

Análisis estadístico de señales

Juan Esteban Pineda Lopera*, Marly Alejandra Ortega†, Elisabeth Gómez Cano‡

*†‡Universidad de Antioquia, Medellín, Colombia

*jesteban.pineda1@udea.edu.co, †alejandra.ortega1@udea.edu.co, ‡elisabeth.gomezcano@udea.edu.co

I. INTRODUCCIÓN

Las bioseñales como el electrocardiograma (ECG) y la electroencefalografía (EEG), permiten identificar cuantitativamente fenómenos fisiológicos, siendo importantes en entornos clínicos. Por su naturaleza estocástica y no estacionaria, estas señales requieren métodos estadísticos específicos para evaluar su comportamiento en el tiempo y su capacidad para diagnosticar. Este trabajo, presenta un análisis estadístico de señales ECG, enfocado en su comportamiento en el tiempo y sus características de ciclos cardíacos y de señales EEG multicanal, comparando sujetos control con pacientes con enfermedad de Parkinson. El objetivo es determinar si la energía promedio por canal, muestra diferencias entre ambos grupos poblacionales, estableciendo bases para la diferenciación entre ambas condiciones, para así apoyar un diagnóstico temprano y el monitoreo de la enfermedad.

I-A. Análisis estadístico de EEG.

Se procesaron señales EEG multicanal para todos los sujetos de la muestra control y la muestra parkinson. Para cada sujeto y canal se calculó la energía promedio definida como la suma de los cuadrados de la señal, promediada entre todas las épocas del sujeto. El procedimiento se repitió para todos los sujetos del grupo control y del grupo parkinson. El objetivo era determinar si habían diferencias significativas en la energía promedio entre ambos grupos poblacionales. Se utilizó la prueba de Shapiro-Wilk para determinar si cada grupo seguía una distribución normal, en la cual al menos uno de los grupos no cumplía con el supuesto de normalidad, ya que los valores de $p < 0.05$, pero en casi todos los canales se cumplió con el supuesto de homocedasticidad, con valores de $p > 0.85$ en la prueba de Levene, debido a esto se hicieron comparaciones mediante la prueba no paramétrica de Mann-Whitney U, la cual dio como resultado no presentar diferencias significativas entre ambos grupos analizados.

	Resultados control	Resultados Parkinson
Canal	Energía promedio	
1	11.007337	12.174836
2	11.893459	12.690556
3	12.922938	14.732362
4	11.441595	13.003361
5	11.913825	12.894515
6	41.759198	56.293409
7	35.174209	47.129171
8	46.265497	61.770295

Cuadro I: Resultados del cálculo de la energía promedio por canal.

	Canal 1	Canal 2
Prueba de Normalidad (Shapiro-Wilk)	No	No
p-valor Normalidad	0.0063	0.0040
Prueba de Homocedasticidad (Levene)	Sí	Sí
p-valor Homocedasticidad	0.8848	0.9591
p-valor Mann-Whitney U	0.4057	0.5705
Diferencia Significativa	No	No

Cuadro II: Comparación de la energía promedio entre grupos de los canales 1 y 2.

La energía promedio, obtenida al calcular la suma de los cuadrados de la señal y promediar a lo largo de todas las épocas, constituye una medida general del contenido energético de la señal. Sin embargo, al tratarse de un valor global, esta métrica podría no ser lo suficientemente sensible para reflejar las alteraciones neurofisiológicas sutiles que caracterizan a enfermedades como el Parkinson. Algunos estudios sugieren que, por su naturaleza integradora, esta medida tiende a suavizar o incluso a ocultar diferencias relevantes entre sujetos sanos y pacientes, lo que limita su utilidad para fines discriminatorios en contextos clínicos [1].

Diversos estudios recientes han resaltado la importancia de analizar características específicas del EEG, más allá de medidas globales como la energía promedio. En particular, el análisis por bandas de frecuencia ha demostrado ser especialmente informativo en poblaciones con enfermedad de Parkinson. Se ha reportado una disminución en la potencia de las bandas beta y gamma, especialmente en regiones frontales y centrales, lo cual sugiere una alteración en la actividad cortical que no necesariamente queda reflejada en indicadores agregados como la energía promedio. Estos hallazgos sugieren que métricas más específicas, tanto espectrales como no lineales, pueden ofrecer una mayor sensibilidad para detectar alteraciones cerebrales sutiles asociadas con esta enfermedad [2].

