



**UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA**

**CLASIFICACIÓN DE SUEÑO MEDIANTE MEDICIÓN DE LA ACTIVIDAD  
MOTORA**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO**

**GABRIEL ANDRÉS ORELLANA LÓPEZ**

**PROFESOR GUÍA:  
CLAUDIO HELD BARRANDEGUY**

**MIEMBROS DE LA COMISION:  
PATRICIO PEIRANO CAMPOS  
PABLO ESTEVEZ VALENCIA**

Este trabajo ha sido parcialmente financiado por CONICYT - Chile, bajo la licencia  
FONDECYT N° 1120319

**SANTIAGO DE CHILE  
2014**

**RESUMEN DE LA MEMORIA PARA OPTAR AL  
TÍTULO DE: Ingeniero Civil Eléctrico  
POR: Gabriel Andrés Orellana López  
FECHA: Noviembre de 2014  
PROFESOR GUÍA: Claudio Held Barrandeguy**

## **CLASIFICACIÓN DE SUEÑO MEDIANTE LA MEDICIÓN DE ACTIVIDAD MOTORA**

El objetivo del presente trabajo de memoria fue desarrollar un sistema capaz de aprender de registros actigráficos nocturnos de adolescentes sanos y, en base a la información obtenida y a registros polisomnográficos simultáneos, generar un algoritmo que permita clasificar automáticamente los diferentes estados y etapas del sueño en base al registro actigráfico.

La actigrafía es una técnica de extracción de información fisiológica que consiste en la ubicación de un dispositivo, denominado actígrafo, en la muñeca de la mano no dominante del paciente. El actígrafo graba la actividad de acelerómetros internos, registrando la actividad motora de la extremidad. Esta técnica ha despertado un creciente interés en la comunidad científica debido a su simplicidad, bajo costo y su carácter no invasivo, por lo que se han desarrollado numerosos estudios para el análisis de dicha señal.

El archivo de registros del Laboratorio de Sueño y Neurobiología Funcional del INTA incluye numerosos registros de actigrafía, y polisomnogramas de los mismos pacientes. Se revisaron los registros, descartando aquellos que presentaban ruido o periodos sin mediciones. Luego se buscó sincronizar los registros actigráficos aptos con su respectivo polisomnograma. Se construyó una base de datos compuesta de 114 registros actigráficos muestreados a un dato por minuto, cada uno con su respectivo hipnograma, obtenido a partir del polisomnograma y validado por expertos en medicina del sueño para ser utilizados como ground truth. Con esta base de datos se construyó un sistema compuesto por dos clasificadores basados en Redes Neuronales Artificiales, el primero es un clasificador que permite distinguir entre sueño y vigilia, mientras que el segundo permite clasificar entre los estados y etapas del sueño.

El clasificador de sueño y vigilia utiliza 11 características extraídas de los registros actigráficos, las que fueron elegidas a través de un proceso de selección de características. Entre éstas se cuentan indicadores estadísticos, que miden tanto tendencia central como dispersión en una ventana de tiempo móvil, así como características de contexto temporal, es decir, en qué tiempo se localiza la ventana en relación con el sueño completo, y el contexto de actividad, es decir, información sobre la actividad previa y posterior a la ventana. Este clasificador tiene una precisión de 93,2%, con una detección de sueño de 96,2% y una detección de vigilia de 79,6%. Estos resultados constituyen una mejora respecto de los publicados a la fecha por otros grupos de investigación.

El segundo clasificador utiliza las mismas características que el primero excepto dos, que fueron desechadas mediante un test estadístico por ser mutuamente redundantes con otras que entregan más información, lo que no es el caso para la clasificación sueño-vigilia. En esta clasificación se busca distinguir entre los minutos pertenecientes a cuatro etapas: Sueño Quieto 1 (SQ 1), SQ 2, SQ 3&4 y Sueño Paradójico (SP). El clasificador de estados y etapas del sueño tiene una precisión global del 61,0%, detectando el 12,2% de los minutos pertenecientes a la etapa SQ 1, 80,7% de los minutos pertenecientes a la etapa SQ 2, 68,8% de los minutos pertenecientes a SQ3&4 y 6,7% de los etiquetados como etapa SP. El bajo rendimiento del sistema de clasificación se debe a que los datos actigráficos no parecen ser adecuados para discriminar estados y etapas de sueño. No existen estudios con los cuales comparar este resultado.

Combinando los ambos clasificadores se alcanzó una precisión de 62,4%.

## **Agradecimientos**

En primer lugar, agradezco a mis padres. Además agradezco a mis profesores y compañeros que conocí y de los que aprendí a lo largo de todos estos años.

Este proyecto fue financiado por el Departamento de Ingeniería Eléctrica, por el Instituto de Nutrición y Tecnología de los Alimentos (INTA) y por CONICYT, bajo la licencia FONDECYT N°1120319.

## Tabla de contenido

Agradecimientos.....	ii
Tabla de contenido.....	iii
Índice de tablas.....	vi
Índice de ilustraciones.....	x
1. Introducción.....	1
1.1. Características del sueño.....	1
1.1.1. Generalidades.....	1
1.1.2. Etapas del sueño.....	4
1.2. Estudios del sueño.....	7
1.2.1. Polisomnograma (PSG).....	7
1.2.2. Actigrafía (ACT).....	8
1.2.3. Variables fisiológicas aisladas.....	9
1.3. Estado del arte en clasificadores automáticos.....	11
1.3.1. Clasificación automática sueño/vigilia.....	11
1.3.2. Clasificación automática de estados y etapas del sueño.....	12
1.4. Clasificación automática basada en actigrafía.....	13
1.4.1. Estado del arte.....	13
1.4.2. Fundamentación fisiológica de clasificación de estados y etapas del sueño mediante información actigráfica.....	14
1.5. Clasificación.....	16
1.5.1. Espacio de características.....	16
1.5.2. Aprendizaje supervisado.....	17
1.6. Objetivos.....	17
1.6.1. Objetivo general.....	17
1.6.2. Objetivos específicos.....	17
1.7. Estructura del documento.....	18
2. Metodología.....	19
2.1. Sujetos y registros.....	19
2.1.1. Procedimiento de adquisición.....	19

2.1.2.	Base de datos y sincronización de registros .....	20
2.2.	Sistema de clasificación.....	22
2.2.1.	Pre-procesamiento .....	23
2.2.1.1.	Eliminación de ruido .....	23
2.2.1.2.	Saturación.....	24
2.2.1.3.	Aplicación de logaritmo natural.....	25
2.2.1.4.	Pre-procesamiento previo a etapa de clasificación .....	26
2.2.2.	Metodología de clasificación.....	26
2.2.2.1.	Indicadores estadísticos de resultados.....	26
2.2.2.1.1.	Clasificación sueño/vigilia.....	26
2.2.2.1.2.	Clasificación de estados y etapas del sueño.....	28
2.2.2.2.	Algoritmo de clasificación .....	28
2.2.2.3.	Balance de base de datos .....	29
2.2.2.3.1.	Balance en base de datos de clasificador sueño/vigilia .....	30
2.2.2.3.2.	Balance en base de datos de clasificador de estados y etapas del sueño.....	31
2.2.2.4.	Test estadístico .....	32
2.2.3.	Algoritmo de clasificación sueño/vigilia.....	34
2.2.3.1.	Extracción de características para clasificación sueño/vigilia .....	34
2.2.3.1.1.	Ventanas de tiempo.....	36
2.2.3.1.2.	Extracción de características de tendencia central.....	37
2.2.3.1.3.	Extracción de características de dispersión de datos .....	38
2.2.3.1.4.	Extracción de características temporales .....	39
2.2.3.1.5.	Extracción de características de tendencia.....	41
2.2.3.2.	Comparación entre características.....	42
2.2.3.2.1.	Comparación entre características de tendencia central .....	43
2.2.3.2.2.	Comparación entre características de dispersión .....	44
2.2.3.2.3.	Comparación entre características temporales.....	45
2.2.3.2.4.	Comparación entre características de tendencia .....	46
2.2.3.3.	Selección final de características.....	47
2.2.4.	Algoritmo de clasificación de estados y etapas del sueño.....	50

2.2.4.1.	Extracción de características para clasificación de estados y etapas del sueño.....	50
3.	Resultados.....	51
3.1.	Clasificación sueño/vigilia .....	51
3.1.1.	Procedimiento de medición de resultados .....	51
3.1.2.	Resultados.....	51
3.2.	Clasificación de estados y etapas del sueño .....	53
3.2.1.	Procedimiento de medición de resultados .....	53
3.2.2.	Resultados.....	55
3.3.	Unificación de clasificadores .....	56
4.	Discusión .....	57
4.1.	Análisis de resultados .....	57
4.1.1.	Clasificación sueño/vigilia .....	57
4.1.1.1.	Utilidad de características desarrolladas .....	59
4.1.2.	Clasificación estados y etapas del sueño .....	60
4.2.	Comparaciones con otros estudios .....	64
5.	Conclusiones.....	67
5.1.	Clasificación sueño/vigilia .....	67
5.2.	Clasificación estados y etapas del sueño .....	68
5.3.	Unificación de clasificadores .....	69
5.4.	Trabajo futuro .....	69
	Bibliografía.....	71

## Índice de tablas

Tabla 1-1. Bandas fisiológicas. Frecuencias límite de las bandas son referenciales, y no son compartidas por todos los estudios.....	4
Tabla 1-2. Patrones de sueño utilizados para clasificación automática [9] basándose en criterios definidos por R&K [10]. Los estados y etapas están divididos en VI (Vigilia), SQ 1, SQ 2, SQ 3-4 (Sueño Quieto) y SP (Sueño Paradójico). Los patrones mencionados son Movimientos Oculares Rápidos (MOR), caracterizados por variaciones bruscas en el EOG; Ondas Theta, caracterizadas por una frecuencia dominante de ondas en dicha banda (entre 3-7 Hz) en el EEG; Husos Sigma, definidos como episodios de aparición de ondas en la banda sigma (entre 10-15 Hz); Ondas Delta Lentas, caracterizadas por predominancia de ondas de frecuencia comprendida en la banda entre 0.5-3 Hz; y Tono muscular, que se detecta por actividad electromiográfica en electrodos ubicados en el mentón. Se marca con ✓ si hay predominancia del patrón en alguna etapa, con ✗ si no hay predominancia y con – si hay ausencia de dicho patrón. ....	6
Tabla 1-3. Porcentaje aproximados de duración de estados y etapas del sueño dentro un periodo de sueño nocturno normal (es decir, desde el SOS hasta que el sujeto se despierta la mañana siguiente). Información extraída de Carskadon et al. [1]......	7
Tabla 2-1. Número de minutos pertenecientes a clases sueño y vigilia y proporción de estas clases dentro de las bases de datos original y balanceada. Además se muestra el factor por el que se replican las clases de la base de datos original para obtener una base de datos equilibrada. ....	31
Tabla 2-2. Número de minutos pertenecientes a clases NREM 1, NREM 2, NREM 3-4 y REM, además de proporción de estas clases dentro de las bases de datos original y balanceada. También se muestra el factor por el que se replican las clases de base de datos original para obtener la base de datos equilibrada.....	32
Tabla 2-3. Características constituyentes de conjunto de características compartidas. Para comparación de características se generan dos conjuntos de características conformados por las características que se quieren comparar y las características compartidas. ....	41
Tabla 2-4. Subgrupos de características de tendencia central sometidos a comparación. Para realizar las comparaciones, se conformaron cuatro grupos de características constituidos por la unión del grupo de características compartidas (excepto las características de tendencia central) y cada uno de los subgrupos presentados. A la derecha puede apreciarse la tasa de mala clasificación, la que indica la capacidad del conjunto conformado para clasificar (mientras más baja, mayor capacidad).....	42
Tabla 2-5. Resultados de comparaciones entre características de tendencia central. Se compara los resultados de un conjunto de características (eje vertical) contra el rendimiento	

utilizando otro conjunto (eje horizontal). No se realizaron todas las comparaciones posibles debido a razones de costo computacional, marcando como  $\emptyset$  las que no se realizaron. Tampoco se realizaron comparación entre grupos idénticos, por lo que se marcó esos espacios como “-”. Un valor negativo en el valor de  $t$  implica que el clasificador entrenado con el otro conjunto obtuvo mejores resultados. .... 43

Tabla 2-6. Subgrupos de características de dispersión sometidos a comparación. Para realizar las comparaciones, se conformaron cuatro grupos de características constituidos por la unión del grupo de características compartidas (excepto las características de dispersión) y cada uno de los subgrupos presentados. A la derecha puede apreciarse la tasa de mala clasificación, la que indica la capacidad del conjunto conformado para clasificar (mientras más baja, mayor capacidad). .... 43

Tabla 2-7. Resultados de comparaciones entre características de dispersión. Se compara los resultados de un conjunto de características (eje vertical) contra el rendimiento utilizando otro conjunto (eje horizontal). No se realizaron todas las comparaciones posibles debido a razones de costo computacional, marcando como  $\emptyset$  las que no se realizaron. Tampoco se realizaron comparación entre grupos idénticos, por lo que se marcó esos espacios como “-”. Un valor negativo en el valor de  $t$  implica que el clasificador entrenado con el otro conjunto obtuvo mejores resultados. .... 44

Tabla 2-8. Subgrupos de características de tendencia central y dispersión sometidos a comparación. Para realizar las comparaciones, se conformaron cuatro grupos de características constituidos por la unión del grupo de características compartidas (excepto las características de tendencia central y las de dispersión) y cada uno de los subgrupos presentados. A la derecha puede apreciarse la tasa de mala clasificación, la que indica la capacidad de cada conjunto conformado para clasificar (mientras más baja, mayor capacidad). .... 48

Tabla 2-9. Resultados de comparaciones entre características de dispersión. Se compara los resultados de un conjunto de características (eje vertical) contra el rendimiento utilizando otro conjunto (eje horizontal). No se realizaron todas las comparaciones posibles debido a razones de costo computacional, marcando como  $\emptyset$  las que no se realizaron. Tampoco se realizaron comparación entre grupos idénticos, por lo que se marcó esos espacios como “-”. Un valor negativo en el valor de  $t$  implica que el clasificador entrenado con el otro conjunto obtuvo mejores resultados. .... 49

Tabla 3-1. Resultados de clasificación sueño y vigilia para la base de datos balanceada utilizando el punto de operación propio de la base de datos balanceada. Los valores de la diagonal en la matriz son las épocas bien clasificadas, las que al dividirse por el número total de épocas entregan el valor de la precisión. .... 51

Tabla 3-2. Resultados de clasificación sueño y vigilia tanto en base de datos balanceada como en base de datos original (cardinalidad de clase vigilia dividida en 4) utilizando el punto de operación calculado para base de datos original. Los valores en las diagonales de las matrices contienen las épocas bien clasificadas. Dividiendo la suma de éstos por el número total de épocas de cada base de datos se obtiene el valor de la precisión en dicha base de datos.....	52
Tabla 3-3. Cardinalidad y proporción de las clases dentro de la base datos utilizada en este estudio.....	53
Tabla 3-5. Resultados de clasificación de etapas del sueño para base de datos balanceada en el punto de operación correspondiente a la base de datos original. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que divididas por el número total de épocas entregan el valor de la precisión para este punto de operación. ....	54
Tabla 3-4. Resultados de clasificación de estados y etapas del sueño para base de datos balanceada en el punto de operación correspondiente a la base de datos balanceada. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que al ser divididas por el número total de épocas entregan el valor de la precisión alcanzada con este punto de operación.....	54
Tabla 3-6. Resultados de clasificación de etapas del sueño para base de datos original con el punto de operación correspondiente a la base de datos original. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que divididas por el número total de épocas entregan el valor de la precisión para este punto de operación.....	55
Tabla 3-7. Resultados de los dos clasificadores combinados para clasificar todos los estados y etapas del ciclo sueño vigilia en la base de datos original utilizando los puntos de operación desarrollados para cada clasificador para la base de datos original. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que divididas por el número total de épocas entregan el valor de la precisión correspondiente a los puntos de operación escogidos.....	56
Tabla 4-1. Resultados de clasificación de toda la base de datos en el punto de operación para base de datos original sin aplicar reglas de posprocesamiento.....	58
Tabla 4-2. Resultados de clasificación sueño vigilia de toda la base de datos en el punto de operación para base de datos original luego de aplicar regla N° 1 de posprocesamiento, que consiste en variar la clasificación de todos los minutos que están rodeados por minutos marcados con la clasificación opuesta.....	58
Tabla 4-3. Resultados de clasificación sueño vigilia de toda la base de datos en el punto de operación para base de datos original luego de aplicar las reglas N° 1 y N° 2 de posprocesamiento. La reglas N° 2 consiste en que todos los pares de minutos consecutivos	

marcados como la misma clase y que están rodeados de pares de minutos consecutivos marcados como la clase opuesta cambian su clasificación a la de sus vecinos..... 59

Tabla 4-4. Comparación de rendimientos de distintos clasificadores con el objetivo de medir la capacidad discriminadora de las características temporales y de contexto. En la prueba A se utilizaron todas las características descritas en la metodología (11 características). En la prueba B se utilizaron todas las características excepto las características I y II. En la prueba C se utilizaron todas las características excepto las características VIII y IX. .... 60

Tabla 4-5. Resultados de clasificación de estados y etapas del sueño con base de datos balanceada en el punto de operación correspondiente a la base de datos balanceada. Se consideraron NREM 2 y REM como una única clase. La precisión se calculó sumando todas las épocas bien clasificadas, ubicadas en la diagonal de la matriz (en negritas) y dividiéndolo por el total de épocas. .... 63

Tabla 4-6. Resultados de clasificación de estados y etapas del sueño con base de datos balanceada en el punto operación correspondiente a la base de datos original. Se consideraron NREM 2 y REM como una única clase. La precisión se calculó sumando todas las épocas bien clasificadas, ubicadas en la diagonal de la matriz (en negritas) y dividiéndolo por el total de épocas. .... 63

Tabla 4-8. Resultados de implementación de algoritmo presentado en Sadeh et al.[42] probado en la base de datos utilizada en este estudio. La precisión se calculó sumando las épocas correctamente clasificadas, ubicadas en la diagonal (en negrita) y dividiedola por el número total de épocas. .... 64

Tabla 4-7. Resultados de clasificación de estados y etapas del sueño con base de datos original en el punto operación correspondiente a la base de datos original. Se consideraron NREM 2 y REM como una única clase. La precisión se calculó sumando todas las épocas bien clasificadas, ubicadas en la diagonal de la matriz (en negritas) y dividiéndolo por el total de épocas. .... 64

Tabla 4-9. Resultados de estudios donde se desarrollaron clasificaciones de sueño/vigilia que utilizaron sujetos adultos o adolescentes sanos basados en información actigráfica. ... 65

## Índice de ilustraciones

- Figura 1-1. Tres fases del espectro de Sleep Onset (SOS), es decir, del proceso de transición entre vigilia y sueño. Sleep Onset se define como el momento en que los expertos señalan que el sujeto se queda dormido. Las fases no tienen duraciones fijas y pueden estar superpuestas. La primera fase hace referencia a que la actigrafía mide la actividad motora, por lo que si se usa actigrafía, se fijará el SOS cuando el sujeto se quede quieto. La segunda fase hace referencia al hecho de que la polisomnografía señala como inicio del sueño la medición señales fisiológicas que coinciden con el estado en que los músculos se relajan. La tercera fase hace referencia un SOS perceptivo, cuando el sujeto no logra responder estímulos externos. Figura basada en trabajo de Tryon [7]..... 2
- Figura 1-2. Movimientos de la articulación del codo que pueden ser detectados por el actígrafo. A: Flexión y extensión. B: Supinación y Pronación. En conjuntos, ambos pares de movimientos permiten describir variaciones de posición de la muñeca en los ejes polar y acimutal, utilizándose coordenadas esféricas y al brazo como una componente radial dentro de éstas. Imagen extraída de Johnson D., Pedowitz R., “Practical Orthopaedic Sports Medicine & Arthroscopy”, 2006, Lippincott Raven, 1ra Edición..... 8
- Figura 1-3. Modos de adquisición de señal actigráfica. El gráfico muestra la señal de un acelerómetro en el tiempo. En puntos blancos se pueden apreciar los cruces por un umbral de la señal del acelerómetro, este umbral es normalmente definido igual a cero. En negro, el tiempo que transcurre la señal sobre un umbral definido por el usuario en función de la tecnología del actígrafo (este umbral puede variar dependiendo del actígrafo). En gris, la integración digital (suma) de la señal. Imagen extraída de Jean-Louis et al. [21]. ..... 9
- Figura 1-4. Evolución de las publicaciones científicas que involucran actigrafía (ACT) en comparación con las que involucran polisomnografía (PSG) a lo largo de las últimas dos décadas. Notar la diferencia de escalas en el eje y para ACT y PSG. Imagen extraída de Sadeh [22]..... 10
- Figura 1-5. Ejemplo de Hipnograma. En el eje Y se muestran el estado vigilia (wake), las cuatro etapas del estado sueño quieto (NREM) y el estado de sueño paradójico (PS). En el eje X se muestra la hora. Extraído de Zoubek et al. [37]. ..... 12
- Figura 1-6. Cambios en el potencial eléctrico de membrana y en la actividad sináptica de las motoneuronas durante sueño y vigilia. En A se muestran las variaciones de la actividad eléctrica entre los estados SQ (NREM), SP (REM) y vigilia (wakefulness) para 5 señales distintas. Las dos últimas son electromiogramas (EMG) tomados en el cuello y en el bíceps femoral respectivamente y revelan claras diferencias en los patrones de activaciones eléctricas. En B se aprecia el potencial eléctrico de una motoneurona lumbar a través de los distintos estados del sueño. Un menor valor indica menor actividad. En C se muestran los patrones del potencial eléctrico de una motoneurona lumbar para cada fase de sueño con

una gran ganancia. Se puede notar en la fase SP importantes peaks inhibitorios. Imagen extraída de Chase [43].	15
Figura 2-1. Sincronización entre actigrafía (en azul) y PSG a través de medición de movimiento en brazo no dominante (en púrpura) y de medición de movimiento en pierna no dominante (en celeste). Pueden apreciarse bastantes coincidencias entre la actividad motora y la actividad grabada en la PSG, lo que es signo de una buena sincronización.	21
Figura 2-2. Diagrama de bloques de sistema de clasificación de estados y etapas del CSV. El primer bloque hace referencia al pre-procesamiento aplicado a los registros actigráficos almacenados en la base de datos. El segundo bloque es un clasificador que separa las épocas marcadas como sueño de las de vigilia por medio de un entrenamiento basado en aprendizaje supervisado utilizando hipnograma. El tercer bloque del sistema es un clasificador que discrimina entre los distintos estados y etapas del sueño por medio de aprendizaje supervisado utilizando hipnogramas para entrenamiento. Luego de todas estas etapas, el sistema entrega una salida indicando la clasificación para cada época.	22
Figura 2-3. Visualización de registro actigráfico. En rojo se puede apreciar ruido consistente en valores de baja intensidad y baja variación. Los registros que presentaron este tipo de ruido fueron eliminados de la base de datos utilizada en este estudio.	23
Figura 2-4. Dos ejemplos de registros actigráficos nocturnos. En rojo, umbral de saturación de 300 cruces por cero por minuto. Puede apreciarse que no es común que se supere este umbral en medio del periodo de sueño, y que sí es común que se supere en los eventos de vigilia ubicados al inicio y al final de periodo de sueño.	24
Figura 2-5. Registro actigráfico. Arriba se aprecia el registro sin pre-procesamiento. Abajo, el mismo registro después de aplicado logaritmo natural.	25
Figura 2-6. Esquema de red neuronal artificial utilizada en la clasificación sueño/vigilia. A la izquierda se aprecian las entradas (Input). En este caso se cuenta con 11 neuronas de entradas. A su derecha se aprecian las dos capas ocultas (Hidden 1 y 2) con una cantidad de neuronas iguales al número de entradas más uno. Puede apreciarse también la función de transferencia utilizada en estas capas (función de tangente sigmoidea). A su derecha se aprecia la capa de salida (Output) con una única neurona, que es la que entrega la salida de la red. Figura extraída de MATLAB 2011a.	28
Figura 2-7. Ejemplo de curva ROC para dos clasificadores (test A y test B). El test A tiene mejor rendimiento que el test B, debido a que tiene mejores valores de sensibilidad y especificidad. La línea ND indica una clasificación binaria completamente al azar.	30
Figura 2-8. Esquema de procedimiento para prueba estadística $t$ para comparación de dos clasificadores (A y B) descrita en Dietterich [49]. La base de datos balanceada se divide al azar en dos. Ambos clasificadores se entrenan utilizando la primera mitad y se prueban	

utilizando la segunda mitad (proceso “(1)”). De la misma forma, ambos clasificadores se entrenan con la segunda mitad de la base de datos, y se prueban con la primera (proceso “(2)”). Los resultados de ambos clasificadores para ambos procesos son comparados. El procedimiento se repite 5 veces..... 33

Figura 2-9. Histogramas superpuestos de clases sueño (en azul) y vigilia (en rojo) para valor ACT del minuto analizado después del pre-procesamiento. Puede apreciarse que la clase vigilia no supera en cantidad de épocas a la clase sueño en ningún valor de la actividad motora (eje X), por lo que no existe un punto donde ubicar un umbral eficiente.... 34

Figura 2-10. Registro actigráfico después de pre-procesamiento. Para extracción de características de un minuto en particular (asterisco rojo) se utiliza una ventana de 11 minutos (rojo) sobre la que se calculan diversos parámetros. .... 35

Figura 2-11. Señal obtenida (rojo) luego de aplicar filtro promediador de ventana móvil de 11 minutos sobre actividad actigráfica luego de pre-procesamiento (azul). ..... 35

Figura 2-12. Distribuciones de la actividad actigráfica (luego de pre-procesamiento) después de ser filtrada por un filtro promediador de ventana de ancho de 11 minutos. Arriba se puede ver la distribución de toda la base de datos. Abajo se puede ver la distribución de los datos catalogados por el experto como vigilia..... 36

Figura 2-13. Superposición de histogramas de distribución de la desviación estándar de los datos de ACT después de pre-procesamiento, considerando una ventana de 11 minutos. En azul, datos pertenecientes a clase sueño, en rojo, datos pertenecientes a clase vigilia..... 39

Figura 2-14. Gráfico de características temporales calculadas. En celeste están los minutos marcados como de “baja actividad”, debido a que no hay actividad importante en su vecindad. En púrpura se aprecia la característica proporcional, que indica la ubicación de cada minuto dentro del episodio de baja actividad, es decir, para agrupaciones de minutos marcados como de baja actividad consecutivos (para efectos de visualización, se limitan los valores de esta característica entre 0 y 100). En rojo se aprecia la característica de tiempo absoluta, consiste en la cantidad de minutos consecutivos de baja actividad. .... 40

Figura 2-15. Segmento de registro actigráfico. En azul se aprecia el promedio de la actividad actigráfica después de pre-procesamiento medido en ventanas móviles de 11 minutos. En rojo se aprecia la derivada del promedio de la actividad actigráfica medido en ventanas móviles de 11 minutos. En negro se muestra el nivel del cero del promedio. .... 45

Figura 2-16. Segmento de registro actigráfico. En azul se aprecia la actividad actigráfica después de pre-procesamiento, donde se remarca el minuto analizado (asterisco). Se pueden apreciar los 20 minutos previos al minuto analizado (morado) y los 20 minutos posteriores (rojo). ..... 45

Figura 2-17. Superposición de histogramas de distribución de promedio de 11 minutos de actividad actigráfica después de pre-procesamiento de las clases sueño y vigilia. En rojo, datos clasificados por el experto como vigilia. En azul, datos clasificados como sueño. Abajo, acercamiento del gráfico de arriba en el punto de interés  $x=2,2$ . ..... 46

Figura 2-18. Funciones de distribución acumulativas empíricas del promedio de ventanas de 11 minutos de la actividad actigráfica después de pre-procesamiento de las clases sueño y vigilia. En rojo se muestra los datos clasificados por el experto como vigilia. En azul, los clasificados como sueño. Se puede apreciar que el 97% de los datos pertenecientes a la clase sueño están bajo el punto 2,2, mientras que el 66% de los datos pertenecientes a la clase vigilia están sobre este punto. .... 47

Figura 3-1. Curva ROC de clasificador sueño/vigilia. En el eje X, 1-Especificidad, en el eje Y, Sensibilidad. Se marca punto de operación calculado para base de datos equilibrada (cuadrado) y para base de datos original (círculo). ..... 52

Figura 3-2. Salidas del clasificador ANN sueño/vigilia (rojo) en función de la actividad motora (azul) y del estado del sujeto (sueño o vigilia) marcado por expertos médicos (negro). En verde se puede apreciar el umbral utilizado por el punto de operación correspondiente a la proporción original entre las clases, umbral igual a 0,28 (verde). ..... 53

Figura 4-1. Evolución de la precisión, la sensibilidad y la especificidad a medida que aumentan los minutos considerados en el periodo de tiempo de tolerancia al error. Se considera una tolerancia al error de “N” si un minuto m está mal clasificado, pero si en su vecindad de ancho N existe un minuto que tiene la misma clasificación predicha por el sistema para m. Un círculo marca el valor del estadístico calculado, mientras que la barra que lo rodea marca la desviación estándar del estadístico. .... 57

Figura 4-2. Salidas del clasificador ANN de estados y etapas del sueño en un registro en función de la actividad motora (azul) y del estado o etapa del sueño en que se encuentra el sujeto (negro), determinada por un experto médico. El punto de operación es el correspondiente a la base de datos balanceada. Se aprecian 4 salidas distintas, correspondientes a las cuatro clases a clasificar por el clasificador, de manera que el que la clase que tiene un valor mayor en un minuto determinado se adjudica la etiqueta de ese minuto. Además se puede apreciar que los valores de las clases NREM 2 y REM tienden a mantenerse altos en los mismos segmentos del registro. .... 61

Figura 4-3. Curvas ROC de clasificadores comparados. En azul, clasificador basado en ANN desarrollado en este estudio. En rojo, algoritmo desarrollado en Sadeh et al. [42]. Los círculos corresponden a los puntos de operación donde la precisión se maximiza para la base de datos original. .... 66

# 1. Introducción

## 1.1. Características del sueño

### 1.1.1. Generalidades

El sueño, según Carskadon et al. [1], es un estado de comportamiento reversible caracterizado por un aislamiento perceptual y una ausencia de respuesta frente al ambiente. El sueño se da naturalmente todos los días y es acompañado típicamente (pero no necesariamente) de una postura horizontal de reposo, marcada disminución de movimientos corporales, ojos cerrados y otros indicadores que se asocian comúnmente con el dormir, siendo claramente distinguible del estado de actividad habitualmente observado durante el día, denominado vigilia (VI) [2].

Los ciclos de sueño se rigen, principalmente, por dos procesos distintos y complementarios [3]. En primer lugar se tienen el circadiano, que comprende los eventos que ocurren dentro de un sistema biológico a periodos medianamente regulares, cercanos a la duración de un día [2]. Estos procesos sincronizan funciones tales como el latido cardiaco, el sistema endocrino, la temperatura corporal, entre otros. El ciclo sueño-vigilia (CSV) es el más evidente de los procesos circadianos. Este proceso opera con propiedades cronométricas similares a las de un reloj, funcionando como un regulador del sueño. En los individuos con patrón sueño-vigilia coincidente con el ciclo noche-día (que es lo más habitual), la necesidad de dormir más intensa aparece al inicio de la franja nocturna.

