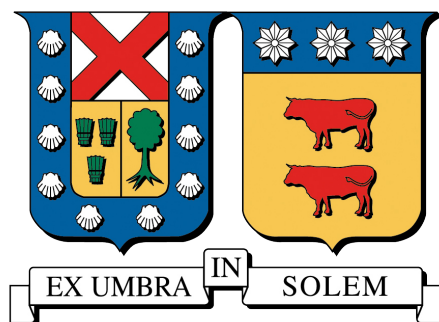


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE ELECTRÓNICA
VALPARAÍSO - CHILE



**" ESTIMACIÓN DE FUNCIÓN RESPUESTA HEMODINÁMICA A UNA
MUESTRA AMPLIA DE SUJETOS SANOS ".**

IGNACIO LUIS GODOY FLORES

**MEMORIA DE TITULACIÓN PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN ELECTRÓNICA**

PROFESOR GUÍA : ALEJANDRO WEINSTEIN
PROFESOR CORREFERENTE : STÉREN CHABERT

OCTUBRE 2025

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo:

Estimación de función respuesta hemodinámica a una muestra amplia de sujetos sanos.

Nombre del candidato(a):

Ignacio Luis Godoy Flores

Carrera / Grado:

Ingeniería Civil Electrónica

Campus: Casa Central

Departamento:

Electrónica

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Alejandro Weinstein O., en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO** contiene información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (marcar una opción):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 4-11-2025

Firma: A. Weinstein O.

Estudiante o Candidato(a):

Fecha: 04-11-2025

Firma: Ignacio Godoy

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

*Dedicado a mis Abuelos, Raúl, Olivia y Elvira
y a mis gatitos Constantine, Agnes y Martina.
Siempre los amaré*

*“Ain’t about how fast I get there
Ain’t about what’s waiting on the other side
It’s the climb”
Hannah Montana*

Agradecimientos

Quisiera agradecer en primera instancia a mis profesores, Alejandro W. y Steren C., por brindarme la oportunidad de indagar en un área de la electrónica totalmente nueva.

A mis padres, Miguel y Karen, por darme todo y más, por apoyarme incondicionalmente durante toda mi vida y por confiar a ciegas en mí para seguir este camino que estoy construyendo.

A mi hermano Diego, por ser mi alegría, por ser el mejor hermano que pude tener y por nuestro mutuo apoyo constante y complicidad. Que no dude que siempre estaré para él incondicionalmente .

A toda mi familia —tías, tíos, primos, primas, sobrinas— por acompañarme en mi crecimiento personal y profesional, por nunca dudar de mis capacidades y por toda la ayuda que me han brindado.

En particular, quiero agradecer a:

Mi prima Claudia, por ser como una hermana mayor, por siempre ayudarme a mí y a mi familia sin esperar nada a cambio, y por hacerse cargo de cosas que yo no podría haber sostenido solo.

Mi prima Karin, por preocuparse siempre por mí, por cada comida y detalle, por siempre pensar en mí, por lo atenta que es conmigo, por siempre estar ahí para escucharme y apoyarme.

Mi prima Catalina, por ser mi amiga, confidente e íntima, y por estar siempre ahí para contenerme en mis penas. porque sé que aunque no nos veamos todos los días nuestra vibra sigue intacta.

Mi tía Ximena, por pensar en nosotros, por sus visitas y por estar siempre disponible para lo que necesitemos. A mi Antito, Amparito y Lorenzita, que me alegran los días con sus locuras, que nada ni nadie les quite esa luz inmensa que tienen.

A mis amigas del colegio, Cynthia, Leslie e Isidora, por hacer mis años escolares divertidos y llenos de anécdotas. A mi profesora Macarena, por ser un modelo a seguir, por impulsarme a seguir el área de la ingeniería y por inspirarme con sus experiencias a conocer el mundo. Al tío Vicente, por seguir saludándome y conversando incluso después de salir del colegio, preguntando por mis estudios y animándome en mi carrera.

Finalmente, agradezco a mis amigos de la universidad, quienes han sido fundamentales para llegar hasta aquí. En particular:

A Mireya, por las tardes de estudio y escuchar a Billie en loop. a Betsabé, por años de amistad y su constante disposición a ayudar. A Sebastián y Paula, por todas las conversaciones que me motivaron a seguir esforzándome y por compartir cafecitos mientras discutíamos inquietudes. A Matías, Diego y Britney, por siempre ayudarme con mis dudas y su increíble disposición para apoyar a quienes lo necesitan. A Annette, por ser una amiga incondicional y una compañera increíble en la inmensidad de ramos que compartimos.

A las Amikas, por haberme acompañado todas las noches de estudio con sus podcasts.

Resumen

La función respuesta hemodinámica (HRF) constituye un elemento central en el análisis de señales de resonancia magnética funcional (fMRI), dado que describe la relación entre la actividad neuronal y las variaciones en el flujo sanguíneo cerebral. Sin embargo, determinar una HRF representativa en sujetos sanos sigue siendo un desafío, debido a la variabilidad interindividual y a las limitaciones de los modelos canónicos.

El objetivo general de este trabajo es estimar un conjunto representativo de funciones de respuesta hemodinámica (HRFs) en una muestra amplia de sujetos sanos. Para ello, se plantearon tres objetivos específicos: (i) seleccionar un conjunto de bases de datos públicas de fMRI, (ii) desarrollar un pipeline de procesamiento para obtener HRFs a partir de dichas bases, y (iii) caracterizar las HRFs en función de los atributos de los sujetos.

La metodología se basa en el preprocesamiento de imágenes fMRI mediante SPM12 en MATLAB, junto con la implementación de un modelo Fuzzy GLM y técnicas de mínimos cuadrados para la estimación de las HRFs. Asimismo, se utiliza la función doble gamma como base analítica para describir la dinámica temporal de la respuesta.

Los resultados lograron evidenciar la variabilidad que hay intersujeto e intrasujeto, tanto en comparaciones como teslaje como estímulo, como entre datasets.

Palabras clave: fMRI, preprocesamiento, Fuzzy GLM, respuesta hemodinámica (HRF), funciones gamma, mínimos cuadrados, activación cerebral.

Abstract

The hemodynamic response function (HRF) is a central element in the analysis of functional magnetic resonance imaging (fMRI) signals, as it describes the relationship between neuronal activity and variations in cerebral blood flow. However, determining a representative HRF in healthy subjects remains a challenge due to interindividual variability and the limitations of canonical models.

The general objective of this work is to estimate a representative set of hemodynamic response functions (HRFs) from a large sample of healthy subjects. To this end, three specific objectives were defined: (i) to select a set of publicly available fMRI datasets, (ii) to develop a processing pipeline to extract HRFs from these datasets, and (iii) to characterize the HRFs as a function of subject attributes.

The methodology is based on fMRI preprocessing using SPM12 in MATLAB, combined with the implementation of a Fuzzy GLM model and least-squares techniques for HRF estimation. In addition, the double-gamma function is employed as an analytical basis to describe the temporal dynamics of the response. The results showed the variability present both between subjects and within subjects, in comparisons such as tessellation and stimulus, as well as across datasets.

Keywords: fMRI, preprocessing, Fuzzy GLM, hemodynamic response function (HRF), gamma functions, least squares, brain activation.

Índice general

Resumen	1
Abstract	2
1. Introducción	8
1.1. Motivación y contexto	8
1.2. Planteamiento del problema	8
1.3. Objetivos	9
1.3.1. Objetivo general	9
1.3.2. Objetivos específicos	9
1.4. Alcances	9
1.5. Resultados esperados	10
2. Estado del Arte	11
2.1. Función de Respuesta Hemodinámica (HRF)	11
2.2. Acoplamiento Neurovascular (NVC)	11
2.2.1. fMRI y Señal BOLD	11
2.2.2. Función de Respuesta Hemodinámica	12
2.3. HRF	13
2.4. General Linear Model (GLM)	13
2.5. Fuzzy General Linear Model (GLM)	14
2.5.1. LR-Fuzzy GLM	15
2.6. Fuzzy HRF	15
2.6.1. Construcción del Fuzzy GLM	15
2.7. Mínimos cuadrados	16
2.8. Tipos de estímulos	16
2.9. Síntesis	16
3. Datasets y Preprocesamiento	17
3.1. Datasets	17
3.1.1. Resumen	21
3.2. Preprocesamiento	22
3.2.1. Realineamiento (Realign: Estimate & Reslice)	22
3.2.2. Corrección por desfase temporal (Slice Timing)	22
3.2.3. Coregistro (Coregister: Estimate)	22
3.2.4. Segmentación (Segment)	22
3.2.5. Normalización (Normalize: Write)	23
3.2.6. Suavizado (Smooth)	23

3.2.7. Preprocesamiento mediante script en <i>MATLAB</i>	23
3.2.8. Verificación preprocesamiento	24
4. Procesamiento	26
4.1. T-umbral	26
4.1.1. Steady-state	27
4.2. Generación de Estímulos	27
4.3. Fuzzy GLM para la obtención de clusters	28
4.4. Time Course y Baseline	28
4.5. Ajuste HRF	29
4.6. Filtrado de Clusters	30
4.6.1. Evaluación de la calidad del ajuste	31
5. Resultados	32
5.1. Análisis General	33
5.1.1. Clusters Válidos	33
5.1.2. Clusters No Válidos	34
5.2. Comparaciones entre características	36
5.2.1. Comparación por Región	36
5.2.2. Comparación entre sexo	38
5.2.3. Comparación entre rangos de edad	39
5.2.4. Comparación entre hemisferios	40
5.2.5. Comparación entre lóbulos	44
5.2.6. Comparación entre teslajes	47
5.2.7. Comparación entre datasets	49
6. Conclusión	53
6.1. Trabajos futuros	54
7. Anexos	55
7.1. Análisis General	55
7.2. Tablas estadísticas	56

Índice de tablas

3.1. Resumen de características de los datasets utilizados	21
4.1. Valores críticos de T obtenidos en el análisis de primer nivel (SPM).	27
4.2. Condiciones de control utilizadas para el cálculo del <i>baseline</i> en cada conjunto de datos.	29
4.3. Valores iniciales y rangos de los parámetros de la doble función gamma según Shan et al. [1]	30
5.1. Encabezados almacenados en las tablas de resultados por cluster	32
5.2. Estadística descriptiva de los parámetros de HRF obtenidos ($N = 5092$).	33
5.3. Estadísticos descriptivos de los parámetros de la HRF	34
5.4. Resultados del test de Kruskal-Wallis para diferentes parámetros	35
5.5. Estadísticos descriptivos de los parámetros de la HRF por región	36
5.6. Shapiro-Wilk test por región	37
5.7. Resultados de la prueba de Kruskal-Wallis para las medidas de HRF entre regiones	37
5.8. Estadísticos descriptivos de los parámetros de la HRF por sexo	38
5.9. Shapiro-Wilk test por sexo	38
5.10. Comparación entre sexos (Mann-Whitney U)	39
5.11. Estadísticos descriptivos de los parámetros de la HRF por rango etario	39
5.12. Shapiro-Wilk test por rango de edad	40
5.13. Resultados del test de Kruskal-Wallis por variable	40
5.14. Hemisferios y lóbulos principales identificados con el atlas de Talairach	41
5.15. Estadísticos descriptivos de los parámetros de la HRF por hemisferio	42
5.16. Shapiro-Wilk test para hemisferios	43
5.17. Resultados del test de Kruskal-Wallis para hemisferios	43
5.18. Estadísticos descriptivos de los parámetros de la HRF por lóbulos anteriores y Midbrain	44
5.19. Estadísticos descriptivos de los parámetros de la HRF por lóbulos posteriores y Sub-lobar	45
5.20. Shapiro-Wilk test para lóbulos	46
5.21. Resultados del test de Kruskal-Wallis para cada variable por lóbulo	46
5.22. Estadísticos descriptivos de los parámetros de la HRF según Teslaje	47
5.23. Shapiro-Wilk test for each teslaje	47
5.24. Resultados del test de Kruskal-Wallis para cada variable por teslaje	48
5.25. Estadísticos descriptivos de los parámetros de la HRF según Experimento (Grupo 1)	49

5.26. Estadísticos descriptivos de los parámetros de la HRF según Experimento (Grupo 2)	50
5.27. Shapiro-Wilk test for each experiment	51
5.28. Resultados del test de Kruskal-Wallis para cada variable por experimento . . .	51
7.1. Descriptivos y pruebas de normalidad de las métricas HRF no válidas	56

Índice de figuras

2.1. HRF adaptada de Glover, citada en Chen et al [2].	12
3.1. Imágenes anatómicas sujeto 1, ds002422.	24
4.1. Figuras representativa de los estímulos generados.	28
5.1. Comparación entre las gráficas promedio y HRF particular.	34
5.2. Figuras representativa de los estímulos generados.	41
7.1. Boxplot que demuestran la distribuciones de 5092 HRFs	55

1 | Introducción

La resonancia magnética funcional (fMRI) permite estudiar de manera no invasiva la actividad funcional cerebral, reflejada por los cambios hemodinámicos asociados a la actividad neuronal. La función de respuesta hemodinámica (HRF) actúa como un biomarcador clave que describe la relación entre la actividad neuronal y la señal Blood Oxygen Level Dependent (BOLD). Sin embargo, la HRF presenta variabilidad inter- e intrasujeto, y los datos de fMRI suelen contener ruido considerable, lo que limita la eficacia de los métodos tradicionales que asumen una HRF fija.

