



## **Trabajo Práctico Final** **Análisis y Procesamiento de Señales**

*“Clasificación automática de las etapas del sueño  
utilizando señales EEG y EOG”*

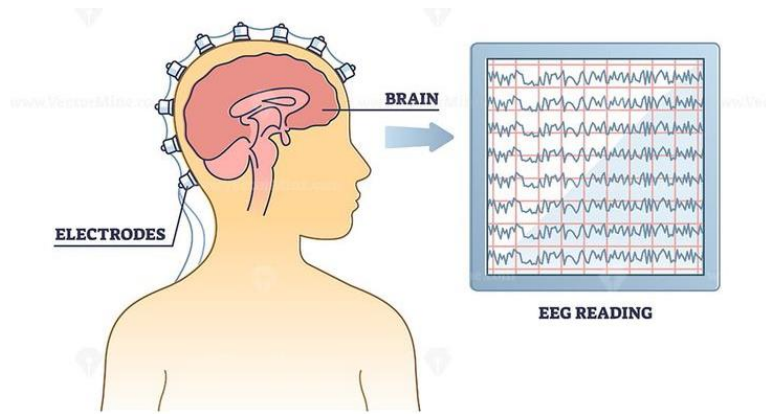
Universidad Nacional de San Martín  
Ingeniería Biomedica

Docentes: Mariano Llamedo Soria, Francisco Hernan Ledesma, David Ezequiel  
Benoit

Camila B. Smith

[cbsmith@estudiantes.unsam.edu.ar](mailto:cbsmith@estudiantes.unsam.edu.ar)

Julio 2025



# Clasificación automática de las etapas del sueño utilizando señales EEG y EOG

Universidad Nacional de San Martín, ECyT

Camila Belen Smith | Analisis y Procesamiento de Señales | Julio 2025

# 1. Introducción

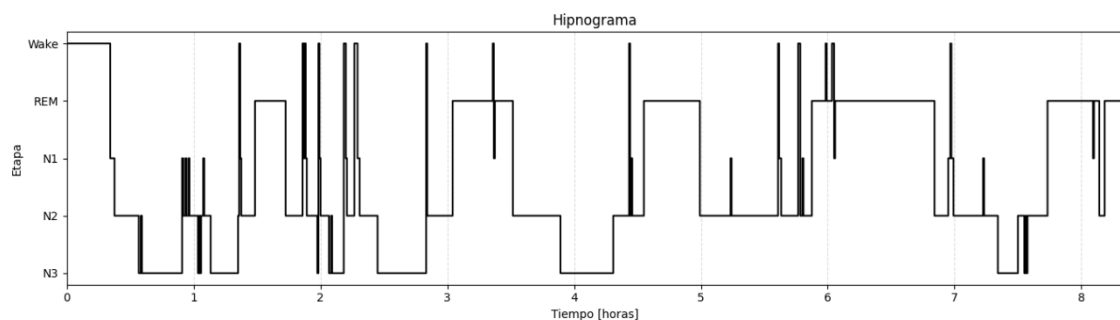
El sueño es un estado fisiológico activo que cumple funciones esenciales en la restauración del cuerpo, la consolidación de la memoria y el desarrollo cerebral. A lo largo de la noche, se alternan ciclos de aproximadamente 90 a 120 minutos entre dos fases principales: el sueño sin movimientos oculares rápidos (NREM) y el sueño con movimientos oculares rápidos (REM) [Brinkman et al., 2023].

La fase NREM se divide en tres etapas progresivas. En N<sub>1</sub>, que representa el inicio del sueño, predominan ondas theta de baja amplitud. En N<sub>2</sub>, emergen eventos característicos como los husos de sueño y los complejos K, que marcan una mayor profundidad. En N<sub>3</sub>, también llamada “sueño de ondas lentas”, se observan ondas delta de alta amplitud, asociadas con procesos de restauración física [Brinkman et al., 2023]. En contraste, el sueño REM presenta un EEG desincronizado, con ondas theta y alfa, movimientos oculares rápidos y parálisis muscular casi completa. Es la fase donde ocurren la mayoría de los sueños vívidos, y cumple un rol clave en el aprendizaje y la memoria emocional. La vigilia, por su parte, se caracteriza por actividad alfa y beta, con mayor tono muscular y mayor frecuencia en el EEG.

La clasificación manual de estas etapas, basada en la inspección visual de señales EEG, EOG y EMG en registros polisomnográficos, es el método estándar en medicina del sueño. No obstante, es un proceso laborioso, requiere formación especializada y presenta una variabilidad interevaluador significativa [Guillot et al., 2020]. En este contexto, el desarrollo de sistemas automáticos basados en procesamiento de señales y aprendizaje automático ofrece una alternativa eficiente y objetiva.

El objetivo de este trabajo es diseñar un sistema automático para la clasificación de etapas del sueño a partir de señales EEG y EOG, mediante la extracción de características fisiológicamente relevantes y la implementación de modelos supervisados.

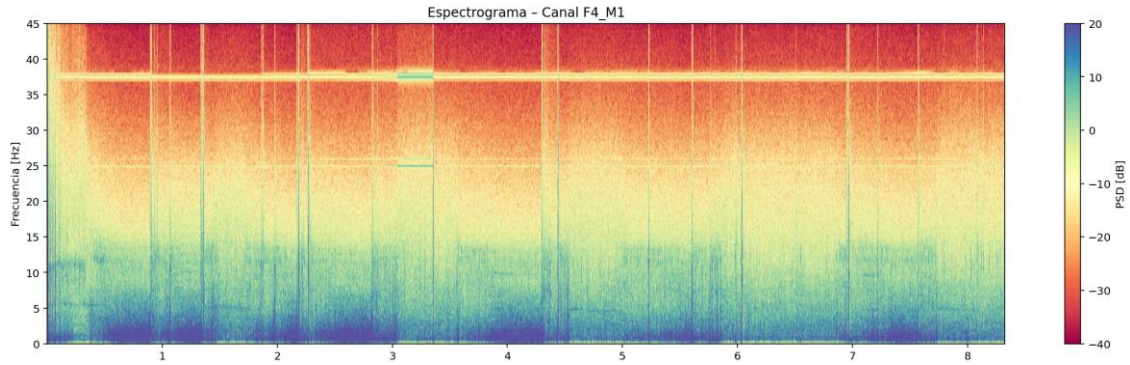
Todo el desarrollo fue implementado en Python y está disponible en el siguiente repositorio de GitHub: [https://github.com/camilabsmith/Final\\_APS](https://github.com/camilabsmith/Final_APS).



**Figura 1.** Hipnograma anotado según criterios AASM para un sujeto del dataset DOD-H. Se observa la alternancia típica entre fases NREM (N<sub>1</sub>, N<sub>2</sub>, N<sub>3</sub>) y REM a lo largo de la noche.

Para ilustrar la dinámica temporal y espectral del sueño humano, se presentan en la Figura 1 y la Figura 2 dos visualizaciones complementarias de un sujeto del dataset. La Figura 1 muestra el hipnograma, que representa la sucesión de etapas del sueño a lo largo de la noche, permitiendo observar la estructura cíclica entre NREM y REM. La Figura 2 muestra

el espectrograma completo calculado sobre el canal EEG F4-M1, donde se destaca la variación en la distribución espectral a lo largo del tiempo: durante las etapas N3 se observa un aumento sostenido de potencia en frecuencias delta ( $<4$  Hz), mientras que en REM predomina la actividad en bandas rápidas como alfa y beta (8–30 Hz), en línea con la desincronización cortical característica de esa etapa.



**Figura 2.** Espectrograma calculado sobre el canal EEG F4-M1 del mismo sujeto mostrado en la Figura 1. Se visualiza la distribución de potencia en frecuencia a lo largo del tiempo. La actividad delta ( $<4$  Hz) predomina durante el sueño profundo (N3), mientras que REM se caracteriza por actividad en bandas rápidas como alfa y beta (8–30 Hz).

El espectrograma mostrado fue obtenido mediante la Transformada de Fourier de Tiempo Corto (STFT), que permite analizar señales no estacionarias como las EEG. A diferencia de la transformada de Fourier tradicional —adecuada solo para señales estacionarias—, la STFT descompone la señal en componentes frecuenciales localizadas en el tiempo, utilizando una ventana deslizante. En este caso, se utilizó una ventana de 5 segundos con 50 % de solapamiento, lo que genera una representación bidimensional (tiempo vs. frecuencia) del contenido espectral.

Este tipo de análisis implica un compromiso entre resolución temporal y frecuencial: una ventana más corta mejora la precisión temporal pero empeora la resolución en frecuencia, mientras que una ventana más larga permite observar con más detalle la distribución espectral, a costa de perder definición temporal. Esta limitación está asociada al principio de incertidumbre, que impide mejorar simultáneamente ambas resoluciones. Por tanto, la elección del tamaño de ventana debe balancear estos factores según el objetivo del análisis. En este trabajo, el objetivo era visualizar cómo varían las bandas de frecuencia predominantes a lo largo del sueño, lo cual justifica el uso de una ventana intermedia (5 segundos).

## 2. Datos

En este trabajo se utilizó el conjunto de datos DOD-H (Dreem Open Dataset - Healthy), publicado por Guillot et al. (2020). Este dataset fue recolectado por la Unidad de Fatiga y Vigilancia del Instituto de Investigación Biomédica de las Fuerzas Armadas Francesas (IRBA), y contiene registros polisomnográficos de 25 voluntarios sanos, de entre 18 y 65 años, sin trastornos del sueño diagnosticados. Las señales fueron adquiridas con dispositivos Siesta PSG (Compumedics), utilizando una frecuencia de muestreo de 250 Hz.

