figure 1.4, AF recording can be obtained through a simple thermistor during a Type 4 study (Collop et al., 2007; Flemons et al., 2003). Based on the considerations aforementioned in section 1.4, this fact would have several advantages (Collop et al., 2007; Ferber et al., 1994): (i) fewer signals are monitored in Type 4 study, therefore the required equipment is less expensive than in PSG (Type 1 study); (ii) it is conducted with a portable monitoring equipment that can be used in the patient's home, thus avoiding the alteration of usual sleep patterns; (iii) a single sensor is used (thermistor), which is less uncomfortable and intrusive for children; and (iv) a single channel is analyzed (AF), resulting in a less complex and less time-consuming test. Hence, this would facilitate access to the diagnostic test, allowing early OSA detection in children and reducing long waiting lists.

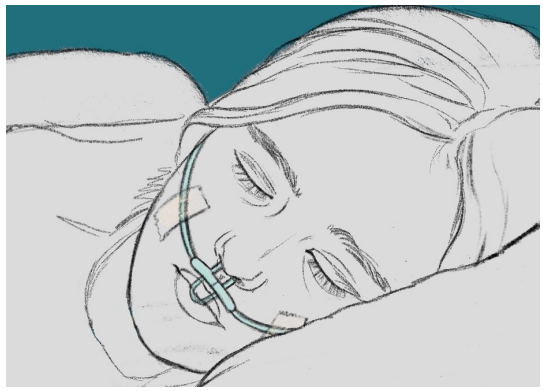


Figure 1.4: Thermistor to record airflow signal.

Chapter 2

Hypothesis and objectives

As previously mentioned, pediatric OSA seriously affects health and quality of life of children who suffer from it. Consequently, it is of the utmost importance that they are timely diagnosed and treated to prevent the negative consequences associated to OSA. Due to the complexity of the standard diagnostic test, the high healthcare costs, as well as the unavailability of required equipment and medical specialists, great efforts have been made to search and develop simpler alternative methods that help diagnose pediatric OSA. Thus, the work conducted during this Doctoral Thesis has focused on characterizing overnight AF by means of different methodological approaches to aid in determining the presence and severity of OSA in children. The starting hypotheses that have motivated this research work, as well as the main and specific objectives to be achieved, are declared in section 2.1 and section 2.2, respectively.

2.1 Hypothesis

Great efforts have been made to simplify the diagnosis of childhood OSA in recent years. As stated in section 1.5, one widely used approach to address this issue has been single-channel automatic analysis of PSG-derived signals (Garde et al., 2019; Hornero et al., 2017; Lazaro et al., 2014). In this Doctoral Thesis, overnight AF analysis is proposed as a potential information source to diagnose OSA in children. This signal reflects the respiratory activity during sleep time, including the breathing pauses associated with OSA (Berry et al., 2012). In this regard, the amplitude of AF manifests a notable reduction in the presence of apneic and hypopneic events (Figure 1.3). Moreover, AF is directly involved in the ap-

nea/hypopnea definition (Berry et al., 2012). Thus, we assume that *the overnight AF gathers useful information to help diagnose pediatric OSA*.

Overnight AF signal has already been characterized in adults and has shown a high ability to diagnose this disease (Gutiérrez-Tobal et al., 2016; Han et al., 2008). The natural way to address the simplification of OSA diagnosis in children is to directly apply the methodologies that have been previous and successfully used in adults, such as conventional non-linear and spectral analyzes (Gutiérrez-Tobal et al., 2015; Jiménez-García et al., 2020). However, pediatric OSA diagnosis presents additional difficulties. On the one hand, the rules for scoring apneas and hypopneas are more conservative than in adults. In this regard, the AASM establishes the duration of event as $\geq$ 2 respiratory cycles for children and $\geq$ 10 seconds for adults (Alonso-Álvarez et al., 2011; Berry et al., 2012). On the other hand, the criteria for determining OSA severity degree are also more restrictive, being 1, 5, and 10 e/h the AHI thresholds commonly used in children and 5, 15 and 30 e/h in adults (Alonso-Álvarez et al., 2011; Kapur et al., 2017; Tan et al., 2014). Thus, an AHI = 5 e/h defines the threshold for mild OSA in adults, while it is often used to recommend surgical treatment in children due to the severe consequences that OSA causes on them (Tan et al., 2014). These diagnostic differences between adults and children hinder the automatic analysis of AF signal, causing that the methods applied in adults are not as effective to characterize the OSA particularities in children. Consequently, it is required other techniques that provide different information from traditional methods previously used in adults, such as changes in underlying dynamics, harmonic components, or energy of AF. Thereby, *novel feature extraction approaches can characterize the behavior of overnight AF in children and improve its diagnostic ability.* In this regard, cardiorespiratory signals, and therefore AF, are dynamic, non-linear, and non-stationary (Martín-González et al., 2018). Hence, it is particularly assumed that *RP, bispectrum, and DWT methods can adapt to the intrinsic properties of pediatric overnight AF and provide useful OSA-related information.*

However, the automatic diagnosis of childhood OSA requires not only extracting new features from AF by means of different methods, but also identifying those that provide relevant and complementary information. Thus, *feature selection methods can identify these features and maximize their diagnostic ability.* In addition, a model that automatically recognizes existing regularities and patterns in extracted data is also needed. Thereby, model must be able to infer behaviors from AF features and use this information to make a prediction of AHI (regression task) or OSA severity degree (classification task) of pediatric subjects. In this

regard, *supervised machine-learning methods can be useful tools to automatically detect the presence and severity of OSA in children.*

Based on the aforementioned considerations, the global hypothesis of this Doctoral Thesis can be summarized as:

> "*The characterization of overnight AF by means of novel approaches can help and simplify pediatric OSA diagnosis.*"

2.2 Objectives

The main objective of this Doctoral Thesis is to design, implement, and assess novel automatic signal processing methods that allow exhaustively characterizing the overnight AF from children and helping in the pediatric OSA diagnosis. In order to achieve this main goal, a four-stage methodology is proposed: (i) preprocessing to resample, automatically remove artifacts, and standardize AF, (ii) feature extraction to characterize pediatric overnight AF, (iii) feature selection to maximize the diagnostic ability of the information extracted from AF, and (iv) application of machine-learning methods to classify children according to the OSA severity degree (classification task), and/or estimate their AHI (regression task). This proposal leads us to the statement of the following specific objectives:

I. To improve the quality of the AF recordings by means of a novel preprocessing method.

II. To characterize the pediatric AF to find the behavioral properties that are intrinsically related to the presence of apneic events.

III. To identify the relevant and non-redundant information from overnight AF that maximize their diagnostic ability.

IV. To develop and evaluate machine-learning models to determine the presence and severity of pediatric OSA with a high diagnostic performance.

Chapter 3

Materials

This chapter describes the databases involved in the research, as well as the signal acquisition process. Thus, section 3.1 presents demographic and clinical characteristics of children that compose the databases under study. Regarding the involved signals and its acquisition process, including recording equipment and sampling rate, they are detailed in section 3.2.

3.1 Subjects under study

In this Doctoral Thesis, we analyzed a database of pediatric AF recordings. This database was provided and expanded by the Comer Children's Hospital, University of Chicago Medicine (Chicago, IL, USA). All the subjects involved in the research were referred to the pediatric sleep unit of this hospital as they presented OSA characteristic symptoms. The common symptoms included daytime hypersomnolence, snoring, respiratory cessations during sleep, and/or overnight awakenings. The children underwent a Type 1 sleep study (PSG) by means of a digital polysomnography device (Polysmith, Nihon Kohden America Inc., Irvine, CA, USA). This device allowed recording up to 32 neurophysiological and cardiorespiratory signals. Afterwards, medical specialists visually inspected the polysomnographic recordings and scored the apnea and hypopnea events following the guidelines of the AASM (Berry et al., 2012). Based on the obtained AHI, the pediatric subjects were diagnosed as no-OSA (AHI $<$ 1 e/h), mild OSA (1 e/h $\leq$ AHI $<$ 5 e/h), moderate OSA (5 e/h $\leq$ AHI $<$ 10 e/h), and severe OSA (AHI $\geq$ 10 e/h) (Alonso-Álvarez et al., 2011; Hornero et al., 2017; Tan et al., 2014). The legal caretakers of all children gave their informed written consent for participat-

ing in the study, whose protocol were approved by the Ethics Committee of the Comer Children's Hospital (approval numbers: 11-0268-AM017, 09-115-B-AM031, IRB14-1241) and conformed to the Declaration of Helsinki.

Firstly, the database was composed by AF recordings from 501 pediatric subjects aged from 0 to 13 years (314 boys and 187 girls). Their demographic and clinical data are presented in Table 3.1. Out of the 501 children, 134 were healthy controls (AHI $<$ 1) and 367 patients with OSA (AHI $\geq$ 1). Thereby, 501 AF recordings were analyzed in our first study (Barroso-García et al., 2017). Afterwards, the database was expanded, involving 946 pediatric subjects (584 boys and 362 girls). Their demographic and clinical data are shown in Table 3.2. The age of the subjects was kept in a range from 0 to 13 years. Out of the 946 children, 163 were healthy controls and 783 patients with OSA. The expanded database was used in the rest of our studies (Barroso-García et al., 2020, 2021a,b).

Table 3.1: Initial database: clinical and demographic characteristics from the involved children.

	All	**No-OSA**	**Mild OSA**	**Moderate OSA**	**Severe OSA**
Subjects (n)	501	134 (26.7%)	187 (37.3%)	76 (15.2%)	104 (20.8%)
Age (years)	6 [6]	7 [6]	6 [4]	5 [5]	4 [5.5]
Males (n)	314 (62.7%)	86 (64.2%)	119 (63.6%)	43 (56.6%)	66 (63.4%)
BMI (kg/m^2)	17.8 [6.5]	17.3 [5.4]	17.9 [6.4]	19.0 [8.9]	17.6 [6.7]
AHI (e/h)	3.2 [7.1]	0.5 [0.6]	2.6 [1.9]	6.8 [2.5]	18.3 [16.1]

Data are presented as median [interquartile range] or number (%). BMI: body mass index, AHI: apnea-hypopnea index, OSA: obstructive sleep apnea.

Table 3.2: Expanded database: clinical and demographic characteristics from the involved children.

	All	**No-OSA**	**Mild OSA**	**Moderate OSA**	**Severe OSA**
Subjects (n)	946	163 (17.2%)	386 (40.8%)	172 (18.2%)	225 (23.8%)
Age (years)	6 [6]	7 [6]	6 [5]	5 [6]	5 [5.3]
Males (n)	584 (61.7%)	98 (60.1%)	242 (62.7%)	106 (61.6%)	138 (61.3%)
BMI (kg/m^2)	17.9 [6.2]	17.4 [5.7]	17.8 [5.5]	18.9 [7.9]	18.3 [7.3]
AHI (e/h)	3.8 [7.8]	0.5 [0.6]	2.5 [1.8]	6.8 [2.4]	19.1 [17.2]

Data are presented as median [interquartile range] or number (%). BMI: body mass index, AHI: apnea-hypopnea index, OSA: obstructive sleep apnea.

3.2 Signals under study

Overnight AF was recorded by a thermal sensor integrated in the polysomnographic system. Thermistor was placed in nostrils and mouth from children to measure the difference of temperature between inspired (cold air) and expired air (warm air). AF recordings that composed the two databases were acquired at sampling frequencies of 200 and 500 Hz. All of them were pre-processed before its analysis. The pre-processing stage details are indicated in section 4.1.1. Signals < 3 hours of duration after pre-processing were excluded from the study (Álvarez et al., 2020). Regarding RRV signals, these were derived from pre-processed AF recordings. As stated above, an algorithm was applied to detect inspiratory peaks in AF, i.e., where its first derivative pass from positive to negative (Korten and Haddad, 1989). Afterwards, the elapsed time between consecutive inspiratory peaks was computed, thus obtaining RRV signal (Cysarz et al., 2008).

In addition, SpO_2-derived $ODI3$ was used for comparison purposes. Overnight SpO_2 signal was recorded by means of a fingertip pulse oximeter integrated in the polysomnographic system. Pulse oximeter included an optical sensor to measure the blood oxygen saturation level by the intensity of light transmitted through finger tissues (the intensity is lower as the blood oxygen saturation is higher). These oximetry recordings were acquired at sampling frequencies of 200 and 500 Hz. All of them were pre-processed before computing $ODI3$ (section 4.1.2). Afterwards, an algorithm was applied to automatically detect drops $\geq 3\%$ from preceding SpO_2 baseline (Taha et al., 1997). The established minimum duration to score an event as oxygen desaturation was 10 seconds. The total number of these events was quantified and then divided by the number of hours of SpO_2 recording, thus obtaining $ODI3$.

Chapter 4

Methods

This chapter is devoted to the methodology conducted during this Doctoral Thesis. As shown in the Figure 4.1, this methodology consists of four stages: pre-processing, feature extraction, feature selection, and application of machine-learning approaches. Firstly, the AF and SpO_2 recordings were subjected to a pre-processing stage to resample them and automatically remove noise and artifacts. Moreover, AF signals were standardized to minimize the effects caused by other OSA-unrelated particularities. Afterwards, a feature extraction stage was performed to comprehensively characterize pediatric overnight AF by means of different techniques. In this regard, cardiorespiratory signals, and therefore AF, are dynamic, non-linear, and non-stationary (Martín-González et al., 2018). Consequently, non-linear, spectral, bispectral, RP, and wavelet analyses have been conducted for adapting to the intrinsic properties of overnight AF and getting useful OSA-related information from it. In order to identify those features from AF that provide relevant and complementary information while maximizing its diagnostic ability, a feature selection stage has been applied. Thus, wrapper and filter algorithms were implemented: forward stepwise logistic regression (FSLR) and fast correlation-based filter (FCBF), respectively. Then, supervised machine-learning methods have been applied to recognize regularities in AF features, infer behaviors from them, and use this information to automatically detect the presence and severity of OSA in children. This stage was conducted from three different approaches: discrimination between OSA-negative and OSA-positive pediatric subjects (binary classification task), classification of children according to their OSA severity degree (multiclass classification task), and AHI estimation of each child (regression task). The binary and multiclass classification tasks were performed

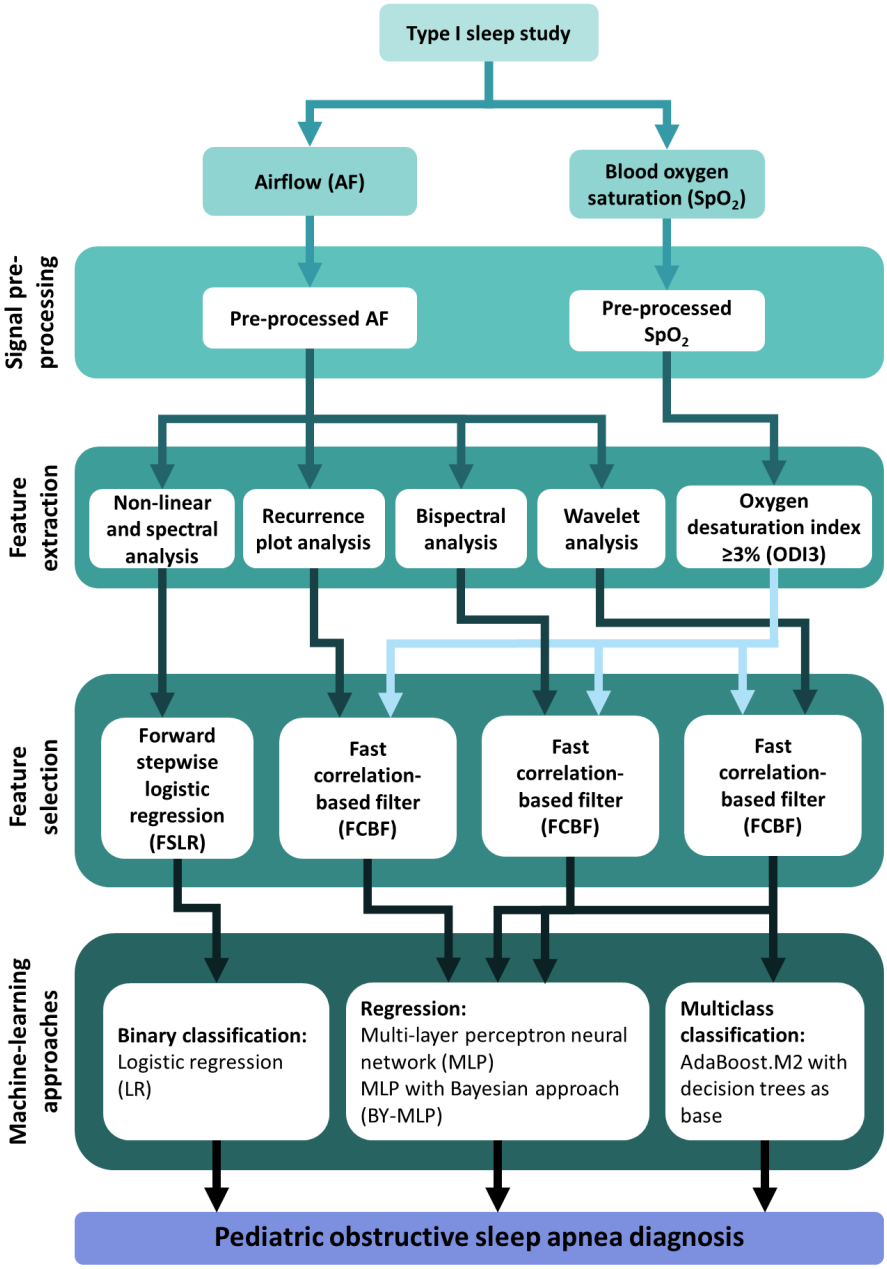


Figure 4.1: Methodological scheme of the Doctoral Thesis.

by means of logistic regression (LR) and adaptive boosting (AdaBoost.M2) algorithms, respectively. Regarding the regression task, it was performed through a multi-layer perceptron neural network (MLP) and a MLP with Bayesian approach (BY-MLP).

The methods applied in each of these stages are comprehensively described in sections 4.1 (signal pre-processing), 4.2 (feature extraction), 4.3 (feature selection), and 4.4 (machine-learning approaches). Finally, statistical techniques used to assess the diagnostic ability of the proposed methods are detailed in section 4.5.

4.1 Signal pre-processing

After reviewing the pediatric AF and SpO_2 recordings, it was observed that they had been recorded at different sampling frequencies, presented different resolution, and contained undesired artifacts caused by sensor contact loss due to movements of the children during sleep time. Thus, these recordings were subjected to a pre-processing stage to improve its quality and increase the effectiveness of subsequent analysis. The techniques used for this purpose are described below.

4.1.1 AF signal

The AF recordings obtained during Type 1 sleep study were recorded at 200 and 500 Hz. Thus, the first pre-processing step consisted of resampling these signals so that all of them had the same sampling frequency. Since the AASM recommends using 100 Hz for AF signal (Berry et al., 2012), this frequency was established to carry out the resampling process.

Afterwards, a Finite Impulse Response (FIR) filter with linear phase property was applied to remove the noise from AF signal without introducing phase distortion (Oppenheim et al., 1999). Thus, a Kaiser windowed low-pass filter was designed to establish a trade-off between the main-lobe width and the side-lobe amplitude, independently defining the transition bandwidth and the stopband attenuation (Oppenheim et al., 1999; Saramäki et al., 1993). A cut-off frequency = 1.5 Hz, transition band = 2 Hz, attenuation = 100 db, and sampling frequency = 100 Hz were the parameters used to design this filter.

In order to minimize inter-individual differences caused by particular features other than OSA, AF recordings were standardized. In this regard, the signal normalization method proposed by Várady et al. (2002) was used for this purpose. Thus, an adaptive baseline correction and scale factor were applied to each sam-

ple of filtered AF signal (Várady et al., 2002), which allowed the acquisition of normalized samples with values in the range [-1, 1].

Finally, a novel method was implemented to automatically remove possible noise sources contained in AF signal, such as noise caused by equipment, spurious, and lost samples due to lack of contact with the sensor (see Figure 4.2). According to the properties of these noisy elements, we have developed an artifact elimination protocol based on statistical moments from AF segments. Particularly, 30-second segments with a standard deviation ≥ 0.550 or ≤ 0.026, or a kurtosis ≤ 1.320 were considered artifacts and removed from AF recordings. Figure 4.3 shows a fragment of AF signal before and after pre-processing. As can be seen, artifacts were removed after signal pre-processing.

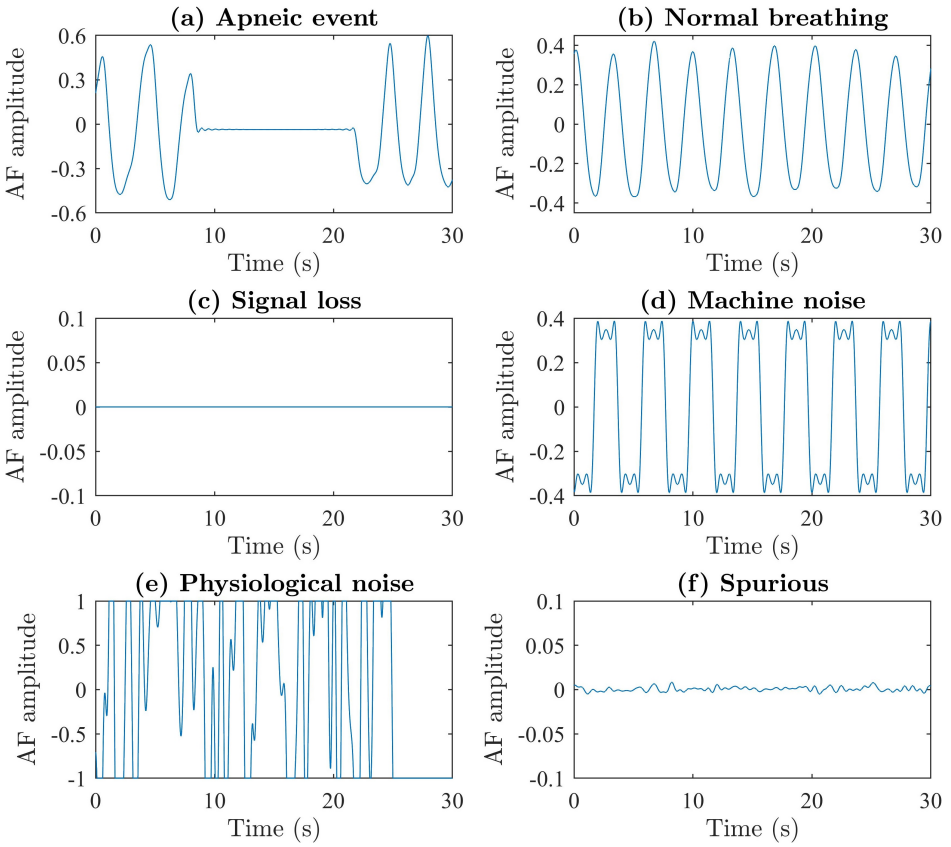


Figure 4.2: Segment types contained in airflow signal (AF).

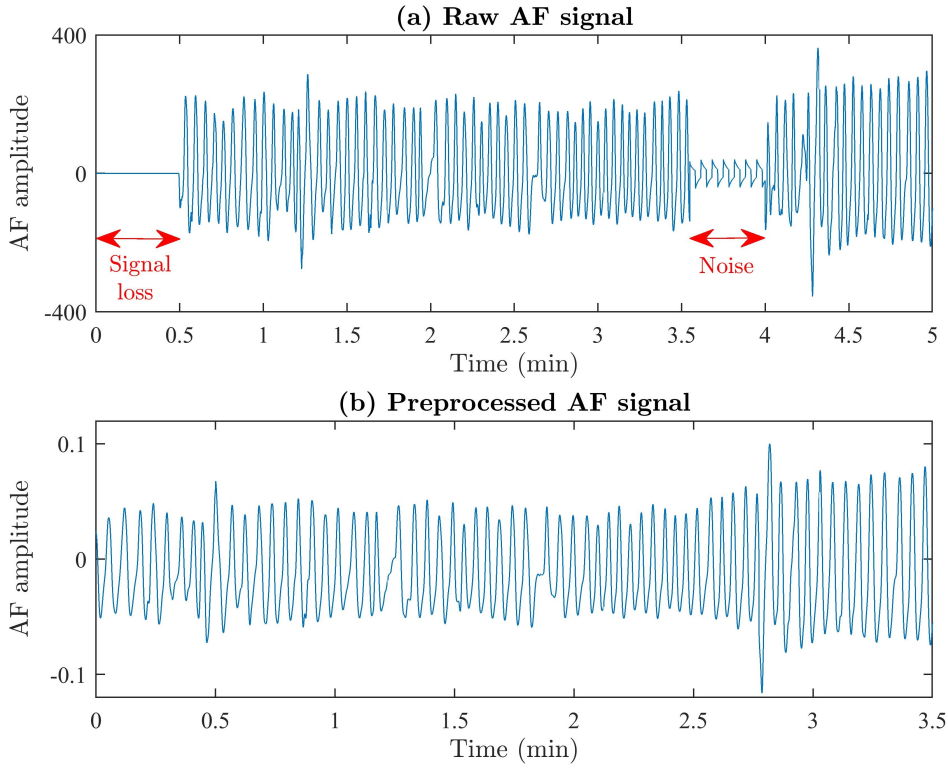


Figure 4.3: Airflow signal (AF) before and after pre-processing stage.

4.1.2 SpO_2 signal

Like AF signals, SpO_2 recordings were also subjected to a pre-processing stage before computing the $ODI3$. It was observed that these recordings contained artifacts and lost samples caused by movements of the children during sleep time and lack of contact with the sensor. These artifacts were automatically removed from SpO_2 signals following the method proposed by Magalang et al. (2003). Thereby, SpO_2 values $< 50\%$ and slope changes $\geq 4\%$ per second were considered artifacts and removed from SpO_2 recordings (Magalang et al., 2003). Regarding the sampling frequency, oximetry signals were recorded at 200 and 500 Hz during Type 1 sleep study. In order to homogenize their sampling frequency, they were resampled at 25 Hz as recommended by the AASM (Berry et al., 2012). In addition, SpO_2 values were rounded to the second decimal so that all oximetry signals had the same resolution (Garde et al., 2014b). Figure 4.4 shows a SpO_2 signal before and after pre-processing. As can be seen, artifacts were removed after signal pre-processing.

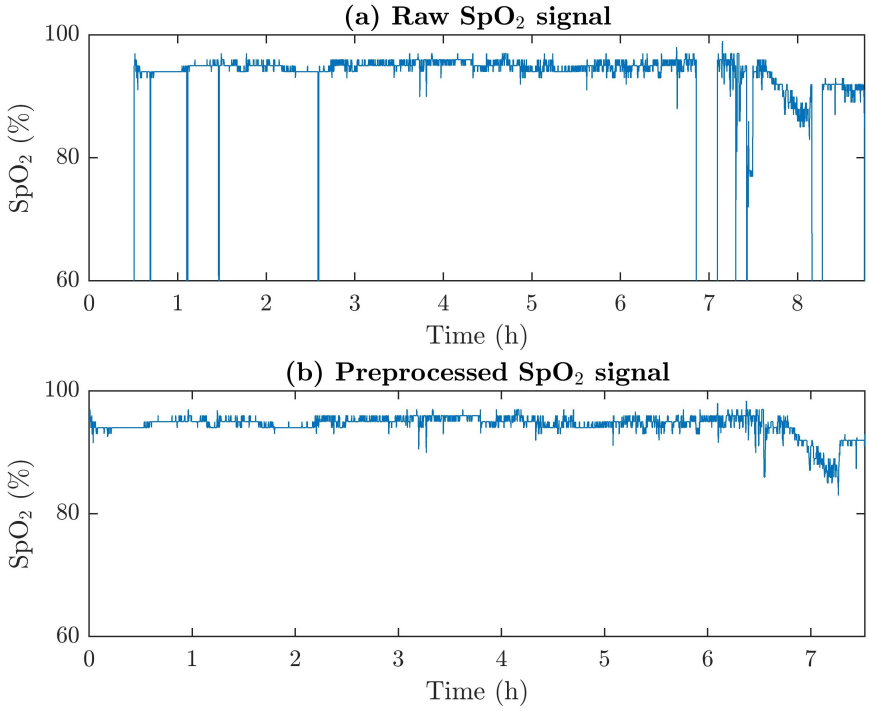


Figure 4.4: Oximetry signal (SpO_2) before and after pre-processing stage.

4.2 Feature extraction

As previously mentioned, the research conducted during the Doctoral Thesis has focused on exhaustively characterizing the behavior of overnight AF in children to obtain relevant and useful information that helps to simplify the diagnosis of pediatric OSA. Therefore, a methodological stage of feature extraction from AF was needed to address its characterization. In this regard, the classical feature extraction methods have been shown to be less effective characterizing the OSA particularities in children than in adults due to the diagnostic differences between them (Álvarez et al., 2020; Gutiérrez-Tobal et al., 2013, 2015, 2016; Jiménez-García et al., 2020). Consequently, novel approaches have been proposed in this work to adapt to the intrinsic properties of AF signal and provide different information from classical methods. These techniques are comprehensively described below.

4.2.1 Central tendency measure

Based on first-order difference plots, the central tendency measure (CTM) is a non-linear method that allows to quantify the variability of a signal (Cohen et al., 1996). A first-order difference plot is a scatter plot where the differences between displaced samples of the original signal are represented: $(x[n+2]-x[n+1])$ versus $(x[n+1]-x[n])$, being $x[n]$ the value of the n^{th} sample of the signal (Abásolo et al., 2006; Cohen et al., 1996). CTM is computed from these graphic representations by selecting a circular region of radius r centered in the plot origin. Its value is determined by the number of graph points that are located within the circular region, divided by the total number of graph points. The total number of graph points is $N-2$, where N is the number of samples of the signal. Thereby, the CTM is mathematically defined as (Cohen et al., 1996):

$$CTM = \frac{\sum_{n=1}^{N-2} \delta(n)}{N-2}, \tag{4.1}$$

where $\delta(n)$ is:

$$\delta(n) = \begin{cases} 1, & \text{if } ((x[n+2]-x[n+1])^2 + (x[n+1]-x[n])^2)^{1/2} \leq r \\ 0, & \text{otherwise} \end{cases} . \tag{4.2}$$

As can be seen, CTM values are within the range [0, 1]. Thereby, its value will tend to 0 when the graph points are widely scattered, which will indicate a high variability of the signal (Abásolo et al., 2006). In contrast, a high density of graph points around the plot origin will lead to CTM values close to 1, indicating a low variability (Abásolo et al., 2006). Note that CTM will not provide accurate information about signal variability if unsuitable values of r are selected. The circular region may not contain graph points if r is too small and almost all of them will be located within this region if r is too large. Consequently, a suitable choice of r is essential. In this regard, there is no guideline to optimize its value since it is chosen according to the character of the data (Abásolo et al., 2006). In this study, CTM was computed with several r values. Then, Spearman's correlation was used to assess the relationship between the AHI and the CTM obtained for each of them in the corresponding training set. Finally, the r value with the highest correlation was chosen as optimal.

4.2.2 Spectral entropies

Power spectral density (PSD) estimation is an approach typically used to analyze the spectral properties of cardiorespiratory signals (Gutiérrez-Tobal et al., 2015; Jiménez-García et al., 2020; Martín-Montero et al., 2020). Note that these signals are not stationary, i.e. their characteristics change over time (Martín-González et al., 2018). In order to deal with this limitation, PSD is estimated using Welch's method (Welch, 1967). Thereby, the signal is divided into K overlapping subsequences of length L using a window function w_f. Then, the discrete Fourier transform (DFT) is applied to compute the modified periodogram $\hat{P}_k$ of each subsequence x_k (Welch, 1967):

$$\hat{P}_k(f_n) = \frac{\left|\sum_{i=0}^{L-1} x_k(i) \cdot w_f(i) \cdot e^{-\frac{2kijn}{L}}\right|^2}{LU}, k = 1, ..., K, \tag{4.3}$$

where f_n and U are defined as:

$$f_n = \frac{n}{L}, n = 0, ..., L/2, \tag{4.4}$$

$$U = \frac{\sum_{i=0}^{L-1} |w_f(i)|^2}{L}. \tag{4.5}$$

Finally, the PSD of the signal is estimated by averaging the K obtained periograms:

$$PSD(f_n) = \frac{\sum_{k=1}^{K} \hat{P}_k(f_n)}{K}. \tag{4.6}$$

Once the PSD has been obtained, spectral entropies of i^{th} order (SE_i) are computed from it. Based on Shannon's entropy, spectral entropy (SE) is a measure that quantifies the irregularity of a signal in the time domain by measuring the uniformity of its spectral distribution (Inouye et al., 1991). Thereby, a flat spectrum with uniform spectral content will present high SE values, indicating a high irregularity of the signal. In the same way, a peaked spectrum with few frequency components will have low SE values, indicating a low irregularity of the signal (Inouye et al., 1991). Due to higher-order SE can potentiate the spectral distribution changes of a signal, spectral entropies of first (SE_1), second (SE_2), and third (SE_3) order are obtained in this study (Barroso-García et al., 2017). Thus, SE_i is computed by means of the following equation (Hornero et al., 2008; Jiménez-García et al., 2020):

$$SE_i = -\frac{\sum_{f=f_1}^{f_2} PSDn_i(f) \cdot log[PSDn_i(f)]}{log(M)}, i = 1, 2, 3, \tag{4.7}$$

where i is the order of entropy, M is the number of frequency components, and $PSDn$ is the normalized PSD in the frequency range f_1–f_2 (Atri and Mohebbi, 2015; Jiménez-García et al., 2020):

$$PSDn_i(f) = \frac{|PSD(f)|^i}{\sum_{f=f_1}^{f_2} |PSD(f)|^i}. \tag{4.8}$$

Note that division by $log(M)$ normalizes the SE_i values between 0 and 1. Thus, values close to 0 will indicate that the signal is very regular, while values close to 1 will indicate an irregularity increase.

4.2.3 Recurrence plot-derived features

Recurrence plot (RP) is a non-linear analysis method that provides information about the underlying dynamics and phase-space of a dynamic system (Zbilut and Webber, 2006). Dynamic systems are those that change and evolve over time, such as physiological signals, so RPs can also be applied to non-stationary data (Martín-González et al., 2018; Marwan et al., 2007). RPs graphically represent the recurrences of these systems, i.e., the times at which they return to previous phase-space states, allowing to visualize periodicity patterns (Marwan et al., 2007). Thus, it is necessary to know the trajectories of the m-dimensional phase-space that define the dynamic system. Since a signal is a one-dimensional system of length N, these trajectories are obtained by reconstructing its phase-space (Martín-González et al., 2018; Marwan et al., 2007). In this regard, the Taken's time-delay method is widely used for this purpose (Takens, 1981). This method consists of generating subsequences of dimension m with τ-delayed samples from the original signal $x(t)$. Thereby, its trajectories $\vec{x}_i$ can be defined as follows (Martín-González et al., 2018; Takens, 1981):

$$\vec{x}_i = [x(i), x(i+\tau), \ldots, x(i+(m-1)\cdot\tau)], i = 1, \ldots, N-(m-1)\cdot\tau. \tag{4.9}$$

Once the phase-space trajectories have been obtained, the recurrence matrix $R_{i,j}$ is calculated for its subsequent graphic representation. This is a binary and symmetric matrix of size $N-(m-1)\cdot\tau \times N-(m-1)\cdot\tau$, whose values are 1 if there is a recurrence and 0 otherwise. Thus, two trajectories $\vec{x}_i$ and $\vec{x}_j$ are recurrent

when their Euclidean distance is less than the distance threshold ε (Marwan et al., 2007; Nguyen et al., 2014):

$$R_{i,j} = \begin{cases} 1, & \text{if } \| \vec{x}_i - \vec{x}_j \| \leq \varepsilon \\ 0, & \text{otherwise} \end{cases} , i, j = 1, \ldots, N - (m-1) \cdot \tau, \tag{4.10}$$

Note that RP will not provide accurate information about the underlying dynamics of the signal if unsuitable values of τ, m, or ε are selected. An inappropriate τ value could lead to auto-correlated state trajectories (Martín-González et al., 2018; Nguyen et al., 2014). In order to avoid it, τ is usually adjusted by means of the auto-mutual information (AMI) function (Jeong et al., 2001; Martín-González et al., 2018; Marwan et al., 2007). This non-linear technique estimates the predictability of a signal from its previous values (Jeong et al., 2001). The dependency degree between the original signal and its τ-delayed version is quantified for each τ value by measuring the mutual information between them (Jeong et al., 2001). Thereby, the optimum τ value is determined as the first AMI local minimum of the signal. Regarding m, an overly high value could reduce isolated recurrences and generate false diagonal structures in a RP (Marwan et al., 2007). In order to avoid it, this parameter is adjusted by means of the well-known false nearest neighbor's (FNN) method (Kennel et al., 1992; Martín-González et al., 2018; Marwan et al., 2007). This method finds the minimum dimension required to reconstruct the signal phase-space by examining the number of neighbors of each trajectory point (Kennel et al., 1992). Widely separated points in the phase-space will be considered nearby (false neighbors) if m value is too low. Thus, m is increased until its optimal value is reached, i.e., when the number of false neighbors is 0 (Kennel et al., 1992). In addition, a suitable choice of ε is crucial to obtain the RP. This will not contain recurrences if ε is too small and almost all of them will be spurious if ε is too large (Marwan et al., 2007). In this regard, the fixed distance method is commonly applied to adjust this parameter due to its suitability for signal detection (Martín-González et al., 2018; Schinkel et al., 2008). Based on the standard deviation (σ) of each signal, this method sets the distance threshold as $\varepsilon = \lambda \cdot \sigma$ (Ramírez Ávila et al., 2013; Schinkel et al., 2008). In this study, RPs was computed with several ε values in the training set. Then, Spearman's correlation was used to assess the relationship between the AHI and the RP features obtained from each of them. Finally, the ε value with which RP features reached the highest average correlation was chosen as fixed distance.

Once the RPs have been computed, these are characterized by means of the recurrence quantification analysis (RQA) (Marwan et al., 2007; Zbilut and Webber, 2006). RPs can contain isolated recurrences, recurrences shaping diagonal lines (deterministic structures) and/or recurrences forming vertical lines (laminar structures). Thus, RQA is based on quantifying the number and duration of these recurrences according to the structure type that they form (Marwan et al., 2007). Thereby, the following features are extracted from RP:

- Features based on the recurrence density:
  - Recurrence rate (REC). It quantifies the variability of a signal by measuring the proportion of recurrences contained in its RP (Marwan et al., 2007; Zbilut and Webber, 2006). Thus, many signal trajectories returning to previous phase-space states will lead to a high REC value, which will indicate a low signal variability (Martín-González et al., 2018):

$$REC = \frac{\sum_{i,j=1}^{L} R_{i,j}}{L^2}, \tag{4.11}$$

    where $L = N - (m-1) \cdot \tau$.

- Features based on diagonal structures:
  - Determinism (DET). This feature quantifies the predictability of a signal by measuring the proportion of recurrences that constitute diagonal lines in the RP (Martín-González et al., 2018; Marwan et al., 2007). A RP with many diagonal lines and few isolated recurrences will result in a high DET value, indicating a high signal predictability (Marwan et al., 2007):

$$DET = \frac{\sum_{l=l_{\min}}^{L} l \cdot p(l)}{\sum_{i,j=1}^{L} R_{i,j}}, \tag{4.12}$$

    where $p(l)$ is the distribution of diagonal lines of length l. Diagonal lines must have a minimum length ($l_{\min}$) to be considered as such. An overly high $l_{\min}$ could lead to inaccurate DET values. Thus, $l_{\min} = 2$ is commonly used to calculate this feature (Martín-González et al., 2018; Schinkel et al., 2008).

