

# Denoising de señales de electroencefalograma utilizando LMS y sus variantes

Francisco Basili  
Ingeniería Electrónica  
ITBA  
Buenos Aires, Argentina  
fbasili@itba.edu.ar

Sergio Andrés Peralta  
Ingeniería Electrónica  
ITBA  
Buenos Aires, Argentina  
speralta@itba.edu.ar

Kevin Wahle  
Ingeniería Electrónica  
ITBA  
Buenos Aires, Argentina  
kwahle@itba.edu.ar

**Resumen**—Este artículo presenta la implementación y resultados de un sistema de cancelación de interferencia diseñado para eliminar artefactos no deseados en señales de electroencefalograma (EEG). Se emplea la técnica de descenso por gradiente en algoritmos como Least Mean Squares (LMS), Normalized Least Mean Squares (NLMS) y Quantized Least Mean Squares (Q-LMS). El enfoque propuesto busca mejorar la calidad de las señales EEG, buscando el camino para poder lograr diagnósticos más precisos.

## I. INTRODUCCIÓN

El electroencefalograma (EEG) es una herramienta de gran utilidad en el estudio de la actividad cerebral y desempeña un papel fundamental en la investigación médica y clínica, particularmente en el diagnóstico y seguimiento de trastornos neurológicos, así como en la comprensión de los procesos cognitivos y emocionales. Sin embargo, las señales de EEG frecuentemente se ven afectadas por interferencias no deseadas, que pueden obstaculizar una interpretación precisa y dificultar su utilidad en aplicaciones clínicas y de investigación. La presencia de artefactos musculares (EMG), artefactos oculares derivados de movimientos oculares y parpadeos (EOG) y artefactos cardíacos (ECG), es común en los registros de EEG, lo que plantea una problemática para la obtención de información cerebral fiable.

El artefacto EOG se origina en el ojo humano debido a la diferencia de carga eléctrica entre la córnea positiva y la retina negativa. Este fenómeno da lugar a un dipolo eléctrico, cuyas características se ven modificadas por los movimientos oculares y los parpadeos. Como resultado, se genera una señal eléctrica conocida como EOG. En registros de EEG, una parte de esta señal EOG se propaga hacia el cuero cabelludo, donde se superpone con las señales cerebrales, especialmente en electrodos frontales.

Por otro lado, el artefacto EMG consiste en la actividad eléctrica generada por la contracción de músculos. Esta actividad eléctrica puede propagarse y superponerse con las señales cerebrales. Las señales EMG se caracterizan por tener una alta amplitud y frecuencias más elevadas en comparación con las señales EEG. A pesar de que generalmente no modifica en gran medida el EEG, la dificultad para verificar y filtrar adecuadamente el artefacto EMG puede ser un desafío importante. Además, es poco común encontrar este tipo de artefacto

en datasets de EEG, lo que lleva a la preferencia de utilizar el ECG en su lugar para evitar confusiones y errores en la interpretación.

A su vez, el artefacto ECG se origina en la actividad eléctrica del corazón y se manifiesta como una señal eléctrica que se propaga por el cuerpo. En comparación con el EOG y el EMG, el artefacto ECG suele ser más fácil de identificar y separar de las señales cerebrales. Por esta razón, en muchos casos, se prefiere utilizar el ECG en lugar del EMG para lograr llegar a resultados más contundentes, especialmente en aplicaciones como la de este proyecto.

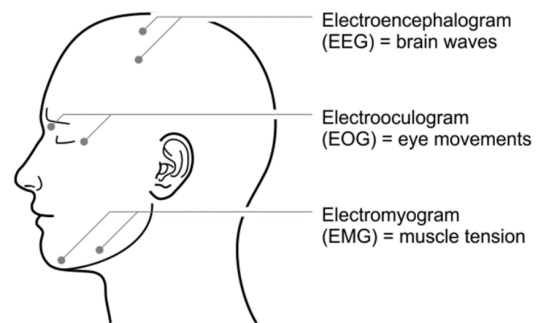


Figura 1: Obtención de señales de EEG, EOG y EMG [1].

El contenido espectral del EEG se refleja en patrones de ondas que están asociados con diferentes estados de alerta y actividad mental. Estos patrones se dividen en distintas bandas espectrales, que son intervalos de frecuencia específicos. Estas bandas incluyen delta ( $\delta$ ) de 0.5 a 4 Hz, theta ( $\theta$ ) de 4 a 8 Hz, alpha ( $\alpha$ ) de 8 a 13 Hz, beta ( $\beta$ ) de 13 a 22 Hz y gamma ( $\gamma$ ) de 22 a 30 Hz.

Una característica crucial del EEG es su naturaleza localmente estacionaria, lo que significa que sus propiedades estadísticas cambian lentamente con el tiempo, pero se mantienen estacionarias en intervalos de tiempo cortos [2]. En el caso del EEG, este período de estacionariedad local se estima típicamente en no más de 30 segundos.

## II. DESARROLLO

### II-A. Dataset

Para llevar a cabo las diferentes pruebas para lograr el correcto denoising de la señal de EEG se utilizó el MIT-BIH Polysomnographic Database [3]. Este dataset cuenta con distintas grabaciones de señales fisiológicas durante el sueño. Los sujetos incluidos en el estudio fueron sometidos a monitoreo en el Laboratorio de Sueño del Hospital Beth Israel en Boston. El propósito principal de este estudio fue evaluar el síndrome de apnea obstructiva del sueño crónico.