La frecuencia de muestreo ( $f_s$ ) indica cuántas muestras por segundo se toman de una señal analógica para convertirla a formato digital. En este caso,  $f_s = 250$  Hz, lo que significa que cada segundo de señal contiene 250 valores discretos. Para evitar aliasing, es decir, la superposición de componentes de alta frecuencia que se pliegan sobre el espectro observable, la teoría del muestreo establece que la frecuencia de muestreo debe ser al menos el doble de la máxima frecuencia presente en la señal (criterio de Nyquist). Como el contenido espectral relevante en EEG típicamente se encuentra por debajo de los 45 Hz, esta condición se cumple ampliamente: 250 Hz permite representar adecuadamente señales con componentes de hasta 125 Hz, lo cual es más que suficiente para capturar las bandas fisiológicas del sueño (delta, theta, alfa, beta y gamma baja).

Cada registro incluye múltiples señales fisiológicas: EEG, EOG, EMG y ECG. En particular, el EEG fue registrado mediante 12 derivaciones distribuidas en regiones frontales, centrales y occipitales del cuero cabelludo, utilizando montajes referenciales y bipolares según el sistema internacional 10–20. Los canales disponibles son: C<sub>3</sub>/M<sub>2</sub>, F<sub>4</sub>/M<sub>1</sub>, F<sub>3</sub>/F<sub>4</sub>, F<sub>3</sub>/M<sub>2</sub>, F<sub>4</sub>/O<sub>2</sub>, F<sub>3</sub>/O<sub>1</sub>, FP<sub>1</sub>/F<sub>3</sub>, FP<sub>1</sub>/M<sub>2</sub>, FP<sub>1</sub>/O<sub>1</sub>, FP<sub>2</sub>/F<sub>4</sub>, FP<sub>2</sub>/M<sub>1</sub> y FP<sub>2</sub>/O<sub>2</sub>. Además, se registraron señales EOG provenientes del ojo izquierdo y derecho.

Para este trabajo se utilizaron exclusivamente las señales EEG y EOG, ya que son las más directamente vinculadas con la arquitectura del sueño. Esta decisión permitió reducir la complejidad computacional sin sacrificar la información necesaria para la clasificación de etapas.

Cada noche fue anotada por cinco técnicos especializados en sueño de tres centros diferentes, siguiendo las reglas de la American Academy of Sleep Medicine (AASM), versión 2.5 (2018). Las etapas del sueño fueron clasificadas en cinco categorías: vigilia, N<sub>1</sub>, N<sub>2</sub>, N<sub>3</sub> y REM, con anotaciones realizadas sobre épocas de 30 segundos. Para cada época se cuenta con cinco anotaciones independientes (una por técnico), y una etiqueta de consenso por mayoría, que fue utilizada como referencia en este estudio.

El conjunto completo de señales y anotaciones se encuentra disponible públicamente en el repositorio de código abierto: <https://github.com/Dreem-Organization/dreem-learning-open>

### 3. Metodología

#### 3.0 Estructura general del código

El procesamiento de los datos y la extracción de características fueron implementados en Python, organizando el flujo en distintos scripts reutilizables:

- `utils_info.py`: contiene funciones comunes como el filtrado, la segmentación de señales y la extracción de características tanto de EEG como de EOG.
- `extraccion_feat.py`: aplica los procedimientos definidos en `utils_info.py` sobre los archivos .h5 del dataset, generando archivos .csv con las features calculadas por ventana de 30 segundos.

- `entrenamiento_modelos.py`: entrena los modelos clasificadores sobre los .csv generados, calcula métricas de evaluación, genera figuras (matrices de confusión, hipnogramas) y guarda resultados por fold.
- `Individual.py` : permite realizar un análisis exploratorio visual e interpretativo de un sujeto específico. Incluye el diseño y verificación de los filtros Butterworth aplicados a EEG y EOG, la carga y preprocesamiento de señales desde un archivo .h5, la visualización del hipnograma, y el trazado simultáneo de señales EEG y EOG para una época representativa de cada etapa del sueño. Además, implementa gráficos de densidad espectral de potencia (PSD) para comparar etapas, el cálculo del espectrograma a lo largo de la noche, y la detección e inspección de complejos K utilizando descomposición wavelet y energía de Teager. Este script fue clave para validar visualmente el comportamiento de las señales y evaluar la calidad del preprocesamiento y la extracción de características.

### 3.1 Filtrado

El filtrado de las señales EEG y EOG se realiza dentro del script `extraccion_feat.py`, utilizando la función `filtrado()` definida en `utils_info.py`. Para cada canal, se aplica un filtro digital IIR (Infinite Impulse Response) del tipo Butterworth, diseñado mediante la función `iirdesign` del módulo `scipy.signal` y aplicado con `sosfiltfilt`. Esta función implementa un filtrado bidireccional (hacia adelante y hacia atrás), lo cual garantiza una fase lineal perfecta (sin desfase temporal) a pesar de tratarse de un filtro IIR.

El diseño Butterworth se caracteriza por tener una respuesta en frecuencia suave y monótonica, sin ondulaciones (ripple) en la banda pasante ni en la de atenuación. La magnitud de su respuesta en frecuencia está definida por la expresión:

$$|H(\Omega)| = \frac{1}{\sqrt{1 + \left(\frac{\Omega}{\Omega_c}\right)^{2N}}}$$

donde  $\Omega$  es la frecuencia angular,  $\Omega_c$  la frecuencia de corte y  $N$  el orden del filtro. A la frecuencia  $\Omega = \Omega_c$ , el módulo de la respuesta se reduce a  $\frac{1}{\sqrt{2}}$ , es decir,  $-3.01$  dB, lo que define el punto de corte. Esta propiedad garantiza una transición suave entre la banda pasante y la de atenuación, con una forma de respuesta particularmente adecuada para preservar la morfología de señales fisiológicas como el EEG.

Los filtros IIR tienen respuesta al impulso infinita, ya que su salida depende tanto de entradas actuales y pasadas, como de la salida previa. Esta retroalimentación permite lograr transiciones abruptas con órdenes más bajos que los filtros FIR, pero los vuelve más sensibles a errores numéricos. Para mitigar estos efectos, se implementó el filtro en su forma factorizada en secciones de segundo orden (SOS), que mejora la estabilidad numérica frente a implementaciones directas. Esta estructura divide el filtro en bloques más simples, reduciendo la sensibilidad a errores y permitiendo una ejecución más robusta en sistemas digitales.

Dado que el filtrado fue aplicado de forma bidireccional (mediante `sosfiltfilt`), las ganancias en dB se duplican. Es decir, el ripple en la banda pasante y la atenuación en la banda de rechazo se ven amplificados. Para obtener una especificación final de, por ejemplo, 1 dB de ripple y 40 dB de atenuación, se debe diseñar el filtro base con 0.5 dB de ripple y 20 dB de atenuación. Este ajuste fue incorporado en el diseño utilizado en este trabajo.

Matemáticamente, `sosfiltfilt` implementa el filtro hacia adelante y luego hacia atrás, obteniendo una señal final equivalente a aplicar el filtro original dos veces. Si la función de transferencia del filtro es  $H(z)$  la salida del sistema es:

$$H_{bidireccional}(z) = H(z) \cdot H(z^{-1})$$

$$\angle H_{bidir}(e^{j\omega}) = \angle H(e^{j\omega}) + \angle H(e^{-j\omega}) = 0$$

Al mismo tiempo, la magnitud se eleva al cuadrado:

$$|H_{bidir}(e^{j\omega})| = |H(e^{j\omega})|^2$$

lo que implica una duplicación en decibeles:

$$|H_{bidir}|_{db} = 2 \cdot |H|_{db}$$

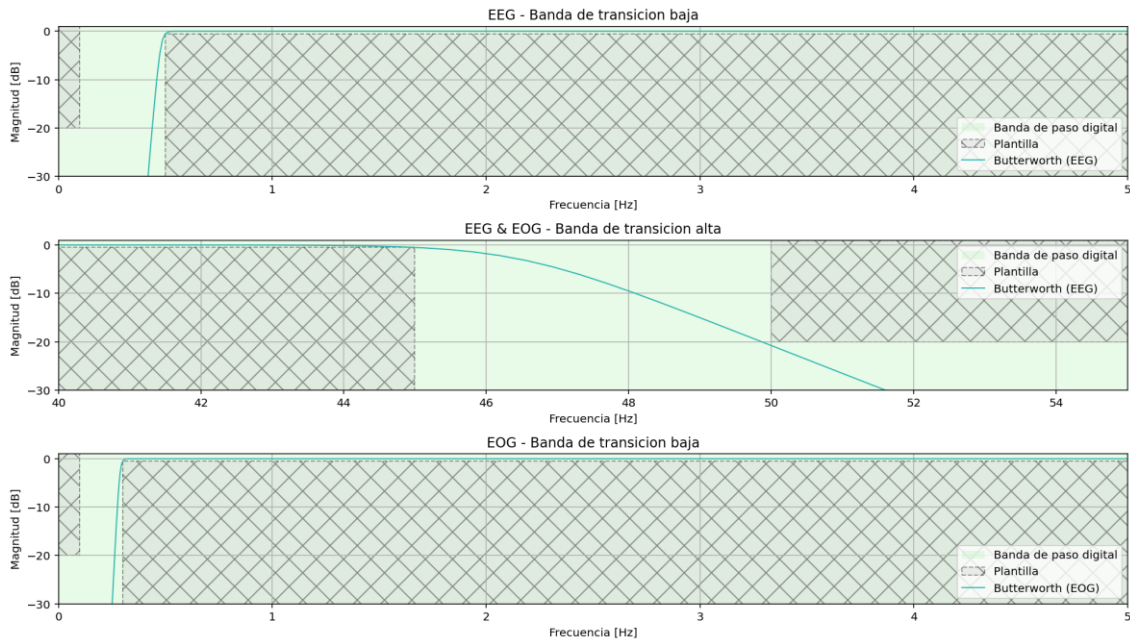
Las especificaciones del diseño se definieron en función de las bandas EEG relevantes para el análisis del sueño. Para EEG se utilizó una banda de paso de 0.5 a 45 Hz, con bandas de stop de 0.0–0.1 Hz y desde 50 Hz hasta  $f_s/2$ , ripple de 0.5 dB y atenuación de 20 dB. Para EOG se emplearon parámetros similares, con una banda de paso ligeramente más estrecha (0.3 a 40 Hz).