Este documento aborda la necesidad de analizar un amplio grupo de sujetos sanos (más de 300), presentando un pipeline que integra preprocesamiento estandarizado, modelado adaptativo de la HRF y estimación probabilística mediante el modelo lineal general difuso (Fuzzy GLM). Se incluye la selección de bases de datos públicas, la caracterización de la HRF en función de los atributos de los sujetos, y la optimización de códigos previos para garantizar reproducibilidad y eficiencia.

1.1. Motivación y contexto

El estudio de la actividad cerebral mediante imágenes de resonancia magnética funcional (fMRI) ha permitido avanzar en la comprensión de los procesos cognitivos y emocionales en humanos. La caracterización de la respuesta hemodinámica (HRF) y la estimación precisa de la activación neuronal son fundamentales para interpretar correctamente los patrones de señal observados y así considerar a HRF como un biomarcador. Sin embargo, la variabilidad intersujeto e intrasujeto de la HRF y la presencia de ruido en los datos de fMRI representan un desafío importante para los métodos de análisis tradicionales. En este contexto, el uso de enfoques avanzados como el modelo lineal general difuso (*Fuzzy GLM*) ofrece una alternativa que considera la incertidumbre de la HRF y mejora la robustez del análisis estadístico.

1.2. Planteamiento del problema

Los métodos convencionales de análisis de fMRI asumen una HRF canónica fija para todos los sujetos y condiciones, lo que puede llevar a sesgos en la estimación de la activación cerebral. Además, la complejidad de los experimentos modernos y la variabilidad fisiológica requieren técnicas que puedan adaptarse a diferentes perfiles de respuesta hemodinámica sin comprometer la sensibilidad estadística. Por tanto, se plantea la necesidad de desarrollar un pipeline de

análisis que integre preprocesamiento robusto, modelado flexible de la HRF y un enfoque probabilístico para la estimación de la actividad neuronal.

1.3. Objetivos

1.3.1. Objetivo general

Estimar un conjunto representativo de funciones de respuesta hemodinámica (HRF) de una muestra amplia de sujetos sanos.

1.3.2. Objetivos específicos

1. Seleccionar un conjunto de bases de datos públicas de resonancia magnética funcional (fMRI).
2. Desarrollar un pipeline de procesamiento para obtener un conjunto de HRFs a partir de las bases de datos.
3. Caracterizar las HRFs en función de los atributos de los sujetos.

1.4. Alcances

El presente estudio se fundamenta en gran medida en el trabajo previo desarrollado por Danilo Lara [3], tomando como base sus códigos y metodologías. Los alcances del estudio, en relación con sus objetivos, incluyen:

- Parametrización de los scripts existentes, permitiendo que los análisis puedan adaptarse a distintos datasets, sujetos y condiciones experimentales.
- Optimización de los códigos originales, mejorando la eficiencia computacional y reduciendo los tiempos de procesamiento en el preprocesamiento y análisis de datos de fMRI.
- Garantizar la reproducibilidad de los resultados mediante la implementación de procedimientos estandarizados, control de versiones y documentación clara de cada paso del análisis.
- Facilitar la generalización de los métodos a nuevos experimentos, permitiendo iterar sobre sujetos, sesiones y runs sin modificaciones manuales extensivas.
- Evaluación preliminar de la calidad de los datos y filtrado de clusters con parámetros no válidos, asegurando la robustez y confiabilidad de los resultados finales.

Estos alcances reflejan cómo el estudio no solo reproduce y optimiza trabajos previos, sino que también amplía su aplicabilidad, robustez y utilidad para futuros análisis de datos de fMRI.

1.5. Resultados esperados

Se espera que la aplicación del pipeline propuesto a un conjunto más amplio de sujetos y bases de datos permita obtener un conjunto de funciones de respuesta hemodinámica (HRF) más representativo y robusto que los reportados previamente. Esto posibilitará validar los hallazgos de estudios previos, como la variación de la HRF según la región cerebral y el tipo de estímulo. Además, se anticipa que la estimación de los parámetros de la HRF (altura, time-to-peak, ancho a media altura, onset y área bajo la curva) será más precisa y confiable, favoreciendo su análisis en relación con atributos de los sujetos como edad, sexo y características del dataset.

2 | Estado del Arte

En el estado del arte se abordan varios puntos relevantes para el proceso de obtención de la función de respuesta hemodinámica (HRF). Este análisis comprende desde el origen de la HRF hasta su estimación mediante mínimos cuadrados, incluyendo el enfoque Fuzzy GLM. A continuación, se presentan las secciones necesarias para la comprensión de la estimación de la HRF.

2.1. Función de Respuesta Hemodinámica (HRF)

La estimación de la función de respuesta hemodinámica (HRF) requiere comprender los fundamentos fisiológicos y técnicos que subyacen a la señal de resonancia magnética funcional (fMRI).

2.2. Acoplamiento Neurovascular (NVC)

El acoplamiento neurovascular (NVC) es el mecanismo mediante el cual el flujo sanguíneo cerebral (CBF) se ajusta a la actividad neuronal [4]. Las neuronas y células gliales generan señales que inducen vasodilatación, mientras que células vasculares como endoteliales, pericitos y musculares lisas traducen estas señales en cambios locales de flujo sanguíneo. Este proceso asegura un suministro adecuado de oxígeno y nutrientes, esencial porque el cerebro no almacena grandes cantidades de estos elementos y depende del flujo sanguíneo dinámico para su correcto funcionamiento.

2.2.1. fMRI y Señal BOLD

La resonancia magnética funcional (fMRI) es una técnica no invasiva que permite estudiar la actividad cerebral midiendo cambios en el flujo sanguíneo y la oxigenación [5]. Introducida a principios de los años 90, ha revolucionado la neurociencia cognitiva y clínica, permitiendo mapear regiones cerebrales activas durante tareas sensoriales, motoras o cognitivas, así como estudiar procesos complejos como memoria, atención o toma de decisiones.

El principio de fMRI se basa en la señal BOLD (Blood Oxygenation Level Dependent), que refleja cambios en la oxigenación sanguínea inducidos por la actividad neuronal. La hemoglobina oxigenada es diamagnética, mientras que la desoxigenada es paramagnética, generando variaciones en la señal magnética detectada por el escáner. Durante la activación

cerebral, el flujo sanguíneo aumenta, diluyendo la hemoglobina desoxigenada y provocando un incremento en la señal BOLD. De esta manera, la fMRI utiliza indirectamente la señal BOLD para mapear la actividad neuronal, combinando información sobre metabolismo del oxígeno, flujo y volumen sanguíneo en las regiones activas [6].

2.2.2. Función de Respuesta Hemodinámica

El cambio en la señal de resonancia magnética (RM) desencadenado por la actividad neural se conoce como respuesta hemodinámica. La relación entre el estímulo $s(t)$ y la señal BOLD $Y(t)$ involucra la función de respuesta hemodinámica (HRF). Estimar o determinar la HRF es importante para la correcta interpretación de estudios neurológicos [2].

Tal como se muestra en la imagen [2.4], la HRF dispone de ciertos componentes, como el peak (una amplitud máxima que ocurre aproximadamente entre los 4 y 6 segundos), seguido de un descenso que corresponde a una disminución de la señal por debajo de la línea base, producto de un flujo sanguíneo reducido y un aumento en el volumen sanguíneo.

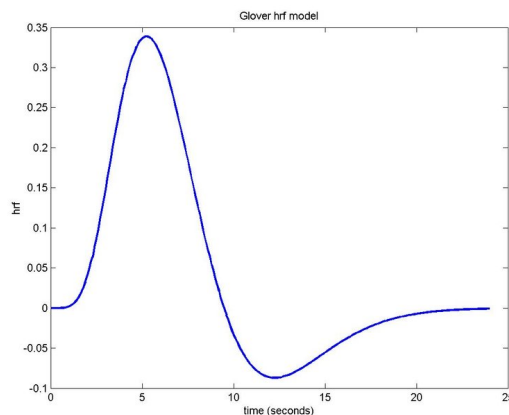


Figura 2.1: HRF adaptada de Glover, citada en Chen et al [2].

Un estímulo puede provocar una respuesta hemodinámica en el cerebro que dure aproximadamente más de 20 segundos, y su duración coincide con el tiempo de respuesta del sujeto. La forma e intensidad de esta respuesta varían según regiones cerebrales y personas. En fMRI, al investigar la actividad cerebral, es importante considerar las propiedades espaciales y temporales de la respuesta hemodinámica (HRF). En los diseños relacionados con eventos, la HRF se vincula estrechamente al estímulo, lo que permite caracterizar la actividad neuronal. Para efectos prácticos y a modo general, se asume linealidad en HRF, esto es, las respuestas de estímulos sucesivos se suman si se superponen en el tiempo.

2.3. HRF

La forma más común para modelar la HRF es la *doble gamma*, propuesta por Glover y utilizada en múltiples estudios, como el de Shan et al [1]. Este modelo combina dos funciones gamma: una excitatoria principal y otra inhibitoria de menor amplitud, lo que permite capturar tanto el pico inicial como el undershoot posterior de la señal BOLD.

En su trabajo, Shan [1] utiliza esta función doble gamma, un ejemplo de las funciones gammas utilizadas es la siguiente ecuación:

$$h(t) = \sum_{i=1}^2 A_i \frac{t^{\alpha_i-1} \beta_i^{\alpha_i} e^{-\beta_i t}}{\Gamma(\alpha_i)} \quad (2.1)$$

- A_i : amplitud del componente i . Por lo general, $i = 1$ representa el pico positivo e $i = 2$ el undershoot.
- α_i : parámetro de forma de la función gamma para el componente i , que controla la curvatura del pico.
- β_i : parámetro de escala de la función gamma para el componente i , que determina la dispersión temporal.
- $\Gamma(\alpha_i)$: función gamma de Euler, utilizada para normalizar cada componente.
- t : tiempo transcurrido desde la presentación del estímulo.

2.4. General Linear Model (GLM)

El primer acercamiento en el análisis de neuroimágenes, en específico fMRI, fue propuesto por Friston et al. (1995), quienes introdujeron el uso del General Linear Model (GLM) como marco estadístico para modelar la señal BOLD. En este enfoque, la serie temporal de cada voxel se representa como una combinación lineal de regresores, más un término de error [7]:

$$X = G\beta + \varepsilon \quad (2.2)$$

- X corresponde a la serie temporal de la señal BOLD en un voxel.
- G es la matriz de diseño que contiene los regresores experimentales, obtenidos a partir de los estímulos.
- β son los parámetros de interés, que indican la contribución de cada regresor en la señal observada.
- ε representa el término de error o ruido no explicado por el modelo.

Posteriormente, se extendió esta metodología al estudio de diseños *event-related*, mostrando cómo el GLM puede incorporar funciones base temporales para modelar la respuesta hemodinámica (HRF), logrando así una mejor caracterización de su dinámica temporal [8]. A partir del modelo general (2.2), la HRF puede estimarse al expresar la señal observada como la convolución entre la función de estímulo $s(t)$ y la respuesta hemodinámica $h(t)$:

$$x(t) = s(t) * h(t) + \varepsilon(t) \quad (2.3)$$

Seguindo a Josephs et al. (1997), la HRF se representa como una combinación lineal de un conjunto de funciones base $g_b(t)$ ponderadas por los coeficientes β_b :

$$h(t) = \sum_{b=1}^{n_{BAS}} g_b(t) \beta_b \quad (2.4)$$

Sustituyendo (2.4) en (2.3), se obtiene una formulación equivalente al modelo lineal general, que permite estimar los parámetros β mediante mínimos cuadrados ordinarios:

$$\hat{\beta} = (G^T G)^{-1} G^T X \quad (2.5)$$

Finalmente, la HRF estimada puede reconstruirse como una combinación ponderada de las funciones base utilizadas:

$$\hat{h}(t) = \sum_{b=1}^{n_{BAS}} g_b(t) \hat{\beta}_b \quad (2.6)$$

De esta forma, se deriva desde el modelo general (2.2) una formulación explícita que permite obtener la HRF a partir de la estimación de los parámetros del GLM.