– Average length of the diagonal lines (LEN). It measures the average prediction time of a signal, i.e., the time during which its trajectories can be predicted from previous phase-space states (Acharya et al., 2011; Martín-González et al., 2018). This is obtained by quantifying the time that the trajectories run in close phase-space states, which is determined by the length of the diagonal lines of the RP (Marwan et al., 2007). Thus, a RP with many long diagonal lines will lead to a high LEN value, indicating a high average prediction time of the signal:

$$LEN = \frac{\sum_{l=l_{\min}}^{L} l \cdot p(l)}{\sum_{l=l_{\min}}^{L} p(l)}. \tag{4.13}$$

– Maximum length of the diagonal lines ($L_{\max}$). This feature estimates the exponential divergence of the signal trajectories in phase-space from the diagonal line lengths of the RP (Martín-González et al., 2018; Zbilut and Webber, 2006). A RP with short diagonal lines, and therefore a low $L_{\max}$ value, will indicate that the signal trajectories diverge rapidly in the phase-space (Marwan et al., 2007):

$$L_{\max} = \max(l_i), i = 1, ..., \sum_{l=l_{\min}}^{L} p(l). \tag{4.14}$$

– Entropy of the diagonal distribution ($ENTR$). Based on Shannon's entropy, $ENTR$ quantifies the irregularity of a signal by measuring the uniformity of the distribution of diagonal line lengths of its RP (Leonardi, 2018; McCamley et al., 2017). Thereby, a RP with uniformly distributed diagonal line lengths will present a high $ENTR$ value, indicating a high irregularity of the signal. In the same way, a distribution concentrated around a few length values will result in a low $ENTR$ value, indicating a low irregularity of the signal (Leonardi, 2018; McCamley et al., 2017):

$$ENTR = -\sum_{l=l_{\min}}^{L} p(l) \cdot log[p(l)]. \tag{4.15}$$

– Trend ($TREND$). It estimates the non-stationarity of a signal from the distribution of recurrences in diagonal lines parallel to the main diagonal of its RP (Marwan et al., 2007). Thus, a RP without drifts and with homogeneously distributed recurrences will result in $TREND \approx 0$, indicating a high stationarity of the signal (Webber and Zbilut, 1994;

Zbilut and Webber, 2006). However, a RP with fading of recurrences towards the corners will have a high absolute value of $TREND$, indicating an increase in the non-stationarity degree of the signal (Webber and Zbilut, 1994; Zbilut and Webber, 2006):

$$TREND = \frac{\sum_{i=1}^{\tilde{L}}(i - \tilde{L}/2) \cdot (REC_i - \langle REC_i \rangle)}{\sum_{i=1}^{\tilde{L}}(i - \tilde{L}/2)^2}, \tag{4.16}$$

where $\tilde{L} < L$ to exclude the RP edges, REC_i is the recurrence density in the i-distant line from the main diagonal, and $\langle REC_i \rangle$ is the average REC_i.

- Features based on vertical structures:
  - Laminarity (LAM). This feature quantifies the complexity of a signal by measuring the proportion of recurrences that constitute vertical lines in the RP (Martín-González et al., 2018). These structures arise when the signal does not change state or changes very slowly (Martín-González et al., 2018; Zbilut and Webber, 2006). Thus, a RP with many vertical lines and few isolated recurrences will result in a high LAM value (Marwan et al., 2007). This will indicate that the signal has few state changes and, therefore, a low complexity (Martín-González et al., 2018):

$$LAM = \frac{\sum_{v=v_{\min}}^{L} v \cdot p(v)}{\sum_{i,j=1}^{L} R_{i,j}}, \tag{4.17}$$

    where $p(v)$ is the distribution of vertical lines of length v. Vertical lines must have a minimum length ($v_{\min}$) to be considered as such. An overly high $v_{\min}$ could lead to inaccurate LAM values. Thus, $v_{min} = 2$ is commonly used to calculate this feature (Marwan et al., 2007; Schinkel et al., 2008).
  - Average length of the vertical lines (trapping time, TT). It measures the average trapping time of a signal, i.e., the time during which its trajectories do not change state (Acharya et al., 2011; Nguyen et al., 2014). This feature is obtained by quantifying the length of the vertical lines of the RP (Marwan et al., 2007). Thereby, a RP with many long vertical lines will lead to a high TT value. This fact will indicate that

the signal abides in a particular state for a long time and, therefore, it has a low complexity (Martín-González et al., 2018):

$$TT = \frac{\sum_{v=v_{\min}}^{L} v \cdot p(v)}{\sum_{v=v_{\min}}^{L} p(v)}. \tag{4.18}$$

– Maximum length of the vertical lines ($V_{\max}$). It quantifies the complexity of a signal from the duration of the laminar structures presented in its RP (Martín-González et al., 2018; Marwan et al., 2007). Thus, a RP with long vertical lines will result in a high $V_{\max}$ value, indicating a low complexity of the signal (Martín-González et al., 2018):

$$V_{\max} = \max(v_i), i = 1, ..., \sum_{v=v_{\min}}^{L} p(v). \tag{4.19}$$

4.2.4 Bispectral features

Bispectrum is a frequency analysis method that provides information about the non-gaussianity, non-linearity, and non-stationarity of a time series (Chua et al., 2010; Zhang et al., 2000). Although these are intrinsic properties of cardiorespiratory signals such as AF (Martín-González et al., 2018), the classical spectral analysis methods ignore them since they are not capable of detecting behavioral changes or trends associated to these properties (Martín-Montero et al., 2021; Vaquerizo-Villar et al., 2018b). In contrast, bispectrum can reveal this type of alterations and is less sensitive to noise, making it a suitable analysis method for physiological signals (Atri and Mohebbi, 2015; Martín-Montero et al., 2021). In addition, the signal amplitude and phase information is kept by means of this technique, which also allows us to discover interactions and phase couplings between the different signal patterns (Chua et al., 2010; Emin Tagluk and Sezgin, 2011).

Computationally, bispectrum represents the third-order cumulant of the spectral probability distribution of a signal and it can be estimated by means of the following equation (Chua et al., 2010; Martín-Montero et al., 2021):

$$B(f_1, f_2) = X(f_1) \cdot X(f_2) \cdot X^*(f_1 + f_2), f_1, f_2 = 0, ..., f_s/2, \tag{4.20}$$

where $X(f)$ is the Fourier transform, f_1 and f_2 are the frequency indices, and f_s is the signal sampling rate. Note that the bispectral symmetry conditions allow the acqusition of the full bispectrum from the non-redundant triangular region defined by $0 \leq f_1 \leq f_2 \leq f_1 + f_2 \leq f_s/2$ (Chua et al., 2010). Then, bispectrum is

normalized to ensure that its magnitude values are bounded in the range [0, 1]:

$$B_N(f_1, f_2) = \frac{B(f_1, f_2)}{\sum_{f_1,f_2=0}^{f_s/2} |B(f_1, f_2)|}. \tag{4.21}$$

This fact allows reducing differences between subjects caused by physiological factors other than OSA, as well as estimating the phase coupling degree in equal conditions (Chua et al., 2010).

Once the bispectrum has been normalized, the bispectral band associated to the individual respiratory rate is determined. This adaptive band (AB) is obtained by defining a range of 0.075 Hz around the point where the maximum bispectral amplitude is reached, i.e., around the normal respiration peak (Fleming et al., 2011; Milagro et al., 2019). Afterwards, AB is characterized by means of the following features:

- Features based on the bispectral amplitude:
  - Maximum bispectral amplitude (B_{max}). This feature quantifies the maximum value of bispectral magnitudes contained within the region under study (Wang et al., 2009). It corresponds to the maximum bispectral power. Therefore, a high B_{max} value will indicate a high maximum bispectral power inside the adaptive band:

$$B_{max} = \max(|B_N(f_1, f_2)|_{f_1,f_2 \in AB}), \tag{4.22}$$

  - Minimum bispectral amplitude (B_{min}). It quantifies the minimum value of bispectral magnitudes contained within the adaptive band (Martín-Montero et al., 2021). It corresponds to the minimum bispectral power of the considered region. Hence, the minimum bispectral power inside the band will be lower as lower is B_{min} value:

$$B_{min} = \min(|B_N(f_1, f_2)|_{f_1,f_2 \in AB}), \tag{4.23}$$

  - Total bispectral power (B_{total}). This feature is computed as the sum of the bispectral amplitudes inside adaptive band (Ning and Bronzino, 1990). It allows quantifying the gaussianity deviations of the signal components that are associated to this region. Thereby, gaussian components (i.e. normally distributed components) will lead to bispectral amplitude values = 0 and, thus, to low B_{total} values (Chua et al., 2010):

$$B_{total} = \sum_{f_1,f_2 \in AB} |B_N(f_1, f_2)|, \tag{4.24}$$

- Features based on entropies:

  – Bispectral entropies of i^{th} order (BE_i). Based on Shannon's entropy, bispectral entropy is a measure that quantifies the irregularity of a signal by measuring the uniformity of its bispectral distribution (Chua et al., 2010; Vaquerizo-Villar et al., 2018b). Thereby, a flat bispectrum with uniform bispectral content will present high BE_i values, indicating a high irregularity of the signal (Martín-Montero et al., 2021). In the same way, a peaked bispectrum with few involved frequency components will have low BE_i values, indicating a low irregularity of the signal (Martín-Montero et al., 2021). Due to higher-order entropies can potentiate the bispectral distribution changes of a signal, bispectral entropies of first (BE_1), second (BE_2), and third (BE_3) order are obtained by means of the following equations (Atri and Mohebbi, 2015; Chua et al., 2010):

$$BE_i = -\sum_{j=1}^{J} p_j \cdot log(p_j), i = 1, 2, 3, \tag{4.25}$$

  where i is the entropy order, J is the number of histogram bins, and p is the magnitude distribution in AB:

$$p_j = \frac{|B_N(f_1, f_2)|^i}{\sum_{f_1,f_2 \in AB} |B_N(f_1, f_2)|^i}. \tag{4.26}$$

  – Phase entropy (PE). This feature is also a variant of the Shannon's entropy. It quantifies the irregularity of the bispectral phase signal in the considered region, basing on the distribution of its phase angles (Chua et al., 2010; Vaquerizo-Villar et al., 2018b). PE is computed as follows:

$$PE = -\sum_{n=1}^{N} p(\psi_n) \cdot log[p(\psi_n)], \tag{4.27}$$

  where N is the number of histogram bins and $p(\psi)$ is the phase angle distribution in AB:

$$p(\psi_n) = \frac{1}{L} \cdot \sum_{f_1,f_2 \in AB} Ind[\phi(B_N(f_1, f_2)) \in \psi_n], \tag{4.28}$$

such that L is the size of the region AB, ϕ is the bispectral phase angle, and $Ind(\cdot)$ is the indicator whose value is 1 if ϕ belongs to the range ψ_n:

$$\psi_n = \left\{ \phi \left| -\pi + \frac{2\pi n}{N} \leq \phi \leq -\pi + \frac{2\pi(n+1)}{N} \right. \right\}. \tag{4.29}$$

As in the previous entropies, uniformly distributed phase angles in a wide range of harmonic components will give rise to high PE values, indicating a high irregularity of the signal phase in the bispectral region under study (Martín-Montero et al., 2021). In contrast, PE value will be low if few harmonic components are affected. This fact will reveal that the signal is less random and has a more regular phase in the bispectral band (Martín-Montero et al., 2021).

- Features based on bispectral moments:
  - Sum of logarithmic bispectral magnitudes inside adaptive band (H_1), sum of logarithmic bispectral magnitudes contained in the adaptive band diagonal (H_2), first and second order spectral moments of bispectral magnitudes contained in the adaptive band diagonal (H_3 and H_4, respectively). They quantify the phase coupling and provide information about the signal non-linearity in the region under study (Martín-Montero et al., 2021). Thus, strong non-linear interactions between the frequency components of the signal will lead to greater phase coupling and therefore high values of these features (Atri and Mohebbi, 2015). As already mentioned, they are calculated according to the bispectral content of the adaptive band and its diagonal AB_{diag} (Atri and Mohebbi, 2015; Zhou et al., 2008):

$$H_1 = \sum\nolimits_{f_1, f_2 \in AB} log(|B_N(f_1, f_2)|). \tag{4.30}$$

$$H_2 = \sum\nolimits_{f_k \in AB_{diag}} log(|B_N(f_k, f_k)|). \tag{4.31}$$

$$H_3 = \sum\nolimits_{f_k \in AB_{diag}} k \cdot log(|B_N(f_k, f_k)|). \tag{4.32}$$

$$H_4 = \sum\nolimits_{f_k \in AB_{diag}} (k - H_3)^2 \cdot log(|B_N(f_k, f_k)|). \tag{4.33}$$

- Features based on bispectral weighted center:
  - Components fm_1 and fm_2 of the weighted center of bispectrum (WCOB). fm_1 and fm_2 are 2 vectors that allow locating the bispectral coupling focus within the region under study (Wang et al., 2014; Zhang et al., 2000). These are obtained by means of the following equations (Wang et al., 2014; Zhang et al., 2000):

$$fm_1 = \frac{\sum_{f_1,f_2 \in AB} f_1 \cdot B_N(f_1, f_2)}{\sum_{f_1,f_2 \in AB} B_N(f_1, f_2)}. \tag{4.34}$$

$$fm_2 = \frac{\sum_{f_1,f_2 \in AB} f_2 \cdot B_N(f_1, f_2)}{\sum_{f_1,f_2 \in AB} B_N(f_1, f_2)}. \tag{4.35}$$

    The WCOB features provide information about the frequency interactions that occur between the signal components (Wang et al., 2014). Thus, low fm_1 and/or fm_2 values will indicate that there is greater activity at low frequencies, while high values of them will reveal an activity shift towards higher frequencies (Martín-Montero et al., 2021; Wang et al., 2014).

4.2.5 Wavelet features

Wavelet decomposition is a multi-resolution analysis method that provides information about the signal behavior at different frequency ranges (Rioul and Vetterli, 1991). Although classical spectral analysis methods also provide information at frequency domain, these are limited by using a fixed spectral and temporal resolution that could hinder the low frequency scans (Rioul and Vetterli, 1991; Vaquerizo-Villar et al., 2018c). Moreover, they take the signal stationarity for granted. In contrast, wavelet transform performs a signal decomposition at different resolutions, providing high spectral resolution at low frequencies while offering high temporal resolution at high frequencies (Figliola and Serrano, 1997; Rioul and Vetterli, 1991). Note that this fact is particularly important in the present thesis context, since OSA-elicited alterations are reflected in low frequency bands of AF signal (Gutiérrez-Tobal et al., 2015; Jiménez-García et al., 2020). In addition, wavelet transform can be applied to non-stationary signals, making it a suitable analysis method for physiological recordings (Figliola and Serrano, 1997; Rosso et al., 2006).

Wavelet transform can be obtained in a discrete or continuous way (Rioul and Vetterli, 1991; Rosso et al., 2006). In this study, discrete wavelet transform (DWT) was used due to its higher computational efficiency compared to continuous wavelet transform (CWT) (Rioul and Vetterli, 1991). As can be seen in Figure 4.5, the algorithm calculates the DWT of N-sample input signal $(x(n), n \in N)$ by recursively applying high-pass $(g(n))$ and low-pass $(h(n))$ filters, followed by a dyadic sampling (Mallat, 1989; Rioul and Vetterli, 1991). These filters are based on scaling and translation of the basis functions. The basis functions are the mother wavelet and its mirror version, which correspond to $g(n)$ and $h(n)$ respectively, such that (Mallat, 1989; Rioul and Vetterli, 1991):

$$g(n) = (-1)^{1-n} \cdot h(1-n). \tag{4.36}$$

Thereby, two signals are obtained at each decomposition level j: detail signal D_j and approximation signal A_j. The coefficients of D_j are obtained through $g(n)$, while those of A_j are obtained by means of $h(n)$ (Rioul and Vetterli, 1991; Vaquerizo-Villar et al., 2018c):

$$D_j(k) = \sum\nolimits_n A_{j-1}(n) \cdot g(2k-n), \tag{4.37}$$

$$A_j(k) = \sum\nolimits_n A_{j-1}(n) \cdot h(2k-n), \tag{4.38}$$

where $k \in Z$ is the position of each coefficient within the detail/approximation signal of level j. Afterwards, A_j signal is decomposed back into a new range of high and low frequencies, giving rise to detail and approximation signals of the next level (D_{j+1} and A_{j+1}). Thus, the process continues until the maximum decomposition level is reached: $M = log_2(N)$.

In this study, DWT was applied to 2^{16}-sample segments of AF signal (Jiménez-García et al., 2020). Therefore, the highest reached decomposition level was $M = 16$. The mother wavelets used were Haar and Daubechies-5. Both wavelets have already been successfully tested on respiratory signals (Kermit et al., 2000; Lee et al., 2011). Some motivations for using Haar mother wavelet are that: (i) it is the simplest basis function, (ii) it does not elicit edge effect, and (iii) its stepped shape could help detect sudden changes in AF (Gogolewski, 2020; Hadaś-Dyduch, 2018). Regarding Daubechies-5 mother wavelet, it was chosen due to: (i) it provides a time-frequency localization of high precision, (ii) it causes a reduced edge effect compared to other higher order Daubechies wavelets, and (iii) its wavy shape could

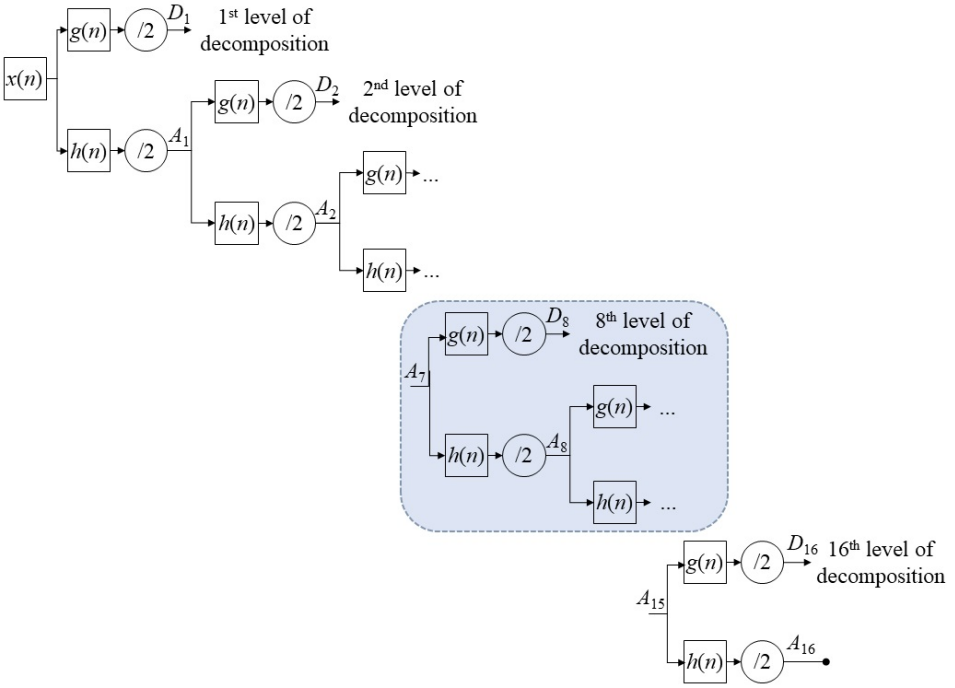


Figure 4.5: Discrete wavelet decomposition process. Figure taken from Barroso-García et al. (2021b).

better fit the AF behavior (Daubechies, 1990, 1992; Gogolewski, 2020). Figure 4.6 shows an example of DWT applying Daubechies-5 mother wavelet to a 10-min segment of AF signal. As can be observed, the 8^{th} detail level covers the frequency range 0.1953–0.3906 Hz. This frequency band corresponds to the usual duration of children's respiratory cycles, i.e., to normal breathing rates (Fleming et al., 2011). Thus, the wavelet features were particularly extracted from D_8 to characterize the detail signal associated to the normal breathing band. In addition, the wavelet entropy was obtained from the full outline to have a global view of the AF behavior.

- Features derived from a particular decomposition level:
  - First to fourth order statistical moments (M_{1D_j}–M_{4D_j}). These features correspond to mean, standard deviation, skewness and kurtosis of the D_j detail level coefficients (M_{1D_j}, M_{2D_j}, M_{3D_j}, and M_{4D_j}, respectively). Thus, they allow quantifying central tendency, variation, asymmetry, and tailedness of its distribution (Vaquerizo-Villar et al., 2018c).

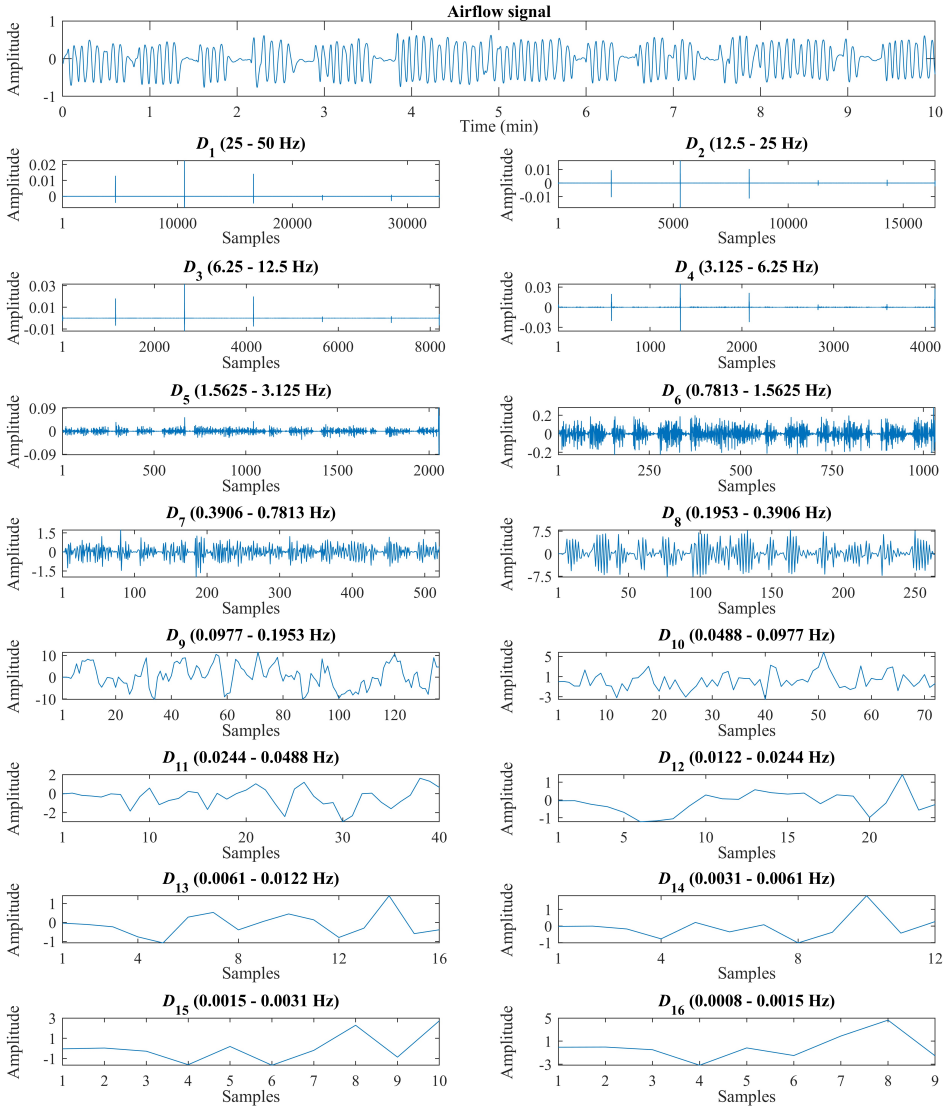


Figure 4.6: Wavelet decomposition process applied to a 10-min segment of airflow. Figure obtained from Barroso-García et al. (2021b).

– Maximum and minimum (Max_{D_j} and Min_{D_j}). These features correspond to maximum and minimum value achieved by D_j detail level coefficients. Thereby, they provide information about the highest and lowest amplitude obtained by the signal components associated to this level. Thus, Max_{D_j} and Min_{D_j} are calculated as follows (Vaquerizo-Villar et al., 2018c):

$$Max_{D_j} = \max[D_j(k)]. \tag{4.39}$$

$$Min_{D_j} = \min[D_j(k)]. \tag{4.40}$$

- Energy (E_{D_j}). Based on the quadratic amplitudes of D_j, this feature quantifies the activity level producted in the frequency component range associated to each detail level (Rosso et al., 2001; Vaquerizo-Villar et al., 2018c). Therefore, a detail level with high E_{D_j} values will reveal that the signal has a high activity in the encompassed frequency band. In order to compute E_{D_j}, the following equation is used (Figliola and Serrano, 1997; Rosso et al., 2006):

$$E_{D_j} = \sum\nolimits_k |D_j(k)|^2, \tag{4.41}$$

- Features derived from the full wavelet outline:

  - Wavelet entropy (WE). Wavelet entropy is a measure that quantifies the irregularity of a signal by measuring the uniformity of its energy distribution in the different decomposition levels (Figliola and Serrano, 1997; Rosso et al., 2001). Thereby, uniformly distributed energy in a wide level range will reach high WE values, indicating a high signal irregularity. In contrast, WE will present low values if signal energy is concentrated in only a few decomposition levels. This fact will reveal that the signal has a more regular behavior. Thus, this feature can be obtained as a variant of the Shannon's entropy by means of the following equation (Rosso et al., 2001; Vaquerizo-Villar et al., 2018c):

$$WE = -\sum\nolimits_{j=1}^{M} p_j \cdot log(p_j), \tag{4.42}$$

    where M is the maximum decomposition level, and p is the energy distribution computed as:

$$p_j = \frac{E_{D_j}}{\sum_{j=1}^{N} E_{D_j}}. \tag{4.43}$$

4.3 Feature selection

A feature selection stage was developed after feature extraction in order to obtain a more compact and explanatory representation of the extracted data, as well as to reduce the input set to classifiers. In this regard, the selection techniques allow us (Guyon and Elisseeff, 2003): (i) improving understanding of feature extraction processes, (ii) obtaining computationally faster and more efficient classifiers, and (iii) improving its prediction ability.

Note that extracting a large number of features does not ensure right pattern recognition results. The collected features could be redundant, thus they would not provide additional information when combined (Guyon and Elisseeff, 2003). It is also possible that some features are not relevant to solve the addressed problem, thus negatively affecting predictability (Guyon and Elisseeff, 2003). A large number of features can even cause the classifier to overfit, which reduces its predictability (Guyon and Elisseeff, 2003; Saeys et al., 2007). Feature selection techniques provide an optimal subset to characterize a certain model (Saeys et al., 2007). Hence the need and importance of performing this selection stage prior to pattern recognition.

It is important to point out that the feature selection methods do not apply any data transformation, they only choose a reduced subset from the input features (Saeys et al., 2007). According to how it works, three technical categories can be differentiated (Saeys et al., 2007): (i) filter methods, which are independent of the predictive model, (ii) wrapper methods, that depend on the predictive model, and (iii) embedded methods, which are dependent on the model and both are jointly constructed. In this regard, we have been implemented and automatically applied forward stepwise logistic regression wrapper method (FSLR) and fast correlation-based filter method (FCBF). Both FSLR and FCBF selection methods have already been shown to be particularly useful in the diagnosis of pediatric OSA (Álvarez et al., 2018; Gutiérrez-Tobal et al., 2015; Hornero et al., 2017; Jiménez-García et al., 2020). Each of them is detailed below.

4.3.1 Forward stepwise logistic regression

Proposed by Hosmer and Lemeshow (2002), FSLR feature selection algorithm is a wrapper method that uses logistic regression (LR) models to obtain a optimum subset of features. Its approach is based on evaluating the LR models built in each algorithm step by including features that contribute with relevant information (forward selection), and excluding those that provide redundant information (back-

ward elimination) (Gutiérrez-Tobal et al., 2015; Hosmer and Lemeshow, 2002). In this regard, statistical significance is used to determine which features should remain in the model and which ones should not. According to Hosmer and Lemeshow (2002), this parameter is defined based on the p-value of the likelihood-ratio χ^2 test that is applied in each algorithm step.

Thereby, the algorithm starts from a null LR model, without features, only with the constant term. Then, a forward selection is performed by computing the significance of each of the N features that were extracted in the previous stage. Thus, their corresponding N LR models are built and compared with the initial model through the likelihood-ratio test. The feature that obtains the highest significance (i.e. lowest p-value) is included in the model if its p-value is less than a certain input threshold p_{input}. Subsequently, the forward search is performanced yet again with the rest of the features, their corresponding LR models are built, and these are compared with the previous model. In addition to the forward selection, a backward elimination process is conducted to test that the M features that constitute the model are still statically significant after each new entry. Hence, M LR models are built and compared with the previous model through the likelihood-ratio test. The feature with lowest significance (i.e. highest p-value) is excluded from the model if its p-value is greater than a certain output threshold p_{output}. The process continues until there are no features to select or until none of them meets the input and/or output criteria from the model. Note that the condition $p_{input} < p_{output}$ should be met to prevent the algorithm from running infinitely. Moreover, they should be carefully chosen. In this regard, important features could stay out of the model if these thresholds are overly restrictive (a too low value of p_{input} and/or p_{output}). In contrast, if p_{input} and p_{output} are inappropriately high, almost all features will enter the model and none will be removed (Hosmer and Lemeshow, 2002).

4.3.2 Fast correlation-based filter

The FCBF algorithm is one of the most popular feature selection methods in the context of pediatric OSA diagnosis (Hornero et al., 2017; Jiménez-García et al., 2020; Martín-Montero et al., 2021; Vaquerizo-Villar et al., 2018a,b,c). It is a filter method that uses the symmetrical uncertainty (SU) to obtain an optimum subset of features (Yu and Liu, 2004). The approach of this algorithm is based on first selecting a subset with the most relevant features and then eliminating from it those that are redundant with respect to other more relevant features. Therefore,

the optimun subset is only composed of relevant and independent features (Yu and Liu, 2004).

In order to determine which features should remain in the model and which ones should not, SU is used. This is a non-linear correlation measure defined by the entropy of the features. Thus, SU can be computed as follows (Yu and Liu, 2004):

$$SU(X_i|X_j) = 2 \cdot \frac{H(X_i) - H(X_i|X_j)}{H(X_i) + H(X_j)}, \tag{4.44}$$

where $H(X_i)$, $H(X_j)$, and $H(X_i|X_j)$ are the Shannon's entropies of the features X_i and X_j:

$$H(X_i) = -\sum\nolimits_{x_i \in X_i} p(x_i) \cdot log(p(x_i)). \tag{4.45}$$

$$H(X_j) = -\sum\nolimits_{x_j \in X_j} p(x_j) \cdot log(p(x_j)). \tag{4.46}$$

$$H(X_i|X_j) = -\sum\nolimits_{x_j \in X_j} p(x_j) \cdot \sum\nolimits_{x_i \in X_i} p(x_i|x_j) \cdot log(p(x_i|x_j)). \tag{4.47}$$

Thereby, FCBF calculates the SU between the variable V to be predicted and each of the N extracted features (i.e. $SU(X_i|V), i \in N$) to quantify its relevance in the first step of the algorithm (Yu and Liu, 2004). The feature X_i that obtain a $SU(X_i|V)$ greater than a certain input threshold p_{input} is considered relevant and included in a preliminary subset. Afterwards, FCBF computes the SU between each pair of features that constitute this subset (i.e. $SU(X_i|X_j), i, j \in N$) to evaluate its redundancy in a second step of the algorithm (Yu and Liu, 2004). The feature X_i that has a higher SU with X_j than with V ($SU(X_i|X_j) \geq SU(X_i|V)$) and that is also less relevant than X_j ($SU(X_j|V) \geq SU(X_i|V)$) is considered redundant and excluded from the final optimum subset.

4.4 Machine-learning approaches

As stated above, this work is also encompassed in the machine-learning field. It is a branch of artificial intelligence focused on developing computer algorithms capable of automatically identifying complex patterns in data to make predictions from them (Alpaydin, 2014; Bishop, 2006). There are different machine-learning approaches: supervised learning, unsupervised learning, and reinforcement learning (Alpaydin, 2014). Due to our original dataset is labeled, supervised machine-learning methods have been applied in the Doctoral Thesis to recognize patterns in AF-extracted features and use this information to automatically detect the presence and severity of OSA in children. Thus, this stage was conducted

from three different approaches: discrimination between OSA-negative and OSA-positive pediatric subjects (binary classification task), classification of children according to their OSA severity degree (multiclass classification task), and AHI estimation of each child (regression task). The binary and multiclass classification tasks were carried out by means of logistic regression (LR) and adaptive boosting (AdaBoost.M2) algorithms, respectively. Regarding the regression task, it was performed through a multi-layer perceptron neural network (MLP) and a MLP with Bayesian approach (BY-MLP). Each of these methods is presented below.

4.4.1 Binary classification: Logistic regression

Regression methods model how an output dependent variable changes when one or more input independent variables (i.e., predictor variables) are modified (Hosmer and Lemeshow, 2002). Depending on the nature of this dependent variable, some models or others can be used. In this regard, linear regression models are applied when the dependent variable is continuous, while LR models are widely used when it is dichotomous (Hosmer and Lemeshow, 2002). The latter allow estimating the posterior probability $p(C_j|x_i)$ that one of the two mutually exclusive options contemplated in the dependent variable $(C_j, j = 1, 2)$ occurs as a function of predictor variable value x_i. Due to this, LR is used as binary classifier (Hosmer and Lemeshow, 2002). In this Doctoral Thesis, LR was applied to discriminate OSA-negative from OSA-positive subjects (our dependent variable) based on AF features obtained in the selection stage (our predictor variables). Thus, the logistic function was used (Hosmer and Lemeshow, 2002):

$$p(C_j|x_i) = \frac{e^{\beta_0+\sum_{i=1}^{N}\beta_i\cdot x_i}}{1+e^{\beta_0+\sum_{i=1}^{N}\beta_i\cdot x_i}}, \tag{4.48}$$

where N is the number of predictor variables x_i, and β_0 and β_i are the model coefficients. Note that each new predictor variable x_i that enters into the logistic function increases the number of required coefficients β_i. Hence, these coefficients should be defined and fitted in order to obtain a suitable model. In this regard, an estimator is used for this purpose (Hosmer and Lemeshow, 2002). The least squares estimator is used in classical linear regression models. However, some of its statistical properties are not met when the output dependent variable is dichotomous (Hosmer and Lemeshow, 2002). Consequently, LR models use maximum likelihood as an estimator, which allows us to find values of β_i by maximizing the probability of obtaining the observed value set (Hosmer and Lemeshow, 2002).

4.4.2 Multi-class classification: Adaptive Boosting

Adaptive Boosting (AdaBoost) is an ensemble learning method used for classification tasks (Freund and Schapire, 1997; Witten et al., 2011). In the particular case of multiclass classification, it is performed by means of the AdaBoost.M2 algorithm. This method has already been shown to be particularly useful with AF signal in the diagnosis of OSA (Gutiérrez-Tobal et al., 2016; Jiménez-García et al., 2020). It is based on building a strong classifier by combining several weak classifiers or learners (Freund and Schapire, 1997; Witten et al., 2011). These weak classifiers are simple and of the same type. In this way, a single weak classifier is trained in each algorithm iteration. AdaBoost initially assigns the same weight to all observations. Then, the first weak classifier makes its prediction. Those correctly classified observations are updated with a weight decrease. In contrast, those that have been misclassified by the first learner are identified and penalized by assigning them a greater weight. In the second iteration, the next weak classifier will focus on correcting the observations that its predecessor has erroneously predicted (i.e., the ones with the greatest weight). This fact allows the new weak classifier to adapt to the observations, thus minimizing its pseudo-loss ε_t (Freund and Schapire, 1997; Witten et al., 2011). Being an iterative process, the algorithm continues adding learners until reaching the established maximum number L. The final classification prediction $H(x)$ will be made based on the weighted vote of the predictions $h_t(x)$ obtained by the weak classifiers (Freund and Schapire, 1997):

$$H(x) = \arg\max \sum_{t=1}^{L} \log(1/\beta_t) \cdot h_t(x), \tag{4.49}$$

where L is the number of weak classifiers and β_t is the weight update coefficient of the weak classifier $t \in L$, which depend on the pseudo-loss ε_t and the learning rate α required to minimize overfitting:

$$\beta_t = \left(\frac{\varepsilon_t}{1 - \varepsilon_t} \right)^{\alpha}. \tag{4.50}$$

Thus, AdaBoost allows to create a strong classifier much more robust and precise than the weak classifiers used in the process. Due to decision trees are widely used as weak classifiers by boosting methods (Witten et al., 2011), they were applied in the present Doctoral Thesis. Note that L and α are hyperparameters that should be adjusted. In this regard, their values were chosen as those that maximize the Cohen's kappa coefficient (k) estimated by 0.632 bootstrap procedure in 1000 replicates from training dataset (Witten et al., 2011).

4.4.3 Regression: Multi-layer perceptron neural network

Artificial neural networks (ANNs) are methods that consist of a large number of interconnected processing units called neurons (Jain et al., 2000). They have the ability to learn complex and non-linear input-output relationships, thus adapting to the intrinsic properties of input data. The ANNs most used in pattern recognition are feed-forward networks, among which are multi-layer perceptron (MLP) networks (Jain et al., 2000). Hence, a MLP is an ANN whose neurons are organized in layers and where its connections are forward unidirectional, no feedback loops. In this regard, three-layer MLP networks (i.e. an input layer, a hidden layer, and an output layer) are capable of providing a universal approximation (Bishop, 1995). Consequenly, this network architecture was implemented throughout this Doctoral Thesis. As Figure 4.7 shows, the first layer of MLP is the input layer. This is composed of N_I neurons, as many features as they have been selected in the

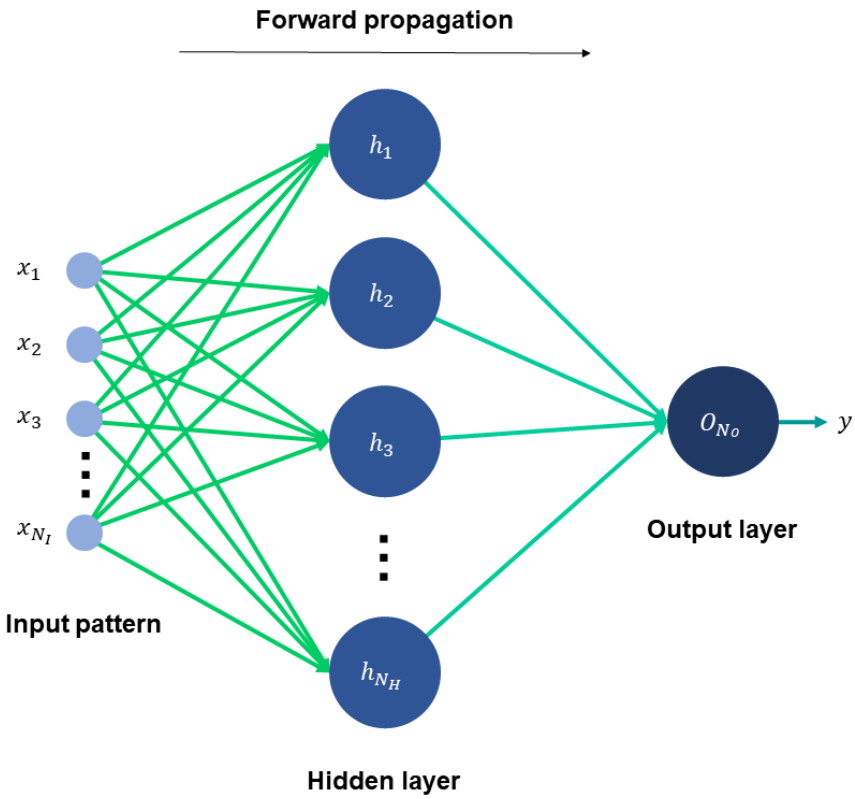


Figure 4.7: Generic MLP network architecture with N_I neurons in the input layer, a single hidden layer with N_H neurons, and one neuron in the output layer ($N_O = 1$).

previous selection stage. Thus, each neuron receives a input feature and spreads it to the next layer neurons (Bishop, 1995). The next layer of MLP is the hidden layer. This is composed of a number of neurons N_H that should experimentally be tuned. Each of its neurons receives as input a linear combination of the previous layer neuron outputs and provides a non-linear function of it as output (Bishop, 1995). Finally, the last layer of MLP is the output layer. Since MLP is intended to estimate the AHI of pediatric subjects, a single neuron is required in this layer. Then, this neuron receives as input the hidden layer neuron outputs and provides a linear function of it as MLP output y (Bishop, 1995):

$$y = g_{out}(\sum_{j=1}^{N_H} w_{jo} \cdot g_{hidden}(\sum_{i=1}^{N_I} w_{ij} \cdot x_i + b_j) + b_{out}), \tag{4.51}$$

where x_i is the input feature i, w_{ij} is the connection weight between the input neuron i and the hidden neuron j, w_{jo} is the connection weight between the hidden neuron j and the single output neuron, b_j is the bias associated to the hidden neuron j, b_{out} is the bias associated to the single output neuron, and g_{hidden} and g_{out} are the activation function of the hidden and output neurons, respectively.