Sabiendo entonces que el filtrado bidireccional produce una duplicación tanto en el ripple como en la atenuación (al elevar la magnitud al cuadrado en frecuencia), es posible prever este efecto desde el diseño del filtro. Por ello, en lugar de exigir directamente 1 dB de ripple y 40 dB de atenuación en la plantilla, se optó por valores base de 0.5 dB y 20 dB, respectivamente. Esta decisión evita sobreexigir el diseño, permite usar órdenes más bajos, y garantiza que el comportamiento final del filtro —tras aplicar `sosfiltfilt`— cumpla con las especificaciones reales de atenuación total y preservación morfológica.

Este diseño fue elegido por su capacidad de preservar la morfolología de eventos neurofisiológicos como complejos K y husos de sueño, sin distorsionar las relaciones de amplitud entre componentes. A diferencia de diseños como Chebyshev o elliptic, el Butterworth no maximiza la selectividad, pero sí garantiza una transición suave y una buena respuesta global en señales fisiológicas.

Dado que el análisis se realiza en forma offline (sobre registros completos ya adquiridos), el uso de filtrado bidireccional resulta apropiado. En un sistema online, este tipo de filtrado no sería viable porque requiere conocer muestras futuras; en ese caso, se recomienda utilizar filtros FIR, cuya fase es lineal por construcción, aunque a costa de un orden mucho mayor.

En la Figura 3 se muestran las plantillas y las respuestas en frecuencia de los filtros utilizados para EEG y EOG, diseñados con las especificaciones detalladas anteriormente.



**Figura 3.** Plantillas y respuestas en frecuencia de los filtros aplicados a las señales EEG y EOG. Se utilizó un diseño IIR Butterworth para preservar las bandas relevantes del sueño y atenuar interferencias, incluyendo los 50 Hz de red eléctrica.

Dado que la clasificación de etapas de sueño se plantea en este trabajo como un procedimiento offline, es decir, sobre registros ya adquiridos y completos, el uso de filtrado bidireccional con `sosfiltfilt` resulta apropiado, ya que garantiza una respuesta en fase lineal perfecta al aplicar el filtro hacia adelante y hacia atrás. Esto permite preservar la alineación temporal de los eventos sin introducir distorsiones de fase, algo fundamental en señales neurofisiológicas.

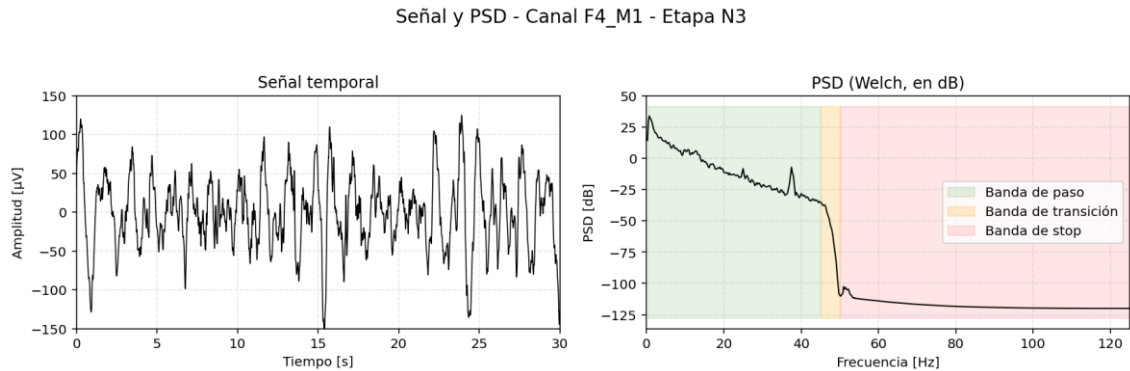
En cambio, si el sistema se orientara a una aplicación en tiempo real (online), este tipo de filtrado no sería viable, ya que requiere conocer muestras futuras. En ese caso, una alternativa recomendable serían los filtros FIR (Finite Impulse Response), cuya salida depende únicamente de valores presentes y pasados de la entrada, sin retroalimentación. Estos filtros poseen varias propiedades deseables: fase lineal inherente, diseño sencillo mediante métodos como ventanas o algoritmos óptimos (por ejemplo, Parks-McClellan), y estabilidad garantizada. La fase lineal implica que todas las frecuencias son retrasadas por igual, por lo que no hay distorsión relativa entre componentes, y el retardo introducido puede corregirse fácilmente.

Sin embargo, en la práctica observada durante el curso, los filtros FIR requieren órdenes elevados para cumplir especificaciones exigentes como bandas de transición angostas o fuerte atenuación, lo que implica mayor carga computacional y memoria. Por esta razón, y dado que el sistema es offline, se prefirió el uso de filtros IIR Butterworth, que logran resultados comparables con menor orden, especialmente adecuados cuando se implementan en estructuras numéricamente estables como las secciones de segundo orden (SOS).

A fin de validar visualmente el comportamiento del filtrado aplicado, en la Figura 4 se muestra un ejemplo representativo correspondiente a una época de 30 segundos del canal



F4\_M1. A la izquierda se observa la señal EEG filtrada en dominio temporal, y a la derecha su densidad espectral de potencia estimada mediante el método de Welch. En el espectro se destacan las bandas de paso, transición y stop especificadas en el diseño, evidenciando que el filtro cumple adecuadamente con los criterios de atenuación esperados, incluyendo la supresión efectiva de componentes por encima de 50 Hz.



**Figura 4.** Señal EEG filtrada (30 s) y su espectro en dB con bandas del filtro marcadas. Canal F4\_M1, etapa N3.

### 3.2 Segmentación

Las señales filtradas fueron segmentadas en ventanas consecutivas de 30 segundos, sin solapamiento, en concordancia con el criterio establecido por la AASM para la anotación de etapas del sueño. Este procedimiento se implementa mediante la función `segmentar()` del archivo `utils_info.py`, y se aplica automáticamente dentro del script `extraccion_feat.py`.

Cada segmento resultante constituye una unidad de análisis sobre la cual se extraen las características temporales y espectrales utilizadas en la clasificación automática.

Para ilustrar visualmente el resultado de este procedimiento, en el Apéndice se incluyen cinco figuras representativas (Figuras A.1 a A.5), una por cada etapa del sueño, que muestran la actividad registrada en los 12 canales EEG y los 2 canales EOG durante una época de 30 segundos.

### 3.3 Extracción de características

A partir de cada segmento de 30 segundos, se extrajeron características de las señales EEG y EOG mediante las funciones `features_eeg()` y `features_eog()` definidas en `utils_info.py`, ejecutadas desde el script `extraccion_feat.py`. Las señales ya habían sido previamente filtradas y segmentadas, por lo que las características reflejan información puramente asociada al contenido de cada época (30 segundos) de sueño.

#### *Características extraídas de EEG*

Para cada canal EEG se extrajeron 14 características. En el dominio temporal se calcularon:

- **Varianza:** cuantifica la dispersión de la señal respecto a su media. En señales EEG, permite estimar la variabilidad de amplitud, útil para distinguir entre estados con oscilaciones intensas (como N3) y más tenues (como REM). La representación matemática de la varianza es:

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^N (x[n] - \bar{x})^2$$

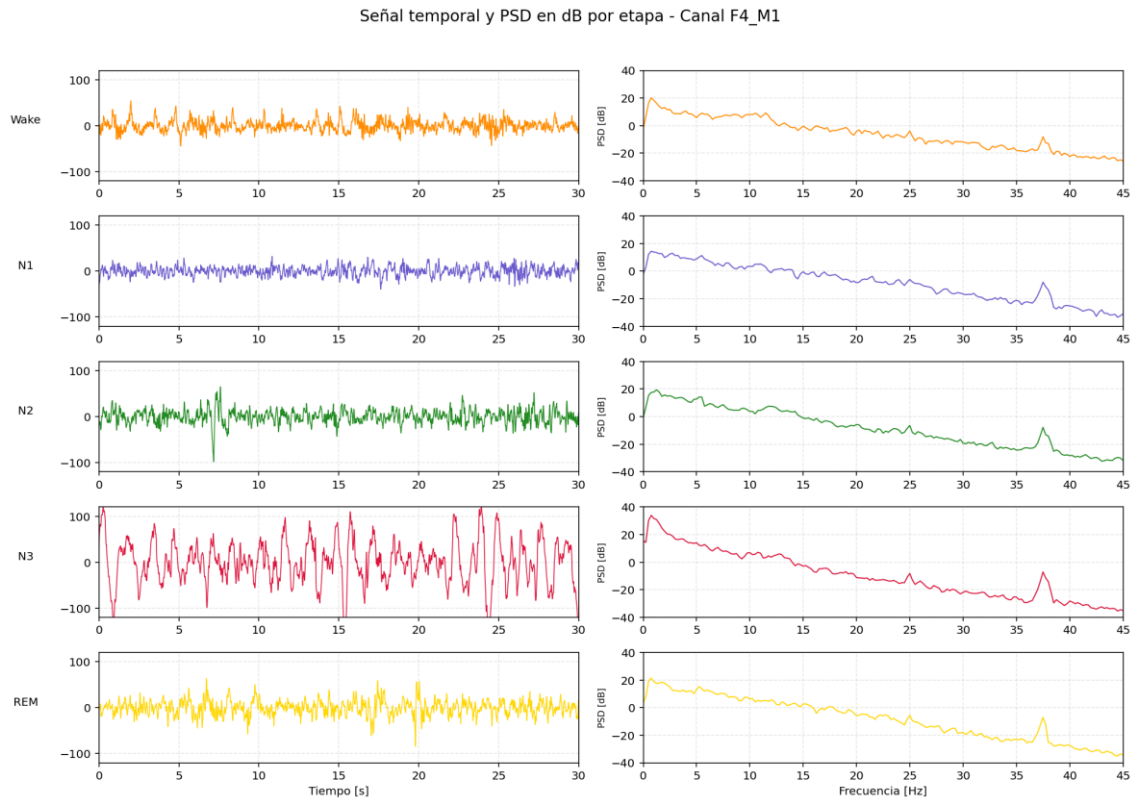
donde  $x[n]$  es la señal de EEG  $\bar{x}$  su media y  $N$  el número de muestras del segmento (serían 7500 ya que son 30 segundos y  $f_s=250\text{Hz}$ ).