2.5. Fuzzy General Linear Model (GLM)

En muchos modelos de estimación de la HRF se utiliza una forma fija de esta función. Sin embargo, dicha aproximación no refleja la variabilidad biológica real, lo que genera errores y pérdidas en la detección de activaciones. Para abordar este problema, Veloz et al. [9] proponen un GLM basado en números difusos, en el cual la convolución entre los estímulos y la HRF (que da origen a la señal BOLD) se representa mediante números difusos.

En este caso, tanto la HRF como la señal resultante se modelan como números difusos. La convolución se extiende de la siguiente forma:

$$\tilde{y}[n] = \sum_{k=-\infty}^{\infty} s[k] \tilde{h}[n - k], \quad (2.7)$$

donde $\tilde{h}[n] : \mathbb{Z} \rightarrow F(\mathbb{R})$ corresponde a la HRF difusa y $\tilde{y}[n]$ a la señal BOLD difusa. Aquí $F(\mathbb{R})$ denota el conjunto de números difusos en $\mathbb{R}$.

2.5.1. LR-Fuzzy GLM

En este estudio se emplea la representación LR de números difusos, debido a su principal ventaja: la reducción del costo computacional. Un número difuso $\tilde{h}$ no se interpreta como un único valor, sino como un conjunto de posibles valores en $\mathbb{R}$, con diferentes grados de pertenencia.

Un caso común es el número difuso triangular, definido por tres parámetros: el valor central o modal, que representa el valor más probable, la extensión izquierda, que indica la dispersión hacia valores menores, y la extensión derecha, que refleja la dispersión hacia valores mayores. La notación LR se emplea para describir la forma de la pertenencia, donde $L(\cdot)$ representa el decrecimiento hacia la izquierda y $R(\cdot)$ el decrecimiento hacia la derecha del valor modal.

2.6. Fuzzy HRF

El primer paso para construir la *fuzzy HRF* consiste en generar un conjunto de curvas plausibles de HRF utilizando la función doble γ propuesta por Glover. Esta formulación permite reproducir diferentes formas observadas empíricamente al variar los parámetros $\{\tau_1, \tau_2, \delta_1, \delta_2, c\}$.

En cada instante temporal n , se construye un intervalo que contiene todas las HRFs seleccionadas:

$$[h_\alpha^L[n], h_\alpha^U[n]] = \left[\min_i h_i[n], \max_i h_i[n] \right], \quad (2.8)$$

donde $h_i[n]$ son las HRFs candidatas. Este intervalo se modela como un número difuso triangular, cuya parte central $b[n]$ corresponde a la HRF canónica de Glover, y cuyos límites $a[n]$ y $c[n]$ se calculan como:

$$a[n] = \frac{-\alpha b[n] + h_\alpha^L[n]}{1 - \alpha}, \quad c[n] = \frac{-\alpha b[n] + h_\alpha^U[n]}{1 - \alpha}. \quad (2.9)$$

El resultado es una HRF difusa que asigna a cada instante un número difuso triangular, representando explícitamente la variabilidad fisiológica. Aunque en este trabajo se emplearon funciones triangulares por simplicidad, podrían adoptarse otras representaciones, como funciones gaussianas.

2.6.1. Construcción del Fuzzy GLM

Considerando lo anterior, la ecuación del GLM difuso queda expresada como:

$$\tilde{y}(v) = \tilde{X}\theta(v) + \varepsilon(v), \quad (2.10)$$

donde $y(v)$ es la señal observada, X la matriz de diseño, $\theta(v)$ los parámetros y $\varepsilon(v)$ el término de error.

Este modelo se denomina **LR-FGLM**, y puede escribirse en forma expandida como:

$$[y^l(v), y^c(v), y^r(v)] = [L\theta(v), X\theta(v), R\theta(v)] + \varepsilon(v), \quad (2.11)$$

donde L , X y R son matrices que organizan las componentes izquierda, central y derecha, respectivamente.

La solución estimada de los parámetros es:

$$\hat{\theta}(v) = (L^T L + X^T X + R^T R)^{-1} (L^T y(v) + X^T y^c(v) + R^T y^r(v)). \quad (2.12)$$

En síntesis, el modelo LR-FGLM permite incorporar explícitamente la variabilidad fisiológica de la HRF en el marco del GLM, reduciendo la dependencia de una forma fija y, con ello, disminuyendo los falsos negativos en la detección de activaciones. Además, gracias a la representación LR, el método mantiene un bajo costo computacional, lo que lo convierte en una alternativa robusta y eficiente frente a los enfoques clásicos basados en HRF canónica.

2.7. Mínimos cuadrados

En el análisis de series temporales de fMRI, [7] adaptan el modelo lineal general para considerar la autocorrelación introducida por la función de respuesta hemodinámica (HRF), descrito en la ecuación 2.2.

El enfoque óptimo, según el teorema de Gauss–Markov, sería *decorrelacionar* los datos multiplicando por K^{-1} y aplicar mínimos cuadrados ordinarios. Sin embargo, debido a la alta sensibilidad de esta inversión, los autores proponen un método más robusto: aplicar mínimos cuadrados directamente sobre los datos suavizados, obteniendo:

$$\hat{\beta} = (G^T G)^{-1} G^T K X, \quad (2.13)$$

con $G = KG$. Aunque no es el estimador más eficiente, resulta insesgado y más robusto ante posibles errores en la modelación de la HRF [10]. En síntesis, los mínimos cuadrados constituyen una herramienta central en el análisis de fMRI, siendo la base de la mayoría de los enfoques actuales.

2.8. Tipos de estímulos

Cabe mencionar que existen dos tipos de diseño experimental: en el **evento relacionado**, los estímulos se presentan de forma discreta y aleatoria, lo que permite analizar la respuesta cerebral a cada evento; mientras que en el **diseño en bloques**, los estímulos de una misma condición se presentan de manera continua, favoreciendo la medición de activaciones sostenidas.

2.9. Síntesis

En síntesis, la estimación de la HRF es un paso crítico en el análisis de fMRI. Aunque el GLM clásico con HRF canónica constituye una herramienta estándar, su rigidez limita la capacidad de modelar variaciones individuales. El Fuzzy GLM representa un avance que permite mayor flexibilidad y adaptabilidad, justificado por la variabilidad observada en la HRF. Por estas razones, en esta memoria se adopta el Fuzzy GLM con mínimos cuadrados como enfoque principal para estimar la respuesta hemodinámica.

3 | Datasets y Preprocesamiento

Para este estudio se utilizaron múltiples conjuntos de datos, en adelante denominados **datasets**, los cuales son de carácter público y fueron obtenidos desde la plataforma OpenNeuro. Este repositorio proporciona acceso abierto a datos de neuroimagen estructurados bajo el estándar **BIDS** (*Brain Imaging Data Structure*), lo que garantiza una organización coherente y compatible con diversas herramientas de análisis.

La selección de los datasets se realizó considerando los siguientes criterios:

- Que los sujetos fuesen sanos, sin diagnósticos neurológicos o psiquiátricos reportados.
- Que se dispusiera de imágenes funcionales (fMRI) junto con imágenes anatómicas (T1).
- Que los datos estuviesen correctamente organizados según el estándar BIDS.
- Que la información sobre los parámetros de adquisición esté completa y bien documentada, especialmente el *Time of Repetition* (TR), el número de cortes (*number of slices*) y demás variables necesarias para el preprocesamiento.
- Que existiera diversidad entre los datasets en cuanto a tareas cognitivas, características demográficas (edad, sexo), y regiones geográficas.

Una vez obtenidos los datasets se deben preprocesar siguiendo los pasos de la sección 3.2

3.1. Datasets

A continuación, se describen en detalle los distintos *datasets* utilizados en este estudio. Para cada uno se presenta una breve caracterización que incluye el tipo de tarea experimental, la cantidad de sujetos disponibles, las características demográficas, los parámetros de adquisición y cualquier particularidad relevante para su posterior procesamiento y análisis.

Dataset: ds002422 – Resting State and Arithmetic Task

Este conjunto de datos contiene mediciones de resonancia magnética funcional (fMRI) obtenidas mientras los participantes realizaban tareas de estado de reposo y aritmética.

Durante la *main condition*, los participantes observaban operaciones aritméticas ya resueltas (suma, resta, multiplicación y división con números de una o dos cifras). En la *control condition*, los estímulos consistían en líneas de dígitos con caracteres no matemáticos. Los participantes

respondían usando un dispositivo MR-compatible de tres botones, indicando si la operación era correcta, incorrecta o no era una operación aritmética.

El diseño experimental es de tipo **evento en bloques** y el dominio cognitivo evaluado corresponde a funciones **cognitivas superiores**, específicamente el procesamiento numérico y aritmético.

Característica	Valor
Cantidad de sujetos	46
Sexo	24 hombres / 22 mujeres
Edad	37.1 años
Región de origen	Rusia
Teslaje del escáner	1.5 Tesla
TR (Repetition Time)	2.1 segundos
Tipo de diseño	Estímulo en Bloques
Dominio cognitivo	Cognitivo Superior
Enlace	https://openneuro.org/datasets/ds002422

Dataset: ds005576 - 7T odor experiment

En este estudio los participantes percibían los estímulos olfativos al mismo tiempo que escuchaban o leían su correspondiente nombre. Los estímulos incluían tanto olores agradables como neutros o sin olor (odorless), y se presentaban de manera puntual a lo largo del escaneo.

El diseño experimental es de tipo **evento relacionado** y el dominio cognitivo evaluado corresponde a **percepción sensorial olfativa**, con componentes auditivos o visuales asociados a la denominación de los estímulos.

Característica	Valor
Cantidad de sujetos	25
Sexo	Desconocido
Edad (media)	Desconocido
Región de origen	Japón
Teslaje del escáner	7 Tesla
TR (Repetition Time)	2.5 segundos
Tipo de diseño	Evento relacionado
Dominio cognitivo	Sensorial (olfativo y auditivo/visual)
Enlace	https://openneuro.org/datasets/ds005576

Dataset: ds005613 - NeuroBehavioural Understanding of Language Aptitude

El dataset NEBULA101 (NeuroBehavioural Understanding of Language Aptitude) reúne datos conductuales y de neuroimagen estructural y funcional de 101 adultos sanos. Su objetivo principal fue investigar los correlatos neuronales y conductuales de la aptitud para el lenguaje y su relación con la estructura y función cerebral. Los participantes completaron diversas tareas cognitivas y lingüísticas distribuidas en cuatro sesiones: una sesión online no supervisada (cuestionarios), una online supervisada, una presencial conductual y **una sesión de resonancia magnética**.

Durante la sesión de resonancia, los participantes realizaron varias tareas de lenguaje natural y lectura en pantalla, incluyendo procesamiento semántico, fonológico y sintáctico, así como tareas de resting state y localizadores funcionales. Los eventos están anotados de manera detallada en los archivos BIDS, y las tareas fueron diseñadas como **eventos discretos** con estimulación visual.

El diseño experimental es de tipo **evento en bloques**, y los dominios cognitivos implicados corresponden al **lenguaje**, específicamente comprensión semántica, procesamiento fonológico y sintáctico, además de atención.

Característica	Valor
Cantidad de sujetos	101
Sexo	33 hombres / 68 mujeres
Edad (media)	23.35 años
Región de origen	Suiza
Teslaje del escáner	3 Tesla
TR (Repetition Time)	2.5 segundos
Tipo de diseño	Evento en bloques
Dominio cognitivo	Lenguaje (semántica, fonología, sintaxis)
Enlace	https://openneuro.org/datasets/ds005613

Dataset: ds005700 - NeuroEmo: An fMRI Dataset for Emotion Recognition

El propósito del estudio es permitir el análisis del reconocimiento de emociones utilizando estímulos culturalmente relevantes, como fragmentos de películas indias de Bollywood, en datos de fMRI recolectados de participantes indios. Este conjunto de datos explora la base neuronal de las emociones de una manera contextual y significativa, con un enfoque en la conectividad funcional y dinámica.

El dataset contiene datos fMRI crudos recolectados de 40 participantes sanos que visualizaron videoclips emocionales de películas indias. Las tareas incluyeron condiciones de reposo (*resting-state*) y tareas de evocación emocional, donde se presentaron clips audiovisuales emocionalmente cargados (por ejemplo, tristes, felices, neutrales). La estructura del dataset sigue el estándar BIDS y los eventos se codifican en archivos TSV asociados a cada run funcional.

El diseño experimental se basa en **bloques**, dado que cada clip tiene una duración prolongada y

constituye un bloque emocional coherente. El dominio cognitivo principal es el **procesamiento emocional**, con componentes visuales y auditivos (multimodalidad sensorial).

Característica	Valor
Cantidad de sujetos	40
Sexo	Desconocido
Edad (media)	Desconocido
Región de origen	India
Teslaje del escáner	3 Tesla
TR (Repetition Time)	3 segundos
Tipo de diseño	En bloques
Dominio cognitivo	Emociones (audiovisual)
Enlace	https://openneuro.org/datasets/ds005700

Dataset: ds006092 - CogRief Study

Este conjunto de datos está asociado con el manuscrito titulado " Activación cerebral en el Trastorno de Duelo Prolongado Post-Pandemia durante una tarea emocional en resonancia magnética ". Incluye datos fMRI BOLD, mapas de campo (field-maps) e imágenes estructurales T1 de 60 participantes sanos.