Luego, el proceso homeostático, que puede ser definido como el conjunto de fenómenos de autorregulación orientados a la preservación de una relativa estabilidad en las composiciones y las propiedades del medio interno de un organismo. Los mecanismos homeostáticos que controlan el equilibrio interno del organismo también controlan los tiempos del CSV y la necesidad de dormir. Este mecanismo mantiene el equilibrio interno, de manera que a más horas estando despierto, mayor será la necesidad de dormir, y a más horas durmiendo, menos fuerte será dicha necesidad. Dentro de los procesos homeostáticos, es ampliamente aceptada la tesis de que el sueño tiene una importante función en diversos procesos mentales. Algunos autores (citados en [4]) sostienen que durante el sueño se realiza una consolidación de los pesos sinápticos dentro de las redes neuronales generados durante el día, contribuyendo, de esta manera, con los procesos de memoria, aprendizaje y plasticidad cerebral.

Los procesos circadiano y homeostático se desarrollan en forma independiente, pero su interacción determina la coordinación, duración y calidad del CSV.

Existen numerosos aspectos del sueño que no son del todo conocidos y que son sujeto de estudio, como las funciones y la necesidad de sueño, y las diversas enfermedades asociadas con el sueño, como la apnea, el insomnio o la narcolepsia [2].

En la investigación del sueño se usan ampliamente procedimientos objetivos y subjetivos [4]. Sin embargo, como las manifestaciones fisiológicas de los fenómenos abarcados por el sueño son de muy diversa naturaleza, se requiere un procedimiento que permita poder medirlas como conjunto, combinando la información obtenida aisladamente en cada uno de las diversas actividades fisiológicas. Es por esto que como herramienta de diagnóstico de referencia (gold standard) para todas las enfermedades conocidas vinculadas al sueño, así como para su estudio científico, se utiliza la Polisomnografía (PSG) convencional en laboratorio. Ésta consiste en registrar simultáneamente grabaciones de distintas actividades fisiológicas, como Electroencefalografía (EEG), Electromiografía (EMG), Electrooculografía (EOG), etc.; además de tomar nota sobre ciertos eventos, como cuando se apagan o se prenden las luces. Todos estos datos son posteriormente analizados por expertos. El número mínimo de canales, así como su ubicación, depende del propósito de la evaluación. Por ejemplo, la recomendación general consiste en más de un canal de EEG, un canal de EMG, dos canales de EOG, Electrocardiografía (EKG), esfuerzo respiratorio torácico y abdominal, de flujo de aire por vías respiratorias, de saturación de oxígeno y de posición corporal [5], sin embargo, para detección de micro-despertares no son necesarias mediciones de EKG, de esfuerzo torácico y abdominal, de flujo de aire por las vías

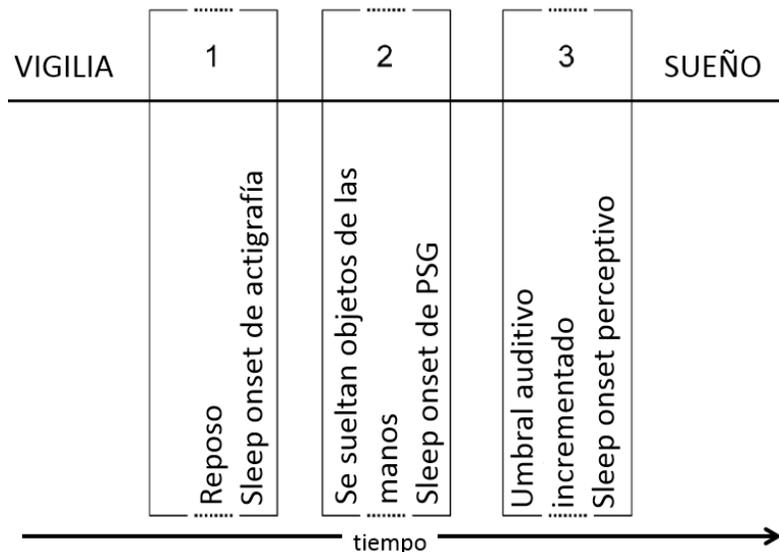


Figura 1-1. Tres fases del espectro de Sleep Onset (SOS), es decir, del proceso de transición entre vigilia y sueño. Sleep Onset se define como el momento en que los expertos señalan que el sujeto se queda dormido. Las fases no tienen duraciones fijas y pueden estar superpuestas. La primera fase hace referencia a que la actigrafía mide la actividad motora, por lo que si se usa actigrafía, se fijará el SOS cuando el sujeto se quede quieto. La segunda fase hace referencia al hecho de que la polisomnografía señala como inicio del sueño la medición de señales fisiológicas que coinciden con el estado en que los músculos se relajan. La tercera fase hace referencia a un SOS perceptivo, cuando el sujeto no logra responder a estímulos externos. Figura basada en el trabajo de Tryon [7].

respiratorias, de saturación de oxígeno, ni de posición corporal [6].

Pese a la precisión de las diversas mediciones antes mencionadas, existen diferentes autores que hacen notar la dificultad de establecer el evento puntual que determine la transición entre vigilia y sueño por medio de diversos métodos de análisis [7]. El momento en que la transición entre vigilia y sueño se establece se denomina Sleep Onset (SOS). Es bien sabido que las clasificaciones expertas subjetivas basadas en mediciones de variables fisiológicas pueden sobreestimar el tiempo total de sueño, detectando de forma prematura el SOS. De hecho algunos pacientes insisten en que no han podido conciliar el sueño después de una noche completa verificada normalmente por la PSG.

En la figura 1-1 se muestra un resumen visual de las tres fases de SOS [7] que componen un espectro de SOS, es decir, el proceso de transición entre vigilia y sueño. Los espacios entre los bloques muestran que no todos los eventos en el proceso de SOS están representados. Los eventos identificados con cada bloque no son de igual duración, ni tienen una duración fija, lo que depende de diversos factores: 1) estado de ánimo del sujeto; 2) procesos vinculados con el desarrollo –se ve un aumento de la duración de las fases con la edad–; 3) varias enfermedades que pueden modificar aguda o crónicamente la duración de las fases (típicamente aumentando la duración del SOS).

El primer indicio constatable del inicio de la transición de vigilia a sueño es la inactividad, perteneciente a la primera fase del proceso. Ningún animal duerme sin antes estar en reposo. Puesto que esta fase se caracteriza por la ausencia de movimiento, las metodologías de detección de SOS basadas en medición de movimientos corporales tienden a indicar naturalmente a esta fase como la que representa el SOS. Personas sanas comúnmente pasan esta fase rápidamente, permitiendo una separación mínima entre el SOS de dispositivos medidores de movimiento corporal y el de la PSG. Personas con problemas de sueño pueden demorar en superar esta fase, redundando en importantes diferencias entre las marcas de los dos métodos.

La segunda fase del SOS corresponde a la disminución del tono muscular, a tal punto que las manos no siguen sosteniendo objetos y los dejan caer. Este criterio fue el que se utilizó como gold standard para identificar la etapa I del sueño quieto (la primera etapa de sueño, cronológicamente hablando), ya que se buscó identificar, mediante EEG, el estado cerebral que coincidía con un tono muscular disminuido. Sin embargo, según Ogilvie et al. [8], individuos que en la grabación de un PSG se les indicó pulsar un interruptor en su mano al momento de escuchar un tono, mostraron gran capacidad para realizar la tarea. Antes del SOS de la PSG el 99,3% de los tonos fueron respondidos correctamente. Luego del SOS de la PSG, en etapa I del sueño quieto, este valor se redujo a 72,2% de los tonos. Incluso en la etapa III y IV del sueño las respuestas correctas fueron 24% y 5,3%, respectivamente. Esto hace evidente la naturaleza flexible de las fases de SOS.

**Tabla 1-1. Bandas fisiológicas. Frecuencias límite de las bandas son referenciales, y no son compartidas por todos los estudios.**

Nombre	Rango de Frecuencia
Delta	0,5 – 3 Hz
Theta	3 – 7 Hz
Alfa	7 – 10 Hz
Sigma	10 – 15 Hz
Beta	15 – 30 Hz

La tercera fase del SOS es el aumento del umbral auditivo, lo que ocurre aproximadamente un minuto después del SOS de PSG en sujetos sanos. Numerosos estudios indican que el SOS reportado por los mismos pacientes coincide estrechamente con el aumento de este umbral. En este estado los sujetos no responden cuando se les dice su nombre a volumen bajo, como tampoco a estímulos externos [8].

### **1.1.2. Etapas del sueño**

Una vez que se supera la tercera fase del espectro de SOS, se puede afirmar que el sujeto está dormido. Se distinguen dos estados de comportamiento asociados al sueño [9], denominados sueño quieto (SQ) y sueño paradójico (SP). Durante VI y SP, el EEG es de baja amplitud y con componentes de alta frecuencia, reflejo de una desincronización de la actividad eléctrica cerebral. Esta desincronización hace que los campos eléctricos generados al interior del cerebro se anulen entre sí por efecto de la interferencia negativa, disminuyendo la amplitud de las diferencias de potencial detectadas en el EEG.

En SQ, por el contrario, aparecen ondas de mayor amplitud cuya frecuencia va disminuyendo a medida que el sueño se hace más profundo.

La clasificación de sueño no está completamente estandarizada y, comúnmente, diferentes expertos tienen enfoques levemente diferentes. En la década de los 60, Rechtschaffen y Kales (R&K) publicaron un manual de clasificación de etapas de sueño [10], estableciéndose durante cuatro décadas como la única guía universalmente aceptada para la clasificación de sueño [11]. En trabajos posteriores a la publicación del manual (citados en [12]) se llamó la atención sobre la necesidad de actualizar estos estándares, mejorar su resolución temporal e incorporar herramientas tecnológicas que permitieran hacer la evaluación de los registros de sueño de una forma más confiable y reproducible.

De acuerdo con R&K, los registros de sueño deben dividirse en intervalos fijos denominados épocas, generalmente de 20 o 30 segundos de duración, a los cuales se les asigna un estado o etapa de sueño según ciertos patrones visuales identificables en los distintos canales de EEG, EOG y EMG. Cada uno de estos estados y etapas debe durar al menos un minuto, es decir 3 o 2 épocas consecutivas, dependiendo si éstas duran 20 o 30 segundos respectivamente. Los patrones se pesquisan en ciertas bandas de frecuencia EEG

fisiológicas (tabla 1-1) dentro de los canales y luego, en función de la presencia o ausencia de éstos, se clasifican las épocas siguiendo una serie de reglas. Las etapas propuestas son: VI, SQ 1, SQ 2, SQ 3, SQ 4, SP y MOV (que se refiere a ruido provocado por movimientos del sujeto).

Dentro de las principales críticas que se le hicieron posteriormente al método de clasificación propuesto por R&K se cuentan (citadas en [9]):

- Debido a que tanto el reconocimiento de los patrones como la aplicación de las reglas están sujetos a interpretaciones subjetivas, la etapa de sueño asociada a una determinada época puede diferir entre distintos expertos.
- Durante un sueño fragmentado y durante períodos con transiciones rápidas entre distintas etapas, las reglas de clasificación pueden resultar ambiguas.
- R&K no fue diseñado originalmente como un estándar universal sino como un método de referencia. Las reglas nunca han sido validadas adecuadamente y se utilizan para clasificar registros que presentan patologías del sueño, algo para lo que no fueron diseñadas.

El año 2007 la Academia Estadounidense de Medicina del Sueño (AASM por sus siglas en inglés) propuso una nueva pauta de clasificación [13]. Entre sus más importantes innovaciones se cuentan la instauración de criterios objetivos (numéricos) en reemplazo de la identificación de patrones visuales, la eliminación de la etapa MOV (algo prácticamente eliminado por los avances técnicos), etiquetar el antiguo estado SP como R y renombrar las antiguas etapas pertenecientes al estado SQ como N1, N2 y N3 (fusionando etapas SQ 3 y SQ 4 en N3). Otro cambio es la estandarización del procedimiento de medición de la PSG en su sección de EEG.

En este informe se hace referencia a la clasificación R&K para designar las etapas del sueño debido a que ese fue el método utilizado por los expertos para marcar los registros utilizados en la base de datos. Tanto los criterios de clasificación R&K y AASM han sido validados mediante su uso en diversos estudios.

La caracterización de estados y etapas se realiza identificando la presencia de ciertos patrones en la época analizada. Tal como puede apreciarse en la tabla 1-2, las etapas SQ 1 y SQ 2 muestran un continuo descenso en la frecuencia de los patrones de las ondas cerebrales. También puede verse esta disminución en la actividad motora de las extremidades, identificándose estas etapas, a grandes rasgos, como una transición entre la vigilia y el sueño. Por su parte, las etapas SQ 3-4, que se asocian al sueño profundo, se caracterizan por ondas de baja frecuencia (Slow wave sleep) en el espectro de la banda delta y estarían vinculadas a la función restauradora somática del dormir [9].

Durante SP se producen movimientos oculares rápidos (MOR) y una inhibición del tono muscular (TM). A este estado se le asocia la actividad onírica más vívida del dormir. Su

Tabla 1-2. Patrones de sueño utilizados para clasificación automática [9] basándose en criterios definidos por R&K [10]. Los estados y etapas están divididos en VI (Vigilia), SQ 1, SQ 2, SQ 3-4 (Sueño Quieto) y SP (Sueño Paradójico). Los patrones mencionados son Movimientos Oculares Rápidos (MOR), caracterizados por variaciones bruscas en el EOG; Ondas Theta, caracterizadas por una frecuencia dominante de ondas en dicha banda (entre 3-7 Hz) en el EEG; Husos Sigma, definidos como episodios de aparición de ondas en la banda sigma (entre 10-15 Hz); Ondas Delta Lentas, caracterizadas por predominancia de ondas de frecuencia comprendida en la banda entre 0.5-3 Hz; y Tono muscular, que se detecta por actividad electromiográfica en electrodos ubicados en el mentón. Se marca con ✓ si hay predominancia del patrón en alguna etapa, con \* si no hay predominancia y con - si hay ausencia de dicho patrón.

Patrones	Dominancia, no-dominancia o ausencia de patrones en etapas de sueño				
	VI	SQ 1	SQ 2	SQ 3-4	SP
Movimientos Oculares Rápidos	✓	*	*	*	✓
Ondas Theta	-	✓	-	-	✓
Husos Sigma	*	*	✓	-	*
Ondas Delta Lentas	*	*	*	✓	*
Tono Muscular	✓	-	-	-	*

función desde el punto de vista fisiológico no está clara ya que personas con supresión de SP no presentarían problemas de salud particulares [14]. Por otro parte, en SP se presenta un aumento del gasto energético cerebral (en comparación con el resto del sueño), así como aumento de la incidencia de patologías cardiacas. Algunas hipótesis acerca de su función son:

- Propiciar la producción de neuronas en ciertas zonas del cerebro que presentan muerte celular habitualmente. Esta neurogénesis desapareció en casos en los que el SP fue inhibido [14].
- Consolidar la redes sinápticas formadas durante el periodo anterior de vigilia. De esta manera tiene un rol clave en la consolidación de la memoria de corto plazo y en el aprendizaje [15].
- Remover ciertos modos de interacción indeseables en redes de células de la corteza cerebral, es decir, eliminar interconexiones que generan comportamientos obsoletos, así como propiciar el aprendizaje de nuevos. Esta hipótesis además plantea que esta reconfiguración neuronal estaría vinculada con la capacidad de plasticidad cerebral [16].

En cuanto a la ocurrencia de los estados y etapas del sueño también se pueden identificar ciertos patrones. Durante el sueño se producen ciclos en los cuales el individuo comienza en VI y va pasando progresivamente a SQ 1, SQ 2, SQ 3, SQ 4 y SP, para recomenzar, normalmente, en SQ 2 y llegar eventualmente a SP, definiéndose, de esta manera dos ciclos [17], el ciclo SP (REMC por sus siglas en ingles) que comprende desde el inicio de un estado REM al inicio del siguiente estado REM; y el ciclo SQ (NREMC, por sus siglas en inglés), que comprende desde el inicio de la etapa SQ 2 hasta el inicio de la siguiente etapa SQ 2. Generalmente estos ciclos no ocurren en forma ordenada y se producen transiciones diferentes a las mencionadas [9]. Es común que las etapas SQ 3 y 4 sean más cortas en el

**Tabla 1-3. Porcentaje aproximados de duración de estados y etapas del sueño dentro un periodo de sueño nocturno normal (es decir, desde el SOS hasta que el sujeto se despierta la mañana siguiente). Información extraída de Carskadon et al. [1].**

<b>Estados o etapas durante el sueño</b>	<b>Porcentaje aproximado</b>
Vigilia	Menos de 5%
Sueño Quieto 1	Entre 2% y 5%
Sueño Quieto 2	Entre 45% y 55%
Sueño Quieto 3	Entre 3% y 8%
Sueño Quieto 4	Entre 10% y 15%
Sueño Paradójico	Entre 20% y 25%

segundo ciclo que en el primero, llegando eventualmente a desaparecer en los ciclos siguientes, lo que lleva a la etapa SQ 2 a expandirse, reemplazando a las etapas SQ 3 y 4 [1]. También se observa que la duración de las etapas de SP es de mayor duración en los ciclos cercanos al final del sueño, mientras que las etapas de SQ duran más al principio del sueño. Otros estados o etapas pueden saltarse o no manifestarse durante algún ciclo, o aparecer muy brevemente. En la tabla 1-3 se muestran los porcentajes de tiempo que ocupa cada una de los estados y etapas [1].

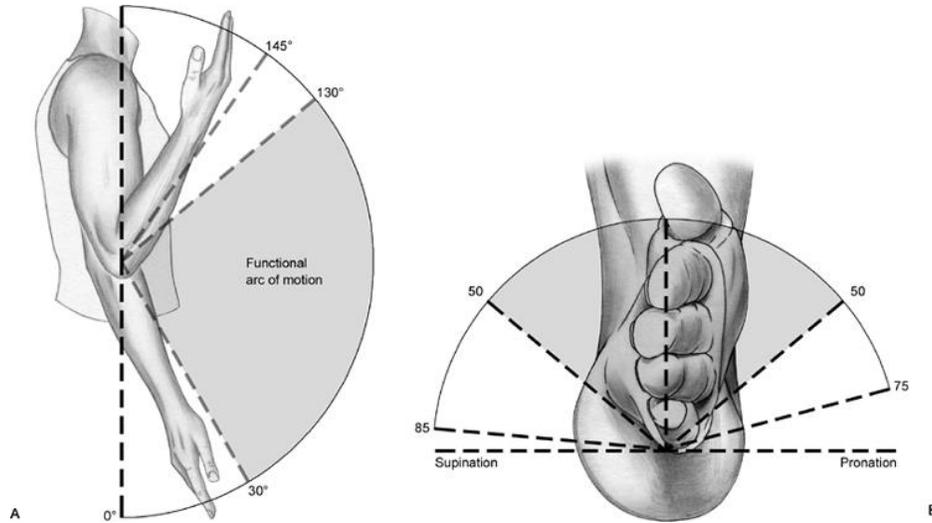
La clasificación de las etapas de sueño es una componente fundamental para cualquier estudio polisomnográfico, tanto para fines científicos como clínicos. En 1990 la Asociación Americana de Desórdenes de Sueño publicó la clasificación internacional de desórdenes de sueño (ICSD), la cual define más de ochenta trastornos de sueño [18]. Además de apoyar el diagnóstico de estas alteraciones, el estudio de los registros polisomnográficos permite hacer una evaluación neurofuncional del sistema nervioso central.

## **1.2. Estudios del sueño**

### **1.2.1. Polisomnograma (PSG)**

El PSG es un examen exploratorio consistente en la obtención y grabación de ciertas señales fisiológicas, así como el cálculo y análisis de ciertos parámetros dependientes de ellas. Fue definido formalmente en 2007 por la AASM [19], previo a esto se utilizaba, pero con distintas configuraciones. Las señales cuya medición se recomienda son: EEG (dos derivaciones), EOG (dos derivaciones), EKG, EMG (en el mentón y en las piernas), medición de flujo aéreo, medición de esfuerzo respiratorio (normalmente EMG en diafragma), saturación de oxígeno en la sangre y posición del cuerpo.

Otros parámetros técnicos definidos por la pauta de la AASM son la impedancia máxima de los electrodos (5 K $\Omega$ ), la mínima resolución digital (12 bits por muestra) así como las frecuencias de muestreo y filtros de frecuencia para la adquisición de las señales fisiológicas [13]. En el caso de este estudio, las frecuencias de muestreo utilizadas están dentro de los rangos aceptables. Con su estandarización en el año 2007, el PSG es



**Figura 1-2. Movimientos de la articulación del codo que pueden ser detectados por el actígrafo. A: Flexión y extensión. B: Supinación y Pronación. En conjuntos, ambos pares de movimientos permiten describir variaciones de posición de la muñeca en los ejes polar y acimutal, utilizándose coordenadas esféricas y al brazo como una componente radial dentro de éstas. Imagen extraída de Johnson D., Pedowitz R., “Practical Orthopaedic Sports Medicine & Arthroscopy”, 2006, Lippincott Raven, 1ra Edición.**

considerado actualmente como el gold standard en cualquier estudio de medicina del sueño, ya sea de carácter clínico o investigativo [7].

### 1.2.2. Actigrafía (ACT)

La actigrafía es un procedimiento que consiste en la ubicación de un actígrafo en la muñeca de la mano no dominante del sujeto, que registra los movimientos realizados por el sujeto. El actígrafo es un dispositivo de apariencia similar a un reloj de pulsera, equipado con acelerómetros que registran los movimientos a los que se ve sometido el dispositivo.

Los actígrafos pueden contar con dos o tres acelerómetros, dos siempre dispuestos en los ejes polar y acimutal, y opcionalmente un tercero en el eje radial considerando al brazo como una componente en el eje radial (ver figura 1-2) visto en coordenadas esféricas. Esto se debe a que los movimientos que puede realizar la articulación del codo corresponden únicamente a supinación y pronación (movimientos posibles en el eje polar, como se aprecia en la figura 1-2, izquierda) y flexión y extensión (movimientos posibles en el eje acimutal, como se aprecia en la figura 1-2, derecha), por lo que un acelerómetro dispuesto en el eje radial no es necesario, ya que no se registra movimiento en ese eje, aunque sí es útil por su carácter complementario, debido a aceleraciones residuales que ocurren en los otros ejes.

Por motivos de capacidad de almacenamiento, los actígrafos almacenan la señal de voltaje de los acelerómetros, sino que la procesan, reduciendo su frecuencia de muestreo a un periodo determinado, denominado época (normalmente de un minuto en los actígrafos anteriores al año 2000), para conseguir grabaciones de mayor duración [20]. Existen tres modos de operación del actígrafo, los que determinarán de qué manera se procesan las

### Mini-Motionlogger Movement Quantification Modalities

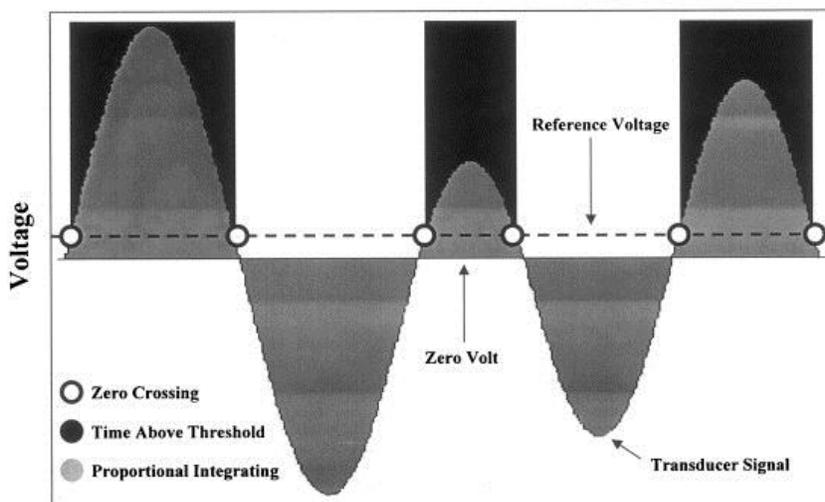


Figura 1-3. Modos de adquisición de señal actigráfica. El gráfico muestra la señal de un acelerómetro en el tiempo. En puntos blancos se pueden apreciar los cruces por un umbral de la señal del acelerómetro, este umbral es normalmente definido igual a cero. En negro, el tiempo que transcurre la señal sobre un umbral definido por el usuario en función de la tecnología del actígrafo (este umbral puede variar dependiendo del actígrafo). En gris, la integración digital (suma) de la señal. Imagen extraída de Jean-Louis et al. [21].

señales de los acelerómetros [21]. Estos modos son: conteo de cruces por cero por época, tiempo en que la señal de voltaje de los acelerómetros sobrepasa un umbral fijado por el operador por época e integración digital (suma de los valores de la señal) por época (ver figura 1-3).

Por la simplicidad y el reducido costo de esta metodología, la actigrafía se ha popularizado en la comunidad científica como una herramienta útil en el análisis de la actividad motora en el sueño (figura 1-4) [22]. En este sentido, mucho se ha debatido sobre la capacidad de este procedimiento para obtener una medición valedera del Sleep Onset (SOS). Comúnmente se ha evaluado a la actigrafía versus la PSG, asumiendo que esta última no entrega errores, por lo que todas las discrepancias entre ambas han sido atribuidas a errores aleatorios propios de la naturaleza de la actigrafía. Sin embargo, estos juicios no toman en consideración que el proceso de transición desde vigilia a sueño no es de carácter discreto, sino más bien continuo, de manera que lo que la actigrafía y la PSG miden son distintas etapas del proceso de SOS, como se explicó anteriormente (figura 1-1).

#### 1.2.3. Variables fisiológicas aisladas

En esta sección se buscó mostrar estudios de sueño distintos a la actigrafía, pero que al igual que ésta, se basan en información sobre un número reducido de sensores, en comparación con la PSG. Estos estudios de sueño, tanto clínicos como científicos, buscan recabar y procesar información útil tanto para la clasificación de estados y etapas del sueño como para la marca de patrones de EEG o respiratorios, usados con fines clínicos. El objetivo de estos estudios es lograr obtener información válida, pero utilizando menos

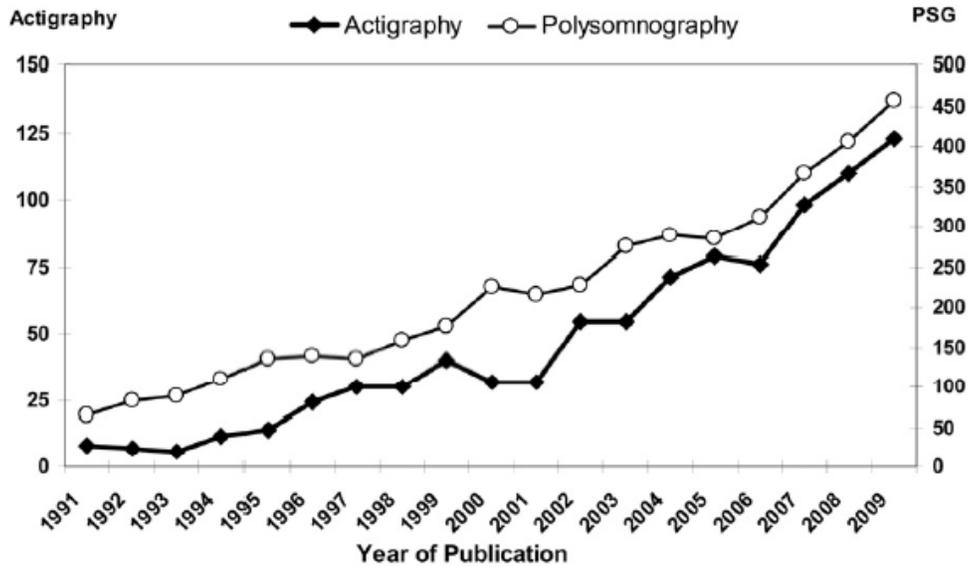


Figura 1-4. Evolución de las publicaciones científicas que involucran actigrafía (ACT) en comparación con las que involucran polisomnografía (PSG) a lo largo de las últimas dos décadas. Notar la diferencia de escalas en el eje y para ACT y PSG. Imagen extraída de Sadeh [22].

canales que la PSG, reduciendo el posible ruido provocado por variables que no aportan información al problema en particular. Además se busca reducir los costos de toma, almacenamiento y procesamiento de datos y permitir un protocolo de toma de muestras menos invasivo para el sujeto que el PSG, realizado ambulatoriamente y sin alterar su rutina habitual. Los modelos desarrollados en estos estudios se comparan, por regla general, contra el análisis del PSG llevado a cabo por expertos (análisis considerado como el ground truth) para evaluar sus resultados.

Por ejemplo, las apneas en el sueño se han analizado de manera importante dentro de la comunidad científica. Existen trabajos que abordan este problema utilizando la PSG [23], sin embargo hay otros trabajos que, basados sólo en la información obtenida de ciertas variables fisiológicas, generan modelos capaces de detectar y analizar este desorden. La variación de la frecuencia cardiaca (HRV, por sus siglas en inglés, calculada a partir del EKG) en el sueño es una de las variables fisiológicas utilizadas. Penzel et al. [24], así como Hayano et al. [25], relacionan directamente la HRV con episodios de apneas del sueño; Khandoker et al. [26] y Shouldice et al. [27] realizan lo mismo, agregando otras variables calculadas a partir del EKG. Por otra parte, Garde et al. [28] estudiaron el comportamiento de la señal de saturación de oxígeno en el sueño de niños con y sin apneas obstructivas, mientras que Karci et al. [29] y Yadollahi et al. [30] utilizaron micrófonos cercanos a las vías respiratorias para detectar apneas.

## **1.3.Estado del arte en clasificadores automáticos**

### **1.3.1. Clasificación automática sueño/vigilia**

Como se explicó en la sección 1.2.1, el gold standard en estudios de sueño es la interpretación de la PSG realizada por personal médico capacitado. La identificación de los diferentes estados y etapas dentro del CSV es muy importante para el diagnóstico de alteraciones del sueño normal. Dentro de los desórdenes del sueño descritos por la Asociación Americana de Medicina del Sueño [18] muchos están asociados directamente a ciertas etapas o estados del CSV. Por ejemplo el insomnio es un trastorno del sueño que consiste en la incapacidad para conciliar y/o mantener el sueño de manera normal, por lo que está asociado a la transición entre los estados de sueño y vigilia. Por otra parte el trastorno de comportamiento del sueño paradójico [31] (REM sleep behaviour disorder, RBD) es una patología en la que durante SP no se produce la atonía muscular (falta de tono muscular) que debe ocurrir en una persona sana. Además este síntoma se ve acompañado de sueños más vívidos y aterradores que lo normal. Esto provoca que el individuo realice movimientos o produzca ruidos mientras sueña, provocando lesiones a sí mismo y a otros. Este trastorno está directamente vinculado al estado SP.

Una segunda razón para la clasificación de los estados y etapas asociados al sueño es el estudio de la calidad restauradora de éste. Se considera [32] que en SQ 3 y 4 se produce la restauración física del organismo, por lo que la presencia, e incluso el aumento en la duración de esta etapa en el sueño debiera ocurrir en periodos de alta actividad física. Lo mismo podría ocurrir con SP [33], estado que se asocia con procesos cognitivos, como el aprendizaje y la memoria, así como la restauración de la capacidad para concentrarse, lo que hace esperable el incremento de esta etapa en periodos de alta actividad intelectual.

Se han desarrollado numerosos clasificadores automáticos para discernir entre los estados de sueño y vigilia basados en diversos métodos y algoritmos. La primera aproximación para esta clasificación es, naturalmente, sistemas basados en información entregada por la PSG. Sin embargo estos sistemas, en su totalidad, están enfocados en discernir entre todos los estados y etapas del sueño, no sólo entre sueño y vigilia, por lo que no son comparables con otros clasificadores de sueño y vigilia descritos en otros estudios.