- Número de cruces por cero: cuenta la cantidad de veces que la señal cambia de signo entre muestras consecutivas. Esta métrica puede discriminar entre estados con oscilaciones lentas (como  $N_3$ ) y más rápidas (como REM o vigilia).

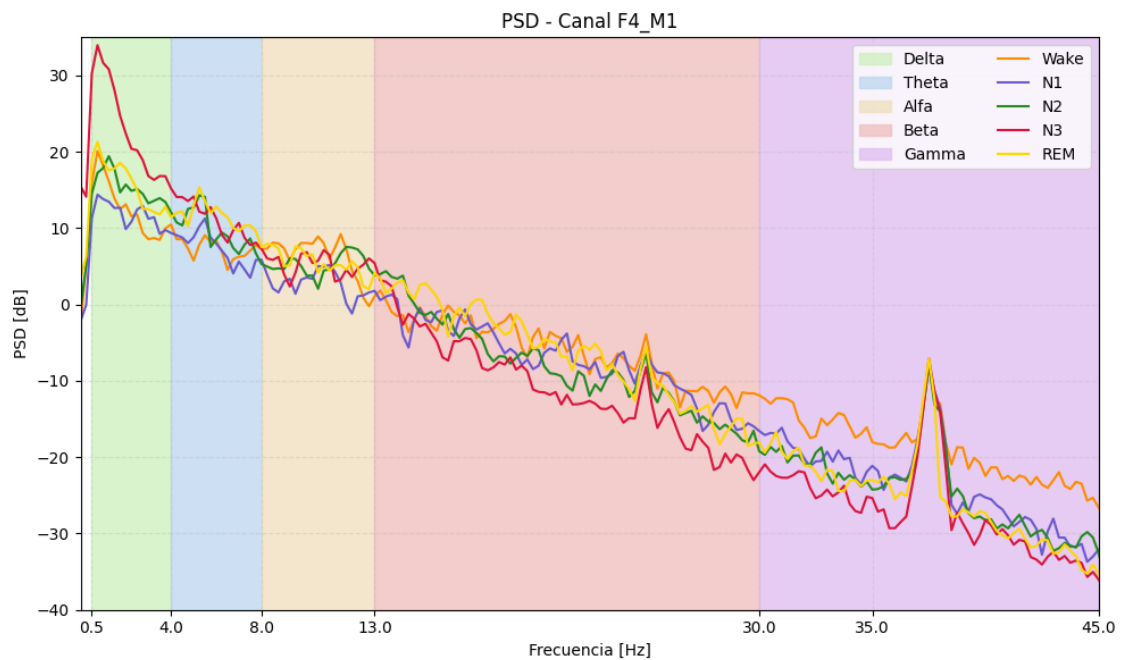
En el dominio frecuencial, se estimaron proporciones de energía en bandas relevantes: delta (0.5–4 Hz), theta baja (4–6 Hz), theta alta (6–8 Hz), theta total (4–8 Hz), alfa (8–13 Hz), beta (13–30 Hz) y gamma (30–45 Hz). Estas proporciones se obtuvieron utilizando el método de Welch, que es una mejora del periodograma clásico.

El periodograma estima la densidad espectral de potencia (PSD) como el módulo al cuadrado de la DFT de la señal. Sin embargo, esta estimación es sesgada y presenta alta varianza, aun cuando se aumenta la duración de la muestra. El método de Welch mejora esta estimación promediando periodogramas modificados, es decir, periodogramas calculados sobre segmentos solapados (aquí con 50 % de solapamiento), suavizados con una ventana de Hann para reducir efectos de borde. En este trabajo se utilizaron ventanas de hasta 4 segundos, según la duración del segmento. La energía en cada banda se obtuvo integrando la PSD sobre el rango correspondiente, y luego se normalizó respecto de la energía total. A partir de estas proporciones se derivaron cuatro relaciones espectrales: alfa/theta total, theta total/delta, beta/alfa y alfa/(delta + theta total).

Para ilustrar visualmente cómo varía el contenido espectral entre las distintas etapas del sueño, se presenta a continuación la densidad espectral de potencia (PSD) de cinco segmentos representativos, uno por clase (Wake,  $N_1$ ,  $N_2$ ,  $N_3$  y REM), obtenidos del canal EEG F4\_M1. La PSD fue estimada en decibeles mediante el método de Welch, utilizando ventanas de 4 segundos con 50 % de solapamiento. Las bandas de frecuencia EEG clásicas (delta, theta, alfa, beta, gamma) se indican con franjas de fondo coloreadas.



En la **Figura 5**, se muestran la señal EEG temporal (izquierda) y su correspondiente densidad espectral de potencia en decibeles (derecha) para una época de 30 segundos de cada etapa del sueño. Se observa cómo varía la morfología de la señal y el contenido espectral según la etapa, destacándose, por ejemplo, la predominancia de ondas delta en N3 o la mayor proporción de energía en bandas más altas durante la vigilia y REM.



En la **Figura 6**, se comparan las PSD de las cinco etapas del sueño en un mismo gráfico para el canal F4\_M1. Las bandas de frecuencia EEG clásicas se indican con colores de fondo,

facilitando la interpretación visual. Esta representación permite contrastar de manera clara las diferencias espectrales entre estados, evidenciando patrones característicos de cada uno.

Por último, se calculó la entropía de Shannon, una medida de la teoría de la información que cuantifica la complejidad o aleatoriedad de la señal. Se aplicó sobre el histograma de amplitudes del segmento, dividiendo el rango de valores en 100 bins equiespaciados. Esta métrica evalúa qué tan distribuida está la señal: si las amplitudes se concentran en pocos valores, la entropía es baja (indica una señal más predecible); si en cambio están repartidas de forma más uniforme, la entropía es alta (refleja mayor complejidad). En el contexto del sueño, esta característica permite diferenciar etapas con oscilaciones regulares como el sueño profundo (N<sub>3</sub>) de aquellas más irregulares, como REM o vigilia.

### ***Características extraídas de EOG***

Para cada canal EOG se extrajeron 8 características, agrupadas en los dominios temporal y frecuencial.

En el dominio temporal, se calcularon:

- Varianza, que cuantifica la dispersión de la señal respecto a su media.
- Número de cruces por cero, que estima la cantidad de veces que la señal cambia de signo entre muestras consecutivas.
- Cantidad total de picos, detectados como máximos y mínimos locales cuya amplitud supera una desviación estándar, con una separación mínima de 100 ms entre eventos.
- Amplitud media de los picos positivos y negativos, calculadas como el promedio de los valores pico en cada dirección.

En el dominio frecuencial, se calcularon tres proporciones de energía espectral mediante el método de Welch, integrando la densidad espectral de potencia (PSD) en las bandas 0.1-1 Hz, 1-4 Hz y 4-8 Hz. Cada proporción se obtuvo dividiendo la energía en la banda correspondiente por la energía total del segmento.

Estas características fueron calculadas para los 12 canales EEG y los 2 canales EOG disponibles por sujeto. En total, cada segmento de 30 segundos quedó representado por 184 variables: 168 provenientes de EEG (14 características por canal) y 16 de EOG (8 por canal).

### **3. 4. Detección de Complejos K mediante Wavelets y Energía de Teager**

Los complejos K son eventos transitorios característicos del sueño N<sub>2</sub>, compuestos por una onda aguda negativa seguida de una onda positiva más lenta, de al menos 0.5 segundos de duración. Su detección automática resulta útil como marcador de transición al sueño profundo y como feature relevante para la clasificación de etapas del sueño. En este trabajo se implementó una estrategia basada en transformada wavelet discreta y energía de Teager, inspirada en el enfoque de detección propuesto por Hegde et al. (2011).

## Transformada wavelet discreta (DWT)

La transformada wavelet discreta (DWT) permite representar una señal en múltiples escalas, descomponiéndola en componentes que capturan la estructura global y los detalles de la señal. Esta descomposición se logra mediante un banco de filtros paso bajo y paso alto, seguido de diezmado (submuestreo). La salida del filtro paso bajo genera los coeficientes de aproximación, que contienen la información de baja frecuencia (tendencia global), mientras que el filtro paso alto produce los coeficientes de detalle, que representan las oscilaciones locales y cambios abruptos.

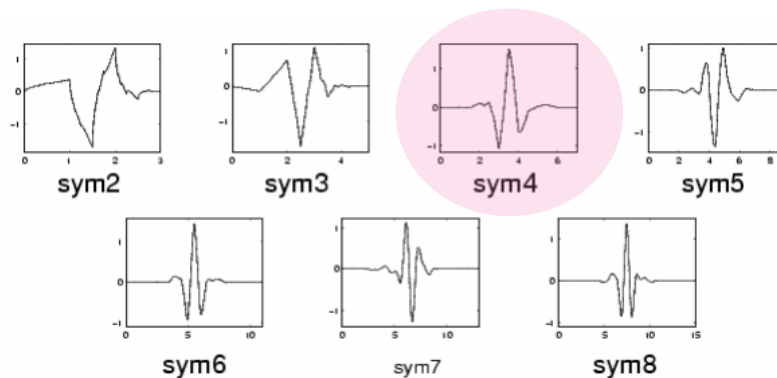
Formalmente, la señal original  $x[n]$  se proyecta sobre una base construida a partir de funciones wavelet  $\psi_{j,k}(t)$  y funciones de escalado  $\Phi_{j,k}(t)$  que son dilataciones y traslaciones discretas de funciones madre:

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k) \quad \Phi_{j,k}(t) = 2^{-\frac{j}{2}} \Phi(2^{-j}t - k)$$

Estas funciones forman una base ortonormal que permite una representación exacta de la señal (sin pérdida de información), siempre que se use una wavelet.