Durante la adquisición funcional, los participantes realizaron una Tarea Stroop Emocional de Conteo (Emotional Counting Stroop Task, ECST). Esta tarea está diseñada para evaluar la interferencia emocional en la atención y el control cognitivo. Consiste en presentar palabras con carga emocional (positiva, negativa, neutra y relacionadas con la muerte), y se solicita al participante que indique cuántas veces aparece la palabra en la pantalla, ignorando su contenido emocional.

El diseño experimental corresponde a un **diseño en bloques**, ya que cada tipo de palabra (condición emocional) se presenta durante bloques de tiempo definidos (duraciones de 29 segundos). El dominio cognitivo principal corresponde a **atención emocional** y **control cognitivo**, involucrando el conflicto atencional generado por el procesamiento de contenido afectivo.

Característica	Valor
Cantidad de sujetos	60
Sexo	26 hombres / 34 mujeres
Edad (media)	25.55 años
Región de origen	España
Teslaje del escáner	3 Tesla
TR (Repetition Time)	2.5 segundos
Tipo de diseño	En bloques
Dominio cognitivo	Atención emocional / Control cognitivo
Enlace	https://openneuro.org/datasets/ds006092

Dataset: ds006105 - Skewed Gambling Task: Deconstructing neural predictors of risky choice

En este estudio, 75 adultos diestros fueron escaneados con fMRI mientras realizaban la Skewed Gambling Task, una tarea de toma de decisiones diseñada para evaluar cómo las personas responden a opciones de apuesta con diferentes niveles de riesgo y recompensa. El objetivo fue analizar la actividad cerebral relacionada con el procesamiento del riesgo y la recompensa durante decisiones económicas.

Los participantes debían elegir entre distintas alternativas de apuestas que variaban en términos de probabilidad y magnitud de ganancia o pérdida, lo que permite estudiar mecanismos de evaluación de riesgo y preferencia por recompensas sesgadas.

El experimento se basa en un **diseño basado en eventos (event-related)**, ya que cada decisión individual (opción de apuesta) constituye un evento aislado con sus propios onsets. El dominio cognitivo principal es la **toma de decisiones bajo riesgo**, involucrando procesos de evaluación económica, anticipación de recompensas y control ejecutivo.

Característica	Valor
Cantidad de sujetos	75
Sexo	50 hombres / 25 mujeres
Edad (media $\pm$ DE)	26.31 años
Región de origen	Reino Unido
Teslaje del escáner	3 Tesla
TR (Repetition Time)	2 segundos
Tipo de diseño	Evento relacionado
Dominio cognitivo	Toma de decisiones / Evaluación de riesgo
Enlace	https://openneuro.org/datasets/ds006105

3.1.1. Resumen

Tabla 3.1: Resumen de características de los datasets utilizados

Dataset	País	Edad promedio	Hombres	Mujeres	Sujetos	Teslaje	Diseño	Dominio Cognitivo
ds002422	Rusia	37.1	52.2 %	47.8 %	46	1.5 T	En bloques	Cognitivo superior
ds005576	Japón	–	–	–	25	7 T	Evento relacionado	Sensorial (olfativo/auditivo)
ds005613	Suiza	23.35	32.7 %	67.3 %	101	3 T	En bloques	Lenguaje / Atención
ds005700	India	–	–	–	40	3 T	En bloques	Emociones (audiovisual)
ds006092	España	25.55	43.3 %	56.7 %	60	3 T	En bloques	Atención emocional / Control cognitivo
ds006105	Reino Unido	26.31	66.7 %	33.3 %	75	3 T	Evento relacionado	Toma de decisiones / Riesgo
Totales / Promedios	–	27.1	47.3 %	52.7 %	347	–	–	–

Finalmente, se tiene una edad promedio de **27.1**, con una distribución de **47.3 % hombres** y **52.7 % mujeres**, y un universo total de **347 participantes** provenientes de estudios realizados en Europa y Asia. Las tareas evaluadas comprenden una diversidad de dominios cognitivos como el control ejecutivo, el reconocimiento emocional, la atención, el aprendizaje por recompensa y la toma de decisiones bajo riesgo.

3.2. Preprocesamiento

Antes de llevar a cabo cualquier análisis funcional, es fundamental someter los datos de resonancia magnética funcional (fMRI) a un proceso riguroso de preprocesamiento. Esta etapa tiene como objetivo corregir artefactos, reducir el ruido y estandarizar los datos, de modo que las variaciones observadas en la señal BOLD reflejen de manera precisa la actividad neuronal subyacente. En particular, cuando se busca estimar la función de respuesta hemodinámica (HRF), es crucial trabajar con datos cuidadosamente preprocesados, ya que dicha función es altamente sensible a distorsiones temporales, espaciales y fisiológicas. A continuación, se describen las principales etapas aplicadas al conjunto de datos, siguiendo las directrices propuestas en el manual oficial de SPM12 y respetando la organización estructural definida por el estándar BIDS [11].

3.2.1. Realineamiento (Realign: Estimate & Reslice)

El realineamiento corrige el movimiento del sujeto entre los distintos volúmenes adquiridos durante la exploración funcional. Esta etapa estima transformaciones rígidas (traslaciones y rotaciones en los tres ejes) para alinear cada volumen con respecto a una imagen de referencia, generalmente el primer volumen de la primera corrida. Posteriormente, las imágenes son reinterpoladas para mantener la coherencia espacial. Además, este proceso genera archivos de parámetros de movimiento que pueden ser utilizados como regresores de confusión en etapas posteriores del análisis.

3.2.2. Corrección por desfase temporal (Slice Timing)

Debido a que los distintos cortes (slices) que componen un volumen funcional se adquieren en diferentes momentos dentro del mismo TR, es necesario aplicar una corrección temporal. El procedimiento de *slice timing* ajusta cada señal de voxel como si todos los cortes hubiesen sido adquiridos simultáneamente, interpolando los datos en función del orden de adquisición, el número de cortes y el TR. Esta etapa es particularmente relevante cuando se requiere una alta precisión temporal, como es el caso del modelamiento de la HRF.

3.2.3. Coregistro (Coregister: Estimate)

En esta etapa, la imagen anatómica estructural (T1) del sujeto se alinea espacialmente con la imagen funcional promedio generada durante el realineamiento. Esto permite que ambas modalidades compartan el mismo sistema de coordenadas, lo cual es crucial para las etapas de segmentación y normalización. El algoritmo de coregistro maximiza la similitud entre ambas imágenes utilizando medidas de entropía cruzada.

3.2.4. Segmentación (Segment)

La imagen anatómica, ya alineada con la funcional, es segmentada en distintos tejidos: sustancia gris, sustancia blanca y líquido cefalorraquídeo. Esta segmentación se realiza mediante un modelo probabilístico basado en plantillas poblacionales. Además, este proceso estima los

parámetros de deformación necesarios para transformar la imagen del espacio nativo del sujeto al espacio estándar MNI (Montreal Neurological Institute).

3.2.5. Normalización (Normalize: Write)

Tanto las imágenes funcionales como anatómicas son transformadas al espacio MNI utilizando los parámetros de deformación obtenidos en la segmentación. Esto permite comparar datos entre sujetos y aplicar máscaras o regiones de interés (ROIs) definidas en espacio estándar. En este estudio, se utilizó una resolución isotrópica de $2 \times 2 \times 2$ mm para las imágenes normalizadas.

3.2.6. Suavizado (Smooth)

Finalmente, se aplicó un filtro gaussiano tridimensional con un *ancho a mitad de altura* (FWHM, por sus siglas en inglés) de 6 mm. Esta operación reduce el ruido de alta frecuencia y mejora la relación señal/ruido, además de cumplir con los supuestos estadísticos del modelo lineal general (GLM) utilizados posteriormente. El suavizado también ayuda a compensar pequeñas discrepancias anatómicas residuales entre sujetos.

3.2.7. Preprocesamiento mediante script en *MATLAB*

Una vez descrito lo anterior se debe establecer que, debido a que cada dataset tiene sus propias características tales como el tiempo de repetición, cantidad de cortes, etc que son esenciales para el procesamiento en conjunto con la cantidad de sujetos y runs, se definen scripts para cada dataset en donde desde los .JSON se extraen datos tales como TR, mientras que datos como Nslices se obtienen desde el head del archivo .nii. En general se define el script como lo siguiente:

Listing 3.1: Pseudocódigo del script de preprocesamiento en SPM12

```
1. Iniciar SPM con configuracion para datos funcionales (fMRI).
2. Definir la carpeta raiz del dataset (BIDS) y listar sujetos.
3. Para cada sujeto:
  a. Identificar imagenes funcionales y anatomicas.
  b. Leer encabezado de la imagen funcional y
    extraer el numero de cortes (slices).
  c. Leer archivo JSON para obtener el TR y
    el orden de adquisicion de cortes.
  d. Construir lote de preprocesamiento
    (matlabbatch) con los siguientes pasos:

    i. Realineamiento
    ii. Slice timing
    iii. Coregistro
    iv. Segmentacion
    v. Normalizacion
```

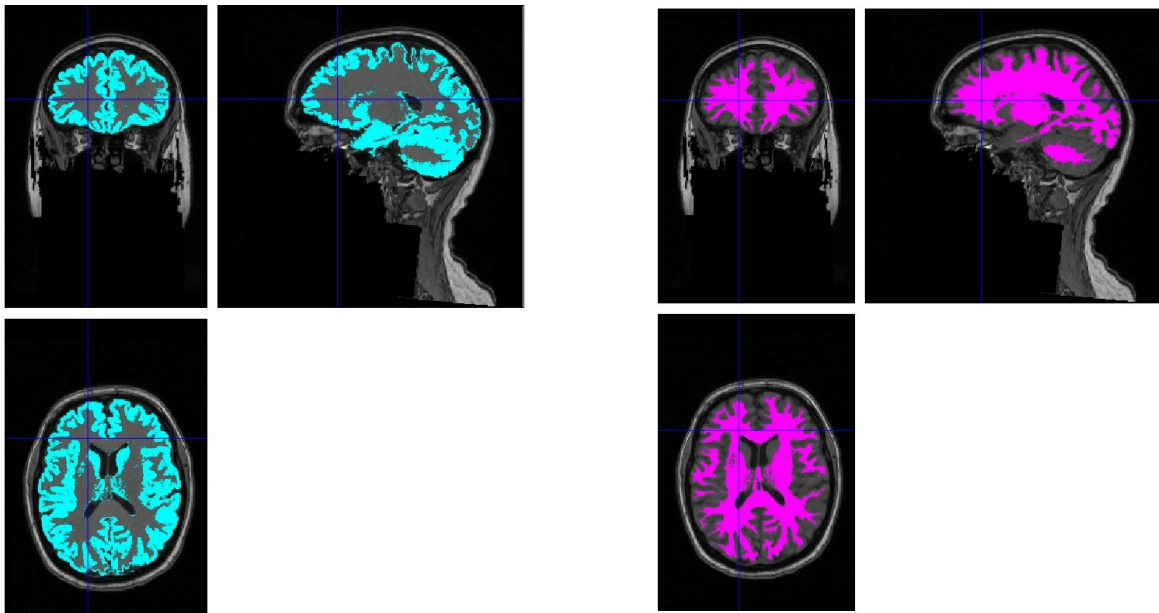
- vi. Suavizado
 - e. Ejecutar el batch con `spm_jobman`.
4. Repetir el proceso para todos los sujetos disponibles.

3.2.8. Verificación preprocesamiento

Para poder comprobar el efectivo preprocesamiento lo que incluye las correcciones de imágenes se evalúan 2 ejes para determinar si el preprocesamiento es correcto.

Verificación visual

Para comprobar visualmente la correcta alineación de las imágenes funcionales, se puede inspeccionar el volumen promedio (mean) junto con los primeros componentes principales de movimiento corregido (c1, c2, etc.). En este caso, se recomienda superponer el volumen mean con las imágenes corregidas c1 y c2 del primer sujeto del conjunto de datos. Si la alineación es adecuada, las estructuras anatómicas principales (sustancia gris, sustancia blanca, etc.) deben coincidir espacialmente entre los volúmenes superpuestos, sin desplazamientos visibles ni artefactos de movimiento. Esta inspección puede realizarse usando la herramienta SPM disponible en matlab junto a su GUI respectiva que permite un manejo cómodo de los datos.



(a) Overlay C1 - sustancia gris.

(b) Overlay C2 - sustancia blanca.

Figura 3.1: Imágenes anatómicas sujeto 1, ds002422.