Los clasificadores basados en PSG están enfocados en clasificar todas las etapas del sueño porque evidentemente poseen mucha más información útil que cualquier otro procedimiento (ya que de esa forma fue concebida, registro simultáneo de todas las señales fisiológicas de interés). Además los estándares utilizados científica y clínicamente, R&K [12] y AASM [13], emplean la PSG para la clasificación de todas las etapas del sueño.

En cuanto a otras metodologías, se puede citar a Adnane et al. [34], donde se utiliza una única derivación de EKG para clasificar entre sueño y vigilia. Para ello se calcula la frecuencia cardíaca y se utiliza la variabilidad de ésta como entrada para un sistema basado

en SVM y árboles de decisión. Sus resultados son sensibilidad de 84,34% y precisión de 79,31% utilizando una base de datos pública con sujetos con y sin apnea del sueño.

### 1.3.2. Clasificación automática de estados y etapas del sueño

Existen diversos estudios que han desarrollado sistemas clasificadores basados en PSG o en parte de ella, como derivaciones puntuales de EEG, EOG, EMG o ECG; así como otros que utilizan otras fuentes de información. En este último grupo se puede mencionar el estudio de Tagliazucchi et al. [35], donde se realiza una segmentación automática del sueño utilizando un procedimiento denominado Imagen de Resonancia Magnética Funcional (fMRI, por sus siglas en inglés), que consiste en obtener múltiples imágenes de la actividad cerebral a modo de grabaciones de video (no sólo imágenes puntuales), obteniendo como resultado una precisión superior al 80% en la configuración de procesamiento de datos más efectiva. El estudio se basa en estándar AASM.

Por su parte, entre los estudios que utilizan canales de la PSG para clasificar se tienen trabajos como el de Becq et al. [36], que presenta una comparación entre 5 clasificadores automáticos basados en EEG y EMG: clasificadores lineal, cuadrático, de k-vecinos cercanos (kNN), de kernel de Parzen y de redes neuronales. Éstos toman como entrada registros de 8 canales y generan automáticamente hipnogramas, que corresponde a una representación gráfica del estado en que se encuentra el individuo en cada momento del periodo de tiempo considerado (figura 1-5). Los resultados de estos clasificadores varían entre el 29% de clasificaciones erróneas, obtenido por el clasificador de redes neuronales, y 42% o 53%, por parte de clasificador de kernel de Parzen y cuadrático respectivamente. Todas las clasificaciones se basan en el estándar R&K.

Otra aproximación se presenta en Zoubek et al. [37] donde se exhiben resultados con una

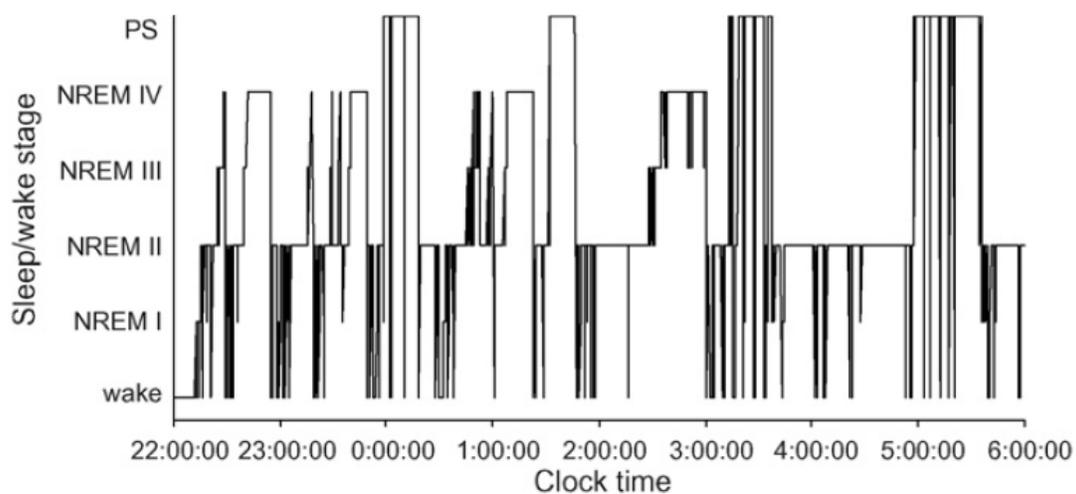


Figura 1-5. Ejemplo de Hipnograma. En el eje Y se muestran el estado vigilia (wake), las cuatro etapas del estado sueño quieto (NREM) y el estado de sueño paradójico (PS). En el eje X se muestra la hora. Extraído de Zoubek et al. [37].

precisión de un 80,11% (épocas correctamente clasificadas dividido total de épocas) utilizando ciertos canales de EEG, EMG y EOG y un clasificador basado en redes neuronales, previo paso de selección de características. Se basa en el estándar R&K.

Sen et al. [38], por su parte, realizan un estudio comparativo entre clasificadores que utilizan todos los canales de EEG considerados en el PSG, mostrando resultados con una precisión cercana al 98%. Las clasificaciones se basan en estándar AASM.

Otro método utilizado es la actigrafía. Sadeh et al. [39] presentan resultados de clasificación entre SQ y SP basado en ACT en niños menores de un año, pero comparándolo con el juicio emitido por observadores entrenados (no por PSG, el gold standard). Se alcanzó un 95,3% de coincidencia entre los observadores y el método planteado para discernir entre sueño y vigilia, mientras que la coincidencia para clasificar entre SQ y SP alcanzó 54% y 87% respectivamente. Se basa en estándar R&K.

## **1.4. Clasificación automática basada en actigrafía**

### **1.4.1. Estado del arte**

Como se mencionó anteriormente, durante las últimas dos décadas la ACT se ha convertido en una importante ayuda para la investigación científica, y particularmente para la medicina del sueño. Trabajos iniciales establecieron la confiabilidad y validez de la detección sueño-vigilia por medio de la ACT, particularmente para poblaciones normales de lactantes y adultos. Una revisión de los resultados de estos estudios se realiza en Tryon [40], donde se mencionan 14 estudios validados de actigrafía de muñeca comparados con PSG para distintos parámetros utilizados en medicina del sueño, entregando resultados de entre 72% y 98% para Tiempo Total de Sueño (TST), entre 82% y 96% para Porcentaje de Sueño, entre 56% y 91% para Eficiencia de Sueño y entre 49% y 87% para Wake After Sleep Onset (WASO, índice que refleja épocas clasificadas como vigilia por el PSG después que se marcó el comienzo del sueño por la actigrafía). En el mismo trabajo se presentan los porcentajes de coincidencia entre actigrafía y PSG identificando sueño y vigilia. Estos porcentajes variaron entre 78,8% y 99,7% para sueño y entre 48,5% y 79,8% para vigilia.

Sin embargo publicaciones han puesto en duda la validez de los algoritmos de detección sueño-vigilia por medio de actigrafía para poblaciones o instrumentos específicos [22]. Además se cuestiona los bajos índices de especificidad (es decir, de detección de vigilia), obteniéndose resultados en este ítem entre 24% y 44%. Es por esto que se puede argüir que la actigrafía presenta un muy buen índice de sensibilidad (épocas de sueño correctamente marcadas por el sistema divididas por épocas de sueño marcadas por el experto) para clasificar el sueño, pero tiene problemas para detectar la vigilia con: a) algunos instrumentos; b) algunos algoritmos; c) algunas poblaciones. Esto puede deberse a que, como se explicaba anteriormente, estas clasificaciones consideran los estados del sueño como discretos y, además, a que consideran a la fase 2 del espectro de SOS como el gold

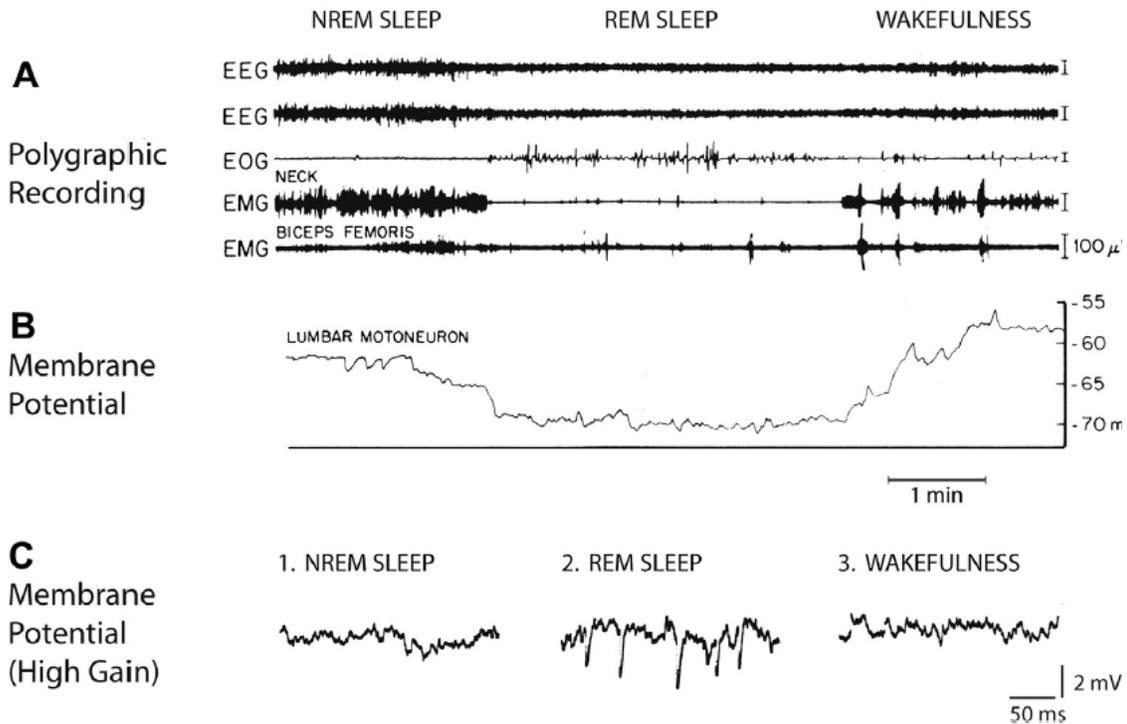
standard (figura 1-1) para marcar el inicio del sueño. Esto puede acarrear errores de clasificación debido a que los algoritmos basados en la actigrafía pueden detectar certeramente sólo la fase 1 del espectro de SOS. Esto lleva a sobrestimar el sueño por sobre la vigilia en pos de aumentar la precisión global del clasificador [41], ya que los minutos marcados por expertos como sueño están en proporción cercana a 4:1 con los minutos marcados como vigilia en los registros utilizados. En respuesta a esto se han desarrollado nuevos sistemas de clasificación en función a estadísticos de orden superior buscando mejorar el balance entre la detección de sueño y de vigilia. Entre estos destaca Domingues et al. [41], que obtiene precisión, sensibilidad y especificidad (épocas de vigilia correctamente marcadas por sistema divididas por épocas de vigilia marcadas por experto) del método allí desarrollado de 77,8%, 75,6% y 81,6% respectivamente.

Según la investigación bibliográfica desarrollada, Sadeh et al. [42] presenta los mejores resultados en un clasificador entre sueño y vigilia. En este trabajo se usan datos de ACT muestreados a una frecuencia de un dato por minuto, comparando clasificaciones minuto a minuto contra el gold standard, el PSG. Su precisión, sensibilidad y especificidad son 92,6%, 97,9% y 74,3% respectivamente.

#### **1.4.2. Fundamentación fisiológica de clasificación de estados y etapas del sueño mediante información actigráfica**

No se encontraron estudios publicados que realicen clasificación de estados y etapas del sueño (SQ 1, SQ 2, SQ 3-4, SP) basados únicamente en la información entregada por la actigrafía comparada con marcas expertas realizadas con PSG.

Zoubek et al. [37] al describir su método de selección de características menciona que una de las características que mejora considerablemente los resultados del sistema es la entropía de EMG. Entropía, en su versión procedente de teoría de la información, es, a grandes rasgos, una medida de la variabilidad de una señal. La selección de la entropía de EMG como una de las características más discriminantes está en concordancia con las reglas R&K [10], que establecen al EMG como uno de los factores más importantes para discernir entre SQ y SP, ya que es un indicador del tono muscular. Durante una etapa de SP, los músculos esqueléticos del paciente se vuelven atónicos, lo que implica una señal EMG plana muy cercana a cero, es decir, una baja entropía de EMG. Por el contrario, durante las primeras etapas de SQ la actividad EMG del paciente es todavía elevada y esto se ve reflejado en una alta entropía. Este fenómeno fisiológico es explicado detalladamente por Chase [43]. Durante la fase de SQ hay un nivel de actividad muscular (y de tono muscular) similar al de vigilia activa, pero de amplitud menor. En contraste con esto, la fase de SP se puede dividir en dos tipos de periodos, en primer lugar están los periodos tónicos, donde la actividad muscular tónica está ausente (i.e. atonía) o muy disminuida (i.e. hipotonía); sin embargo existen también breves periodos fásicos de actividad muscular. Prácticamente todos los músculos esqueléticos exhiben actividad con este patrón de control muscular durante SP. Esto se debe a que, durante el periodo tónico de SP, los procesos pos-sinápticos



**Figura 1-6.** Cambios en el potencial eléctrico de membrana y en la actividad sináptica de las motoneuronas durante sueño y vigilia. En A se muestran las variaciones de la actividad eléctrica entre los estados SQ (NREM), SP (REM) y vigilia (wakefulness) para 5 señales distintas. Las dos últimas son electromiogramas (EMG) tomados en el cuello y en el bíceps femoral respectivamente y revelan claras diferencias en los patrones de activaciones eléctricas. En B se aprecia el potencial eléctrico de una motoneurona lumbar a través de los distintos estados del sueño. Un menor valor indica menor actividad. En C se muestran los patrones del potencial eléctrico de una motoneurona lumbar para cada fase de sueño con una gran ganancia. Se puede notar en la fase SP importantes peaks inhibitorios. Imagen extraída de Chase [43].

inhibitorios que regulan a las motoneuronas (células encargadas de transmitir los impulsos nerviosos para los movimientos de los músculos) generan señales de alta amplitud y frecuencia, provocando la desactivación de estas células. Por su parte en el periodo fásico de SP los procesos pos-sinápticos inhibitorios se ven reforzados, al igual que los excitatorios. Los procesos excitatorios intentan activar la motoneurona mientras que los inhibitorios la desactivan, provocando una competencia entre ambos procesos. Cuando los procesos excitatorios ganan se generan movimientos rápidos oculares y algunas contracciones y sacudidas se producen en los músculos posturales (o tónicos), en la musculatura de la mandíbula, del sistema digestivo, de la lengua, del diafragma y otros. Cuando los procesos inhibitorios ganan hay una profunda supresión de la actividad muscular [43].

Al analizar la figura 1-6 se pueden apreciar claras diferencias estadísticas entre las series de tiempo que representan a la actividad EMG en el bíceps femoral para las tres etapas de sueño. Se podría ver a simple vista una mayor varianza en el periodo SQ (NREM), mientras

que se pueden apreciar peaks de mayor amplitud en el periodo de vigilia, y de algo menor amplitud en el periodo de SP (REM), lo que implicaría un valor promedio mayor.

### **1.5. Clasificación**

El objetivo del trabajo descrito en este informe se puede reducir al de etiquetar cada minuto de las noches de sueño analizadas con una etiqueta perteneciente a una clase. Las etiquetas hacen referencia a las clases que, en el caso de este trabajo, son los distintos estados y etapas del sueño, además del estado de vigilia. Este tipo de problema cabe dentro de la definición de problema de clasificación, es decir, asociar un dato (un minuto del registro actigráfico, en este caso) a una clase, ajustando los parámetros de la clasificación a la información extraída de un conjunto de entrenamiento, compuesto por datos de la misma naturaleza.

#### **1.5.1. Espacio de características**

El dato a clasificar puede estar representado, al momento de hacerse la clasificación, por el dato mismo, o por características que le representen. Por ejemplo, si se quisieran clasificar flores según su color, la forma más obvia sería introducir el color de cada a flor al sistema y entrenarlo para que lo reconozca y pueda separar los individuos de diferente color. Sin embargo al momento de implementar el entrenamiento surge la pregunta ¿cómo representar el color de la flor?, una solución sería ingresar los 3 valores del estándar RGB del color de la flor, lo que significaría, que lo que se ingresa al sistema no es un valor que representa el color de la flor, sino tres valores que representan la intensidad de color rojo, verde y azul, respectivamente, dentro del color de la flor. A estos valores representantes de un dato, y que contienen la información que se utilizará para la clasificación de ese dato, se los denomina “características”.

De la misma forma, y de acuerdo a cual sea el dato a clasificar, se pueden extraer numerosas características, por ejemplo, si se quiere clasificar un pez, se pueden extraer su peso, dimensiones, color, tamaño de la cola, de los dientes, etc. Si se trata de clasificar un proceso, como por ejemplo una señal de radio o el precio de una acción en la bolsa, lo que se busca son parámetros que puedan caracterizar el proceso completo o partes de éste. Características posibles para estos casos son estadísticos aplicados tanto en el dominio del tiempo como de la frecuencia.

El espacio de características es, por lo tanto, un espacio de N dimensiones (con N igual al número de características) donde se ubican las representaciones de los datos a clasificar. De esta manera los algoritmos de clasificación no hacen más que establecer, bajo distintos criterios, ciertas fronteras geométricas (comúnmente hiperplanos) separadoras que delimitan las representaciones de datos de una misma clase a un sector del espacio de características, separándolas de las representaciones de otras clases.

### **1.5.2. Aprendizaje supervisado**

Debido a la naturaleza de la base de datos disponible, es decir, como se cuenta tanto con las entradas del sistema (registros actigráficos), como con las salidas (hipnogramas) que equivalen a las etiquetas, el curso de acción más lógico es implementar un sistema basado en aprendizaje supervisado. Este paradigma para resolución del problema de clasificación consiste en un sistema que “aprende” de un conjunto de entrenamiento constituido por pares de datos conformados por el dato a clasificar (o la representación de éste en el espacio de características definido) y la etiqueta del dato, teniéndose así un aprendizaje “supervisado”, ya que se le entrega al sistema la entrada y luego se corrige su salida (clasificación empírica) para acercarse lo más posible a la salida esperada (clasificación esperada). Por lo tanto el objetivo de todo método de entrenamiento bajo el paradigma de aprendizaje supervisado es reducir el error producido entre la clasificación empírica y la esperada [44].

Finalmente, la etapa de operación del sistema entrenado con aprendizaje supervisado se lleva a cabo ingresando al sistema las entradas que se presenten durante la operación; entradas en forma de características, que buscarán representar e individualizar a cada dato; a lo que el sistema responderá entregando clasificaciones según lo aprendido durante la fase de entrenamiento.

Es importante señalar que los resultados de un clasificador no pueden ser medidos mediante su acción sobre el conjunto de entrenamiento, ya que eso implicaría un sobreajuste de los parámetros a ese conjunto, restándole generalidad al sistema. Para esto se destina una porción de la base de datos que actúa como conjunto de prueba, utilizándose sólo para extraer resultados.

## **1.6. Objetivos**

### **1.6.1. Objetivo general**

Desarrollar un sistema clasificador de estados y etapas del ciclo sueño vigilia (VI, SQ 1, SQ 2, SQ 3-4 y SP) época a época (con épocas de 1 minuto) utilizando únicamente la información actigráfica de los sujetos muestreada a un dato por minuto y usando para entrenamiento (como ground truth) las clasificaciones generadas por expertos usando PSG.

### **1.6.2. Objetivos específicos**

- Seleccionar, dentro de toda la base de datos disponible en el INTA, los registros que contengan información confiable, de manera de poder utilizarlos en el desarrollo de este sistema.
- Procesar los registros para sincronizar los datos actigráficos con las marcas expertas plasmadas en hipnogramas.
- Generar un sistema de clasificación entre las clases sueño y vigilia, de manera de separar la clase vigilia del resto de las clasificaciones posteriores.

- Determinar qué características y qué procedimientos de clasificación logran discriminar mejor entre las épocas correspondientes a sueño y las correspondientes a vigilia.
- Generar un sistema clasificador de estados y etapas del sueño. La cantidad de estados y etapas a clasificar está sujeta a los resultados preliminares alcanzados en el sistema de clasificación sueño/vigilia.
- Determinar qué características y qué procedimientos de clasificación logran discriminar mejor entre las épocas correspondientes a los diferentes estados y etapas del sueño.
- Establecer procedimientos certeros, consistentes y de significancia para medir los resultados obtenidos por los sistemas desarrollados, así como para comparar su rendimiento con el de otros trabajos.

### **1.7. Estructura del documento**

En la presente sección se describe la estructura de este documento de memoria de ingeniero para los capítulos que siguen.

En el capítulo 2 se describe la metodología empleada en la generación de la base de datos, en la construcción del sistema de clasificación sueño/vigilia y en el sistema clasificador entre etapas del sueño.

En el capítulo 3 se muestran los resultados obtenidos, así como la metodología para obtenerlos.

En el capítulo 4 se discuten los resultados y la metodología seguida para el desarrollo del sistema de clasificación, y se comparan con lo expuesto en estudios similares publicados anteriormente.

En el capítulo 5 se exponen las conclusiones extraídas del trabajo realizado.

## **2. Metodología**

### **2.1. Sujetos y registros**

Los registros conformantes de la base de datos utilizada en este trabajo fueron obtenidos de adolescentes sanos de 15 de años de edad de ambos sexos en el Laboratorio de Sueño y Neurobiología Funcional, perteneciente al Instituto de Nutrición y Tecnología de los Alimentos (INTA) de la Universidad de Chile.

El protocolo de investigación fue aprobado por el Comité de Ética del INTA, Universidad de Chile, y por la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT). Los padres de los adolescentes implicados en este estudio fueron informados sobre los aspectos metodológicos involucrados, firmando un consentimiento informado.

Se tuvo acceso a 197 registros actigráficos y 825 hipnogramas, que fueron elaborados por expertos analizando 825 polisomnogramas. De esta totalidad de registros se seleccionaron sólo aquellos que contenían información sobre una sesión de adquisición que tuviera tanto registro actigráfico como hipnograma, descartando la mayoría de los hipnogramas por esta razón. Otros motivos por lo que no se agregaron algunos registros a la base de datos fueron el excesivo ruido en la señal y la ausencia de señal.

Al finalizar esta etapa de comparación entre ambos tipos de registros se llegó a que un total de 140 sesiones estaban representadas de manera aceptable tanto en un hipnograma como en una actigrafía, lo que permitía su análisis.

#### **2.1.1. Procedimiento de adquisición**

Los registros de PSG se adquirieron durante el periodo de sueño nocturno de los adolescentes en una habitación especialmente acondicionada para ello en el Laboratorio de Sueño y Neurobiología Funcional del INTA. El procedimiento se estandarizó para limitar las potenciales influencias del ambiente, los ritmos circadianos y/o de la ingesta de alimentos en los registros. El niño y su madre fueron transportados desde sus casas al laboratorio, arribando una a dos horas antes de su horario normal de inicio de sueño, periodo durante el cual se les alimentaba y el niño realizaba sus rutinas diarias previas a acostarse (lavarse los dientes, ponerse pijama, etc.); una vez en la cama, se procedía a realizar la colocación de los electrodos, y luego el niño se disponía a dormir. La madre también permanecía en el laboratorio durante la noche y dormía en una habitación contigua. La temperatura se mantenía constante (20-22°C) a lo largo de la sesión de registro. Durante la adquisición un experto permanecía en la habitación monitoreando todo el proceso [2]. El protocolo de toma de muestras considera la aplicación de la PSG por dos noches consecutivas en cada sujeto.

Los hipnogramas se obtuvieron por el análisis de expertos sobre los polisomnogramas utilizando el software proporcionado por el fabricante de los polígrafos usados [2]. Los polisomnogramas se obtuvieron mediante un polígrafo modelo Easy EEG-II de 32 canales

(Cadwell, WA, USA, 2000). Se utilizaron 21 canales distribuidos de la siguiente manera: cinco para diferentes regiones del EEG, cuyos electrodos se posicionaron según el sistema internacional 10-20, utilizándose las derivaciones F4-C4, C4-O2, F3-C3, C3-O1 y C4-C3; dos para el EOG; cuatro para el EMG, cuyas medidas se realizaron en el mentón, diafragma y en ambas piernas utilizando electrodos de superficie; uno para el EKG (derivación V1); dos para la detección de actividad motora (extremidades superior e inferior, lado no dominante) utilizando transductores piezoeléctricos; uno para la medición del flujo de aire nasal; uno para el esfuerzo respiratorio abdominal; uno para la medición del pulso; uno para la temperatura corporal; uno para la oximetría; uno para el registro de la posición del sujeto y uno para el registro de ronquidos. Todas las señales se muestrearon a una tasa de 200 Hz y se respaldaron en discos en formato EDF (*European Data Format*) para un análisis y procesamiento posterior.

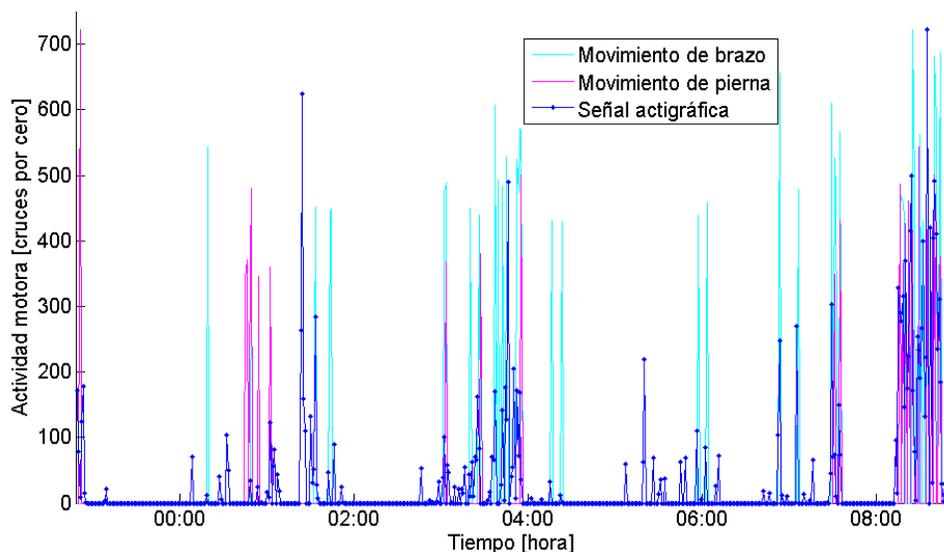
Los registros actigráficos fueron obtenidos por medio de un actígrafo modelo Actiwatch 64 (AW-64, Mini-mitter, Bend, OR, USA) de dos acelerómetros, con una capacidad de almacenamiento de 64 kB. Los acelerómetros tienen una frecuencia de muestreo de 20 [Hz], pero los archivos generados contienen un dato por minuto para aumentar el periodo pesquizado, pudiéndose almacenar hasta 22 días de información. Existen diversos modos de funcionamiento que emplean distintos algoritmos para calcular el valor del dato de cada época en función de la señal de los acelerómetros (figura 1-5). El modo utilizado en la adquisición de los datos de este estudio es “cruces por cero”, elección basada en la literatura [42], [45] . El actígrafo se coloca en la muñeca de la extremidad superior no dominante del sujeto varios días antes de que se someta al PSG, y se le retira al finalizar la segunda noche de registro de PSG. De esta manera se puede disponer de información acerca de la actividad motora de sujeto mientras lleva una rutina normal y se puede la comparar con la de las noches analizadas.

### **2.1.2. Base de datos y sincronización de registros**

Los registros actigráficos están en archivos formato AWD, siendo éste el estándar para los registros actigráficos generados por los actígrafo de modelo Actiwatch. El formato AWD almacena los valores de la actividad motora por épocas junto con información relacionada con la toma de muestra (hora y fecha de inicio, duración de la época, modelo de actígrafo) y con el sujeto de estudio (nombre, sexo). El fabricante proporciona un software que permite la visualización y el análisis de estos registros, pero dicho software no fue empleado en el desarrollo de este trabajo.

Los hipnogramas fueron generados por el software entregado por el fabricante de los polígrafos a una frecuencia de 2 datos por minuto (épocas de 30 segundos, comúnmente utilizadas en análisis de PSG).

Para lograr comparar los registros actigráficos y los hipnogramas se debieron seguir dos pasos. En primer lugar, como la señal de los hipnogramas estaba muestreada al doble de



**Figura 2-1. Sincronización entre actigrafía (en azul) y PSG a través de medición de movimiento en brazo no dominante (en púrpura) y de medición de movimiento en pierna no dominante (en celeste). Pueden apreciarse bastantes coincidencias entre la actividad motora y la actividad grabada en la PSG, lo que es signo de una buena sincronización.**

frecuencia que los registros actigráficos, se construyeron nuevos hipnogramas con épocas de un minuto. La clasificación de cada época de 1 minuto se calculó en base a las dos épocas de 30 segundos que la componen, de manera que si ambas épocas de 30 segundos tienen originalmente la misma clasificación, la época de 1 minuto que las contiene se clasificará de la misma manera. Por el contrario, si las dos épocas de 30 segundos diferían en su clasificación original, la época de 1 minuto se marcó de acuerdo a un orden de prioridades entre los estados y etapas definido de acuerdo a la proporción de éstos en el CSV (dándole más prioridad a aquellos con menor proporción, ver tabla 1-3).

En segundo lugar, se procedió a sincronizar los registros actigráficos con los hipnogramas, ya que no había seguridad de que los relojes utilizados para marcar el inicio de las grabaciones estuviesen correctamente sincronizados. Para llevar a cabo este paso se utilizaron los canales de la PSG correspondientes a la medición de actividad motora (mediante transductores piezoeléctricos) en el brazo y la pierna del lado no dominante. Estas señales están sincronizadas con el hipnograma debido a que son obtenidas junto a las demás señales del PSG, y el hipnograma se marca de acuerdo a la información obtenida del PSG.

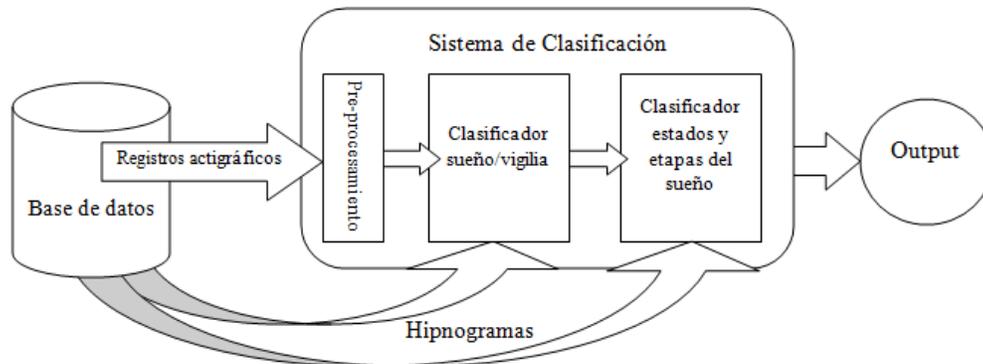
Como criterio de sincronización se buscó el número minutos debía desplazarse el registro actigráfico en el tiempo (en cualquiera de los sentidos temporales) de manera que los picos de actividad del registro actigráfico coincidieran temporalmente con los picos de aceleración de los canales de actividad motora del PSG en piernas y brazos (figura 2-1). Esta tarea se realizó mediante una inspección visual.