El conjunto de datos contiene más de 80 horas de grabaciones polysomnográficas con entre cuatro y siete canales, que incluyen señales de EEG, EOG (electrooculograma derecho), EMG y ECG, entre otras. Todos los sujetos incluidos en la base de datos eran hombres, con edades comprendidas entre 32 y 56 años y un peso promedio de 119 kg.

De particular interés en este estudio es el conjunto de datos correspondiente al paciente 41 (slp41). Este conjunto cuenta con una frecuencia de muestreo de 250 Hz y tiene una duración de 6.5 horas (5,850,000 muestras de datos).

### II-B. Análisis y preprocesamiento de las señales

Para realizar el denoising se tomó un fragmento de 90 segundos de las señales originales. Si bien no cumple con el criterio de estacionaridad local establecido en la sección previa, se realizó un test de estacionaridad pertinente (Augmented Dickey-Fuller) para comprobar la correcta selección de valores. El tiempo de los fragmentos debió ser necesariamente mayor al previamente establecido (30 segundos), debido al tiempo de convergencia de la señal resultante.

Debido a la condición de estacionaridad local de las señales, resultó apropiado para aplicar los algoritmos normalizarlas, restándole la media, la cual se puede observar en la Tabla I.

	Media	Varianza
EEG	-6.382844	52.472913
ECG	-51.683022	13633.971791
EOG	4.944258	518.197179

Tabla I: Media y varianza de las distintas señales a analizar.

A su vez fue necesario adelantar temporalmente las señales de ECG y EOG para asegurar causalidad a la hora de aplicar el algoritmo. Los adelantos fueron de 10 y 400 muestras respectivamente. Las señales temporales resultantes y sus respectivos espectrogramas se pueden observar en las Figuras 2 y 3.

En la Figura 2, se evidencia una notoria correlación entre las diversas señales. Se destaca la presencia de picos en el registro del EOG que se reflejan claramente en la señal del EEG, con especial énfasis en un pico destacado que se produce alrededor del segundo 65. Además, es claramente discernible la presencia de los picos del ECG en la señal de EEG, y estos picos mantienen su frecuencia característica.

El espectrograma de las diversas señales se presenta en la Figura 3. En este gráfico, se destaca una notable relación entre las señales, particularmente entre el EEG y el EOG. Se aprecia

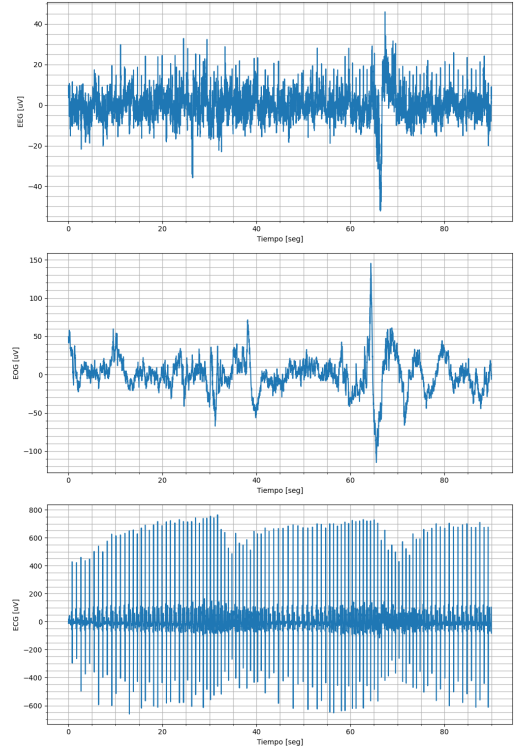


Figura 2: Señales temporales de EEG, ECG y EOG.

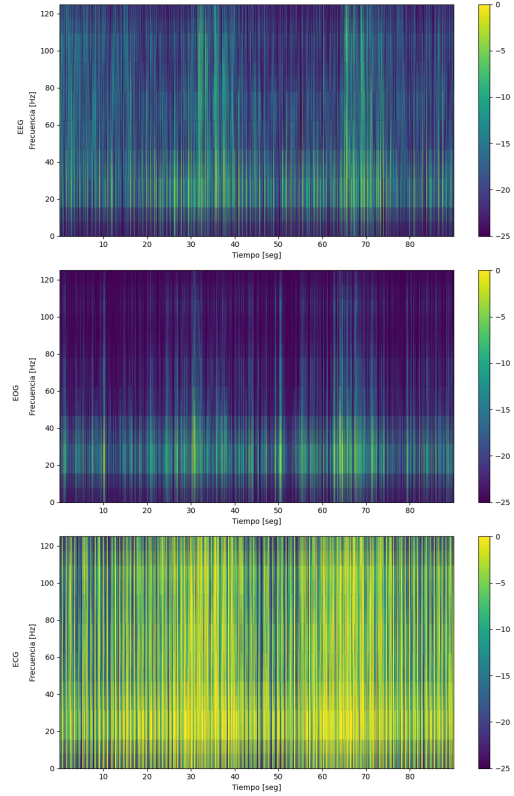


Figura 3: Espectrogramas de EEG, ECG y EOG.

cómo comparten picos de energía en las mismas frecuencias simultáneamente.

El modelo a utilizar para lograr el denoising será el de cancelación de interferencia del tipo cascada [4], tal como se lo puede observar en la Figura 4.

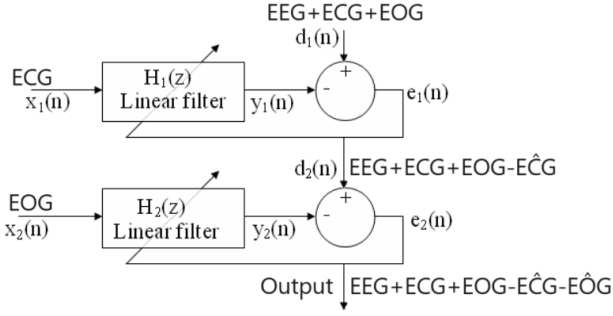


Figura 4: Modelo de cancelación de interferencia utilizado.