Note that weights influence the MLP learning process, thus they should be carefully adjusted. In this regard, two different approaches were applied in the studies that this Doctoral Thesis encompasses:

- **Classic MLP**. Initially, weights are randomly assigned. Afterwards, they are adjusted by jointly using scaled conjugate gradient to minimize the error function and weight decay regularization to avoid overfitting (Bishop, 1995). This last method adds a regulatory parameter α that penalizes the use of high weights at the neuron input. Like N_H, α should also be adjusted in the training set.

- **MLP with Bayesian approach (BY-MLP)**. In this case, an alternative method is used to set the weights. It is based on modeling the distribution function of the weight space instead of finding an optimum value for them (Bishop, 1995). Thus, the posterior probability is computed and the network weights are adjust by means of Bayesian inference. This technique is self-regulated, so it does not introduce any regularization parameters and only N_H should be tuned (Bishop, 1995).

Regarding the hyperparameter optimization, leave-one-out cross-validation (loo-cv) method was applied to the training dataset and the averaged Cohen's kappa coefficient (k) was obtained for each N_H value in Barroso-García et al.

(2020). The N_H that obtained the highest k was chosen as optimum. A similar process was applied to jointly optimize N_H and α in Barroso-García et al. (2021a). Loo-cv method was also used and the maximum k determined the optimum values of N_H-α pair. In Barroso-García et al. (2021b), N_H was chosen as the value that maximized the k estimated by 0.632 bootstrap procedure in 1000 replicates from training dataset (Witten et al., 2011).

4.5 Statistical analysis

The analysis and interpretation of the obtained results, as well as the performance and reliability evaluation of the used methods, is a key issue in any research study. In order to address these essential tasks, suitable statistical tests, diagnosis performance measures, agreement measures, and validation methods were applied in the present Doctoral Thesis. Thus, the techniques used in each case are presented below.

4.5.1 Statistical tests

Statistical tests are tools for accepting or rejecting probabilistic modeling hypotheses. Thereby, statistical tests were used in this research work to assess whether: (i) the data under study come from a normal distribution, (ii) its population variances are equal (homoscedasticity), and (iii) the distribution of independent variables is the same.

Firstly, the Lilliefors test (Lilliefors, 1967) was applied to evaluate the normality of the clinical and sociodemographic data of the subjects under study, as well as the normality of each feature extracted from AF signal. These data were also subjected to the Leneve test to assess their homoscedasticity (Levene, 1961). It was observed that their distributions were neither normal nor homoscedastic. Therefore, the clinical and sociodemographic data of the subjects belonging to the training set were compared with those of the test set using the non-parametric Mann-Whitney U, χ^2, and Fisher's exact tests (Fisher, 1922; Mann and Whitney, 1947). In this regard, the Mann-Whitney U test was applied to continuous variables (age, BMI, and AHI), while the χ^2 test and Fisher's exact test were applied to categorical variables (gender and OSA severity groups). In order to determine whether there were statistically significant differences between the training and test data, the significance level $\alpha = 0.01$ was established. Moreover, we evaluate possible differences between the severity groups (i.e. non-OSA, mild, moderate, and

severe) in AF-derived features. Thus, the Mann-Whitney U test was applied for pairwise comparisons between gropus and the Kruskal-Wallis test for comparisons among 4 groups (Kruskal and Wallis, 1952; Mann and Whitney, 1947). According to the number of subjects involved in these comparisons, the significance levels $\alpha = 0.01$ and $\alpha = 0.05$ were established. Note that the probability of making a type I error by chance not only depends on the α value. This probability noticeably increases when the number of performed tests (T) increases: $1-(1-\alpha)^T$ (Narsky and Porter, 2013). Consequently, p-values should be corrected. In this Thesis, the correction was carried out using the Bonferroni method, that is, multiplying the p-value by the number of applied tests or, alternatively, α/T (Farcomeni, 2008). Therefore, those features with a p-value $< \alpha$ after Bonferroni correction were considered to have statistically significant differences between severity groups. In the latest studies (Barroso-García et al., 2021a,b), it was also assessed whether there were differences in any performance and/or agreement metric between the different predictive models that were applied to the test set. In this case, the Mann-Whitney U test for pairwise comparisons with Bonferroni correction was applied. Due to the large number of values that were compared (1000 samples of each metric after applying bootstrapping, see subsection 4.5.4), 0.001 was used as significance level to minimize the probability of making a type I error.

In order to complement the information provided by the statistical tests, other techniques were also applied. In this regard, the Spearman's correlation coefficient was used to evaluate the relationship between variables (Zar, 1972), particularly between the features extracted from AF and the AHI. Thus, this tool allowed us to detect those that had the highest correlation with the AHI from the subjects under study. This information was also very useful to optimize some methods developed and applied in this Thesis: (i) radius value of the CTM, (ii) fixed distance threshold value used to compute the RPs, and (iii) type of wavelet mother used to performance the multi-resolution analysis of AF signal. In addition, box and violin plots were used to analyze the distribution and tendencies of the extracted features (Hintze and Nelson, 1998). Box plots show the median, 25^{th} and 75^{th} percentiles, maximum and minimum, as well as outliers of the data distribution. The violin plots provide addicional information about the data arrangement by showing the density of its distribution. Hence, the use of these diagrams allowed us to obtain a visual summary of the distribution of AF features and to complete the information provided by statistical tests.

4.5.2 Diagnostic performance measures

In order to evaluate the usefulness of the AF features, as well of the models built from them, several diagnostic performance measures are computed. These metrics are derived from the number of subjects correctly and/or wrongly classified. Taking into account that the OSA severity degree of pediatric subjects is considered in this Thesis according to the AHI thresholds 1, 5, and 10 e/h, the following terms are defined based on these cut-off points (Fawcett, 2006):

- **True Positives (TP)**. Subjects with an AHI≥cut-off who are classified as ≥cut-off by the proposed model (i.e. positives correctly classified).
- **True Negatives (TN)**. Subjects with an AHI<cut-off who are classified as <cut-off by the proposed model (i.e. negatives correctly classified).
- **False Positives (FP)**. Subjects with an AHI<cut-off who are classified as ≥cut-off by the proposed model (i.e. negatives wrongly classified).
- **False Negatives (FN)**. Subjects with an AHI≥cut-off who are classified as <cut-off by the proposed model (i.e. positives wrongly classified).

According to these terms, the following diagnostic performance metrics are calculated (Fawcett, 2006; Flemons and Littner, 2003; Zweig and Campbell, 1993):

- **Sensitivity (Se)**. Rate of subjects with AHI≥cut-off correctly classified:

$$Se = \frac{TP}{TP + FN} \cdot 100. \tag{4.52}$$

- **Specificity (Sp)**. Rate of subjects with AHI<cut-off correctly classified:

$$Sp = \frac{TN}{TN + FP} \cdot 100. \tag{4.53}$$

- **Accuracy (Acc)**. Rate of subjects with an AHI<cut-off or AHI≥cut-off correctly classified:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100. \tag{4.54}$$

- **Positive predictive value (PPV)**. Probability that a subject classified as ≥cut-off is actually a subject with an AHI≥cut-off:

$$PPV = \frac{TP}{TP + FP} \cdot 100. \tag{4.55}$$

- **Negative predictive value (NPV)**. Probability that a subject classified as <cut-off is actually a subject with an AHI<cut-off:

$$NPV = \frac{TN}{TN + FN} \cdot 100. \tag{4.56}$$

- **Positive likelihood ratio (LR+)**. Rate of subjects with AHI≥cut-off correctly classified with respect to the rate of subjects with AHI<cut-off wrongly classiffied:

$$LR+ = \frac{Se}{100 - Sp}. \tag{4.57}$$

- **Negative likelihood ratio (LR-)**. Rate of subjects with AHI≥cut-off wrongly classified with respect to the rate of subjects with AHI<cut-off correctly classiffied:

$$LR- = \frac{100 - Se}{Sp}. \tag{4.58}$$

- **Area under the Receiver Operating Characteristic curve (AUC)**. A Receiver Operating Characteristic (ROC) curve is a parametric curve in the XY plane representing 100−Sp versus Se (Zweig and Campbell, 1993). The point [(100−Sp) = 0, Se = 100] of this curve represents the ideal classification in which FP = 0 and FN = 0, while Se, Sp, Acc, PPV, and NPV reach their maximum value 100. Thereby, higher classification results will be obtained as the ROC curve is closer to this ideal classification point (Fawcett, 2006; Zweig and Campbell, 1993). One way to summarize the information provided by the ROC curves is to use AUC measure, i.e., the area enclosed between the ROC curve and the abscissa axis (Fawcett, 2006; Zweig and Campbell, 1993). Thus, AUC values close to 1 will reveal that the predictive model is highly effective.

4.5.3 Agreement measures

In addition to the diagnostic performance measures, global metrics are required to evaluate the agreement between observers and determine the reliability of the proposal. Thus, we need measures that allow evaluating the agreement between the actual AHI (i.e. the one derived from the PSG) and the AHI estimated by regression models, as well as the agreement between the severity group diagnosed by medical specialists and the one predicted by binary or multiclass classification models. In this regard, the suitable way to approach the issue depends closely on

the data nature. If the data are continuous, the intraclass correlation coefficient is commonly used as agreement measure. However, when the data are categorical, the most applied metrics are the multiclass precision and the Cohen's kappa index. Each of these measures is detailed below:

- **Intraclass correlation coefficient (ICC)**. It is a global measure for evaluating the reliability of regression models (Chen and Barnhart, 2008). Unlike other conventional correlation coefficients, the notion underlying in the ICC formulation is to consider systematic error (Bartko, 1966; Chen and Barnhart, 2008). This fact prevents the correlation from being perfect when two observers systematically measure different quantities from each other and the agreement is null. Moreover, ICC estimates the average of the correlations between all possible orderings by pairs of observations (Bartko, 1966). Therefore, this avoids the order dependency limitation that arises with other approaches such as Pearson's. A common way to calculate the ICC is without making suppositions about the analysis of variance (ANOVA) or using repeated measures (Chen and Barnhart, 2008). The basis idea is that the total variability of measurements can be decomposed into two components: the variability due to differences between instances, and the variability caused by differences between observers. The latter, in turn, depends on the variability between observations and a residual variability associated to the error that any measurement entails. Thus, ICC can be computed as follows (Chen and Barnhart, 2008):

$$ICC = \frac{MS_I - MS_E}{MS_I - (O-1) \cdot MS_E + O \cdot (MS_O - MS_E)/N}, \tag{4.59}$$

  where N is the total number of instances, O is the number of observers, MS_I is the instances mean square, MS_E is the error mean square, and MS_O is the observers mean square. Thus, ICC values can range from 0 to 1, being an ICC = 1 the maximum possible agreement (Weir, 2005). In this case, all observed variability would be explained by the differences between subjects and not by the differences between the estimation methods or the different observers. In contrast, a value ICC = 0 is obtained when the agreement occurs only by chance (Weir, 2005).

- **Multiclass accuracy ($\mathbf{Acc_M}$)**. It is a global measure for evaluating the reliability of multiclass classification models. In multiclass prediction, the result is often presented in a 2-dimensional array known as confusion matrix

(Witten et al., 2011). This matrix has as many rows and columns as number of categories to classify (M). Each matrix element n_{ij} shows the number of instances classified as the class indicated in row i by a observer (actual class) and as the class indicated in column j according to other observer (predicted class) (Witten et al., 2011). Extending the definition of 2-class accuracy previously indicated (subsection 4.5.2), multiclass accuracy would be the probability that both observers agree on the assigned class. Thus, Acc_M corresponds to the percentage of instances of each class correctly classified and it is computed based on the main diagonal elements of confusion matrix:

$$Acc_M = \frac{\sum_{i=1}^{M} n_{i,i}}{N} \cdot 100, \tag{4.60}$$

where M is the number of class and N is the total number of instances.

- **Cohen's Kappa index (k).** It is a global measure for evaluating the reliability of binary and multiclass classification models (Cohen, 1960). k quantifies the agreement between two observers in their corresponding classifications of N instances in M mutually exclusive categories. Thereby, its computation is based on the probability that the classification made by both observers agrees ($P(A)$), and the hypothetical probability that the agreement occurs by chance ($P(C)$) (Cohen, 1960):

$$k = \frac{P(A) - P(C)}{1 - P(C)}, \tag{4.61}$$

where $P(A)$ corresponds to the proportion of instances of each class correctly classified:

$$P(A) = \frac{\sum_{i=1}^{M} n_{i,i}}{N}, \tag{4.62}$$

and $P(C)$ includes the marginal distributions in its formula for be able to exclude the agreement happened just by chance when computating k:

$$P(C) = \sum_{i=1}^{M} \left(\frac{\sum_{j=1}^{M} n_{i,j}}{N} \cdot \frac{\sum_{j=1}^{M} n_{j,i}}{N} \right). \tag{4.63}$$

Thus, the maximum agreement corresponds to $k = 1$, while the maximum disagreement corresponds to $k = -1$. In the same way, a value $k = 0$ is obtained when the agreement is caused only by chance (Cohen, 1960).

4.5.4 Validation methods

In order to check the validity of the proposed methods and the results obtained with them, several validation techniques have been used throughout this Doctoral Thesis. Firstly, hold-out validation method was applied to obtain the training and test sets. Depending on the study, the training set was used to optimize parameters of extraction methods (e.g. radius value of CTM, computation parameters of RPs, wavelet type to perform the multi-resolution analysis), selected feature subsets, and hyperparameters of predictive models (e.g. number of hidden layer perceptrons of BY-MLP, number of hidden layer perceptrons and regularization parameter of MLP, learning rate and number of weak classifiers of AdaBoost.M2), as well as to train predictive models. So as to minimize overfitting and to minimize the bias produced by using the same dataset for parameter optimization and predictive model training, leave-one-out cross-validation (loo-cv) and bootstrapping procedures were also applied to the training set (Witten et al., 2011). In addition, k-fold cross-validation and bootstrapping methods were used in the test set to ensure that the results of the predictive models are independent of the training-test partition and thus increase its generalizability (Bishop, 2006). Each of the aforementioned techniques is detailed below:

- **Hold-out**. This method consists of dividing the original dataset into two independent subsets: training and testing (Bishop, 2006; Witten et al., 2011). In this regard, the training set is used to build the proposed model, while the test set is used to evaluate its performance. Hold-out is a computationally fast validation method, but it depends a lot on how the partition is done and the number of subjects used for it (Refaeilzadeh et al., 2016; Witten et al., 2011). Thus, the data should be randomly divided into independent sets. Consequently, statistical tests were applied in this Thesis to ensure that the training and test samples were completely independent and that there were no statistically significant differences (p-value ≥ 0.01) in any sociodemographic or clinical characteristics from the subjects that compose each group. Regarding the size of the subsets, this is an arbitrary issue, although usually the number of subjects belonging to the test set is less than or equal to the number of subjects belonging to the training set. This is because the training set should be large enough to be representative and to allow a suitable fitting of the models (Witten et al., 2011). According to each of our studies, the training set was used to optimize parameters of extraction, selection and pattern recognition stages, as well as to train predictive models.

However, the amount of available data did not allow us to obtain a subset for each of these tasks with a sufficiently representative size. Consequently, hold-out method was combined with additional validation methods to avoid overfitting and to minimize the bias produced by using the same subset of data for several tasks (Witten et al., 2011).

- **Leave-one-out cross-validation (loo-cv)**. It is an iterative method that uses as training set all available observations except one, which is used to test the model (Bishop, 2006; Witten et al., 2011). If a single observation is used to measure the performance and/or agreement if the model, this measure could significantly vary depending on which observation is selected. To avoid this, the process is repeated N times, as many as available observations. Thereby, a different observation is excluded in each iteration, the model is trained with the rest, and then it is tested with the observation that had been excluded (Bishop, 2006; Witten et al., 2011). At the end of the process, we will have N estimates of each performance and/or agreement measure. Its estimated final value will be computed as the average of the values obtained in the N iterations, thus increasing the reliability and generalizability of these results (Witten et al., 2011). In this Thesis, loo-cv was used with the training set to optimize hyperparameters of MLP (Barroso-García et al., 2021a) and BY-MLP (Barroso-García et al., 2020) according to the agreement measure k obtained in the N iterations. Therefore, its application allowed us to build more robust predictive models. Note that this validation method has a high computational cost since the model is refitted and tested as many times as available observations, which can become a time-consuming and tedious task (Witten et al., 2011). However, it allows obtaining models more generalizable by using all available data, as well as reducing the variability that arises when the observations are allocated only into two groups (Refaeilzadeh et al., 2016; Witten et al., 2011). Moreover, loo-cv results are completely reproducible as there is no random partition.

- ***K*-fold cross-validation**. This is also an iterative method. It consists of randomly allocating the data into k sets of approximately the same size (Refaeilzadeh et al., 2016). Thereby, $k-1$ sets are used to train the model and one of them is used as testing. This process is repeated k times using a different test set in each iteration. At the end of the process, we will have k estimates of each performance and/or agreement measure. Its estimated final value will be the average of the values obtained in the k iterations,

which increases its reliability and generalizability (Refaeilzadeh et al., 2016; Witten et al., 2011). Regarding its computational cost, this is determined by the number of required iterations, i.e, the chosen k value (Witten et al., 2011). Overall, its choice depends on the dataset, although a value of $k = 10$ is commonly used (Refaeilzadeh et al., 2016). Consequently, 10-fold cross-validation method was applied to the test set in order to obtain more generalizable results with the logistic regression models proposed in our first study (Barroso-García et al., 2017). Note that loo-cv is a particular case of k-fold, in which the number of folds is equal to the number of available observations. The main advantage of k-fold is that it provides a better balance between bias and variance than loo-cv (Refaeilzadeh et al., 2016). Loo-cv uses $N - 1$ observations to train the model, i.e., practically the whole available dataset. This fact maximizes the model fit to data and reduces the bias (Refaeilzadeh et al., 2016). However, the final estimate of performance and/or agreement measures is computed with practically the same data (there is only one different data between each training set). This fact leads to a higher risk of overfitting and,therefore, a high variance (Refaeilzadeh et al., 2016). In contrast, the k sets used in k-fold method as training are much less overlapping, which leads to a lower variance. Consequently, this provides more accurate estimates of performance measures and/or agreements. (Refaeilzadeh et al., 2016; Witten et al., 2011).

- **Bootstrapping**. It is based on applying the proposed method to M bootstrap replicates (Efron and Tibshirani, 1994). A bootstrap replicate is a sample obtained from the original set of observations by random resampling with replacement. Resampling with replacement consists in that once an observation has been extracted, it is made available again for subsequent extractions. As a result of this resampling type, some observations will appear multiple times in the bootstap replicate and others will not be chosen (Witten et al., 2011). Each bootstrap replicate will be the same size as the original set of observations. Thus, a new replicate is used at each bootstrapping iteration, the model is fitted with the observations from this replicate, and then it is evaluated with the observations that are left out (Efron and Tibshirani, 1994; Witten et al., 2011). The process is repeated M times, as many as bootstrap replicates have been created. Thereby, M estimates will be obtained for each statistical measure of performance and/or agreement. Note that bootstrapping method generates certain bias in these estimates,

which can be a limitation when the training set is too small (Efron and Tibshirani, 1994). In order to deal with this issue, the following correction is used in the estimation of each statistical measure:

$$S = \frac{\sum_{i=1}^{M} 0.368 \cdot S_{train_i} + 0.632 \cdot S_{test_i}}{M}, \tag{4.64}$$

where S_{train_i} and $Stest_i$ are the value of the statistic obtained with the training and test sets of the i^{th} boostrap replicate, respectively. In this Thesis, bootstrapping method was applied to obtain the optimum feature subset in the selection stage (Barroso-García et al., 2020, 2021a,b), to optimize the hyperparameters of AdaBoost.M2 and MLP (Barroso-García et al., 2021b), and to increase the generalization of the results obtained with the test set (Barroso-García et al., 2021a,b). In all of them, 1000 bootstrap replicates were used. In those studies where bootstrapping was applied to the training set in selection and classification stage, the bootstrap replicate sets used for each stage were different to keep data independence.

Chapter 5

Results

This chapter is devoted to the most relevant results obtained throughout the Doctoral Thesis. Since the main objective is to characterize overnight AF to help in the pediatric OSA diagnosis, the chapter has been organized according to the different feature extraction approaches used to characterize the pediatric AF signal: 5.1 combination of CTM and spectral entropies, 5.2 RP-derived features, 5.3 bispectral features, and 5.4 wavelet features. Note that the conducted characterization is directly related to the studies included in the compendium of publications of this Doctoral Thesis (please see Appendix A).

5.1 Central tendency measure and spectral entropies

AF and RRV respiratory signals were characterized and their complementarity was also evaluated. The characterization was raised from two different perspectives, the variability study and the irregularity study of both signals. In order to conduct these studies, CTM was used to quantify signal variability and first to third order spectral entropies were used to quantify signal irregularity.

CTM was calculated for 5-min segments. Then, the values obtained for these segments were averaged to obtain the CTM of each subject. CTM computation involved defining the optimum value of the radius r. Thus, the training set was used to find the value that maximized the correlation between CTM and AHI. The Spearman's correlation coefficient (RHO) was used for this. The results obtained for AF and RRV ($r = 0.0004$ and $r = 7$, respectively) are shown in Figure 5.1.

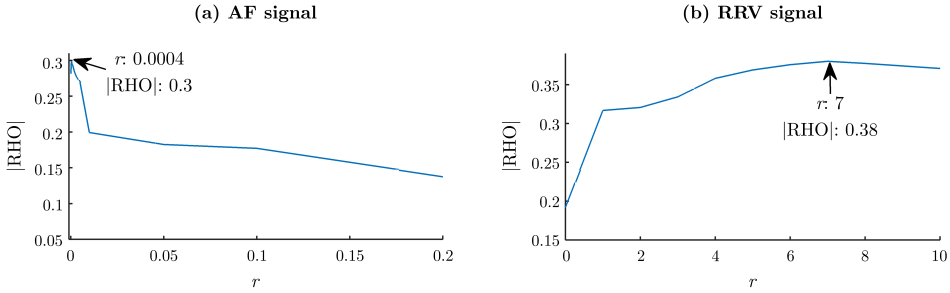


Figure 5.1: Spearman's correlation coefficient (RHO) obtained in the optimization of the radius r with the training set for (a) airflow signal (AF) and (b) respiratory rate variability signal (RRV). Figure taken from Barroso-García et al. (2017).

Regarding the spectral entropies, these were directly obtained from the PSD of each subject. PSD was computed by means of 2^{15}-sample Hamming windows (≈5-min), 50% overlap, and 2^{16}-point DFT. Figure 5.2 displays the averaged PSD by OSA severity groups in the training set. As can be seen, almost all spectral information of AF is located in the frequency range 0–0.6 Hz. Consequently, the spectral entropies from AF were computed for this frequency range. Similarly, not enough spectral content was found above 0.2 Hz for RRV signal. Thus, its spectral entropies were computed in the range 0–0.2 Hz.

Afterwards, CTM and spectral entropies of AF and RRV from the training set were subjected to a statistical analysis. The obtained results are presented in Table 5.1. As can be seen, the CTM and the spectral entropies showed an increasing tendency in AF signal, particularly in the severely affected children. Statistically significant differences (p-value < 0.01) were found among severity groups in CTM, SE_1, and SE_2 by means of Kruskal-Wallis test. When Mann-Whitney U test was applied, the severe OSA group showed statistically significant differences (p-value < 0.01) with the no-OSA and mild OSA groups in CTM, SE_1, and SE_2 from AF, but only SE_2 also obtained significant differences (p-value < 0.05, according to the number of involved subjects) with the moderate OSA group. In contrast, RRV did not show clear tendencies in any spectral entropy and no statistically significant differences were found among severity groups in them. However, the CTM from RRV showed a clear decreasing tendency as AHI increased, as well as significant differences (p-value < 0.01) when Kruskal-Wallis test was applied. According to Mann-Whitney U test, the severe OSA group showed statistically significant differences (p-value < 0.01 or p-value < 0.05, depending on the number of involved subjects) with the rest of the groups. Moreover, significant differences

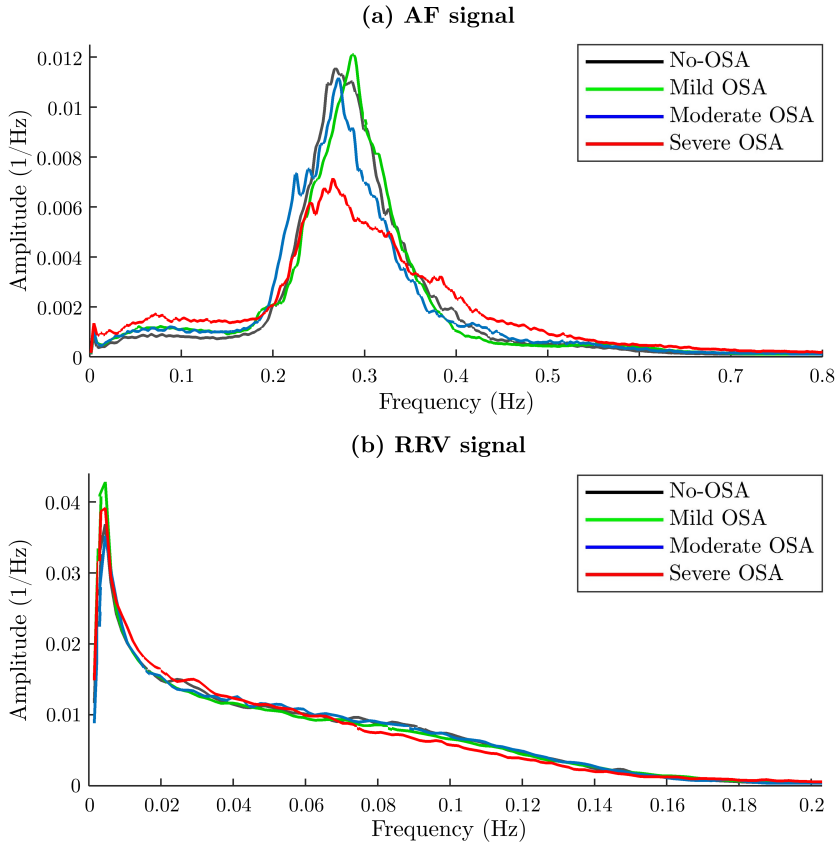


Figure 5.2: Averaged PSD of the four OSA severity groups (i.e. no-OSA, mild, moderate, and severe) in the training set for (a) airflow signal (AF) and (b) respiratory rate variability signal (RRV). Figure extracted from Barroso-García et al. (2017).

Table 5.1: Value of the central tendency measure and the spectral entropies by OSA severity group.

Feature	No-OSA	Mild OSA	Moderate OSA	Severe OSA	RHO	p-value
CTM^{AF}	0.06 [0.04]	0.06 [0.04]	0.06 [0.04]	0.10 [0.09]	0.2992	< 0.01
${SE_1}^{\mathrm{AF}}$	0.81 [0.06]	0.83 [0.08]	0.83 [0.12]	0.88 [0.10]	0.2769	< 0.01
${SE_2}^{\mathrm{AF}}$	0.64 [0.08]	0.66 [0.08]	0.65 [0.15]	0.73 [0.14]	0.1957	< 0.01
${SE_3}^{\mathrm{AF}}$	0.58 [0.07]	0.58 [0.10]	0.56 [0.10]	0.63 [0.17]	0.1166	0.0249
CTM^{RRV}	0.99 [0.02]	0.98 [0.02]	0.97 [0.03]	0.96 [0.08]	−0.3800	< 0.01
${SE_1}^{\mathrm{RRV}}$	0.91 [0.05]	0.90 [0.07]	0.90 [0.06]	0.89 [0.10]	−0.0943	0.2453
${SE_2}^{\mathrm{RRV}}$	0.81 [0.14]	0.78 [0.16]	0.81 [0.11]	0.75 [0.19]	−0.0412	0.2357
${SE_3}^{\mathrm{RRV}}$	0.69 [0.22]	0.65 [0.25]	0.74 [0.25]	0.64 [0.23]	0.0126	0.4132

Data are presented as median [interquartile range]. RHO: Spearman's correlation with the apnea-hypopnea index, p-value: result of Kruskal-Wallis test with Bonferroni correction.

(p-value < 0.05) were also found between the no-OSA and moderate OSA groups. That is, significant differences were obtained with CTM from RRV in 4 of 6 possible comparisons. In addition, this feature obtained the highest Spearman's correlation in absolute value with the AHI in the training set.

The results statistically obtained for each feature are consistent with the results achieved in the feature selection stage. In this regard, FSLR was used as selection method. Thus, 8 features (4 from AF signal and 4 from RRV signal) of the training set formed the FSLR algorithm input. According to the common AHI cut-off points, SE_1 from AF and CTM from RRV were selected by FSLR for 1 e/h, while it selected SE_2 from AF and CTM from RRV for 5 and 10 e/h. Thereby, three binary LR classifiers were constructed and trained with the training set, one for each AHI cut-off point: LR_1 fed with SE_1 from AF and CTM from RRV, LR_5 fed with SE_2 from AF and CTM from RRV, and LR_{10} fed with SE_2 from AF and CTM from RRV. Afterwards, each of these models was evaluated with the test set. The diagnostic performance obtained with each of them is shown in Table 5.2. As can be observed, a moderate-to-high diagnostic performance was reached for 5 and 10 e/h, as well as a fairly balanced Se-Sp pair for the three cut-off points.

5.2 Features derived from recurrence plots

The next step was to characterize the underlying dynamics and the phase-space of pediatric overnight AF. Thus, RPs were used for this purpose. As stated in the methodology section 4.2.3, the computation parameters τ, m, and ε were optimized with the training set using AMI, FNN, and Spearman's correlation methods, respectively. The optimum values obtained by these methods were the following: $\tau = 0.9$ s, $m = 3$, and $\varepsilon = 10^{-2} \cdot \sigma$. Once the calculation parameters were optimized, RPs were obtained for 30-s segments of AF signal. Afterwards, RPs obtained by segments were averaged to obtain the global RP from each subject.

Table 5.2: Diagnostic performance obtained by means of LR models for 1, 5, and 10 e/h.

AHI	Model	Se(%)	Sp(%)	Acc(%)	PPV(%)	NPV(%)	LR+	LR–	AUC
1	LR_1	60.5	58.6	60.0	81.2	25.0	1.1	0.9	0.59
5	LR_5	65.0	80.6	76.0	70.7	78.2	3.6	0.4	0.78
10	LR_{10}	83.3	79.0	80.0	52.8	93.5	4.0	0.2	0.80

AHI: apnea-hypopnea index, Se: sensitivity, Sp: specificity, Acc: accuracy, PPV: positive predictive value, NPV: negative predictive value, LR+: positive likelihood ratio, LR–: negative likelihood ratio, AUC: area under receiver-operating characteristic curve.

An exploratory analysis was conducted to study the behavior of AF-extracted RPs according to the OSA severity degree of the subjects. In this regard, Figure 5.3 shows the averaged RPs by severity group in the training set. As can be seen in this figure, the highest concentration of recurrences (red-orange region) is located close to the main diagonal of RP. However, there is a fading of recurrences towards the corners of RP (dark blue region) as the distance to the main diagonal is greater, which is a common evidence of non-stationary signals (Marwan et al., 2007). Note that there is a greater density of recurrences around the diagonal as the OSA severity increases, thus its fading towards the corners is slower in severely affected children. Consequently, a greater number of diagonal and vertical structures appear in these cases.

This led us to perform a recurrence quantification analysis (RQA) by means of the extraction of 9 RP-based features, as detailed in section 4.2.3. These features were computed from RPs obtained for each 30-s segment. The values calculated

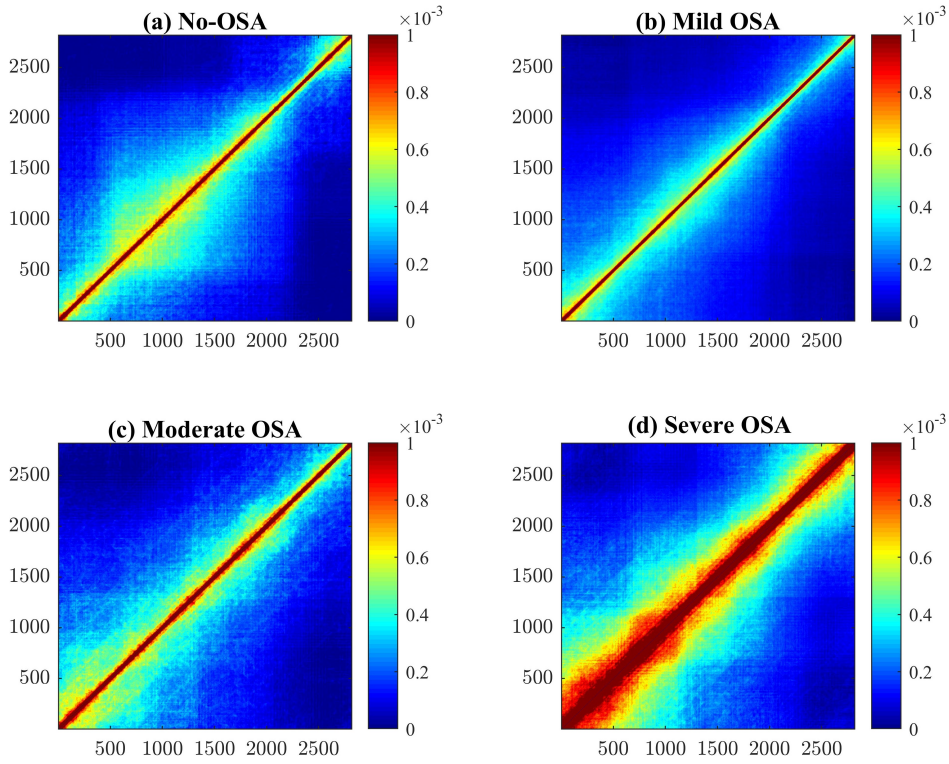


Figure 5.3: Averaged RP of the groups (a) no-OSA, (b) mild OSA, (c) moderate OSA, and (d) severe OSA in the training set. Figure taken from Barroso-García et al. (2020).

from each segment were averaged to obtain the RP features of each subject. Then, the 9 features were subjected to a statistical analysis. According to the results presented in Table 5.3, the features $ENTR$, LAM, LEN, $L_{\max}$, REC, TT, and $V_{\max}$ showed an increasing tendency as the OSA severity degree increased, while this tendency was decreasing for $TREND$. In the case of DET, it did not show a clear tendency. It did not also present statistically significant differences (p-value ≥ 0.01) between the OSA severity groups when Kruskal-Wallis and Mann-Whitney tests were applied with Bonferroni correction. The remaining RP features showed significant differences (p-value < 0.01) between the OSA severity groups when they underwent the Kruskal-Wallis test. In this regard, the severe OSA group presented significant differences (p-value < 0.01) with the other groups in all other features when Mann-Whitney test was applied. It was observed that LAM and $L_{\max}$ also manifested statistical differences between no-OSA and moderate OSA, as well as between mild and moderate OSA groups. In addition, these 2 features obtained higher Spearman's correlation with the AHI in the training set.

The results presented statistically for each feature are consistent with those obtained in the next methodological stage, the feature selection. In this case, FCBF was used as selection method together with the bootstrapping validation method. Thus, the algorithm was applied to 1000 replicates obtained from the training set and the features selected at least 500 times formed the optimum subset. In this regard, the only feature selected more than 500 times was $L_{\max}$ (Figure 5.4a), which was one of those that showed the most differences between OSA severity groups. Moreover, the complementarity between the information extracted from the RPs of AF and $ODI3$ was also evaluated in this study. Hence, this was incorporated into the selection process together with the 9 RP-derived features. In this case, the 2 features selected were $L_{\max}$ and $ODI3$ (Figure 5.4b).

Table 5.3: Value of the RP-derived features by OSA severity group.

Feature	No-OSA	Mild OSA	Moderate OSA	Severe OSA	RHO	p-value
DET	0.79 [0.17]	0.80 [0.20]	0.78 [0.16]	0.81 [0.19]	0.0754	0.3017
$ENTR$	1.05 [0.59]	1.03 [0.66]	1.10 [0.58]	1.40 [0.84]	0.3003	< 0.01
LAM	0.11 [0.07]	0.11 [0.09]	0.13 [0.10]	0.22 [0.15]	0.4649	< 0.01
LEN	10.28 [4.65]	10.01 [4.93]	11.15 [6.22]	16.23 [12.44]	0.3712	< 0.01
$L_{\max}$	55.06 [39.47]	57.25 [39.40]	68.23 [43.80]	140.76 [159.53]	0.4436	< 0.01
$REC(10^{-4})$	1.24 [2.33]	1.10 [2.61]	1.47 [3.37]	2.87 [5.23]	0.2415	< 0.01
$TREND(10^{-4})$	−6.67 [12.17]	−6.13 [12.76]	−7.72 [14.73]	−16.78 [25.99]	−0.2717	< 0.01
TT	1.07 [0.85]	1.07 [0.82]	1.24 [0.80]	1.70 [1.24]	0.3568	< 0.01
$V_{\max}$	2.92 [2.46]	2.81 [2.43]	3.44 [3.04]	5.25 [5.52]	0.3364	< 0.01

Data are presented as median [interquartile range]. RHO: Spearman's correlation with the apnea-hypopnea index, p-value: result of Kruskal-Wallis test with Bonferroni correction.

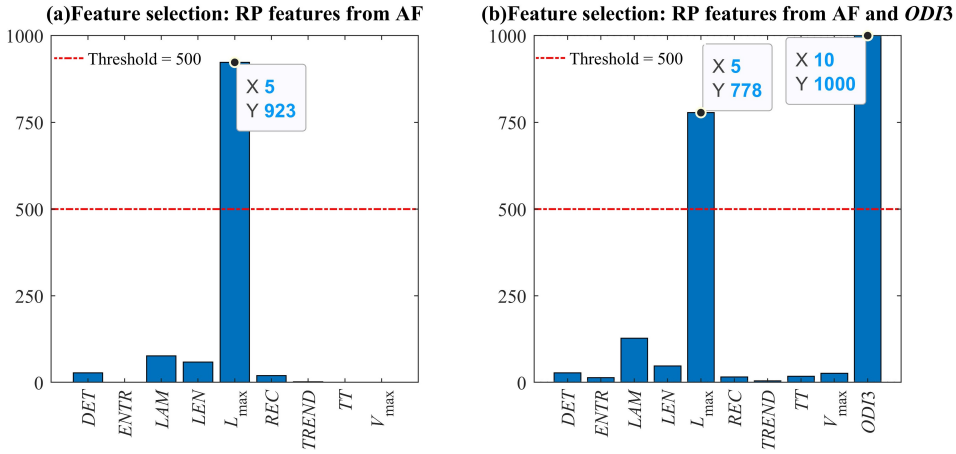


Figure 5.4: Results obtained with fast correlation-based filter for (a) RP-derived features from AF and (b) RP-derived features from AF and $ODI3$.

Therefore, 2 BY-MLP predictive models were built with the training set to estimate AHI of pediatric subjects: BY-MLP^{AF} fed only with AF features ($L_{\max}$) and BY-MLP^{AF,ODI3} fed with the optimum subset that combines both approaches ($L_{\max}$ and $ODI3$). In order to optimize N_H, this was varying from 1 to 30 in steps of 1. Loo-cv validation method was applied to the training set and the averaged k was obtained for each value of N_H. The optimal value of N_H was established as the one with which the highest value of k was obtained: $N_H = 17$ for BY-MLP^{AF} and $N_H = 16$ for BY-MLP^{AF,ODI3}.

Once the hyperparameters were optimized, the models were trained with the training set and then evaluated with the test set. The diagnostic performance achieved by each model for the AHI cut-off points 1, 5, and 10 e/h in test set is shown in Table 5.4. As can be observed, BY-MLP^{AF} model reached a moderate diagnostic performance but improving the one individually obtained by $ODI3$ in most of the diagnostic metrics for 1 e/h. These results, both those of AF and $ODI3$, were outperformed by the BY-MLP^{AF,ODI3} model that combines both approaches. Note that in addition to obtain higher diagnostic performance, BY-MLP^{AF,ODI3} overcame the limitations of BY-MLP^{AF} (OSA severity overestimation) and $ODI3$ (OSA severity underestimation) in 1 and 5 e/h. Thus, it is notable the LR– value ≤ 0.01 reached for 1 e/h, the balanced Se-Sp pair achieved for 5 e/h, as well as the LR+ value ≥ 10 obtained for 10 e/h by BY-MLP^{AF,ODI3} model.

Table 5.4: Diagnostic performance obtained by means of BY-MLP models and $ODI3$ for 1, 5, and 10 e/h.