Verificación automatizada mediante *script*

Además de la inspección visual, se implementó una verificación automática del preprocesamiento mediante un *script* en MATLAB. Este procedimiento permite evaluar de forma sistemática todos los sujetos del conjunto de datos, verificando parámetros clave asociados a las etapas de realineamiento, filtrado temporal y corrección de *slice timing*.

El *script* recorre las carpetas de cada sujeto (`sub-01`, `sub-02`, ...), localiza los archivos relevantes generados durante el preprocesamiento (`rp_*.txt`, `SPM.mat`, `a*.nii`), y extrae de ellos información cuantitativa. Para cada sujeto se calculan las traslaciones y rotaciones máximas estimadas durante el realineamiento, verificando si estas se mantienen dentro de un rango aceptable en relación con el tamaño de voxel. Posteriormente, se evalúa el valor del filtro de alta frecuencia (*High-pass filter*) configurado en el modelo GLM, comprobando que sea igual o superior a 128 segundos. Finalmente, se revisa la existencia de imágenes corregidas temporalmente, lo que confirma la aplicación del *slice timing*.

Los resultados obtenidos se almacenan en un archivo `.csv` que resume, para cada sujeto, las métricas de control de calidad, incluyendo movimiento máximo, parámetros del filtrado temporal y confirmación de corrección temporal. Esta estrategia permite identificar rápidamente posibles errores de preprocesamiento sin necesidad de revisar cada sujeto de forma manual.

Listing 3.2: Pseudocódigo del proceso de verificación automatizada del preprocesamiento.

```
1:  Listar carpetas de sujetos (sub-*)
2:  Para cada sujeto:
3:      Detectar tamaño de voxel desde un archivo
        funcional (.nii)
4:      Cargar rp_*.txt y calcular desplazamientos y
        rotaciones máximas
5:      Evaluar si el movimiento es aceptable
        (<= tamaño de voxel)
6:      Cargar SPM.mat y leer el parámetro High-pass filter
7:      Verificar existencia de imágenes "a*.nii"
        (Slice Timing aplicado)
8:      Guardar resultados en una fila de la tabla
9:  Fin Para
10: Guardar la tabla completa en un archivo CSV.
```

El resultado final es un archivo tabular que resume para todos los sujetos los indicadores de control de calidad: desplazamientos y rotaciones máximas, validez del parámetro de filtrado temporal y confirmación de la corrección temporal de los cortes. Este enfoque automatizado complementa la verificación visual, ofreciendo una evaluación cuantitativa y reproducible del preprocesamiento.

4 | Procesamiento

Una vez definidos los *datasets* utilizados y aplicado el preprocesamiento estándar de fMRI, el siguiente paso consiste en el procesamiento orientado a la estimación de la HRF. Este proceso busca caracterizar cómo la actividad neuronal, evocada por los estímulos experimentales, se traduce en cambios en la señal BOLD en cada sujeto y condición.

A diferencia del preprocesamiento general, que asegura la calidad y comparabilidad de los datos, este capítulo se centra en las etapas específicas necesarias para modelar la relación entre estímulos y señal BOLD. En particular, se describen cuatro etapas generales: la creación de estímulos, la construcción del *Fuzzy GLM*, el ajuste de la HRF mediante mínimos cuadrados y, finalmente, el almacenamiento de métricas asociadas a sus características relevantes.

Siguiendo el método propuesto por Shan et al. [1], este flujo de procesamiento inicia con la obtención de mapas de valores- t , que posteriormente son umbralizados con el fin de identificar las regiones activadas. A partir de dichas regiones se estima la HRF voxel a voxel, garantizando robustez frente a la variabilidad intersujeto e intercondición.

4.1. T-umbral

Los *t-umbrales* (o *t-thresholds*) permiten decidir qué valores de un mapa- t (*t-map*) en un análisis estadístico de fMRI son lo suficientemente altos como para considerarse estadísticamente significativos y no simplemente producto del ruido.

Tras estimar el *fuzzy GLM*, se obtiene un mapa tridimensional en el que cada voxel posee un valor- t que refleja su relación con la señal BOLD bajo la condición experimental. Para utilizar este mapa es necesario fijar un umbral que determine qué voxeles se consideran activos o significativos.

Si bien existen distintos tipos de umbrales, Shan et al. [1] centran su estudio en el uso del *Family-Wise Error Rate* (FWE), un enfoque más estricto que reduce la probabilidad de falsos positivos, con un nivel de significancia de $\alpha = 5\%$. Para la obtención de este umbral FWE, se realiza un análisis de primer nivel en la interfaz de SPM, procesando dos sujetos por cada dataset con el fin de calcular un umbral robusto.

La siguiente tabla resume los valores de los *t-umbrales* obtenidos bajo este criterio:

Tabla 4.1: Valores críticos de T obtenidos en el análisis de primer nivel (SPM).

Dataset	T (FWE, $p < 0.05$)
ds005576	4.79
ds002422	4.95
ds005588	4.95
ds005613	5.1
ds005700	5.0
ds006092	5.17
ds006105	4.3

4.1.1. Steady-state

Otro punto a considerar en la estimación de la HRF es asegurar que la señal BOLD se encuentre en un estado estable (*steady-state*). Los primeros volúmenes adquiridos en una sesión de fMRI suelen presentar inestabilidades lo que puede afectar el ajuste posterior del modelo. Por esta razón, se descartan los primeros scans hasta el último scan antes del primer estímulo.

En el código, esto se implementa de la siguiente forma:

Listing 4.1: código steady state

```
...
first_onset = min(te.onset);
first_vol   = floor(first_onset / TR) + 1;
ssnum       = first_vol - 1;
if ssnum < 1
    ssnum = 1;
end
...
```

Lo que hace el código 4.1 es tomar del *events.tsv* el primer dato de la columna de los onset para luego dividirlo por el TR respectivo y sumarle uno, luego para determinar la cantidad de scans a descartar al valor obtenido correspondiente al primer volumen se le resta uno, así obteniendo el valor *ssnum* el cual corresponde a la cantidad de volúmenes o scans a descartar, en el caso de que este valor sea cero, *ssnum* es igual a 1.

4.2. Generación de Estímulos

La función de estímulo genera una señal binaria en el tiempo (definida sobre vectores de tiempo en segundos y milisegundos), que toma el valor 1 durante la ocurrencia de un estímulo y 0 en ausencia de este, es decir, en los períodos libres de estimulación teniendo en consideración su partida desde el scans después de los descartados. La siguiente figura ilustra este comportamiento.

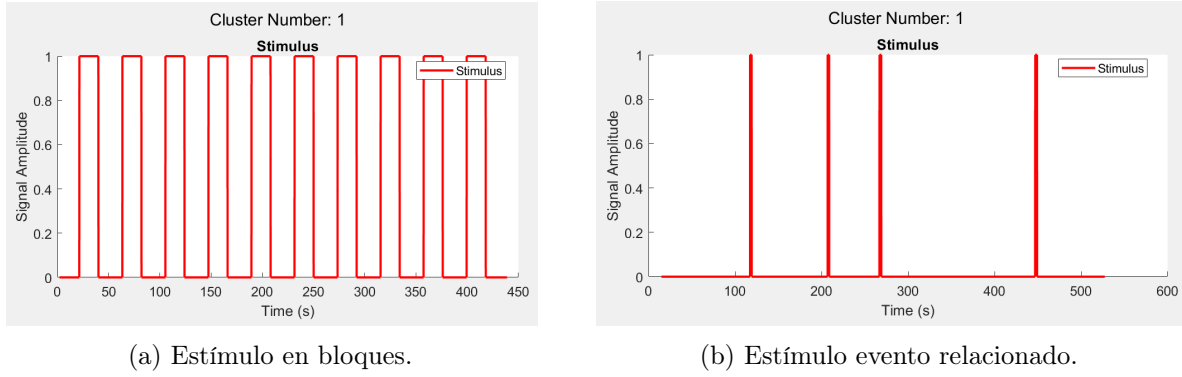


Figura 4.1: Figuras representativa de los estímulos generados.

La figura 5.2a corresponde Función de estímulo del **sujeto 1** del dataset **ds002422**, para la condición **main condition** en el **run 1 - Estímulo en bloques**, mientras que la figura 5.2b corresponde a la Función de estímulo **sujeto 2** del dataset **ds005576**, para la condición **Lemon_Lim** en el **run 3 - Estímulo evento relacionado**.

4.3. Fuzzy GLM para la obtención de clusters

El procedimiento implementado utiliza un enfoque de *Fuzzy GLM* [9] para la estimación de la respuesta hemodinámica (HRF). Este enfoque parte de la idea de que la HRF no es fija, sino que puede variar en su forma y dispersión. Por ello, en lugar de emplear una única HRF (canónica), se genera un conjunto de funciones que delimitan un rango de variabilidad de la respuesta BOLD. A partir de estas funciones, y tras su convolución con la función de estímulo, se construyen las señales BOLD esperadas en tres versiones: una HRF central, una HRF inferior y una HRF superior.

Estas señales BOLD modeladas se integran en la matriz de diseño del GLM, junto con un intercepto, lo que permite representar de manera flexible la relación entre los estímulos presentados y la señal cerebral registrada.

Finalmente, el ajuste del GLM sobre los datos observados de la señal BOLD permite estimar coeficientes asociados a cada una de estas representaciones, a partir de los cuales se calcula un estadístico t . Este valor cuantifica el grado de asociación entre la actividad cerebral medida y el estímulo experimental, considerando la incertidumbre derivada de la variabilidad en la forma de la HRF. Los mapas resultantes de valores t (t -maps) permiten posteriormente identificar regiones o clusters significativamente activados.

4.4. Time Course y Baseline

El análisis de la señal BOLD comienza con la extracción del *Time Course* (TC) de cada voxel dentro de los clusters identificados. A partir del TC de los voxels que componen cada cluster, se calcula el promedio, denominado TC_m , que representa la señal promedio del cluster.

Para identificar los clusters relevantes, se aplicó un umbral de significancia correspondiente a un alfa del 5 %, descartando aquellos clusters cuyo tamaño fuera menor a 80 [mm³]. Los clusters que cumplieron con estos criterios fueron analizados en detalle, registrando variables como el volumen del cluster, la localización de los voxels, el valor t y el time course de los voxels.

A continuación, se definió la línea basal (baseline) para normalizar la señal del cluster. En este estudio, el **baseline** se calculó a partir de los momentos asociados a la condición de control definida para cada experimento (por ejemplo, *odorless*, *intact* o *control condition*). Para cada "run", se extrajeron los valores del TC correspondientes a la condición de control y se calculó su promedio, el cual se utilizó como valor de baseline.

La normalización del time course promedio del cluster se realizó mediante la siguiente transformación:

$$TCmb = \frac{TCm - base}{base} \cdot 100 \quad (4.1)$$

donde TCm corresponde al time course promedio de los voxels del cluster y $base$ al promedio de la señal durante la condición de control.

Tabla 4.2: Condiciones de control utilizadas para el cálculo del *baseline* en cada conjunto de datos.

Dataset	Condición de control
ds002422	control_condition
ds005576	odorless
ds005613	intact
ds005700	white noise
ds006092	neutras
ds006105	symmetric

4.5. Ajuste HRF

Utilizando el método de mínimos cuadrados [1], una vez obtenido el *Time Course*, se procede a ajustar la HRF modelada mediante la doble función gamma (ver ecuación 2.1). Shan et al. [1] propone valores iniciales para cada parámetro de esta función, así como los rangos en los que pueden variar, los cuales se resumen en la siguiente tabla:

El ajuste se realiza de forma iterativa mediante un método de mínimos cuadrados con restricciones de límite para los parámetros. En cada iteración, se genera una HRF simulada a partir de los parámetros actuales y se compara con el *time course* observado, calculando un error residual. A continuación, los parámetros se ajustan dentro de los límites definidos con el objetivo de minimizar dicho error. Este proceso se repite hasta que se identifican los valores que mejor reproducen la señal medida. La **función objetivo** a minimizar corresponde al **error cuadrático total** entre la señal observada y la señal modelada, expresado como:

Tabla 4.3: Valores iniciales y rangos de los parámetros de la doble función gamma según Shan et al. [1]

Parámetro	Valor inicial	Rango
A_1	6	(0, 15)
a_1	7	(2, 10)
b_1	1	(0.5, 2)
A_2	1	(0, 10)
a_2	16	(6, 25)
b_2	1	(0, 1.5)

$$E(\theta) = \sum_{t=1}^N [y(t) - \hat{y}(t, \theta)]^2$$

donde θ representa el conjunto de parámetros libres de la HRF. Este error cuantifica la discrepancia entre la señal BOLD medida y la respuesta hemodinámica modelada.

Finalmente, los parámetros óptimos obtenidos se utilizan para construir la HRF definitiva, a partir de la cual se genera la señal estimada mediante la convolución con la secuencia de estímulos, obteniéndose el *time series* de los primeros 30 segundos con un muestreo de 0,1s.