Se procedió a analizar las correlaciones cruzadas entre el EEG y el EOG, y entre el EEG y el ECG. Esto permitió observar la relación entre las señales de manera cuantitativa e identificar para qué valores de lags (delay entre muestras) sucedía. También permitió verificar la causalidad de las señales y obtener un valor adecuado de orden de filtro  $M$ . Los resultados obtenidos se pueden observar en las Figuras 5 y 6. Allí se verifica que la correlación cruzada respecto al EOG se extiende a mayores valores de lags en comparación a la obtenida respecto al ECG, por lo que se decidió utilizar un orden de filtro  $M_{ECG} = 15$  y  $M_{EOG} = 750$ .

Posteriormente se estimaron los autovalores de la matriz de autocorrelación del ECG y del EOG. Para ello, haciendo uso de la condición de estacionariedad previamente mencionada, se asumió que la matriz era Toeplitz para facilitar el cálculo. Una vez obtenidos los autovalores máximos, y teniendo en cuenta las relaciones en la Ecuación 1, se obtuvieron los valores de la Ecuación 2, donde  $\lambda$  son los autovalores de la matriz de autocorrelación,  $\chi$  es el spread de autovalores y  $\mu$  es el parámetro de paso.

$$\chi = \frac{\lambda_{max}}{\lambda_{min}} \quad \mu_{max} = \frac{2}{\lambda_{max}} \quad (1)$$

$$\begin{aligned} \lambda_{maxECG} &= 7,64 \times 10^4 & \lambda_{maxEOG} &= 1,24 \times 10^5 \\ \mu_{maxECG} &= 2,62 \times 10^{-5} & \mu_{maxEOG} &= 1,62 \times 10^{-5} \\ \chi_{ECG} &= 2,94 \times 10^3 & \chi_{EOG} &= 1,5 \times 10^6 \end{aligned} \quad (2)$$

### II-C. Cuestiones a analizar

Durante el desarrollo del proyecto se destacaron ciertas cuestiones a tener en cuenta. Principalmente se notó una fuerte relación entre el orden del filtro elegido y la dificultad para hallar el valor de  $\mu_{optimo}$ : cuanto mayor era el valor de  $M$ , menor era el paso necesario para pasar de un estado de no alteración de la señal a la divergencia.

También se debieron tener en cuenta cuestiones relacionadas con las métricas. En esta aplicación, como la señal de interés

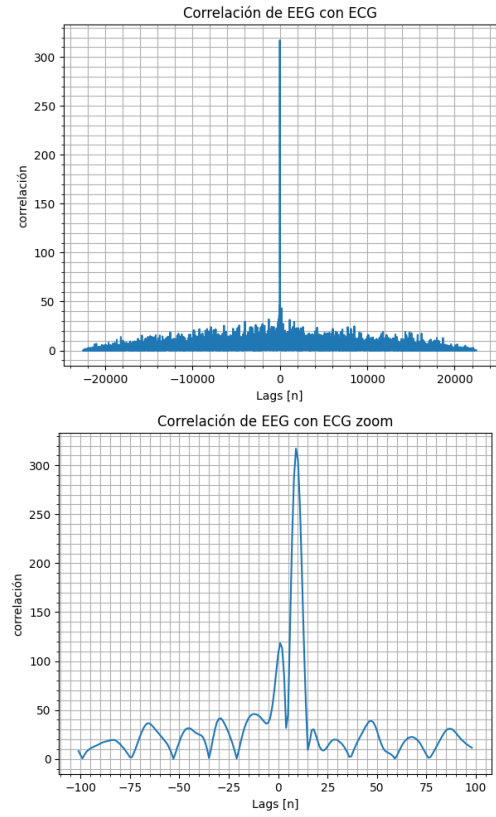


Figura 5: Correlación cruzada entre el EEG y el ECG.

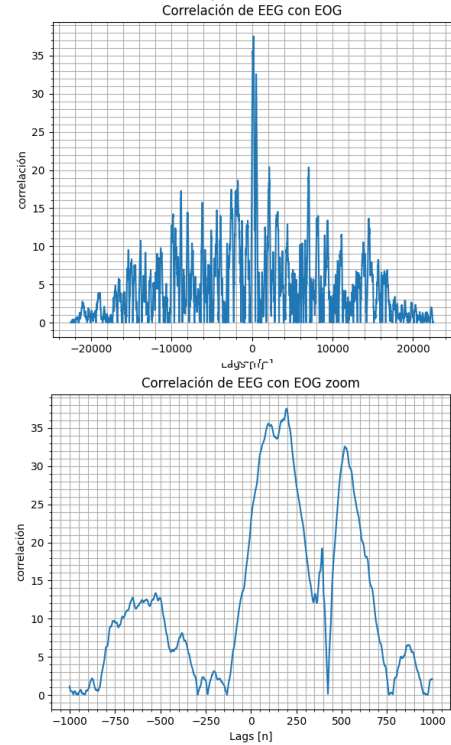


Figura 6: Correlación cruzada entre el EEG y el EOG.