A nivel práctico, la DWT se implementa mediante un algoritmo multietapa, donde en cada nivel se separa la señal en una parte de aproximación (baja frecuencia) y otra de detalle (alta frecuencia), y se continúa la descomposición sobre la componente de aproximación. Este proceso genera una jerarquía de resoluciones, conocida como descomposición multirresolución, que resulta especialmente útil para el análisis de señales fisiológicas como el EEG, donde coexisten eventos lentos con otros más rápidos.

En este trabajo se utilizó la wavelet sym4, perteneciente a la familia Symmlet. Esta wavelet tiene soporte compacto, es casi simétrica, y posee cuatro momentos de desvanecimiento, lo que la hace adecuada para representar eventos EEG con cierta estructura morfológica. La descomposición se realizó hasta el nivel 6, por lo que la componente de aproximación A6 retiene principalmente frecuencias entre 0 y 1.95 Hz (para una frecuencia de muestreo de 250 Hz). Este rango incluye las oscilaciones lentas típicas de los complejos K, sin eliminar completamente su morfología.



**Figura 7.** Funciones wavelet de la familia Symmlet desde sym2 hasta sym8.

## Energía de Teager -Kaiser (TEO)

Para estimar la energía de la señal de aproximación A6, se aplicó el operador de energía de Teager, una técnica desarrollada originalmente por H. Teager y formalizada por Kaiser (1990). Este operador calcula la energía instantánea generada por el fenómeno físico que origina la señal tomando únicamente tres muestras sucesivas y se define matemáticamente como:

$$\Psi\{x[n]\} = x[n]^2 - x[n-1] \cdot x[n+1]$$

A diferencia de la energía cuadrática clásica  $x[n]^2$ , que depende únicamente de la amplitud, el TEO captura también la variación interna de la señal, ya que combina el valor actual con sus vecinos adyacentes. Esto le permite responder de manera más precisa a oscilaciones de corta duración, lo que lo hace particularmente adecuado para detectar eventos transitorios como los complejos K.

Desde una perspectiva física, Kaiser demostró que este operador se relaciona con la energía de oscilaciones simples, siendo proporcional al producto del cuadrado de la amplitud y del cuadrado de la frecuencia:  $E \propto A^2 \omega^2$ . Esta propiedad no está presente en la energía clásica, que no distingue entre una onda de 10 Hz y otra de 1000 Hz si ambas tienen la misma amplitud. El TEO, en cambio, asigna mayor energía a frecuencias más altas, reflejando mejor el esfuerzo necesario para generar esa señal.

Además, el operador es no lineal, lo que resulta beneficioso: su estructura le permite ignorar señales constantes, rampas o componentes lentas, respondiendo solo ante estructuras oscilatorias localizadas. También es invariante ante inversiones temporales y robusto frente a cruces por cero, ya que no requiere divisiones ni derivadas.

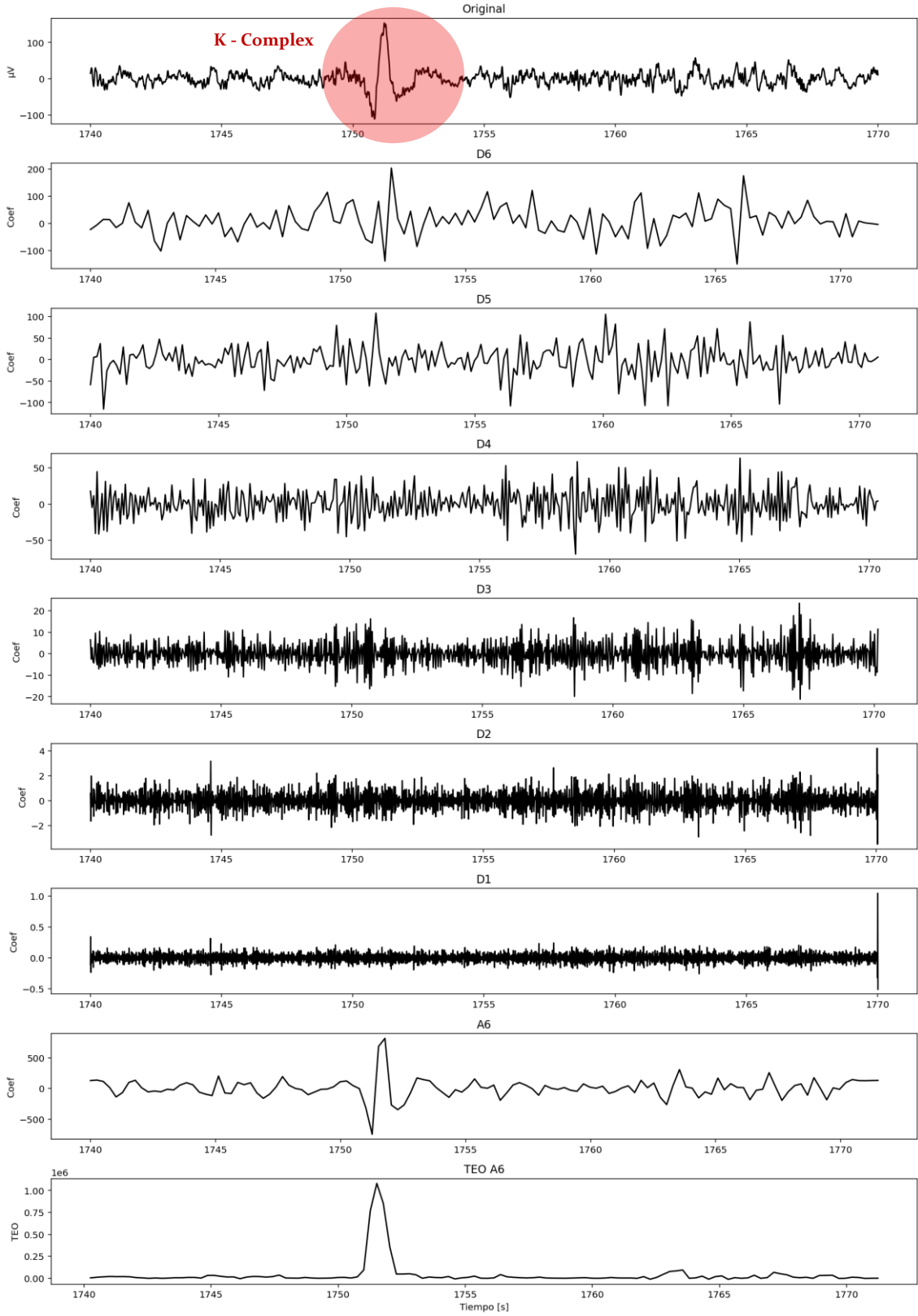
Dado que el TEO introduce términos cruzados si se aplica a señales con múltiples componentes frecuenciales, su uso se restringe a señales previamente filtradas. Por eso, en este trabajo se aplica únicamente sobre la componente A6 de la descomposición wavelet, que concentra la banda de interés. La señal resultante se normaliza y se le aplica un detector de picos para identificar eventos candidatos. Cada pico detectado se interpreta como un posible complejo K, y su cantidad se registra como una característica adicional por canal.

## Detección y cuantificación de eventos

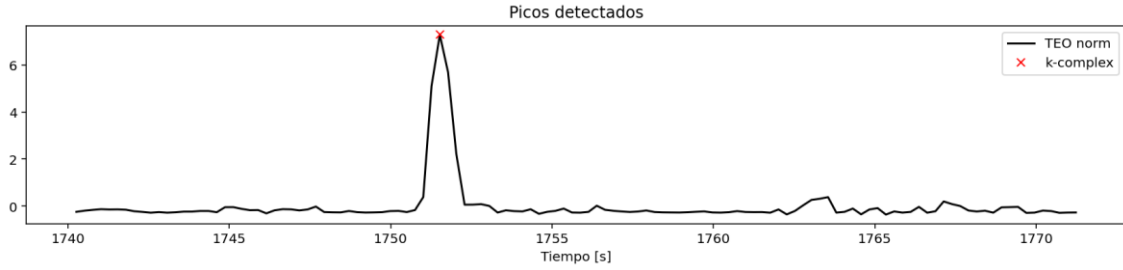
Luego de aplicar el operador de energía de Teager sobre la componente A6, la señal resultante se normaliza (media cero, desviación estándar uno) para hacer comparables sus amplitudes entre segmentos. Sobre esta señal normalizada se aplica un detector de picos, basado en condiciones de altura mínima, prominencia y distancia temporal, calibradas empíricamente para evitar detecciones erróneas.

Cada pico detectado se interpreta como un evento candidato a complejo K, y se cuantifica su número dentro del segmento de 30 segundos. Este valor se incorpora como una característica adicional por canal EEG, enriqueciendo el conjunto de descriptores morfológicos y espectrales disponibles para el modelo de clasificación de etapas de sueño. Esta métrica aporta información sobre la presencia y densidad de actividad transitoria característica del estadio N2, pero también permite detectar eventos aislados en otras

etapas. En la Figura 9 se presenta la misma señal TEO normalizada, junto con los picos detectados.



**Figura 8.** Descomposición wavelet de un segmento EEG de 30 segundos del canal F4\_M1 correspondiente a la etapa N2. Se muestran la señal original, los detalles D6 a D1, la aproximación A6 y la energía de Teager (TEO) calculada sobre A6. Las líneas rojas marcan el intervalo donde ocurre un complejo K.



**Figura 9.** Energía de Teager (TEO) normalizada del componente A6 (wavelet sym6, nivel 6), con picos detectados indicados en rojo. Estos picos corresponden a eventos con la morfología y prominencia características de los complejos K.

### 3.5 Clasificación

#### *Validación cruzada por grupos*

El conjunto completo de 25 sujetos del dataset DOD-H fue dividido en 5 grupos aproximadamente equilibrados, de forma que en cada iteración se utilizó un grupo diferente como conjunto de test y los cuatro restantes como conjunto de entrenamiento. Este esquema de validación cruzada Leave-One-Group-Out permite evaluar la capacidad de generalización del modelo a sujetos no vistos durante el entrenamiento, una condición más exigente y realista. En cada fold, se entrenó el modelo sobre el 80 % de los datos del grupo de entrenamiento y se evaluó sobre el 20 % restante del grupo de test.