AHI	Model	Se(%)	Sp(%)	Acc(%)	PPV(%)	NPV(%)	LR+	LR–	AUC
1	BY-MLP^{AF}	99.3	4.2	81.1	81.4	60.0	1.0	0.2	0.67
	BY-MLP^{AF,ODI3}	97.7	22.2	83.2	84.1	69.6	1.3	0.1	0.81
	$ODI3$	59.9	86.1	64.9	94.8	33.7	4.3	0.5	0.82
5	BY-MLP^{AF}	80.9	48.9	60.9	48.7	81.0	1.6	0.4	0.74
	BY-MLP^{AF,ODI3}	78.7	78.3	78.5	68.5	86.0	3.6	0.2	0.88
	$ODI3$	69.5	89.4	81.9	79.7	83.0	6.5	0.3	0.88
10	BY-MLP^{AF}	63.8	85.1	80.6	53.7	89.7	4.3	0.4	0.79
	BY-MLP^{AF,ODI3}	78.8	94.3	91.0	78.8	94.3	13.7	0.2	0.93
	$ODI3$	81.3	88.5	87.0	65.7	94.6	7.1	0.2	0.92

AHI: apnea-hypopnea index, Se: sensitivity, Sp: specificity, Acc: accuracy, PPV: positive predictive value, NPV: negative predictive value, LR+: positive likelihood ratio, LR–: negative likelihood ratio, AUC: area under receiver-operating characteristic curve.

5.3 Bispectral features

The next step was to characterize the gaussianity and the phase coupling of pediatric overnight AF, as well as the possible non-linear interactions between its different frequency components. Thus, bispectrum was used for this purpose. It was computed by means of 2^{12}-sample Hamming windows (≈30-s), 50% overlap, and 2^{13}-point DFT. Figure 5.5 displays the averaged bispectrum by OSA severity groups in the training set. Once the bispectrum was computed, the band corresponding to the particular respiratory rate of each pediatric subject was extracted from it. The bandwidth used to determine this band was 0.15 Hz. Figure 5.6 shows the averaged bispectral band by severity group in the training set. As can be seen in Figure 5.5, bispectrum reaches a greater amplitude in the frequency range 0.2–0.4 Hz (yellow region), which corresponds to the normal breathing band shown in Figure 5.6. However, bispectral power decreases in this band as the OSA severity degree increases (from 0.017 Hz^{-1} for no-OSA to 0.008 Hz^{-1} for severe OSA) and it is redistributed into other wider frequency ranges. Consequently, a new coupling focus appears around 0.05 Hz as the AHI increases, possibly associated with the occurrence of apneic events.

As stated in the methodology section 4.2.4, 13 bispectral features were extracted from the adaptive band of each subject in order to quantify the information contained therein. These 13 features were subjected to a statistical analysis. The results are presented in Table 5.5. As can be seen, the features B_{max}, B_{total}, H_1, H_2, H_3, H_4, and fm_2 showed a decreasing tendency as the OSA severity degree increased, while this tendency was increasing for BE_1, BE_2,

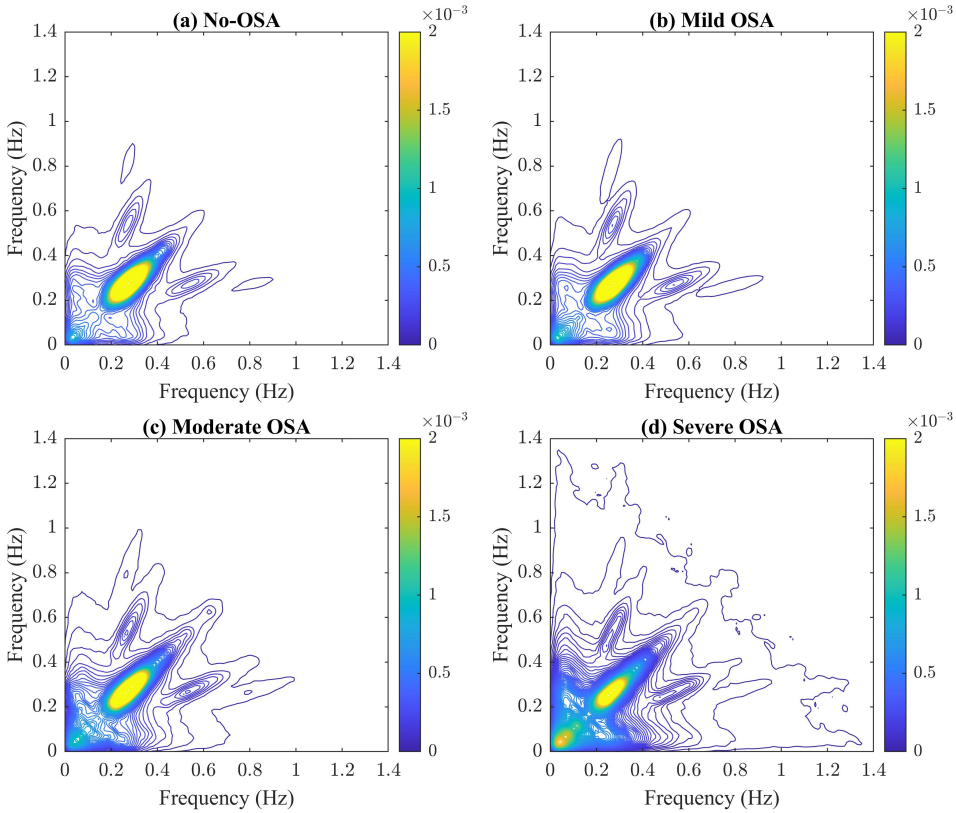


Figure 5.5: Averaged bispectrum of the groups (a) no-OSA, (b) mild OSA, (c) moderate OSA, and (d) severe OSA in the training set. Figure derived from Barroso-García et al. (2021a).

BE_3, PE, and fm_1. In the case of B_{min}, it experienced an increasing tendency in the less severe groups. Regarding the obtained p-values, B_{min} and fm_1 did not present statistically significant differences (p-value ≥ 0.01) between the OSA severity groups when Kruskal-Wallis and Mann-Whitney tests were applied with Bonferroni correction. In contrast, the rest of the bispectral features showed significant differences (p-value < 0.01) between the OSA severity groups when they underwent the Kruskal-Wallis test. Thus, the severe OSA group presented significant differences (p-value < 0.01) with the rest of the groups in these features when Mann-Whitney test was applied. In addition, it was observed that B_{max}, B_{total}, H_2, and H_4 also manifested statistical differences between no-OSA and moderate OSA groups, and B_{total} between mild and moderate OSA groups.

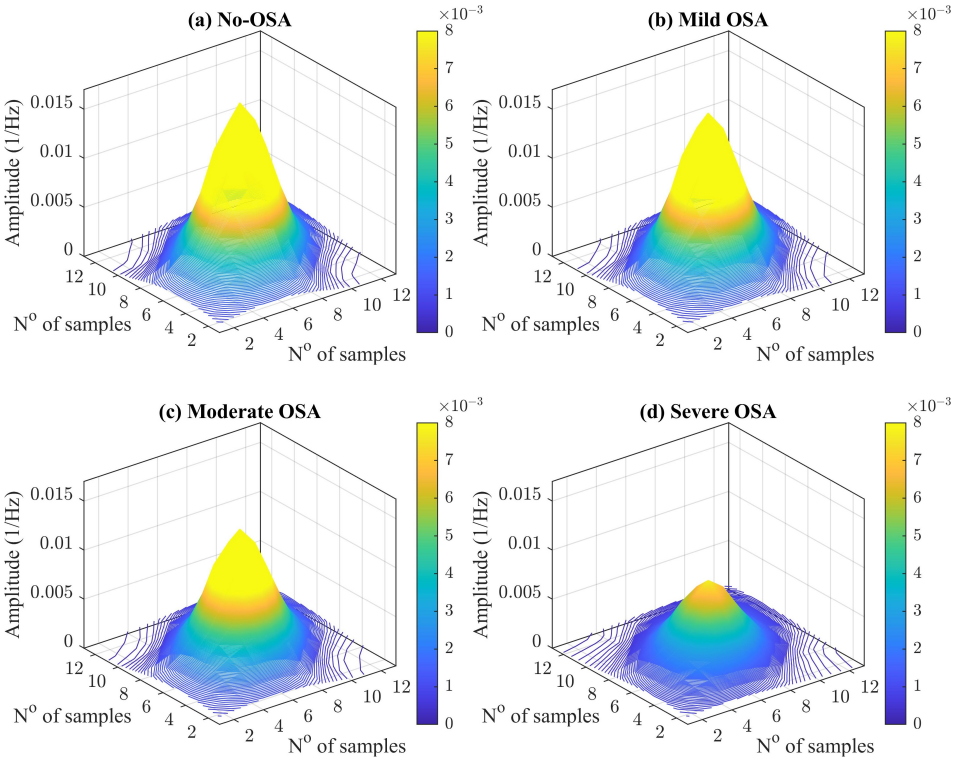


Figure 5.6: Averaged bispectral adaptive band of the groups (a) no-OSA, (b) mild OSA, (c) moderate OSA, and (d) severe OSA in the training set. Figure obtained from Barroso-García et al. (2021a)

Table 5.5: Value of the bispectral features by OSA severity group.

Feature	No-OSA	Mild OSA	Moderate OSA	Severe OSA	RHO	*p*-value
$Bmax$	0.03 [0.02]	0.03 [0.02]	0.02 [0.02]	0.01 [0.02]	−0.4231	< 0.01
$Bmin(10^{-4})$	0.74 [1.02]	0.76 [0.95]	0.99 [1.25]	0.93 [1.54]	0.1147	0.0224
$Btotal$	0.54 [0.27]	0.52 [0.31]	0.43 [0.28]	0.22 [0.29]	−0.4659	< 0.01
BE_1	5.61 [0.33]	5.61 [0.36]	5.68 [0.38]	5.89 [0.48]	0.3235	< 0.01
BE_2	4.60 [0.54]	4.59 [0.67]	4.65 [0.70]	4.95 [0.89]	0.2283	< 0.01
BE_3	3.92 [0.75]	3.96 [0.78]	3.94 [0.68]	4.20 [0.98]	0.1725	< 0.01
PE	1.78 [0.56]	1.81 [0.57]	1.77 [0.58]	1.95 [0.64]	0.1695	< 0.01
H_1	−551.47 [50.71]	−551.76 [60.39]	−568.31 [52.97]	−611.34 [115.67]	−0.3338	< 0.01
H_2	−68.54 [7.35]	−69.83 [9.04]	−71.75 [8.67]	−80.15 [16.14]	−0.4357	< 0.01
H_3	−484.66 [76.22]	−494.51 [86.71]	−510.21 [70.66]	−571.55 [121.66]	0.3989	< 0.01
$H_4(10^7)$	−1.70 [0.73]	−1.75 [0.81]	−1.92 [0.80]	−2.71 [1.66]	−0.4173	< 0.01
fm_1	7.72 [0.90]	7.87 [0.88]	7.88 [0.86]	7.93 [0.89]	0.0641	0.5911
fm_2	5.91 [0.84]	5.93 [0.92]	5.86 [0.77]	5.61 [1.06]	−0.1576	< 0.01

Data are presented as median [interquartile range]. RHO: Spearman's correlation with the apnea-hypopnea index, *p*-value: result of Kruskal-Wallis test with Bonferroni correction.

The results statistically presented for each feature are consistent with those obtained in the next methodological stage, the feature selection. FCBF was used as selection method together with a bootstrap method. Thus, the algorithm was applied to 1000 replicates obtained from the training set. Figure 5.7 displays the SU between each pair of features as the median value of the 1000 iterations. Note that the features from the same bispectral approach share a lot of information with each other ($SU \geq 0.5$). B_{max} and B_{total} also share sustancial informacion with the entropy-based features and moment-based features, respectively. However, lower SU values were found comparing B_{min} and PE with the rest of bispectral features. Thereby, the features selected as relevant and non-redundant a number of times equal to or greater than the average significance formed the optimum subset. In this regard, B_{min}, BE_1, H_2, and fm_2 were selected more than 182.9 times (Figure 5.8a). Moreover, the complementarity between the information extracted from the bispectrum of AF and $ODI3$ was also evaluated in this study. Hence, this was incorporated into the selection process together with the 13 bispectral features. In this case, the features selected more than 212.6 times were BE_1, H_2, fm_2, and $ODI3$ (Figure 5.8b).

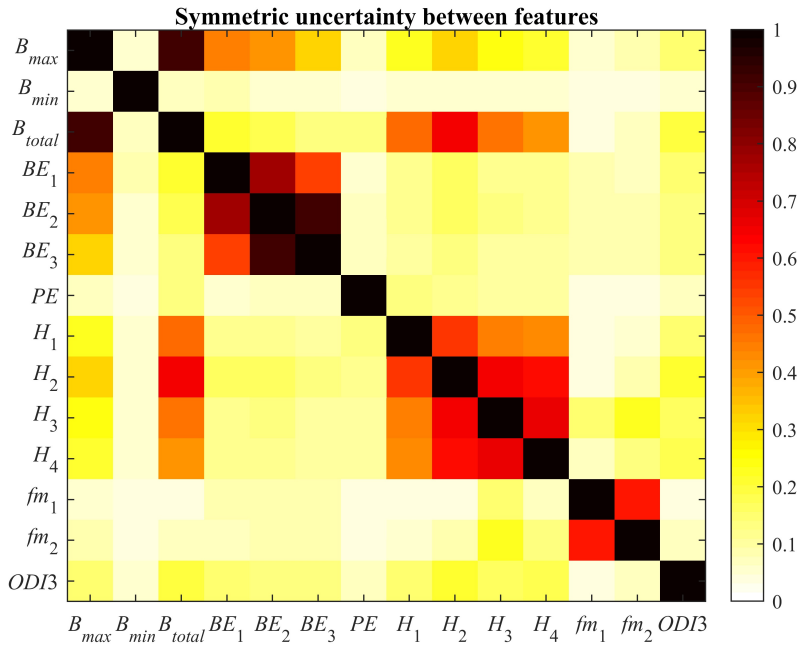


Figure 5.7: Heat map of the symmetric uncertainty between bispectral features and $ODI3$. The color scale represents the median value of 1000 bootstrap replicates. Figure derived from Barroso-García et al. (2021a).

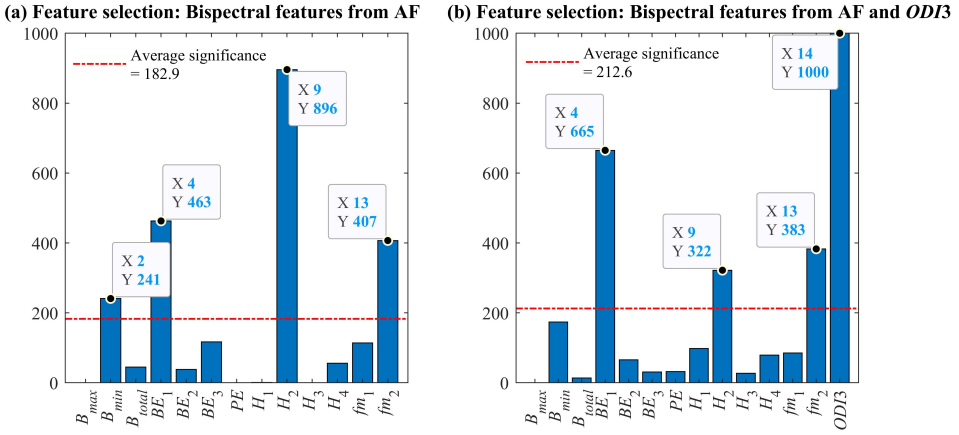


Figure 5.8: Results obtained with fast correlation-based filter for (a) bispectral features from AF and (b) bispectral features from AF and $ODI3$.

Therefore, 2 MLP predictive models were built with the training set to estimate AHI of pediatric subjects: MLP^{AF} fed only with AF features (B_{min}, BE_1, H_2, and fm_2) and MLP^{AF,ODI3} fed with the optimum subset that combines both approaches (BE_1, H_2, fm_2, and $ODI3$). In order to optimize the N_H-α pair, N_H ranged from 1 to 40 in steps of 1 and α ranged from 1 to 20 also in steps of 1. The loo-cv validation method was applied to the training set and the averaged k was obtained for each value of N_H-α pair. The optimal values of N_H and α were established as those with which the highest value of k was reached: $N_H = 39$ and $\alpha = 7$ for MLP^{AF}, and $N_H = 2$ and $\alpha = 8$ for MLP^{AF,ODI3}.

Once the hyperparameters were optimized, the models were trained with the training set and then evaluated with the test set. In order to obtain more generalizable results, the evaluation was performed in 1000 bootstrap replicates. The performance and agreement metrics were computed according to the bootstrap 0.632 procedure. Moreover, possible significant differences (p-value < 0.001) between models were assessed by means of Mann-Whitney U test together Bonferroni correction. The diagnostic performance achieved by each model for the AHI cut-off points 1, 5, and 10 e/h in test set is shown in Table 5.6. In addition, its global performance is shown in Table 5.7. Data are presented as median and 95% confidence interval in both tables.

As can be observed, MLP^{AF} model reached a moderate diagnostic performance but significantly improving (p-value < 0.001) the one individually obtained by $ODI3$ in Se and Acc for 1 e/h, as well in Se for 5 e/h. Even so, the accuracies obtained with MLP^{AF} and $ODI3$ were significantly outperformed (p-value<0.001)

Table 5.6: Diagnostic performance obtained by means of MLP models and $ODI3$ for 1, 5, and 10 e/h. Data presented as median [95% confidence interval].

AHI	Model	Se(%)	Sp(%)	Acc(%)	PPV(%)	NPV(%)	LR+	LR–	AUC
1	MLP^{AF}	94.1 ^{a,b} [92.4, 95.6]	11.2 ^{a,b} [6.7, 15.7]	78.1 ^{a,b} [75.7, 80.6]	81.7 ^{a,b} [79.3, 84.2]	30.8 ^{a,b} [17.7, 42.7]	1.1 ^{a,b} [1.0, 1.1]	0.6 ^{a,b} [0.4, 1.9]	0.72 ^{a,b} [0.68, 0.76]
	$MLP^{AF,ODI3}$	98.0 ^{a,c} [97.1, 98.9]	15.3 ^{a,c} [9.9, 20.5]	82.2 ^{a,c} [79.8, 84.4]	83.0 ^{a,c} [80.6, 85.3]	65.0 ^{a,c} [44.9, 83.5]	1.2 ^{a,c} [1.1, 1.3]	0.1 ^{a,c} [0.1, 0.4]	0.82 ^{a,c} [0.79, 0.84]
	$ODI3$	59.8 ^{b,c} [56.7, 63.3]	86.1 ^{b,c} [80.8, 90.8]	64.8 ^{b,c} [61.9, 68.0]	94.8 ^{b,c} [92.8, 96.6]	33.7 ^{b,c} [29.5, 38.1]	4.6 ^{b,c} [3.5, 10.8]	0.5 ^{b,c} [0.4, 0.5]	0.82 ^{b,c} [0.79, 0.85]
5	MLP^{AF}	78.7 ^{a,b} [74.6, 83.2]	50.6 ^{a,b} [47.0, 54.7]	61.2 ^{a,b} [58.2, 64.2]	48.9 ^{a,b} [45.2, 52.8]	79.8 ^{a,b} [76.0, 83.9]	1.6 ^{a,b} [1.5, 1.8]	0.4 ^{a,b} [0.3, 0.5]	0.72 ^{a,b} [0.68, 0.75]
	$MLP^{AF,ODI3}$	81.6 ^{a,c} [77.6, 85.7]	83.0 ^{a,c} [80.0, 85.9]	82.5 ^{a,c} [80.1, 84.8]	74.2 ^{a,c} [70.1, 78.5]	88.3 ^{a,c} [85.5, 91.0]	4.9 ^{a,c} [4.2, 6.4]	0.2 ^{a,c} [0.2, 0.3]	0.88 ^{a} [0.86, 0.91]
	$ODI3$	69.5 ^{b,c} [64.6, 74.2]	89.4 ^{b,c} [86.9, 91.7]	81.9 ^{b,c} [79.5, 84.3]	79.8 ^{b,c} [75.0, 84.0]	83.0 ^{b,c} [80.3, 85.8]	6.7 ^{b,c} [5.6, 10.2]	0.3 ^{b,c} [0.3, 0.4]	0.88 ^{b} [0.86, 0.90]
10	MLP^{AF}	55.9 ^{a,b} [49.6, 62.8]	83.2 ^{a,b} [80.6, 85.6]	77.4 ^{a,b} [74.7, 80.0]	47.4 ^{a,b} [41.1, 53.6]	87.5 ^{a,b} [85.2, 89.8]	3.4 ^{a,b} [2.8, 4.3]	0.5 ^{a,b} [0.5, 0.6]	0.76 ^{a,b} [0.72, 0.79]
	$MLP^{AF,ODI3}$	72.3 ^{a,c} [66.3, 78.1]	95.0 ^{a,c} [93.5, 96.4]	90.2 ^{a,c} [88.4, 92.0]	79.6 ^{a,c} [73.7, 85.6]	92.7 ^{a,c} [91.1, 94.3]	15.0 ^{a,c} [11.8, 30.7]	0.3 ^{a,c} [0.2, 0.4]	0.93 ^{a,c} [0.91, 0.95]
	$ODI3$	81.1 ^{b,c} [75.7, 86.1]	88.6 ^{b,c} [86.3, 90.8]	87.0 ^{b,c} [84.9, 89.1]	65.8 ^{b,c} [60.3, 71.7]	94.6 ^{b,c} [93.0, 96.1]	7.2 ^{b,c} [6.1, 10.0]	0.2 ^{b,c} [0.2, 0.3]	0.92 ^{b,c} [0.90, 0.94]

AHI: apnea-hypopnea index, Se: sensitivity, Sp: specificity, Acc: accuracy, PPV: positive predictive value, NPV: negative predictive value, LR+: positive likelihood ratio, LR–: negative likelihood ratio, AUC: area under receiver-operating characteristic curve, ^{a}Significant differences (p-value < 0.001) between MLP^{AF} and $MLP^{AF,ODI3}$, ^{b}Significant differences (p-value < 0.001) between MLP^{AF} and $ODI3$, ^{c}Significant differences (p-value < 0.001) between $MLP^{AF,ODI3}$ and $ODI3$. All differences were evaluated using Mann-Whitney U test with Bonferroni.

Table 5.7: Global performance obtained by means of MLP models and $ODI3$ for 1, 5, and 10 e/h. Data presented as median [95% confidence interval].

Model	k_4	ICC	Acc_4 (%)
MLP^{AF}	0.14 [0.11, 0.18] ^{a,b}	0.52 [0.41, 0.59] ^{a,b}	37.08 [34.17, 40.15] ^{a,b}
$MLP^{AF,ODI3}$	0.38 [0.34, 0.42] ^{a,c}	0.88 [0.81, 0.91] ^{a}	57.94 [55.02, 61.09] ^{a,c}
$ODI3$	0.29 [0.26, 0.33] ^{b,c}	0.88 [0.82, 0.90] ^{b}	46.21 [43.39, 49.60] ^{b,c}

k_4: four-class Cohen's kappa, ICC: intra-class correlation coefficient, Acc_4: four-class accuracy, ^{a}Significant differences (p-value < 0.001) between MLP^{AF} and $MLP^{AF,ODI3}$, ^{b}Significant differences (p-value < 0.001) between MLP^{AF} and $ODI3$, ^{c}Significant differences (p-value < 0.001) between $MLP^{AF,ODI3}$ and $ODI3$. All differences were evaluated using Mann-Whitney U test with Bonferroni correction.

by the model that combines both approaches ($MLP^{AF,ODI3}$). Note that in addition to get a statistically higher diagnostic performance, $MLP^{AF,ODI3}$ overcame the limitations of MLP^{AF} (OSA severity overestimation) and $ODI3$ (OSA severity underestimation) in 1 and 5 e/h. Moreover, it is notable the LR– value ≤ 0.01 reached for 1 e/h, the balanced Se-Sp pair achieved for 5 e/h, as well as the LR+ value ≥ 10 obtained for 10 e/h by $MLP^{AF,ODI3}$ model. This model also achieved a significantly higher global performance (p-value < 0.001) than MLP^{AF} and $ODI3$ in terms of k and Acc_4.

5.4 Wavelet features

The final step was to characterize the pediatric overnight AF by means of DWT. Thus, DWT was applied to 2^{16}-sample segments (≈10-min) of AF signal. Haar and Daubechies-5 mother wavelets were used to compute the coefficients with and without sign of the 8^{th} level detail signal (0.1953–0.3906 Hz, corresponding to normal breathing band in children). The wavelet features were extracted for each of them. Thus, the average Spearman's correlation of the D_8-extracted features with Haar mother wavelet was 0.21 for coefficients with sign and 0.32 for coefficients without sign. In the same way, the average Spearman's correlation of the D_8-extracted features with Daubechies-5 mother wavelet was 0.24 for coefficients with sign and 0.34 for coefficients without sign. Therefore, the D_8-extracted features with Daubechies-5 mother wavelet and without sign were used in subsequent analysis.

Thus, Figure 5.9 and Figure 5.10 show the averaged D_8 detail signal by OSA severity groups in the training set, as well as the distribution of its coefficients (computed without sign through Daubechies-5 mother wavelet), respectively. As can be seen, D_8 detail signal amplitude notably decreases as the AHI increases.

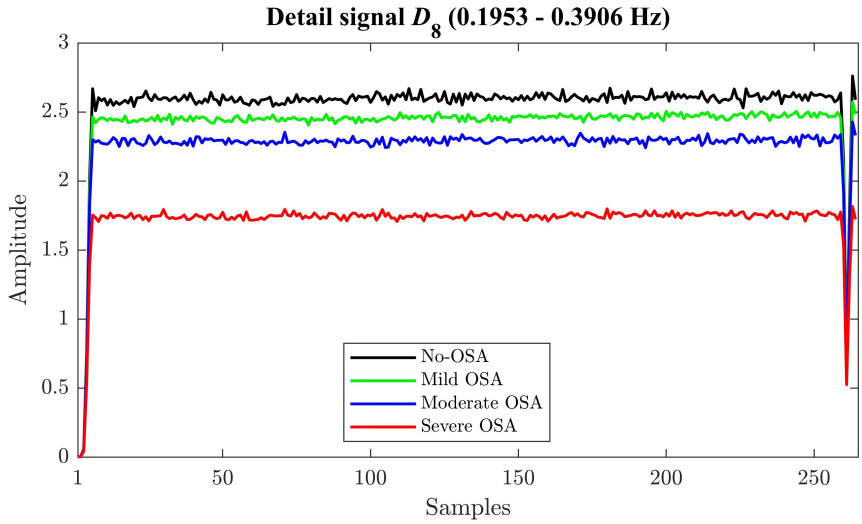


Figure 5.9: Averaged 8$^{\text{th}}$ level detail signal of the groups no-OSA, mild OSA, moderate OSA, and severe OSA in the training set. Figure obtained from Barroso-García et al. (2021b).

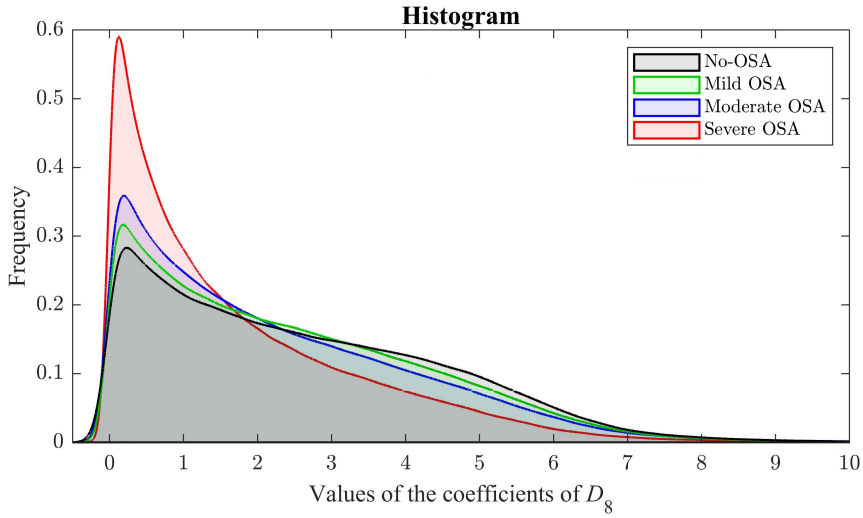


Figure 5.10: Coefficient distribution of 8$^{\text{th}}$ level detail signal for the groups no-OSA, mild OSA, moderate OSA, and severe OSA in the training set. Figure extractes from Barroso-García et al. (2021b).

This fact leads to the coefficients of D_8 from the most severely-affected subjects being concentrated close to zero value. Consequently, its distribution is more asymmetrical and pointed in these cases.

As stated in the methodology section 4.2.5, 8 wavelet features were extracted in order to quantify the information contained in the D_8 detail signal of each subject, as well in the full wavelet outline. These features were calculated from the detail coefficients obtained for each 2^{16}-sample segment. The values computed by segments were averaged to obtain the wavelet features of each subject. Then, they features were subjected to a statistical analysis to study its behavior according to the OSA severity degree of the subjects. Thus, Table 5.8 displays the value of each wavelet feature as median and interquartile range from the no-OSA, mild, moderate, and severe OSA groups, as well as its Spearman's correlation with the AHI, and the result from Kruskal-Wallis test with Bonferroni correction to evaluate possible differences among OSA severity groups in the training set. In this regard, the features M_{1D_8}, M_{2D_8}, Max_{D_8}, Min_{D_8}, and E_{D_8} showed a decreasing tendency as the AHI increased, while this tendency was increasing for M_{3D_8} and M_{4D_8}. In the case of WE, it did not show a clear tendency. Regarding the obtained p-values, all wavelet features showed significant differences (p-value < 0.01) between the OSA severity groups when the Kruskal-Wallis test was applied. In addition, M_{3D_8} and Min_{D_8} were the wavelet features that obtained highest correlation in absolute value with the AHI.

These results presented visually and statistically for each wavelet feature are consistent with those obtained in the feature selection stage. In this regard, FCBF was used as selection method together with a bootstrapping method. Thus, the algorithm was applied to 1000 replicates obtained from the training set and the features selected as relevant and non-redundant a number of times equal to or greater than the average significance formed the optimum subset. As can be seen in Figure 5.11a, M_{3D_8} was the only feature selected more than 125.25 times. Moreover, the complementarity between the information provided by the wavelet

Table 5.8: Value of the wavelet features by OSA severity group.

Feature	No-OSA	Mild OSA	Moderate OSA	Severe OSA	RHO	p-value
M_{1D_8}	2.62 [0.97]	2.46 [0.86]	2.29 [1.03]	1.67 [1.08]	-0.4024	< 0.01
M_{2D_8}	2.64 [1.27]	2.34 [1.20]	2.27 [1.34]	1.61 [1.28]	-0.3058	< 0.01
M_{3D_8}	0.25 [0.69]	0.29 [0.54]	0.50 [0.79]	1.05 [1.06]	0.4413	< 0.01
M_{4D_8}	2.87 [2.76]	2.98 [1.96]	3.58 [2.92]	5.32 [4.34]	0.3666	< 0.01
Max_{D_8}	6.67 [1.09]	6.61 [0.99]	6.59 [1.36]	6.21 [1.44]	-0.1662	< 0.01
$Min_{D_8}(10^{-3})$	2.87 [0.84]	2.60 [0.93]	2.52 [1.14]	1.86 [1.06]	-0.4154	< 0.01
$E_{D_8}(10^3)$	2.68 [1.56]	2.36 [1.41]	2.19 [1.67]	1.33 [1.44]	-0.3809	< 0.01
WE	0.26 [0.04]	0.25 [0.04]	0.26 [0.05]	0.28 [0.05]	0.2793	< 0.01

Data are presented as median [interquartile range]. RHO: Spearman's correlation with the apnea-hypopnea index, p-value: result of Kruskal-Wallis test with Bonferroni correction.

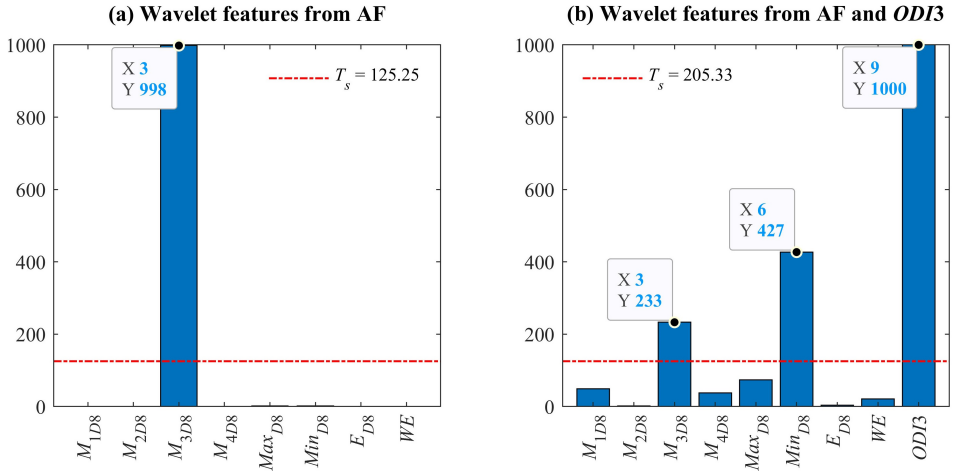


Figure 5.11: Results obtained with fast correlation-based filter for (a) wavelet features from AF and (b) wavelet features from AF and $ODI3$.

features from AF and $ODI3$ was also evaluated in this study. Hence, this was incorporated into the selection process together with the 8 wavelet features. In this case, the features selected more than 205.33 times were M_{3D_8}, Min_{D_8}, and $ODI3$ (Figure 5.11b).

Therefore, 2 AdaBoost.M2 multi-class classifiers were built with the training set to discriminate pediatric subjects according to their OSA severity degree: AB^{AF} fed with wavelet features from AF (M_{3D_8}) and $\text{AB}^{AF,ODI3}$ fed with the optimum subset that combines both approaches (M_{3D_8}, Min_{D_8}, and $ODI3$). In order to optimize the L-α pair, L varied in the range [1:9 10:10:90 100:100:900 1000:1000:10000] and α ranged from 0.1 to 1 in steps of 0.1. Bootstrapping validation method was applied to the training set (1000 replicates) and k value for each L-α pair was estimated by means of 0.632 bootstrap procedure. The optimal values of L and α were established as those with which the highest k value was reached: $L = 8000$ and $\alpha = 1$ for AB^{AF}, and $L = 3000$ and $\alpha = 1$ for $\text{AB}^{AF,ODI3}$. In the same way, 2 BY-MLP predictive models were built with the training set to estimate AHI of pediatric subjects: BY-MLP^{AF} fed only with M_{3D_8} and $\text{BY-MLP}^{AF,ODI3}$ fed M_{3D_8}, Min_{D_8}, and $ODI3$. In order to optimize N_H value, it ranged from 1 to 40 in steps of 1. Bootstrapping was applied to the training set (1000 replicates) and k value was estimated by means of 0.632 bootstrap procedure. The maximum value of k was obtained with $N_H = 1$ for BY-MLP^{AF} and $N_H = 36$ for $\text{BY-MLP}^{AF,ODI3}$.

Once the hyperparameters were optimized, the models were trained with the training set and then evaluated with the test set. In order to obtain more generalizable results, the evaluation was carried out in 1000 bootstrap replicates. The performance and agreement metrics were computed according to the bootstrap 0.632 procedure. Moreover, possible significant differences (p-value < 0.001) between models were assessed by means of Mann-Whitney U test together Bonferroni correction. The diagnostic performance achieved by each model for the AHI cut-off points 1, 5, and 10 e/h in test set is shown in Table 5.9. In addition, its global performance is shown in Table 5.10. Data are presented as median and 95% confidence interval in both tables.

As can be observed, AB^{AF} and BY-MLP^{AF} models reached a moderate diagnostic performance but significantly improving (p-value < 0.001) the result obtained by $ODI3$ in Se and Acc for 1 e/h, as well in Se for 5 e/h. Even so, the diagnostic accuracies individually achieved with AF and $ODI3$ were significantly outperformed (p-value < 0.001) by the models that combines both approaches: AB^{AF,ODI3} and BY-MLP^{AF,ODI3}. In this regard, BY-MLP^{AF,ODI3} reached the highest diagnostic accuracy for 1 and 5 e/h, while AB^{AF,ODI3} achieved it for 10 e/h. Note that in addition to get a statistically higher diagnostic performance, these models overcame the limitations of AB^{AF} and BY-MLP^{AF} (OSA severity overestimation), as well as the limitations of $ODI3$ (OSA severity underestimation) in 1 and 5 e/h. Moreover, AB^{AF,ODI3} and BY-MLP^{AF,ODI3} also achieved a significantly higher global performance (p-value < 0.001) than the individual approaches in terms of k_2, k_4, and Acc_4.

Table 5.9: Diagnostic performance obtained by means of AB and BY-MLP models, as well with $ODI3$ for 1, 5, and 10 e/h.