está incluida en el error, ciertas métricas como la curva de aprendizaje no son aplicables. Dado que no disponíamos de una señal sin ruido como referencia, no fue posible calcular métricas estándar como el Error Cuadrático Medio (MSE) o la Relación Señal-Ruido (SNR), entre otras. Fue por ello que resultó necesario establecer una métrica propia para evaluar la calidad del proceso de filtrado. Esta métrica, denominada Factor de Mejora, se puede obtener a partir de la Ecuación 3, donde  $r_{x-y}$  es la correlación cruzada entre  $x$  e  $y$ , y  $\|a\|^2$  representa la energía de  $a$ .

$$FM_y = \frac{\|r_{EEG-y\{\text{prefiltrado}\}}\|^2}{\|r_{EEG-y\{\text{postfiltrado}\}}\|^2} \quad (3)$$

Además, se identificó una relación significativa entre el valor de  $\mu$  y la energía de la señal. Esto resultó ser de suma importancia para garantizar la condición de  $\mu$  pequeño. Este fenómeno se vincula con  $\mu_{max}$ , que a su vez está condicionado por el valor de  $\lambda$ , y este último depende de la matriz de autocorrelación. A su vez, la matriz de autocorrelación se ve influenciada por la energía de la señal de entrada. Por lo tanto, la escala de la señal afecta directamente al valor de  $\mu_{max}$ , por lo que debe ser cuidadosamente considerada (si es que es expresada en  $\mu V$ ,  $mV$ ,  $V$ , etc.).

#### II-D. LMS

Se aplicó LMS teniendo en cuenta el modelo de la Figura 4 y los  $\mu_{max}$  calculados, aplicando una actualización de pesos dada por la Ecuación 4. Fue necesario realizar un ajuste minucioso de los  $\mu$ . Para ello se evaluó el factor de mejora para distintos valores como se puede observar en la figura 7. Allí se puede observar como el factor de mejora FM aumenta a medida que  $\mu$  lo hace, hasta llegar a un estado de divergencia, momento a partir del cual FM cae abruptamente.

$$w(n+1) = w(n) + \mu \cdot x(n) \cdot e^*(n) \quad (4)$$

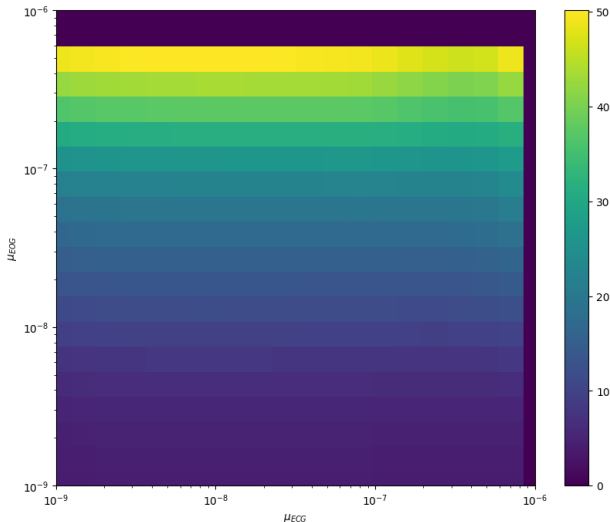


Figura 7: Variación del FM para distintos valores de  $\mu$ .

Finalmente se obtuvieron los mejores resultados tomando valores de  $\mu$  que se encontrasen a cierto margen del estado de divergencia, resultando  $\mu_{ECG} = \mu_{EOG} = 1 \times 10^{-7}$ .

#### II-E. NLMS

Se aplicó NLMS teniendo en cuenta el modelo de la Figura 4. A diferencia de LMS, se independizó a la actualización de los pesos de la energía de la señal, siendo la nueva actualización de pesos la expresada en la Ecuación 5. Este algoritmo parecería tener una gran utilidad a priori, debido a las grandes variaciones de energía perteneciente al ECG.

$$w(n+1) = w(n) + \frac{\tilde{\mu}}{\|x(n)\|^2 + \delta} x(n) e^*(n) \quad (5)$$

Finalmente se tomaron valores de  $\tilde{\mu}$  que se encontrasen con cierto margen del estado divergencia, siendo  $\mu_{ECG} = 2 \times 10^{-4}$ ,  $\mu_{EOG} = 5 \times 10^{-2}$ ,  $\delta_{ECG} = \delta_{EOG} = 1 \times 10^{-10}$ . Como era de esperarse, debido a la normalización aplicada, los valores de  $\tilde{\mu}$  resultaron ser mayores a los valores de  $\mu$  utilizados en LMS.

#### II-F. Q-LMS

Q-LMS es una variante LMS que incorpora una cuantización en sus actualizaciones de peso. Esta cuantización reduce la complejidad computacional en comparación con LMS. Al trabajar con valores cuantizados en lugar de continuos, Q-LMS disminuye la cantidad de operaciones aritméticas necesarias para actualizar los pesos del filtro adaptativo. Esto puede resultar en una implementación más eficiente en términos de recursos computacionales y velocidad de procesamiento.

Además, una característica significativa de Q-LMS es su capacidad para manejar señales ruidosas de manera efectiva. La cuantización puede actuar como un mecanismo de filtrado adicional, ayudando a mitigar el impacto del ruido en las actualizaciones de peso. Esto lo convierte en una herramienta valiosa en entornos donde la señal de interés está contaminada por interferencias o perturbaciones.