## 2.2.Sistema de clasificación

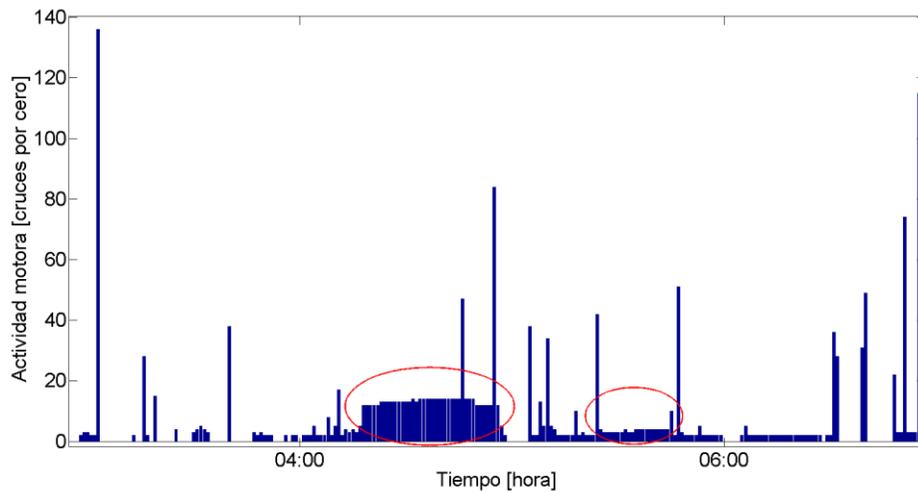
El sistema de clasificación de estados y etapas del CSV que se introduce en este capítulo consta, inicialmente, de un primer bloque de pre-procesamiento de los datos (tanto de registros actigráficos como de hipnogramas) con los objetivos de eliminar el ruido presente en ellos y de modificarlos de tal manera que optimicen el desempeño de las etapas posteriores (figura 2-2). El segundo bloque dentro del sistema es un clasificador “preliminar”, que se entrena para discernir minutos pertenecientes a la clase sueño y a la clase vigilia. Finalmente considera un tercer y último bloque dentro del sistema de clasificación, donde se desarrolla una clasificación entre los distintos estados y etapas del sueño considerados.

El esquema de resolución de problema (figura 2-2), utilizando dos bloques de clasificación, tiene dos razones de ser. Primero, no hay resultados publicados de un clasificador de estados y etapas del CSV basado únicamente en información actigráfica, y sí hay publicaciones referentes a clasificaciones entre sueño y vigilia basada en dicha información, por lo que un clasificador de dos etapas permite evaluar el desempeño del sistema descrito en este informe, en comparación con los descritos en la literatura, al menos en su primera etapa. Segundo, permite determinar si el sistema de clasificación en su fase de entrenamiento (es decir, las características usadas, el algoritmo de clasificación, los parámetros de ese algoritmo, etc.) son los adecuados para cada uno de los bloques y si deben hacerse modificaciones particulares para alguno de ellos.

Un punto en el que debe hacerse hincapié es que para el tercer bloque, se busca generar un clasificador que no ha sido desarrollado antes (no se encontraron metodologías ni resultados publicados), por lo que se permitirán ciertas concesiones, tales como fusionar



**Figura 2-2. Diagrama de bloques de sistema de clasificación de estados y etapas del CSV. El primer bloque hace referencia al pre-procesamiento aplicado a los registros actigráficos almacenados en la base de datos. El segundo bloque es un clasificador que separa las épocas marcadas como sueño de las de vigilia por medio de un entrenamiento basado en aprendizaje supervisado utilizando hipnograma. El tercer bloque del sistema es un clasificador que discrimina entre los distintos estados y etapas del sueño por medio de aprendizaje supervisado utilizando hipnogramas para entrenamiento. Luego de todas estas etapas, el sistema entrega una salida indicando la clasificación para cada época.**



**Figura 2-3. Visualización de registro actigráfico. En rojo se puede apreciar ruido consistente en valores de baja intensidad y baja variación. Los registros que presentaron este tipo de ruido fueron eliminados de la base de datos utilizada en este estudio.**

etapas de sueño o alcanzar resultados que no serían destacables en otro contexto. Esto se debe a la dificultad intrínseca del problema, ya que ciertos criterios de diferenciación entre determinados estados o etapas no tienen relación directa con el movimiento según los estándares de clasificación de estados y etapas del CSV (R&K y AASM), por lo que su discriminación, utilizando sólo esta herramienta, es incierta.

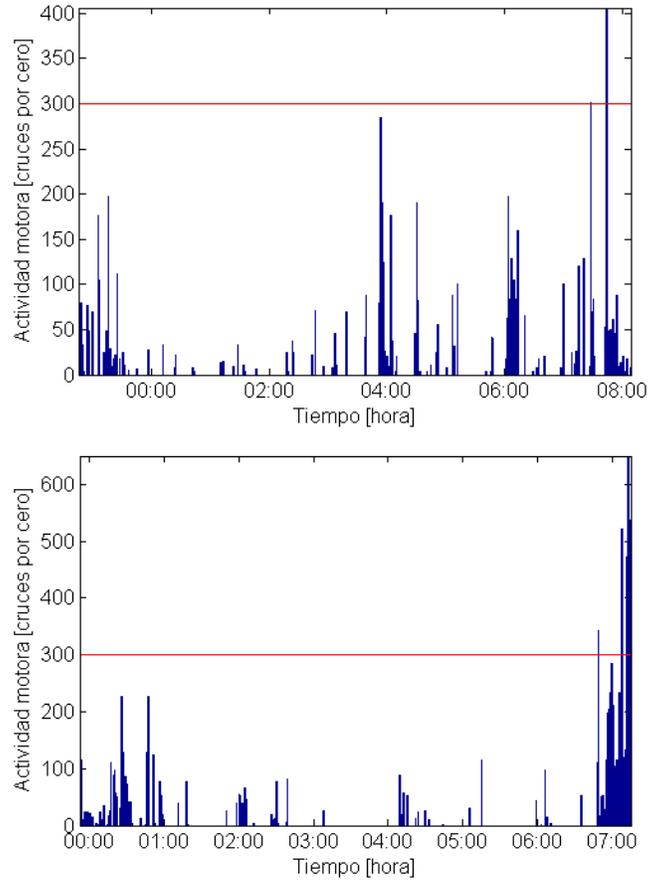
El sistema se implementó en la plataforma MATLAB R2011a, utilizando, para simulaciones y cálculos, un computador con procesador Intel® Core™ i7-3770 con 8 GB de memoria RAM.

### **2.2.1. Pre-procesamiento**

En esta sección (representada por el primer bloque del sistema de clasificación, en la figura 2-2) se detallan los procedimientos aplicados a los datos antes de ser ingresados al sistema de clasificación del segundo bloque. Muchas de las decisiones tomadas en este bloque están relacionadas con la elección de Redes Neuronales Artificiales (ANN) como el algoritmo de clasificación, elección que se fundamentará en las secciones 2.2.3 y 2.2.4.

#### **2.2.1.1. Eliminación de ruido**

Al visualizar los registros actigráficos, los expertos del INTA notaron que muchas grabaciones estaban contaminadas por un ruido de poca amplitud presente durante largos periodos. Este ruido era normalmente constante (se apreciaba una cantidad constante de cruces por cero por minuto), aunque había casos en los que este valor constante variaba levemente a lo largo del tiempo (como en figura 2-3), siempre con valores bajo los 15 cruces por cero por minuto. Se aplicó un algoritmo que elimina los valores constantes y de baja amplitud, pero no se desarrolló un algoritmo para eliminar el ruido que presentaba



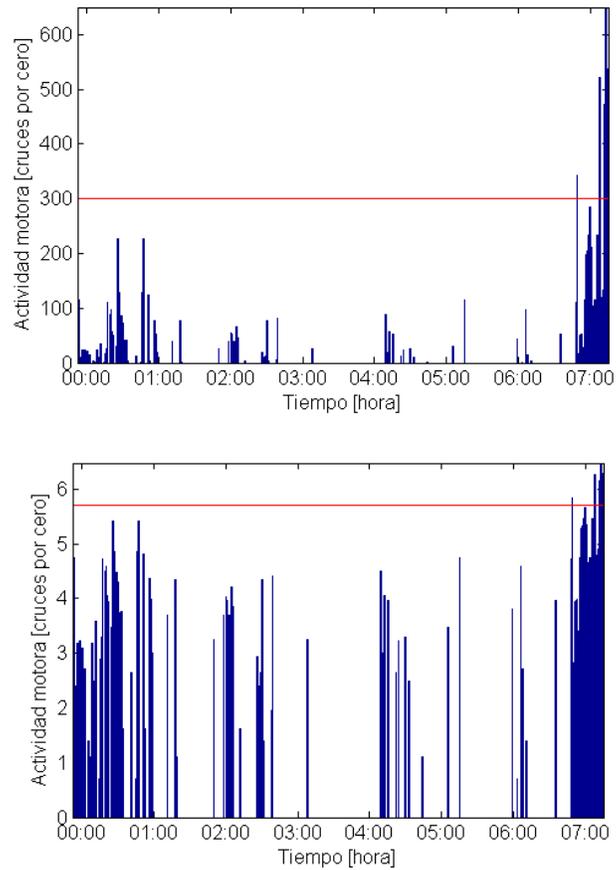
**Figura 2-4. Dos ejemplos de registros actigráficos nocturnos. En rojo, umbral de saturación de 300 cruces por cero por minuto. Puede apreciarse que no es común que se supere este umbral en medio del periodo de sueño, y que sí es común que se supere en los eventos de vigilia ubicados al inicio y al final de periodo de sueño.**

variaciones, por lo que se removieron de la base de datos final los 6 registros que lo presentaban.

### **2.2.1.2. Saturación**

El siguiente paso en el pre-procesamiento de los datos antes de ingresarlos al sistema de clasificación es la saturación de los datos a un valor máximo. Esto significa que se escoge arbitrariamente un valor máximo (de saturación) permitido para los valores de la actividad actigráfica para cada minuto, saturando los que superen ese valor, es decir, se les asigna finalmente el valor de saturación (figura 2-4).

El valor elegido para saturar fue 300 cruces por cero por minuto, ya que se observó que valores cercanos a 500 tendrían a influir negativamente en las clasificaciones de todo el registro que los contenía. Esto puede deberse a que al aumentar el máximo valor de los valores de ACT de un registro, se tiende a relativizar la importancia de los valores intermedios, los que también son indicio importante actividad. Otra razón por la que se



**Figura 2-5. Registro actigráfico. Arriba se aprecia el registro sin pre-procesamiento. Abajo, el mismo registro después de aplicado logaritmo natural.**

escogió el valor de saturación de 300 fue que se apreció que los eventos que superaban este valor estaban vinculados comúnmente a vigilia. Considerando, además, la posterior aplicación de logaritmo natural a los datos (fundamentada en la sección 2.2.1.3), se concluyó que no se perdía información relevante al saturar a 300 los valores de la actividad motora que excedían esta cota, es decir, los valores entre 300 y el máximo valor registrado después del SOS en la base de datos (cerca de 600), ya que éstos son bastante similares al valor de saturación después de aplicárseles logaritmo natural ( $\ln(300 + 1) = 5,707$  y  $\ln(600 + 1) = 6,399$ ).

### **2.2.1.3. Aplicación de logaritmo natural**

Al analizar la tendencia de los datos se hizo notorio que la diferencia de amplitud de los valores de ACT podría inducir a errores en el proceso de entrenamiento de una red neuronal. Esto se debe a que una ANN puede realizar, por medio de variación de sus pesos en el entrenamiento, solamente combinaciones lineales afines de sus entradas.

La aplicación de logaritmo natural al *raw data* de actigrafía (después de sumarle 1, para que no se indeterminen los valores iguales a 0) ya fue utilizada en otros estudios como una de sus características [42], [46]. Ésta tiene como objetivo relativizar la importancia de la amplitud de los eventos dentro del proceso de clasificación, dando relevancia similar a valores disímiles, pero igualmente altos, como 100 y 300, por ejemplo, pero que de todas

maneras reflejan una alta actividad (figura 2-5). Una razón para esto es que la frecuencia de muestreo del registro actigráfico es un dato cada minuto, por lo que la magnitud del movimiento realizado por el sujeto puede haber sido la misma en los dos casos antes mencionados, sólo que en el primer caso esta actividad se extendió por más tiempo, mientras que en el segundo por menos. De la misma manera, dos valores de actividad de baja intensidad (por ejemplo 5 y 15) debiesen tener el mismo efecto para la clasificación.

#### **2.2.1.4. Pre-procesamiento previo a etapa de clasificación**

Por defecto, MATLAB realiza un pre-procesamiento de los datos antes de ser ingresados como entradas al entrenamiento y a la operación de una red neuronal con la configuración “patternnet”, que fue la utilizada en este trabajo (más adelante se explicarán las razones de su elección). En la práctica la única función que se aplicó en esta etapa fue “mapminmax”, la que mapea los valores de entrada a la ANN linealmente, haciendo coincidir el mínimo con 0 y el máximo con 1, de manera de optimizar entrenamiento de la ANN.

### **2.2.2. Metodología de clasificación**

El objetivo de este trabajo es generar un sistema de clasificación de los diferentes estados y etapas del sueño. Para lograr esto fue necesario definir una metodología que entregara resultados de clasificación que fuesen estandarizados y comparables. Esto fue necesario por dos motivos: 1) Era necesario una metodología que entregara alguna forma de indicador sobre la idoneidad de una característica por sobre otra para clasificar etapas o estados de sueño 2) Era necesario un protocolo de comparación entre distintos algoritmo de clasificación.

A continuación se detallan las partes de esta metodología.

#### **2.2.2.1. Indicadores estadísticos de resultados**

Como primer punto, fue necesario estandarizar los indicadores de resultados, es decir, cómo serán medidos los resultados de los indicadores.

##### **2.2.2.1.1. Clasificación sueño/vigilia**

La nomenclatura utilizada para analizar los problemas de clasificación binarios ha sido adaptada desde la nomenclatura de test estadísticos binarios. Un ejemplo de esto pueden ser los test estadísticos aplicados a exámenes médicos diseñados para detección de una enfermedad puntual, los que, bajo el paradigma de clasificación, podrían considerarse como clasificadores entre los estados enfermo y sano. De esta manera el objetivo de estos problemas de clasificación se reduce a diagnosticar la enfermedad cuando el paciente realmente la tiene y a indicar que el paciente no la posee, cuando realmente no la posee. Siguiendo este protocolo, existen 4 posibles escenarios:

1. El paciente tiene la enfermedad y el sistema indica que el paciente tiene la enfermedad. Esto se denomina Verdadero Positivo (VP), ya que se diagnosticó que la enfermedad existía (positivo) y el diagnóstico fue verdadero.

2. El paciente no tiene la enfermedad y el sistema indica que el paciente tiene la enfermedad. Esto se denomina Falso Positivo (FP), ya que se diagnosticó la enfermedad (positivo), pero el diagnóstico fue falso.
3. El paciente no tiene la enfermedad y el sistema indica que el paciente no tiene la enfermedad. Esto se denomina Verdadero Negativo (VN), debido a que el sistema infirió que la enfermedad no estaba presente (negativo) y el diagnóstico fue verdadero.
4. El paciente tiene la enfermedad, pero el sistema indica que el paciente no tiene la enfermedad. Esto se denomina Falso Negativo (FN), ya que no se diagnosticó la enfermedad (negativo) y el diagnóstico fue falso.

Los sistemas de clasificación entre sueño y vigilia disponibles en la literatura han sido presentados como detectores de sueño [41], [42], [47], es por esto que se atribuye el diagnóstico positivo cuando se declara sueño y el diagnóstico negativo cuando se declara vigilia. Esto implica que se considera como VP a las épocas de sueño correctamente clasificadas, como FP a las épocas de vigilia clasificadas como sueño, como VN a las épocas de vigilia correctamente detectadas y como FN a las épocas de sueño clasificadas como vigilia.

Adicionalmente se definen indicadores que informan sobre el desempeño del clasificador. Para evaluar la detección de sueño se define la sensibilidad, mientras que para evaluar detección de vigilia se define la especificidad. Finalmente se define la precisión para evaluar la detección sobre toda la base de datos.

$$sensibilidad = \frac{VP}{VP + FN} \quad 2.1$$

$$especificidad = \frac{VN}{VN + FP} \quad 2.2$$

$$precisión = \frac{VP + VN}{VP + VN + FP + FN} \quad 2.3$$

También se hace referencia a un indicador denominado tasa de mala clasificación, que da cuenta de la proporción de épocas que fueron mal clasificadas. Este indicador está definido como:

$$tasa\ mala\ clasificación = \frac{FP + FN}{VP + VN + FP + FN} = 1 - precisión \quad 2.4$$

El valor predictivo sobre una clase (positivo o negativo) es un indicador que entrega información acerca de cuán certero es el clasificador clasificando como una clase particular. Para el caso de la clase positiva, el índice se define como:

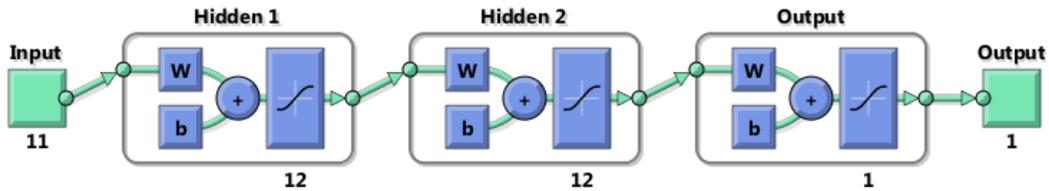


Figura 2-6. Esquema de red neuronal artificial utilizada en la clasificación sueño/vigilia. A la izquierda se aprecian las entradas (Input). En este caso se cuenta con 11 neuronas de entradas. A su derecha se aprecian las dos capas ocultas (Hidden 1 y 2) con una cantidad de neuronas iguales al número de entradas más uno. Puede apreciarse también la función de transferencia utilizada en estas capas (función de tangente sigmoidea). A su derecha se aprecia la capa de salida (Output) con una única neurona, que es la que entrega la salida de la red. Figura extraída de MATLAB 2011a.

$$\text{valor predictivo positivo} = \frac{VP}{VP + FP} \quad 2.5$$

#### 2.2.2.1.2. Clasificación de estados y etapas del sueño

Al no ser una clasificación binaria, la clasificación de estados y etapas del sueño no puede definirse en función de un positivo y un negativo. Tampoco tiene sentido hablar de sensibilidad o especificidad. En su lugar se hablará de “detección de la clase X” para referirnos a la razón entre el número de épocas correctamente detectadas de la clase X dividido por el número de épocas marcadas como de la clase X por el experto.

El término precisión mantiene su utilidad, calculándose como el número de correctas detecciones en todas las clases, dividido por el número total de épocas en la base de datos. A su vez la tasa de mala clasificación se calcula como 1 menos la precisión. También el valor predictivo para cada clase mantiene su fórmula.

#### 2.2.2.2. Algoritmo de clasificación

Se eligió como algoritmo de aprendizaje a las redes neuronales artificiales (Artificial Neural Networks, ANN) debido a dos factores:

- La salida de la ANN es una variable numérica, no categórica (como sí lo es en el caso del *Support Vector Machine*, por ejemplo). Esto permite realizar análisis sobre distintas decisiones relacionadas con el desarrollo del sistema, tales como la selección de las características o la efectividad del pre-procesamiento. Esto es posible debido a que se puede observar el efecto que un cambio en las entradas genera sobre las salidas.
- Relacionado con el anterior, al tener una salida numérica, el proceso de clasificación al utilizar una ANN se realiza imponiendo un umbral (originalmente 0,5), es decir, clasificando como la clase positiva todas los elementos cuyas salidas sobrepasen este umbral, y como negativa a aquellos cuya salida sea menor que este umbral. Esto permite modificar la inclinación del clasificador por una u otra clase por medio de la modificación del umbral, sin modificar lo aprendido en el entrenamiento.

Las ANN tienen múltiples parámetros que deben ser configurados para un óptimo funcionamiento. Entre estos parámetros se cuentan la cantidad de capas de neuronas, la cantidad de neuronas por capa, las funciones de activación de cada neurona, el algoritmo de aprendizaje y la función que mide el desempeño del sistema.

A través de pruebas preliminares se llegó a la conclusión de que la configuración de la red que se comporta mejor en este problema es “patternnet” (figura 2-6) implementada en MATLAB 2011a. Ésta que cuenta con funciones de activación de tangente sigmoidea (*tansig*) y con función de costos de error cuadrático medio (*Mean Square Error*), por su parte el algoritmo de aprendizaje fue modificado, usándose el método Levenberg-Marquardt. En las mismas pruebas se llega a la conclusión de que la mejor configuración, en cuanto a número de capas de neuronas y a número de neuronas por capa, es con dos capas ocultas de igual cantidad de neuronas, cantidad igual al número de entradas de la ANN más 1.

### 2.2.2.3. *Balance de base de datos*

Un punto importante al desarrollar un clasificador es la cardinalidad de las clases que pretenden clasificarse, es decir, la cantidad de minutos que corresponden a cada estado y etapa del CSV. Todos los algoritmos de aprendizaje implementados para ANN, al ser utilizados para clasificación, tienden a sobre-representar la clase que contiene más datos [48]. Esto se debe a que el entrenamiento de la ANN tiene por objetivo maximizar la precisión de clasificación global, no por clase. Esto implica que si una clase está mucho más representada en la base de datos de entrenamiento que otra, la ANN va a tender a *sobre-clasificar* la clase más común, generándose una convergencia más rápida en el entrenamiento en comparación con una base de datos donde las clases estuviesen balanceadas. Esto se refleja en un menor rendimiento en la clasificación.

Para compensar este contratiempo se optó por equilibrar las bases de datos utilizadas para entrenar los dos clasificadores. El protocolo utilizado para balancear las bases de datos consta, en primer lugar, de la detección de la clase con mayor cardinalidad (denominémosla clase  $A$ ). En segundo lugar el resto de las clases (en este caso, las clases  $B_i$ ) se replican un número  $n_i$  de veces, siendo  $n_i$  el número entero que minimiza la diferencia de la cardinalidad de la clase  $A$  con la multiplicación de cardinalidad de la clase  $B_i$  por el número  $n_i$ .

$$\min_{n_i \in \mathbb{N} - \{0\}} (|A_i - B_i \times n_i|) \quad 2.6$$

Siguiendo este procedimiento se puede obtener un clasificador con mejor rendimiento, es decir, con una curva ROC más cercana a la esquina izquierda superior, lo que significa una mayor sensibilidad y especificidad, si se considera un clasificador binario (figura 2-7). Sin embargo, se debe responder la pregunta de cómo se pueden obtener resultados de clasificaciones representativas de la base de datos con las proporciones originales si es que

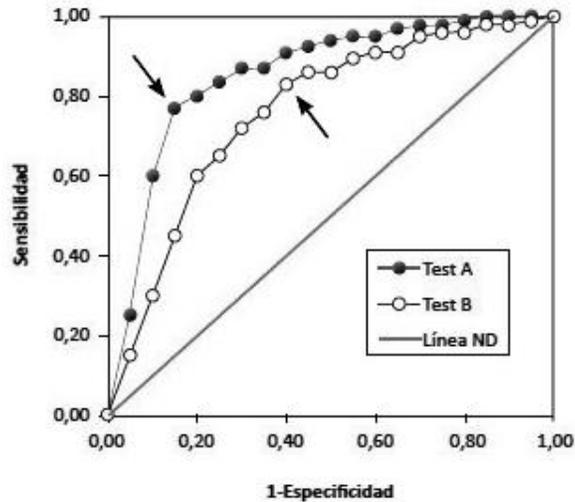


Figura 2-7. Ejemplo de curva ROC para dos clasificadores (test A y test B). El test A tiene mejor rendimiento que el test B, debido a que tiene mejores valores de sensibilidad y especificidad. La línea ND indica una clasificación binaria completamente al azar.

el clasificador fue entrenado con una proporción diferente (balanceada). La solución a esta pregunta surgió al observar nuevamente la figura 2-7: cada uno de los puntos de la curva ROC representa un punto de operación generado, en el caso de clasificadores ANN, por un umbral aplicado a la salida del clasificador. De esta manera, al variarse el umbral de clasificación pueden obtenerse distintos puntos de la curva ROC. Estos puntos presentan distintos balances entre sensibilidad y especificidad, lo que implica que para cada posible par de valores de cardinalidad de las clases existe un punto de la curva ROC con un distinto valor de precisión total.

Para ubicar el punto de operación óptimo se desarrolló una metodología que consiste en buscar a través de toda la curva ROC el umbral que maximiza la precisión global utilizando el conjunto entrenamiento. Cuando se ubicó el umbral óptimo, se ingresaron los datos del conjunto de prueba a la ANN, utilizando como umbral de clasificación el encontrado previamente, obteniéndose los resultados de clasificación.

#### 2.2.2.3.1. Balance en base de datos de clasificador sueño/vigilia

Las cardinalidades de las dos clases, una vez realizado el proceso de balanceo, se pueden ver en la tabla 2-1. Se puede apreciar que el número  $n_i$  de replicaciones del conjunto de vigilia es igual a 4.

El procedimiento para obtener el punto de operación que maximiza la precisión global con la proporción original de la base de datos del clasificador sueño/vigilia es el siguiente: Como el número de datos en vigilia se cuadruplicó, si la red discrimina a uno de estos datos como vigilia o sueño, lo va a hacer también con los otros 3 idénticos, por lo tanto, aunque se aplique el clasificador sobre la base de datos original o sobre la balanceada, el valor de la especificidad va a seguir siendo el mismo.

**Tabla 2-1. Número de minutos pertenecientes a clases sueño y vigilia y proporción de estas clases dentro de las bases de datos original y balanceada. Además se muestra el factor por el que se replican las clases de la base de datos original para obtener una base de datos equilibrada.**

Clase	Base de datos original		Factor	Base de datos balanceada	
	min	%		min	%
Sueño	48990	81,96	<b>1</b>	48990	53,18
Vigilia	10783	18,04	4	43132	46,82
Total	59773	100		92122	100

Para la base de datos con una determinada proporción entre las clases 1 y 2, y conociendo los valores de sensibilidad y especificidad, el valor de la precisión total viene dado por:

$$Acc = \frac{S_1}{(S_1 + S_2)} \times sens + \frac{S_2}{(S_1 + S_2)} \times spec \quad 2.7$$

Con  $Acc$  igual a precisión,  $sens$  igual a sensibilidad y  $spec$  igual especificidad. Siendo además  $S_1$  la cantidad total de épocas pertenecientes a la clase 1 (en este caso, sueño), y  $S_2$  la cantidad de épocas pertenecientes a la clase 2. Para obtener la fórmula de la precisión con la base de datos con la proporción original, basta con dividir la cantidad de datos pertenecientes a una determinada clase por el factor utilizado para balancear la base de datos, en este caso,  $S_2$  debe dividirse por 4 (tabla 2-1).

En el caso de esta base de datos balanceada, se tiene que la fórmula para calcular la precisión con un determinado par de valores para sensibilidad y especificidad es la siguiente:

$$Acc' = \frac{S_S/1}{(S_S/1 + S_W/4)} \times sens + \frac{S_W/4}{(S_S/1 + S_W/4)} \times spec \quad 2.8$$

Con  $Acc'$  como el nuevo valor de precisión. Siendo  $S_S$  la cantidad de datos clasificados como sueño por el experto en la base de datos balanceada y  $S_W$  la cantidad de datos clasificados como vigilia por el experto en la base de datos balanceada, es decir, con 4 veces más datos que la clase vigilia original.

#### 2.2.2.3.2. Balance en base de datos de clasificador de estados y etapas del sueño

Las cardinalidades de las clases utilizadas en esta clasificación pueden verse en la tabla 2-2. Se pueden apreciar los múltiples valores que adoptan los  $n_i$  para cada clase (bajo el rótulo de factores).

Tabla 2-2. Número de minutos pertenecientes a clases NREM 1, NREM 2, NREM 3-4 y REM, además de proporción de estas clases dentro de las bases de datos original y balanceada. También se muestra el factor por el que se replican las clases de base de datos original para obtener la base de datos equilibrada.

Clase	Base de datos original		Factor	Base de datos balanceada	
	min	%		min	%
NREM 1	3083	6,3	8	24664	24,7
NREM 2	25393	51,8	1	25393	25,4
NREM 3-4	11601	23,7	2	23202	23,2
REM	8913	18,2	3	26739	26,7
Total	48990	100		99998	100

El procedimiento para obtener el punto de operación es similar al del primer clasificador, pero aplicado a 4 clases. Como se dijo en la sección 2.2.2.1, en este clasificador se obtuvo un índice de clasificación para cada clase. La precisión corregida para las proporciones originales se obtuvo al multiplicar ese índice por su porcentaje de épocas en la base de datos original. De esta forma, la precisión del método viene dada por:

$$\begin{aligned}
 Acc' = & \frac{\frac{S_{N1}}{8}}{\left(\frac{S_{N1}}{8} + \frac{S_{N1}}{1} + \frac{S_{N3-4}}{2} + \frac{S_R}{3}\right)} \times ind_{N1} + \frac{\frac{S_{N2}}{1}}{\left(\frac{S_{N1}}{8} + \frac{S_{N1}}{1} + \frac{S_{N3-4}}{2} + \frac{S_R}{3}\right)} \times ind_{N2} \\
 & + \frac{\frac{S_{N3-4}}{2}}{\left(\frac{S_{N1}}{8} + \frac{S_{N1}}{1} + \frac{S_{N3-4}}{2} + \frac{S_R}{3}\right)} \times ind_{N3-4} + \frac{\frac{S_R}{3}}{\left(\frac{S_{N1}}{8} + \frac{S_{N1}}{1} + \frac{S_{N3-4}}{2} + \frac{S_R}{3}\right)} \times ind_R
 \end{aligned} \tag{2.9}$$

Considerando  $S_{N1}$  la cantidad de épocas clasificadas como NREM 1 por el experto en la base de datos balanceada,  $S_{N2}$  el número de épocas etiquetadas por el experto como NREM 2,  $S_{N3-4}$  la cantidad de épocas etiquetadas por el experto como NREM 3-4 y  $S_R$  las épocas etiquetadas por el experto como REM. Usando el mismo protocolo de identificación de las cuatro clases se designan los índices de clasificación ( $ind_{CLASE}$ ).

#### 2.2.2.4. Test estadístico

Para determinar la idoneidad de una característica por sobre otras se hizo necesario poder comparar su capacidad para discernir entre las clases a clasificar. Para esto se recurrió al trabajo de Dietterich [49], donde se presenta una metodología para comparar dos algoritmos de clasificación que emplean aprendizaje supervisado. Esta metodología se basa en definir un estadístico  $\tilde{t}$  que responde a la distribución t de Student para determinar si la diferencia entre los rendimientos de dos clasificadores es estadísticamente significativa. Para comparar la idoneidad de dos características puntuales se generaron dos conjuntos de características iguales, excepto por las características comparadas, y se entrenó a dos clasificadores utilizando cada uno de estos conjuntos. Las características compartidas por ambos conjuntos fueron aquellas que ya habían demostrado una buena capacidad de discriminación.