I-B. Análisis estadístico de ECG.

Se tomaron 2 señales de ECG, antes y después de un proceso de filtrado. Fueron tomadas con una frecuencia de muestreo de 1024 Hz durante 30 segundos, correspondientes a 30.720 muestras y aproximadamente 36 ciclos cardíacos.

Primero, se desea conocer la naturaleza del filtro aplicado, por lo que se analizan ambas señales en función del tiempo, como se ven en la Figura 1. Se evidencia la eliminación de las tendencias lineales, centrando las amplitudes en torno al cero y eliminando oscilaciones de baja frecuencia. Esto se puede evidenciar más claramente al observar el comportamiento de las señales en el dominio de la frecuencia en la Figura

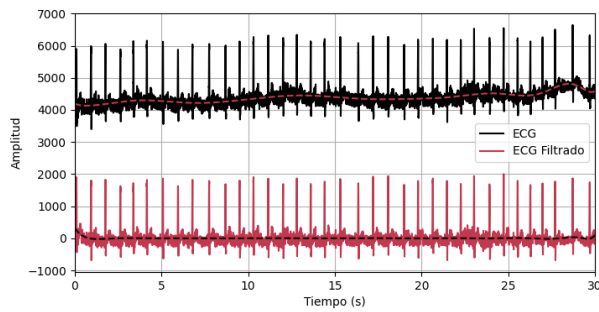


Figura 1: Comparación en función del tiempo de la señal de ECG antes y después del filtrado.

2, donde se evidencia el uso de un filtro pasa bandas con frecuencias de corte 0.5Hz y 150Hz, conservando todas las frecuencias intermedias, incluyendo el característico pico de 60Hz correspondiente al ruido de la red eléctrica.

El ancho de banda resultante entra dentro de lo recomendado por el Instituto Nacional Estadounidense de Estandares (ANSI). La frecuencia de corte superior de 150Hz es la mínima para mantener los errores de amplitud por debajo de $25\mu V$ en más del 95 % de los casos en sujetos adolescentes y adultos. Similarmente, se recomienda una frecuencia de corte inferior de 0,05Hz, sin embargo, este puede subir hasta 0,67 Hz para filtros digitales lineales con cero distorsión de fase [3].

Estas frecuencias de corte son necesarias para eliminar los artefactos de ECG más comunes de las señales de interés. Entre dichos artefactos, se encuentran las interferencias de baja frecuencia (*Baseline Wander* o desviación de la línea base), causadas por mala conexión de electrodos, electrolito seco, cremas en la piel o excesivo movimiento del pecho durante la respiración. Por otro lado, los ruidos de alta frecuencia corresponden a interferencias por la línea de poder (que no se eliminaron en este filtrado), artefactos musculares, ruido gaussiano e interferencias electromagnéticas [4].

Posteriormente, se analizan los estadísticos (Promedio, Valor Cuadrático Medio (RMS), Varianza y Desviación estándar) para un ciclo cardíaco de ambas señales, y se registran en el Cuadro III. Se evidencia mayor cambio en las medidas de tendencia central (Promedio y valor RMS), esto debido a la eliminación de tendencias lineales y la consecuente centralización de la señal respecto a cero, mientras que las medidas de dispersión (Varianza y desviación estándar) conservan valores muy similares, demostrando que el filtro altera la posición del centro de la señal, pero no tanto la manera en que los datos se agrupan respecto a este.

A continuación, se estudia si la señal filtrada es estacionaria, inicialmente, desde sus valores estadísticos y sus gráficos. Al observar los estadísticos segregados por ciclos cardíacos, evidenciados en el Cuadro IV, se nota poco rango de variabilidad entre estos, que mantienen valores en escalas muy similares. De manera similar, al observar sus respectivos gráficos en la Figura 3, aparentan tener picos en instantes cercanos, con