#### **Modelos utilizados**

Para abordar la tarea de clasificación automática de etapas del sueño, se entrenaron tres modelos supervisados utilizando los vectores de características extraídos por ventana de 30 segundos: Ridge Classifier, Regresión Logística y Random Forest.

#### **Ridge Classifier**

Es un modelo de clasificación lineal que intenta encontrar una fórmula matemática capaz de separar las distintas clases (etapas del sueño) en función de las características numéricas extraídas de las señales. Para ello, construye una combinación lineal de las características:

$$f(x) = w_1x_1 + w_2x_2 + \dots + w_nx_n$$

donde:

- $x = (x_1, x_2, \dots, x_n)$  es el vector de características de una muestra (por ejemplo, varianza, energía en delta, etc.),
- $w_1, \dots, w_n$  son los "pesos" que el modelo asigna a cada característica.



Luego, para evitar que estos pesos tomen valores muy grandes (lo que podría indicar sobreajuste), se agrega una penalización sobre su magnitud. El modelo busca minimizar:

$$J(w) = \sum (y_i - w^t x_i)^2 + \lambda ||w||^2$$

- $y_i$  es la clase real de la muestra  $i$ ,
- $\lambda$  es un número positivo que controla cuánto se penalizan los pesos grandes.

Este modelo es útil cuando hay muchas variables correlacionadas, como sucede con los canales EEG.

## Regresión Logística

Este modelo también combina linealmente las características de entrada (como el Ridge), pero en lugar de predecir directamente una clase, calcula la probabilidad de que una muestra pertenezca a cada clase. Esta probabilidad se obtiene aplicando una función llamada softmax:

$$P(y = k | x) = \frac{\exp(w_k^t x + b_k)}{\sum_j \exp(w_j^t x + b_j)}$$

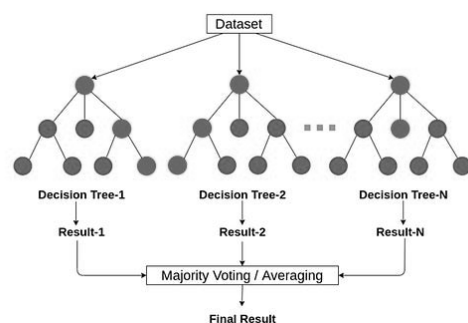
Donde:

- $w_k$  es el conjunto de pesos para la clase  $k$ ,
- $b_k$  es un número adicional llamado sesgo (permite desplazar la función),
- La fracción asegura que todas las probabilidades sumen 1.

El modelo asigna como clase final aquella con mayor probabilidad estimada. La regresión logística es ampliamente usada por su simplicidad y buena interpretación.

## Random Forest

Es un modelo que no se basa en fórmulas matemáticas lineales, sino en árboles de decisión. Un árbol de decisión clasifica una muestra haciendo preguntas secuenciales sobre sus características (por ejemplo, "¿la energía en delta es mayor a cierto valor?"). En Random Forest, se entrenan muchos árboles diferentes, cada uno con distintos subconjuntos de datos y variables. Luego, cada árbol vota por una clase, y la clase más votada es la predicción final.



Este método es robusto, tolerante al ruido, y muy útil cuando las relaciones entre variables y clases son complejas o no lineales.

Todos los modelos fueron entrenados utilizando la biblioteca scikit-learn.

## Métricas de evaluación

Para cada modelo y fold se calcularon las siguientes métricas:

- Accuracy: proporción de épocas correctamente clasificadas sobre el total:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

- Recall por clase: fracción de instancias verdaderamente positivas que fueron correctamente identificadas:

$$Recall = \frac{TP}{(TP + FN)}$$

- Precisión (Precision) por clase: fracción de predicciones positivas que fueron correctas:

$$Precision = \frac{TP}{(TP + FP)}$$

- F1-score por clase: promedio armónico entre precisión y sensibilidad:

$$F1 = 2 \times \frac{Precision \times Recall}{(Precision + Recall)}$$

- Matriz de confusión: tabla que muestra la cantidad de instancias clasificadas correctamente e incorrectamente por clase. Se reporta en dos versiones: absoluta (conteo bruto) y normalizada por fila (proporción respecto a la clase real).

## 4. Resultados

### Resultados

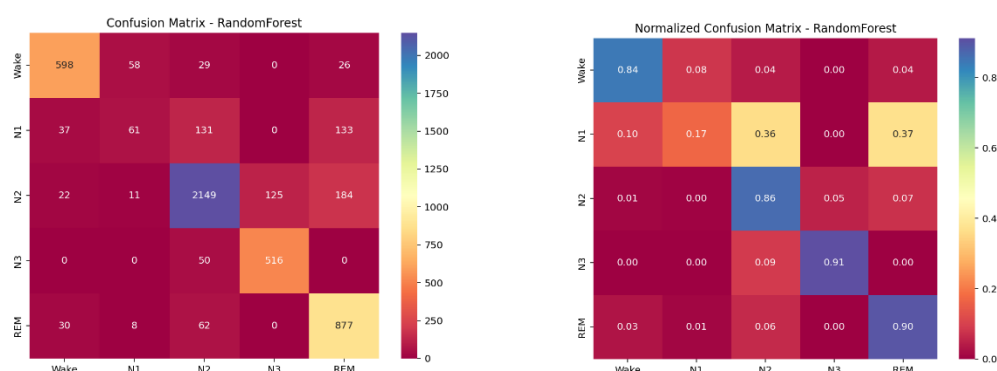
En la Tabla 1 se presentan las métricas de desempeño obtenidas para los cinco folds de validación cruzada, evaluando tres clasificadores: Random Forest, Regresión Logística y Ridge Classifier. Se incluyen los valores de exactitud (Accuracy), sensibilidad (Recall), precisión (Precision) y F1-score por fold, así como el promedio y desvío estándar de cada métrica.

Fold	RandomForest				LogisticRegression				RidgeClassifier			
	Acc	Recall	Pre	F1	Acc	Rec	Pre	F1	Acc	Rec	Pre	F1
1	0.77	0.71	0.68	0.68	0.74	0.72	0.67	0.68	0.65	0.67	0.61	0.61
2	0.80	0.67	0.77	0.69	0.74	0.66	0.68	0.67	0.71	0.65	0.65	0.64
3	0.80	0.70	0.72	0.70	0.70	0.69	0.63	0.63	0.67	0.66	0.59	0.60
4	0.80	0.67	0.73	0.70	0.71	0.70	0.66	0.66	0.70	0.67	0.62	0.63
5	0.82	0.74	0.74	0.73	0.76	0.78	0.71	0.73	0.69	0.72	0.65	0.66
Mean	0.80	0.70	0.73	0.70	0.73	0.71	0.67	0.67	0.68	0.67	0.62	0.63
	± 0.02	± 0.03	± 0.03	± 0.02	± 0.02	± 0.04	± 0.03	± 0.04	± 0.02	± 0.03	± 0.03	± 0.02

**Tabla 1:** Resultados de validación cruzada (5 folds) para los tres modelos evaluados: Random Forest, Regresión Logística y Ridge Classifier. Se reportan los valores de exactitud (Accuracy), sensibilidad (Recall), precisión (Precision) y F1-score para cada fold, junto con el promedio y el desvío estándar al pie de la tabla.

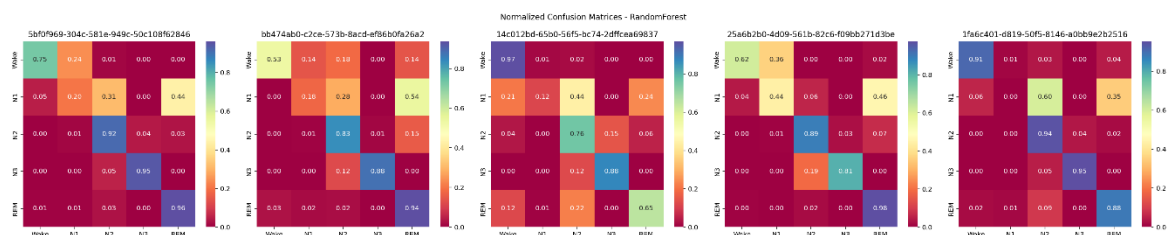
Se observa que el modelo Random Forest obtuvo consistentemente el mejor rendimiento, superando en todas las métricas a los otros dos clasificadores. Por este motivo, en esta sección se presentan únicamente los resultados detallados de este modelo, mientras que las figuras correspondientes a los otros dos se incluyen en el Apéndice A.

A continuación, se muestran las matrices de confusión generales del fold 5 para Random Forest, ya que este fold presentó los valores más altos de desempeño de forma consistente entre sujetos.



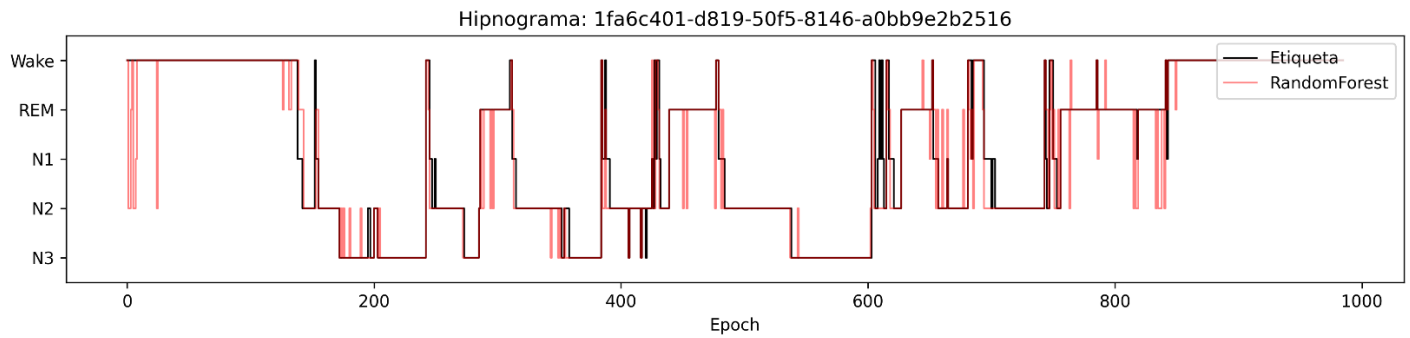
**Figura 10:** Matrices de confusión generales para el modelo Random Forest en el fold 5. A la izquierda se muestra la matriz con conteos absolutos, y a la derecha su versión normalizada por fila (proporción respecto a la clase verdadera).