AHI	Model	Se(%)	Sp(%)	Acc(%)	PPV(%)	NPV(%)	LR+	LR−	k_2
1	AB^{AF}	79.9 ^{a,b,c,d} [77.1, 82.5]	47.2 ^{a,b,c,d} [39.4, 54.8]	73.6 ^{a,b,c,d} [70.9, 76.3]	86.4 ^{a,b,c,d} [83.9, 88.9]	35.6 ^{a,b,c,d} [29.3, 41.5]	1.5 ^{a,b,c,d} [1.4, 1.9]	0.4 ^{a,b,c,d} [0.4, 0.6]	0.24 ^{a,b,c,d} [0.17, 0.31]
	$AB^{AF,ODI3}$	80.3 ^{a,e,f,g} [77.6, 83.1]	68.1 ^{a,e,f,g} [61.9, 74.5]	78.0 ^{a,e,f,g} [75.5, 80.5]	91.5 ^{a,e,f,g} [89.3, 93.3]	44.9 ^{a,e,f,g} [39.0, 51.3]	2.6 ^{a,e,f,g} [2.2, 3.5]	0.3 ^{a,e,f,g} [0.3, 0.4]	0.40 ^{a,e,f,g} [0.34, 0.47]
	BY-MLP^{A}F	100.0 ^{b,e,h,i} [100.0, 100.0]	0.0 ^{b,e,h,i} [0.0, 0.0]	80.9 ^{b,e,h,i} [78.5, 83.2]	80.9 ^{b,e,h,i} [78.5, 83.2]	ND ^{b,e,h,i}	1.0 ^{b,e,h,i} [1.0, 1.0]	ND ^{b,e,h,i}	0.00 ^{b,e,h,i} [0.00, 0.00]
	BY-MLP$^{A}F, ODI3$	91.2 ^{c,f,h,j} [89.1, 93.0]	43.3 ^{c,f,h,j} [36.5, 50.6]	82.0 ^{c,f,h,j} [79.5, 84.3]	87.2 ^{c,f,h,j} [84.9, 89.2]	53.6 ^{c,f,h,j} [45.7, 61.9]	1.6 ^{c,f,h,j} [1.5, 1.9]	0.2 ^{c,f,h,j} [0.2, 0.3]	0.37 ^{c,f,h,j} [0.29, 0.44]
	$ODI3$	59.8 ^{d,g,i,j} [56.7, 63.3]	86.1 ^{d,g,i,j} [80.8, 90.8]	64.8 ^{d,g,i,j} [61.9, 68.0]	94.8 ^{d,g,i,j} [92.8, 96.6]	33.7 ^{d,g,i,j} [29.5, 38.1]	4.6 ^{d,g,i,j} [3.5, 10.8]	0.5 ^{d,g,i,j} [0.4, 0.5]	0.29 ^{d,g,i,j} [0.24, 0.34]
5	AB^{AF}	74.4 ^{a,b,c,d} [70.1, 79.1]	47.2 ^{a,b,c,d} [43.3, 51.2]	57.5 ^{a,c,d} [54.4, 60.5]	45.8 ^{a,c,d} [42.1, 49.9]	75.6 ^{a,b,c,d} [71.6, 79.9]	1.4 ^{a,c,d} [1.3, 1.6]	0.5 ^{a,b,c,d} [0.4, 0.7]	0.19 ^{a,c,d} [0.14, 0.25]
	$AB^{AF,ODI3}$	68.0 ^{a,e,f,g} [63.1, 72.8]	90.3 ^{a,e,f,g} [87.9, 92.5]	81.9 ^{a,e,f} [79.5, 84.4]	80.8 ^{a,e,f,g} [76.2, 85.2]	82.5 ^{a,e,f,g} [79.7, 85.3]	7.2 ^{a,e,f,g} [5.9, 11.1]	0.4 ^{a,e,f,g} [0.3, 0.4]	0.60 ^{a,e,f} [0.55, 0.65]
	BY-MLP^{AF}	77.3 ^{b,e,h,i} [73.1, 81.5]	45.1 ^{b,e,h,i} [41.4, 49.1]	57.1 ^{e,h,i} [54.2, 60.3]	45.8 ^{e,h,i} [42.0, 49.8]	76.9 ^{b,e,h,i} [72.8, 81.2]	1.4 ^{e,h,i} [1.3, 1.6]	0.5 ^{b,e,h,i} [0.4, 0.6]	0.20 ^{e,h,i} [0.15, 0.25]
	BY-MLP^{AF,ODI3}	79.3 ^{c,f,h,j} [74.9, 83.5]	83.8 ^{c,f,h,j} [80.9, 86.6]	82.1 ^{c,f,h,j} [79.8, 84.4]	74.6 ^{c,f,h,j} [70.4, 79.0]	87.2 ^{c,f,h,j} [84.5, 89.9]	5.0 ^{c,f,h,j} [4.3, 6.5]	0.3 ^{c,f,h,j} [0.2, 0.3]	0.62 ^{c,f,h,j} [0.58, 0.67]
	$ODI3$	69.5 ^{d,g,i,j} [64.6, 74.2]	89.4 ^{d,g,i,j} [86.9, 91.7]	81.9 ^{d,i,j} [79.5, 84.3]	79.8 ^{d,g,i,j} [75.0, 84.0]	83.0 ^{d,g,i,j} [80.3, 85.8]	6.7 ^{d,g,i,j} [5.6, 10.2]	0.3 ^{d,g,i,j} [0.3, 0.4]	0.60 ^{d,i,j} [0.55, 0.66]
10	AB^{AF}	41.1 ^{a,b,c,d} [34.7, 47.7]	85.5 ^{a,b,c,d} [83.1, 87.8]	76.1 ^{a,b,c,d} [73.5, 78.4]	43.3 ^{a,b,c,d} [36.9, 50.5]	84.3 ^{a,b,c,d} [81.9, 86.6]	2.9 ^{a,b,c,d} [2.4, 3.8]	0.7 ^{a,b,c,d} [0.6, 0.8]	0.27 ^{a,b,c,d} [0.20, 0.34]
	$AB^{AF,ODI3}$	72.4 ^{a,e,f,g} [66.6, 77.9]	96.0 ^{a,e,f,g} [94.6, 97.3]	91.0 ^{a,e,f,g} [89.3, 92.6]	83.0 ^{a,e,f,g} [77.4, 88.5]	92.8 ^{a,e,f,g} [91.0, 94.4]	19.0 ^{a,e,f,g} [14.6, 51.8]	0.3 ^{a,e,f,g} [0.2, 0.4]	0.72 ^{a,e,g} [0.66, 0.77]
	BY-MLP^{AF}	50.0 ^{b,e,h,i} [43.0, 56.7]	76.0 ^{b,e,h,i} [73.2, 78.8]	70.5 ^{b,e,h,i} [67.8, 73.1]	36.0 ^{b,e,h,i} [30.6, 41.7]	84.9 ^{b,e,h,i} [82.3, 87.3]	2.1 ^{b,e,h,i} [1.8, 2.6]	0.7 ^{b,e,h,i} [0.6, 0.8]	0.23 ^{b,e,h,i} [0.16, 0.29]
	BY-MLP^{AF,ODI3}	74.9 ^{c,f,h,j} [68.8, 80.5]	95.0 ^{c,f,h,j} [93.4, 96.5]	90.7 ^{c,f,h,j} [88.9, 92.5]	80.0 ^{c,f,h,j} [74.5, 85.8]	93.3 ^{c,f,h,j} [91.7, 94.9]	15.6 ^{c,f,h,j} [12.2, 30.3]	0.3 ^{c,f,h,j} [0.2, 0.3]	0.71 ^{c,h,j} [0.66, 0.77]
	$ODI3$	81.1 ^{d,g,i,j} [75.7, 86.1]	88.6 ^{d,g,i,j} [86.3, 90.8]	87.0 ^{d,g,i,j} [84.9, 89.1]	65.8 ^{d,g,i,j} [60.3, 71.7]	94.6 ^{d,g,i,j} [93.0, 96.1]	7.2 ^{d,g,i,j} [6.1, 10.0]	0.2 ^{d,g,i,j} [0.2, 0.3]	0.64 ^{d,g,i,j} [0.59, 0.70]

AHI: apnea-hypopnea index, Se: sensitivity, Sp: specificity, Acc: accuracy, PPV: positive predictive value, NPV: negative predictive value, LR+: positive likelihood ratio, LR−: negative likelihood ratio, AUC: area under receiver-operating characteristic curve, ^{a}Significant differences (p-value < 0.001) between AB^{AF} and $AB^{AF,ODI3}$, ^{b}Significant differences (p-value < 0.001) between AB^{AF} and BY-MLP^{AF}, ^{c}Significant differences (p-value < 0.001) between AB^{AF} and BY-MLP^{AF,ODI3}, ^{d}Significant differences (p-value < 0.001) between AB^{AF} and $ODI3$, ^{e}Significant differences (p-value < 0.001) between $AB^{AF,ODI3}$ and BY-MLP^{AF}, ^{f}Significant differences (p-value < 0.001) between $AB^{AF,ODI3}$ and BY-MLP^{AF,ODI3}, ^{g}Significant differences (p-value < 0.001) between $AB^{AF,ODI3}$ and $ODI3$, ^{h}Significant differences (p-value < 0.001) between BY-MLP^{AF} and BY-MLP^{AF,ODI3}, ^{i}Significant differences (p-value < 0.001) between BY-MLP^{AF} and $ODI3$, ^{j}Significant differences (p-value < 0.001) between BY-MLP^{AF,ODI3} and $ODI3$. All differences were evaluated using Mann-Whitney U test with Bonferroni correction.

Table 5.10: Global performance obtained by means of AB and BY-MLP models, as well with $ODI3$ for 1, 5, and 10 e/h. Data presented as median [95% confidence interval].

Model	k_4	$\mathbf{Acc_4}$ (%)
AB^{AF}	0.11 [0.08, 0.15] ^{a,b,c,d}	30.52 [27.90, 33.37] ^{a,b,c,d}
$AB^{AF,ODI3}$	0.40 [0.36, 0.45] ^{a,e,f,g}	57.46 [54.47,60.60] ^{a,e,f}
BY-MLP^{AF}	0.07 [0.03, 0.10] ^{b,e,h,i}	32.53 [29.87, 35.20] ^{b,e,h,i}
BY-$MLP^{AF,ODI3}$	0.41 [0.36, 0.45] ^{c,f,h,j}	58.57 [55.36, 61.47] ^{c,f,h,j}
$ODI3$	0.38 [0.34, 0.43] ^{d,g,i,j}	57.23 [53.95, 60.22] ^{d,i,j}

k_4: four-class Cohen's kappa, Acc_4: four-class accuracy, ^{a}Significant differences (p-value < 0.001) between AB^{AF} and $AB^{AF,ODI3}$, ^{b}Significant differences (p-value < 0.001) between AB^{AF} and BY-MLP^{AF}, ^{c}Significant differences (p-value < 0.001) between AB^{AF} and BY-$MLP^{AF,ODI3}$, ^{d}Significant differences (p-value < 0.001) between AB^{AF} and $ODI3$, ^{e}Significant differences (p-value < 0.001) between $AB^{AF,ODI3}$ and BY-MLP^{AF}, ^{f}Significant differences (p-value < 0.001) between $AB^{AF,ODI3}$ and BY-$MLP^{AF,ODI3}$, ^{g}Significant differences (p-value < 0.001) between $AB^{AF,ODI3}$ and $ODI3$, ^{h}Significant differences (p-value < 0.001) between BY-MLP^{AF} and BY-$MLP^{AF,ODI3}$, ^{i}Significant differences (p-value < 0.001) between BY-MLP^{AF} and $ODI3$, ^{j}Significant differences (p-value < 0.001) between BY-$MLP^{AF,ODI3}$ and $ODI3$. All differences were evaluated using Mann-Whitney U test with Bonferroni correction.

Chapter 6

Discussion

The study presented in this Doctoral Thesis focuses on characterizing overnight AF to help determine the presence and severity of OSA in children. In order to carry out this characterization, single-channel AF signal was automatically analyzed using different approaches: CTM, spectral entropies, RP, bispectrum, and wavelet. Each of these approaches was able to adapt to the intrinsic properties of the signal, such as non-linearity and non-stationarity, allowing us to uncover behaviors of pediatric nocturnal AF that were previously unknown in OSA context. Thus, the main findings of the study developed throughout this Thesis are discussed in the present chapter. Afterwards, the results obtained with each approach are discussed from the complementarity and performance viewpoint. In addition, a comparison with other state-of-the-art studies is also carried out. Finally, the main study limitations are exposed in the last section of this chapter.

6.1 Characterization of nocturnal AF in children

6.1.1 Central tendency measure and spectral entropies

Each of the different methodological approaches used in this Doctoral Thesis revealed OSA-related changes in the behavior of pediatric overnight AF. In this regard, the CTM from AF experienced an increasing tendency as severity increased, while it was decreasing for the CTM from RRV. This fact suggests that the OSA presence decreases the variability of AF and increases the one of RRV. This was supported by the statistical analysis, where the CTM from AF of the severe OSA group experienced statistically significant differences with the no-OSA and mild

OSA groups. Thus, the variability of AF would be especially useful to reflect the severe OSA particularities. Regarding the CTM from RRV, this showed significant differences between the severe OSA and the rest of the groups, as well as between the no-OSA and moderate OSA groups. Hence, the variability could be a common characteristic of the disease in RRV signal and, consequently, a useful indicator for its diagnosis.

In the case of the spectral entropies of AF, these showed an increasing tendency as the AHI was higher. This agrees with the Figure 5.2(a), where it was observed that the subjects with greater severity presented a flatter spectrum in the band 0–0.6 Hz. This tendency suggests that OSA could increase the irregularity of AF, causing its oscillatory behavior to evolve without a specific way. In addition, differences between SE_1, SE_2, and SE_3 could be observed. While SE_3 only manifested significant differences between the no-OSA and severe OSA groups, SE_1 also showed those between the mild and severe OSA groups. In contrast, only SE_2 presented differences between the severe OSA and all other groups, including the moderate OSA group. This fact would indicate that SE_2 can better reflect changes of irregularity caused by apneic events in AF and, thus, be particularly useful in severely affected cases. Regarding the spectral entropies of RRV, these showed high values, but without statistically significant differences between the OSA severity groups. This agrees with the Figure 5.2(b), since no visual differences could be appreciated in the flatness of the PSDs in the band 0–0.2 Hz. This suggests that RRV signal has a mostly irregular behavior regardless of the presence and severity degree of OSA.

6.1.2 Features derived from recurrence plots

The variability and irregularity changes experienced by AF in the presence of apneic events also agree with what was revealed by the RP-derived features REC and $ENTR$. These features showed an increasing tendency as the severity degree increased. Hence, OSA would reduce the variability (high REC) and increase the irregularity (high $ENTR$) of AF signal, which supports the findings made by means of the CTM and the spectral entropies of AF. In addition, the high recurrence density (REC) achieved by the most severe groups was consistent with the exploratory analysis conducted (Figure 5.3), where it was observed that the RP presented a greater number of recurrences in those cases. Since a recurrence indicated the existence of a relationship between 2 phase-space trajectories (same state in 2 different trajectories), this increase of recurrences revealed a greater

phase coupling in the groups with higher OSA severity. It was also observed that the recurrences faded towards the corners of the RP, which is a common manifestation of non-stationary time series. Such an effect occurred in the four OSA groups, suggesting that AF signal has a mostly non-stationary behavior. However, higher absolute values of $TREND$ were reached as the AHI increased, indicating that apneic events could increase the non-stationarity degree of AF. Regarding the complexity, LAM, TT, and $V_{\max}$ experienced a clear increasing tendency as the OSA severity was greater. This fact suggests that OSA increases the duration of the laminar structures of AF (high $V_{\max}$), causing the signal to not change or change very slowly (high LAM) since its trajectories remain longer trapped in the same phase state (high TT). Consequently, apneic events would be reducing the complexity of AF signal. In the case of $L_{\max}$ and LEN, these also experienced higher values with increasing AHI. Hence, OSA could increase the time during which the AF trajectories run in close phase-space states, causing these to diverge more slowly (high $L_{\max}$) and increasing the average prediction time of AF (high LEN). Regarding the predictability, DET did not show clear tendencies or statistical differences between severity groups. This fact suggests that AF signal has a mostly predictable behavior regardless of the presence and severity degree of OSA. However, REC, $ENTR$, $TREND$, TT, $V_{\max}$, and LEN presented significant differences between the severe OSA group and the rest of the groups. Hence, these would be especially useful to reflect the severe OSA particularities in AF. In addition, it was observed that $L_{\max}$ and LAM also manifested statistical differences between mild and moderate OSA groups, respectively. Thus, the exponential divergence and the laminarity could be common manifestations of pediatric OSA in AF signal and, consequently, useful indicators for its diagnosis.

6.1.3 Bispectral features

This approach provided information about the phase coupling, the gaussianity, and the non-linear interaction of the harmonic components of AF signal. As seen in the exploratory analysis (Figure 5.5 and 5.6), there was a coupling focus around the normal breathing band (0.20–0.40 Hz). As the AHI increased, this focus faded while another appeared at low frequencies (around 0.05 Hz). This fact suggests that apnea and hypopnea events reduce the bispectral amplitude and phase coupling of AF in the normal breathing band and reallocate it in other OSA-related frequency components. According to this bispectral amplitude reduction, B_{max} and B_{total} experienced a decrease with increasing severity. Due to the gaussian

components are nullified in high order spectrums (Chua et al., 2010), this decreasing tendency suggests that OSA reduces the non-gaussianity degree of AF signal. Regarding the bispectral entropies (BE_1, BE_2, and BE_3) and the phase entropy (PE), these experienced an increasing tendency as the AHI was higher. Thus, OSA could increase the irregularity of AF signal, disturbing its oscillatory behavior and causing it to evolve without following a specific amplitude and phase way. Moreover, H_1, H_2, H_3, and H_4 presented a decreasing tendency as the severity increased. This suggests that OSA reduces the non-linear interaction between the AF frequency components in the normal breathing band (low H_1). In addition, apneic events would also reduce the non-linear interaction between the harmonic components of AF signal (low H_2), leading to less phase coupling and fewer affected components (low H_3 and H_4). Furthermore, $WCOB$ revealed that OSA modify the coupling focus location. The slightly increasing and decreasing tendency of fm_1 and fm_2, respectively, suggests that OSA shift the activity focus toward lower frequency components. Regarding the minimum coupling of the adaptive band, B_{min} did not show clear tendencies or statistical differences between severity groups. However, BE_1, BE_2, BE_3, PE, H_1, H_3, and fm_2 presented significant differences between the severe OSA group and the rest of the groups. Thus, these could be regarded as a characteristic of the highest development of the disease. Moreover, B_{max}, B_{total}, H_2, and H_4 also showed differences between the no-OSA and moderate OSA groups, and B_{total} between the mild and moderate OSA groups. Thus, these could be common manifestations of pediatric OSA in AF signal and, consequently, useful indicators for its diagnosis.

6.1.4 Wavelet features

Concerning the wavelet analysis, it was observed that the pediatric subjects presented an AF detail signal with lower amplitude at the frequency range associated to normal respiration as the OSA severity increased (Figure 5.9). This is consistent with the decreasing tendency presented by M_{1D_8} and M_{2D_8}. Thus, the activity reduction caused by apneic events in the normal breathing band could decrease the value and dispersion range of the coefficients of the 8^{th} level detail signal from AF. Regarding the distribution of these coefficients, the asymmetry and peakedness increased with severity (Figure 5.10). Consequently, M_{3D_8} and M_{4D_8} showed an increasing tendency. This would indicate that OSA disturbs the frequency distribution of pediatric AF in the normal breathing band. Particularly, the recurrence of apneic events would reduce the frequency components of AF in

this region, resulting in almost all its coefficients are close to 0. A decreasing tendency of Max_{D_8}, Min_{D_8}, and En_{D_8} could also be observed. This suggests that OSA not only reduces the amplitude of D_8 from AF, but also its energy at this level of resolution. Moreover, WE experienced an increasing tendency. Hence, apneas and hypopneas would reduce the energy produced in the normal respiration frequency range and reallocate it into other frequency ranges associated with the recurrence of these events. Consequently, OSA would cause AF signal to be also more irregular in terms of energy. Note that all wavelet features showed significant differences between severity groups. The greatest differences arose in M_{1D_8}, M_{3D_8}, and Min_{D_8}, which also showed the highest absolute correlation with the AHI. This fact highlighted the usefulness of these features to reflect the OSA-related particularities in AF signal.

Based on the aforementioned considerations, the different methodological approaches proposed in this Doctoral Thesis allow adapting to the intrinsic properties of pediatric overnight AF, characterizing its behavior, and providing useful OSA-related information.

6.2 Complementarity

These OSA-related findings obtained from the pediatric AF characterization are consistent with the results achieved in the selection stage.

As explained in Barroso-García et al. (2017), SE_1, SE_2, and CTM from AF signal, as well as CTM from RRV signal, showed high separability between OSA severity groups. When the information provided by both signals was combined, FSLR algorithm automatically selected the features SE_1 from AF and CTM from RRV for 1 e/h, and SE_2 from AF and CTM of RRV for 5 e/h and 10 e/h. This fact allowed us to discover the following findings: (i) CTM from AF provides redundant information despite its high separability, (ii) the variability information (quantified by CTM) completes and is additional to the irregularity information (quantified by SE) from these respiratory signals, and (iii) there is complementarity between the information provided by AF and that extracted from RRV.

When pediatric overnight AF was characterized by means of RP-derived features in Barroso-García et al. (2020), LAM and $L_{\max}$ showed the highest separability between OSA severity groups despite clear tendencies of the rest. This supported that FCBF only selected $L_{\max}$ as relevant and non-redundant feature of the 9 extracted. Moreover, $L_{\max}$ continued to be selected along with $ODI3$ when it was incorporated into the selection process. These facts revealed that:

(i) despite its high separability, LAM provides redundant information, (ii) the information about the AF phase-space exponential divergence provided by means $L_{\max}$ is more useful for characterizing the pediatric OSA particularities on its own than other individual or combined RP features, and (iii) there is complementarity between the RP-extracted information from AF signal and that provided by the commonly used $ODI3$.

Regarding the bispectral analysis of AF signal conducted in Barroso-García et al. (2021a), it was observed that features from the same bispectral approach were redundant with each other (high SU values). Consequently, FCBF selected one bispectral feature of each approach as optimum subset: B_{min}, BE_1, H_2, and fm_2, i.e., the most relevant ones. In addition, BE_1, H_2, and fm_2 continued to be selected along with $ODI3$ when it was incorporated into the selection process. These facts allowed us to make the following findings: (i) despite not having individual utility (p-value ≥ 0.01 after Bonferroni correction), B_{min} contributes with additional information when it is combined with other bispectral features from AF, (ii) the different bispectral approaches provide additional and complementary information to each other, and (iii) there is complementarity between the bispectral information from AF signal and that provided by the oximetric index $ODI3$.

In the case of the AF characterization using DWT in Barroso-García et al. (2021b), M_{3D_8} and Min_{D_8} showed greater separability between OSA severity groups than the rest of wavelet features. This supported that FCBF only selected M_{3D_8} as relevant and non-redundant feature of the 8 extracted. However, both M_{3D_8} and Min_{D_8} were selected along with $ODI3$ when it was incorporated into the selection process. These facts revealed that: (i) the information about the distribution asymmetry of AF coefficients provided by M_{3D_8} is more useful for characterizing the pediatric OSA particularities than that individual and jointly provided by other wavelet features, (ii) although Min_{D_8} could be redundant to M_{3D_8}, it has shown to complement $ODI3$ with additional and different information from that supplied by M_{3D_8}, and (iii) there is complementarity between the information obtained from wavelet analysis of pediatric AF and the information provided by the $ODI3$.

All in all, the pediatric AF characterization through the different methodological approaches showed complementarity with the $ODI3$. Note that the hypopnea definition is based on a significant AF reduction that occurs along with a blood oxygen desaturation $\geq 3\%$ or an arousal (Berry et al., 2012). In this regard, the AASM recommends the use of a thermistor sensor and a nasal pressure sensor to score apneas and hypopneas, respectively (Berry et al., 2012). All AF recordings

used in this research were acquired by thermistor. Although the results obtained only with AF did not show signs of OSA severity underestimation due to sensor type, this could have conditioned the detection of hypopneas. This effect could have been reduced by incorporating the information about desaturations provided by the $ODI3$. Consequently, an improvement would emerge in the hypopneic event characterization. In the same way, the AF analysis from different perspectives would provide the necessary information to reduce the well-known OSA severity underestimation of the $ODI3$ (Kirk et al., 2003; Oeverland et al., 2002). Hence, the information provided by AF is additional and complete to the information about the occurrence of blood oxygen desaturations provided by this widely used oximetric index.

6.3 Diagnostic performance

This complementarity was not only manifested in the selection stage, but also in the pattern recognition stage. In this regard, moderate-to-high accuracies were achieved when the predictive models based on AF-extracted information were evaluated: 60.0%–81.1% for 1 e/h, 57.1%–76.0% for 5 e/h, and 70.5%–80.6% for 10 e/h (Barroso-García et al., 2017, 2020, 2021a,b). The AHI threshold 10 e/h reached the highest global diagnostic performance by jointly considering all metrics. This agrees with what was found in the conducted exploratory and statistical analyses, where severe OSA group presented more differences with respect to other OSA groups in all methodological approaches applied in this research. Thus, the highest Acc for 1 e/h and 10 e/h was obtained by means of RP-derived features (Barroso-García et al., 2020). Although the variability and irregularity analysis of AF and RRV obtained the lowest Acc for 1 e/h, this approach reached the highest Acc for 5 e/h, as well as a much more balanced Se-Sp pair than the other approaches for the three AHI thresholds (Barroso-García et al., 2017). It should be noted that the predictive models only based on RP, bispectrum, or wavelets features manifested an overestimation of the pediatric OSA severity degree for 1 and 5 e/h (low Sp values), as well as an underestimation of it for 10 e/h (low Se values) (Barroso-García et al., 2020, 2021a,b). In contrast, $ODI3$ presented a severity underestimation for 1 and 5 e/h.

These unwanted under and overestimation effects were reduced when the AF-extracted features were used along with the oximetric index. Combining both approaches, the severity overestimation remained for 1 e/h. This fact was offset as the models fed with AF (RP, bispectral, or wavelet features) and $ODI3$

significantly outperformed the diagnostic accuracy achieved only by the $ODI3$. In addition, the models built with RP-derived features from AF and $ODI3$, as well as with bispectral features from AF and $ODI3$, reached a LR– ≤ 0.1 for this threshold, providing a strong indicator to discard the pediatric OSA presence (Deeks and Altman, 2004), even more reliably than $ODI3$ according to the reported results. Regarding 5 events/h, the combination of bispectral features from AF and $ODI3$, as well as wavelet features from AF and $ODI3$, outperformed the individual diagnostic accuracy of the $ODI3$. Moreover, a higher Se and an almost perfectly balanced Se-Sp pair were obtained thanks to the information provided by the RP, bispectrum, and wavelet from AF. It is particularly remarkable, since the proposed models can provide similar importance to the classification of pediatric subjects above and below the threshold of this intermediate illness degree. $ODI3$ was also significantly outperformed for 10 events/h in almost all metrics. The models fed with AF features and $ODI3$ achieved higher diagnostic accuracy for all the methodological approaches proposed in this Thesis. In addition, higher values for LR+ (≥ 10) were also reached, which provides a strong indicator to establish the presence of severe OSA in children (Deeks and Altman, 2004), even more reliably than $ODI3$ according to the results. Thus, the combined use of AF and $ODI3$ achieved a significantly higher diagnostic performance for all AHI thresholds, outperforming those individually obtained.

In order to facilitate the comparison of the results jointly obtained by the $ODI3$ and each of the methodologies applied to AF in this research, these are summarized in Table 6.1. As can be seen, the RPs from AF signal and the $ODI3$

Table 6.1: Summary of the results jointly obtained by the $ODI3$ and each of the methodologies applied to AF in this research.

Study	AHI	Se(%)	Sp(%)	Acc(%)	PPV(%)	NPV(%)	LR+	LR–
Barroso-García et al. (2020)	1	97.7	22.2	83.2	84.1	69.6	1.3	0.1
	5	78.7	78.3	78.5	68.5	86.0	3.6	0.2
	10	78.8	94.3	91.0	78.8	94.3	13.7	0.2
Barroso-García et al. (2021a)	1	98.0	15.3	82.2	83.0	65.0	1.2	0.1
	5	81.6	83.0	82.5	74.2	88.3	4.9	0.2
	10	72.3	95.0	90.2	79.6	92.7	15.0	0.3
Barroso-García et al. (2021b) - AdaBoost	1	80.3	68.1	78.0	91.5	44.9	2.6	0.3
	5	68.0	90.3	81.9	80.8	82.5	7.2	0.4
	10	72.4	96.0	91.0	83.0	92.8	19.0	0.3
Barroso-García et al. (2021b) - BY-MLP	1	91.2	43.3	82.0	87.2	53.6	1.6	0.2
	5	79.3	83.8	82.1	74.6	87.2	5.0	0.3
	10	74.9	95.0	90.7	80.0	93.3	15.6	0.3

AHI: apnea-hypopnea index, Se: sensitivity, Sp: specificity, Acc: accuracy, PPV: positive predictive value, NPV: negative predictive value, LR+: positive likelihood ratio, LR–: negative likelihood ratio.

obtained the highest Acc for 1 e/h (Barroso-García et al., 2020). Although the wavelet-based models obtained a more balanced Se-Sp pair for this threshold, the RP-based model achieved a remarkably lower LR– than the rest of approaches. It should be noted that this diagnostic metric is considered as a reliable indicator to confirm the disease absence when its value is ≤ 0.1 (Deeks and Altman, 2004). Hence, the RP-derived information from AF signal could be used together with the $ODI3$ as a robust tool to discard the pediatric OSA presence. In the case of 5 e/h, the bispectral analysis of AF and the $ODI3$ achieved higher Acc than the other methodological approaches (Barroso-García et al., 2021a). In addition, this approach also obtained the most balanced Se-Sp pair, as well as the highest Se for this AHI threshold. Thus, the bispectral information from AF signal and the $ODI3$ could be particularly useful for discriminating pediatric subjects mildly affected by OSA from those moderately-to-severely affected. Regarding 10 e/h, the RP and wavelet features from AF obtained the highest diagnostic accuracy when these were combined with the $ODI3$ (Barroso-García et al., 2020, 2021b). Although the RP-based model obtained a slightly more balanced Se-Sp pair for this threshold, the AdaBoost model reached a remarkably higher LR+ using wavelet features. This diagnostic metric is considered as a reliable indicator to confirm the disease presence when its value is ≥ 10.0 (Deeks and Altman, 2004). Consequently, the DWT-derived information from AF signal and the $ODI3$ could be jointly used as a robust tool to determine the severe OSA presence in children. Considering all the thresholds, the AdaBoost model fed with wavelet features from AF and $ODI3$ provided the highest overall performance for 10 e/h, a much more balanced Se-Sp pair for 1 e/h, and higher Sp, PPV and LR+ for 5 e/h, at the expense of a slightly lower Acc at these last two thresholds. Thus, this could be proposed as our final model for the pediatric OSA diagnosis.

All in all, these proposals based on AF analysis would be an innovative and effective way to early diagnose children with OSA, as well as to automatically identify those who do not suffer from it. On the one hand, being able to automatically detect children without OSA would avoid referring them to hospital for undergoing to PSG. Consequently, the long waiting lists would be streamlined and the medical costs, such as those derived from hospitalization and equipment acquisition, would be reduced. On the other hand, children with moderate-to-severe OSA have a higher risk of developing comorbidities and major adverse health consequences (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). These include increases in C-reactive protein level and blood pressure, cardiac strain, cor pulmonale, as well as neurocognitive deficits (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). Due to

the severity of these consequences, children with AHI ≥ 5 are usually subjected to surgical treatment (adenotonsillectomy), while those with an AHI ≥ 10 can present residual OSA even after treatment (Alonso-Álvarez et al., 2011). Consequently, the tools proposed in this research would help to early diagnose these cases in order that they are treated in a timely fashion and thus avoid worse and irreversible sequels. In addition, this would allow the medical team to focus on doubtful pediatric OSA cases.

6.4 Comparison with state-of-the-art studies

Due to the high prevalence of pediatric OSA and its relationship with other severe pathologies, great efforts have been made to simplify the diagnosis of this disease in recent years. In this regard, several studies have focused their research on applying signal processing techniques to a reduced set of physiological recordings to automatically detect OSA in children. Tables 6.2 and 6.3 summarize the main methodological characteristics of these studies and the diagnostic performance achieved in them, respectively. In order to facilitate the comparison of results, Table 6.3 also shows those obtained by the methodological approaches with highest diagnostic performance of this research.

As can be observed, cardiorespiratory signals such as ECG, PPG, SpO_2, and AF are widely used in these state-of-the-art studies. Shouldice et al. (2004) carried out a temporal and spectral analysis of ECG signal, while Martín-Montero et al. (2020, 2021) used it to extract and characterize the pediatric HRV signal by spectral and bispectral techniques. These studies reached accuracies that ranged between 52.6%–84.0% for 1 e/h, 76.4%–80.0% for 5 e/h, and 82.8%–89.3% for 10 e/h. Other studies conducted an automatic analysis of PPG signal to diagnose OSA in children. In these works, PPG signal was characterized in both time domain and frequency domain (Dehkordi et al., 2016; Gil et al., 2009, 2010; Lazaro et al., 2014). However, they did not consider the OSA severity degree and they only assessed their proposal for 5 e/h.Thus, the results obtained for this threshold were within the range 71.0%–86.7%, being the study performed by Lazaro et al. (2014) the one that obtained the highest diagnostic performance with PPG signal. Regarding the oximetry signal, its use predominates in the studies of this context (Gutiérrez-Tobal et al., 2021). As can be seen in Garde et al. (2014a, 2019), Hornero et al. (2017), Xu et al. (2018), and Calderón et al. (2020), the characterization of SpO_2 by spectral, temporal, and non-linear techniques,

Table 6.2: Methodological summary of other state-of-the-art studies focused on the automatic OSA diagnosis in children.

Study	Signal	Extraction	Selection	Pattern recognition	Validation	#Total/ #Test
Shouldice et al. (2004)	ECG	Temporal, Spectral	–	QDA	Training /Test	50/25
Martín-Montero et al. (2020)	ECG	Spectral of HRV	–	LDA	Training /Test	1738/757
Martín-Montero et al. (2021)	ECG	Bispectral of HRV	FCBF	MLP	Training /Bootstrap /Test	1738/757
Gil et al. (2010)	PPG	DAP events, Spectral of HRV	Wrapper	QDA	–	21/21
Gil et al. (2010)	PPG	DAP events, PTTV, HRV	Wrapper	LDA	Loo-cv	21/21
Lazaro et al. (2014)	PPG	DAP events, Spectral of PRV	Wrapper	LDA	Loo-cv	21/21
Dehkordi et al. (2016)	PPG	Temporal, Spectral, Detrended fluctuation	LASSO	LASSO	10-fold-cv	146/146
Garde et al. (2014a)	SpO_2, PRV	Temporal, Non-linear, Spectral	Optimizing the AUC	LDA	Loo-cv /4-fold-cv	146/146
Hornero et al. (2017)	SpO_2	Temporal, Spectral, Non-linear, $ODI3$	FCBF	MLP	Training /loo-cv /Test	4191/3602
Álvarez et al. (2018)	SpO_2	Anthropometrics, Temporal, Symbolic dynamics, Oximetric indexes	FSLR	LR	Bootstrap	142/142
Vaquerizo-Villar et al. (2018b)	SpO_2	Anthropometrics, Spectral, Bispectral, $ODI3$	FCBF	MLP	Training /Validation /Test	298/75
Vaquerizo-Villar et al. (2018c)	SpO_2	Temporal, Spectral, Wavelet, $ODI3$	FCBF	SVM	Training /10-fold-cv /5-fold-cv /Test	981/392
Xu et al. (2018)	SpO_2	Temporal, Spectral, Non-linear, $ODI3$	–	MLP	Validation study	432/432
Garde et al. (2019)	SpO_2, PRV	Temporal, Spectral	Stepwise-selection	LR	Loo-cv	207/207
Calderón et al. (2020)	SpO_2	Oximetric indexes	–	LR	10-fold-cv	453/453
Gutiérrez-Tobal et al. (2015)	AF, SpO_2	Spectral, $ODI3$	FSLR	LR	Bootstrap	50/50
Jiménez-García et al. (2020)	AF, SpO_2	Temporal, Spectral, Non-linear, $ODI3$	FCBF	AdaBoost.M2	Training /Bootstrap /Test	974/390

ECG: electrocardiogram, PPG: photoplethysmography, SpO_2: blood oxygen saturation signal, AF: airflow signal, PRV: pulse rate variability signal, DAP: decreases in amplitude fluctuations of the PPG signal, HRV: heart rate variability, PTTV: pulse transit time variability, $ODI3$: 3% oxygen desaturation index, AUC: area under the receiver operating characteristic curves, FSLR: forward stepwise logistic regression, LASSO: least absolute shrinkage and selection operator, FCBF: fast correlation based filter, QDA: quadratic discriminant analysis, LDA: linear discriminant analysis, LR: logistic regression, MLP: multi-layer perceptron, SVM: support vector machine, AdaBoost: adaptive boosting, Loo-cv: leave-one-out cross validation.

Table 6.3: Comparison of the diagnostic performance obtained in other state-of-the-art studies focused on the automatic OSA diagnosis in children.

Study	AHI	Se(%)	Sp(%)	Acc(%)	PPV(%)	NPV(%)	LR+	LR–
Shouldice et al. (2004)	1	85.7	81.8	84.0	85.7	81.8	4.7	0.2
Martín-Montero et al. (2020)	1	42.5	72.3	52.6	75.0	39.1	1.5	0.8
	5	50.0	80.9	76.4	31.3	90.3	2.6	0.6
	10	63.8	84.7	82.8	29.5	95.9	4.2	0.4
Martín-Montero et al. (2021)	1	76.3	38.3	63.4	70.7	45.5	1.2	0.6
	5	62.5	84.2	81.0	40.7	92.8	4.0	0.4
	10	66.7	91.6	89.3	44.2	96.5	7.9	0.4
Gil et al. (2009)	5	87.5	71.4	80.0	–	–	3.1	0.2
Gil et al. (2010)	5	75.0	85.7	80.0	–	–	5.2	0.3
Lazaro et al. (2014)	5	100	71.4	86.7	–	–	3.5	0
Dehkordi et al. (2016)	5	76.0	68.0	71.0	–	–	2.4	0.4
Garde et al. (2014a)	5	88.4	83.6	84.9	76.9	92.6	5.4	0.1
Hornero et al. (2017)	1	84.0	53.2	75.2	81.6	53.7	1.8	0.3
	5	68.2	87.2	81.7	68.6	87.0	5.3	0.4
	10	68.7	94.1	90.2	67.7	94.3	11.6	0.3
Álvarez et al. (2018)	5	73.5	89.5	83.3	82.0	84.3	10.4	0.3
Vaquerizo-Villar et al. (2018b)	5	61.8	97.6	81.3	95.5	75.5	25.3	0.4
	10	60.0	94.5	85.3	80.0	86.7	11.0	0.4
Vaquerizo-Villar et al. (2018c)	5	71.9	91.1	84.0	83.8	84.5	14.6	0.3
Xu et al. (2018)	1	95.3	19.1	79.6	82.0	51.5	1.2	0.2
	5	77.8	80.5	79.4	72.3	84.7	4.0	0.3
	10	73.5	92.7	88.2	75.8	91.9	10.1	0.3
Garde et al. (2019)	1	80.0	65.0	75.0	–	–	2.3	0.3
	5	85.0	79.0	82.0	–	–	4.1	0.2
	10	82.0	91.0	89.0	–	–	9.1	0.2
Calderón et al. (2020)	5	62.0	96.0	79.0	94.3	–	15.5	0.4
Gutiérrez-Tobal et al. (2015)	3	85.9	87.4	86.3	88.4	85.8	6.8	0.2
Jiménez-García et al. (2020)	1	92.1	36.0	81.3	85.8	51.9	1.4	0.2
	5	76.0	85.7	82.1	76.0	85.7	5.3	0.3
	10	62.7	97.7	90.3	88.1	90.6	27.5	0.4
Barroso-García et al. (2020)	**1**	**97.7**	**22.2**	**83.2**	**84.1**	**69.6**	**1.3**	**0.1**
	5	**78.7**	**78.3**	**78.5**	**68.5**	**86.0**	**3.6**	**0.2**
	10	**78.8**	**94.3**	**91.0**	**78.8**	**94.3**	**13.7**	**0.2**
Barroso-García et al. (2021a)	**1**	**98.0**	**15.3**	**82.2**	**83.0**	**65.0**	**1.2**	**0.1**
	5	**81.6**	**83.0**	**82.5**	**74.2**	**88.3**	**4.9**	**0.2**
	10	**72.3**	**95.0**	**90.2**	**79.6**	**92.7**	**15.0**	**0.3**
Barroso-García et al. (2021b) - AdaBoost	**1**	**80.3**	**68.1**	**78.0**	**91.5**	**44.9**	**2.6**	**0.3**
	5	**68.0**	**90.3**	**81.9**	**80.8**	**82.5**	**7.2**	**0.4**
	10	**72.4**	**96.0**	**91.0**	**83.0**	**92.8**	**19.0**	**0.3**

AHI: apnea-hypopnea index, Se: sensitivity, Sp: specificity, Acc: accuracy, PPV: positive predictive value, NPV: negative predictive value, LR+: positive likelihood ratio, LR–: negative likelihood ratio.

as well as oximetric indices extracted from it, is a commonly used approach. Other works combined this with more novel techniques such as symbolic dynamics, bispectrum, or wavelet (Álvarez et al., 2018; Vaquerizo-Villar et al., 2018b,c). These SpO_2-based studies achieved accuracies that ranged between 75.0%–79.6% for 1

e/h, 79.0%–84.9% for 5 e/h, and 85.3%–90.2% for 10 e/h. Regarding AF signal, only 2 studies have been found that use this signal combined with the oximetry to help in pediatric OSA diagnosis. Gutiérrez-Tobal et al. (2015) characterized the overnight AF in frequency domain by defining new specific interest bands to children and extracting spectral features. Although their study achieved a high diagnostic accuracy for 3 e/h (86.3%), the evaluated threshold hinders a direct comparison with the performance results obtained in this Thesis. Jiménez-García et al. (2020) also characterized AF along with SpO_2 by means of spectral, temporal, and non-linear techniques, reaching accuracies of 81.3% for 1 e/h, 82.1% for 5 e/h, and 90.3% for 10 e/h.

The studies that comprise this compendium of publications reached accuracies ranged between 60.0%–83.2% for 1 e/h, 76.0%–82.5% for 5 e/h, and 80.0%–91.0% for 10 e/h. Thus, the research conducted in Barroso-García et al. (2020, 2021a,b) outperformed the Acc obtained for 1 e/h by the state-of-the-art studies that used SpO_2 and/or AF signal. Regarding ECG, only the study of Shouldice et al. (2004) slightly outperformed our results for this threshold (84.0% Acc versus 83.2% Acc). However, the sample size in their study is considerably smaller than that used in our research (50 subjects versus 946 subjects), limiting the generalization of their results. In addition, it should be noted that Barroso-García et al. (2020, 2021a) obtained a LR– = 0.1, the lowest value found in the literature for 1 e/h. As previously mentioned, this metric is a robust indicator to confirm the disease absence (Deeks and Altman, 2004). Therefore, our proposals, and particularly the one conducted in Barroso-García et al. (2020), are more robust discarding the pediatric OSA presence than the rest of the studies considered here. Regarding the AHI threshold 5 e/h, all our studies obtained results within the performance ranges achieved with ECG, PPG, SpO_2, and/or AF signals by other research works. The highest Acc reached by our studies for this threshold was with bispectrum in Barroso-García et al. (2021a) (82.5%), which outperformed most of the state-of-the-art studies. In addition, it achieved a very balanced Se-Sp pair compared to other works. Note that the generalization of the results from other studies reaching very meritorious performances is limited by the smaller number of their observations (Álvarez et al., 2018; Garde et al., 2014a; Lazaro et al., 2014). On the other hand, the studies based on RPs, bispectrum, and wavelet outperformed the diagnostic accuracies obtained for 10 e/h by all previously indicated state-of-the-art studies. Moreover, our research achieved high LR+ values for this threshold, particularly with the proposal based on wavelet and AdaBoost (LR+ = 19.0) (Barroso-García et al., 2021b). This metric was only improved by the

work of Jiménez-García et al. (2020). However, 3 of the 4 studies included in this compendium of publications outperformed their Acc and Se and obtained a more balanced Se-Sp pair for 10 e/h. Hence, our proposals, and particularly the one conducted in Barroso-García et al. (2021b), would be more useful discriminating the severe OSA degree than the rest of the studies considered here.