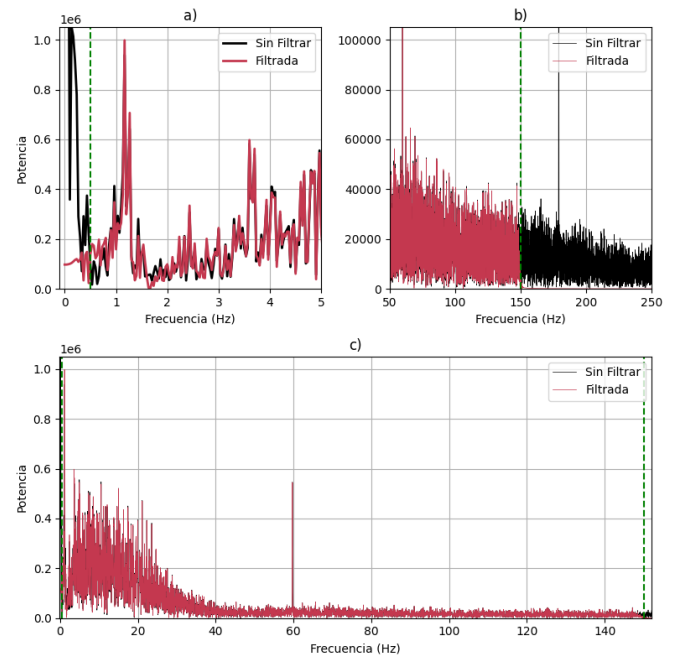


Figura 2: a) Evidencia del filtro pasa altas, con marcador en 0.5Hz. b) Evidencia del filtro pasa bajas, con marcador en 150Hz. c) Comparativa de ambas señales en su totalidad, con marcas indicando el filtro pasa banda.

Estadísticos	Señal cruda	Señal filtrada
Promedio	4262,22	-0,32
Valor RMS	4272,98	303,04
Varianza	91835,53	91832,35
Desviación estándar	303,04	303,04

Cuadro III: Comparación de estadísticos para la un ciclo de la señal de ECG antes y después del filtrado.

amplitudes y formas parecidas, lo que indica estacionariedad en la señal.

A continuación se aplican pruebas de hipótesis entre 2 ciclos aleatorios de la señal para verificar el supuesto de estacionariedad de la misma. Primero se realiza la prueba de normalidad de Shapiro-Wilk, debido a la poca cantidad de datos [5]. A continuación, se aplica la prueba de Levene de homocedasticidad y, finalmente, si se cumple normalidad en ambos ciclos y homocedasticidad, se aplica una prueba t de Student, de lo contrario, se utiliza una U de Mann-Whitney. Se aplica este procedimiento dos veces con 4 señales distintas, llegando a resultados no concluyentes: En una de las comparaciones, la U de Mann-Whitney arroja una $p = 5,04 \times 10^{-5}$, por lo que con un $\alpha = 0,05$, se rechaza la hipótesis nula, aceptando diferencia estadísticamente significativa entre las medianas de los dos ciclos y rechazando estacionariedad; sin embargo, en la otra comparación, se obtuvo una $p = 0,80$, llegando a conclusiones totalmente contrarias e impidiendo definir la naturaleza de la señal.

Por lo tanto, es necesario aplicar una prueba Dickey-Fuller, cuya hipótesis nula es que no hay estacionariedad. Se requiere

Ciclo	Promedio	Varianza	RMS	Desviación estándar
5	9.95	78658.08	280.64	280.46
6	12.07	87415.92	295.91	295.66
7	-0.32	91832.36	303.04	303.04
8	-0.88	75442.26	274.67	274.67
9	13.02	85922.43	293.41	293.12
11	-1.29	77666.65	278.69	278.69
15	1.58	81697.77	285.83	285.83
24	12.32	83893.60	289.91	289.64
25	0.94	88099.79	296.82	296.82
26	-3.42	83559.66	289.09	289.07
27	-0.51	78862.06	280.82	280.82
32	-12.76	75019.07	274.19	273.9
33	7.80	74090.75	272.31	272.2
34	-29.90	83672.58	290.80	289.26
36	-9.99	97263.27	312.03	311.87

Cuadro IV: Comparación de estadísticos para 15 ciclos cardíacos aleatorios de la señal filtrada.