4.6. Filtrado de Clusters

Una vez procesados los datasets y estimadas las funciones de respuesta hemodinámica (HRF) a nivel voxel, se agruparon las señales con características similares en **clusters**. Cada *cluster* representa un conjunto de voxels que comparten una morfología comparable en su respuesta hemodinámica, lo que permite reducir la variabilidad y facilitar la caracterización de patrones espaciales y funcionales comunes.

Estos *clusters* se obtuvieron mediante un proceso de agrupamiento basado en la similitud de los parámetros de la HRF (altura, tiempo al pico, ancho a media altura, área bajo la curva, y onset). De este modo, cada *cluster* puede interpretarse como una unidad representativa de un tipo particular de respuesta cerebral frente al estímulo.

Se procedió a aplicar un **filtrado de calidad** para eliminar aquellos que no cumplieran con los criterios esperados, debido principalmente a ruido, artefactos o estimaciones inestables. Los criterios de exclusión se detallan a continuación:

- **Altura mayor a 10 o igual a 0:** Se descartaron HRF que presentaran valores extremos, los cuales podrían deberse a artefactos, ruido u otras anomalías.
- **Time to Peak mayor a 10 s:** Se eliminaron funciones que tardaban excesivamente en alcanzar su máximo, indicativo de posibles artefactos o errores de estimación.
- **Time to Peak menor a 1 s y altura mayor a 4:** Se descartaron HRF con ascensos demasiado rápidos y altos, que probablemente correspondan a anomalías fisiológicas o de

registro.

- **Presencia de dos peaks significativos:** Se excluyeron funciones con un pico secundario cuya altura fuera igual o superior al 50 % del pico principal.
- **Curvas negativas predominantes al inicio:** Se eliminaron aquellas HRF con deflexiones negativas iniciales que alcanzaran al menos el 75 % de la magnitud del pico principal.

Este filtrado permitió conservar únicamente los *clusters* con HRF confiables y representativas, garantizando un análisis posterior más robusto y fisiológicamente interpretable.

Adicionalmente, se generó una tabla con las HRF que fueron **rechazadas** por alguno de los criterios anteriores. Cada registro incluye la etiqueta del *cluster* correspondiente y el motivo de exclusión, lo que permite mantener trazabilidad sobre los datos descartados.

4.6.1. Evaluación de la calidad del ajuste

La calidad del ajuste de la HRF se evaluó mediante la comparación entre el tiempo de curva observado en cada cluster y el generado por la HRF obtenida al convolucionarla con la función estímulo correspondiente. Esta discrepancia se cuantificó mediante los errores residuales, que reflejan la diferencia entre los valores observados y simulados, y su suma de cuadrados (E) indica la precisión del ajuste: valores menores de E corresponden a una HRF más precisa. Adicionalmente, se calculó el Criterio de Información de Akaike (AIC) para evaluar el desempeño del modelo considerando tanto la calidad del ajuste como su complejidad; un menor valor de AIC indica un modelo óptimo y evita el sobreajuste. Los resultados se presentan mediante estadísticas descriptivas de las HRF obtenidas, permitiendo comparar su forma y características entre clusters y condiciones experimentales.

5 | Resultados

En esta sección se presentan los resultados obtenidos tras la depuración de los *clusters* de las funciones de respuesta hemodinámica (HRF), proceso en el cual se aplicaron criterios de exclusión que consideraron valores anómalos o fisiológicamente poco plausibles. De esta manera, se logró conformar una base de datos limpia y consistente por cada *dataset*, la cual posteriormente fue integrada para obtener un análisis global. Debido a la envergadura de los datos (un total de 5.092 HRFs válidas y 4.195 HRFs no válidas), se optó por seleccionar ciertas características para generar un archivo `.tsv` según la descripción de los parámetros y luego realizar la estadística descriptiva utilizando el software *Jamovi* [12]. Este programa de análisis estadístico de código abierto permite calcular de manera eficiente medidas de tendencia central, dispersión y otras métricas descriptivas, así como generar tablas y gráficos de forma reproducible.

Tabla 5.1: Encabezados almacenados en las tablas de resultados por cluster

Encabezado	Descripción
SubjectID	Identificador del sujeto analizado
Sesion	Número o etiqueta de la sesión experimental
Run	Número del <i>run</i> o bloque funcional
Condicion	Condición experimental (p. ej., <i>PositiveSkew</i> , <i>Symmetric</i>)
Experimento	Nombre o código del experimento
Altura	Amplitud máxima de la HRF
TimeToPeak	Tiempo hasta el pico máximo de la HRF (s)
AnchoMitadAltura	Anchura temporal al 50 % de la altura máxima
Onset	Latencia de inicio de la respuesta
AreaBajoCurva	Área bajo la curva de la HRF
X	Coordenada espacial (eje izquierda–derecha)
Y	Coordenada espacial (eje posterior–anterior)
Z	Coordenada espacial (eje inferior–superior)

Del total de HRFs analizadas (9.287), aproximadamente el 54.9 % fueron consideradas válidas, mientras que el 45.1 % fueron descartadas por no cumplir los criterios de calidad.

5.1. Análisis General

5.1.1. Clusters Válidos

En la Tabla 5.2 se presentan los resultados válidos descriptivos de los parámetros de la función de respuesta hemodinámica (HRF) estimados a partir de un total de 5.902 observaciones. Estos parámetros resumen la morfología temporal de la respuesta BOLD y permiten caracterizar de manera global las diferencias entre regiones y sujetos.

En términos generales, los valores medios observados son consistentes con los rangos típicos reportados en la literatura para estudios de fMRI en sujetos sanos. La amplitud promedio de la respuesta (*Altura*) fue de 3.31, con una desviación estándar considerable ($DS = 2.14$), lo que sugiere variabilidad interregional e intersujeto en la magnitud de la activación. El tiempo al pico (*TimeToPeak*) presentó una media de 3.91 s ($DS = 2.57$), dentro del rango fisiológicamente esperado para la respuesta hemodinámica (aproximadamente 4–7 s). El ancho a media altura (*AnchoMitadAltura*) alcanzó una media de 4.45 s, reflejando la duración típica de la respuesta BOLD antes de retornar a la línea base.

El parámetro *Onset*, con una media de 1.27 s, indica el retardo promedio entre el estímulo y el inicio de la respuesta, valor coherente con los tiempos de latencia neurovascular conocidos. Por su parte, el área bajo la curva (*AreaBajoCurva*) mostró una media de 11.32, aunque con alta dispersión ($DS = 4.79$), lo que podría asociarse con diferencias en la magnitud de la activación según la región cortical o el tipo de tarea.

En conjunto, estos resultados permiten establecer un perfil promedio de la HRF a nivel global, sirviendo como referencia para comparaciones entre grupos, regiones cerebrales o condiciones experimentales en los análisis posteriores.

Tabla 5.2: Estadística descriptiva de los parámetros de HRF obtenidos ($N = 5092$).

Parámetro	Media	DS	Mínimo	Máximo	N
Altura	3.31	2.14	1.25e-4	8.14	5092
TimeToPeak [s]	3.91	2.57	0.60	10.00	5092
AnchoMitadAltura [s]	4.45	2.64	1.30	14.20	5092
Onset [s]	1.27	1.07	0.10	6.80	5092
AreaBajoCurva	11.32	4.79	3.60e-4	15.00	5092

La siguiente gráfica muestra una representación de la HRF con los valores descriptivos generales y se compara con una de las HRFs que más cerca del promedio se encuentra. Esta HRF corresponde al cluster 15 del sujeto 20 del dataset **ds005700**

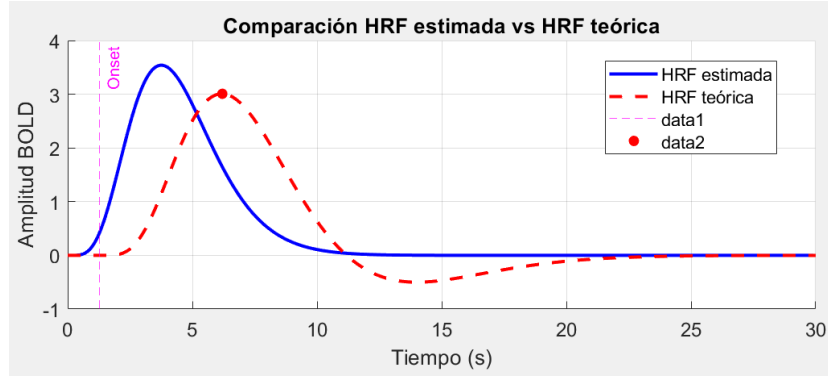


Figura 5.1: Comparación entre las gráficas promedio y HRF particular.

5.1.2. Clusters No Válidos

Tal como se mencionó al inicio de este capítulo, además de las HRFs que cumplen con los criterios establecidos, se identificaron aquellas que no satisfacen los requisitos definidos para ser consideradas válidas. Estas funciones fueron clasificadas como *no válidas* y, por tanto, descartadas del análisis principal. En total, se eliminaron 4.195s HRFs que presentaban parámetros fuera de los rangos fisiológicos esperados.

La siguiente tabla presenta las estadísticas descriptivas de estas HRFs no válidas, permitiendo observar su comportamiento general en términos de los parámetros característicos de la respuesta hemodinámica:

Tabla 5.3: Estadísticos descriptivos de los parámetros de la HRF

Parámetro	N	Media	Mediana	SD	Mínimo	Máximo
Altura	4195	0.977	0.841	1.39	1.09e-16	11.0
TimeToPeak	4195	13.052	13.500	2.85	0.500	23.6
AnchoMitadAltura	4195	11.771	12.200	2.19	1.200	18.6
Onset	4195	5.114	5.200	1.47	0.100	16.7
AreaBajoCurva	4195	9.972	10.842	4.36	0.000	15.0

Como se aprecia en los resultados, el valor promedio del parámetro **TimeToPeak** es de **13.05 segundos**, notablemente superior al rango típico de entre 4 y 7 segundos reportado para respuestas hemodinámicas normales y al umbral aplicado de **10 segundos** para filtrar activaciones tardías. Este desplazamiento temporal indica una activación anómalamente demorada de la señal BOLD, posiblemente asociada a ruido fisiológico, artefactos de movimiento o errores en la estimación de los eventos experimentales.

Por su parte, la amplitud promedio de la señal (*Altura* = **0.98**) es relativamente baja, lo que sugiere que varias respuestas presentan una intensidad reducida o una forma distorsionada respecto a la línea base. En conjunto, estos patrones respaldan la decisión de excluir dichos *clusters* del análisis final, garantizando que las HRFs consideradas válidas representen de manera más fidedigna la dinámica neurovascular esperada. Del total de HRFs analizadas, se descartaron 19 HRFs por presentar altura menor a 0.005 con mala forma, 99 HRFs por tener un time-to-peak menor a 1 con altura superior a 4, 4.062 HRFs por time-to-peak mayor a

10 y 15 HRFs cuya área bajo la curva era igual a cero. La mayoría de las HRFs descartadas correspondió al criterio de time-to-peak mayor a 10, mientras que los otros motivos afectaron a un número relativamente reducido de HRFs.

De la tabla anexada 7.1, que evidencia la curtosis, la asimetría y la normalidad de las métricas altura, TimeToPeak, AnchoMitadAltura, Onset y ÁreaBajoCurva, se observa cómo estas características se relacionan con los motivos por los cuales ciertos datos fueron descartados (por ejemplo, TimeToPeak mayor a 10, altura muy baja, etc.).

En general, la altura muestra tanto asimetrías positivas como negativas, lo que indica que su distribución puede estar sesgada hacia valores altos en algunos subgrupos y hacia valores bajos en otros. Esto refleja una variabilidad considerable en la forma de las HRF: algunas presentan picos más pronunciados, mientras que otras son muy planas o deformadas.

Por su parte, el TimeToPeak presenta asimetrías positivas en todos los subgrupos, pero con magnitudes relativamente bajas (no superiores a 1.2), lo que indica que, aunque la distribución está sesgada hacia valores más tardíos, la desviación de la simetría no es muy pronunciada. Esto sugiere que la mayoría de los TimeToPeak se concentran cerca del valor medio, con pocos casos extremadamente tempranos.

En conjunto, estas diferencias indican que la altura es mucho más variable y propensa a extremos que el TimeToPeak, cuya distribución es más homogénea y concentrada.

Finalmente, los resultados de la prueba de normalidad de Shapiro–Wilk muestran que, en la mayoría de los contextos problemáticos como altura muy baja, TimeToPeak muy rápido o muy tardío y Área = 0 *se puede rechazar la hipótesis de normalidad* (p l.t. 0.05¹). Esto sugiere que estas métricas no siguen una distribución normal y que, por lo tanto, los análisis paramétricos estándar podrían no ser totalmente adecuados para estos datos.