La actualización de pesos se define como [5] [6]:

$$w(n+1) = w(n) + \mu \cdot (q(n) + 1) \cdot x(n) e^*(n)$$

donde:

$$q_{upper} = \frac{2}{\mu \cdot \lambda_{m\acute{a}x}}$$

$$\psi(n+1) = \beta \psi(n) + \gamma e(n)^2, \quad (0 < \beta < 1, \gamma > 0)$$

$$q(n+1) = \begin{cases} q_{upper} & \text{if } \psi(n+1) > q_{upper} \\ 1 & \text{if } \psi(n+1) < 1 \\ \psi(n+1) & \text{otherwise} \end{cases}$$

Se puede observar que  $\psi(n)$  es el parámetro que se utiliza como regla de decisión para la actualización de  $q(n)$ , y consiste en un modelo en el que  $\beta$  es el parámetro autoregresivo y  $\gamma$

se encarga de realizar la actualización en relación al cuadrado del error en el instante previo.

Finalmente se tomaron valores de  $\mu$  que se encontrasen con cierto margen del estado de divergencia, siendo  $\mu_{ECG} = 1 \times 10^{-8}$ ,  $\mu_{EOG} = 1 \times 10^{-7}$ ,  $\beta = 0,5$  y  $\gamma = 5 \times 10^{-3}$ .

### III. RESULTADOS

Se procedió a realizar una comparación entre la señal de error y la señal de EEG original utilizando los tres algoritmos, como se puede apreciar en la Figura 8.

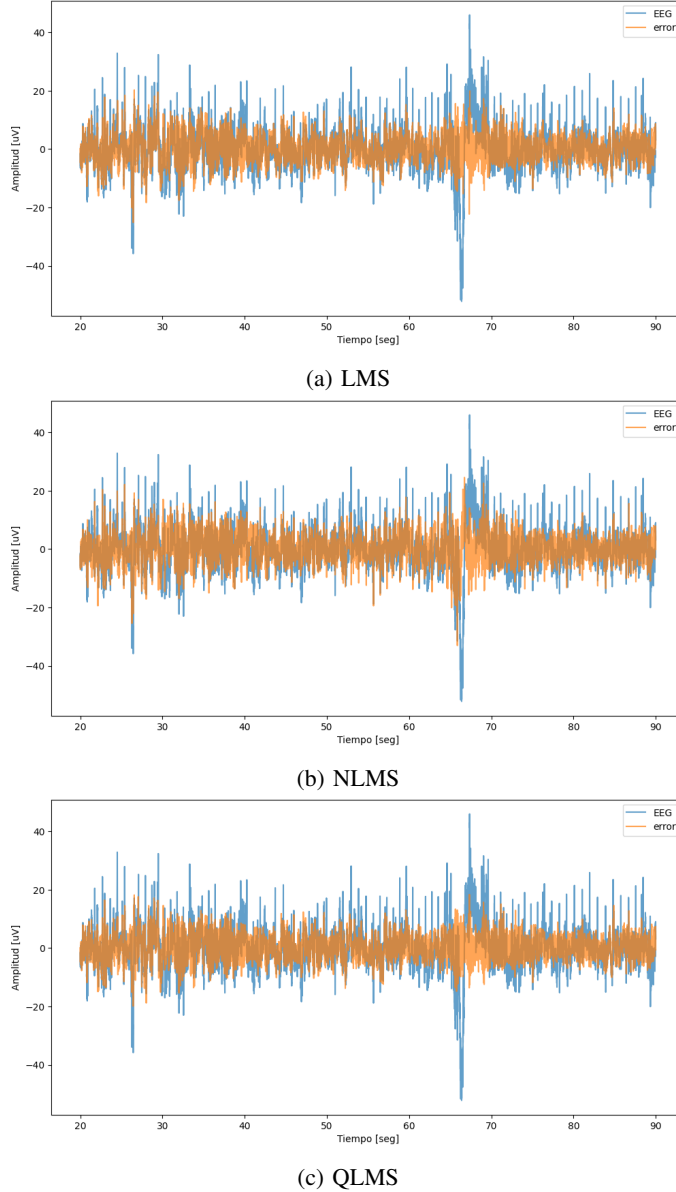


Figura 8: Error y EEG original para distintos algoritmos.

En estas representaciones, es evidente que los tres algoritmos logran un filtrado exitoso de los artefactos presentes en la señal original. A medida que transcurre el tiempo, se observa un progresivo aumento en el desempeño del filtrado de estos