Other studies carried out in our research group also applied bispectrum to help diagnose OSA in children. In them, the bispectral features were extracted from SpO_2 (Vaquerizo-Villar et al., 2018b) and ECG-derived HRV (Martín-Montero et al., 2021). The use of these features notably improved the performance obtained by means of classic spectral approaches, thus increasing the diagnostic ability of these signals. However, the results achieved in Barroso-García et al. (2021a) with bispectral features from AF and $ODI3$ improved those reached by Vaquerizo-Villar et al. (2018b) and Martín-Montero et al. (2021). In other study, wavelet features were extracted from SpO_2 signal together with temporal and spectral features (Vaquerizo-Villar et al., 2018c). This approach achieved a high diagnostic accuracy for 5 e/h, improving the performance previously obtained with SpO_2 by means of conventional methods. In Barroso-García et al. (2021b), wavelet analysis was applied to AF signal and evaluated for 1, 5, and 10 e/h, i.e., considering the severity degree of the disease. This allowed us to detect pediatric subjects with severe OSA, as well as those who do not suffer from the disease. In addition, their performance was outperformed in terms of Se, NPV, and LR– for 5 e/h.

In summary, the characterization of pediatric overnight AF through the different methodological approaches proposed in this research allowed to obtain a high overall diagnostic performance for the three AHI thresholds in comparison with other state-of-the-art studies.

6.5 Limitations of the study

After conducting this research, it is necessary to point out and discuss some limitations that could have conditioned the reached results.

One of these limitations is the database size. Although the conducted research has involved a large and incremental number of subjects (from 501 to 946 children), it would have been desirable to have a greater number of recordings. Thus, the database would be more representative of OSA characteristics in children and the results would be more generalizable. In this regard, several validation techniques have been used throughout this Doctoral Thesis, such as loo-cv, k-fold, and bootstrapping, which have increased the validity and generalization of our results.

Another limitation is that the number of children with OSA predominates in this study (163 no-OSA, 386 mild, 172 moderate, and 225 severe OSA). This fact could influence the predictive model building that identify pathological and normality patterns, which would tend to overestimate and introduce a bias in the results. However, the number of subjects belonging to each OSA severity group reflects the actual disease prevalence in the pediatric population referred to sleep units, since all subjects showed previous symptoms of suffering from OSA.

It should also be noted that the AF analyses were not performed by age ranges, BMI, and/or gender of the pediatric subjects. This fact hinders the identification of the population subgroups where our diagnostic proposal would be more efficient, as well as the detection of characteristics or phenotypes that are conditioning our results. Nevertheless, pediatric AF recordings have been normalized in this research to minimize the effects generated by the age difference and by any other physiological characteristic other than OSA that could improperly influence the results.

Another limitation concerns the acquisition of AF signal. All pediatric signals used in this research have been recorded in a hospital environment. This ensured that sleep studies were conducted in a controlled environment. However, one of the main advantages of using single-channel AF for the pediatric OSA diagnosis is that it can be recorded using a portable device with thermistor during a type 4 study (Collop et al., 2007; Flemons et al., 2003). This offers the ability of recording AF outside the hospital setting. Therefore, it would be interesting to validate the efficacy of the proposed methodology in AF signals directly acquired at the patient's home.

In addition, the AASM recommends using thermistor sensor to adequately score apneas and nasal pressure sensor to score hypopneas (Berry et al., 2012). In this Doctoral Thesis, AF recordings acquired by means of a thermistor sensor have been used as only information source since nasal pressure sensor proved to be very noisy in our database. Consequently, this could have influenced the hypopnea detection and thus have conditioned the results of this research. However, Se values achieved through the different methodological approaches did not reveal a disease underestimation by using only thermistor. Moreover, previous studies in OSA adults have shown that very similar diagnostic performances can be reached when applying automatic analysis to AF, regardless it is acquired with a thermistor or a nasal prong sensore (Gutierrez-Tobal et al., 2012; Gutiérrez-Tobal et al., 2013, 2016).

As stated above, the AASM distinguishes between obstructive, central, and mixed apnea depending on the respiratory event origin and the respiratory effort criteria (Berry et al., 2012). Thus, information about the thoracic and/or abdominal movement is essential to be able to identify these types of respiratory events. However, only single-channel AF recordings were used in this research to simplify the pediatric OSA diagnosis, and no data about respiratory effort were available. Consequently, it was not possible to discriminate among obstructive, central, and mixed apneic events.

Another limitation is that the total recording time (TRT) was used to obtain the $ODI3$. It is important to point out that information about the EEG is essential to be able to distinguish between sleep and wake stages and accurately estimate the total sleep time (TST) (Iber et al., 2007). Alternatively, overnight actigraphy data could be used as a suboptimal option to calculate the TST (Grandner and Rosenberger, 2019). However, no reference EEG and/or actigraphy signals were available in our study. Consequently, it was not possible to compute the $ODI3$ using the TST, which could have generated a negative effect of underestimation for this clinical parameter.

Note that there are additional factors related to AF signal acquisition: the sensor could not suitably fit the child, it could lose contact with nose-mouth at night, or it could even record undesired artifacts caused by movements. This leads to AF recordings containing many noise sources, hindering its analysis. In this regard, a novel automatic method has been implemented in the present Doctoral Thesis to deal with this limitation.

There are also limitations concerning feature engineering methods. Previous studies have shown promising results by automatically analyzing AF signal through feature extraction and selection techniques (Gutiérrez-Tobal et al., 2015; Jiménez-García et al., 2020). Nevertheless, traditional approaches require a high knowledge degree to identify a priori which features should be extracted (Goodfellow et al., 2016). Moreover, they provide a low abstraction level, which limits its ability to detect complex patterns in physiological data (Goodfellow et al., 2016). Consequently, relevant information from AF signal could go unnoticed. Hence, it would be interesting to apply pattern recognition methods based on novel deep-learning techniques to address these limitations.

Finally, it should be noted that explainable artificial intelligence (XAI) methods have not been applied during the development of this research. Commonly, machine and deep learning models behave like black boxes, since it is hard to understand why they reach their results. In order to address this inherent drawback

and obtain more understandable results that increase the reliability of automatic OSA diagnostics, an exhaustive AF characterization work has been performed by means of different methodological techniques. This characterization has allowed to discover novel and useful OSA-related information from the AF signal, which has helped to better understand the results obtained in subsequent automatic processing stages.

Chapter 7

Conclusions

The four papers included in the present compendium of publications share a common research focus: the comprehensive characterization of nocturnal AF behavior in children to obtain relevant and useful information that helps to simplify the pediatric OSA diagnosis. In order to address this characterization of AF, the design, implementation, and evaluation of novel automatic signal processing algorithms were performed. Thus, AF was analyzed using different approaches adapted to the signal intrinsic properties: no-linear, spectral, bispectral, RP, and wavelet analyses. These approaches allowed us to discover novel and useful OSA-related information from AF signal, which had not been studied so far in the pediatric OSA context. The obtained results revealed that the overnight AF characterization by means of the methodologies proposed in this research improves its performance and helps to automatically diagnose the pediatric OSA.

In this chapter, the contributions that our research provides to the pediatric OSA diagnosis are indicated in section 7.1. Afterwards, the conclusions drawn from the studies conducted in this compendium are exposed in section 7.2. Finally, possible future research lines are suggested in section 7.3, which could continue and complement the findings of this Doctoral Thesis.

7.1 Contributions

Once the study is completed, it is important to highlight the original contributions that our work brings to this research field. Thus, the main ones are listed below:

1) **The AF signal pre-processing algorithm developed in this Doctoral Thesis**. As previously mentioned, there are many external factors related to the signal acquisition that originate noise sources and hinder its analysis. In order to deal with this drawback, the noise were carefully studied. As a result, a novel algorithm was designed and implemented to automatically remove artifacts from AF signal. In addition, this was combined with filtering and normalization techniques to complete the pre-processing. Consequently, the proposed algorithm allows to improve the quality of AF recordings, as well as to increase the effectiveness of its subsequent analysis.

2) **The exhaustive characterization of overnight AF in the pediatric OSA diagnosis context**. To the best of our knowledge, pediatric AF signal has only been analyzed by conventional spectral or non-linear methods previously used in adults. However, the clinical and diagnostic differences between both adults and children limit a suitable characterization of AF through these techniques. In contrast, we have extracted novel and useful information from pediatric AF signal, which has been able to reflect the OSA-related particularities and to improve its diagnostic ability.

3) **Characterization of AF and RRV signals by means of CTM and spectral entropies**. Although some of these features were previously used in adults, this is the first research where they are used and combined to help in pediatric OSA diagnosis. Thus, its use provided information about the variability and the irregularity of these respiratory signals. It is also important to note that this is the first time that AF-derived RRV signal is analyzed in the pediatric setting.

4) **Characterization of AF signal using RP-derived features**. To the best of our knowledge, this approach had not been previously evaluated to characterize AF signal. Hence, the RP-based analysis revealed novel information about the underlying dynamics and the phase-space of overnight AF.

5) **Characterization of pediatric AF through bispectral features**. This is the first research in which bispectrum is used for this purpose. It allowed us to obtain useful information related to the gaussianity, the non-linear interaction, and the phase coupling of AF signal.

6) **Characterization of overnight AF by DWT analysis**. This methodological approach had also not been used to characterize AF in the pediatric OSA context. In this regard, wavelet features revealed changes in the frequency components and the energy distribution of AF signal.

7) **Identification of optimum features subsets from AF**. To the best of our knowledge, this is the first time that relevant and non-redundant features from AF have been found in children.

7.2 Main conclusions

Based on the research developed throughout this Doctoral Thesis, the following conclusions are drawn:

1) CTM, spectral entropies, RP, bispectrum, and DWT analyses can characterize the behavior of overnight pediatric AF, adapting to its intrinsic properties and providing useful OSA-related information.

2) After applying CTM and spectral entropies, we concluded that OSA reduces the variability and increases the irregularity of AF. In addition, it also increases the variability of RRV signal.

3) Based on the information provided by RPs, OSA modifies the underlying dynamics and the phase-space of AF. Concretely, the occurrence of apneic events decreases the variability, the stationarity, and the complexity of AF signal, as well as the exponential divergence of its phase-space. It can also be concluded that OSA increases the dwell time at a certain phase state of AF (i.e., it does not change, or changes very slowly). Moreover, its average prediction time and its irregularity are higher as the AHI increases.

4) The bispectral features led us to conclude that OSA reduces the non-gaussianity of AF, as well as the non-linear interaction of its harmonic components. Childhood OSA also decreases the phase coupling in the normal breathing band, shifting the coupling focus towards low frequency components related to the occurrence apneic events. In addition, the irregularity of AF signal increases in terms of amplitude and phase when the OSA severity is higher.

5) Derived from the wavelet analysis, we can conclude that OSA disturbs the energy distribution and the frequency components of AF signal. Concretely, apneic events reduce the AF detail signal amplitude and the energy produced in the normal breathing band. In addition, the frequency components of AF decrease, while its irregularity increases in terms of energy as the AHI is higher.

6) The OSA-related information provided by AF signal is additional and complementary to that obtained from RRV signal. Particularly, there is complementarity between the irregularity of AF and the variability of RRV, thus highlighting its ability to reflect different OSA particularities.

7) The information about the changes that experiments the exponential divergence of AF phase-space is more useful to characterize the pediatric OSA-related particularities than the information individual and jointly provided by other RP-derived features.

8) The different bispectral approaches contribute with complementary information about the effects that apneic and hypopneic events cause in AF signal. Hence, its joint use provides a more complete characterization of the behavior of this signal.

9) The information about the changes that experiments the distribution asymmetry of AF coefficients is more useful to characterize the pediatric OSA-related particularities than the information individual and jointly provided by other wavelet features.

10) The information provided by AF through the different methodological approaches is complementary to the information provided by the classic $ODI3$ about the occurrence of blood oxygen desaturations. Thus, its joint use is particularly useful for accurately diagnosing the pediatric OSA.

11) The proposed AF processing methodologies allow obtaining highly accurate predictive models to automatically diagnose the presence and severity of childhood OSA. The highest diagnostic performance was reached by BY-MLP fed with RP-features and $ODI3$ for discarding the OSA presence (1 e/h), MLP fed with bispectral features and $ODI3$ for discriminating children mildly affected by OSA from those moderately-to-severely affected (5 e/h), and AdaBoost fed with wavelet features and $ODI3$ for confirming the severe

OSA presence (10 e/h). Considering together the performance achieved for all thresholds, the latter could be proposed as final model.

Based on the aforementioned considerations, we can conclude that the characterization of overnight AF by means of these novel methods can help to simplify the OSA diagnosis in children. In addition, the high performance of the proposed models suggests that they could be incorporated into clinical practice as reliable automatic screening methods for pediatric OSA.

7.3 Future research lines

During the development of this research, several new ideas have arisen. Although they fall outside the scope of this Doctoral Thesis, they could be the starting point of future research and contribute to our findings. Thus, the following ideas could be addressed in future works:

1) It would be interesting to apply our proposal to pediatric AF recordings by groups of age, BMI, gender, or combinations of these and other clinical and demographic characteristics. The comparison of these results could help to identify the population subgroups where our diagnostic proposal is most effective. It would also help to detect physiologic and demographic factors that influence our results, as well as to establish a future improvement plan. In addition, it would allow to find OSA-related sociodemographic particularities not yet covered in the pediatric context.

2) The validation of the proposed methodology in other larger databases from different hospital centers would be a direct continuation of our study. This would not only increase the reliability and generalizability of our results, but also help to detect possible conditioning factors (e.g., the use of different signal recording devices).

3) The validation of our proposal in domiciliary AF recordings would also be a natural way to continue our research. This is a key point as the simplification of childhood OSA diagnosis by automatic AF analysis includes that it can be acquired at the patient's home during a type 4 study.

4) Another interesting research would be the combination of overnight AF acquired by means of a thermistor sensor and the one obtained through a nasal pressure sensor. This would be very helpful to determine which of the two

AF recording types has higher ability to diagnose pediatric OSA. In addition, it could be evaluated whether both provide complementary information about the occurrence of apneas and hypopneas in children.

5) The combination of AF signal with other PSG-derived signals could be another novel future work. In this regard, it would be particularly interesting to combine AF signal with thoracic and/or abdominal movement signal in order to discriminate between the different types of apneic events (i.e., obstructive, central, and mixed events). Moreover, the analysis proposed in this research could be applied together with EEG or actigraphy signals in order to improve the $ODI3$ estimation.

6) Advanced pattern recognition methods based on novel deep-learning techniques could be also applied, such as long short-term memory (LSTM) and inception neural networks. Thus, these methods could be used with raw AF signals and/or AF-derived images (e.g., RP or bispectrum) to improve and maximize its diagnostic performance in future research.

7) Finally, it would be interesting to use explainable artificial intelligence (XAI) methods, such as layer-wise relevance propagation (LRP) and local interpretable model-agnostic explanations (LIME). In this way, the results derived from the predictive models would be more understandable, which would increase medical reliability in the OSA diagnostics that these models report.

Appendix A

Papers included in this Doctoral Thesis

Irregularity and Variability Analysis of Airflow Recordings to Facilitate the Diagnosis of Paediatric Sleep Apnoea-Hypopnoea Syndrome.

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, Félix del Campo, David Gozal, and Roberto Hornero. *Entropy*, vol. 19 (9), pp. 447, 2017. Impact factor in 2017: 2.305, Q2 in "PHYSICS, MULTIDISCIPLINARY" (JCR-WOS).

DOI: https://doi.org/10.3390/e19090447.

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

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Pablo Núñez, Félix del Campo, David Gozal, and Roberto Hornero. *Computer Methods and Programs in Biomedicine*, vol. 183, pp. 105083, 2020. Impact factor in 2020: 5.428, Q1 in "COMPUTER SCIENCE, THEORY & METHODS" (JCR-WOS).

DOI: https://doi.org/10.1016/j.cmpb.2019.105083.

Bispectral Analysis of Overnight Airflow to Improve the Pediatric Sleep Apnea Diagnosis.

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, David Gozal, and Roberto Hornero. *Computers in Biology and Medicine*, vol. 129, pp. 104167, 2021. Impact factor in 2020: 4.589, Q1 in "MATHEMATICAL & COMPUTATIONAL BIOLOGY" (JCR-WOS).

DOI: https://doi.org/10.1016/j.compbiomed.2020.104167.

Wavelet Analysis of Overnight Airflow to Detect Obstructive Sleep Apnea in Children.

Verónica Barroso-García, Gonzalo C. Gutiérrez-Tobal, David Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, Leila Kheirandish-Gozal, and Roberto Hornero. *Sensors*, vol. 21 (4), pp. 1491, 2021. Impact factor in 2020: 3.576, Q1 in "INSTRUMENTS & INSTRUMENTATION" (JCR-WOS).

DOI: https://doi.org/10.3390/s21041491.

Appendix B

Scientific achievements

B.1 Publications

B.1.1 Papers indexed in the JCR

1. Andrea Crespo, Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, María L. Alonso-Álvarez, Joaquín Terán-Santos, Roberto Hornero, Félix del Campo, "Multiscale Entropy Analysis of Unattended Oximetric Recordings to Assist in the Screening of Paediatric Sleep Apnoea at Home", *Entropy*, vol. 19 (6), pp. 284, June, 2017, DOI: 10.3390/e19060284.

2. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Irregularity and Variability Analysis of Airflow Recordings to Facilitate the Diagnosis of Paediatric Sleep Apnoea-Hypopnoea Syndrome", *Entropy*, vol. 19 (9), pp. 447, August, 2017, DOI: 10.3390/e19090447.

3. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Utility of bispectrum in the screening of pediatric sleep apnea-hypopnea syndrome using oximetry recordings", *Computer Methods and Programs in Biomedicine*, vol. 156, pp. 141–149, March, 2018, DOI: 10.1016/j.cmpb.2017.12.020.

4. Daniel Álvarez, Andrea Crespo, Fernando Vaquerizo-Villar, Gonzalo C. Gutiérrez-Tobal, Ana Cerezo-Hernández, **Verónica Barroso-García**, J. Mark Ansermino, Guy A. Dumont, Roberto Hornero, Félix del Campo, Ainara Garde, "Symbolic dynamics to enhance diagnostic ability of portable oximetry from the Phone Oximeter in the detection of paediatric sleep apnoea", *Physiological Measurement*, vol. 39 (10), pp. 104002, October, 2018, DOI: 10.1088/1361-6579/aae2a8.

5. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Detrended fluctuation analysis of the oximetry signal to assist in paediatric sleep apnoea-hypopnoea syndrome diagnosis", *Physiological Measurement*, vol. 39 (11), pp. 114006, November, 2018, DOI: 10.1088/1361-6579/aae66a.

6. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Wavelet Analysis of Oximetry Recordings to Assist in the Automated Detection of Moderate-to-Severe Pediatric Sleep Apnea-Hypopnea Syndrome", *PLOS One*, vol. 13 (12), e0208502, December, 2018, DOI: 10.1371/journal.pone.0208502.

7. Daniel Álvarez, Ana Sánchez, Ana Andrés, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Roberto Hornero, Félix del Campo, "Influence of chronic obstructive pulmonary disease and moderate-to-severe sleep apnoea in overnight cardiac autonomic modulation: time, frequency and non-linear analyses", *Entropy*, vol. 21 (4), pp. 381, April, 2019, DOI: 10.3390/e21040381.

8. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Pablo Núñez, Félix del Campo, David Gozal, Roberto Hornero, "Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis", *Computer Methods and Programs in Biomedicine*, vol. 183, pp. 105083, January, 2020, DOI: 10.1016/j.cmpb.2019.105083.

9. Pablo Núñez, Jesús Poza, Carlos Gómez, **Verónica Barroso-García**, Aarón Maturana-Candelas, Miguel A. Tola-Arribas, Mónica Cano, Roberto Hornero, "Characterization of the dynamic behavior of neural activity in

Alzheimer's disease: exploring the non-stationarity and recurrence structure of EEG resting-state activity", *Journal of Neural Engineering*, vol. 17 (1), pp. 016071, February, 2020,DOI: 10.1088/1741-2552/ab71e9.

10. Daniel Álvarez, Ana Cerezo-Hernández, Andrea Crespo, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Fernando Moreno, Ainhoa Arroyo, Tomás Ruiz-Albi, Félix del Campo, Roberto Hornero, "A machine learning-based test for adult sleep apnoea screening at home using oximetry and airflow", *Scientific Reports*, vol. 10, pp. 5332, March, 2020, DOI: 10.1038/s41598-020-62223-4.

11. Daniel Álvarez, Carmen A. Arroyo, Julio de Frutos, Andrea Crespo, Ana Cerezo-Hernández, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Fernando Moreno, Tomás Ruiz-Albi, Roberto Hornero, Félix del Campo, "Assessment of Nocturnal Autonomic Cardiac Imbalance in Positional Obstructive Sleep Apnea. A Multiscale Nonlinear Approach", *Entropy*, vol. 22 (12), pp. 1404, December, 2020, DOI: 10.3390/e22121404.

12. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Bispectral Analysis of Overnight Airflow to Improve the Pediatric Sleep Apnea Diagnosis", *Computers in Biology and Medicine*, vol. 129, pp. 104167, February, 2021, DOI: 10.1016/j.compbiomed.2020.104167.

13. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, David Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, Leila Kheirandish-Gozal, Roberto Hornero, "Wavelet Analysis of Overnight Airflow to Detect Obstructive Sleep Apnea in Children", *Sensors*, vol. 21 (4), pp. 1491, February, 2021, DOI: 10.3390/s21041491.

14. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Eduardo Santamaría-Vázquez, Félix del Campo, David Gozal, Roberto Hornero, "A convolutional neural network architecture to enhance oximetry ability to diagnose pediatric obstructive sleep apnea", *IEEE Journal of Biomedical and Health Informatics*, vol. 25 (8), August, 2021, DOI: 10.1109/JBHI.2020.3048901.

15. Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, David Gozal, **Verónica Barroso-García**, Daniel Álvarez, Félix del Campo, Leila Kheirandish-Gozal, Roberto Hornero, "Bispectral Analysis of Heart Rate Variability to Characterize and Help Diagnose Pediatric Sleep Apnea", *Entropy*, vol. 23 (8), pp. 1016, August, 2021, DOI: 10.3390/e23081016.

B.1.2 Book chapters

1. **Verónica Barroso-García**, Jorge Jiménez-García, Gonzalo C. Gutiérrez-Tobal, Roberto Hornero, "Airflow analysis in the context of sleep apnea", in *Advances in the Diagnosis and Treatment of Sleep Apnea: Filling the Gap Between Physicians and Engineers*, Cham, Switzerland: Springer Nature, Editors: Thomas Penzel and Roberto Hornero, 2022 (Under Review).

2. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Javier Gómez-Pilar, Félix del Campo, Roberto Hornero, "Conventional machine-learning methods applied to the automatic diagnosis of sleep apnea", in *Advances in the Diagnosis and Treatment of Sleep Apnea: Filling the Gap Between Physicians and Engineers*, Cham, Switzerland: Springer Nature, Editors: Thomas Penzel and Roberto Hornero, 2022 (Under Review).

B.1.3 International conferences

1. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Andrea Crespo, Ainhoa Arroyo, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Félix del Campo, Roberto Hornero, "Multi-Class AdaBoost to Detect Sleep Apnea-Hypopnea Syndrome Severity from Oximetry Recordings Obtained at Home", *Global Medical Engineering Physics Exchanges and Pan American Health Care Exchanges 2016 (GMEPE/PAHCE 2016)*, ISBN: 978-1-5090-2485-8, pp. 95–99, Madrid (Spain), April 4 - April 9, 2016, DOI: 10.1109/GMEPE-PAHCE.2016.7504632.

2. Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Andrea Crespo, Ainhoa Arroyo, Félix del Campo, Roberto Hornero, "Automated Analysis of Unattended Portable Oximetry by means of Bayesian Neural Networks to Assist in the Diagnosis of Sleep Apnea", *Global Medical Engineering Physics Exchanges and Pan American Health Care Exchanges 2016 (GMEPE/PAHCE 2016)*, ISBN:

978-1-5090-2485-8, pp. 79–82, Madrid (Spain), April 4 - April 9, 2016, DOI: 10.1109/GMEPE-PAHCE.2016.7504628.

3. Gonzalo C. Gutiérrez-Tobal, Julio de Frutos, Daniel Álvarez, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, Roberto Hornero, "A Bayesian Neural Network Approach to Compare the Spectral Information from Nasal Pressure and Thermistor Airflow in the Automatic Sleep Apnea Severity Estimation", *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2017)*, ISBN: 978-1-5090-2809-2, pp. 3741–3744, Jeju (South Korea), July 11 - July 15, 2017, DOI: 10.1109/EMBC.2017.8037670.

4. Fernando Vaquerizo-Villar, Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Leila Kheirandish-Gozal, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Usefulness of discrete wavelet transform in the analysis of oximetry signals to assist in childhood sleep apnea-hypopnea syndrome diagnosis", *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2017)*, ISBN: 978-1-5090-2809-2, pp. 3753–3756, Jeju (South Korea), July 11 - July 15, 2017, DOI: 10.1109/EMBC.2017.8037673.

5. Andrea Crespo, Fernando Vaquerizo-Villar, Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Ana Cerezo-Hernández, Graciela López-Muñiz, Leila Kheirandish-Gozal, David Gozal, Roberto Hornero, Félix del Campo, "Automated detection of childhood sleep apnea using discrete wavelet transform of nocturnal oximetry and anthropometric variables", *European Respiratory Society International Congress 2017 (ERS 2017)*, ISSN: 1399-3003, PA1308, Milán (Italy), September 9 - September 13, 2017, DOI: 10.1183/1393003.congress-2017.PA1308.

6. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Adrián Martín-Montero, Andrea Crespo, Félix del Campo, Roberto Hornero, "Pulse rate variability analysis to enhance oximetry as at-home alternative for sleep apnea diagnosing", *World Congress on Medical Physics and Biomedical Engineering 2018 (IUPESM 2018)*, ISBN: 978-981-10-9038-7, pp. 213–217, Praga (Czech Republic), June 3 - June 8, 2018, DOI: 10.1007/978-981-10-9038-7_39.

7. Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Bispectral Analysis to Enhance Oximetry as a Simplified Alternative for Pediatric Sleep Apnea Diagnosis", *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2018)*, ISBN: 978-1-5386-3646-6, pp. 175–178, Honolulu (United States), July 17 - July 21, 2018, DOI: 10.1109/EMBC.2018.8512248.

8. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Improving the Diagnostic Ability of Oximetry Recordings in Pediatric Sleep Apnea-Hypopnea Syndrome by Means of Multi- Class AdaBoost", *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2018)*, ISBN: 978-1-5386-3646-6, pp. 167–170, Honolulu (United States), July 17 - July 21, 2018, DOI: 10.1109/EMBC.2018.8512264.

9. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Félix del Campo, David Gozal, Roberto Hornero, "Usefulness of Spectral Analysis of Respiratory Rate Variability to Help in Pediatric Sleep Apnea-Hypopnea Syndrome Diagnosis", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2019)*, ISBN: 978-1-5386-1311-5, pp. 4580–4583, Berlin (Germany), July 23 - July 27, 2019, DOI: 10.1109/EMBC.2019.8857719.

10. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Félix del Campo, David Gozal, Roberto Hornero, "Convolutional Neural Networks to Detect Pediatric Apnea- Hypopnea Events from Oximetry", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2019)*, ISBN: 978-1-5386-1311-5, pp. 3555–3558, Berlin (Germany), July 23 - July 27, 2019, DOI: 10.1109/EMBC.2019.8857934.

11. Pablo Núñez, Jesús Poza, Carlos Gómez, **Verónica Barroso-García**, Saúl J. Ruiz-Gómez, Aarón Maturana-Candelas, Miguel A. Tola-Arribas, Mónica Cano, Roberto Hornero, "Characterization of EEG Resting-state Activity in Alzheimer's Disease by Means of Recurrence Plot Analyses", *41st Annual*

International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2019), ISBN: 978-1-5386-1311-5, pp. 5786–5789, Berlin (Germany), July 23 - July 27, 2019, DOI: 10.1109/EMBC.2019.8856600.

12. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Andrea Crespo, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Fernando Moreno, Félix del Campo, Roberto Hornero, "A bagging-based automatic method to estimate apnea-hypopnea index from home-oximetry recordings", *European Respiratory Society International Congress 2019 (ERS 2019)*, ISSN 1399-3003, PA822, Madrid (Spain), September 28 - October 2, 2019, DOI: 10.1183/13993003.congress-2019.PA822.

13. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Andrea Crespo, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Fernando Moreno, Félix del Campo, Roberto Hornero, "Assessment of least square boosting to estimate apnea-hypopnea index from at-home oximetry recordings", *8th European Medical and Biological Engineering Conference (EMBEC 2020)*, ISBN: 978-961-243-411-3, pp. 67, Portoroz (Slovenia), November 29 - December 3,2020.

B.1.4 National conferences

1. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Andrea Crespo, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Leila Kheirandish-Gozal, David Gozal, Félix del Campo, Roberto Hornero, "Modelos de máquinas de vector soporte aplicados sobre la oximetría nocturna para la detección automática de niños con síndrome de la apnea-hipopnea del sueño severo", *XXIV Reunión Anual de la Sociedad Española del Sueño (SES 2016)*, Valladolid (Spain), March 31 - April 2, 2016.

2. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Análisis espectral de la señal de flujo aéreo como ayuda al diagnóstico del síndrome de apnea-hipopnea del sueño en niños", *XXXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2016)*, ISBN: 978-84-9048-531-6, pp. 228–231, Valencia (Spain), October 23 - October 25, 2016.

3. Fernando Vaquerizo-Villar, Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Leila Kheirandish-Gozal, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Análisis de la señal de oximetría mediante la densidad espectral de potencia y bispectrum en la ayuda al diagnóstico de la apnea infantil", *XXXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2016)*, ISBN: 978-84-9048-531-6, pp. 202–205, Valencia (Spain), October 23 - October 25, 2016.

4. Gonzalo C. Gutiérrez-Tobal, Julio de Frutos, Daniel Álvarez, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, Roberto Hornero, "Estimación de la severidad de la apnea del sueño mediante redes neuronales bayesianas entrenadas con información espectral del flujo aéreo de sonda de presión y termistor", *XXV Reunión Anual de la Sociedad Española del Sueño (SES 2017)*, Santander (Spain), April 20 - April 22, 2017.

5. Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Ana Cerezo-Hernández, Graciela López-Muñiz, Leila Kheirandish-Gozal, David Gozal, Roberto Hornero, "Transformada wavelet de la señal de oximetría nocturna y variables antropométricas en la ayuda al diagnóstico automático de la apnea del sueño infantil", *XXV Reunión Anual de la Sociedad Española del Sueño (SES 2017)*, Santander (Spain), April 20 - April 22, 2017.

6. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Roberto Romero-Oraá, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Análisis de diferencias de segundo orden aplicado a la señal de flujo aéreo monocanal para la ayuda al diagnóstico del síndrome de la apnea-hipopnea del sueño en niños", *XXXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2017)*, ISBN: 978-84-9082-797-0, pp. 481–484, Bilbao (Spain), November 29 - December 1, 2017.

7. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Roberto Romero-Oraá, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Análisis de fluctuaciones sin tendencias (DFA) en los registros de oximetría para la ayuda en el diagnóstico del síndrome de la apnea-hipopnea del sueño

infantil", *XXXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2017)*, ISBN: 978-84-9082-797-0, pp. 209–212, Bilbao (Spain), November 29 - December 1, 2017.

8. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, **Verónica Barroso-García**, Fernando Vaquerizo-Villar, Adrián Martín-Montero, Andrea Crespo, Félix del Campo, Roberto Hornero, "Aplicación de la entropía espectral a la señal de variabilidad de pulso para incrementar el potencial de la oximetría en el diagnóstico de la apnea del sueño a domicilio", *XXVI Reunión Anual de la Sociedad Española del Sueño (SES 2018)*, Barcelona (Spain), April 26 - April 28, 2018.

9. Fernando Vaquerizo-Villar, Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, **Verónica Barroso-García**, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, "Utilidad de los patrones binarios locales aplicados a la señal de oximetría en la ayuda al diagnóstico del síndrome de la apnea-hipopnea del sueño en niños", *XXVI Reunión Anual de la Sociedad Española del Sueño (SES 2018)*, Barcelona (Spain), April 26 - April 28, 2018.

10. Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, Jesús Poza, Daniel Álvarez, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Saúl J. Ruiz-Gómez, Leila Kheirandish-Gozal, Félix del Campo, David Gozal, Roberto Hornero, "Caracterización de la apnea del sueño infantil mediante nuevas bandas espectrales del EEG", *XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2018)*, ISBN: 978-84-09-06253-9, pp. 249–252, Ciudad Real (Spain), November 21 - November 23, 2018.

11. Daniel Álvarez, Ana Cerezo-Hernández, Andrea Crespo, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Carmen A. Arroyo, Tomás Ruiz-Albi, Roberto Hornero, Félix del Campo, "Comparación de tests de screening automático de estimación del índice de apnea-hipopnea basados en oximetría y flujo aéreo no supervisados frente a la polisomnografía domiciliaria", *XXVII Reunión Anual de la Sociedad Española del Sueño (SES 2019)*, Vitoria (Spain), April 11 - April 13, 2019.

12. Ana Cerezo-Hernández, Daniel Álvarez, Ana Sánchez, Ana Andrés, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Roberto Hornero, Tomás Ruiz-Albi, Félix del Campo, "Estudio

del efecto combinado de la enfermedad pulmonar obstructiva crónica y del síndrome de apnea obstructiva del sueño moderado-a-severo en la modulación nocturna de la función autonómica cardiaca", *XXXVIII Congreso de la Sociedad Castellano-Leonesa y Cántabra de Patología Respiratoria (SOCALPAR 2019)*, Zamora (Spain), May 10 - May 11, 2019.

13. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Andrea Crespo, Fernando Moreno, Fernando Vaquerizo-Villar, Ana Cerezo-Hernández, **Verónica Barroso-García**, Félix del Campo, Roberto Hornero, "Screening automático de la apnea del sueño a partir de la pulsioximetría domiciliaria", *52º Congreso Nacional de la Sociedad Española de Neumología y Cirugía Torácica (SEPAR 2019)*, ISSN: 0300-2896, pp. 456–457, Santiago de Compostela (Spain), June 13 - June 16, 2019.

14. Daniel Álvarez, Ana Cerezo-Hernández, Andrea Crespo, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Fernando Moreno, Ainhoa Arroyo, Tomás Ruiz-Albi, Roberto Hornero, Félix del Campo, "Estimación automática del índice de apnea-hipopnea mediante los registros de oximetría y flujo aéreo realizados en el domicilio", *52º Congreso Nacional de la Sociedad Española de Neumología y Cirugía Torácica (SEPAR 2019)*, ISSN: 0300-2896, pp. 432, Santiago de Compostela (Spain), June 13 - June 16, 2019.

15. Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Jorge Jiménez-García, Leila Kheirandish-Gozal, Félix del Campo, David Gozal, Roberto Hornero, "Utilidad de nuevas bandas espectrales en la señal de HRV para ayudar en el diagnóstico de la apnea del sueño infantil", *XXXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2019)*, ISBN: 978-84-09-16707-4, pp. 295–298, Santander (Spain), November 27 - November 29, 2019.

16. Jorge Jiménez-García, Gonzalo C. Gutiérrez-Tobal, María García, Daniel Álvarez, **Verónica Barroso-García**, Fernando Vaquerizo-Villar, Adrián Martín-Montero, Félix del Campo, Leila Kheirandish-Gozal, David Gozal, Roberto Hornero, "Evaluación de la información espectral de las señales de flujo aéreo y saturación de oxígeno en sangre para la ayuda al diagnóstico de la apnea del sueño infantil", *XXXVII Congreso Anual de la Sociedad*

Española de Ingeniería Biomédica (CASEIB 2019), ISBN: 978-84-09-16707-4, pp. 25–28, Santander (Spain), November 27 - November 29, 2019.

17. Daniel Álvarez, Andrea Crespo, Ana Cerezo-Hernández, Gonzalo C. Gutiérrez-Tobal, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Fernando Moreno, Carmen A. Arroyo, Roberto Hornero, Félix del Campo, "Estudio de la asociación entre el índice de calidad del sueño de Pittsburgh y los índices polisomnográficos de severidad del SAHS", *53º Congreso Nacional de la Sociedad Española de Neumología y Cirugía Torácica (SEPAR 2020)*, ISSN: 0300-2896, pp. 390, Virtual conference (Spain), November 12 - November 14, 2020.

18. **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Félix del Campo, David Gozal, Roberto Hornero, "Análisis del flujo aéreo nocturno mediante wavelets para la ayuda en el diagnóstico de la apnea del sueño infantil", *XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2020)*, ISBN: 978-84-09-25491-0, pp. 252–255, Virtual conference (Spain), November 25 - November 27, 2020.

19. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Eduardo Santamaría-Vázquez, Félix del Campo, Roberto Hornero, "Modelo de deep learning basado en la arquitectura Inception para el diagnóstico de la apnea del sueño infantil mediante la señal de oximetría", *XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2020)*, ISBN: 978-84-09-25491-0, pp. 340–343, Virtual conference (Spain), November 25 - November 27, 2020.

B.2 International internship

Three-month research internship at the Sleep Technology and Analytics Group (STAG), University of Eastern Finland (UEF), Kuopio, Finland.

i. Purpose of the internship

The main purpose of the research stay was to develop novel and advanced signal processing methods based on deep-learning techniques to help in the pediatric OSA diagnosis. In order to reach this objective, the developed study encompassed: (1) state-of-the-art revision of deep-learning techniques

and hyperparameter optimization, (2) data collection of biomedical recordings of subjects with and without OSA (patients and controls, respectively), (3) design, development, and application of deep-learning methods to the biomedical recordings, and (4) analysis and evaluation of the results obtained.

ii. Methodological summary

The experimental study that was conducted during the research stay foused on the automatic deep learning-based estimation of the AHI from raw airflow signals in pediatric cohort. In this regard, the Childhood Adenotonsillectomy Trial dataset (CHAT, https://sleepdata.org/datasets/chat) was used. Thereby, 1,639 raw airflow recordings from thermistor were preparated and resampled prior to the processing and analysis process. Afterwards, an exhaustive review of the literature state was carried out in order to find the most suitable deep learning-based methods for our purpose. Thus, the use of convolutional neural networks (CNN) together with a long short-term memory (LSTM) architecture were considered to make a AHI regession of each child from their airflow signal. These deep learning methods were programmed and their hyperparameters were optimized according to the intrinsic properties of pediatric overnight AF. The results showed that the diagnostic performance measures were generally above those achieved in AF with traditional machine learning approaches, suggesting that the proposed methodology was suitable for helping diagnose pediatric OSA.

iii. Quality indicators of the institution

UEF is an university training institution with more than 15500 students and 2700 staff members. The University participates in several international networks, such as the Campus Europae, and cooperates with international universities, such as Nanjing University of China. Since launching its operations in 2010, UEF has reached suitable positions in international rankings (Shanghai: 501-600 and Times High Education: 401-500 in 2020). The STAG of UEF has expert staff in sleep disorders, as well as in advanced biomedical signal processing techniques. This research group cooperates with national and international institutions, conducting globally recognized research in the development of OSA diagnosis techniques. It is also important to highlight that their work is supported by 70 research projects and grants of public and private funding, all related to sleep research. The head of the STAG, Prof. Dr. Juha Töyräs, who was also the supervisor of the internship, is Profes-

sor at the UEF, Professor at the University of Queensland (Australia), and Chief Physicist at the Diagnostic Imaging Centre of the Kuopio Univeristy Hospital (Finland). He has supervised 45 Doctoral Theses and is currently supervising more than 20 in several fields of medical physics. In addition, he has published more than 282 JCR articles, reaching 7353/10783 citations and a H index of 45/54 according to Scopus/Google Scholar. Thus, these indicators reflect the high quality of the University and the research group, as well as the researcher who supervised the internship.