Tabla 5.4: Resultados del test de Kruskal-Wallis para diferentes parámetros

Parámetro	χ^2	df	p
Altura	386	3	l.t. .001
TimeToPeak	337	3	l.t. .001
AnchoMitadAltura	348	3	l.t. .001
Onset	328	3	l.t. .001
AreaBajoCurva	310	3	l.t. .001

Se aplicó el test de Kruskal-Wallis para evaluar diferencias entre los grupos de HRFs en cada parámetro medido: Altura, TimeToPeak, AnchoMitadAltura, Onset y AreaBajoCurva. Dado que los datos no cumplen con la hipótesis de normalidad, se utilizó este test no paramétrico. Los resultados muestran valores de χ^2 elevados y p “l.t. .001” en todos los parámetros, lo que indica diferencias significativas entre los grupos. Esto respalda la decisión de descartar aquellas HRFs que no cumplieran con los criterios establecidos, ya que al menos un grupo presentaba valores atípicos o distribuciones distintas.

¹l.t. significa menor que

5.2. Comparaciones entre características

5.2.1. Comparación por Región

Tabla 5.5: Estadísticos descriptivos de los parámetros de la HRF por región

Región	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
España	3.13	3.94	5.48	0.925	11.4
India	3.35	5.06	5.52	1.65	13.8
Japón	3.90	4.83	3.95	2.13	15.0
Reino Unido	4.63	3.44	3.91	1.12	15.0
Rusia	4.48	3.47	3.08	1.42	15.0
Suiza	3.41	2.28	3.34	0.553	9.54
<i>Mediana</i>					
España	1.64	3.50	6.20	0.500	12.5
India	4.74	3.10	3.00	1.20	15.0
Japón	3.89	4.60	3.60	2.10	15.0
Reino Unido	5.46	2.30	2.60	0.800	15.0
Rusia	4.52	3.40	3.00	1.40	15.0
Suiza	3.16	1.40	2.30	0.300	9.64
<i>Desviación estándar</i>					
España	2.65	2.47	2.84	0.864	3.77
India	1.74	2.63	3.22	0.723	2.70
Japón	1.24	1.78	1.21	0.860	4.91e-5
Reino Unido	1.57	2.35	2.63	0.718	0.00841
Rusia	0.343	0.569	0.723	0.146	0.00261
Suiza	1.92	2.09	2.20	0.791	3.65
<i>Mínimo</i>					
España	0.220	1.00	1.70	0.100	2.31
India	0.325	3.10	3.00	0.700	3.68
Japón	1.94	1.10	1.70	0.200	15.0
Reino Unido	1.31	2.30	2.60	0.800	15.0
Rusia	1.32	3.40	3.00	1.00	15.0
Suiza	0.240	0.900	1.90	0.100	2.03
<i>Máximo</i>					
España	8.14	10.0	14.2	3.50	15.0
India	4.74	10.0	13.2	3.70	15.0
Japón	7.85	9.20	7.20	4.10	15.0
Reino Unido	5.46	10.0	10.6	4.00	15.0
Rusia	4.52	9.90	10.6	3.20	15.0
Suiza	7.34	10.0	13.7	4.20	15.0

En el contexto del análisis por regiones o países, se dispone de un total de 1.320 HRFs correspondientes a España, 810 a India, 199 a Japón, 154 al Reino Unido, 882 a Rusia y 1.339 a Suiza.

Tabla 5.6: Shapiro-Wilk test por región

Region	Altura	TimeToPeak	AnchoMitadAltura	Onset	AUC
<i>Shapiro-Wilk W</i>					
Espana	0.799	0.893	0.909	0.738	0.849
India	0.690	0.708	0.693	0.684	0.508
Japon	0.829	0.895	0.869	0.898	0.724
Reino Unido	0.534	0.525	0.536	0.498	0.428
Rusia	0.0981	0.0933	0.0936	0.0912	0.0610
Suiza	0.956	0.592	0.680	0.556	0.945
<i>Shapiro-Wilk p</i>					
Espana	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
India	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Japon	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Reino Unido	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Rusia	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Suiza	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001

La tabla 5.6 (anexada) presenta los resultados del test de Shapiro-Wilk, en los que todos los valores de p son l.t. 0.001. Esto indica que se rechaza la hipótesis de normalidad para todas las regiones y medidas evaluadas, por lo que los datos no siguen una distribución normal. En consecuencia, para los análisis posteriores se deben considerar métodos estadísticos no paramétricos o aplicar transformaciones a los datos que permitan aproximarse a la normalidad.

Tabla 5.7: Resultados de la prueba de Kruskal-Wallis para las medidas de HRF entre regiones

Medida	χ^2	gl	p
Altura	270	5	l.t. 0.001
TimeToPeak	1276	5	l.t. 0.001
AnchoMitadAltura	802	5	l.t. 0.001
Onset	2004	5	l.t. 0.001
AreaBajoCurva	1649	5	l.t. 0.001

Se realizó la prueba de Kruskal-Wallis para comparar las medidas de la HRF entre las distintas regiones. Como se observa en la Tabla 5.7, todos los valores de p son l.t. 0.05, lo que indica que se rechaza la hipótesis nula y que existen diferencias significativas entre regiones para Altura, TimeToPeak, AnchoMitadAltura, Onset y Área Bajo la Curva.

5.2.2. Comparación entre sexo

Tabla 5.8: Estadísticos descriptivos de los parámetros de la HRF por sexo

Sexo	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
F	3.67	3.10	3.97	0.877	11.6
M	3.55	3.36	4.20	0.969	12.0
<i>Mediana</i>					
F	4.16	2.30	3.00	0.500	13.0
M	4.52	3.40	3.00	0.800	15.0
<i>Desviación estándar</i>					
F	2.10	2.22	2.50	0.827	3.81
M	2.03	2.02	2.46	0.740	3.85
<i>Mínimo</i>					
F	0.24	0.90	1.70	0.10	2.03
M	0.22	0.90	1.70	0.10	2.31
<i>Máximo</i>					
F	8.14	10.0	14.2	4.20	15.0
M	8.14	10.0	13.8	4.00	15.0

En el conjunto de datos que incluye la variable sexo, se registraron 2.148 HRFs correspondientes a participantes femeninos y 1.547 HRFs correspondientes a participantes masculinos.

Tabla 5.9: Shapiro-Wilk test por sexo

Sexo	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Shapiro-Wilk W</i>					
F	0.944	0.809	0.787	0.815	0.825
M	0.925	0.841	0.812	0.867	0.773
<i>Shapiro-Wilk p</i>					
F	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
M	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001

La tabla anterior (5.9) indica que, si bien la variable Altura muestra cierta aproximación a la normalidad, el resto de las características presentan mayores desviaciones. Sin embargo, en todos los casos, los valores de p son menores a 0.05, lo que lleva a rechazar la hipótesis de normalidad. Por lo tanto, para comparar las diferencias entre los dos grupos independientes, se recomienda utilizar un test no paramétrico, como el Mann-Whitney U.

El test de Mann-Whitney U muestra que no hay diferencias significativas entre sexos en Altura ($p = 0.359$). En cambio, TimeToPeak, AnchoMitadAltura, Onset y AreaBajoCurva presentan diferencias significativas (p l.t. .001), indicando que hombres y mujeres difieren en estas características.

Tabla 5.10: Comparación entre sexos (Mann-Whitney U)

Variable	Statistic	p
Altura	1.63e+6	0.359
TimeToPeak	1.45e+6	1.t. .001
AnchoMitadAltura	1.50e+6	1.t. .001
Onset	1.44e+6	1.t. .001
AreaBajoCurva	1.53e+6	1.t. .001

Note. $H_a \mu_F \neq \mu_M$

5.2.3. Comparación entre rangos de edad

Tabla 5.11: Estadísticos descriptivos de los parámetros de la HRF por rango etario

Rango etario	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
18-25	3.54	3.03	4.03	0.810	11.2
26-35	3.64	3.51	4.21	1.07	12.3
36-60	3.84	3.59	4.17	1.13	13.2
<i>Mediana</i>					
18-25	3.48	2.30	3.00	0.500	12.1
26-35	4.52	3.40	3.00	1.30	15.0
36-60	4.52	3.40	3.00	1.40	15.0
<i>Desviación estándar</i>					
18-25	2.14	2.17	2.50	0.800	3.91
26-35	2.03	2.18	2.58	0.786	3.82
36-60	1.94	1.98	2.45	0.713	2.92
<i>Mínimo</i>					
18-25	0.220	0.900	1.70	0.100	2.03
26-35	0.255	0.900	1.70	0.100	2.71
36-60	0.559	1.00	1.70	0.100	3.08
<i>Máximo</i>					
18-25	8.14	10.0	14.2	4.20	15.0
26-35	8.13	10.0	13.8	4.20	15.0
36-60	8.14	10.0	10.7	3.90	15.0

El número de HRFs por rango etario es de 3.456 para el grupo de 18-25 años, 661 para el grupo de 26-35 años y 459 para el grupo de 36-60 años.

La tabla 5.12 sugiere que para todos los casos se rechaza la hipótesis de normalidad lo que sugiere un análisis Kruskal-Wallis.

El test de Kruskal-Wallis indica diferencias significativas entre rangos etarios en todas las variables debido a que p es menor a 0.05 en todos los casos.

Tabla 5.12: Shapiro-Wilk test por rango de edad

Rango	Altura	TimeToPeak	AnchoMitadAltura	Onset	AUC
<i>Shapiro-Wilk W</i>					
18-25	0.942	0.812	0.809	0.786	0.850
26-35	0.912	0.827	0.777	0.879	0.732
36-60	0.884	0.781	0.758	0.856	0.667
<i>Shapiro-Wilk p</i>					
18-25	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
26-35	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
36-60	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001

Tabla 5.13: Resultados del test de Kruskal-Wallis por variable

Variable	χ^2	df	p
Altura	7.09	2	0.029
TimeToPeak	91.96	2	l.t. .001
AnchoMitadAltura	34.62	2	l.t. .001
Onset	179.03	2	l.t. .001
AreaBajoCurva	113.38	2	l.t. .001

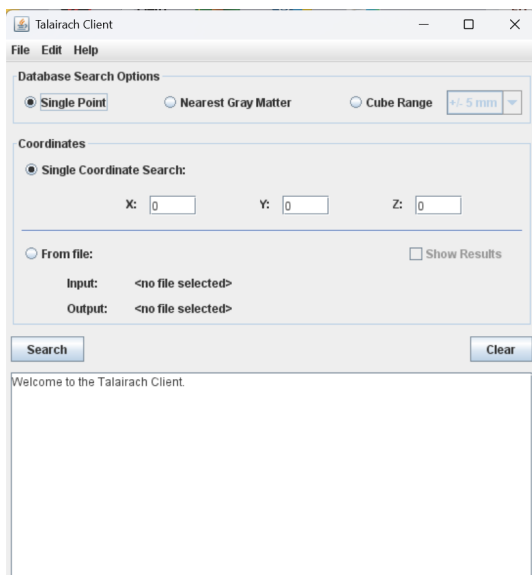
5.2.4. Comparación entre hemisferios

Para identificar los hemisferios y lóbulos en los que se localizan los *clusters* con activación cerebral, en primera instancia se consideran los vóxeles con mayor nivel de activación, ya que estos permiten determinar con precisión la ubicación anatómica de dichas regiones. Para esta clasificación se emplea la herramienta *Talairach* [13], que una vez instalada puede ejecutarse desde la línea de comandos de la siguiente forma:

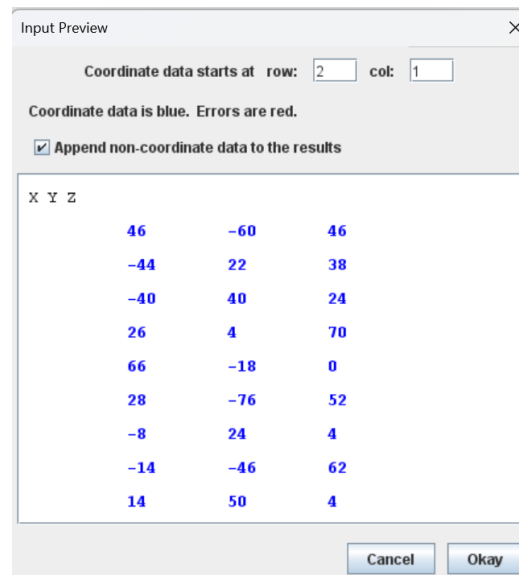
```
cd "C:\Users\godoy\Downloads\talairach"
java -jar talairach.jar
```

Cabe destacar que, para su correcta ejecución, la aplicación debe iniciarse desde la carpeta donde se ubica el archivo .jar de Talairach.

Una vez ejecutada la aplicación, se despliega la interfaz gráfica de usuario (GUI) de Talairach, la cual constituye la ventana principal de trabajo. Desde esta interfaz es posible acceder a las distintas funciones del programa, entre ellas la incorporación de tablas de dimensiones anatómicas y la consulta de localizaciones cerebrales específicas. Esta GUI permite una interacción más intuitiva con la herramienta, facilitando la clasificación y el análisis de los *clusters* de activación cerebral en función de su ubicación en el cerebro.