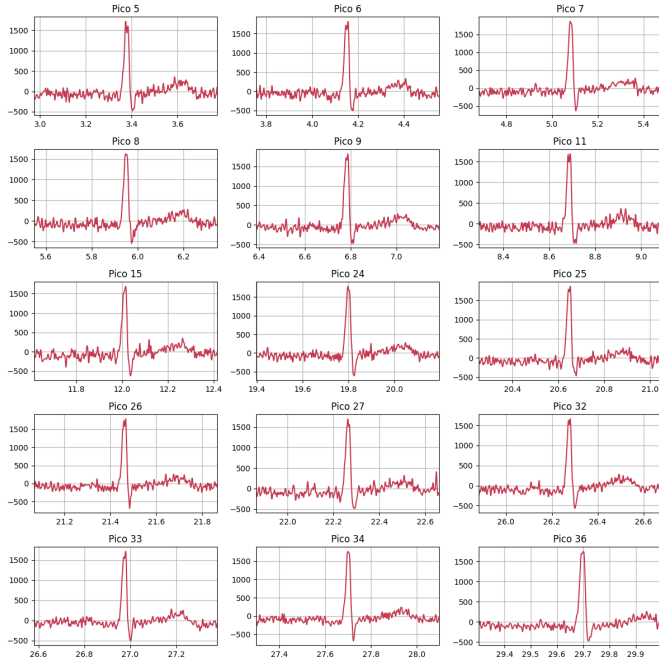


Figura 3: Gráficos de 15 ciclos cardíacos aleatorios de la señal filtrada.

que su estadístico sea bastante negativo, que se refleja en un valor p cercano a cero, para rechazar la hipótesis nula [6]. En este caso, se obtuvo un estadístico $ADF = -26,61$ y un $p \approx 0$, por lo que se rechaza la hipótesis nula y queda comprobado que la señal de ECG filtrada es estacionaria.

II. CONCLUSIONES

- El proceso de filtrado de la señal ECG fue eficaz para eliminar el ruido de alta frecuencia, manteniendo la forma y estructura de los ciclos cardíacos, como se evidenció tanto en las gráficas como en los valores estadísticos.
- Los análisis estadísticos hechos demostraron que la señal ECG filtrada es estacionaria, con características estables entre ciclos, lo cual puede garantizar confiabilidad.

- Los resultados indican que aunque se cumple el supuesto de homocedasticidad, los datos no siguen una distribución normal, lo que limita el uso de pruebas paramétricas.
- A pesar de la homogeneidad de las varianzas, las pruebas no paramétricas mostraron que no hay diferencias significativas entre los grupos estudiados, siendo necesario realizar la exploración de diferentes métricas y bandas de frecuencia para una comprensión más completa.
- Dado que los resultados no mostraron diferencias significativas en la energía promedio entre los grupos analizados, es necesario realizar estudios con un tamaño de muestras mayor para aumentar la potencia estadística y la generalización de los resultados.

REFERENCIAS

- [1] A. Pal, M. Behari, V. Goyal, and R. Sharma, "Study of EEG microstates in Parkinson's disease: a potential biomarker?" *Cognitive Neurodynamics*, vol. 15, no. 3, pp. 463–471, Jun. 2021. [Online]. Available: <https://doi.org/10.1007/s11571-020-09643-0>
- [2] M. Aljalal, S. A. Aldosari, K. AlSharabi, A. M. Abdurraqueeb, and F. A. Alturki, "Parkinson's Disease Detection from Resting-State EEG Signals Using Common Spatial Pattern, Entropy, and Machine Learning Techniques," *Diagnostics*, vol. 12, no. 5, p. 1033, Apr. 2022. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9139946/>
- [3] P. Kligfield, L. S. Gettes, J. J. Bailey, R. Childers, B. J. Deal, E. W. Hancock, G. Van Herpen, J. A. Kors, P. Macfarlane, D. M. Mirvis, O. Pahlm, P. Rautaharju, and G. S. Wagner, "Recommendations for the Standardization and Interpretation of the Electrocardiogram: Part I: The Electrocardiogram and Its Technology: A Scientific Statement From the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society *Endorsed by the International Society for Computerized Electrocardiology*," *Circulation*, vol. 115, no. 10, pp. 1306–1324, Mar. 2007. [Online]. Available: <https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.106.180200>
- [4] T. Anbalagan, M. K. Nath, D. Vijayalakshmi, and A. Anbalagan, "Analysis of various techniques for ECG signal in healthcare, past, present, and future," *Biomedical Engineering Advances*, vol. 6, p. 100089, Nov. 2023. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2667099223000191>
- [5] "Contrastes de Normalidad." [Online]. Available: <http://www.ub.edu/aplica;nforspss/cap5-6.htm>
- [6] Z. Bobbitt, "Augmented Dickey-Fuller Test in Python (With Example)," May 2021. [Online]. Available: <https://www.statology.org/dickey-fuller-test-python/>