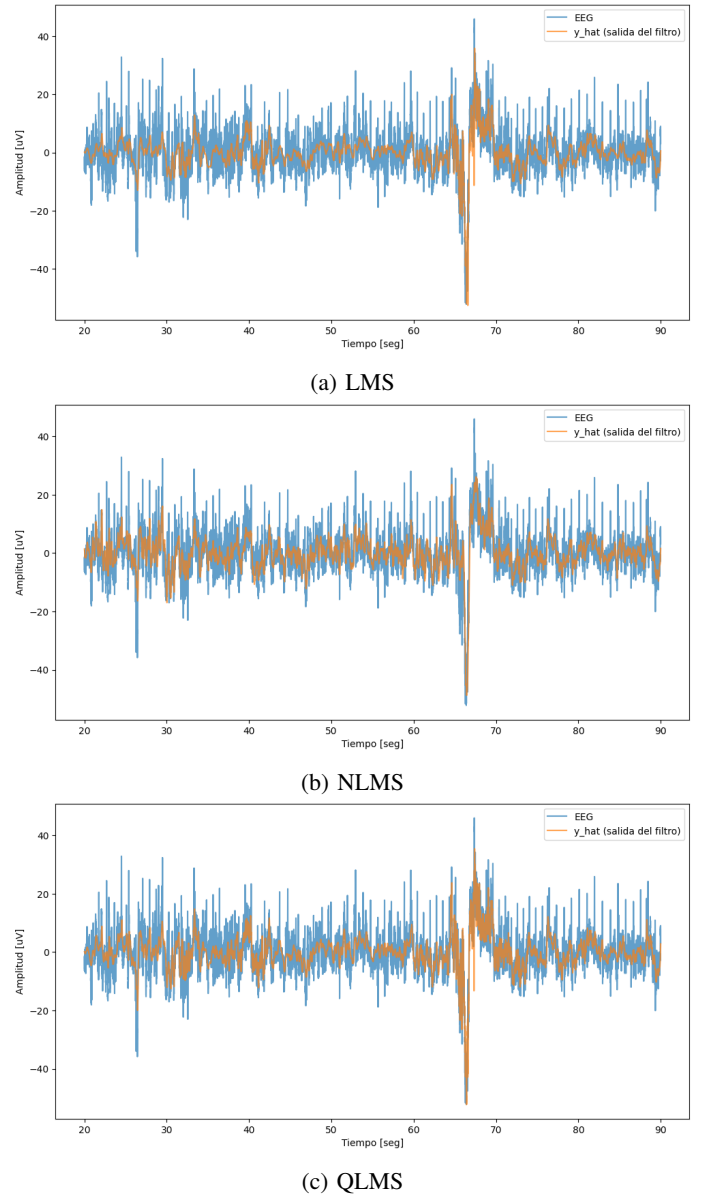


Figura 9:  $\hat{y}$  y EEG para distintos algoritmos.

artefactos, especialmente en lo que respecta a la eliminación de las interferencias del ECG.

Es relevante destacar que, en la Figura 8b, se aprecia que el algoritmo NLMS no logra reducir los artefactos del EOG al mismo nivel que los otros algoritmos. Esta diferencia se hace más evidente en la Figura 9, donde se presenta tanto la señal original de EEG como la salida del filtro aplicado al EOG, la cual denota la señal correlacionada que va a ser eliminada.

No obstante, es importante destacar que el algoritmo NLMS posee una ventaja inobjetable. Al estar normalizado por la energía de la señal, los coeficientes correspondientes al ECG exhiben variaciones menos abruptas en comparación con los otros algoritmos, tal como se puede apreciar en la Figura 9. Se notó que todos los algoritmos aportan un mayor peso al coeficiente 350 del EOG, correspondiente a valores de lag



de 400 muestras, lo que coincide con el adelanto temporal mencionado anteriormente. Sin embargo, curiosamente, este valor de lag no es el de mayor correlación observado en la Figura 6.

con el algoritmo Q-LMS.

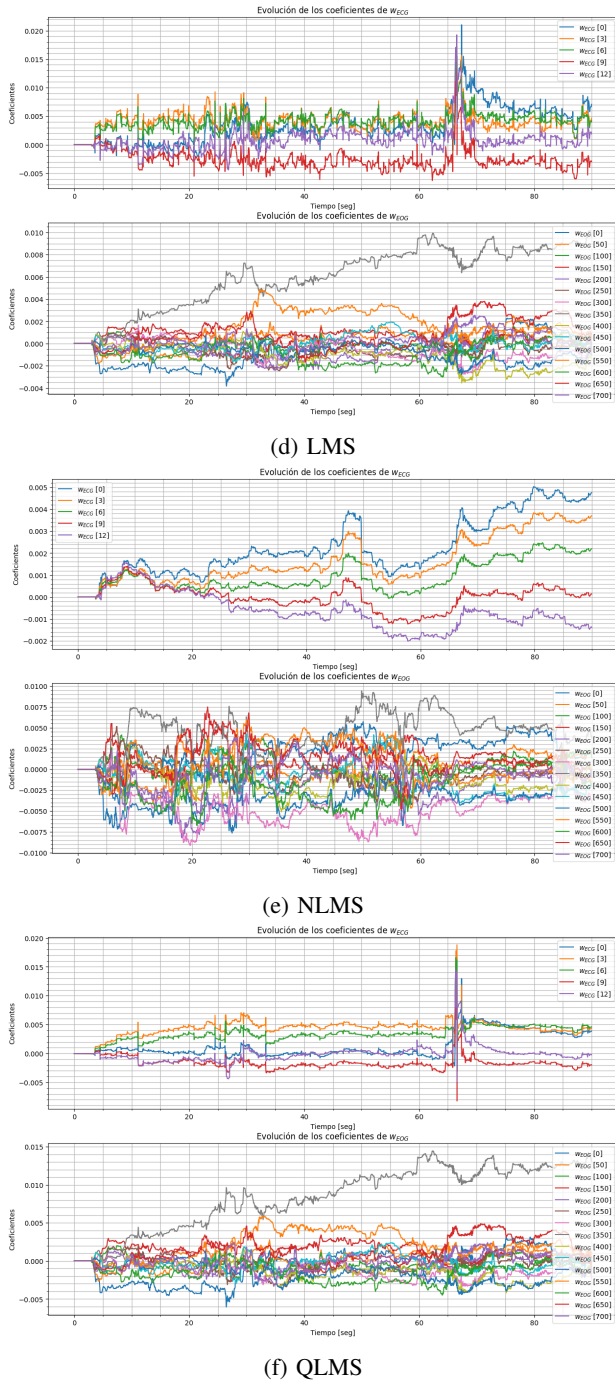


Figura 9: Variación de los coeficientes para ambas señales.