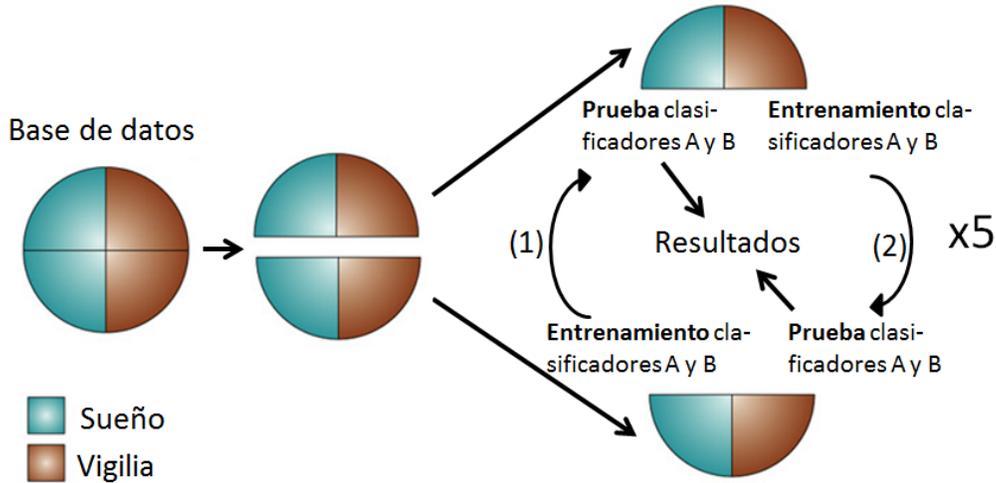


Figura 2-8. Esquema de procedimiento para prueba estadística  $\tilde{t}$  para comparación de dos clasificadores (A y B) descrita en Dietterich [49]. La base de datos balanceada se divide al azar en dos. Ambos clasificadores se entrenan utilizando la primera mitad y se prueban utilizando la segunda mitad (proceso “(1)”). De la misma forma, ambos clasificadores se entrenan con la segunda mitad de la base de datos, y se prueban con la primera (proceso “(2)”). Los resultados de ambos clasificadores para ambos procesos son comparados. El procedimiento se repite 5 veces.

Siguiendo a Dietterich [49] las comparaciones entre los clasificadores (A y B) se realizaron dentro de una validación cruzada 5x2 (figura 2-8), es decir, 5 iteraciones de validación cruzada de dos caminos. En la primera validación cruzada de dos caminos (iteración  $i = 1$ , con  $i \in [1,5]$ ) la base de datos se divide en dos mitades (primera y segunda mitad). En un primer paso (paso “(1)” de la figura 2-8), la primera mitad de la base de datos se utiliza para generar los dos conjuntos de características, los que sirven para entrenar los dos clasificadores. Estos clasificadores luego son probados sobre las características generadas a partir de los datos de la segunda mitad, obteniéndose un resultado para cada clasificador, es decir, por cada conjunto de características. Estos resultados están en la forma de la proporción de épocas mal por clasificadas por el clasificador A o B, en el paso “(1)”. A estos valores se los denomina  $p_A^{(1)}$  o  $p_B^{(1)}$ . En el paso “(2)”, la segunda mitad se utiliza para generar el conjunto de características con la que se entrena los clasificadores, los que son probados sobre el conjunto de características generado a partir de primera mitad. Desde estas pruebas se extraen otros dos resultados,  $p_A^{(2)}$  y  $p_B^{(2)}$ . En el protocolo de validación cruzada 5x2, este procedimiento se repite 5 veces, obteniéndose un total de 10 resultados de clasificación por cada algoritmo. Para cada paso, “(1)” o “(2)”, de cada iteración  $i$ , se calcula una comparación ( $p_i^{(1)}$  para el paso “(1)”) entre los resultados de los clasificadores:

$$p_i^{(1)} = p_A^{(1)} - p_B^{(1)} \quad 2.10$$

Además se calcula la varianza  $s_i^2$  entre los resultados de los clasificadores para cada iteración  $i$  del método:

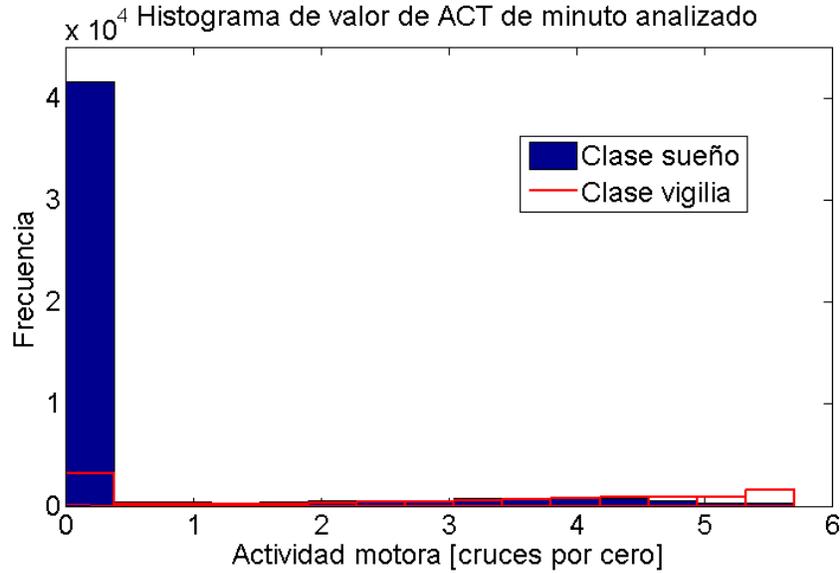


Figura 2-9. Histogramas superpuestos de clases sueño (en azul) y vigilia (en rojo) para valor ACT del minuto analizado después del pre-procesamiento. Puede apreciarse que la clase vigilia no supera en cantidad de épocas a la clase sueño en ningún valor de la actividad motora (eje X), por lo que no existe un punto donde ubicar un umbral eficiente.

$$s_i^2 = \left(p_i^{(1)} - \bar{p}_i\right)^2 + \left(p_i^{(2)} - \bar{p}_i\right)^2 \quad 2.11$$

Siendo  $\bar{p}_i$  el promedio entre  $p_i^{(1)}$  y  $p_i^{(2)}$ .

La fórmula para calcular el estadístico  $\tilde{t}$  viene dada por:

$$\tilde{t} = \frac{p_1^{(1)}}{\sqrt{\frac{1}{5} \sum_{i=1}^5 s_i^2}} \quad 2.12$$

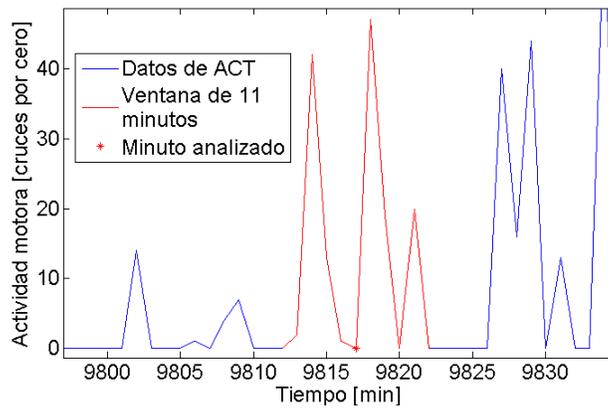
El estadístico  $\tilde{t}$  tiene una distribución estadística t de Student con 5 grados de libertad.

En la implementación de este test para comparar algoritmos de ANN cada conjunto de entrenamiento se subdivide en uno de entrenamiento y otro de validación (80% y 20% de los datos, respectivamente). Para cada prueba se entrenan diez redes distintas (entrenadas con el conjunto de entrenamiento), eligiéndose la red que mejores resultados presenta sobre el conjunto de validación para posterior evaluación en el conjunto de prueba.

### 2.2.3. Algoritmo de clasificación sueño/vigilia

#### 2.2.3.1. Extracción de características para clasificación sueño/vigilia

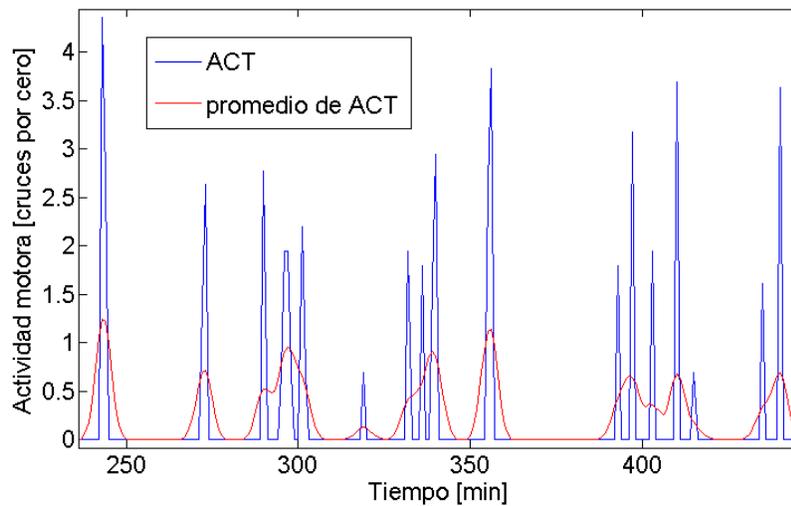
La más obvia fuente de información para clasificar un determinado minuto como sueño o vigilia es la actividad motora registrada en ese minuto. Sin embargo al utilizar esta característica para discernir entre ambas clases en un histograma (figura 2-9) se observa



**Figura 2-10. Registro actigráfico después de pre-procesamiento. Para extracción de características de un minuto en particular (asterisco rojo) se utiliza una ventana de 11 minutos (rojo) sobre la que se calculan diversos parámetros.**

que la clase vigilia está casi completamente contenida en la clase sueño. Es por esto que no es posible encontrar un umbral que entregue buenos resultados de clasificación, por lo que deben calcularse nuevas características.

Para fines explicativos, las características extraídas de los datos de actividad actigráfica pueden dividirse en dos grandes grupos, las de carácter temporal o circadiano y las de naturaleza estadística. Las características de naturaleza temporal o circadiana son las que buscan caracterizar la ubicación temporal del minuto analizado dentro del ciclo completo de sueño, indicando, por ejemplo, si se está analizando un minuto cercano al inicio del sueño o uno cercano al final. Por su parte, las características estadísticas señalan las tendencias de los valores del registro actigráfico dentro de una ventana de tiempo acotada y



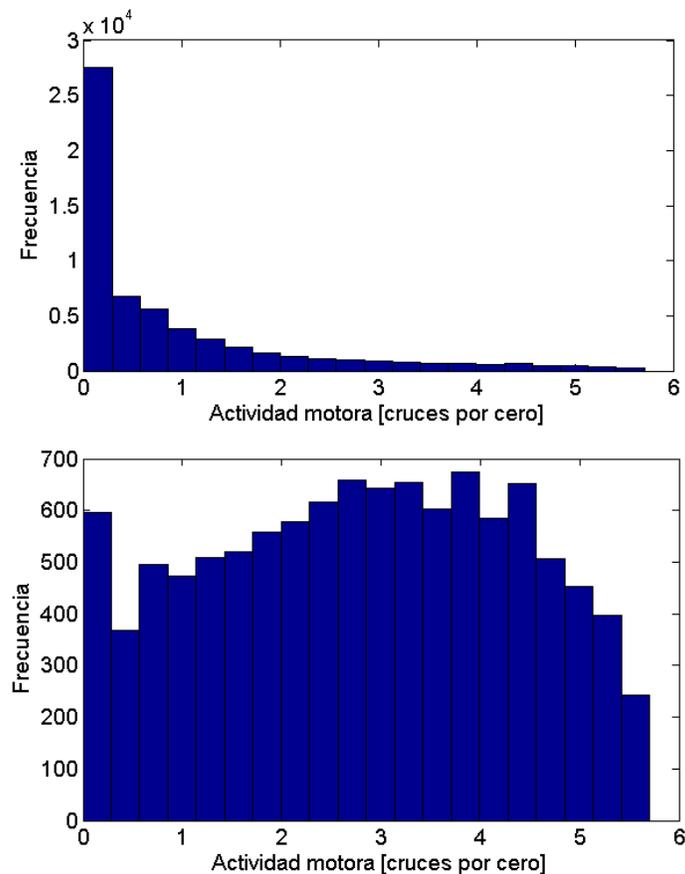
**Figura 2-11. Señal obtenida (rojo) luego de aplicar filtro promediador de ventana móvil de 11 minutos sobre actividad actigráfica luego de pre-procesamiento (azul).**

fija, permitiendo obtener información acerca de, por ejemplo, la tendencia central o la dispersión de los datos dentro de la ventana.

### 2.2.3.1.1. Ventanas de tiempo

Muchas de las características utilizadas en este trabajo son calculadas sobre ventanas móviles de tiempo ubicadas en el registro actigráfico. Se utilizan tres tipos de ventanas de tiempo: 1) una ventana que contiene la actividad actigráfica de 11 minutos, con el minuto analizado en el centro (figura 2-10); 2) una ventana con la actividad actigráfica de 20 minutos, considerando el minuto inmediatamente anterior al minuto analizado y los 19 anteriores; 3) una ventana con la actividad actigráfica de 20 minutos, tomando en cuenta el minuto inmediatamente posterior al analizado y los 19 posteriores a este.

También se emplean 2 secciones de la ventana de 11 minutos para extracción de características. Se usa una sección de ventana que contiene los primeros 6 minutos de la ventana 11, es decir, el minuto analizado y los 5 anteriores, y otra sección que contiene los últimos 6, es decir, el minuto analizado y los 5 posteriores.



**Figura 2-12. Distribuciones de la actividad actigráfica (luego de pre-procesamiento) después de ser filtrada por un filtro promediador de ventana de ancho de 11 minutos. Arriba se puede ver la distribución de toda la base de datos. Abajo se puede ver la distribución de los datos catalogados por el experto como vigilia.**

### 2.2.3.1.2. Extracción de características de tendencia central

La fuente de información más básica acerca de los valores de la ACT en la ventana analizada son los estadísticos que entregan mediciones de la tendencia central de estos valores, esto es, de la tendencia de los datos para agruparse en torno a un valor central.

Se evaluaron distintos estadísticos sobre tres segmentos de la ventana de 11 minutos: 1) sobre la ventana completa; 2) sobre el segmento que comprende el minuto analizado y los 5 minutos posteriores; y 3) sobre el segmento que comprende el minuto analizado y los 5 anteriores. Los estadísticos analizados fueron:

1) Promedio (media) aritmético:

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} \quad 2.13$$

Es el estadístico más común para medir tendencia central de los datos, debido a que entrega una buena medida de ésta para conjuntos de datos pertenecientes a una distribución normal o alguna de forma similar. Sin embargo no se comporta de buena manera si los datos pertenecen a una mezcla de dos o más distribuciones de probabilidad diferentes o si hay una presencia importante de valores atípicos (*outliers*).

El promedio puede utilizarse, a su vez, para generar un filtro pasabajo al aplicarse por ventanas móviles, obteniéndose una nueva señal al aplicarse en una ventana de 11 minutos (figura 2-11). En el caso particular de los datos aquí analizados se puede observar que los promedios de la ventana de 11 minutos para datos catalogados por el experto como vigilia presentan una distribución unimodal semejante a una gaussiana (figura 2-12). Esta distribución que presentó un valor de skewness de 0,027. Para tener un punto de comparación, se desarrolló un programa que generara 1000 datos al azar con distribución normal y luego se midió el skewness (que debiese ser 0 para distribución normal perfecta). El valor obtenido en esta prueba fue 0,037, lo que muestra que la distribución de los datos de vigilia, luego del filtro pasabajo, presenta una mediana, una moda y una media muy cercanas.

2) Mediana:

Es un estadístico vinculado a la distribución de probabilidad de los datos. La mediana es igual al dato que separa la mitad superior de la distribución de probabilidad de la mitad inferior.

La mediana es un estadístico más robusto que el promedio [50] debido a que siempre entrega un valor representativo de la tendencia central de los datos, ya que su valor no se ve afectado por la presencia de outliers.

3) Media truncada:

Es un estadístico que tiene como parámetro un determinado porcentaje  $p$  menor a 50%. Consiste en eliminar los valores que estén ubicados en el  $p\%$  mayor y en el  $p\%$  menor. A los valores restantes se les calcula el promedio aritmético.

Comúnmente se aplica con  $p = 25\%$ . Sin embargo en este estudio se aplicó con un  $p = 10\%$ , lo que permitía, según la implementación de MATLAB 2011a, eliminar el mayor y el menor valor dentro de ventanas tanto de 11 como de 6 valores (que son la cardinalidades utilizadas).

Esta medida permite obtener valores mucho más robustos para la tendencia central debido a la eliminación de outliers implícita en su cálculo.

### 2.2.3.1.3. Extracción de características de dispersión de datos

Es necesaria una medida de la dispersión de los valores de la ACT medida en el periodo comprendido en la duración de las ventanas usadas. Como candidatos a cumplir esta tarea se analizaron 3 estadísticos: la desviación estándar, la desviación mediana absoluta y la cantidad de minutos con valores mayores que un umbral.

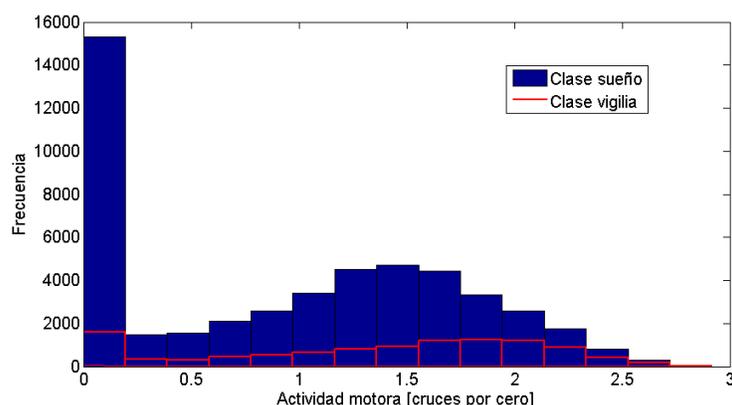
1) Desviación Estándar:

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad 2.14$$

Es el estadístico más común para medir dispersión, debido a que presenta un comportamiento muy estable y estudiado para datos con distribución normal. Sin embargo esta condición no se cumple para el *raw data* de ACT ni para los registros pre-procesados (ver figura 2-9). Sin embargo el histograma del ACT luego de que se aplicara logaritmo y ser sometido a filtro promediador (figura 2-11), da a entender que, a ventanas de tiempo mayores a un minuto, la cantidad de movimiento medido por el actígrafo puede satisfacer una distribución normal con un promedio mayor que cero. Para efectos prácticos la desviación estándar está relacionada con las diferencias entre los valores de los datos y el promedio de los datos, por lo que comparte con el promedio aritmético tanto las fortalezas (buen comportamiento en distribuciones normales) como las debilidades (poca robustez frente a outliers, pobre comportamiento para distribuciones bimodales, entre otras). Al analizarla por medio de un histograma (figura 2-13) no muestra indicios claros de poseer una capacidad importante para discriminar entre las clases por sí sola.

2) Desviación Absoluta Media, *Median Absolute Deviation* (MAD):

$$MAD = mediana_i(|x_i - mediana_j(x_j)|) \quad 2.15$$



**Figura 2-13. Superposición de histogramas de distribución de la desviación estándar de los datos de ACT después de pre-procesamiento, considerando una ventana de 11 minutos. En azul, datos pertenecientes a clase sueño, en rojo, datos pertenecientes a clase vigilia.**

Es un estadístico más robusto que la desviación estándar [50] ya que utiliza la mediana en vez de la media aritmética para calcular tanto el valor característico contra el que se comparan los datos para calcular el error, como para calcular el valor característico de estos errores.

3) Número de minutos sobre un umbral (NMSU) o *Number of minutes Above a Threshold* (NAT):

Es un indicador que mide la cantidad de minutos (dentro de la ventana de 11 minutos) que tienen valores de actividad motora por sobre un umbral determinado. Este umbral fue utilizado en Sadeh et al. [42] exitosamente, pero su aplicación a la base de datos utilizada en este estudio no fue posible debido a la ausencia de información respecto a si ambas base de datos (la utilizada en el estudio citado anteriormente y la utilizada en este estudio) son comparables. Se optó finalmente por fijar este umbral en 1 cruce por cero, valor que dio los mejores resultados a la hora de discriminar entre las dos clases en las pruebas preliminares.

2.2.3.1.4. Extracción de características temporales

Al revisar el estado del arte en clasificación de sueño y vigilia mediante registros actigráficos se puede apreciar la ausencia de características relacionadas a la ubicación temporal del minuto analizado dentro del sueño completo. Por ejemplo, si el minuto analizado junto con sus minutos inmediatamente próximos presentan una baja actividad, un clasificador tendería a clasificar ese minuto como sueño. Sin embargo, si este minuto se encuentra al inicio del periodo de sueño (por ejemplo a las 23hrs, asumiendo horarios de sueño normales), es muy posible que esta baja actividad se deba a movimientos leves propios de alguien que está tratando de conciliar el sueño. De esta forma, al utilizar sólo características estadísticas sobre ventanas temporales de corto alcance, se pierde información útil que permitiría modificar los criterios de clasificación para distintas ubicaciones temporales dentro del periodo de sueño.

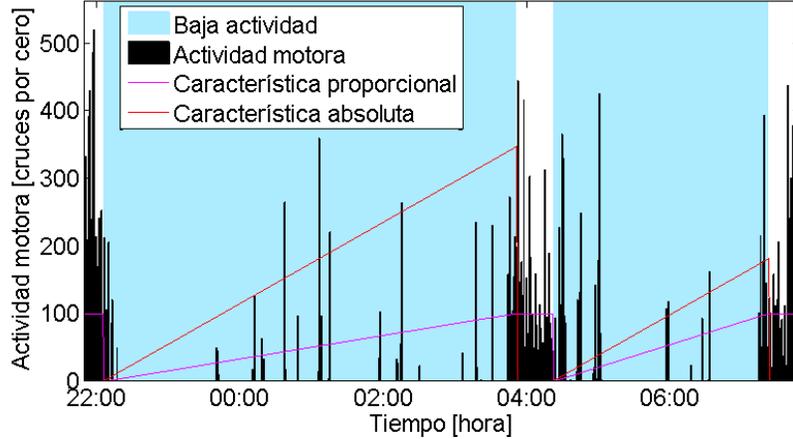


Figura 2-14. Gráfico de características temporales calculadas. En celeste están los minutos marcados como de “baja actividad”, debido a que no hay actividad importante en su vecindad. En púrpura se aprecia la característica proporcional, que indica la ubicación de cada minuto dentro del episodio de baja actividad, es decir, para agrupaciones de minutos marcados como de baja actividad consecutivos (para efectos de visualización, se limitan los valores de esta característica entre 0 y 100). En rojo se aprecia la característica de tiempo absoluta, consiste en la cantidad de minutos consecutivos de baja actividad.

El primer candidato desarrollado para cumplir esta función fue una característica que consistía en la hora del día en minutos. De esta manera, a cada minuto del registro se le asociaba un número, que correspondía al minuto del día en el que el minuto analizado había sido grabado (desde 1 hasta  $60 \times 24 = 1440$ ). Los registros utilizados, al ser todos de carácter nocturno y con sujetos con horarios de sueño nocturno, poseen cierta homogeneidad en las horas de inicio y fin del sueño, por lo que es posible establecer correspondencias entre la hora del minuto analizado y su ubicación temporal en el periodo de sueño. Sin embargo esta característica adolece que es dependiente de la regularidad de los hábitos y de los horarios de sueño de los sujetos de estudio. En el caso particular de este estudio, la regularidad en estos aspectos era esperable debido a los sujetos utilizados y los protocolos considerados, pero esta regularidad no era esperable en el resto de la población, por lo que fue necesario independizarse de esta presunción. Para ellos se intentó generar una nueva característica que entregara la misma información sin basarse en suposiciones sobre los horarios de sueño de los sujetos.

Se desarrollaron dos características que entregan información sobre en qué parte del periodo de sueño se encuentra el minuto analizado. En primer paso para calcularlas es marcar todos los minutos del registro que tienen “poca actividad”. El criterio para definir si minuto  $i$  es de baja actividad es el siguiente:

$$\sum_{j=i-5}^{i+5} \ln(ACT_j) < 10 \tag{2.16}$$

**Tabla 2-3. Características constituyentes de conjunto de características compartidas. Para comparación de características se generan dos conjuntos de características conformados por las características que se quieren comparar y las características compartidas.**

Tipo de característica	Nº	Descripción de característica
Tendencia central	1	$\log(\text{mediana}(\text{ACT}_{11\text{min}}) + 1)$
	2	$\log(\text{mediana}(\text{ACT}_{6\text{min iniciales}}) + 1)$
	3	$\log(\text{mediana}(\text{ACT}_{6\text{min finales}}) + 1)$
Dispersión	4	$\log(\text{MAD}(\text{ACT}_{6\text{min iniciales}}) + 1)$
	5	$\log(\text{MAD}(\text{ACT}_{6\text{min finales}}) + 1)$
	6	$\text{NMSU}(\text{ACT}_{6\text{min iniciales}})$
	7	$\text{NMSU}(\text{ACT}_{6\text{min finales}})$
Temporal	8	Hora del día [min]
Tendencia	9	$\log(\text{ACT} + 1)$

Definiéndose como “episodios de baja actividad” aquellos lapsos de tiempo que albergan épocas consecutivas catalogadas únicamente como de “baja actividad”. En la práctica esto se traduce en que los episodios de baja actividad abarcan, en parte importante de los registros, desde que el sujeto intenta quedarse dormido hasta cuando comienza a moverse de manera importante en la mañana siguiente. Además brinda información sobre despertares prolongados durante la noche, de manera de detectar posibles eventos de insomnio.

Luego de que se identifican los episodios de baja actividad, se procede a calcular dos características (figura 2-14):

- I. Característica temporal absoluta: valor igual al número de minutos de duración del episodio de baja actividad al medirse el minuto actual (posibles valores enteros entre 1 e infinito).
- II. Característica temporal proporcional: indicador de ubicación temporal del minuto analizado. Su valor es la proporción del minuto analizado dentro del episodio de baja actividad (posibles valores decimales entre 0, para el inicio del episodio de baja actividad, y 1, para el final).

#### 2.2.3.1.5. Extracción de características de tendencia

Como último tipo de características, y una vez que se confirmó al promedio del logaritmo natural de valores de ACT como un buen indicador para discernir entre sueño y vigilia, se consideró la utilización de esta característica como una señal extra (figura 2-9), de manera de calcular otras características basadas en ella. Estas nuevas características están vinculadas tanto a la forma de la onda (la derivada de la señal), como a las tendencias que sigue la onda en tiempos prolongados (promedios de 20 minutos anteriores y posteriores al minuto analizado).

### 2.2.3.2. Comparación entre características

Todas las pruebas realizadas para comparar características siguieron el procedimiento de cálculo de estadístico  $\tilde{t}$  con validación cruzada 5x2 [49]. Este protocolo sirve para comparar entre dos clasificadores, por lo que, para comparar el poder discriminatorio de una característica en el problema de clasificación es necesario entrenar dos ANN con conjuntos de características idénticos, a excepción de las características que se comparan. Esto se realiza eliminando de un conjunto de características compartidas previamente definido las características redundantes a las que se quieren comparar y luego agregando las que sí se quieren comparar, por ejemplo, si se quieren comparar dos características de tendencia central, en primer lugar se eliminan todas las características de tendencia central que hayan en el conjunto de características compartidas (conjunto S), luego se crean dos conjuntos que consisten en la unión del conjunto S y una de las características de tendencia central que se desea comparar. Esta característica puede estar evaluada en distintas ventanas de tiempo (por ejemplo, en la ventana de 11 minutos completa y en la sección correspondiente a los primeros 6 minutos de la ventana de 11 minutos), por lo que podría aportar más de un valor por cada época. Las ventanas de tiempo usadas para comparar cada tipo de características se eligieron de acuerdo a resultados preliminares.

Para todas las pruebas se utilizó como conjunto de características compartidas uno conformado por las características utilizadas en una publicación de un estado de avance anterior en este trabajo [51], entre estas características se cuentan la mediana, como medida de tendencia central, el MAD, como medida de dispersión, la hora del día, como

**Tabla 2-4. Subgrupos de características de tendencia central sometidos a comparación. Para realizar las comparaciones, se conformaron cuatro grupos de características constituidos por la unión del grupo de características compartidas (excepto las características de tendencia central) y cada uno de los subgrupos presentados. A la derecha puede apreciarse la tasa de mala clasificación, la que indica la capacidad del conjunto conformado para clasificar (mientras más baja, mayor capacidad).**

Subgrupo	Características	Tasa de mala clasificación
A	$\log(\text{mediana}(\text{ACT}_{11\text{min}}) + 1)$	0,1207±0,0017
	$\log(\text{mediana}(\text{ACT}_{6\text{min iniciales}}) + 1)$	
	$\log(\text{mediana}(\text{ACT}_{6\text{min finales}}) + 1)$	
B	$\text{mediana}(\log(\text{ACT}_{11\text{min}} + 1))$	0,1213±0,0014
	$\text{mediana}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{mediana}(\log(\text{ACT}_{6\text{min finales}} + 1))$	
C	$\text{trimean}(\log(\text{ACT}_{11\text{min}} + 1))$	0,1197±0,0011
	$\text{trimean}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{trimean}(\log(\text{ACT}_{6\text{min finales}} + 1))$	
D	$\text{promedio}(\log(\text{ACT}_{11\text{min}} + 1))$	0,1194±0,0009
	$\text{promedio}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{promedio}(\log(\text{ACT}_{6\text{min finales}} + 1))$	

**Tabla 2-5. Resultados de comparaciones entre características de tendencia central. Se compara los resultados de un conjunto de características (eje vertical) contra el rendimiento utilizando otro conjunto (eje horizontal). No se realizaron todas las comparaciones posibles debido a razones de costo computacional, marcando como  $\emptyset$  las que no se realizaron. Tampoco se realizaron comparación entre grupos idénticos, por lo que se marcó esos espacios como “-”. Un valor negativo en el valor de  $\tilde{t}$  implica que el clasificador entrenado con el otro conjunto obtuvo mejores resultados.**

	A	B	C	D
A	-	$\tilde{t} = 2,414$	$\tilde{t} = -0,275$	$\emptyset$
	-	$p - valor = 0,061$	$p - valor = -0,794$	$\emptyset$
B	$\tilde{t} = -2,414$	-	$\emptyset$	$\emptyset$
	$p - valor = -0,061$	-	$\emptyset$	$\emptyset$
C	$\tilde{t} = 0,275$	$\emptyset$	-	$\tilde{t} = 1,719$
	$p - valor = 0,794$	$\emptyset$	-	$p - valor = 0,146$
D	$\emptyset$	$\emptyset$	$\tilde{t} = -1,719$	-
	$\emptyset$	$\emptyset$	$p - valor = -0,146$	-

característica temporal, y el logaritmo de la actividad motora en el minuto analizado (tabla 2-3).

#### 2.2.3.2.1. Comparación entre características de tendencia central

Como se explicó anteriormente, para comparar la capacidad de las distintas características de tendencia central para discriminar entre sueño y vigilia, se procedió a generar distintos conjuntos de características que contengan las características a comparar. Tal como se explicó anteriormente, estos conjuntos contienen las características compartidas (tabla 2-3) con excepción de las de tendencia central. Además se les agregan las diferentes características de tendencia central a evaluar calculadas sobre distintas ventanas de tiempo

**Tabla 2-6. Subgrupos de características de dispersión sometidos a comparación. Para realizar las comparaciones, se conformaron cuatro grupos de características constituidos por la unión del grupo de características compartidas (excepto las características de dispersión) y cada uno de los subgrupos presentados. A la derecha puede apreciarse la tasa de mala clasificación, la que indica la capacidad del conjunto conformado para clasificar (mientras más baja, mayor capacidad).**

Subgrupos	Características	Tasa de mala clasificación
A	$\log(MAD(ACT_{6 \min \text{ iniciales}}) + 1)$	0,1211±0,0021
	$\log(MAD(ACT_{6 \min \text{ finales}}) + 1)$	
B	$MAD(\log(ACT_{6 \min \text{ iniciales}} + 1))$	0,1221±0,0023
	$MAD(\log(ACT_{6 \min \text{ finales}} + 1))$	
C	$DE(\log(ACT_{6 \min \text{ iniciales}} + 1))$	0,1211±0,0014
	$DE(\log(ACT_{6 \min \text{ finales}} + 1))$	
D	$NMSU(\log(ACT_{6 \min \text{ iniciales}} + 1))$	0,198±0,0011
	$NMSU(\log(ACT_{6 \min \text{ finales}} + 1))$	

(tabla 2-4).