B.3 Awards and honors

B.3.1 Prizes

05/2016: **SOCALPAR 2016 prize**, awarded by the Sociedad Castellano-Leonesa y Cántabra de Patología Respiratoria (SOCALPAR), for the project entitled "Utilidad de una red neuronal basada en características demográficas y de oximetría nocturna como método de ayuda al diagnóstico del síndrome de apnea-hipopnea obstructiva del sueño en niños", conducted by Félix del Campo, Daniel Álvarez, Andrea Crespo, Tania Álvaro, Gonzalo C. Gutiérrez-Tobal, Ainhoa Arroyo, Julio de Frutos, Tomás Ruiz-Albi, **Verónica Barroso-García**, Fernando Vaquerizo-Villar, David Gozal, and Roberto Hornero.

11/2016: **Third José María Ferrero Corral prize in the XXXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2016)**, awarded by the Sociedad Española de Ingeniería Biomédica (SEIB), for the study entitled "Análisis espectral de la señal de flujo aéreo como ayuda al diagnóstico del síndrome de apnea-hipopnea del sueño en niños", conducted by **Verónica Barroso-García**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, Félix del Campo, David Gozal, and Roberto Hornero.

04/2017: **Prize to the second best conference paper in the XXV Reunión Anual de la Sociedad Española del Sueño (SES 2017)**, awarded by the Sociedad Española del Sueño (SES), for the study entitled "Transformada wavelet de la señal de oximetría nocturna y variables antropométricas en la ayuda al diagnóstico automático

de la apnea del sueño infantil", conducted by Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, Gonzalo C. Gutiérrez-Tobal, **Verónica Barroso-García**, Ana Cerezo-Hernández, Graciela López-Muñiz, Leila Kheirandish-Gozal, David Gozal, Roberto Hornero, and Félix del Campo.

04/2017: **Prize to the best innovative proposal in the Taller de preparación de propuestas INNOvadoras para participar en el proyecto europeo INNOLABS**, awarded by the Clúster de Salud de Castilla y León (BIOTECYL), conducted by **Verónica Barroso-García**, Fernando Vaquerizo-Villar, Gonzalo C. Gutiérrez-Tobal, and Roberto Hornero.

11/2018: **First José María Ferrero Corral prize in the XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2018)**, awarded by the Sociedad Española de Ingeniería Biomédica (SEIB), for the study entitled "Caracterización de la apnea del sueño infantil mediante nuevas bandas espectrales del EEG", conducted by Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, Jesús Poza, Daniel Álvarez, Fernando Vaquerizo-Villar, **Verónica Barroso-García**, Saúl J. Ruiz-Gómez, Leila Kheirandish-Gozal, Félix del Campo, David Gozal, and Roberto Hornero.

03/2021: **Winner of the 1st edition of the "#Cuéntame11F – Cuéntame3min" scientific divulgation competition**, awarded and organized by "Iniciativa 11 de Febrero" to make visible the work of women scientists, promote technological vocations in girls and adolescents, and remove gender stereotypes in the science and technology field. In this competition, I explained to non-expert audience what sleep apnea is, how it is diagnosed, and what my contribution is in this field of study.

06/2021: **Accésit prize, "Premio de Innovación Educativa del Consejo Social de la Universidad de Valladolid, Edición 2021"**, awarded by the University of Valladolid for the teaching innovation project entitled "InGenias: fomentando las vocaciones tecnológicas y la divulgación científica", coordinated by Noemí Merayo Álvarez.

B.3.2 Grants

07/2017: **"Ayuda para financiar la contratación predoctoral de personal investigador (Orden EDU/602/2016)"**, grant from the Consejería de Educación de la Junta de Castilla y León and cofunded by European Social Fund. Destination place: Biomedical Engineering Group, University of Valladolid, Valladolid, Spain. Duration: July 25, 2017 – December 04, 2021.

07/2017: **"Ayudas Financieras Erasmus+ destinadas al Personal Docente e Investigador de la Universidad de Valladolid para la realización de un curso de Inglés en Instituciones de Enseñanza Superior y Empresas durante el curso académico 2016/2017"**, grant from the University of Valladolid and cofunded by European Funds. Destination place: Atlantic Language School, Dublin, Ireland. Duration: July 31, 2017 – August 11, 2017.

07/2018: **"Ayudas Financieras Erasmus+ destinadas al Personal Docente e Investigador de la Universidad de Valladolid para la realización de un curso de Inglés en Instituciones de Enseñanza Superior y Empresas durante el curso académico 2017/2018"**, grant from the University of Valladolid and cofunded by European Funds. Destination place: Caledonian Language School Ltd., Edinburgh, Scotland. Duration: July 23, 2018 – August 03, 2018.

07/2019: **"Movilidad de Doctorandos. Ayudas para la asistencia a cursos, congresos y jornadas relevantes para el desarrollo de Tesis Doctorales (convocatoria 2019)"**, grant from the University of Valladolid. Destination place: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2019), Berlin, Germany. Duration: July 23, 2019 – July 27, 2019.

08/2019: **"Ayudas Financieras Erasmus+ destinadas al Personal Docente e Investigador de la Universidad de Valladolid para la realización de un curso de Inglés en Instituciones de Enseñanza Superior y Empresas durante el curso académico 2018/2019"**, grant from the University of Valladolid and cofunded by European Funds. Destination place: British School of Grammar & Speaking (BSGS) College, London, England. Duration: August 05, 2019 – August 16, 2019.

06/2021: **"Movilidad de Doctorandos. Ayudas para estancias breves en el desarrollo de tesis doctorales (convocatoria 2021)"**, grant from the University of Valladolid. Destination place: Sleep Technology and Analytics Group, University of Eastern Finland, Kuopio, Finland. Duration: September 01, 2021 – December 03, 2021.

06/2021: **"Ayudas financieras destinadas a estudiantes o recién titulados de la Universidad de Valladolid para la realización de prácticas Erasmus+ en empresas extranjeras con sede en el espacio europeo de educación superior (EEES) y países asociados del programa durante el curso académico 2021/2022"**, grant from the University of Valladolid and cofunded by European Funds. Destination place: Sleep Technology and Analytics Group, University of Eastern Finland, Kuopio, Finland. Duration: September 01, 2021 – December 03, 2021.

11/2021: **"Acciones de Movilidad para estancias del personal CIBER-BBN en grupos externos. Año 2021"**, "Centro de Investigación Biomédica en Red en el área temática de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)". Destination place: Sleep Technology and Analytics Group, University of Eastern Finland, Kuopio, Finland. Duration: September 01, 2021 – December 03, 2021.

Apéndice C

Resumen en castellano

C.1 Introducción

La apnea obstructiva del sueño (AOS) es una patología con alta prevalencia, que se caracteriza por presentar pausas respiratorias durante el sueño (Alonso-Álvarez et al., 2011; Berry et al., 2012). Estas pausas respiratorias pueden ser completas (apneas) o parciales (hipopneas). En niños, una apnea se define como una reducción del flujo aéreo $\geq$90 % durante al menos 2 ciclos respiratorios (Berry et al., 2012). En el caso de una hipopnea, esta se define como una reducción entre 30 % y 90 % del flujo aéreo que ocurre durante al menos 2 ciclos respiratorios y que va acompañada de una desaturación de oxígeno en sangre $\geq$3 % o de un arousal (Berry et al., 2012). La presencia de estos eventos respiratorios conducen a anomalías en el intercambio de gases, hipoxia, microdespertares y sueño fragmentado, lo que afecta negativamente a las funciones fisiológicas y cognitivas de los niños (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). En este sentido, este trastorno puede ocasionarles graves deficiencias neurocognitivas a largo plazo, desordenes de comportamiento, así como disfunciones cardiovasculares, metabólicas y endocrinas, lo que reduce drásticamente su salud y calidad de vida (Alonso-Álvarez et al., 2011; Kaditis et al., 2016). Por lo tanto, es esencial que los niños afectados de AOS sean diagnosticados y tratados lo antes posible.

La AOS pediátrica tiene una alta prevalencia, ya que afecta al 5.7 % de los niños entre 2 y 18 años (Lumeng and Chervin, 2008; Marcus et al., 2012). Según los datos clínicos de atención primaria solicitados a la Subdirección General de Información Sanitaria Española en marzo de 2021, se estima que en España esta enfermedad afecta al 9.56‰ de los niños menores de 15 años. A pesar de su alta prevalencia

la AOS sigue siendo una enfermedad infradiagnosticada, estimándose que el 90 % de los niños afectados aún no cuentan con un diagnóstico médico (Kheirandish-Gozal, 2010). Para diagnosticarla, los niños son remitidos a una unidad pediátrica del sueño especializada donde se les realiza la polisomnografía nocturna (PSG) (Jon, 2009). Este estudio del sueño se basa en registrar simultáneamente varias señales neurofisiológicas y cardiorrespiratorias mientras el niño duerme. Después, los registros son inspeccionados visualmente por especialistas médicos para puntuar los eventos apneicos y calcular el índice de apnea-hipopnea (IAH). Así, el IAH derivado de la PSG es el parámetro clínico utilizado para diagnosticar la AOS pediátrica (Spruyt, 2012). No obstante, la PSG es técnicamente compleja y requiere mucho tiempo, ya que el niño debe permanecer hospitalizado al menos una noche y después todos sus registros fisiológicos deben ser manualmente analizados (Collop et al., 2007; Ryan et al., 1995). También hay que tener en cuenta que el hecho de pernoctar en un ambiente diferente al habitual y con múltiples sensores conectados, puede ser especialmente incómodo para los niños y afectar en sus patrones de sueño (Jon, 2009). Además, el personal, los equipos y las instalaciones necesarias para realizar una PSG no siempre están disponibles, lo que da lugar a largas listas de espera y retrasos en el diagnóstico.

Todas estas limitaciones inherentes de la PSG han llevado a explorar y desarrollar técnicas más simples, que puedan servir como alternativa y agilizar el diagnóstico de la AOS pediátrica (Alonso-Álvarez et al., 2015). En este sentido, diversos estudios han centrado sus investigaciones en analizar automáticamente un conjunto reducido de señales cardiorrespiratorias que, si bien están implicadas en la PSG, podrían ser adquiridas con dispositivos de monitorización portátil en la casa del paciente. Entre estas señales está la saturación de oxígeno en sangre (SpO_2) (Calderón et al., 2020; Garde et al., 2019), la fotopletismografía (*photoplestismography*, PPG) (Dehkordi et al., 2016; Lazaro et al., 2014), el electrocardiograma (ECG) (Shouldice et al., 2004), la variabilidad de la frecuencia cardiaca (*heart rate variability*, HRV) derivada del ECG (Martín-Montero et al., 2020, 2021) o el flujo aéreo respiratorio (FA) (Gutiérrez-Tobal et al., 2015; Jiménez-García et al., 2020). En esta Tesis Doctoral se propone caracterizar exhaustivamente el comportamiento del FA nocturno pediátrico para obtener información relevante que ayude a simplificar el diagnóstico de la AOS en niños. El FA es capaz de reflejar la actividad respiratoria de un sujeto mientras duerme, incluidas las pausas respiratorias asociadas con la AOS (Berry et al., 2012). De hecho, esta señal interviene directamente en la de definición de apnea e hipopnea (Berry et al., 2012). Además, el FA puede ser fácilmente adquirido en la casa del paciente mediante un disposi-

tivo de monitorización portátil dotado de termistor (Collop et al., 2007; Flemons et al., 2003). Todo esto hace que el análisis del FA sea un enfoque prometedor para simplificar el diagnóstico de la AOS infantil.

En esta investigación el FA pediátrico se caracterizó mediante novedosas técnicas de análisis. Los resultados obtenidos dieron lugar a cuatro artículos publicados en revistas indexadas en el *Journal Citation Reports* (JCR), lo que ha permitido presentar esta Tesis Doctoral por compendio de publicaciones. El primer artículo se centró en el análisis del FA y de la variabilidad de la frecuencia respiratoria (*respiratory rate variability*, RRV) utilizando la medida de la tendencia central (*central tendency measure*, CTM) y entropías espectrales (Barroso-García et al., 2017). Este enfoque nos permitió caracterizar la variabilidad e irregularidad de estas señales respiratorias. En el segundo artículo (Barroso-García et al., 2020), la caracterización del FA se llevó a cabo mediante la extracción de nueve características derivadas de los gráficos de recurrencia (*recurrence plots*, RPs). Estas características proporcionaron información novedosa sobre la dinámica subyacente y el espacio fase de la señal de FA. En cuanto al tercer artículo (Barroso-García et al., 2021a), este se centró en caracterizar el FA usando trece características bispectrales. Así, el análisis bispectral permitió obtener información útil sobre la gaussianidad, el acoplamiento de fase y la interacción no lineal de los componentes armónicos del FA. Por último, en el cuarto artículo se utilizó la transformada wavelet discreta (*discrete wavelet transform*, DWT) para analizar el FA pediátrico (Barroso-García et al., 2021b). De este modo, las características derivadas del análisis wavelet ofrecieron información sobre la distribución de energía y los componentes frecuenciales de la señal de FA. Así, cada uno de los enfoques metodológicos propuestos en esta investigación nos permitió descubrir comportamientos del FA nocturno pediátrico que hasta el momento eran desconocidos en el contexto de la AOS. Además, se incorporó el índice de desaturación de oxígeno en sangre del 3 % ($ODI3$), un parámetro clínico utilizado como alternativa subóptima a la PSG cuando esta no está disponible (Kaditis et al., 2015; Van Eyck and Verhulst, 2018). Esto nos permitió evaluar su complementariedad con la información obtenida del FA mediante los diferentes enfoques metodológicos.

C.2 Hipótesis y objetivos

El FA es capaz de reflejar la actividad respiratoria durante el tiempo de sueño, incluidas las pausas respiratorias asociadas con la AOS (Berry et al., 2012). Además, este interviene directamente en la de definición de apnea e hipopnea (Berry et al.,

2012). Por lo tanto, *la señal de FA nocturno recopila información útil para ayudar a diagnosticar la AOS pediátrica.* Algunos estudios previos ya han intentado caracterizar esta señal en el contexto pediátrico (Gutiérrez-Tobal et al., 2015; Jiménez-García et al., 2020). Sus investigaciones se basaron en la aplicación directa de técnicas que habían sido usadas con éxito en adultos. Sin embargo, las diferencias clínicas y diagnósticas entre adultos y niños dificultan el análisis automático de la señal de FA. Este hecho provoca que los métodos aplicados hasta ahora en adultos no sean tan efectivos para caracterizar las particularidades de la AOS en niños. De este modo, se requieren otras técnicas que proporcionen información diferente de los métodos tradicionales previamente usados en adultos. Por ello, asumimos que *nuevas técnicas de extracción de características pueden caracterizar el comportamiento del FA nocturno en niños y mejorar su capacidad diagnóstica.* En este sentido, y dado que FA es una señal dinámica, no lineal y no estacionaria (Martín-González et al., 2018), *los métodos RP, bispectrum y DWT pueden adaptarse a las propiedades intrínsecas del FA nocturno pediátrico y proporcionar información útil relacionada con la AOS.* Además, también asumimos que *los métodos de selección de características pueden identificar las características relevantes y complementarias de FA y maximizar así su capacidad diagnóstica.* No obstante, para que se pueda realizar un diagnóstico automático de la enfermedad se necesita un modelo predictivo que reconozca automáticamente los patrones existentes en los datos extraídos de FA. Así, *los métodos de aprendizaje supervisado pueden ser herramientas útiles para detectar automáticamente la presencia y la severidad de la AOS infantil.* Teniendo en cuenta todas las consideraciones anteriormente mencionadas, esta Tesis Doctoral parte de la hipótesis general de que *la caracterización del FA nocturno mediante novedosos enfoques metodológicos puede ayudar a simplificar el diagnóstico de la AOS pediátrica.*

De acuerdo con esta hipótesis, el objetivo principal de nuestra investigación es *diseñar, implementar y evaluar nuevos métodos automáticos de procesado de señal que permitan caracterizar exhaustivamente el FA nocturno de los niños y ayudar a diagnosticar la AOS infantil.* Para alcanzar este objetivo principal, se propone una metodología en 4 etapas: (i) preprocesado, (ii) extracción de características, (iii) selección de características y (iv) aplicación de métodos de aprendizaje automático. Esta propuesta nos lleva a formular los siguientes objetivos específicos:

I. Mejorar la calidad de los registros de FA mediante novedosos métodos de preprocesado.

II. Caracterizar el FA pediátrico para encontrar las propiedades conductuales que están intrínsecamente relacionadas con la presencia de eventos apneicos.

III. Identificar la información relevante y no redundante del FA nocturno que maximiza su capacidad diagnóstica.

IV. Desarrollar y evaluar modelos de aprendizaje automático para determinar la presencia y gravedad de AOS pediátrica con un alto rendimiento diagnóstico.

C.3 Sujetos

En esta Tesis Doctoral se ha analizado una base de datos de registros de FA pediátrico procedentes del *Comer Children's Hospital* de la Universidad de Chicago (Chicago, IL, EE.UU.). Todos los sujetos involucrados en la investigación fueron remitidos a la unidad pediátrica del sueño de este hospital por presentar síntomas característicos de la AOS, como hipersomnolencia diurna, ronquidos, interrupciones respiratorias y despertares durante la noche. Los niños fueron sometidos a un estudio del sueño tipo 1 (PSG) mediante un dispositivo polisomnográfico digital (Polysmith, Nihon Kohden America Inc., Irvine, CA, EE.UU.). Posteriormente, los especialistas médicos inspeccionaron visualmente los registros polisomnográficos y puntuaron los eventos de apnea e hipopnea siguiendo las reglas de la AASM (Berry et al., 2012). Según el IAH obtenido, cada sujeto pediátrico fue diagnosticado como no AOS (IAH $<$ 1 e/h), AOS leve (1 e/h $\leq$ AHI $<$ 5 e/h), AOS moderada (5 e/h $\leq$ AHI $<$ 10 e/h) o AOS severa (AHI $\geq$ 10 e/h) (Alonso-Álvarez et al., 2011; Hornero et al., 2017; Tan et al., 2014). Los tutores legales de todos los niños dieron su consentimiento informado por escrito para participar en el estudio. Los protocolos de estudio fueron aprobados por el Comité Ético del *Comer Children's Hospital* (números de aprobación: 11-0268-AM017, 09-115-B-AM031, IRB14-1241) conforme a la Declaración de Helsinki.

Inicialmente, la base de datos estaba compuesta por 501 registros de FA pediátrico. Todos ellos fueron adquiridos con un termistor a 200 y 500 Hz durante la realización de la PSG. Así, 501 registros de FA fueron analizados en nuestro primer estudio (Barroso-García et al., 2017). Después, la base de datos fue ampliada e involucró a 946 niños. Estos 946 registros de FA fueron analizados en nuestros otros tres estudios (Barroso-García et al., 2020, 2021a,b). Además de los registros de FA obtenidos mediante termistor, también se incorporó registros de SpO_2 adquiridos con un pulsioxímetro a 200 y 500 Hz. Los datos demográficos y clínicos de los sujetos involucrados son mostrados en las Tablas C.1 y C.2.

Tabla C.1: Base de datos inicial: datos clínicos y demográficos de los niños involucrados.

	Todos	**Sin AOS**	**AOS leve**	**AOS moderada**	**AOS severa**
Sujetos (n)	501	134 (26.7 %)	187 (37.3 %)	76 (15.2 %)	104 (20.8 %)
Edad (años)	6 [6]	7 [6]	6 [4]	5 [5]	4 [5.5]
Varones (n)	314 (62.7 %)	86 (64.2 %)	119 (63.6 %)	43 (56.6 %)	66 (63.4 %)
IMC (kg/m^2)	17.8 [6.5]	17.3 [5.4]	17.9 [6.4]	19.0 [8.9]	17.6 [6.7]
IAH (e/h)	3.2 [7.1]	0.5 [0.6]	2.6 [1.9]	6.8 [2.5]	18.3 [16.1]

Datos presentados como mediana [rango intercuartil] o número (%). IMC: índice de masa corporal, IAH: índice de apnea-hipopnea, AOS: apnea obstructiva del sueño.

Tabla C.2: Base de datos ampliada: datos clínicos y demográficos de los niños involucrados.

	Todos	**Sin AOS**	**AOS leve**	**AOS moderada**	**AOS severa**
Sujetos (n)	946	163 (17.2 %)	386 (40.8 %)	172 (18.2 %)	225 (23.8 %)
Edad (years)	6 [6]	7 [6]	6 [5]	5 [6]	5 [5.3]
Varones (n)	584 (61.7 %)	98 (60.1 %)	242 (62.7 %)	106 (61.6 %)	138 (61.3 %)
IMC (kg/m^2)	17.9 [6.2]	17.4 [5.7]	17.8 [5.5]	18.9 [7.9]	18.3 [7.3]
IAH (e/h)	3.8 [7.8]	0.5 [0.6]	2.5 [1.8]	6.8 [2.4]	19.1 [17.2]

Datos presentados como mediana [rango intercuartil] o número (%). IMC: índice de masa corporal, IAH: índice de apnea-hipopnea, AOS: apnea obstructiva del sueño.

C.4 Métodos

La metodología aplicada en esta investigación consta de cuatro etapas: preprocesado, extracción de características, selección de características y aplicación de métodos de aprendizaje automático.

Se observó que los registros de FA y SpO_2 habían sido registrados a diferentes frecuencias de muestreo, presentaban diferente resolución y contenían artefactos no deseados causados por la pérdida de contacto con el sensor y los movimientos de los niños durante el sueño. Por lo tanto, la primera etapa metodológica consistió en preprocesar los registros para remuestrearlos y eliminar automáticamente los artefactos que contenían. Además, las señales de FA fueron normalizadas para minimizar las posibles diferencias ocasionadas por otras particularidades fisiológicas no relacionadas con la AOS (Várady et al., 2002).

Después, se realizó una etapa de extracción de características para caracterizar exhaustivamente el comportamiento del FA pediátrico nocturno. Teniendo que en cuenta que el FA es una señal cardiorrespiratoria dinámica, no estacionaria y no lineal (Martín-González et al., 2018), se buscaron técnicas de extracción de características que se adaptasen a las propiedades intrínsecas de esta señal. De este modo, los registros de FA fueron analizados mediante CTM, entropías espectrales,

bispectrum, RPs y DWT para obtener información relevante relacionada con la AOS. Así, cada uno de los enfoques propuestos en esta investigación nos permitió descubrir comportamientos del FA nocturno pediátrico que hasta el momento eran desconocidos en el contexto de la AOS infantil. En este sentido, la CTM y las entropías espectrales caracterizaron la variabilidad e irregularidad del FA respiratorio, respectivamente (Abásolo et al., 2006; Jiménez-García et al., 2020). Las características derivadas de los RPs proporcionaron información novedosa sobre la dinámica subyacente y el espacio fase del FA pediátrico (Marwan et al., 2007; Zbilut and Webber, 2006). Respecto al análisis bispectral, este permitió obtener información útil sobre la gaussianidad y el acoplamiento de fase de la señal de FA, así como de la interacción no lineal de sus componentes armónicos (Chua et al., 2010; Martín-Montero et al., 2021). Por último, las características wavelet detectaron cambios en la distribución de energía y los componentes frecuenciales del FA (Rioul and Vetterli, 1991; Vaquerizo-Villar et al., 2018c).

La siguiente etapa metodológica que se llevó a cabo fue la selección de características. Puede resultar que las características extraídas en la etapa anterior no proporcionen información complementaria o e incluso que no sean relevantes para resolver el problema en cuestión (Guyon and Elisseeff, 2003). Además, el uso de una alta e inadecuada cantidad de características puede conducir a un sobreajuste de los modelos, lo que afecta negativamente a su predictibilidad (Guyon and Elisseeff, 2003; Saeys et al., 2007). Por ello, antes de realizar el reconocimiento de patrones, se necesita identificar los subconjuntos óptimos de características que maximizan la capacidad diagnóstica del FA. Así, los métodos *forward stepwise logistic regression* (FSLR) (Hosmer and Lemeshow, 2002) y *fast correlation based filter* (FCBF) (Yu and Liu, 2004) fueron implementados y aplicados para seleccionar automáticamente los subconjuntos óptimos de características relevantes y no redundantes del FA.

Finalmente, en la última etapa se aplicaron métodos de aprendizaje automático para identificar patrones en las características extraídas del FA, inferir comportamientos a partir de ellos y utilizar esta información para detectar automáticamente la presencia y severidad de AOS en niños. Esta etapa se llevó a cabo desde tres enfoques diferentes: clasificación de los sujetos pediátricos como AOS negativo o AOS positivo (tarea de clasificación binaria), clasificación de los niños según su grado de severidad de AOS (tarea de clasificación multiclase) y estimación del IAH de cada sujeto (tarea de regresión). Las tareas de clasificación binaria y multiclase se realizaron mediante algoritmos de regresión logística (RL) y *adaptive boosting* (AdaBoost.M2), respectivamente (Freund and Schapire, 1997; Hosmer and Le-

meshow, 2002). En cuanto a la tarea de regresión, está se realizó a través de redes neuronales perceptrón multicapa (*Multi-layer perceptron*, MLP) y redes MLP con enfoque bayesiano (BY-MLP) (Bishop, 1995).

Además, se aplicaron técnicas estadísticas para analizar e interpretar los resultados obtenidos, así como para evaluar la capacidad diagnóstica y la fiabilidad de los métodos propuestos en esta investigación. De este modo, en esta Tesis Doctoral se utilizaron test estadísticos (Lilliefors, Leneve, χ^2, Fisher, Mann-Whitney y Kruskal-Wallis), técnicas visuales (diagramas de caja y de violín), medidas de rendimiento diagnóstico (sensibilidad, especificidad, precisión, valor predictivo positivo y negativo, razón de verosimilitud positiva y negativa y área bajo la curva *receiver operating characteristic*), medidas de concordancia (coeficiente de correlación intraclase, precisión multiclase e índice kappa de Cohen) y técnicas de validación (*hold-out*, *leave-one-out cross-validation*, *k-fold cross-validation* y *bootstrapping*).

C.5 Resultados y discusión

Cada uno de los diferentes enfoques metodológicos empleados en esta Tesis Doctoral revelaron cambios en el comportamiento del FA nocturno pediátrico que estaban relacionados con la ocurrencia de eventos apneicos. En este sentido, los resultados obtenidos con la CTM sugieren que la AOS podría reducir la variabilidad de la señal de FA e incrementar la del RRV. Así, la variabilidad sería especialmente útil reflejando las particularidades de la AOS severa en FA. Además, podría ser un rasgo común de la enfermedad en la señal RRV y, por lo tanto, un útil indicador para su diagnóstico. En el caso de las entropías espectrales, sus tendencias crecientes en FA indican que los eventos apneicos podrían incrementar la irregularidad de esta señal. No obstante, el análisis estadístico realizado sugiere que la entropía cuadrática refleja mejor los cambios de irregularidad causados por la AOS en FA y que ésta podría ser particularmente más útil en los casos severamente afectados. Respecto a las entropías espectrales de RRV, estas mostraron valores elevados, pero sin diferencias estadísticamente significativas entre grupos de severidad. Este hecho indicaría que la señal RRV tiene un carácter mayoritariamente irregular independientemente de la presencia y grado de AOS.

Estos cambios de variabilidad e irregularidad que experimenta el FA en presencia de eventos apneicos coinciden con los hallazgos realizados mediante el análisis de los RPs. De acuerdo con este análisis, la AOS no sólo reduciría la variabilidad e incrementaría la irregularidad de la señal de FA, sino que también reducirían su complejidad y su divergencia exponencial. Además, la alta densidad de recu-

rrencias alcanzada por los grupos más severos reveló que los eventos apneicos podrían aumentar el acoplamiento de fase de FA. También se observó que, aunque el FA tienen un comportamiento mayoritariamente no estacionario, la AOS es capaz de incrementar el grado de no estacionariedad de esta señal. Además, los eventos apneicos podrían aumentar la duración de los estados laminares de FA, haciendo que ésta no cambie o cambie muy lentamente ya que sus trayectorias permanecerían más tiempo atrapadas en el mismo estado de fase. Así, el tiempo de predicción promedio de la señal también se incrementaría con la severidad. Respecto a la predictibilidad, los resultados obtenidos sugieren que la señal de FA tiene un comportamiento mayoritariamente predecible independientemente de la presencia y grado de AOS. Por el contrario, la laminaridad y la divergencia exponencial podrían ser manifestaciones comunes de la AOS en FA y, por lo tanto, ser unos útiles indicadores para su diagnóstico.

En cuanto al análisis bispectral, se pudo observar que había un foco de acoplamiento entorno de la banda de respiración normal (0.20–0.40 Hz). A medida que el IAH aumentaba, este foco se desvanecía mientras otro aparecía a bajas frecuencias (0.05 Hz). Este hecho sugiere que las apneas e hipopneas reducen la amplitud bispectral y el acoplamiento de fase del FA en la banda de respiración normal y lo redistribuyen en otros componentes frecuenciales relacionados con la AOS. De este modo, los eventos apneicos podrían desplazar el foco de actividad hacia componentes de baja frecuencia. Además, los resultados obtenidos indican que los eventos apneicos podrían reducir el grado de no gaussianidad de la señal de FA, así como la interacción no lineal entre sus componentes armónicos. Respecto a la irregularidad, la tendencia creciente de las entropías bispectrales y de fase sugiere que la AOS podría alterar el comportamiento oscilatorio de FA haciendo que su amplitud y fase cambien sin seguir un determinado patrón. Como consecuencia, la señal de FA sería más irregular al incrementar la severidad de la enfermedad. El análisis estadístico realizado también reveló que la irregularidad de amplitud y fase, el grado de acoplamiento en la banda de respiración normal, la localización del foco de acoplamiento y la centralización de los componentes armónicos podrían ser características propias de un alto desarrollo de la enfermedad. Sin embargo, la potencia bispectral total y máxima, la interacción no lineal entre componentes armónicos y la dispersión de estos, podrían reflejar mejor el comportamiento del FA según los diferentes grados de severidad y, por lo tanto, ser más útiles para diagnosticar la AOS.

Respecto al análisis wavelet, se pudo observar que a medida que se incrementaba la severidad de AOS los sujetos presentaban una señal de detalle de menor amplitud en el rango de frecuencias asociado a la respiración normal. De este modo, la reducción de actividad causada por los eventos apneicos en la banda de respiración normal podría disminuir el valor y el rango de dispersión de los coeficientes de la señal de detalle de FA en su 8º nivel de resolución (D_8). Los resultados también indicaron que la recurrencia de apneas e hipopneas podría perturbar la distribución frecuencial de FA y reducir sus componentes frecuenciales en esta banda. Esto daría lugar a que casi todos los coeficientes de D_8 estuvieses próximos a 0. Además, la AOS podría reducir la energía producida en el rango de frecuencias de respiración normal y redistribuirla en otros rangos frecuenciales asociados con la ocurrencia de eventos apneicos. De este modo, la señal de FA también sería más irregular en términos de energía a medida que aumenta la severidad de la enfermedad. Todas las características wavelet mostraron diferencias significativas entre grupos de severidad. Las mayores se alcanzaron con la mínima amplitud, la centralización y la asimetría de los coeficientes de D_8, lo que destacó su utilidad para reflejar las particularidades de la AOS en FA.

En base a estas consideraciones, los diferentes enfoques metodológicos propuestos en esta Tesis Doctoral permitieron adaptarse a las propiedades intrínsecas del FA nocturno pediátrico, caracterizar su comportamiento y proporcionar información útil relacionada con la AOS. En cuanto a su rendimiento diagnóstico, estos enfoques obtuvieron precisiones moderadas–altas: 60.0 %–81.1 % para 1 e/h, 57.1 %–76.0 % para 5 e/h y 70.5 %–80.6 % para 10 e/h (Barroso-García et al., 2017, 2020, 2021a,b). Así, la mayor precisión alcanzada para 1 y 10 e/h fue mediante las características derivadas de los RPs, mientras que para 5 e/h la precisión mas alta se alcanzó con la CTM y las entropías espectrales. Cabe destacar que este último enfoque obtuvo un par sensibilidad - especificidad mucho más balanceado para los tres umbrales. En este sentido, los modelos alimentados con características de RPs, bispectrum o wavelets, mostraron una sobrestimación de la severidad en 1 y 5 e/h (valores bajos de especificidad) y una ligera subestimación en 10 e/h (valores bajos de sensibilidad). Por el contrario, el $ODI3$ presentó una subestimación de la severidad en 1 y 5 e/h.

Estos efectos indeseados de subestimación y sobreestimación se redujeron cuando la información extraída del FA se combinó con el $ODI3$. De este modo, su uso conjunto consiguió un rendimiento diagnóstico significativamente superior al obtenido individualmente, con precisiones del 78.0 %–83.2 % para 1 e/h, 78.5 %–82.5 % para 5 e/h y 90.2 %–91.0 % para 10 e/h (Barroso-García et al., 2020, 2021a,b). En

cuanto al umbral de 1 e/h, la precisión más alta se volvió a obtener con los RPs. Además, este enfoque alcanzó una razón de verosimilitud negativa notablemente más baja que el resto. Por lo tanto, la información derivada de los RPs de FA podría usarse junto con el $ODI3$ como una técnica sólida para descartar la presencia de AOS pediátrica. En el caso de 5 e/h, la precisión más alta se alcanzó mediante el análisis bispectral. Además, este enfoque obtuvo un par sensibilidad - especificidad más balanceado que el resto. Por lo tanto, la información bispectral del FA podría usarse junto con el $ODI3$ como una herramienta eficaz para discriminar a los niños levemente afectados de aquellos con AOS moderada-severa. Respecto a 10 e/h, tanto el enfoque de RPs como el de wavelets obtuvieron un 91.0 % de precisión. No obstante, el modelo AdaBoost alcanzó una razón de verosimilitud positiva notablemente más alta usando características wavelet. En consecuencia, la información wavelet de FA podría usarse junto con el $ODI3$ como una técnica robusta para determinar la presencia de AOS severa en niños. Teniendo en cuenta todos los umbrales, el modelo AdaBoost alimentado con características wavelet de AF y $ODI3$ proporcionó el mayor rendimiento global para 10 e/h, un par sensibilidad - especificidad mucho más balanceado para 1 e/h y mayor especificidad, valor predictivo positivo y ratio de verosimilitud positiva para 5 e/h, a costa de una precisión ligeramente inferior en estos dos últimos umbrales. Por lo tanto, este podría ser propuesto como nuestro modelo final para diagnosticar la AOS pediátrica.

Así, el mayor rendimiento diagnóstico alcanzado con cada uno de los enfoques propuestos en esta Tesis Doctoral es mostrado en la Tabla C.3. La Tabla C.4 resume las principales características metodológicas y el rendimiento diagnóstico obtenido por otros estudios del estado del arte basados en el diagnóstico automático de la AOS infantil. Como puede observarse, los enfoques propuestos en esta investigación alcanzaron un alto rendimiento diagnóstico en comparación con otros estudios de vanguardia. Aunque en 5 e/h obtuvimos unos resultados dentro de los rangos de rendimiento de estos estudios, nuestra propuesta logró superarlos en 1 y 10 e/h. De este modo, la caracterización del FA nocturno mediante estos novedosos métodos podría incrementar su capacidad diagnóstica y ayudar a detectar la AOS pediátrica. Además, el alto rendimiento de los modelos propuestos sugiere que estos se podrían incorporar en la práctica clínica para identificar automáticamente la presencia y severidad de AOS en niños.

Tabla C.3: Rendimiento diagnóstico obtenido con cada uno de los enfoques metodológicos propuestos en este compendio de publicaciones.

Estudio	#Total /#Test	Señal	Métodos (Extracción /Selección/ Clasificación)	IAH	S (%)	E (%)	P (%)	VPP (%)	VPN (%)	LR+	LR–
Barroso-García et al. (2017)	501/251	FA	Entropías espectrales, *CTM* /FSLR/LR	1	60.5	58.6	60.0	81.2	25.0	1.1	0.9
				5	65.0	80.6	76.0	70.7	78.2	3.6	0.4
				10	83.3	79.0	80.0	52.8	93.5	4.0	0.2
Barroso-García et al. (2020)	946/376	FA, SpO_2	Gráficos de recurrencia, *ODI*3/FCBF/ BY-MLP	1	97.7	22.2	83.2	84.1	69.6	1.3	0.1
				5	78.7	78.3	78.5	68.5	86.0	3.6	0.2
				10	78.8	94.3	91.0	78.8	94.3	13.7	0.2
Barroso-García et al. (2021a)	946/376	FA, SpO_2	Bispectrum, *ODI*3/FCBF/ MLP	1	98.0	15.3	82.2	83.0	65.0	1.2	0.1
				5	81.6	83.0	82.5	74.2	88.3	4.9	0.2
				10	72.3	95.0	90.2	796	92.7	15.0	0.3
Barroso-García et al. (2021b)	946/376	FA, SpO_2	Wavelet, *ODI*3/FCBF/ AdaBoost.M2	1	80.3	68.1	78.0	91.5	44.9	2.6	0.3
				5	68.0	90.3	81.9	80.8	82.5	7.2	0.4
				10	72.4	96.0	90.1	83.0	92.8	19.0	0.3
Barroso-García et al. (2021b)	946/376	FA, SpO_2	Wavelet, *ODI*3/FCBF/ BY-MLP	1	91.2	43.3	82.0	87.2	53.6	1.6	0.2
				5	79.3	83.8	82.1	74.6	87,2	5.0	0.3
				10	74.9	95.0	90.7	80.0	93.3	15.6	0.3

IAH: índice de apnea-hipopnea, S: sensibilidad, E: especificidad, P: precisión, VPP: valor predictivo positivo, VPN: valor predictivo negativo, LR+: razón de verosimilitud positiva, LR–: razón de verosimilitud negativa, SpO_2: señal de oximetría, FA: señal de flujo aéreo, *ODI*3: índice de desaturación de oxígeno en sangre del 3 %, *CTM*: *central tendency measure*, FSLR: *forward stepwise logistic regression*, FCBF: *fast correlation based filter*, LR: *logistic regression*, MLP: *multi-layer perceptron*, BY-MLP: *Bayesian MLP*, AdaBoost: *adaptive boosting*.

C.6 Conclusiones

De acuerdo con los resultados obtenidos a lo largo de esta investigación, se pueden derivar las siguientes conclusiones:

1) El *CTM*, las entropías espectrales, los RPs, el bispectrum y el análisis wavelet pueden caracterizar el comportamiento del FA nocturno pediátrico, adaptarse a sus propiedades intrínsecas y proporcionar información útil relacionada con la AOS.

2) Tras aplicar el *CTM* y las entropías espectrales podemos concluir que la AOS reduce la variabilidad e incrementa la irregularidad de la señal de FA. Además, los eventos apneicos también incrementan la variabilidad de la señal RRV.

Tabla C.4: Rendimiento diagnóstico obtenido en otros estudios del estado del arte.