Finalmente se calculó el factor de mejora (FM) para cada uno de los algoritmos. En las Figuras 10 y 11, se puede observar la correlación cruzada y la energía obtenida con la señal ya filtrada. Se notó una gran reducción de la misma, respecto a la observada en las Figuras 5 y 6, especialmente

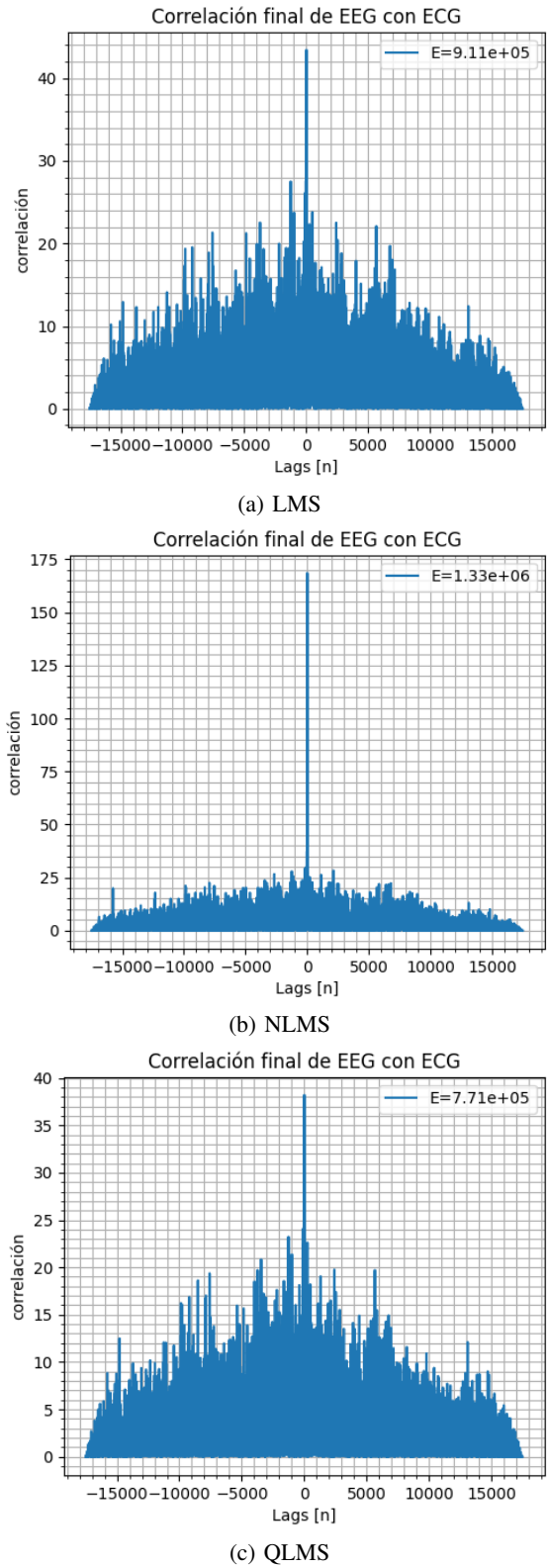
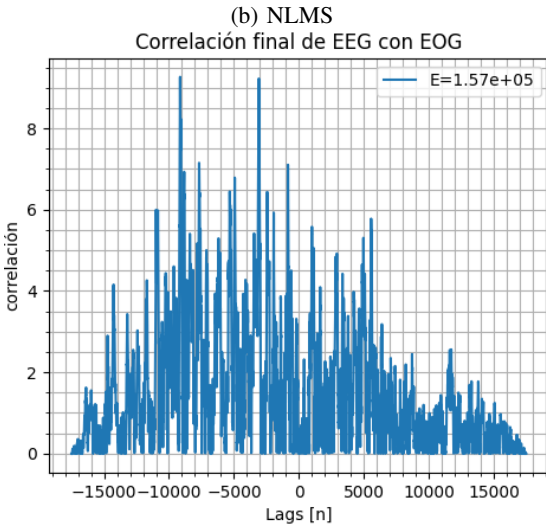
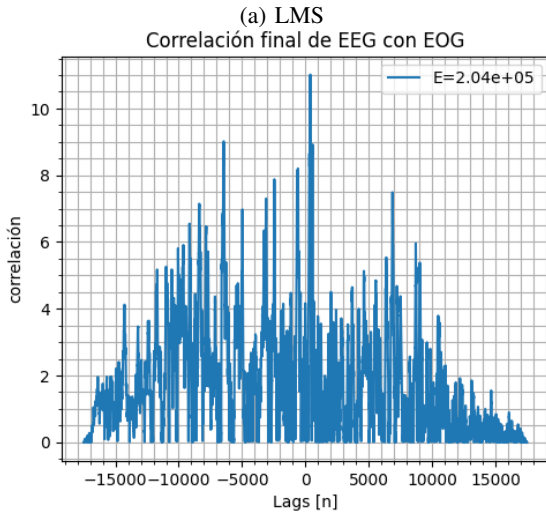
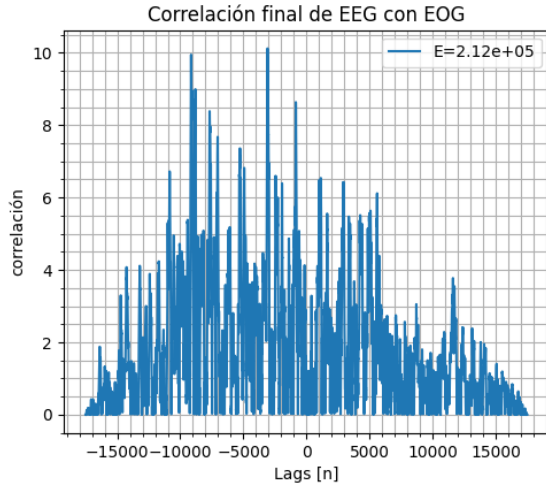


Figura 10: Correlación cruzada entre el EEG y el ECG.