Comparando los rendimientos de los clasificadores entrenados con las características a evaluar se pueden calcular los valores de certeza para los t-test. Estos tienen como hipótesis nula que las diferencias entre los rendimientos de los distintos clasificadores no son estadísticamente significativas (tabla 2-5). Se considera estadísticamente significativa cuando los valores del t-test conducen a un p-valor menor a 0,05. Utilizando esta herramienta estadística se puede concluir si la diferencia rendimiento entre los clasificadores se puede atribuir a la utilización de diferentes características de tendencia central con una certeza de 95%.

Al ver la tabla 2-5 y comparar la característica A con la B se puede apreciar que A tiene mayor poder discriminatorio, pero sin llegar a exhibir una diferencia con una significancia estadística. La característica A también se compara con la C, obteniéndose mejores resultados con ésta última, con una significancia estadística muy baja. También se realizó la comparación entre las características C y D, obteniéndose mejores resultados en la D, pero con diferencias no significativas estadísticamente.

#### 2.2.3.2.2. Comparación entre características de dispersión

Para la comparación sobre características de dispersión se siguió el mismo procedimiento empleado para las características de tendencia central. Las características de dispersión consideradas, con sus respectivos tiempos de evaluación, así como los resultados obtenidos por los conjuntos de características que los contenían pueden verse en la tabla 2-6.

Los resultados de las comparaciones pueden apreciarse en la tabla 2-7. En primer lugar se comparó las características A y B, obteniéndose mejores resultados en la primera con una diferencia estadísticamente significativa. Luego se compararon las características A y C, obteniendo mejores resultados el conjunto que contenía la característica C, aunque por un

**Tabla 2-7. Resultados de comparaciones entre características de dispersión. Se compara los resultados de un conjunto de características (eje vertical) contra el rendimiento utilizando otro conjunto (eje horizontal). No se realizaron todas las comparaciones posibles debido a razones de costo computacional, marcando como  $\emptyset$  las que no se realizaron. Tampoco se realizaron comparación entre grupos idénticos, por lo que se marcó esos espacios como “-”. Un valor negativo en el valor de  $\tilde{t}$  implica que el clasificador entrenado con el otro conjunto obtuvo mejores resultados.**

	A	B	C	D
A	-	$\tilde{t} = 4.267$	$\tilde{t} = -0.669$	$\emptyset$
	-	<b><math>p - valor = 0.008</math></b>	$p - valor = -0.533$	$\emptyset$
B	$\tilde{t} = -4.267$	-	$\emptyset$	$\emptyset$
	<b><math>p - valor = -0.008</math></b>	-	$\emptyset$	$\emptyset$
C	$\tilde{t} = 0.669$	$\emptyset$	-	$\tilde{t} = -1,072$
	$p - valor = 0.533$	$\emptyset$	-	$p - valor = -0.333$
D	$\emptyset$	$\emptyset$	$\tilde{t} = 1.072$	-
	$\emptyset$	$\emptyset$	$p - valor = 0.333$	-

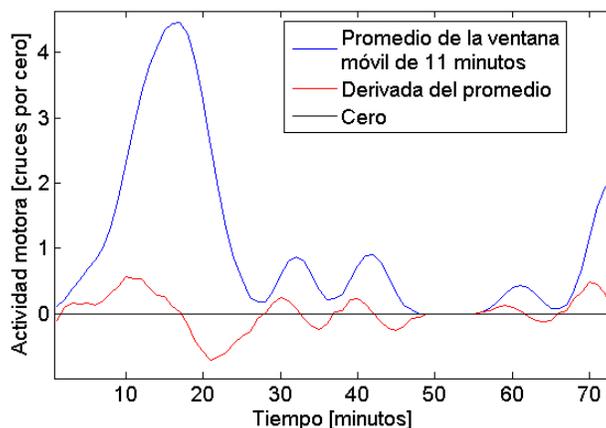


Figura 2-15. Segmento de registro actigráfico. En azul se aprecia el promedio de la actividad actigráfica después de pre-procesamiento medido en ventanas móviles de 11 minutos. En rojo se aprecia la derivada del promedio de la actividad actigráfica medido en ventanas móviles de 11 minutos. En negro se muestra el nivel del cero del promedio.

margen muy menor, lejos de ser significativo. Por último se compararon las características C y D, siendo la característica D la que mejores resultados presenta, aunque también por un margen pequeño.

#### 2.2.3.2.3. Comparación entre características temporales

Se comparó la capacidad de discriminación de la característica “hora del día” versus las dos características temporales definidas anteriormente.

Para el clasificador con hora del día se obtuvo una tasa de mala clasificación de  $0,1077 \pm 0,0026$ , mientras que el clasificador con las dos características temporales presentó una tasa de mala clasificación de  $0,1012 \pm 0,0045$ . Esto arrojó un  $\tilde{t} = 2,0394$ , lo que entregó

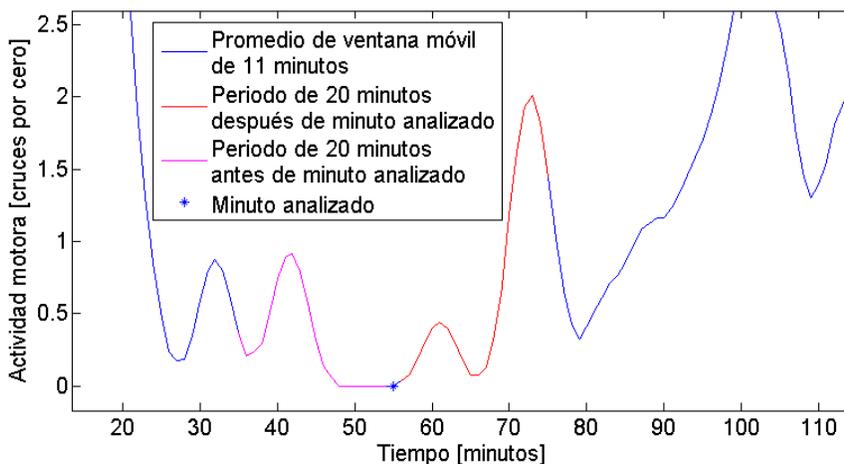


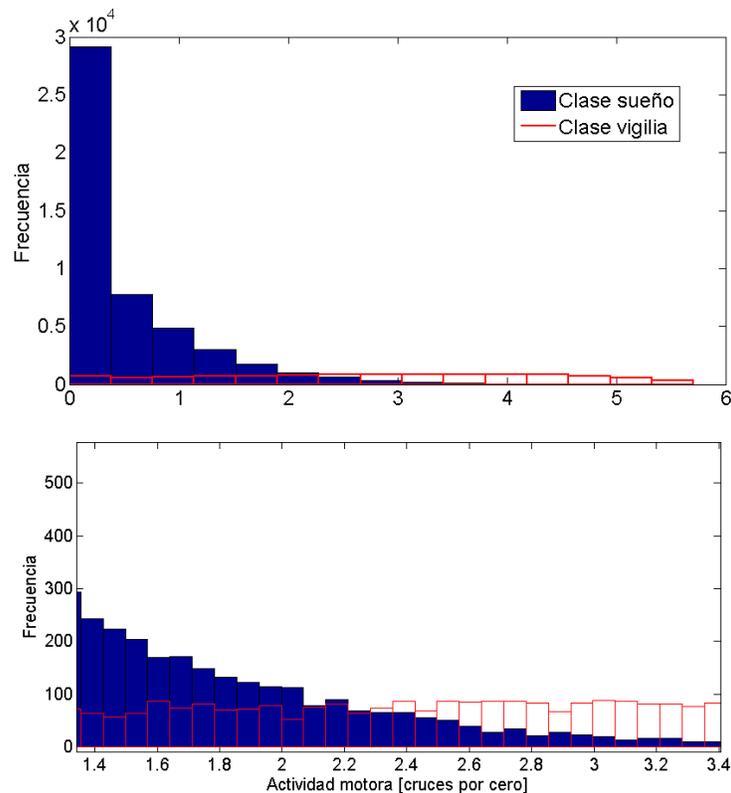
Figura 2-16. Segmento de registro actigráfico. En azul se aprecia la actividad actigráfica después de pre-procesamiento, donde se remarca el minuto analizado (asterisco). Se pueden apreciar los 20 minutos previos al minuto analizado (morado) y los 20 minutos posteriores (rojo).

un  $p - valor = 0,097$ , el que no alcanza a ser concluyente. Esto no permite concluir la existencia de una diferencia estadísticamente significativa en el rendimiento de los clasificadores entrenados con las distintas variables. Sin embargo se puede concluir que las características desarrolladas para reemplazar a la hora del día funcionaron de buena forma. Por medio de su utilización se obtuvieron incluso mejores resultados que mediante la hora del día, aunque esta diferencia no es estadísticamente significativa.

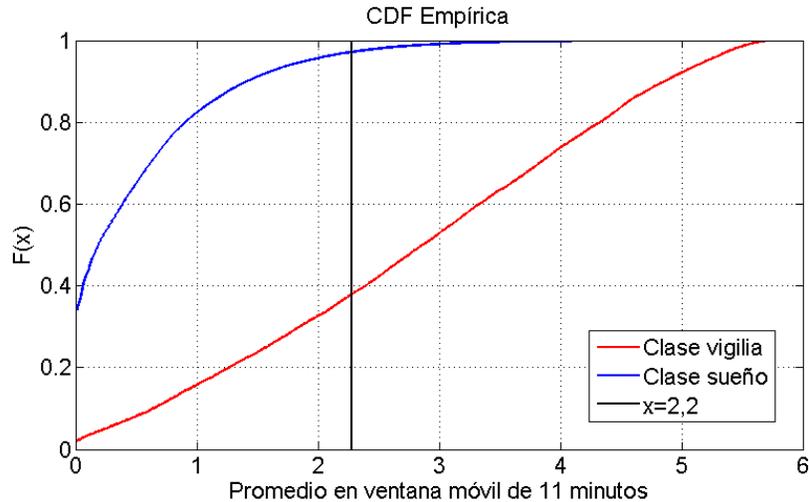
#### 2.2.3.2.4. Comparación entre características de tendencia

Debido a la utilización del promedio de la ventana de 11 minutos como un filtro pasa bajo para generar una segunda señal (figura 2-11), se considera incluir dos nuevos tipos de características extraídas de esta señal filtrada. Estas son: 1) características de forma de la onda filtrada y 2) características de contexto de actividad.

La característica de forma considerada es la derivada de la señal filtrada (figura 2-15), debido a que permite obtener información acerca de si el minuto analizado está en la parte ascendente o descendente de un evento de actividad registrado en el promedio de la actividad actigráfica.



**Figura 2-17. Superposición de histogramas de distribución de promedio de 11 minutos de actividad actigráfica después de pre-procesamiento de las clases sueño y vigilia. En rojo, datos clasificados por el experto como vigilia. En azul, datos clasificados como sueño. Abajo, acercamiento del gráfico de arriba en el punto de interés  $x=2,2$ .**



**Figura 2-18. Funciones de distribución acumulativas empíricas del promedio de ventanas de 11 minutos de la actividad actigráfica después de pre-procesamiento de las clases sueño y vigilia. En rojo se muestra los datos clasificados por el experto como vigilia. En azul, los clasificados como sueño. Se puede apreciar que el 97% de los datos pertenecientes a la clase sueño están bajo el punto 2,2, mientras que el 66% de los datos pertenecientes a la clase vigilia están sobre este punto.**

Para mejorar la información del contexto se utilizaron características de contexto a más largo alcance. Estas características son: 1) el promedio de los valores correspondientes a los 20 minutos previos al actual en la señal filtrada y 2) el promedio de los valores correspondientes a los 20 minutos posteriores al actual en la señal filtrada (figura 2-16).

### 2.2.3.3. Selección final de características

Como se aprecia en la figura 2-12 la distribución de los datos pertenecientes a la clase vigilia es cercano a una de naturaleza unimodal, muy similar a una distribución normal. En la figura 2-17 pueden apreciarse las distribuciones de las clases vigilia (en rojo) y sueño (en azul) superpuestas. De éstas se desprende que aproximadamente en el punto  $x = 2,2$  la distribución de la clase vigilia tiene más datos que la clase sueño, manteniéndose esta tendencia para  $x$  mayores.

Se analizó la función de distribución acumulativa para estas dos distribuciones (figura 2-18), y se llegó a la conclusión que, utilizando sólo la característica del promedio los valores de actigrafía (después del pre-procesamiento) de la ventana de 11 minutos se puede ubicar un umbral que logre clasificarse entre sueño y vigilia con una precisión de 91%, con una sensibilidad de 97% y una especificidad de 66%. Es por esto que se decide incluir esta característica en el sistema de clasificación, ya que sirve como clasificador preliminar, pero también entrega una segunda señal (filtrada) que puede entregar información adicional.

Analizando criterios de configuración de redes neuronales, se arribó a la conclusión de que la cantidad de características empleadas en el clasificador debe ser la menor posible. Es por esto que se decidió que el conjunto final de características consistiría en dos características

**Tabla 2-8. Subgrupos de características de tendencia central y dispersión sometidos a comparación. Para realizar las comparaciones, se conformaron cuatro grupos de características constituidos por la unión del grupo de características compartidas (excepto las características de tendencia central y las de dispersión) y cada uno de los subgrupos presentados. A la derecha puede apreciarse la tasa de mala clasificación, la que indica la capacidad de cada conjunto conformado para clasificar (mientras más baja, mayor capacidad).**

Subgrupos	Características	Tasa de mala clasificación
A	$\log(\text{mediana}(\text{ACT}_{11\text{min}}) + 1)$	0,1219±0,0011
	$\log(\text{mediana}(\text{ACT}_{6\text{min iniciales}}) + 1)$	
	$\log(\text{mediana}(\text{ACT}_{6\text{min finales}}) + 1)$	
	$\log(\text{MAD}(\text{ACT}_{6\text{min iniciales}}) + 1)$	
	$\log(\text{MAD}(\text{ACT}_{6\text{min finales}}) + 1)$	
B	$\text{mediana}(\log(\text{ACT}_{11\text{min}} + 1))$	0,1265±0,0028
	$\text{mediana}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{mediana}(\log(\text{ACT}_{6\text{min finales}} + 1))$	
	$\text{MAD}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{MAD}(\log(\text{ACT}_{6\text{min finales}} + 1))$	
C	$\text{trimean}(\log(\text{ACT}_{11\text{min}} + 1))$	0,1201±0,0011
	$\text{trimean}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{trimean}(\log(\text{ACT}_{6\text{min finales}} + 1))$	
	$\text{DE}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{DE}(\log(\text{ACT}_{6\text{min finales}} + 1))$	
D	$\text{promedio}(\log(\text{ACT}_{11\text{min}} + 1))$	0,1198±0,0011
	$\text{promedio}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{promedio}(\log(\text{ACT}_{6\text{min finales}} + 1))$	
	$\text{NMSU}(\log(\text{ACT}_{6\text{min iniciales}} + 1))$	
	$\text{NMSU}(\log(\text{ACT}_{6\text{min finales}} + 1))$	

de naturaleza temporal, tres de tendencia central, dos de dispersión, una de forma y dos de actividad de contexto, sumado además al valor de ACT del minuto analizado.

Las características de tendencia central fueron aplicadas a la ventana de 11 minutos completa, al segmento de 6 minutos iniciales y al segmento de 6 minutos finales. Por su parte, las características de dispersión fueron aplicadas a la ventana de 11 minutos, y se incluirá una característica igual a alguna medida de dispersión aplicada al segmento de 6 minutos iniciales menos esta medida de dispersión aplicada al segmento de 6 minutos finales.

Como las pruebas individuales entre estadísticos (considerando sólo los de tendencia central o sólo los de dispersión, dejando todo lo demás sin modificar) no arrojaron resultados concluyentes, se procedió a realizar pruebas mezclando un estadístico de dispersión con otro de tendencia central. Las características que fueron agrupadas, así como su resultado, pueden verse en la tabla 2-8. De un análisis de estas pruebas se desprende que

**Tabla 2-9. Resultados de comparaciones entre características de dispersión. Se compara los resultados de un conjunto de características (eje vertical) contra el rendimiento utilizando otro conjunto (eje horizontal). No se realizaron todas las comparaciones posibles debido a razones de costo computacional, marcando como  $\emptyset$  las que no se realizaron. Tampoco se realizaron comparación entre grupos idénticos, por lo que se marcó esos espacios como “-”. Un valor negativo en el valor de  $\tilde{t}$  implica que el clasificador entrenado con el otro conjunto obtuvo mejores resultados.**

	A	B	C	D
A	-	$\tilde{t} = 1,355$	$\tilde{t} = -3,978$	$\emptyset$
	-	$p - valor = 0,239$	$p - valor = -0.011$	$\emptyset$
B	$\tilde{t} = -1,355$	-	$\emptyset$	$\emptyset$
	$p - valor = -0,239$	-	$\emptyset$	$\emptyset$
C	$\tilde{t} = 3,978$	$\emptyset$	-	$\tilde{t} = -1,153$
	$p - valor = 0.011$	$\emptyset$	-	$p - valor = -0.301$
D	$\emptyset$	$\emptyset$	$\tilde{t} = 1.153$	-
	$\emptyset$	$\emptyset$	$p - valor = 0.301$	-

los mejores pares de características para clasificar entre sueño y vigilia son  $SD(\log(1 + ACT_{xmin}))$  y  $trimean(\log(1 + ACT_{xmin}))$ ; y  $NAT(\log(1 + ACT_{xmin}))$  y  $prom(\log(1 + ACT_{xmin}))$ , ambos mostrando diferencias estadísticamente significativas con los restantes dos pares de características (tabla 2-9). Se optó por elegir la primera pareja de características debido a que entrega una mayor robustez al medir la tendencia central de los datos a través de la media truncada, permitiendo además obtener información de grandes fluctuaciones mediante la desviación estándar, algo que puede realizar el NAT.

Para representar a un minuto  $i$  se eligieron las siguientes características (aplicadas sobre los datos ACT pre-procesados):

- I. Conteo de minutos dentro de episodios de baja actividad
- II. Ubicación del minuto analizado dentro del episodio de baja actividad
- III. Promedio de los valores de actividad de la ventana que comprende los minutos desde  $i - 5$  a  $i + 5$ .
- IV. Media truncada de los valores de actividad desde minuto  $i$  al  $i + 5$ .
- V. Media truncada de los valores de actividad desde minuto  $i - 5$  al  $i$ .
- VI. Desviación estándar de los valores de actividad de la ventana que comprende los minutos desde  $i - 5$  a  $i + 5$ .
- VII. Desviación estándar de los valores de actividad de la ventana que va desde el minuto  $i$  hasta el  $i + 5$ , menos la desviación estándar de los valores de actividad de la ventana que va desde el minuto  $i - 5$  hasta el  $i$ .
- VIII. Promedio de la característica III en el segmento comprendido entre  $i - 20$  e  $i - 1$ .
- IX. Promedio de la característica III en el segmento comprendido entre  $i + 1$  e  $i + 20$ .
- X. Valor de función derivada de la característica III evaluada en el minuto  $i$ .
- XI. Actividad motora en el minuto  $i$ .

#### 2.2.4. Algoritmo de clasificación de estados y etapas del sueño

Se optó por evaluar en primera instancia la capacidad con que contaba el conjunto de características escogido para clasificar entre sueño y vigilia para diferenciar entre los estados y etapas del sueño.

##### 2.2.4.1. Extracción de características para clasificación de estados y etapas del sueño

Mediante Stepwise Linear Regression (SLR) se llegó a la conclusión de que dos características usadas para determinar sueño y vigilia no eran útiles para clasificar etapas del sueño. El método SLR utiliza pruebas estadísticas (t-test) para determinar el conjunto de características que entrega un valor de coeficiente de determinación ( $R^2$ ) más cercano a 1, asumiendo que éste es el que predice de mejor forma las variaciones de la variable objetivo (en este caso, un vector que contiene las clasificaciones de estados y etapas del sueño).

Las dos características que no fueron útiles para clasificar etapas del sueño fueron: 1) la media truncada de los valores de actividad desde minuto  $i - 5$  al  $i$  (característica V) y 2) la desviación estándar de los valores de actividad de la ventana que comprende los minutos desde  $i - 5$  a  $i + 5$  (característica VI). La característica V tiene un valor en el t-test de 0,095, lo que entrega un p-valor de 0,924, lo que claramente indica que no es útil. La característica VI obtiene un valor en su t-test de 1,279, con un p-valor de 0,201.

Las características que más discriminaron entre los estados y etapas del sueño son: 1) el promedio de la característica III en el segmento comprendido entre  $i - 20$  e  $i - 1$  (característica VIII), con t-test con valor de 27,521; 2) el promedio de los valores de actividad de la ventana que comprende los minutos desde  $i - 5$  a  $i + 5$  (característica III), con un t-test de valor 12,711; 3) la desviación estándar de los valores de actividad de la ventana que va desde el minuto  $i$  hasta el  $i + 5$ , menos la desviación estándar de los valores de actividad de la ventana que va desde el minuto  $i - 5$  hasta el  $i$  (característica VII), con un valor de t-test de 12,019 y 4) el valor de función derivada de la característica III evaluada en el minuto  $i$  (característica X), con un valor de 11,197 en su t-test.

Finalmente, para corroborar los resultados entregados por SLR se realizó un test de validación cruzada (CV, por sus siglas en inglés) 5x2, obteniéndose un índice de mala clasificación de  $0,5145 \pm 0,0056$  para la prueba con 11 características, mientras que el índice de mala clasificación para la prueba con 9 características fue de  $0,5124 \pm 0,0065$  entregando un  $\tilde{t}$  de comparación de los rendimientos igual a 1,456, al que se le asocia un p-valor de 0,205, lo que implica que no hay diferencias significativas entre los clasificadores. Sin embargo, y aunque las diferencias entre clasificadores sean mínimas, es deseable utilizar la menor cantidad de entradas posibles al algoritmo de clasificación, ya que de esta manera se gasta menos tiempo en entrenarlos o en calcular los parámetros del modelo, según la naturaleza del algoritmo.

### 3. Resultados

Tal como se mencionó anteriormente, los resultados se dividieron entre clasificaciones sueño/vigilia y clasificaciones para etapas y estados del sueño, dejando para una etapa final la unión de ambos clasificadores. Esto se corresponde con la manera en que el problema fue abordado y facilita su análisis.

#### 3.1. Clasificación sueño/vigilia

En esta sección se muestran la metodología para medir los resultados y se presentan los resultados.

##### 3.1.1. Procedimiento de medición de resultados

Debido al reducido tamaño de la base de datos se optó por extraer resultados de cada clasificador empleando validación cruzada. Dentro de los diferentes tipos de validaciones cruzadas, se eligió usar validación cruzada de dos caminos repetida 5 veces, ya que, si bien no es el más apropiado para este caso si se aplica una sola vez, al repetirse sí entrega buenas aproximaciones de lo que sería la utilización de conjuntos de entrenamiento y prueba separados (que es lo que se busca) [49]. Además se tiene la ventaja que, mediante el mismo mecanismo utilizado anteriormente para comparaciones de características (t-test de 5x2 CV), se pueden realizar comparaciones entre los resultados obtenidos por el clasificador de sueño y vigilia desarrollado en este estudio y otros sistemas similares disponibles en la literatura.

Todos los resultados presentados en esta sección fueron obtenidos sobre conjuntos de prueba creados durante la validación cruzada 5x2, por lo que corresponden a promedios de 10 mitades de una base de datos balanceada.

##### 3.1.2. Resultados

La precisión obtenida para la base de datos balanceada fue  $89,9 \pm 0,4\%$ , con una sensibilidad de  $92,2 \pm 0,3\%$  y una especificidad de  $87,3 \pm 1,0\%$  (tabla 3-1).

**Tabla 3-1. Resultados de clasificación sueño y vigilia para la base de datos balanceada utilizando el punto de operación propio de la base de datos balanceada. Los valores de la diagonal en la matriz son las épocas bien clasificadas, las que al dividirse por el número total de épocas entregan el valor de la precisión.**

		Juicio experto usando PSG		Total	
		<i>Sueño</i>	<i>Vigilia</i>		
Resul. Clasif.	<i>Sueño</i>	<i>minutos</i>	<b>22580±87</b>	2738±217	25318
		<i>% columna</i>	92,2	12,7	
	<i>Vigilia</i>	<i>minutos</i>	1915±79	<b>18828±183</b>	20743
		<i>% columna</i>	7,8	87,3	
<b>Total</b>		24495	21566	<b>46061</b>	
<b>Precisión</b>				89,9	

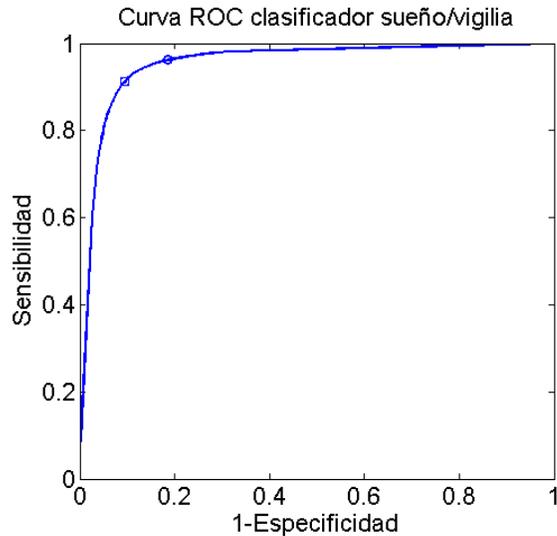
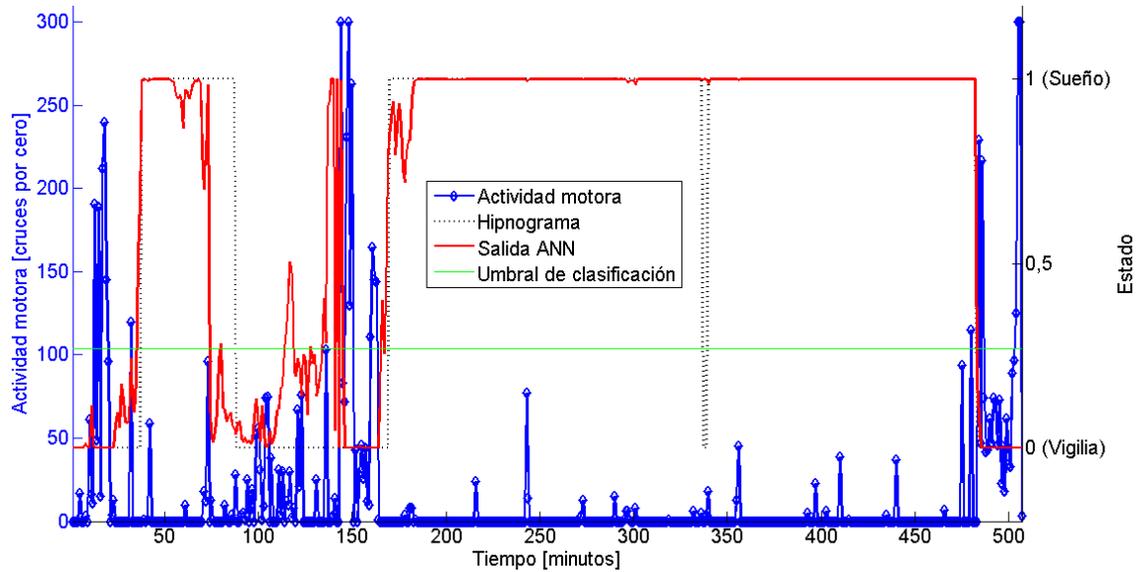


Figura 3-1. Curva ROC de clasificador sueño/vigilia. En el eje X, 1-Especificidad, en el eje Y, Sensibilidad. Se marca punto de operación calculado para base de datos equilibrada (cuadrado) y para base de datos original (círculo).

Considerando la distinta proporción entre las cardinalidades de las clases vigilia y sueño se pudieron calcular modificaciones a los umbrales de clasificación aplicados a la salida de la ANN que maximicen la precisión de clasificación para este desbalance. Para esto se calcula la curva ROC, caracterizada por pares de valores de sensibilidad y especificidad. Cada uno de estos pares corresponden a diferentes puntos de operación (figura 3-1), esto es, diferentes valores del umbral de clasificación aplicado a la salida de la ANN (figura 3-2), que originalmente es 0,5. Uno de estos puntos de la curva ROC maximiza la precisión de la base de datos original, la que se calculó siguiendo los pasos descritos en la sección 2.2.2.3.1.

Tabla 3-2. Resultados de clasificación sueño y vigilia tanto en base de datos balanceada como en base de datos original (cardinalidad de clase vigilia dividida en 4) utilizando el punto de operación calculado para base de datos original. Los valores en las diagonales de las matrices contienen las épocas bien clasificadas. Dividiendo la suma de éstos por el número total de épocas de cada base de datos se obtiene el valor de la precisión en dicha base de datos.

		Juicio experto usando PSG					
				Base de datos balanceada		Base de datos original	
				Sueño	Vigilia	Sueño	Vigilia/4
		Resul. Clasif.	Sueño	minutos	23554±90	4409±279	23554±90
% columna	96,2			20,4	96,2	20,4	
Resul. Clasif.	Vigilia	minutos	941±88	17157±312	941±88	4289±78	
		% columna	3,8	79,6	3,8	79,6	
Total			46061		29886		
Precisión				88,4		93,2	



**Figura 3-2.** Salidas del clasificador ANN sueño/vigilia (rojo) en función de la actividad motora (azul) y del estado del sujeto (sueño o vigilia) marcado por expertos médicos (negro). En verde se puede apreciar el umbral utilizado por el punto de operación correspondiente a la proporción original entre las clases, umbral igual a 0,28 (verde).

La precisión total de esta clasificación en la base de datos balanceada fue  $88,4 \pm 0,6\%$ , con una sensibilidad de  $96,2 \pm 0,4\%$  y una especificidad de  $79,6 \pm 1,3\%$  (tabla 3-2). Sin embargo, al calcular la precisión sobre la base de datos original esta aumentó a  $93,2\%$ , conservándose los valores de sensibilidad y especificidad.

### 3.2. Clasificación de estados y etapas del sueño

#### 3.2.1. Procedimiento de medición de resultados

Los resultados de la clasificación de estados y etapas del sueño se llevan a cabo sin minutos clasificados como vigilia por el criterio experto. Esto es debido a que se intentó aislar los funcionamientos de los clasificadores.

**Tabla 3-3.** Cardinalidad y proporción de las clases dentro de la base datos utilizada en este estudio.

Clase	Minutos	%
NREM 1	3083	6,3
NREM 2	25393	51,8
NREM 3	5081	10,4
NREM 4	6520	13,3
REM	8913	18,2
<b>Total</b>	<b>48990</b>	<b>100</b>

Tabla 3-5. Resultados de clasificación de estados y etapas del sueño para base de datos balanceada en el punto de operación correspondiente a la base de datos balanceada. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que al ser divididas por el número total de épocas entregan el valor de la precisión alcanzada con este punto de operación.