También se incluyen las matrices de confusión individuales por sujeto de test en el mismo fold. Estas permiten observar la variabilidad interindividual en la predicción de las etapas del sueño.



**Figura 11:** Matrices de confusión por sujeto para el modelo Random Forest en el fold 5.

Finalmente, se presenta un hipnograma predicho por Random Forest para el sujeto con mejor desempeño en el fold 5. La figura permite comparar visualmente la secuencia de etapas del sueño predicha por el modelo con la anotación manual.



**Figura 14.** Hipnograma predicho por el modelo Random Forest (fold 5). Se muestra la secuencia temporal de etapas a lo largo de la noche, junto con la anotación manual de referencia.

## Discussion:

Los resultados obtenidos muestran que el sistema de clasificación desarrollado funciona muy bien, sobre todo si se considera que se basa en un pipeline relativamente simple y solo usa EEG y EOG. Las tasas de acierto en etapas como N3 (91 %), REM (90 %) y N2 (86 %) son muy buenas, y están en línea —o incluso por encima— de lo que suelen lograr muchos clasificadores con señales adicionales o arquitecturas más complejas. Wake también tuvo buen rendimiento (84 %). Como era esperable, el rendimiento más bajo se dio en N1 (17 %), una etapa difícil incluso para anotadores humanos, por su ambigüedad y transición con otras fases.

Desde el punto de vista del análisis de señales, se aplicaron herramientas sólidas: filtrado IIR con buena preservación de fase, estimación espectral con Welch, medidas de variabilidad, y también análisis morfológico con wavelets. Las características elegidas cubren bien lo que se espera en señales de sueño, tanto en términos de ritmo como de complejidad.

Dicho eso, hay varias cosas que podrían mejorarse. Por ejemplo, si bien la detección de complejos K funcionó en la mayoría de los casos, también puede generar falsos positivos, sobre todo en canales donde no suelen aparecer. Una mejora posible sería restringir la detección a canales más representativos (como C3/M2 o F3/M2), en lugar de aplicarla a todos por igual.

Tampoco se incluyó un detector de husos de sueño, que son muy relevantes para identificar N2. Sumarlos podría mejorar la separación entre N1 y N2, que es una de las transiciones más difíciles de clasificar.

En cuanto al EOG, se usó de manera bastante acotada. Por cuestiones de tiempo y foco, se calcularon solo algunas características básicas, pero hay espacio para explorar mucho más: movimientos oculares rápidos, diferencias entre canales, direccionalidad, etc. Esto sería especialmente útil para diferenciar mejor REM de otras etapas.

También sería interesante incorporar otras señales disponibles, como EMG, que permitiría distinguir mejor entre REM (sin tono muscular) y Wake. Con eso se ganaría información fisiológica complementaria sin necesidad de cambiar el resto del pipeline.

En lo que respecta al preprocesamiento, si bien se aplicaron buenos filtros, el trabajo no abordó la detección de artefactos. Usar herramientas como MNE para limpiar las señales (por ejemplo, con ICA) podría ayudar a evitar que se entrenen modelos sobre tramos ruidosos o contaminados.

Y por último, aunque Welch es una elección estándar para el análisis espectral, no se comparó con otras opciones como Blackman-Tukey. Probar diferentes métodos podría revelar si hay diferencias significativas en la forma en que se representa la energía por banda y cómo eso impacta en la clasificación.

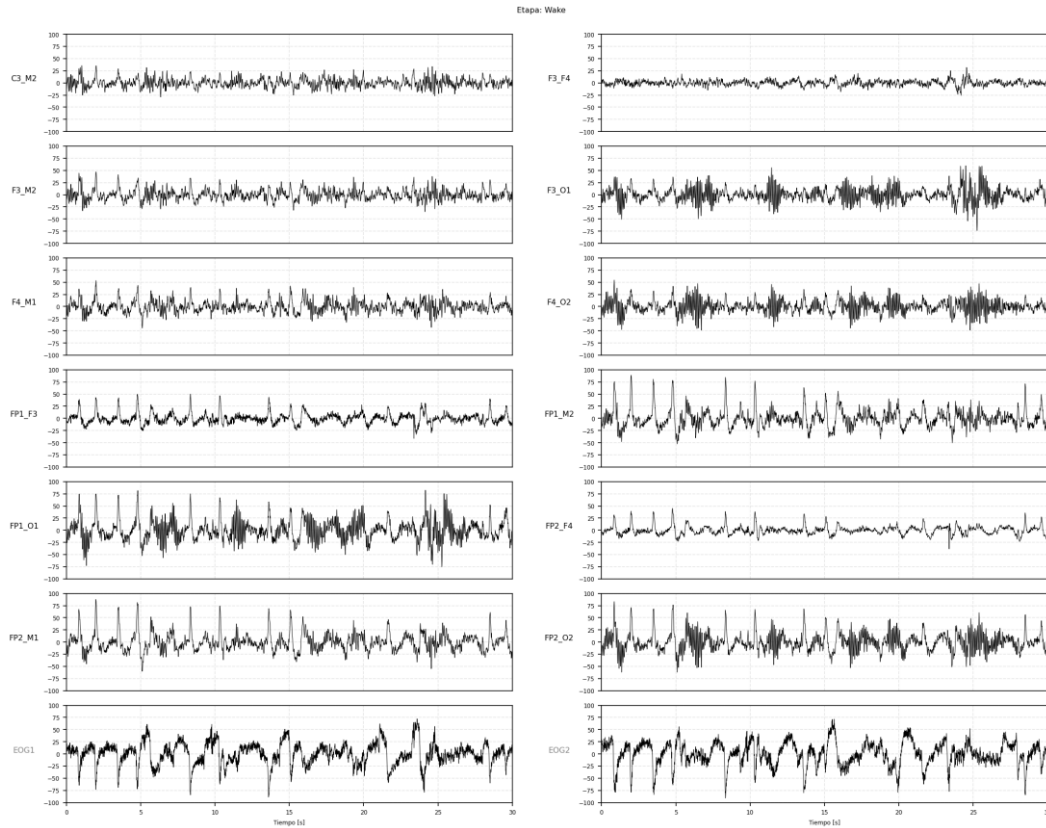
En resumen, el rendimiento del sistema es muy bueno, pero hay margen claro para seguir mejorando: no tanto en la arquitectura del modelo, sino en la calidad y profundidad del análisis de las señales. Mejorando las características, agregando información fisiológica relevante y refinando el preprocesamiento, es probable que se pueda obtener un sistema más robusto y generalizable.

## Bibliografía

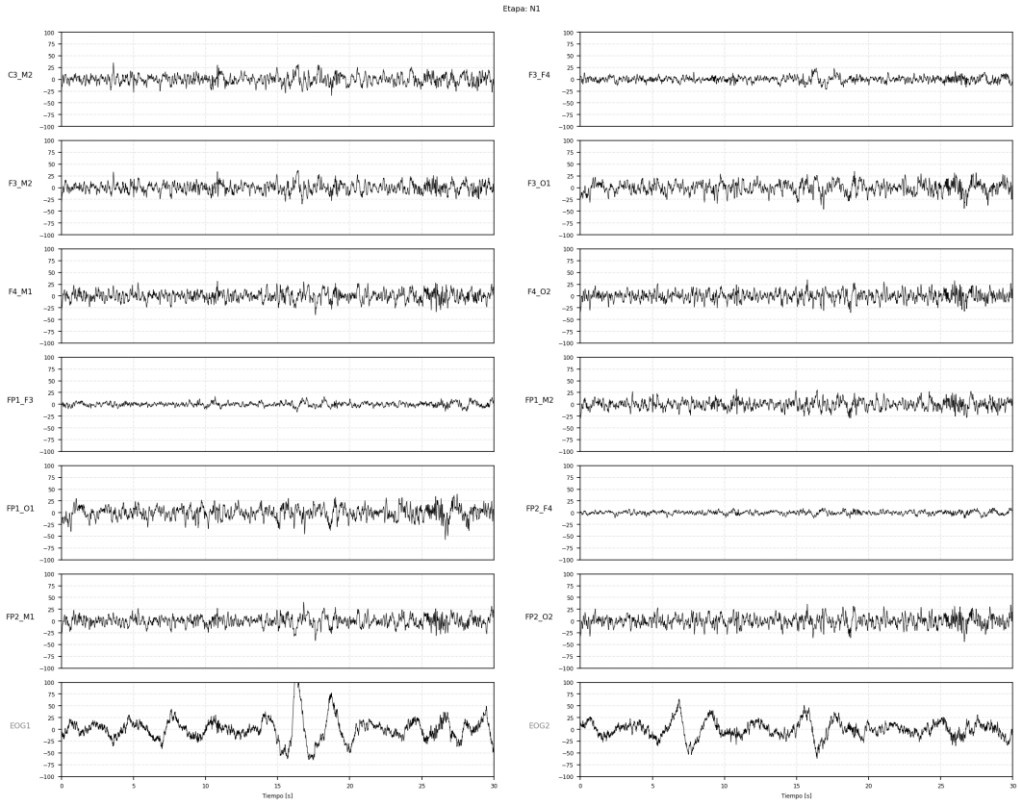
1. Brinkman, J. E., Reddy, V., & Sharma, S. (2023). *Physiology, Sleep*. StatPearls Publishing. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK482512/>
2. Guillot, A., Sauvet, F., During, E. H., & Thorey, V. (2020). Dreem Open Datasets: Multi-Scored Sleep Datasets to Compare Human and Automated Sleep Staging. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. <https://doi.org/10.1109/TNSRE.2020.3011181>
3. Permana, K. E., & Iramina, K. (2024). *Single Channel Sleep Stage Classification Using Random Forest and Feature Extraction with SMOTE Balancing on Fpz-Cz EEG Data*. Mathematical Modelling of Engineering Problems. <https://doi.org/10.18280/mmep.111204>
4. Devuyst, S., Dutoit, T., Stenuit, P., Kerkhofs, M. (2011). *Automatic K-complexes detection in sleep EEG recordings using the Teager energy operator*. Expert Systems with Applications. <https://doi.org/10.1016/j.eswa.2010.10.081>
5. Kaiser, J. F. (1990). *On a simple algorithm to calculate the 'energy' of a signal*. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 381–384. Recuperado de: <https://sci-hub.se/https://ieeexplore.ieee.org/document/115702>