Estudio	#Total /#Test	Señal	Métodos (Extracción /Selección/ Clasificación)	IAH	S (%)	E (%)	P (%)	VPP (%)	VPN (%)	LR+	LR–
Shouldice et al. (2004)	50/25	ECG	Temporal, espectral/–/QDA	1	85.7	81.8	84.0	85.7	81.8	4.7	0.2
Martín-Montero et al. (2020)	1738/757	ECG	Espectral de HRV/–/LDA	1 5 10	42.5 50.0 63.8	72.3 80.9 84.7	52.6 76.4 82.8	75.0 31.3 29.5	39.1 90.3 95.9	1.5 2.6 4.2	0.8 0.6 0.4
Martín-Montero et al. (2021)	1738/757	ECG	Bispectral de HRV/FCBF /MLP	1 5 10	76.3 62.5 66.7	38.3 84.2 91.6	63.4 81.0 89.3	70.7 40.7 44.2	45.5 92.8 96.5	1.2 4.0 7.9	0.6 0.4 0.4
Gil et al. (2009)	21/21	PPG	Espectral de HRV, DAP /$Wrapper$/QDA	5	87.5	71.4	80.0	–	–	3.1	0.2
Lazaro et al. (2014)	21/21	PPG	Espectral de PRV, DAP /$Wrapper$/LDA	5	100	71.4	86.7	–	–	3.5	0
Dehkordi et al. (2016)	146/146	PPG	Temporal, espectral, DFA /LASSO/LASSO	5	76.0	68.0	71.0	–	–	2.4	0.4
Garde et al. (2014a)	146/146	SpO_2 PRV	Temporal, no lineal, espectral /Optimización AUC/LDA	5	88.4	83.6	84.9	76.9	92.6	5.4	0.1
Hornero et al. (2017)	4191 /3602	SpO_2	Temporal, no lineal, espectral $ODI3$/FCBF /MLP	1 5 10	84.0 68.2 68.7	53.2 87.2 94.1	75.2 81.7 90.2	81.6 68.6 67.7	53.7 87.0 94.3	1.8 5.3 11.6	0.3 0.4 0.3
Vaquerizo-Villar et al. (2018b)	298/75	SpO_2	Espectral, bispectral, $ODI3$, antropométricas/FCBF/MLP	5 10	61.8 60.0	97.6 94.5	81.3 85.3	95.5 80.0	75.5 86.7	25.3 11.0	0.4 0.4
Vaquerizo-Villar et al. (2018c)	981/392	SpO_2	Temporal, espectral, wavelet, $ODI3$ /FCBF/SVM	5	71.9	91.1	84.0	83.8	84.5	14.6	0.3
Xu et al. (2018)	432/432	SpO_2	Temporal, no lineal, espectral $ODI3$/–/MPL	1 5 10	95.3 77.8 73.5	19.1 80.5 92.7	79.6 79.4 88.2	82.0 72.3 75.8	51.5 84.7 91.9	1.2 4.0 10.1	0.2 0.3 0.3
Garde et al. (2019)	207/207	SpO_2, PRV	Temporal, espectral/ $Stepwise$/LR	1 5 10	80.0 85.0 82.0	65.0 79.0 91.0	75.0 82.0 89.0	– – –	– – –	2.3 4.1 9.1	0.3 0.2 0.2
Calderón et al. (2020)	453/453	SpO_2	Índices oximetría/–/LR	5	62.0	96.0	79.0	94.3	–	15.5	0.4
Gutiérrez-Tobal et al. (2015)	50/50	FA, SpO_2	Espectral, $ODI3$/FSLR/LR	3	85.9	87.4	86.3	88.4	85.8	6.8	0.2
Jiménez-García et al. (2020)	974/390	FA, SpO_2	Temporal, no lineal, espectral $ODI3$/FCBF/ AdaBoost.M2	1 5 10	92.1 76.0 62.7	36.0 85.7 97.7	81.3 82.1 90.3	85.8 76.0 88.1	51.9 85.7 90.6	1.4 5.3 27.5	0.2 0.3 0.4

IAH: índice de apnea-hipopnea, S: sensibilidad, E: especificidad, P: precisión, VPP: valor predictivo positivo, VPN: valor predictivo negativo, LR+: razón de verosimilitud positiva, LR–: razón de verosimilitud negativa, ECG: electrocardiograma, PPG: fotopletismografía, SpO_2: señal de oximetría, FA: señal de flujo aéreo, PRV: *pulse rate variability*, DAP: *decreases in amplitude fluctuations of PPG*, HRV: *heart rate variability*, $ODI3$: índice de desaturación de oxígeno en sangre del 3 %, AUC: área bajo la *receiver operating characteristic curve*, FSLR: *forward stepwise logistic regression*, LASSO: *least absolute shrinkage and selection operator*, FCBF: *fast correlation based filter*, QDA: *quadratic discriminant analysis*, LDA: *linear discriminant analysis*, LR: *logistic regression*, MLP: *multi-layer perceptron*, SVM: *support vector machine*, AdaBoost: *adaptive boosting*.

3) Según la información proporcionada por los RPs, la AOS modifica la dinámica subyacente y el espacio fase del FA nocturno. Concretamente, los eventos apneicos disminuyen la variabilidad, la estacionariedad y la complejidad de la señal de FA, así como la divergencia exponencial de su espacio fase. También se puede concluir que la AOS aumenta el tiempo durante el cual el FA permanece en un determinado estado del espacio fase. Además, su tiempo medio de predicción y su irregularidad son mayores a medida que aumenta el IAH.

4) La interpretación de las características bispectrales nos llevaron a concluir que la AOS reduce la no gaussianidad del FA, así como la interacción no lineal de sus componentes armónicos. La AOS infantil también disminuye el acoplamiento de fase en la banda de respiración normal y desplaza el foco de acoplamiento hacia componentes de baja frecuencia relacionados con los eventos apneicos. Además, la irregularidad de la señal de FA aumenta en términos de amplitud y fase cuando aumenta la severidad de la enfermedad.

5) De acuerdo con la información proporcionada por el análisis wavelet podemos concluir que la AOS perturba la distribución de energía y de los componentes frecuenciales de la señal de FA. Concretamente, los eventos apneicos reducen la amplitud de la señal de detalle de FA y la energía producida en la banda de respiración normal. En esta banda, los componentes frecuenciales del FA también disminuyen con la AOS. Además, su irregularidad aumenta en términos de energía a medida que aumenta el IAH.

6) La información que proporciona la señal de FA sobre la AOS es adicional y complementaria a la que aporta la señal RRV. Concretamente, existe complementariedad entre la irregularidad del FA y la variabilidad del RRV, lo que destaca su capacidad para caracterizar diferentes rasgos de la AOS infantil.

7) La información sobre la divergencia exponencial del espacio fase del FA es más útil para caracterizar las particularidades de la AOS pediátrica que la que proporcionan otras características derivadas de los RPs.

8) Los diferentes enfoques bispectrales ofrecen información complementaria sobre los efectos que los eventos apneicos provocan en la señal de FA. Por lo tanto, su uso conjunto proporciona una caracterización más completa del comportamiento de esta señal.

9) La información sobre los cambios que experimenta la asimetría de la distribución de los coeficientes del FA es más útil para caracterizar la AOS infantil que la que proporcionan otras características wavelet.

10) La información proporcionada por el FA a través de los diferentes enfoques metodológicos es complementaria a la información proporcionada por el clásico $ODI3$. Por lo tanto, su uso conjunto sería especialmente útil para diagnosticar con precisión la AOS pediátrica.

11) Los métodos de procesado de FA propuestos permiten obtener modelos predictivos de alta precisión para diagnosticar automáticamente la presencia y severidad de la AOS infantil. El modelo BY-MLP diseñado con características RP y el $ODI3$ alcanzó el rendimiento diagnóstico más alto para descartar la presencia de AOS (1 e/h), la MLP diseñada con características bispectrales y el $ODI3$ lo obtuvo para distinguir entre niños afectados levemente de AOS y aquellos con un grado moderado-a-severo (5 e/h), y el modelo AdaBoost construido con características wavelet y el $ODI3$ logró el mayor rendimiento para confirmar la presencia de AOS severa (10 e/h). Considerando conjuntamente el rendimiento alcanzado para todos los umbrales, este último modelo podría ser propuesto como el modelo final.

De acuerdo con las consideraciones anteriormente expuestas, podemos concluir que la caracterización del FA nocturno mediante estos novedosos métodos puede ayudar a simplificar el diagnóstico de la AOS pediátrica. Además, el alto rendimiento de los modelos propuestos sugiere que estos se podrían incorporar en la práctica clínica para identificar automáticamente la presencia y severidad de AOS infantil.

Bibliography

Abásolo, D., Hornero, R., Gómez, C., García, M., López, M., 2006. Analysis of eeg background activity in alzheimer's disease patients with lempel–ziv complexity and central tendency measure. Medical Engineering & Physics 28 (4), 315–322.

Acharya, U. R., Sree, S. V., Chattopadhyay, S., Yu, W., Ang, P. C. A., 2011. Application of recurrence quantification analysis for the automated identification of epileptic EEG signals. International Journal of Neural Systems 21 (3), 199–211.

Alonso-Álvarez, M. L., Canet, T., Cubell-Alarco, M., Estivill, E., Fernández-Julián, E., Gozal, D., Jurado-Luque, M. J., Lluch-Roselló, M. A., Martínez-Pérez, F., Merino-Andreu, M., Pin-Arboledas, G., Roure, N., Sanmartí, F. X., Sans-Capdevila, Ó., Segarra-Isern, F., Tomás-Vila, M., Terán-Santos, J., 2011. Consensus document on sleep apnea-hypopnea syndrome in children. Archivos de Bronconeumologia 47 (SUPPL. 5), 2–18.

Alonso-Álvarez, M. L., Terán-Santos, J., Ordax Carbajo, E., Cordero-Guevara, J. A., Navazo-Egüia, A. I., Kheirandish-Gozal, L., Gozal, D., 2015. Reliability of home respiratory polygraphy for the diagnosis of sleep apnea in children. Chest 147 (4), 1020–1028.

Alpaydin, E., 2014. Introduction to Machine Learning, 3rd Edition. MIT Press.

Álvarez, D., Crespo, A., Vaquerizo-Villar, F., Gutiérrez-Tobal, G. C., Cerezo-Hernández, A., Barroso-García, V., Ansermino, J. M., Dumont, G. A., Hornero, R., del Campo, F., Garde, A., Gutierrez-Tobal, G. C., Cerezo-Hernández, A., Barroso-García, V., Ansermino, J. M., Dumont, G. A., Hornero, R., Campo, F. D., Garde, A., 2018. Symbolic dynamics to enhance diagnostic ability of portable oximetry from the phone oximeter in the detection of paediatric sleep apnoea. Physiological Measurement 39 (10), 104002.

Álvarez, D., Cerezo-Hernández, A., Crespo, A., Gutiérrez-Tobal, G. C., Vaquerizo-Villar, F., Barroso-García, V., Moreno, F., Arroyo, C. A., Ruiz, T., Hornero, R., del Campo, F., 2020. A machine learning-based test for adult sleep apnoea screening at home using oximetry and airflow. Scientific Reports 10 (1), 5332.

Arens, R., Muzumdar, H., 2010. Sleep, sleep disordered breathing, and nocturnal hypoventilation in children with neuromuscular diseases. Paediatric Respiratory Reviews 11 (1), 24–30.

Atri, R., Mohebbi, M., 2015. Obstructive sleep apnea detection using spectrum and bispectrum analysis of single-lead ECG signal. Physiological Measurement 36 (9), 1963–1980.

Barroso-García, V., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Álvarez, D., Vaquerizo-Villar, F., Crespo, A., del Campo, F., Gozal, D., Hornero, R., 2017. Irregularity and Variability Analysis of Airflow Recordings to Facilitate the Diagnosis of Paediatric Sleep Apnoea-Hypopnoea Syndrome. Entropy 19 (9), 447.

Barroso-García, V., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Álvarez, D., Vaquerizo-Villar, F., Núñez, P., del Campo, F., Gozal, D., Hornero, R., 2020. Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis. Computer Methods and Programs in Biomedicine 183, 105083.

Barroso-García, V., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Vaquerizo-Villar, F., Álvarez, D., del Campo, F., Gozal, D., Hornero, R., 2021a. Bispectral analysis of overnight airflow to improve the pediatric sleep apnea diagnosis. Computers in Biology and Medicine 129, 104167.

Barroso-García, V., Gutiérrez-Tobal, G. C., Gozal, D., Vaquerizo-Villar, F., Álvarez, D., del Campo, F., Kheirandish-Gozal, L., Hornero, R., 2021b. Wavelet Analysis of Overnight Airflow to Detect Obstructive Sleep Apnea in Children. Sensors 21 (4), 1491.

Bartko, J. J., 1966. The intraclass correlation coefficient as a measure of reliability. Psychological reports 19 (1), 3–11.

Berry, R. B., Budhiraja, R., Gottlieb, D. J., Gozal, D., Iber, C., Kapur, V. K., Marcus, C. L., Mehra, R., Parthasarathy, S., Quan, S. F., Redline, S., Strohl, K. P., Ward, S. L. D., Tangredi, M. M., 2012. Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Journal of Clinical Sleep Medicine 08 (05), 597–619.

Bishop, C. M., 1995. Neural networks for pattern recognition, 1st Edition. Oxford university Press.

Bishop, C. M., 2006. Pattern recognition and machine learning, 1st Edition. Springer.

Bitners, A. C., Arens, R., 2020. Evaluation and Management of Children with Obstructive Sleep Apnea Syndrome. Lung 198 (2), 257–270.

Blechner, M., Williamson, A. A., 2016. Consequences of Obstructive Sleep Apnea in Children. Current Problems in Pediatric and Adolescent Health Care 46 (1), 19–26.

Bronzino, J. D., Peterson, D. R., 2014. Biomedical engineering fundamentals, 4th Edition. CRC press.

Calderón, J. M., Álvarez-Pitti, J., Cuenca, I., Ponce, F., Redon, P., 2020. Development of a minimally invasive screening tool to identify obese pediatric population at risk of obstructive sleep apnea/hypopnea syndrome. Bioengineering 7 (4), 131.

Carroll, J. L., Loughlin, G. M., 1992. Diagnostic criteria for obstructive sleep apnea syndrome in children. Pediatric Pulmonology 14 (2), 71–74.

Chang, L., Wu, J., Cao, L., 2013. Combination of symptoms and oxygen desaturation index in predicting childhood obstructive sleep apnea. International Journal of Pediatric Otorhinolaryngology 77 (3), 365–371.

Chen, C.-C., Barnhart, H. X., 2008. Comparison of icc and ccc for assessing agreement for data without and with replications. Computational Statistics & Data Analysis 53 (2), 554–564.

Chiner, E., Cánovas, C., Molina, V., Sancho-Chust, J. N., Vañes, S., Pastor, E., Martinez-Garcia, M. A., 2020. Home Respiratory Polygraphy is Useful in the Diagnosis of Childhood Obstructive Sleep Apnea Syndrome. Journal of Clinical Medicine 9 (7), 2067.

Chua, K. C., Chandran, V., Acharya, U. R., Lim, C. M., 2010. Application of higher order statistics/spectra in biomedical signals-A review. Medical Engineering and Physics 32 (7), 679–689.

Clínica Dental Ceballos, 2017. Confirman que la apnea del sueño favorece el crecimiento y la propagación del Cáncer. https://www.clinicadentalceballos.com/apnea-del-sueno-cancer/, last access: 03–01–2022.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20 (1), 37–46.

Cohen, M. E., Hudson, D. L., Deedwania, P. C., 1996. Applying continuous chaotic modeling to cardiac signal analysis. IEEE Engineering in Medicine and Biology Magazine 15 (5), 97–102.

Collop, N. A., Anderson, W. M. D., Boehlecke, B., Claman, D., Goldberg, R., Gottlieb, D. J., Hudgel, D., Sateia, M., Schwab, R., 2007. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients.

Cysarz, D., Zerm, R., Bettermann, H., Frühwirth, M., Moser, M., Kröz, M., 2008. Comparison of Respiratory Rates Derived from Heart Rate Variability, ECG Amplitude, and Nasal/Oral Airflow. Annals of Biomedical Engineering 36 (12), 2085–2094.

Daubechies, I., 1990. The Wavelet Transform, Time-Frequency Localization and Signal Analysis. IEEE Transactions on Information Theory 36 (5), 961–1005.

Daubechies, I., 1992. Ten Lectures on Wavelets, 1st Edition. Society for Industrial and Applied Mathematics.

Deeks, J. J., Altman, D. G., 2004. Diagnostic tests 4: likelihood ratios. BMJ (Clinical research ed.) 329 (7458), 168–169.

Dehkordi, P., Garde, A., Karlen, W., Petersen, C. L., Wensley, D., Dumont, G. A., Mark Ansermino, J., 2016. Evaluation of cardiac modulation in children in response to apnea/hypopnea using the Phone Oximeter™. Physiological Measurement 37 (2), 187–202.

Deng, Z. D., Poon, C. S., Arzeno, N. M., Katz, E. S., 2006. Heart rate variability in pediatric obstructive sleep apnea. In: Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. New York, NY, USA, pp. 3565–3568.

Efron, B., Tibshirani, R. J., 1994. An introduction to the bootstrap. CRC press.

Emin Tagluk, M., Sezgin, N., 2011. A new approach for estimation of obstructive sleep apnea syndrome. Expert Systems with Applications 38 (5), 5346–5351.

Farcomeni, A., 2008. A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. Statistical methods in medical research 17 (4), 347–388.

Fawcett, T., 2006. An introduction to roc analysis. Pattern Recognition Letters 27 (8), 861–874.

Ferber, R., Millman, R., Coppola, M., Fleetham, J., Murray, C. F., Iber, C., McCall, V., Pressman, M., Sanders, M., Strohl, K., Votteri, B., Williams, A., 1994. ASDA standards of practice: Portable recording in the assessment of obstructive sleep apnea. Sleep 17 (4), 378–392.

Figliola, A., Serrano, E., 1997. Analysis of physiological time series using wavelet transforms. IEEE Engineering in Medicine and Biology Magazine 16 (3), 74–79.

Fisher, R. A., 1922. On the interpretation of χ 2 from contingency tables, and the calculation of p. Journal of the Royal Statistical Society 85 (1), 87–94.

Fleming, S., Thompson, M., Stevens, R., Heneghan, C., Plüddemann, A., MacOnochie, I., Tarassenko, L., Mant, D., 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 377 (9770), 1011–1018.

Flemons, W. W., Littner, M. R., Rowley, J. A., Gay, P., Anderson, W. M. D., Hudgel, D. W., McEvoy, R. D., Loube, D. I., 2003. Home Diagnosis of Sleep Apnea: A Systematic Review of the Literature - An Evidence Review Cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. Chest 124 (4), 1543–1579.

Flemons, W. W., Littner, M. R., 2003. Measuring Agreement Between Diagnostic Devices. Chest 124 (4), 1535–1542.

Freund, Y., Schapire, R. E., 1997. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer and System Sciences 55 (1), 119–139.

Garde, A., Dehkordi, P., Karlen, W., Wensley, D., Ansermino, J. M., Dumont, G. A., nov 2014a. Development of a Screening Tool for Sleep Disordered Breathing in Children Using the Phone Oximeter™. PLoS ONE 9 (11), e112959.

Garde, A., Karlen, W., Dehkordi, P., Ansermino, J. M., Dumont, G. A., 2014b. Oxygen saturation resolution influences regularity measurements. In: Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. Chicago, IL, USA, pp. 2257–2260.

Garde, A., Hoppenbrouwer, X., Dehkordi, P., Zhou, G., Rollinson, A. U., Wensley, D., Dumont, G. A., Ansermino, J. M., 2019. Pediatric pulse oximetry-based OSA screening at different thresholds of the apnea-hypopnea index with an expression of uncertainty for inconclusive classifications. Sleep Medicine 60, 45–52.

Gil, E., Mendez, M., Vergara, J. M., Cerutti, S., Bianchi, A. M., Laguna, P., 2009. Discrimination of sleep-apnea-related decreases in the amplitude fluctuations of ppg signal in children by hrv analysis. IEEE transactions on biomedical engineering 56 (4), 1005–1014.

Gil, E., Bailon, R., Vergara, J. M., Laguna, P., 2010. PTT Variability for Discrimination of Sleep Apnea Related Decreases in the Amplitude Fluctuations of PPG Signal in Children. IEEE Transactions on Biomedical Engineering 57 (5), 1079–1088.

Gogolewski, D., 2020. Influence of the edge effect on the wavelet analysis process. Measurement: Journal of the International Measurement Confederation 152, 107314.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep learning. MIT press.

Grandner, M. A., Rosenberger, M. E., 2019. Actigraphic sleep tracking and wearables: Historical context, scientific applications and guidelines, limitations, and considerations for commercial sleep devices. In: Sleep and Health. Elsevier Inc., pp. 147–157.

Gutierrez-Tobal, G. C., Hornero, R., Álvarez, D., Marcos, J., del Campo, F., 2012. Linear and nonlinear analysis of airflow recordings to help in sleep apnoea–hypopnoea syndrome diagnosis. Physiological Measurement 33 (7), 1261–75.

Gutiérrez-Tobal, G. C., Álvarez, D., Marcos, J. V., Del Campo, F., Hornero, R., 2013. Pattern recognition in airflow recordings to assist in the sleep apnoea-hypopnoea syndrome diagnosis. Medical and Biological Engineering and Computing 51 (12), 1367–1380.

Gutiérrez-Tobal, G. C., Alonso-Álvarez, M. L., Álvarez, D., del Campo, F., Terán-Santos, J., Hornero, R., 2015. Diagnosis of pediatric obstructive sleep apnea: Preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients' home. Biomedical Signal Processing and Control 18, 401–407.

Gutiérrez-Tobal, G. C., Member, S., Alvarez, D., del Campo, F., Hornero, R., Member, S., 2016. Utility of AdaBoost to Detect Sleep Apnea-Hypopnea Syndrome From Single-Channel Airflow. IEEE Transactions on Biomedical Engineering 63 (3), 636–646.

Gutiérrez-Tobal, G. C., Álvarez, D., Kheirandish-Gozal, L., Del Campo, F., Gozal, D., Hornero, R., 2021. Reliability of machine learning to diagnose pediatric obstructive sleep apnea: Systematic review and meta-analysis. Pediatric Pulmonology.

Guyon, I., Elisseeff, A., 2003. An Introduction to Variable and Feature Selection. Journal of Machine Learning Research 3 (Mar), 1157–1182.

Hadaś-Dyduch, M., 2018. Nonlinear Alleviation of Edge Effects in the Context of Minimizing Prediction Errors. International Journal of Economics and Finance 10 (2), 161–168.

Han, J., Shin, H. B., Jeong, D. U., Park, K. S., 2008. Detection of apneic events from single channel nasal airflow using 2nd derivative method. Computer Methods and Programs in Biomedicine 91 (3), 199–207.

Hintze, J. L., Nelson, R. D., 1998. Violin plots: a box plot-density trace synergism. The American Statistician 52 (2), 181–184.

Hornero, R., Escudero, J., Fernández, A., Poza, J., Gómez, C., 2008. Spectral and nonlinear analyses of meg background activity in patients with alzheimer's disease. IEEE Transactions on Biomedical Engineering 55 (6), 1658–1665.

Hornero, R., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Philby, M. F., Alonso-Álvarez, M. L., Álvarez, D., Dayyat, E. A., Xu, Z., Huang, Y.-S., Tamae Kakazu, M., Li, A. M., Van Eyck, A., Brockmann, P. E., Ehsan, Z., Simakajornboon, N., Kaditis, A. G., Vaquerizo-Villar, F., Crespo Sedano, A., Sans Capdevila, O., von Lukowicz, M., Terán-Santos, J., Del Campo, F., Poets, C. F., Ferreira, R., Bertran, K., Zhang, Y., Schuen, J., Verhulst, S., Gozal, D., 2017. Nocturnal Oximetry-based Evaluation of Habitually Snoring Children. American Journal of Respiratory and Critical Care Medicine 196 (12), 1591–1598.

Hosmer, D. W., Lemeshow, S., 2002. Applied logistic regression. John Wiley & Sons.

Iber, C., Ancou-Israel, S., Chesson, A. L., Quan, S. F., 2007. The AASM Manual for the Scoring of Sleep and Associated Events Rules, Terminology and Technical Specifications. American Academy of Sleep Medicine (AASM).

Inouye, T., Shinosaki, K., Sakamoto, H., Toi, S., Ukai, S., Iyama, A., Katsuda, Y., Hirano, M., 1991. Quantification of eeg irregularity by use of the entropy of the power spectrum. Electroencephalography and Clinical Neurophysiology 79 (3), 204–210.

Jain, A. K., Duin, R. P. W., Mao, J., 2000. Statistical pattern recognition: A review. IEEE Transactions on pattern analysis and machine intelligence 22 (1), 4–37.

Jeong, J., Gore, J. C., Peterson, B. S., 2001. Mutual information analysis of the EEG in patients with Alzheimer's disease. Clinical Neurophysiology 112 (5), 827–835.

Jiménez-García, J., Gutiérrez-Tobal, G. C., García, M., Kheirandish-Gozal, L., Martín-Montero, A., Álvarez, D., del Campo, F., Gozal, D., Hornero, R., 2020. Assessment of Airflow and Oximetry Signals to Detect Pediatric Sleep Apnea-Hypopnea Syndrome Using AdaBoost. Entropy 22 (6), 670.

Jon, C., 2009. Polysomnography in Children. In: Pediatric Otolaryngology for the Clinician. Humana Press, pp. 35–47.

Kaditis, A., Kheirandish-Gozal, L., Gozal, D., 2015. Pediatric OSAS: Oximetry can provide answers when polysomnography is not available. Sleep Medicine Reviews 27, 96–105.

Kaditis, A. G., Alvarez, M. L. A., Boudewyns, A., Alexopoulos, E. I., Ersu, R., Joosten, K., Larramona, H., Miano, S., Narang, I., Trang, H., Tsaoussoglou, M., Vandenbussche, N., Villa, M. P., Waardenburg, D. V., Weber, S., Verhulst, S., 2016. Obstructive sleep disordered breathing in 2- to 18-year-old children: Diagnosis and management. European Respiratory Journal 47 (1), 69–94.

Kapur, V. K., Auckley, D. H., Chowdhuri, S., Kuhlmann, D. C., Mehra, R., Ramar, K., Harrod, C. G., 2017. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: An American academy of sleep medicine clinical practice guideline. Journal of Clinical Sleep Medicine 13 (3), 479–504.

Kennel, M. B., Brown, R., Abarbanel, H. D. I., 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Physical Review A 45 (6), 3403–3411.

Kermit, M., Eide, Å. J., Lindblad, T., Waldemark, K., 2000. Treatment of obstructive sleep apnea syndrome by monitoring patients airflow signals. Pattern Recognition Letters 21 (3), 277–281.

Kheirandish-Gozal, L., 2010. What is "abnormal" in pediatric sleep? Respiratory care 55 (10), 1366–1376.

Kirk, V. G., Bohn, S. G., Flemons, W. W., Remmers, J. E., 2003. Comparison of Home Oximetry Monitoring with Laboratory Polysomnography in Children. Chest 124 (5), 1702–1708.

Koley, B. L., Dey, D., 2013. Automatic detection of sleep apnea and hypopnea events from single channel measurement of respiration signal employing ensemble binary SVM classifiers. Measurement: Journal of the International Measurement Confederation 46 (7), 2082–2092.

Korten, J. B., Haddad, G. G., 1989. Respiratory waveform pattern recognition using digital techniques. Computers in Biology and Medicine 19 (4), 207–217.

Kruskal, W. H., Wallis, W. A., 1952. Use of ranks in one-criterion variance analysis. Journal of the American statistical Association 47 (260), 583–621.

Lazaro, J., Gil, E., Vergara, J. M., Laguna, P., 2014. Pulse Rate Variability Analysis for Discrimination of Sleep-Apnea-Related Decreases in the Amplitude Fluctuations of Pulse Photoplethysmographic Signal in Children. IEEE Journal of Biomedical and Health Informatics 18 (1), 240–246.

Lee, J., Steele, C. M., Chau, T., 2011. Classification of healthy and abnormal swallows based on accelerometry and nasal airflow signals. Artificial Intelligence in Medicine 52 (1), 17–25.

Leonardi, G., 2018. A method for the computation of entropy in the recurrence quantification analysis of categorical time series. Physica A: Statistical Mechanics and its Applications 512, 824–836.

Levene, H., 1961. Robust tests for equality of variances. Contributions to probability and statistics. Essays in honor of Harold Hotelling, 279–292.

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

Lumeng, J. C., Chervin, R. D., 2008. Epidemiology of pediatric obstructive sleep apnea. Proceedings of the American Thoracic Society 5 (2), 242–252.

Magalang, U. J., Dmochowski, J., Veeramachaneni, S., Draw, A., Mador, M. J., El-Solh, A., Grant, B. J., 2003. Prediction of the Apnea-Hypopnea Index From Overnight Pulse Oximetry. Chest 124 (5), 1694–1701.

Mallat, S. G., 1989. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (7), 674–693.

Mann, H. B., Whitney, D. R., 1947. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics 18 (1), 50 – 60.

Marcus, C. L., 2000. Obstructive sleep apnea syndrome: Differences between children and adults. Sleep 23 (SUPPL. 4), S140–S141.

Marcus, C. L., Brooks, L. J., Ward, S. D., Draper, K. A., Gozal, D., Halbower, A. C., Jones, J., Lehmann, C., Schechter, M. S., Sheldon, S., Shiffman, R. N., Spruyt, K., 2012. Diagnosis and management of childhood obstructive sleep apnea syndrome. Pediatrics 130 (3), e714–e755.

Martín-González, S., Navarro-Mesa, J. L., Juliá-Serdá, G., Ramírez-Ávila, G. M., Ravelo-García, A. G., 2018. Improving the understanding of sleep apnea characterization using Recurrence Quantification Analysis by defining overall acceptable values for the dimensionality of the system, the delay, and the distance threshold. PLOS ONE 13 (4), e0194462.

Martín-Montero, A., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Jiménez-García, J., Álvarez, D., del Campo, F., Gozal, D., Hornero, R., 2020. Heart rate variability spectrum characteristics in children with sleep apnea. Pediatric Research, 1–9.

Martín-Montero, A., Gutiérrez-Tobal, G. C., Gozal, D., Barroso-García, V., Álvarez, D., Del Campo, F., Kheirandish-Gozal, L., Hornero, R., 2021. Bispectral analysis of heart rate variability to characterize and help diagnose pediatric sleep apnea. Entropy 23 (8), 1016.

Marwan, N., Carmen Romano, M., Thiel, M., Kurths, J., 2007. Recurrence plots for the analysis of complex systems. Physics Reports 438 (5-6), 237–329.

McCamley, J., Denton, W., Lyden, E., Yentes, J. M., 2017. Measuring coupling of rhythmical time series using cross sample entropy and cross recurrence quantification analysis. Computational and mathematical methods in medicine 2017, 1–11.

Milagro, J., Gracia, J., Seppa, V.-P., Karjalainen, J., Paassilta, M., Orini, M., Bailon, R., Gil, E., Viik, J., 2019. Noninvasive Cardiorespiratory Signals Analysis for Asthma Evolution Monitoring in Preschool Children. IEEE Transactions on Biomedical Engineering 67 (7), 1863–1871.

Moffa, A., Rinaldi, V., Costantino, A., Cassano, M., Gelardi, M., Fiore, V., Lopez, M. A., Baptista, P., Campisi, P., Casale, M., 2020. Childhood Obstructive Sleep Apnea: from Diagnosis to Therapy—an Update. Current Sleep Medicine Reports 6 (3), 157–162.

Nakano, H., Tanigawa, T., Furukawa, T., Nishima, S., 2007. Automatic detection of sleep-disordered breathing from a single-channel airflow record. European Respiratory Journal 29 (4), 728–736.

Narsky, I., Porter, F. C., 2013. Statistical analysis techniques in particle physics: Fits, density estimation and supervised learning, 1st Edition. John Wiley & Sons.

Nguyen, H. D., Wilkins, B. A., Cheng, Q., Benjamin, B. A., 2014. An Online Sleep Apnea Detection Method Based on Recurrence Quantification Analysis. IEEE Journal of Biomedical and Health Informatics 18 (4), 1285–1293.

Ning, T., Bronzino, J. D., 1990. Autoregressive and Bispectral Analysis Techniques: EEG Applications. IEEE Engineering in Medicine and Biology Magazine 9 (1), 47–50.

Oeverland, B., Skatvedt, O., Kvarner, K. J., Akre, H., 2002. Pulseoximetry: Sufficient to diagnose severe sleep apnea. Sleep Medicine 3 (2), 133–138.

Oppenheim, A. V., Schafer, R. W., Buck, J. R., 1999. Discrete–Time Signal Processing, 2nd Edition. Prentice Hall.

Pu, Y., Patterson, R. P., CramerBornemann, M. A., 2005. Nocturnal cardio-respiratory indices - A novel screening tool for pediatric obstructive sleep disordered breathing. In: Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. Shanghai, China, pp. 2575–2578.

Ramírez Ávila, G., Gapelyuk, A., Marwan, N., Stepan, H., Kurths, J., Walther, T., Wessel, N., 2013. Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods. Autonomic Neuroscience 178 (1-2), 103–110.

Refaeilzadeh, P., Tang, L., Liu, H., 2016. Cross-validation. Encyclopedia of Database Systems, 1–7.

Rioul, O., Vetterli, M., 1991. Wavelets and Signal Processing. IEEE Signal Processing Magazine 8 (4), 14–38.

Rosso, O. A., Blanco, S., Yordanova, J., Kolev, V., Figliola, A., Schürmann, M., Ba ar, E., 2001. Wavelet entropy: A new tool for analysis of short duration brain electrical signals. Journal of Neuroscience Methods 105 (1), 65–75.

Rosso, O. A., Martin, M. T., Figliola, A., Keller, K., Plastino, A., 2006. EEG analysis using wavelet-based information tools. Journal of Neuroscience Methods 153 (2), 163–182.

Ryan, P. J., Hilton, M. F., Boldy, D. A., Evans, A., Bradbury, S., Sapiano, S., Prowse, K., Cayton, R. M., 1995. Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnoea/hypopnoea syndrome: can polysomnography be avoided? Thorax 50 (9), 972–975.

Saeys, Y., Inza, I., Larranaga, P., 2007. A review of feature selection techniques in bioinformatics. Bioinformatics 23 (19), 2507–2517.

Saramäki, T., Mitra, S., Kaiser, J., 1993. Finite impulse response filter design. In: Handbook for digital signal processing. Vol. 4. John Wiley & Sons, pp. 155–277.

Schinkel, S., Dimigen, O., Marwan, N., 2008. Selection of recurrence threshold for signal detection. The European Physical Journal Special Topics 164 (1), 45–53.

Shouldice, R. B., O'Brien, L. M., O'Brien, C., de Chazal, P., Gozal, D., Heneghan, C., 2004. Detection of Obstructive Sleep Apnea in Pediatric Subjects using Surface Lead Electrocardiogram Features. Sleep 27 (4), 784–792.

Sörnmo, L., Laguna, P., 2005. Bioelectrical signal processing in cardiac and neurological applications, 1st Edition. Academic Press.

Spruyt, K., 2012. Pediatric Sleep-Disordered Breathing: Criteria and Spectrum of Disease. In: Sleep Disordered Breathing in Children. Humana Press, pp. 245–260.

Taha, B. H., Dempsey, J. A., Weber, S. M., Badr, M. S., Skatrud, J., Young, T. B., Jacques, A. J., Seow, K. C., 1997. Automated Detection and Classification of Sleep-Disordered Breathing From Conventional Polysomnography Data. Sleep 20 (11), 991–1001.

Takens, F., 1981. Detecting strange attractors in turbulence. In: Dynamical Systems and Turbulence, Warwick 1980. Lecture Notes in Mathematics, vol 898. Springer, pp. 366–381.

Tan, H.-L., Gozal, D., Ramirez, H. M., Bandla, H. P. R., Kheirandish-Gozal, L., 2014. Overnight Polysomnography versus Respiratory Polygraphy in the Diagnosis of Pediatric Obstructive Sleep Apnea. Sleep 37 (2), 255–260.

Tsai, C.-M., Kang, C.-H., Su, M.-C., Lin, H.-C., Huang, E.-Y., Chen, C.-C., Hung, J.-C., Niu, C.-K., Liao, D.-L., Yu, H.-R., 2013. Usefulness of desaturation index for the assessment of obstructive sleep apnea syndrome in children. International Journal of Pediatric Otorhinolaryngology 77 (8), 1286–1290.

Van Eyck, A., Verhulst, S. L., 2018. Improving the diagnosis of obstructive sleep apnea in children with nocturnal oximetry-based evaluations. Expert Review of Respiratory Medicine 12 (3), 165–167.

Vaquerizo-Villar, F., Álvarez, D., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Barroso-García, V., Crespo, A., del Campo, F., Gozal, D., Hornero, R., 2018a. Detrended fluctuation analysis of the oximetry signal to assist in paediatric sleep apnoea–hypopnoea syndrome diagnosis. Physiological Measurement 39 (11), 114006.

Vaquerizo-Villar, F., Álvarez, D., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Barroso-García, V., Crespo, A., del Campo, F., Gozal, D., Hornero, R., 2018b. Utility of bispectrum in the screening of pediatric sleep apnea-hypopnea syndrome using oximetry recordings. Computer Methods and Programs in Biomedicine 156, 141–149.

Vaquerizo-Villar, F., Álvarez, D., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Barroso-García, V., Crespo, A., del Campo, F., Gozal, D., Hornero, R., 2018c. Wavelet analysis of oximetry recordings to assist in the automated detection of moderate-to-severe pediatric sleep apnea-hypopnea syndrome. PLOS ONE 13 (12), e0208502.

Várady, P., Micsik, T., Benedek, S., Benyó, Z., 2002. A novel method for the detection of apnea and hypopnea events in respiration signals. IEEE Transactions on Biomedical Engineering 49 (9), 936–942.

Wang, R., Wang, J., Li, S., Yu, H., Deng, B., Wei, X., 2014. Multiple feature extraction and classification of electroencephalograph signal for Alzheimers' with spectrum and bispectrum. Chaos 25 (1), 013110.

Wang, Y., Wang, W., Liu, Y., Wang, D., Liu, B., Shi, Y., Gao, P., 2009. Feature extracting of weak signal in real-time sleeping EEG with approximate entropy and bispectrum analysis. In: 3rd International Conference on Bioinformatics and Biomedical Engineering, ICBBE 2009. Beijing, China, pp. 1–4.

Webber, C. L., Zbilut, J. P., 1994. Dynamical assessment of physiological systems and states using recurrence plot strategies. Journal of Applied Physiology 76 (2), 965–973.

Weir, J. P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the sem. The Journal of Strength & Conditioning Research 19 (1), 231–240.

Welch, P., 1967. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Transactions on Audio and Electroacoustics 15 (2), 70–73.

Witten, I. H., Frank, E., Hall, M. A., 2011. Data mining: practical machine learning tools and techniques, 3rd Edition. Morgan Kaufmann/Elsevier.

Xu, Z., Gutiérrez-Tobal, G. C., Wu, Y., Kheirandish-Gozal, L., Ni, X., Hornero, R., Gozal, D., 2018. Cloud Algorithm-Driven Oximetry-Based Diagnosis of Obstructive Sleep Apnea in Symptomatic Habitually-Snoring Children. The European respiratory journal 53 (2), 1801788.

Yu, L., Liu, 2004. Efficient Feature Selection via Analysis of Relevance and Redundancy. Journal of Machine Learning Research 5, 1205–24.

Zar, J. H., 1972. Significance testing of the spearman rank correlation coefficient. Journal of the American Statistical Association 67 (339), 578–580.

Zbilut, J. P., Webber, C. L., 2006. Recurrence Quantification Analysis. In: Wiley Encyclopedia of Biomedical Engineering. John Wiley & Sons, Inc.

Zhang, J. W., Zheng, C. X., Xie, A., 2000. Bispectrum analysis of focal ischemic cerebral EEG signal using third-order recursion method. IEEE Transactions on Biomedical Engineering 47 (3), 352–359.

Zhou, S.-M., Gan, J. Q., Sepulveda, F., 2008. Classifying mental tasks based on features of higher-order statistics from EEG signals in brain-computer interface. Information Sciences 178 (6), 1629–1640.

Zweig, M. H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical chemistry 39 (4), 561–577.

Index

Pediatric obstructive sleep apnea (OSA) is a highly prevalent and underdiagnosed breathing disease that can adversely affect the physiological and cognitive functions of children, causing serious long-term neurocognitive deficiencies, behavioral disorders, cardiometabolic malfunctioning, and endocrine dysfunctions. Consequently, it is of the utmost importance that they are timely diagnosed and treated to prevent these negative consequences.

The gold standard test for OSA diagnosis is the hospital overnight polysomnography (PSG). It is highly effective, but also uncomfortable to children, expensive, complex, intensively laborious, and often unavailable, which lead to long waiting lists and diagnostic delays. Thus, great efforts have been made to search and develop simpler diagnostic methods as alternative to PSG. One of these alternatives is the automatic analysis of readily acquirable cardiorespiratory signals.

In this context, the Doctoral Thesis focuses on designing, implementing, and assessing novel automatic signal processing methods that allow comprehensively characterizing the overnight airflow (AF) from children and helping in the pediatric OSA diagnosis. Particularly, non-linear, spectral, bispectral, recurrence plot, and wavelet analyses are proposed for adapting to the intrinsic properties of overnight AF and getting useful OSA-related information from it. In view of the results, we can conclude that the characterization of overnight AF by means of these novel methods can help to simplify the OSA diagnosis in children. In addition, the high performance of the proposed models suggests that they could be incorporated into clinical practice as reliable automatic screening methods for pediatric OSA.

UVa

Doctoral Thesis
Compendium of publications
International Mention

ISBN 978-84-09-39400-5