		Juicio experto usando PSG				
		<i>NREM 1</i>	<i>NREM 2</i>	<i>NREM 3-4</i>	<i>REM</i>	Total
<b>NREM 1</b>	<i>minutos</i>	<b>7660±179</b>	2476±163	609±81	1982±162	12727
	<i>% colum</i>	62,1	19,5	5,2	14,8	
<b>NREM 2</b>	<i>minutos</i>	1398±415	<b>3765±584</b>	1485±70	2383±504	9031
	<i>% colum</i>	11,3	29,7	12,8	17,8	
<b>NREM 3-4</b>	<i>minutos</i>	706±63	2304±161	<b>8277±178</b>	1603±129	12890
	<i>% colum</i>	5,7	18,1	71,4	12,0	
<b>REM</b>	<i>minutos</i>	2568±313	4151±655	1231±273	<b>7402±712</b>	15352
	<i>% colum</i>	20,8	32,7	10,6	55,4	
<b>Total</b>		12332	12696	11602	13370	<b>50000</b>
<b>Precisión</b>						54,2

Estas clasificaciones, entregadas por los expertos del INTA, tienen 5 posibles etiquetas de acuerdo a la clasificación R&K [10] (clasificación distribuidas según tabla 3-3). Estas etiquetas son sueño quieto 1 (NREM 1), sueño quieto 2 (NREM 2), sueño quieto 3 (NREM 3), sueño quieto 4 (NREM 4) y sueño paradójico (REM). Sin embargo, tal como se mencionó en la sección 2.2.2.3.2, se llegó a la conclusión de que la unión de las clases NREM 3 y 4 sería beneficiosa para el sistema por dos motivos: 1) el sistema automático utilizado en las pruebas (con las 9 características definidas anteriormente) tiene problemas en distinguir entre los minutos pertenecientes a NREM 3 y NREM 4, por lo que si estas clases se fusionan, el rendimiento del sistema tiende a mejorar y 2) la unión de las etapas NREM 3 y NREM 4 está considerada en el nuevo manual de clasificación de estados y etapas del sueño publicado por la academia estadounidense de medicina del sueño (AASM)

Tabla 3-4. Resultados de clasificación de etapas del sueño para base de datos balanceada en el punto de operación correspondiente a la base de datos original. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que divididas por el número total de épocas entregan el valor de la precisión para este punto de operación.

		Juicio experto usando PSG				
		<i>NREM 1</i>	<i>NREM 2</i>	<i>NREM 3-4</i>	<i>REM</i>	Total
<b>NREM 1</b>	<i>minutos</i>	<b>1500±380</b>	148±61	19±20	57±36	1725
	<i>% colum</i>	12,2	1,2	0,2	0,4	
<b>NREM 2</b>	<i>minutos</i>	9920±690	<b>10422±1092</b>	3918±483	11098±853	35359
	<i>% colum</i>	79,5	80,6	30,7	82,1	
<b>NREM 3-4</b>	<i>minutos</i>	622±220	1854±350	<b>7624±482</b>	1321±315	11422
	<i>% colum</i>	6,0	16,1	68,8	10,9	
<b>REM</b>	<i>minutos</i>	290±335	272±316	39±47	<b>893±1048</b>	1494
	<i>% colum</i>	2,3	2,1	0,3	6,6	
<b>Total</b>		12332	12696	11602	13370	<b>50000</b>
<b>Precisión</b>						41,2

Tabla 3-6. Resultados de clasificación de etapas del sueño para base de datos original con el punto de operación correspondiente a la base de datos original. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que divididas por el número total de épocas entregan el valor de la precisión para este punto de operación..

		Juicio experto usando PSG				Total
		<i>NREM 1</i>	<i>NREM 2</i>	<i>NREM 3-4</i>	<i>REM</i>	
<b>NREM 1</b>	<i>minutos</i>	<b>188±48</b>	148±61	10±10	19±12	365
	<i>% colum</i>	12,2	1,2	0,2	0,4	
<b>NREM 2</b>	<i>minutos</i>	1240±86	<b>10422±1092</b>	1959±242	3699±284	17320
	<i>% colum</i>	79,5	80,6	30,7	82,1	
<b>NREM 3-4</b>	<i>minutos</i>	78±28	1854±350	<b>3812±241</b>	440±105	6184
	<i>% colum</i>	6,0	16,1	68,8	10,9	
<b>REM</b>	<i>minutos</i>	36±42	272±316	20±24	<b>298±349</b>	626
	<i>% colum</i>	2,3	2,1	0,3	6,6	
<b>Total</b>		1542	12696	5801	4456	<b>24495</b>
<b>Precisión</b>					61,0	

[13], por lo que la nueva clasificación es aún útil.

Las cardinalidades de cada clase se detallan en la tabla 3-3, mientras que los factores utilizados para balancear la base de datos fueron expuestos en la tabla 2-2. Los resultados de la clasificación para el punto de operación calculado a partir de la base de datos balanceada se aprecian en la tabla 3-4.

A diferencia del clasificador de sueño/vigilia el clasificador de estados y etapas se tienen 4 salidas, cada una asociada a una clase. La clasificación dada a cada dato corresponde a la clase asociada a la salida con el valor más alto.

### 3.2.2. Resultados

Para el punto de operación calculado considerando la base de datos balanceada (tabla 3-4) se tuvo una precisión de  $54.2 \pm 0.5\%$ , con un porcentaje de clasificación correcta para NREM 1 (número de minutos donde sistema decide NREM 1 dividido por el número de minutos que experto marca como NREM 1) de  $62.1 \pm 1.8\%$ , un porcentaje de clasificación correcta para NREM 2 de  $29.7 \pm 4.6\%$ , un porcentaje de clasificación correcta para la clase que une NREM 3 con NREM 4 de  $71.4 \pm 1.7\%$  y un porcentaje de clasificación correcta para REM de  $55.4 \pm 5.5\%$ . El punto de operación que entregó estos resultados se obtuvo al sumarle  $-0,05$ ,  $0$ ,  $0$  y  $-0,05$  a las salidas de las 4 clases respectivamente.

En cuanto a los resultados obtenidos para la base de datos balanceada, pero con el punto de operación correspondiente a base original (tabla 3-5), la precisión total fue de  $41,2 \pm 2,0\%$ , con un porcentaje de clasificación correcta de NREM 1 de  $12,2 \pm 3,2\%$ , un porcentaje de clasificación correcta de NREM 2 de  $80,7 \pm 1,4\%$ , un porcentaje de clasificación correcta de la clase que une NREM 3 con NREM 4 de  $68,8 \pm 4,4\%$  y un porcentaje de clasificación

correcta de REM de  $6,7\pm 7,9\%$ . Al aplicarse este punto de operación a la base de datos original (tabla 3-6), se obtuvo un valor de precisión de  $61,0\pm 0,3\%$ , manteniéndose los valores de clasificación para todas las clases. Las modificaciones a las salidas de la ANN definidas para este punto de operación fueron  $-0,45$ ,  $0,3$ ,  $0,1875$  y  $-0,2625$  respectivamente.

### 3.3.Unificación de clasificadores

Se procedió a unificar los dos clasificadores que componen el sistema. Para hacerlo se tomó toda la base de datos y se probó sobre las redes generadas por los procedimientos anteriormente descritos. Finalmente se introdujeron como entradas del segundo clasificador los minutos clasificados como sueño según la salida del primer clasificador.

Los resultados se muestran en la tabla 3-7. Se obtuvo una precisión total de  $62,4\pm 0,2\%$ . Es importante notar que los minutos etiquetados como sueño por los expertos y clasificados como vigilia por el sistema de clasificación no aparecen explícitamente en la tabla 3-7. Esto se debe a que no es posible determinar cuál de las etiquetas se le asignó, ya que no fueron ingresados al clasificador de estados y etapas del sueño.

**Tabla 3-7. Resultados de los dos clasificadores combinados para clasificar todos los estados y etapas del ciclo sueño vigilia en la base de datos original utilizando los puntos de operación desarrollados para cada clasificador para la base de datos original. En la diagonal de la matriz (en negritas) se aprecian las épocas bien clasificadas, las que divididas por el número total de épocas entregan el valor de la precisión correspondiente a los puntos de operación escogidos.**

		Juicio experto usando PSG					
		<i>Vigilia</i>	<i>NREM 1</i>	<i>NREM 2</i>	<i>NREM 3-4</i>	<i>REM</i>	<b>Total</b>
<b>Vig</b>	<i>minutos</i>	<b>8901±171</b>	0*	0*	0*	0*	10881*
	<i>% columna</i>	82.5	0*	0*	0*	0*	
<b>N1</b>	<i>minutos</i>	103±70	<b>130±690</b>	142±1092	12±483	24±853	411
	<i>% columna</i>	1.0	5.4	5.8	0.1	0.3	
<b>N2</b>	<i>minutos</i>	1568±196	2118±690	<b>20132±1092</b>	3478±483	7611±853	34906
	<i>% columna</i>	14.5	87.6	82.6	30.4	86.9	
<b>N3-4</b>	<i>minutos</i>	196±33	155±220	4005±350	<b>7948±482</b>	948±315	13252
	<i>% columna</i>	1.8	6.4	16.4	69.4	10.8	
<b>R</b>	<i>minutos</i>	15±22	14±335	104±316	11±47	<b>179±1048</b>	322
	<i>% columna</i>	0.1	0.6	0.4	0.1	2.0	
<b>Total</b>		10783	2417	24382	11449	8762	<b>59773</b>
						<b>Precisión</b>	62,4

\*: El primer clasificador (sueño/vigilia) no emite opinión acerca de a qué estado o etapa pertenece algún minuto, sólo indica si pertenece a sueño o vigilia. Por lo tanto no hay modo de saber a qué clase se adjudican los minutos que el primer clasificador clasificó como vigilia, pero que estaban etiquetados por el experto como sueño.

## 4. Discusión

### 4.1. Análisis de resultados

#### 4.1.1. Clasificación sueño/vigilia

Analizando los resultados que consideran una base de datos balanceada se puede apreciar que, a pesar de haber pasado de una proporción 82-18 a una 53-47, el sistema mantiene una tendencia a favor de la clase sueño por sobre la clase vigilia (25318 a 20743). Esto puede deberse al hecho de que es más difícil caracterizar un minuto de vigilia que uno de sueño.

Otro punto importante a analizar en esta clasificación es la precisión temporal del sistema, es decir, cuanto demora el sistema en detectar un cambio de estado. Para obtener información sobre este punto se realizó una prueba que consta de reconsiderar malas clasificaciones como buenas si es que éstas cumplen la condición de que a una distancia  $n$  del minuto mal clasificado haya un minuto clasificado como lo indicó el sistema (erróneamente), con  $n$  que varía entre 1 y 5. Este rango fue considerado debido a que es el rango de tiempo utilizado para calcular directamente las características desde la III hasta la VII e indirectamente desde la VIII hasta la X, por lo que podría considerarse éste como el rango de sensibilidad del sistema.

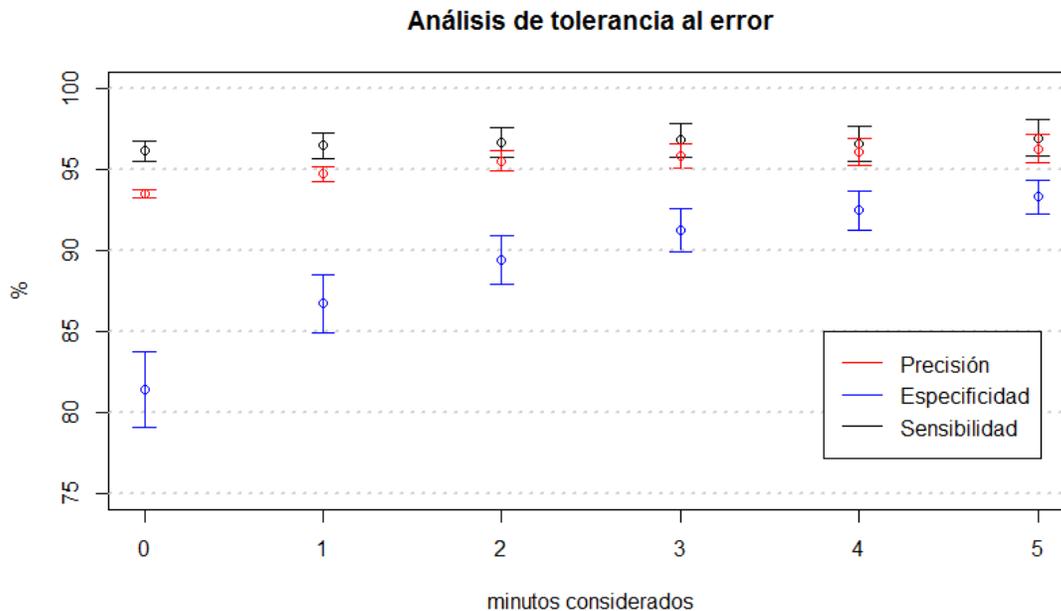


Figura 4-1. Evolución de la precisión, la sensibilidad y la especificidad a medida que aumentan los minutos considerados en el periodo de tiempo de tolerancia al error. Se considera una tolerancia al error de “N” si un minuto  $m$  está mal clasificado, pero si en su vecindad de ancho  $N$  existe un minuto que tiene la misma clasificación predicha por el sistema para  $m$ . Un círculo marca el valor del estadístico calculado, mientras que la barra que lo rodea marca la desviación estándar del estadístico.

Tabla 4-1. Resultados de clasificación de toda la base de datos en el punto de operación para base de datos original sin aplicar reglas de posprocesamiento.

			Juicio experto usando PSG		Total
			<i>Sueño</i>	<i>Vigilia</i>	
Resul. Clasif.	<i>Sueño</i>	<i>minutos</i>	<b>47109±215</b>	1945±176	49054
		<i>% columna</i>	96,2	18,0	
	<i>Vigilia</i>	<i>minutos</i>	1881±215	<b>8838±176</b>	10719
		<i>% columna</i>	3,8	82,0	
<b>Total</b>			48990	10783	<b>59773</b>
<b>Precisión</b>					93,6

Los resultados de esta prueba pueden verse en la figura 4-1, donde se aprecia que la precisión alcanzó el 96% considerando 5 minutos, mientras que la sensibilidad llegó al 97%. La especificidad, por su parte alcanzó 93%.

Una aplicación para el procedimiento anteriormente descrito es la maximización de la detección de vigilia otorgando cierta cantidad de minutos de tolerancia. Esta aplicación consistiría en generar un detector especializado en episodios de vigilia que permita, con cierto margen de error, advertir episodios de vigilia a lo largo de la noche, permitiendo obtener información acerca de si el sujeto tiene problemas para conciliar el sueño. Para esto se necesita encontrar un punto de operación que presente un nivel de detección de vigilia alto, pero sin descuidar el valor de predicción de la clase vigilia, de manera de no aumentar el índice de falsas alarmas. Lo anterior, unido a una tolerancia del error de ciertos minutos solamente aplicada a los datos de la clase vigilia, puede incrementar drásticamente la especificidad sin descuidar la sensibilidad.

Otro análisis posible es el del rendimiento del sistema si se le imponen reglas simples a la salida de la ANN para minimizar ruidos o clasificaciones aisladas. Reglas de esta especie ya fueron usadas anteriormente [47]. Se implementaron y probaron dos reglas: 1) Si una época *E* está clasificada con una etiqueta A, pero las dos épocas vecinas están clasificadas

Tabla 4-2. Resultados de clasificación sueño vigilia de toda la base de datos en el punto de operación para base de datos original luego de aplicar regla N° 1 de posprocesamiento, que consiste en variar la clasificación de todos los minutos que están rodeados por minutos marcados con la clasificación opuesta.

			Juicio experto usando PSG		Total
			<i>Sueño</i>	<i>Vigilia</i>	
Resul. Clasif.	<i>Sueño</i>	<i>minutos</i>	<b>47225±208</b>	1916±176	49141
		<i>% columna</i>	96,4	17,8	
	<i>Vigilia</i>	<i>minutos</i>	1765±208	<b>8867±176</b>	10632
		<i>% columna</i>	3,6	82,2	
<b>Total</b>			48990	10783	<b>59773</b>
<b>Precisión</b>					93,8

como B, se cambiará la etiqueta de la época  $E$ , dejándola finalmente como B. 2) Si dos épocas consecutivas  $E_1$  y  $E_2$  están clasificadas con una misma etiqueta A, pero las dos épocas anteriores y las dos posteriores a la pareja  $E_1$  y  $E_2$  están clasificadas con una etiqueta B, entonces las épocas  $E_1$  y  $E_2$  se etiquetan como pertenecientes a la clase B.

Para probar estas reglas de posprocesamiento de la señal de salida de la ANN se deben considerar la naturaleza de serie de tiempo del registro actigráfico. Esta propiedad fue considerada al extraer las características utilizadas en el clasificador, pero no al evaluar sus resultados, por lo que las pruebas desarrolladas para verificar el funcionamiento de estas reglas no pudieron llevarse a cabo sobre las salidas de las pruebas de validación cruzadas 5x2. Debido a esto las pruebas fueron llevadas a cabo sobre toda la base de datos (es decir, la unión del conjunto de entrenamiento y de prueba de cada validación cruzada). Esto implica que los resultados mostrados para estas pruebas no son comparables con los mostrados en el capítulo “Resultados” (capítulo 3), y sólo sirven a manera de punto de comparación de la mejoría del rendimiento del sistema.

Tal como puede verse al comparar las tablas 4-1 y 4-2, la regla N°1 logra agregar cerca de 250 épocas al conjunto de las épocas bien clasificadas, mejorando en 0,2% la sensibilidad, la especificidad y la precisión. Por su parte la regla N°2 (tabla 4-3) agrega cerca de 70 épocas al conjunto de épocas bien clasificadas, mejorando cerca de un 0,1% la sensibilidad y la especificidad, además de incrementar en un 0,2% la precisión.

#### 4.1.1.1. Utilidad de características desarrolladas

Una de las innovaciones presentadas en este trabajo es la utilización de características destinadas a medir el contexto de la actividad en tiempos mayores a los 5 minutos. Estas características se dividieron en dos tipos, 1) las características temporales, las que indican la ubicación temporal dentro del sueño donde se encuentra el minuto analizado; y 2) las características de tendencia, que indican la tendencia de la actividad en periodos de tiempo

**Tabla 4-3. Resultados de clasificación sueño vigilia de toda la base de datos en el punto de operación para base de datos original luego de aplicar las reglas N° 1 y N° 2 de posprocesamiento. La reglas N° 2 consiste en que todos los pares de minutos consecutivos marcados como la misma clase y que están rodeados de pares de minutos consecutivos marcados como la clase opuesta cambian su clasificación a la de sus vecinos.**

		Juicio experto usando PSG		Total	
		Sueño	Vigilia		
Resul. Clasif.	Sueño	minutos	<b>47285±200</b>	1908±179	49193
		% columna	96,5	17,7	
	Vigilia	minutos	1705±200	<b>8875±179</b>	10580
		% columna	3,5	82,3	
Total		48990	10783	<b>59773</b>	
Precisión				94,0	

**Tabla 4-4. Comparación de rendimientos de distintos clasificadores con el objetivo de medir la capacidad discriminadora de las características temporales y de contexto. En la prueba A se utilizaron todas las características descritas en la metodología (11 características). En la prueba B se utilizaron todas las características excepto las características I y II. En la prueba C se utilizaron todas las características excepto las características VIII y IX.**

	Sensibilidad[%]	Especificidad[%]	Precisión[%]
Prueba A	95,9±0,5	80,7±1,4	93,1±0,3
Prueba B	96,8±0,4	73,5±1,2	92,1±0,2
Prueba C	96,8±0,3	73,7±1,6	92,1±0,2

de 20 minutos anteriores o posteriores al minuto analizado.

Para medir el impacto de estas características en el rendimiento del clasificador se realizó una prueba en la que se medía el rendimiento del clasificador utilizando todas las características (prueba A), comparándolo con el rendimiento del clasificador entrenado con un conjunto de características que no contenía a las de carácter temporal (prueba B) y, a su vez, comparándolo con un clasificador entrenado con un conjunto de características que no contiene las características de contexto, pero sí las temporales (prueba C).

En la tabla 4-4 pueden apreciarse los resultados de estas pruebas optimizadas para la proporción de datos de la base original. Se puede apreciar que la prueba A tiene mejores resultados que las otras dos pruebas, lo que muestra la importancia de ambas características. Se puede apreciar también que las pruebas B y C tienen resultados similares, lo que indica que ambas son igualmente importantes para mejorar el rendimiento del clasificador, mejora que puede observarse en los resultados de la prueba A.

#### **4.1.2. Clasificación estados y etapas del sueño**

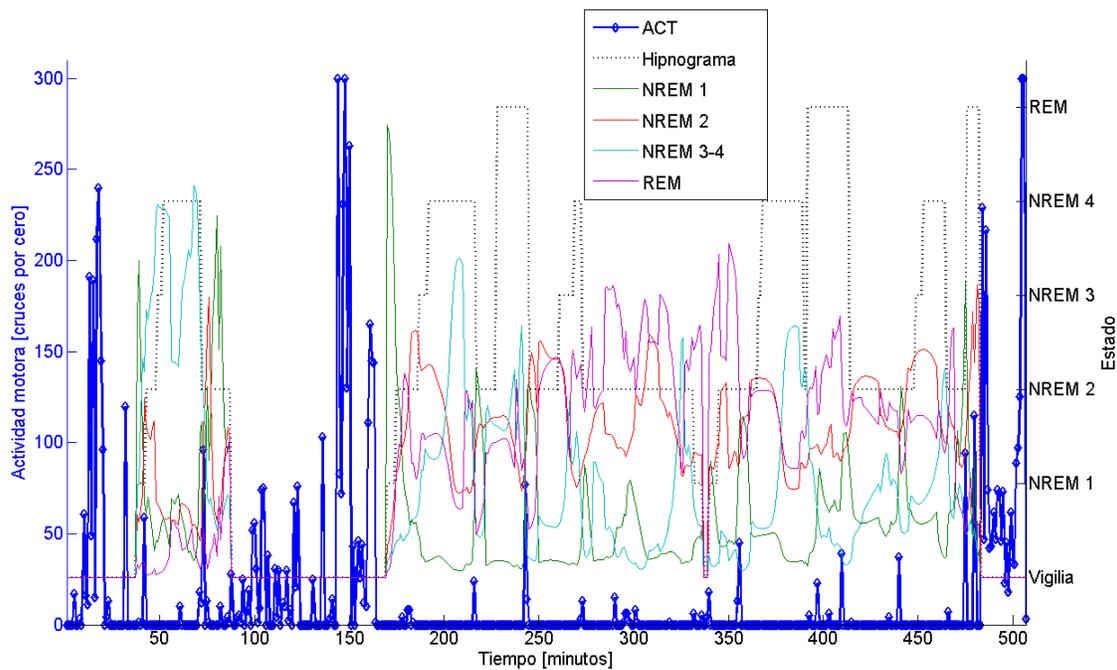
Los resultados obtenidos en la clasificación de estados y etapas del sueño no son del mismo nivel que los obtenidos para la clasificación de sueño/vigilia. Esto se debe a múltiples factores, entre los que se pueden mencionar:

- Factor fisiológico: no está claro si se puede determinar la etapa del sueño de un sujeto únicamente con la información referente a su movimiento y posición. Se pueden suponer generalidades, como que el estado NREM 1 presenta más movimiento que el NREM 2, que el NREM 3 y NREM 4 se caracterizan por una ausencia de movimiento y que la etapa REM presenta, en ciertas ocasiones, movimiento. Sin embargo no está claro si este tipo de información es suficiente para generar buenos resultados en una clasificación.
- Factor de fuente de datos: como se dijo anteriormente, no hay certeza sobre si el movimiento puede arrojar luz sobre una clasificación de etapas de sueño, sin embargo podría obtenerse más información si se contara con mediciones de movimiento en otras partes del cuerpo, como los otros miembros o la cabeza, por ejemplo. En contraposición a esto puede argumentarse que, de extraerse estas

señales, la actigrafía perdería algunas ventajas comparativas, como el bajo costo y la nula invasividad.

- Factor de frecuencia de muestreo: es posible que muchos de los eventos vinculados con el movimiento del cuerpo que distinguen una etapa de otra no sean posibles de apreciar con datos muestreados a un dato por minuto. No es posible, por ejemplo, diferenciar entre movimientos bruscos de corta duración y movimientos lentos de duración cercana al minuto.

Se puede apreciar en los resultados con base de datos equilibrada (tabla 3-4) que el clasificador logra distinguir relativamente bien las clases NREM 1 y NREM 3-4 (con un 62,1% y un 71,4% de detección respectivamente), lo que puede apreciarse en la figura 4-2, mientras que tiene dificultades en distinguir entre las clases NREM 2 y REM (55,4% y 29,7% de detección, respectivamente), ya que ningún otro par de clases genera tantos errores cruzados (minutos clasificados por el experto como una clase que el sistema clasifica como la otra y viceversa). Esto puede significar que ambas etapas son identificables como un paso intermedio entre NREM 1 y NREM 3-4 en términos de actividad motora. Estos fenómenos pueden apreciarse de mejor forma al analizar la figura 4-2, ya que se puede avistar que las clases NREM 1 y NREM 3-4 tienen altos valores



**Figura 4-2.** Salidas del clasificador ANN de estados y etapas del sueño en un registro en función de la actividad motora (azul) y del estado o etapa del sueño en que se encuentra el sujeto (negro), determinada por un experto médico. El punto de operación es el correspondiente a la base de datos balanceada. Se aprecian 4 salidas distintas, correspondientes a las cuatro clases a clasificar por el clasificador, de manera que el que la clase que tiene un valor mayor en un minuto determinado se adjudica la etiqueta de ese minuto. Además se puede apreciar que los valores de las clases NREM 2 y REM tienden a mantenerse altos en los mismos segmentos del registro.

cuando las demás clases disminuyen sus valores. Una situación contraria ocurre con los valores representantes de las clases NREM 2 y REM, los que tienden a tener sus valores más altos en los mismos segmentos del registro, superponiéndose entre ellos. Esta situación es más grave aún si se tiene en consideración que la clase con valores de menor magnitud es NREM 2, la clase con más cardinalidad en la base de datos original.

Analizando la clasificación modificada para ser fiel a la proporción original de los datos se puede apreciar que los umbrales que maximizan la precisión global sobre las 4 clases no hacen más que favorecer la clasificación de la clase NREM 2 en mayor medida, y la clase NREM 3-4. Esto se debe a que la primera es la más abundante dentro de la base de datos (tabla 3-3), por lo que, para aumentar la precisión global, es necesario aumentar la detección de ésta. Por su parte NREM 3-4 es la etapa que el clasificador identifica más confiablemente. Para corroborar esto puede calcularse el valor predictivo del sistema sobre esta clase en particular para la base de datos balanceada en el punto de operación correspondiente a la misma base de datos balanceada, donde se tiene un valor predictivo en la clase NREM 3-4 de 71,8%, y luego de cambiarse el punto de operación este valor baja sólo hasta 68,8%.

Otro punto destacable es el aumento manifiesto en el valor predictivo del clasificador en las clases NREM 1 y REM, la primera aumenta desde 60,2% a 87,0% y la segunda desde 48,2% a 59,9%. Esto se debe a que el sistema, en esta nueva configuración, presenta menos tendencia a clasificar estas clases por sobre la más numerosas (por el fenómeno mencionado anteriormente, vinculado a la cardinalidad de las clases). Esto conduce a que los minutos finalmente clasificados con una de estas etiquetas poseen altos valores de pertenencia a estas clases en la salida del clasificador ANN, es decir, son mejores candidatos para pertenecer a estas clases.

Debido a la dificultad para distinguir entre las clases REM y NREM 2, surgió la idea de fusionar estas clases de manera de generar un clasificador que considere esta problemática y mejore los resultados de detección de las demás clases. De todas maneras, hay que considerar que un clasificador de esta índole no tiene sentido práctico, debido a que, como se muestra en la tabla 3-3, las etapas REM y NREM 2 son las más comunes, por lo que unir las implicaría una clasificación que entrega poca información. De cualquier forma, y a modo de confirmación de estas suposiciones, se realizan pruebas uniando las clases REM y NREM 2.

Tabla 4-5. Resultados de clasificación de estados y etapas del sueño con base de datos balanceada en el punto de operación correspondiente a la base de datos balanceada. Se consideraron NREM 2 y REM como una única clase. La precisión se calculó sumando todas las épocas bien clasificadas, ubicadas en la diagonal de la matriz (en negritas) y dividiéndolo por el total de épocas.

		Juicio experto usando PSG			Total	Valor Pred. [%]	
		NREM 1	NREM 2 – REM	NREM 3 – 4			
Resul. Clasif.	NREM 1	<i>minutos</i>	<b>11503±105</b>	3882±302	958±144	16343	70.4
		<i>% colum</i>	67.9	22.6	5.5		
	NREM 2 – REM	<i>minutos</i>	4208±557	<b>9887±658</b>	3206±466	17301	57.2
		<i>% colum</i>	24.8	57.6	18.4		
	NREM 3 – 4	<i>minutos</i>	1246±216	3384±333	<b>13238±445</b>	17868	74.1
		<i>% colum</i>	7.4	19.7	76.1		
	<b>Total</b>		16957	17153	17402	51512	
	<b>Precisión</b>					67,2	

Para la prueba con una base de datos equilibrada en un punto de operación correspondiente a una base equilibrada (tabla 4-5) se tiene una precisión de  $67,2 \pm 0,3\%$ , con un porcentaje de clasificación correcta de NREM 1 de  $67,9 \pm 1,5\%$ , un porcentaje de clasificación correcta de NREM 2 unido a REM de  $57,6 \pm 3,8\%$  y un porcentaje de clasificación correcta de la clase que une NREM 3 con NREM 4 de  $76,1 \pm 2,3\%$ . Para las clases anteriormente mencionadas se utilizan modificaciones a los umbrales 0, 0,01 y 0,03 respectivamente, modificaciones muy cercanos a 0, lo que refleja que el punto de operación descrito por éstos es muy cercano al obtenido en el entrenamiento de la ANN.

Si se considera el punto de operación calculado para las proporciones originales entre las clases de manera de generar un clasificador que maximice su precisión (tabla 4-6), se obtiene un sistema con una precisión total de  $47,4 \pm 2,0\%$ , con un porcentaje de clasificación correcta de NREM 1 de  $11,6 \pm 3,2\%$ , un porcentaje de clasificación correcta de NREM 2 y REM de  $95,6 \pm 1,0\%$ , y un porcentaje de clasificación correcta de la clase que une NREM 3

Tabla 4-6. Resultados de clasificación de estados y etapas del sueño con base de datos balanceada en el punto operación correspondiente a la base de datos original. Se consideraron NREM 2 y REM como una única clase. La precisión se calculó sumando todas las épocas bien clasificadas, ubicadas en la diagonal de la matriz (en negritas) y dividiéndolo por el total de épocas.

		Juicio experto usando PSG			Total	Valor Pred. [%]	
		NREM 1	NREM 2 – REM	NREM 3 – 4			
Resul. Clasif.	NREM 1	<i>minutos</i>	<b>1961±105</b>	182±302	19±144	2162	90.7
		<i>% colum</i>	11,6	1,1	0,1		
	NREM 2 – REM	<i>minutos</i>	14886±557	<b>16401±658</b>	11308±466	42595	38.5
		<i>% colum</i>	87,8	95,6	65,0		
	NREM 3 – 4	<i>minutos</i>	110±216	570±333	<b>6075±445</b>	6755	89.9
		<i>% colum</i>	0,6	3,3	34,9		
	<b>Total</b>		16957	17153	17402	51512	
	<b>Precisión</b>					47,4	

Tabla 4-8. Resultados de clasificación de estados y etapas del sueño con base de datos original en el punto operación correspondiente a la base de datos original. Se consideraron NREM 2 y REM como una única clase. La precisión se calculó sumando todas las épocas bien clasificadas, ubicadas en la diagonal de la matriz (en negritas) y dividiéndolo por el total de épocas.