## Apéndice A – Visualización de épocas representativas por etapa del sueño

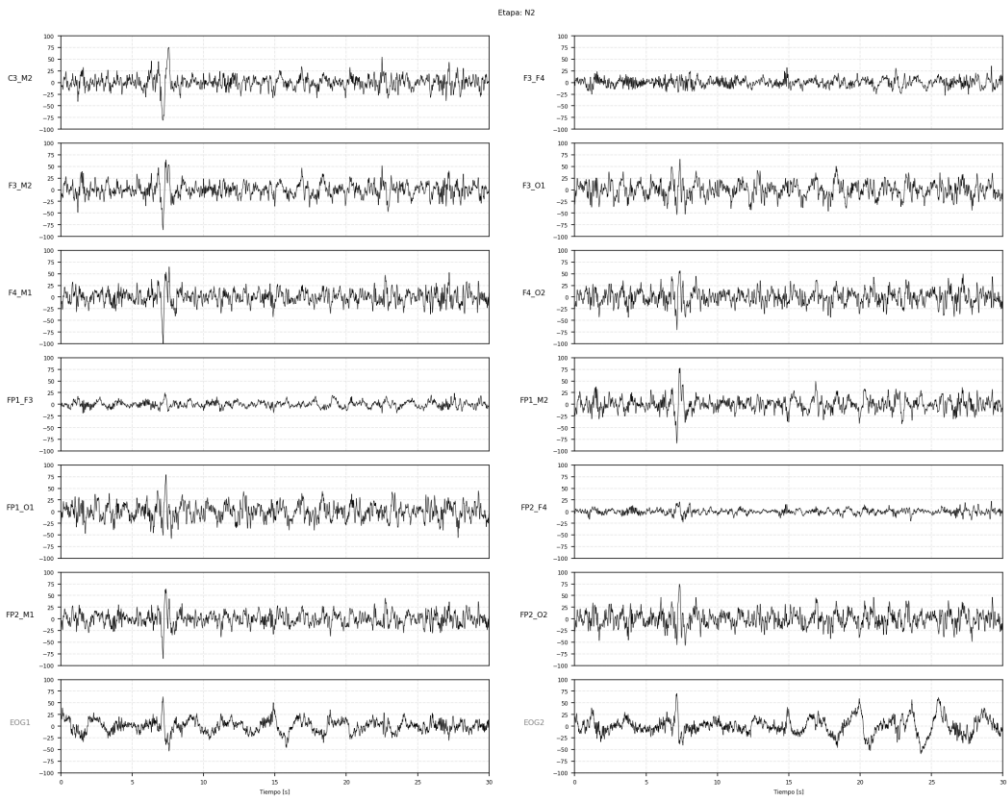
A continuación se presentan cinco figuras, una por cada etapa del sueño (Wake, N1, N2, N3 y REM), que muestran los 12 canales EEG y los 2 canales EOG durante una época de 30 segundos seleccionada como representativa.



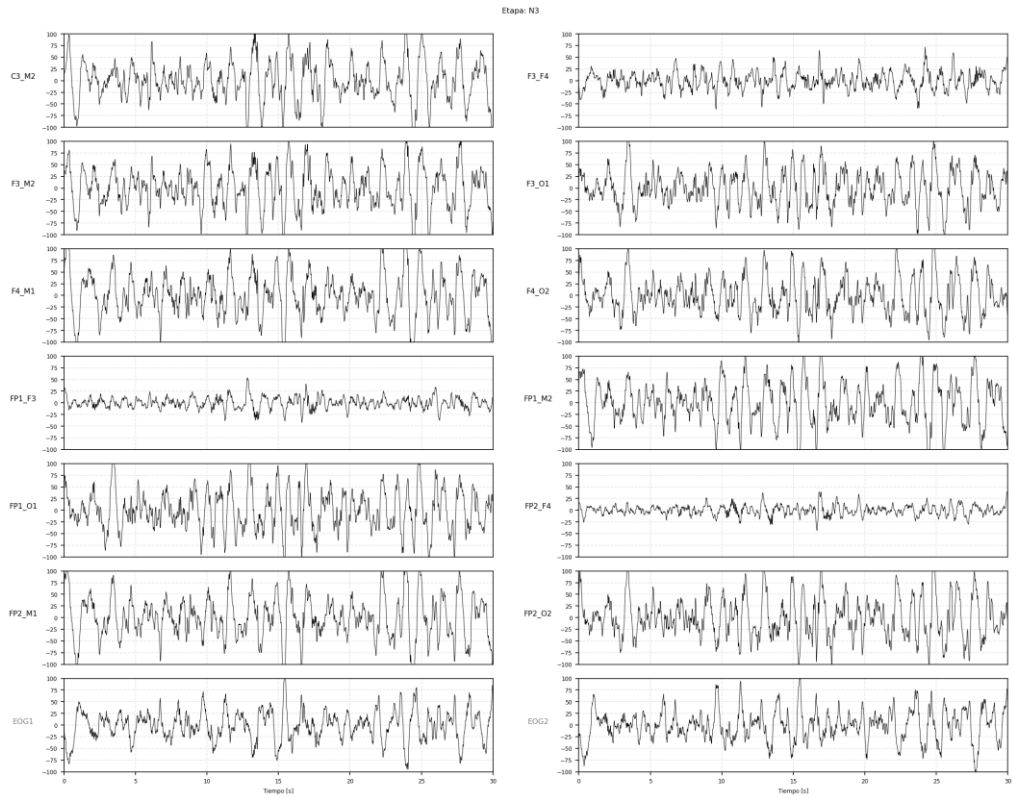
**Figura A.1:** Época (30 segundos) del estadio Wake (vigilia) ya preprocesado.



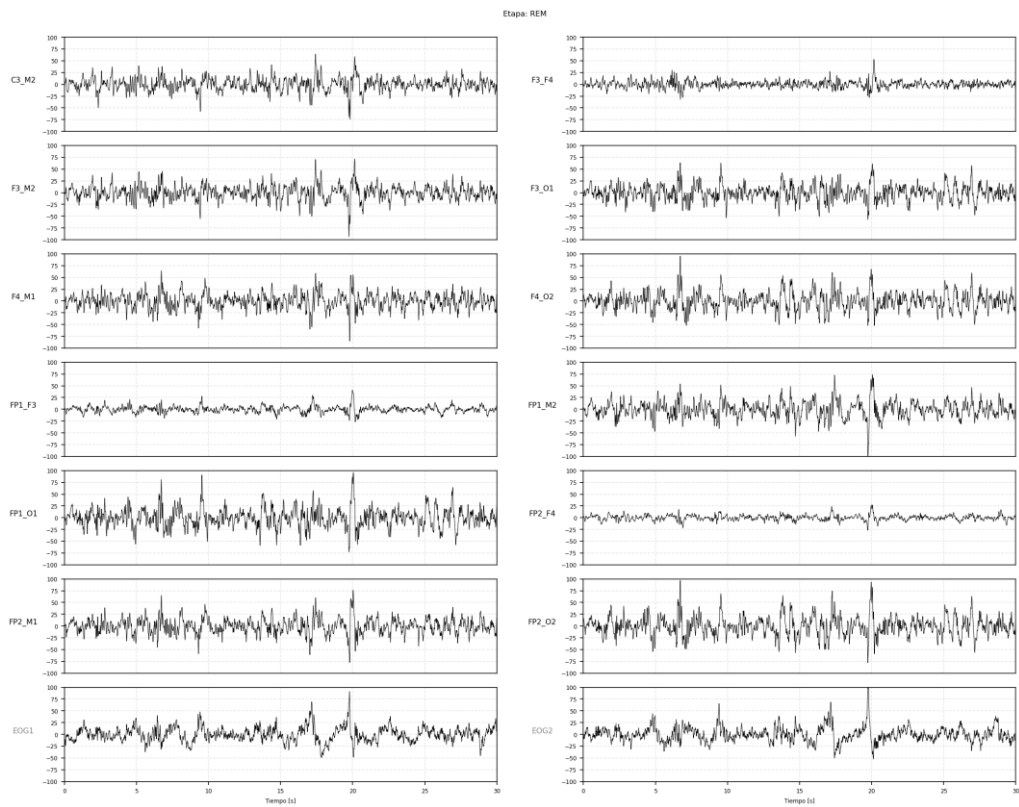
**Figura A.2:** *Epoca (30 segundos) del estadio N1 ya preprocesado.*



**Figura A.3:** *Epoca (30 segundos) del estadio N2 ya preprocesado.*



**Figura A.4:** *Epoca (30 segundos) del estadio N<sub>3</sub> ya preprocesado.*



**Figura A.5:** *Epoca (30 segundos) del estadio REM ya preprocesado.*



## Apéndice B. Resultados complementarios: Logistic Regression y Ridge Classifier

En esta sección se incluyen las matrices de confusión y los hipnogramas correspondientes a los modelos Regresión Logística y Ridge Classifier, evaluados en el fold 5 del esquema de validación cruzada.

Estas figuras se presentan como complemento de los resultados principales mostrados en la Sección 4, y permiten analizar el comportamiento comparativo de estos clasificadores frente a Random Forest, que fue el modelo con mejor rendimiento general.

Se muestran para cada modelo:

- La matriz de confusión general (en versión absoluta y normalizada),
- Las matrices individuales por sujeto de test, y
- El hipnograma predicho para el sujeto con mejor desempeño.