(c) QLMS

Figura 11: Correlación cruzada entre el EEG y el EOG.

En las Tablas II y III se presenta una comparación de los valores utilizados y los resultados obtenidos. En esta tabla, se destaca que el algoritmo que demostró un mejor rendimiento

fue el de Q-LMS, ya que su Factor de Mejora (FM) resultó notablemente superior al de los otros algoritmos.

ECG	$M_{ECG}$	$\mu_{ECG}$	$\delta_{ECG}$	$\beta$	$\gamma$	$FM_{ECG}$
LMS	15	$1 \times 10^{-7}$	-	-	-	2,96
NLMS	15	$2 \times 10^{-4}$	$1 \times 10^{-10}$	-	-	2,03
Q-LMS	15	$1 \times 10^{-8}$	-	0,5	$5 \times 10^{-3}$	3,51

Tabla II: Parámetros utilizados y resultados para ECG

EOG	$M_{EOG}$	$\mu_{EOG}$	$\delta_{EOG}$	$\beta$	$\gamma$	$FM_{EOG}$
LMS	750	$1 \times 10^{-7}$	-	-	-	13,54
NLMS	750	$5 \times 10^{-2}$	$1 \times 10^{-10}$	-	-	14,09
Q-LMS	750	$1 \times 10^{-7}$	-	0,5	$5 \times 10^{-3}$	18,26

Tabla III: Parámetros utilizados y resultados para EOG

#### IV. CONCLUSIONES

En el curso de este análisis, se evaluaron tres algoritmos de cancelación de interferencia, Q-LMS, NLMS y LMS, en el contexto de la eliminación de artefactos en señales de EEG. La comparación entre los algoritmos se realizó a través de una métrica propia que resultó ser muy útil en el análisis y mostró gran coherencia con los resultados obtenidos.

Uno de los hallazgos clave de este estudio es que el algoritmo Q-LMS demostró un rendimiento superior, lo cual es evidente tanto a través de la observación visual de las señales procesadas como a través de la evaluación cuantitativa mediante el factor de mejora.

Una característica distintiva de NLMS es su capacidad para mantener la estabilidad de los coeficientes ante cambios bruscos en la señal. Esto se debe a su normalización con la energía de la señal, lo que evita alteraciones significativas.

En cuanto a las comparaciones entre LMS y Q-LMS, los resultados revelaron una similitud destacada en términos de valores y formas de las curvas resultantes. Además, las tres variantes, Q-LMS, NLMS y LMS, exhibieron una velocidad de convergencia similar en el proceso de cancelación de interferencia.

Sin embargo, es importante destacar que, aunque NLMS mostró una buena adaptación a cambios bruscos en la señal de EOG en comparación con LMS, también demostró una capacidad de adaptación inferior cuando se trataba de picos más pequeños generados por el ECG, en base al FM.

Esto resalta la complejidad de la cancelación de interferencia en señales de EEG y la importancia de seleccionar cuidadosamente el algoritmo adecuado en función de las características específicas de la señal.

Una dirección prometedora para futuras investigaciones es la implementación de un algoritmo de cancelación de interferencia que combine los resultados sobresalientes de los tres algoritmos evaluados en este estudio: Q-LMS, NLMS y LMS. Esto permitiría aprovechar las ventajas de cada algoritmo en un enfoque integrado para mejorar la calidad de las señales de EEG.

Específicamente, se sugiere aplicar Q-LMS con una normalización de energía para el ECG. Esto podría ayudar a mantener la estabilidad de los coeficientes del filtro, evitando variaciones abruptas.

## REFERENCIAS

- [1] L. Cutchen, "Diagnostic testing for sleep disorders," *University of New Mexico Hospitals Sleep Disorders Center*, 2016, november 11, 2016. [Online]. Available: <https://hsc.unm.edu/medicine/departments/pediatrics/divisions/continuum-of-care/pdf/cutchen--diagnostic-testing-for-sleep-disorders.pdf>
- [2] V. Ingle, S. Kogon, and D. Manolakis, *Statistical and Adaptive Signal Processing*, 2005.
- [3] G. Moody. (1999) Mit-bih polysomnographic database. Available at <https://physionet.org/content/slpdb/1.0.0/>. PhysioNet. Published: Aug. 3, 1999. Version: 1.0.0. [Online]. Available: <https://physionet.org/content/slpdb/1.0.0/>
- [4] A. G. Correa, E. Laciari, H. D. Patiño, and M. E. Valentinuzzi, "Artifact removal from eeg signals using adaptive filters in cascade," *Journal of Physics: Conference Series*, vol. 90, p. 012081, 2007, published under licence by IOP Publishing Ltd, 16TH ARGENTINE BIOENGINEERING CONGRESS (SABI 2007) AND THE 5TH CONFERENCE OF CLINICAL ENGINEERING, 26–28 September 2007, Alkazar Hotel, San Juan, Argentina. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1742-6596/90/1/012081>
- [5] M. Arif, I. Naseem, M. Moinuddin, S. S. Khan, and M. M. Ammar, "Adaptive noise cancellation using q-lms," in *2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT)*, 2017, pp. 1–4. [Online]. Available: <https://ieeexplore.ieee.org/document/7916527>
- [6] J. Hatala, "Artefacts removal from brain eeg signals using adaptive algorithms," 2021. [Online]. Available: <https://dspace.vutbr.cz/bitstream/handle/11012/211095/final-thesis.pdf>