		Juicio experto usando PSG			Total	Valor Pred. [%]	
		NREM 1	NREM 2 – REM	NREM 3 – 4			
Resul. Clasif.	NREM 1	<i>minutos</i>	<b>174±10</b>	182±302	6±48	362	48,1
		<i>% colum</i>	11,6	1,1	0,1		
	NREM 2 – REM	<i>minutos</i>	1353±51	<b>16401±658</b>	3769±155	21523	76,2
		<i>% colum</i>	87,8	95,6	65,0		
	NREM 3 – 4	<i>minutos</i>	10±20	570±333	<b>2025±148</b>	2605	77,7
		<i>% colum</i>	0,6	3,3	34,9		
	<b>Total</b>		1537	17153	5800	24490	
	<b>Precisión</b>					75,9	

con NREM 4 de  $34,9 \pm 3.3\%$ . Sin embargo, si se consideran las cardinalidades de clases en la base de datos original (tabla 4-7) se obtiene un valor de precisión de  $75,9 \pm 0,6\%$ , conservando los valores de detección para cada una de las clases obtenidas para ese mismo punto de operación en la base de datos balanceada (tabla 4-6). Para las clases mencionadas se utilizan modificaciones a los umbrales  $-0,45$ ,  $0,3$  y  $-0,2625$  respectivamente, mostrando valores bastante lejanos de 0, lo que implica que el punto de operación correspondiente a los umbrales y el calculado en el entrenamiento son considerablemente diferentes.

Como es de esperar, la precisión aumenta sustancialmente, confirmando las suposiciones iniciales, pero se llega a un sistema sin aplicación real.

## 4.2. Comparaciones con otros estudios

Como se mencionó en la sección 1.4.1, existen numerosos estudios relacionados con el desarrollo de sistemas de clasificación de sueño y vigilia basados en información actigráfica.

De los estudios recopilados en la revisión bibliográfica, Sadeh et al. [42] es el que presenta

Tabla 4-7. Resultados de implementación de algoritmo presentado en Sadeh et al.[42] probado en la base de datos utilizada en este estudio. La precisión se calculó sumando las épocas correctamente clasificadas, ubicadas en la diagonal (en negrita) y dividiedola por el número total de épocas.

		Clasificación Experta		Total	Valor Pred. [%]	
		<i>Sueño</i>	<i>Vigilia</i>			
Resul. Clasif.	<i>Sueño</i>	<i>minutos</i>	<b>23798±139</b>	1817±106	25615	92.9
		<i>% columna</i>	97,2	33,7		
	<i>Vigilia</i>	<i>minutos</i>	697±128	<b>3574±116</b>	4271	83.7
		<i>% columna</i>	2,8	66,3		
	<b>Total</b>		24495	5391	29886	
	<b>Precisión</b>				91,6	

los mejores resultados en el apartado de precisión total (tabla 4-5). Este estudio consiste en un clasificador construido a través de Stepwise Discriminant Analysis, en el que se utilizan 4 características, el promedio de la actividad actigráfica dentro de la ventana de 11 minutos, la desviación estándar de la ventana que conforma el minuto actual y los 5 anteriores, el logaritmo natural del valor de actividad actigráfica del minuto analizado y un indicador que cuenta cuantos minutos, dentro de la ventana de 11 minutos, están sobre un umbral determinado (NMSU), indicador explicado en la sección (2.2.3.1.3).

En el mencionado estudio se utiliza un conjunto de entrenamiento compuesto por datos obtenidos de adultos sano. El sistema clasificador tiene una precisión de 92,58% en el conjunto de prueba de adultos, con una sensibilidad de 97,92% y una especificidad de 74,29%; mientras que en el conjunto de prueba de adolescentes el sistema alcanza una precisión de 91,16%, con una sensibilidad de 94,95% y una especificidad de 74,50%.

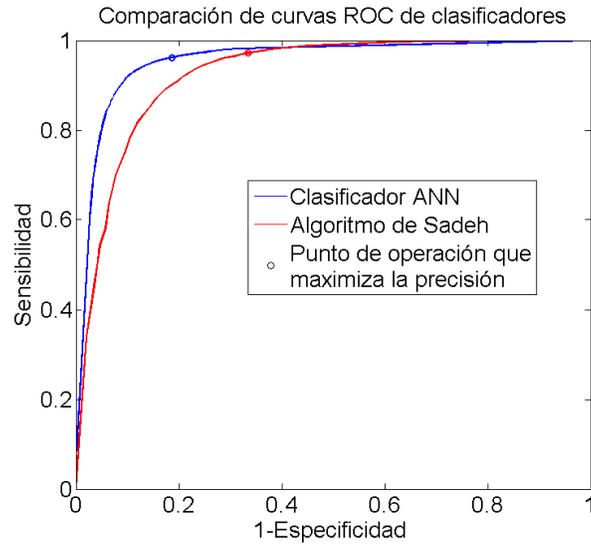
Para comparar el rendimiento del sistema clasificador presentado en Sadeh et al. [42] con el desarrollado en este estudio, se implementó el algoritmo descrito para ser probado sobre la base de datos empleada en este trabajo. Los resultados de la aplicación del algoritmo descrito en Sadeh et al. pueden verse en la tabla 4-8. Alcanzó una precisión de  $91,6 \pm 0,1\%$ , una sensibilidad de  $97,2 \pm 0,5\%$  y una especificidad de  $66,3 \pm 2,0\%$ .

Se evaluó la significancia estadística de las diferencias entre los rendimientos obtenidos por el algoritmo de clasificación de Sadeh y por el sistema expuesto en este trabajo. Se recurrió al 5x2 CV t-test [48] obteniéndose un  $\tilde{t} = 3,6004$  y un  $p - valor = 0,0078$ , por lo que se puede concluir que el sistema descrito en este trabajo tiene un rendimiento significativamente mejor.

Las diferencias entre los valores de precisión publicados por Sadeh et al. [42] y los obtenidos mediante la implementación de este algoritmo (tabla 4-8) pueden deberse a que cada prueba utiliza un punto de operación distinto del clasificador, pero los clasificadores tienen el mismo rendimiento. Este punto parece tener sustento al analizarse los valores de

**Tabla 4-9. Resultados de estudios donde se desarrollaron clasificaciones de sueño/vigilia que utilizaron sujetos adultos o adolescentes sanos basados en información actigráfica.**

<b>Estudio</b>	<b>Sujetos</b>	<b>Prec. [%]</b>	<b>Sens. [%]</b>	<b>Espec. [%]</b>
Sadeh et al.[42]	Entrenado en adultos sanos, probado en adultos sanos	92,6	<b>97,9</b>	74,3
	Entrenado en adulto sanos, probado en adolescentes sanos	91,2	95,0	74,5
Domingues et al.[41]	Entrenado en adultos sanos, probado en adultos sanos	77,8	75,6	<b>81,6</b>
Este estudio	Entrenado en adolescentes sanos, probado en adolescentes sanos	<b>93,2</b>	96,2	79,6



**Figura 4-3. Curvas ROC de clasificadores comparados. En azul, clasificador basado en ANN desarrollado en este estudio. En rojo, algoritmo desarrollado en Sadeh et al. [42]. Los círculos corresponden a los puntos de operación donde la precisión se maximiza para la base de datos original.**

sensibilidad y especificidad, ya que de ellos se desprenden las proporciones entre las cardinalidades de los conjuntos sueño y vigilia. La proporción entre sueño y vigilia para la base de datos utilizada en este trabajo es 82,0% sueño, 18,0% vigilia. Para la prueba en adultos publicada por Sadeh et al. esta proporción es 77,4% a 22,6%, mientras que en adolescentes es de 81,5% a 18,5%. Viendo la similitud entre esta última proporción y la proporción en la base de datos aquí utilizada, además de la similitud en los resultados obtenidos en estas dos bases de datos, podría decirse que el clasificador de Sadeh et al. aplicado a esta base de datos tiene un proporción entre sensibilidad y especificidad similar a la obtenida con la base de datos utilizada en la publicación (tabla 4-9).

El rendimiento de ambos clasificadores para un espectro más amplio de puntos de operación puede apreciarse en la figura 4-3.

En cuanto a la clasificación para etapas del sueño, no existen estudios publicados que intenten esta clasificación basado sólo en información actigráfica.

## **5. Conclusiones**

Las primeras dificultades enfrentadas en este proyecto se presentaron en la construcción de la base de datos. Es importante recalcar la gran importancia que tiene la accesibilidad y la confiabilidad de una base de datos al momento de intentar extraer conocimiento de ella, por lo que era de vital importancia asegurar que los registros actigráficos no fueran extremadamente ruidosos y que estuviesen sincronizados con los hipnogramas.

El mayor desafío en este estudio fue la escasez de información inherente a la señal analizada, ya que se requiere realizar clasificaciones complejas para un periodo de tiempo utilizando un único valor por periodo (la cantidad de veces que una onda de voltaje pasa por cero), teniéndose una frecuencia de muestreo muy baja en comparación con las frecuencias utilizadas para analizar otras señales fisiológicas. Esto puede explicar, en parte, el poco progreso que se ha realizado en este tema en los últimos años con datos de actigrafía.

La principal conclusión es el buen resultado obtenido con las nuevas características generadas, las características temporales y de contexto, las que permitieron incrementar sustancialmente los resultados del clasificador sueño/vigilia.

### **5.1. Clasificación sueño/vigilia**

El sistema desarrollado mostró tener un mejor rendimiento que todos los algoritmos consultados en la literatura, logrando además, mejorar la detección de vigilia sin perjudicar mayormente la detección de sueño en comparación con la gran mayoría de los clasificadores.

El buen rendimiento del sistema de clasificación es atribuible a varios factores, en primer lugar, al uso de redes neuronales artificiales (ANN), las que brindan la posibilidad de aprender de forma consistente un comportamiento requerido, pero sin llegar a sobreentrenarse, como otros tipos de algoritmos de aprendizaje, ya que su entrenamiento puede ser evaluado constantemente contra un conjunto de validación. Para un buen funcionamiento es, además de vital importancia entrenar al sistema con una base de datos balanceada, dado que, por los algoritmos de aprendizaje disponibles, si la base de datos no presenta equilibrio entre la cardinalidad de las clases a clasificar, la ANN tendería a sobreestimar la clase más representada en la base de datos para obtener un mejor rendimiento global, obteniéndose un clasificador sesgado.

Otro de los factores que pueden explicar el buen rendimiento del clasificador sueño/vigilia es la utilización de variables temporales. Este es un ámbito no explorado anteriormente y que aporta información muy valiosa para el clasificador, debido a que le permite conocer sobre el contexto temporal en el que se desarrolla determinado evento actigráfico. Es valioso saber, por ejemplo, si un evento de actividad menor dentro de un contexto sin mayor actividad se encuentra al inicio del periodo de sueño (lo que podría significar un

intento por conciliar el sueño), al final (lo que podría significar que el sujeto despertó, y está tratando de volver a dormir) o si se encuentra en medio del periodo de sueño (lo que sería, más bien, signo de actividad normal durante el sueño).

También se puede señalar la utilidad de la señal generada a partir del filtrado pasabajo hecho a la señal actigráfica, ya que ésta permite generar otras tres características, la derivada y los promedios de esta señal 20 minutos antes y 20 minutos después del punto que está siendo evaluado. Sobre todo las últimas dos características son muy importantes en el rendimiento del sistema (tabla 4-4) ya que ofrecen una medida de contexto adicional a las características temporales.

## **5.2. Clasificación estados y etapas del sueño**

Las características no utilizadas en esta clasificación fueron la media truncada de los valores de actividad desde minuto  $i - 5$  al  $i$  (característica V) y la desviación estándar de los valores de actividad de la ventana que comprende los minutos desde  $i - 5$  a  $i + 5$  (característica VI), mientras que dentro de las más útiles se encontraron el promedio de valores de actividad de la ventana que comprende los minutos desde  $i - 5$  a  $i + 5$  (característica III), el promedio de la característica III en el segmento comprendido entre  $i - 20$  e  $i - 1$  (característica VIII) y la desviación estándar de los valores de actividad de la ventana que va desde el minuto  $i$  hasta el  $i + 5$  menos la desviación estándar de los valores de actividad de la ventana que va desde el minuto  $i - 5$  hasta el  $i$  (característica VII).

Esto parece indicar que en la clasificación de estados y etapas del sueño hay características que se superponen entre ellas, como el caso de la V con la III y la VIII, y más notablemente en caso de la VI con la VII. Es decir, como la capacidad para discriminar de cada una de las variables se redujo en comparación con la clasificación sueño/vigilia (ya que el rendimiento del clasificador en su totalidad disminuyó), la inclusión de variables que entregan información similar parece no ser necesaria, es decir, la información útil que brindan no es estadísticamente significativa.

Al evaluar los resultados obtenidos por el clasificador en el apartado de precisión es claro que el rendimiento no es bueno. Esto se hace patente en el cambio en los índices de clasificación de cada clase cuando se pasa del punto de operación calculado para base de datos balanceada, al punto de operación para base de datos con la proporción original. Se puede observar, en el primer caso, que las clases NREM 1 y NREM 3-4 logran buenos índices (cercaos al 70%), pero las clases NREM 2 y REM obtienen resultados bajos, sumado además a una pobre capacidad de discriminar entre estas clases. Al aplicar los umbrales correspondientes al punto de operación calculado para la proporción original entre las clases, el sistema relaja las condiciones para emitir una clasificación NREM 2 debido a la gran cardinalidad de la clase NREM 2. Esto provoca que el sistema tienda a clasificar esta clase, pero en el proceso se daña severamente la clasificación de las restantes clases. La clase REM parece ser la más afectada debido a que el sistema parece confundir

las clases NREM 2 y REM. Este fenómeno se provoca por la pobre calidad discriminatoria que tiene el sistema, como queda claro cuando se le compara con el clasificador sueño/vigilia, ya que este cuando se aplican los umbrales calculados para el punto de operación correspondiente a la proporción original, el índice de clasificación de la clase con menos cardinalidad (en este caso vigilia) no sufre un descenso tan dramático. Esto puede ser causa de que ese sistema de clasificación es más robusto, lo que puede deberse tanto a la elección de características, a un buen o mal uso del algoritmo de clasificación o al hecho de que la información disponible en los registros actigráficos no es suficiente para discriminar las clases anteriormente nombradas.

En lo concerniente a NREM 2, a través del experimento mostrado en la discusión se fusionaron las clases NREM 2 y REM (tablas 4-2, 4-3 y 4-4). Se pudo apreciar la similitud entre estas dos clases, donde muchas épocas clasificadas por el experto como NREM 2 son clasificadas por el sistema como REM y viceversa, además de obtenerse altos valores en las desviaciones estándar en estos ítems. El experimento mencionado no tiene aplicación práctica, como se explicó anteriormente, pero puede servir como punto de partida para diseñar una metodología que detecte este problema y lo evite.

### **5.3.Unificación de clasificadores**

Los resultados del sistema completo, que considera los dos clasificadores trabajando en conjunto, pueden ser considerados mediocres. Esto es explicable, principalmente, por el bajo rendimiento del segundo clasificador. Sin embargo algo destacable es que los minutos marcados como vigilia por el experto y que no son correctamente clasificados por el primer clasificador, tienden a ser posteriormente clasificados en NREM 2 (tabla 3-7). Esto contraviene la intuición basada en la fisiología que adjudica a la etapa NREM 1 un papel de transición entre la vigilia y el sueño, mientras que la etapa NREM 2 es considerada como una etapa propiamente de sueño. Esto es explicable, nuevamente, por el pobre desempeño del segundo clasificador, ya que tiende a sobreestimar la clase NREM 2, pasando a llevar con ello, las clasificaciones de las clases NREM 1 y REM principalmente. También es posible que esta mala clasificación se deba al hecho de que estos datos no entregan la información suficiente para poder distinguir entre vigilia y NREM 2.

### **5.4.Trabajo futuro**

Existen algunas áreas en las que existen posibles mejoras al sistema expuesto así como posible usos alternativos de los resultados obtenidos:

- Podría realizarse un análisis más exhaustivo de posibles características estadísticas u otros indicadores para el segundo clasificador, así como del poder de éstos para discriminar las diferentes clases, de manera de intentar mejorar el rendimiento del sistema.
- Podría generarse un análisis de acuerdo al ancho de ventana móvil utilizada para extraer características, lo que podría entregar información relacionada con la

fisiología del sueño en el sentido de saber cuánto afecta la actividad ocurrida X minutos antes o X minutos después al estado en que se encuentra el sujeto en el momento actual.

- Podría modificarse la manera de representar las épocas en las pruebas de clasificación, de manera de permitir aplicar reglas de posprocesamiento a las salidas de la ANN.
- Podría optimizarse el clasificador sueño/vigilia para priorizar la detección de episodios de vigilia durante la noche, buscando determinar las veces que el sujeto se despierta.

También se podría trabajar con datos con más información que los disponibles:

- Un posible curso de acción para el mejoramiento del sistema sería contar con una base de datos de registros actigráficos muestreados a una mayor frecuencia. Esto, eventualmente, podría permitir perfeccionar el clasificador de sueño/vigilia y lograr generar alguna característica que ayude al sistema a discernir entre las clases NREM 2 y REM, lo que mejoraría considerablemente el segundo clasificador.
- Otro posible camino para mejorar el rendimiento del sistema sería incluir información proveniente de otra señal fisiológica. Hay estudios científicos que utilizan la señal actigráfica en conjunto con EKG [52], EOG [53], algunas derivaciones de EEG [54], entre otros. Esto podría servir para perfeccionar el clasificador sueño/vigilia, pero además sería muy útil al entregar nueva información, distinta a la actividad motora. Esta nueva información, según la fisiología del sueño, podría estar más ligada a los procesos que diferencian a las etapas NREM 2 y REM, como lo son los movimientos oculares, potenciales de activación muscular, sincronización de frecuencia cardíaca con algún patrón determinado, etc.

Como último aspecto a trabajar está la implementación de este algoritmo con fines médicos. Las aplicaciones de estas técnicas son diversas:

- Debido a su reducido costo y su mínima invasividad, este sistema de clasificación puede ser implementado para utilizarse como examen exploratorio para grandes universos de personas (*screening*), de manera de diagnosticar con niveles relativamente altos de certeza diversas enfermedades, para posteriormente ratificar o descartar estos diagnósticos con otras metodologías.
- Hacer seguimientos por periodos de tiempo prolongados sobre la actividad motora de personas con estados de lenta evolución, tales como embarazos o convalecencias de cirugías importantes, brindando una posibilidad de analizar el comportamiento en largos periodos de tiempo. Con este procedimiento los sujetos no deben descargar la información del dispositivo regularmente ya que necesita de una capacidad de almacenamiento mínima en comparación con las tecnologías actuales.

## Bibliografía

- [1] M. Carskadon and W. Dement, “Normal human sleep: an overview.,” *Princ. Pract. sleep Med.*, vol. 18, pp. 16–25, 1994.
- [2] L. Causa, C. Held, P. Estevez, and C. Perez, “Detección Automática y caracterización de patrones de husos sigma en registros polisonmográficos de niños,” Tesis para optar a los grado de magister en ciencias de la ingeniería mención ingeniería biomédica y al título de ingeniero civil electricista, Departamento de Ingeniería Eléctrica, Universidad de Chile, Enero 2012.
- [3] G. Tononi and C. Cirelli, “Sleep function and synaptic homeostasis.,” *Sleep Med. Rev.*, vol. 10, no. 1, pp. 49–62, Feb. 2006.
- [4] A. T. M. Van de Water, A. Holmes, and D. a Hurley, “Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography-- a systematic review.,” *J. Sleep Res.*, vol. 20, no. 1 Pt 2, pp. 183–200, Mar. 2011.
- [5] C. A. Kushida, M. R. Littner, T. Morgenthaler, C. A. Alessi, D. Bailey, J. Coleman, L. Friedman, M. Hirshkowitz, S. Kapen, M. Kramer, T. Lee-Chiong, D. L. Loubé, J. Owens, J. P. Pancer, and M. Wise, “Practice parameters for the indications for polysomnography and related procedures: an update for 2005.,” *Sleep*, vol. 28, pp. 499–521, 2005.
- [6] M. H. Bonnet, D. Carley, M. Carskadon, P. Easton, C. Guilleminault, R. Harper, B. Hayes, M. Hirshkowitz, P. Ktonas, S. Keenan, M. Pressman, T. Roehrs, J. Smith, J. Walsh, S. Weber, and P. Westbrook, “EEG arousals: scoring rules and examples,” *Sleep*, vol. 15, pp. 173–184, 1992.
- [7] W. W. Tryon, “Issues of validity in actigraphic sleep assessment.,” *Sleep*, vol. 27, no. 1, pp. 158–165, 2004.
- [8] R. D. Ogilvie, R. T. Wilkinson, and S. Allison, “The detection of sleep onset: behavioral, physiological, and subjective convergence.,” *Sleep*, vol. 12, pp. 458–474, 1989.
- [9] J. Heiss, C. M. Held, P. A. Estévez, and C. A. Pérez, “Sistema Neurodifuso aplicado a la Clasificación de Estados de Sueño,” Tesis para optar al grado de magíster en ciencias de la ingeniería, mención ingeniería biomédica y al título de ingeniero civil electricista, Departamento de Ingeniería Eléctrica, Universidad de Chile, Enero 2003.
- [10] A. Rechtschaffen and A. Kales, “A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Los Angeles: UCLA Brain Information Service, Brain Research Institute,” 1968.

- [11] L. Novelli, R. Ferri, and O. Bruni, "Sleep classification according to AASM and Rechtschaffen and Kales: Effects on sleep scoring parameters of children and adolescents," *J. Sleep Res.*, vol. 19, pp. 238–247, 2010.
- [12] S. L. Himanen and J. Hasan, "Limitations of Rechtschaffen and Kales," *Sleep Medicine Reviews*, vol. 4. pp. 149–167, 2000.
- [13] C. Iber, S. Ancoli-Israel, A. Chesson, and S. Quan, "The AASM manual for the scoring of sleep and associated events American Academy of Sleep Medicine," Westchester, IL: American Academy of Sleep Medicine, pp. 48-9, 2007.
- [14] J. M. Siegel, "REM sleep: A biological and psychological paradox," *Sleep Medicine Reviews*, vol. 15, pp. 139–142, 2011.
- [15] S. Diekelmann and J. Born, "The memory function of sleep.," *Nat. Rev. Neurosci.*, vol. 11, pp. 114–126, 2010.
- [16] A. D. Grosmark, K. Mizuseki, E. Pastalkova, K. Diba, and G. Buzsáki, "REM Sleep Reorganizes Hippocampal Excitability," *Neuron*, vol. 75, pp. 1001–1007, 2012.
- [17] I. Feinberg and T. Floyd, "Systematic trends across the night in human sleep cycles," *Psychophysiology*, vol. 16, pp. 283-291, 1979.
- [18] American Academy of Sleep Medicine, "International classification of sleep disorders, 2nd ed.: diagnostic and coding manual" American Academy of Sleep Medicine, Westchester, Illinois, USA, 2005.
- [19] M. M. Grigg-Damberger, "The AASM scoring manual: a critical appraisal," *Current Opinion in Pulmonary Medicine*, vol. 15. pp. 540–549, 2009.
- [20] N. Cellini, M. P. Buman, E. a McDevitt, A. a Ricker, and S. C. Mednick, "Direct comparison of two actigraphy devices with polysomnographically recorded naps in healthy young adults.," *Chronobiol. Int.*, vol. 30, pp. 691–8, 2013.
- [21] G. Jean-Louis, D. F. Kripke, W. J. Mason, J. A. Elliott, and S. D. Youngstedt, "Sleep estimation from wrist movement quantified by different actigraphic modalities," *J. Neurosci. Methods*, vol. 105, pp. 185–191, 2001.
- [22] A. Sadeh, "The role and validity of actigraphy in sleep medicine: An update," *Sleep Medicine Reviews*, vol. 15. pp. 259–267, 2011.
- [23] C. M. Held, L. Causa, F. Jaillet, R. Chamorro, M. Garrido, C. Algarin, and P. Peirano, "Automated detection of apnea/hypopnea events in healthy children polysomnograms: Preliminary results," in *Proceedings of the 35<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Osaka, Japan, pp. 5373–5376, July 3-7, 2013.

- [24] T. Penzel, J. W. Kantelhardt, L. Grote, J.-H. Peter, and A. Bunde, "Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea.," *Biomedical Engineering, IEEE Transactions*, vol. 50, no 10, p. 1143-1151, 2003
- [25] J. Hayano, E. Watanabe, Y. Saito, F. Sasaki, K. Fujimoto, T. Nomiyama, K. Kawai, I. Kodama, and H. Sakakibara, "Screening for obstructive sleep apnea by cyclic variation of heart rate," *Circ. Arrhythmia Electrophysiol.*, vol. 4, pp. 64–72, 2011.
- [26] A. H. Khandoker, J. Gubbi, and M. Palaniswami, "Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, pp. 1057–1067, 2009.
- [27] R. B. Shouldice, L. M. O'Brien, C. O'Brien, P. de Chazal, D. Gozal, and C. Heneghan, "Detection of obstructive sleep apnea in pediatric subjects using surface lead electrocardiogram features.," *Sleep*, vol. 27, pp. 784–792, 2004.
- [28] A. Garde, W. Karlen, P. Dehkordi, D. Wensley, J. M. Ansermino, and G. A. Dumont, "Oxygen saturation in children with and without obstructive sleep apnea using the phone-oximeter," in *Proceedings of the 35<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Osaka, Japan, pp. 2531–2534, July 3-7, 2013.
- [29] E. Karci, Y. S. Dogrusoz, and T. Ciloglu, "Detection of post apnea sounds and apnea periods from sleep sounds," in *Proceedings of the 33<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Boston, Massachusetts, USA, pp. 6075–6078, August 30-September 3, 2011.
- [30] A. Yadollahi and Z. Moussavi, "Apnea detection by acoustical means," in *Proceedings of the 28<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, New York City, New York, USA, pp. 4623–4626, August 31-September 3, 2006.
- [31] B. F. Boeve, "REM sleep behavior disorder: Updated review of the core features, the REM sleep behavior disorder-neurodegenerative disease association, evolving concepts, controversies, and future directions," *Annals of the New York Academy of Sciences*, vol. 1184. pp. 15–54, 2010.
- [32] H. S. Driver and S. R. Taylor, "Exercise and sleep," *Sleep Medicine Reviews*, vol. 4. pp. 387–402, 2000.
- [33] J. M. Siegel, "The REM sleep-memory consolidation hypothesis.," *Science*, vol. 294, pp. 1058–1063, 2001.

- [34] M. Adnane, Z. Jiang, and S. Choi, "Development of QRS detection algorithm designed for wearable cardiorespiratory system," *Comput. Methods Programs Biomed.*, vol. 93, pp. 20–31, 2009.
- [35] E. Tagliazucchi, F. von Wegner, A. Morzelewski, S. Borisov, K. Jahnke, and H. Laufs, "Automatic sleep staging using fMRI functional connectivity data," *Neuroimage*, vol. 63, pp. 63–72, 2012.
- [36] G. Becq, S. Charbonnier, F. Chapotot, A. Buguet, L. Bourdon, and P. Baconnier, "Comparison Between Five Classifiers for Automatic Scoring of Human Sleep Recordings," in *Classification and Clustering for Knowledge Discovery*, vol. 127, pp. 113–127, 2005.
- [37] L. Zoubek, S. Charbonnier, S. Leseq, A. Buguet, and F. Chapotot, "Feature selection for sleep/wake stages classification using data driven methods," *Biomed. Signal Process. Control*, vol. 2, pp. 171–179, 2007.
- [38] B. Sen, M. Peker, A. Cavuşoğlu, and F. V Celebi, "A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms.," *J. Med. Syst.*, vol. 38, p. 18, 2014.
- [39] A. V. I. Sadeh and E. P. B. Hospitavbrown, "Activity-Based Assessment of Sleep Wake Patterns During the 1 st Year o P- Life," vol. 337, pp. 329–337, 1995.
- [40] W. W. Tryon, "Nocturnal activity and sleep assessment," *Clinical Psychology Review*, vol. 16, pp. 197–213, 1996.
- [41] A. Domingues, T. Paiva, and J. M. Sanches, "Sleep and wakefulness state detection in nocturnal actigraphy based on movement information," *IEEE Trans. Biomed. Eng.*, vol. 61, pp. 426–434, 2014.
- [42] A. Sadeh, K. M. Sharkey, and M. A. Carskadon, "Activity-based sleep-wake identification: an empirical test of methodological issues.," *Sleep*, vol. 17, pp. 201–207, 1994.
- [43] M. H. Chase, "Motor control during sleep and wakefulness: clarifying controversies and resolving paradoxes.," *Sleep Med. Rev.*, vol. 17, no. 4, pp. 299–312, Aug. 2013.
- [44] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas. "Supervised machine learning: A review of classification techniques.," *Informatica*, vol. 31, pp. 249-268, 2007.
- [45] A. Sadeh, P. Lavie, A. Scher, E. Tirosh, and R. Epstein, "Actigraphic home-monitoring sleep-disturbed and control infants and young children: a new method for pediatric assessment of sleep-wake patterns," *Pediatrics*, vol. 87, pp. 494-499, 1991.

- [46] J. Tilmanne, J. Urbain, M. V. Kothare, A. Vande Wouwer, and S. V. Kothare, “Algorithms for sleep-wake identification using actigraphy: a comparative study and new results,” *J. Sleep Res.*, vol. 18, pp. 85–98, 2009.
- [47] R. J. Cole, D. F. Kripke, W. Gruen, D. J. Mullaney, and J. C. Gillin, “Automatic sleep/wake identification from wrist activity.,” *Sleep*, vol. 15, pp. 461–469, 1992.
- [48] M. Muhlbaier , A. Topalis and R. Polikar "Incremental learning from unbalanced data", *IEEE Int. Joint Conf. on Neural Networks*, Budapest, Hungary, vol. 2, pp.1057 -1062, July 25-29, 2004
- [49] T. Dietterich, “Approximate Statistical Test For Comparing Supervised Classification Learning Algorithms,” *Neural Comput.*, vol. 10, pp. 1895-1923, 1998.
- [50] C. Croux and C. Dehon, “Robust estimation of location and scale,” *Encycl. Environmetrics*, 5, 2013.
- [51] G. Orellana, C. M. Held, P. Estevez, C. Perez, S. Reyes, C. Algarín, and P. Peirano, “A balanced sleep/wakefulness classification method based on actigraphic data in adolescents,” in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Chicago, Illinois, USA, August 26-30, 2014.
- [52] W. Karlen, C. Mattiussi, and D. Floreano, “Sleep and wake classification with ECG and respiratory effort signals,” in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, pp. 71–78, 2009.
- [53] J. Virkkala, J. Hasan, A. Värri, S. L. Himanen, and K. Müller, “Automatic sleep stage classification using two-channel electro-oculography,” *J. Neurosci. Methods*, vol. 166, pp. 109–115, 2007.
- [54] A. Flexer, G. Gruber, and G. Dorffner, “A reliable probabilistic sleep stager based on a single EEG signal,” *Artif. Intell. Med.*, vol. 33, pp. 199–207, 2005.