(a) GUI Talairach.



(b) Importación de tabla con coordenadas

Figura 5.2: Figuras representativa de los estímulos generados.

Este resulta en un archivo el cual a cada punto se le otorga una etiqueta de hemisferio, lobulo y más características, el foco de este trabajo como se menciona anteriormente es en el estudio de hrfs por hemisferios y lobulos, la siguiente tabla refleja las posibles etiquetas entregadas por Talairach.

Tabla 5.14: Hemisferios y lóbulos principales identificados con el atlas de Talairach

Nivel	Categorías
Hemisferios	Izquierdo (Left), Derecho (Right)
Regiones globales	Cerebrum, Cerebellum, Brainstem
Lóbulos	Frontal Lobe, Parietal Lobe, Temporal Lobe, Occipital Lobe, Limbic Lobe, Sub-lobar, Anterior Lobe (cerebelo), Posterior Lobe (cerebelo), Midbrain

Tabla 5.15: Estadísticos descriptivos de los parámetros de la HRF por hemisferio

Hemisferio	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
Left Brainstem	0.500	9.55	9.50	3.35	5.64
Left Cerebellum	2.88	4.06	4.65	1.31	10.4
Left Cerebrum	3.45	3.92	4.40	1.29	11.8
Right Brainstem	2.46	6.33	5.57	2.67	5.11
Right Cerebellum	2.68	4.51	5.07	1.48	10.2
Right Cerebrum	3.37	3.80	4.34	1.24	11.3
<i>Mediana</i>					
Left Brainstem	0.500	9.55	9.50	3.35	5.64
Left Cerebellum	2.75	3.10	3.00	1.20	12.4
Left Cerebrum	4.50	3.40	3.00	1.20	15.0
Right Brainstem	0.0313	8.50	7.00	3.60	0.236
Right Cerebellum	2.74	3.10	4.10	1.20	11.6
Right Cerebrum	3.95	3.40	3.00	1.20	14.2
<i>Desviación estándar</i>					
Left Brainstem	0.697	0.495	1.41	0.212	7.89
Left Cerebellum	1.98	2.53	2.50	1.10	5.07
Left Cerebrum	2.13	2.52	2.61	1.05	4.67
Right Brainstem	4.23	4.46	3.20	2.06	8.57
Right Cerebellum	1.90	2.81	2.77	1.21	5.19
Right Cerebrum	2.12	2.59	2.65	1.08	4.74
<i>Mínimo</i>					
Left Brainstem	0.00677	9.20	8.50	3.20	0.0617
Left Cerebellum	0.00568	1.00	1.70	0.100	0.0337
Left Cerebrum	0.00103	0.900	1.70	0.100	0.00685
Right Brainstem	0.00978	1.20	1.90	0.300	0.0950
Right Cerebellum	0.00160	1.00	1.90	0.100	0.0110
Right Cerebrum	1.25e-4	0.600	1.30	0.100	3.60e-4
<i>Máximo</i>					
Left Brainstem	0.993	9.90	10.5	3.50	11.2
Left Cerebellum	8.13	9.90	10.9	4.90	15.0
Left Cerebrum	8.14	10.0	13.7	6.80	15.0
Right Brainstem	7.34	9.30	7.80	4.10	15.0
Right Cerebellum	7.34	10.0	14.2	4.50	15.0
Right Cerebrum	8.13	10.0	13.8	6.50	15.0

En el conjunto de datos, el número de HRFs asociadas a cada hemisferio y estructura es el siguiente: Left Brainstem: 2, Left Cerebellum: 235, Left Cerebrum: 2.266, Right Brainstem: 3, Right Cerebellum: 258 y Right Cerebrum: 1.905. Estas cifras reflejan la cantidad de HRFs consideradas en cada región para el análisis estadístico.

Tabla 5.16: Shapiro-Wilk test para hemisferios

Region	Altura	TimeToPeak	AnchoMitadAltura	Onset	AUC
<i>Shapiro-Wilk W</i>					
Left Cerebellum	0.909	0.863	0.860	0.862	0.827
Left Cerebrum	0.936	0.858	0.817	0.874	0.723
Right Brainstem	0.752	0.823	0.850	0.847	0.757
Right Cerebellum	0.896	0.865	0.884	0.876	0.826
Right Cerebrum	0.939	0.851	0.809	0.866	0.775
<i>Shapiro-Wilk p</i>					
Left Cerebellum	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Left Cerebrum	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Right Brainstem	0.005	0.171	0.239	0.232	0.016
Right Cerebellum	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Right Cerebrum	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001

La tabla 5.16 muestra que, en la mayoría de los casos, los valores de p son menores a 0.05, indicando que se rechaza la hipótesis de normalidad. Sin embargo, en el Right Brainstem, algunas variables como TimeToPeak, AnchoMitadAltura y Onset presentan valores de p mayores a 0.05, lo que sugiere que no se rechaza la hipótesis de normalidad para esos casos. Esto indica que, en general, los datos no siguen una distribución normal, aunque ciertas estructuras pueden aproximarse a la normalidad en algunas medidas.

Tabla 5.17: Resultados del test de Kruskal-Wallis para hemisferios

Variable	χ^2	df	p
Altura	30.6	5	l.t. .001
TimeToPeak	18.1	5	0.003
AnchoMitadAltura	35.1	5	l.t. .001
Onset	14.4	5	0.013
AreaBajoCurva	34.6	5	l.t. .001

Todas las variables muestran que p l.t. 0.05 por lo que difieren de características por lo que afirma su variabilidad.

5.2.5. Comparación entre lóbulos

Tabla 5.18: Estadísticos descriptivos de los parámetros de la HRF por lóbulos anteriores y Midbrain

Lóbulo	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
Anterior Lobe	2.65	4.99	5.05	1.83	9.36
Frontal Lobe	3.32	3.76	4.32	1.22	11.2
Limbic Lobe	3.84	3.59	3.97	1.23	11.6
Midbrain	1.68	7.62	7.14	2.94	5.32
<i>Mediana</i>					
Anterior Lobe	3.05	3.85	3.65	1.20	13.5
Frontal Lobe	3.76	3.40	3.00	1.20	14.2
Limbic Lobe	4.41	3.10	3.00	0.80	14.1
Midbrain	0.0313	9.20	7.80	3.50	0.236
<i>Desviación estándar</i>					
Anterior Lobe	2.19	2.67	2.59	1.19	6.45
Frontal Lobe	2.07	2.48	2.58	1.05	4.71
Limbic Lobe	2.38	2.71	2.41	1.20	4.88
Midbrain	3.19	3.62	3.20	1.51	7.23
<i>Mínimo</i>					
Anterior Lobe	0.00588	1.00	1.90	0.100	0.0379
Frontal Lobe	0.00138	0.600	1.30	0.100	0.00906
Limbic Lobe	0.00491	1.00	1.70	0.100	0.0303
Midbrain	0.00677	1.20	1.90	0.300	0.0617
<i>Máximo</i>					
Anterior Lobe	7.34	9.90	13.2	4.50	15.0
Frontal Lobe	8.14	10.0	13.7	6.80	15.0
Limbic Lobe	8.13	10.0	10.6	5.30	15.0
Midbrain	7.34	9.90	10.5	4.10	15.0

Tabla 5.19: Estadísticos descriptivos de los parámetros de la HRF por lóbulos posteriores y Sub-lobar

Lóbulo	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
Occipital Lobe	3.35	4.16	4.56	1.40	11.6
Parietal Lobe	4.16	3.24	3.64	1.08	13.0
Posterior Lobe	2.80	4.10	4.83	1.27	10.5
Sub-lobar	3.53	4.10	4.19	1.50	11.3
Temporal Lobe	2.73	4.63	5.35	1.44	10.8
<i>Mediana</i>					
Occipital Lobe	4.52	3.40	3.00	1.20	15.0
Parietal Lobe	4.52	3.40	3.00	1.20	15.0
Posterior Lobe	2.68	3.10	3.40	1.20	11.5
Sub-lobar	4.52	3.40	3.00	1.20	15.0
Temporal Lobe	1.94	3.40	4.20	1.20	12.5
<i>Desviación estándar</i>					
Occipital Lobe	2.19	2.72	2.70	1.18	5.00
Parietal Lobe	1.93	2.09	2.18	0.859	3.77
Posterior Lobe	1.85	2.67	2.68	1.12	4.65
Sub-lobar	2.25	2.63	2.39	1.19	5.53
Temporal Lobe	2.02	2.76	2.97	1.08	4.80
<i>Mínimo</i>					
Occipital Lobe	0.00556	1.00	1.70	0.100	0.0273
Parietal Lobe	0.00103	0.900	1.70	0.100	0.00685
Posterior Lobe	0.00160	1.00	1.70	0.100	0.0110
Sub-lobar	1.25e-4	1.00	1.70	0.100	3.60e-4
Temporal Lobe	0.00516	0.900	1.70	0.100	0.0331
<i>Máximo</i>					
Occipital Lobe	8.13	10.0	10.6	6.40	15.0
Parietal Lobe	8.13	10.0	11.1	5.20	15.0
Posterior Lobe	8.13	10.0	14.2	4.90	15.0
Sub-lobar	8.13	10.0	10.5	4.50	15.0
Temporal Lobe	8.13	10.0	13.8	4.60	15.0

El número de HRFs consideradas por lóbulo es el siguiente: Anterior Lobe: 112, Frontal Lobe: 1.768, Limbic Lobe: 222, Midbrain: 5, Occipital Lobe: 516, Parietal Lobe: 780, Posterior Lobe: 377, Sub-lobar: 182 y Temporal Lobe: 701. Estos valores representan la cantidad de registros incluidos en cada región para el análisis estadístico.

Tabla 5.20: Shapiro-Wilk test para lóbulos

Lóbulo	Altura	TimeToPeak	AnchoMitadAltura	Onset	AUC
<i>Shapiro-Wilk W</i>					
Anterior Lobe	0.824	0.876	0.871	0.895	0.738
Frontal Lobe	0.943	0.856	0.823	0.863	0.785
Limbic Lobe	0.936	0.839	0.805	0.837	0.723
Midbrain	0.640	0.680	0.900	0.753	0.760
Occipital Lobe	0.913	0.845	0.818	0.864	0.696
Parietal Lobe	0.912	0.768	0.690	0.825	0.605
Posterior Lobe	0.913	0.854	0.871	0.852	0.854
Sub-lobar	0.910	0.875	0.808	0.882	0.674
Temporal Lobe	0.903	0.900	0.876	0.912	0.820
<i>Shapiro-Wilk p</i>					
Anterior Lobe	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Frontal Lobe	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Limbic Lobe	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Midbrain	0.002	0.006	0.409	0.032	0.037
Occipital Lobe	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Parietal Lobe	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Posterior Lobe	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Sub-lobar	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
Temporal Lobe	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001

La tabla 5.20 que describe el test Shapiro-Wilk evidencia que para casi todos los casos se puede rechazar la hipótesis de normalidad, excepto para el Ancho Mitad Altura que sugiere a p mayor a 0.05.

Tabla 5.21: Resultados del test de Kruskal-Wallis para cada variable por lóbulo

Variable	χ^2	df	p
Altura	188.3	8	l.t. .001
TimeToPeak	106.3	8	l.t. .001
AnchoMitadAltura	185.0	8	l.t. .001
Onset	70.5	8	l.t. .001
AreaBajoCurva	151.3	8	l.t. .001

El test Kruskal-Wallis indica que la distribución de estas medidas difiere significativamente entre los distintos lóbulos. Esto sugiere un efecto relevante del lóbulo sobre las características estudiadas de la señal.

5.2.6. Comparación entre teslajes

Tabla 5.22: Estadísticos descriptivos de los parámetros de la HRF según Teslaje

Teslaje	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
1.5	4.48	3.47	3.08	1.42	15.0
3.0	3.35	3.56	4.63	0.958	11.4
7.0	3.90	4.83	3.95	2.13	15.0
<i>Mediana</i>					
1.5	4.52	3.40	3.00	1.40	15.0
3.0	3.05	3.10	3.00	0.500	12.6
7.0	3.89	4.60	3.60	2.10	15.0
<i>Desviación estándar</i>					
1.5	0.343	0.569	0.723	0.146	0.00261
3.0	2.19	2.60	2.90	0.901	3.85
7.0	1.24	1.78	1.21	0.860	4.91e-5
<i>Mínimo</i>					
1.5	1.32	3.40	3.00	1.00	15.0
3.0	0.220	0.900	1.70	0.100	2.03
7.0	1.94	1.10	1.70	0.200	15.0
<i>Máximo</i>					
1.5	4.52	9.90	10.6	3.20	15.0
3.0	8.14	10.0	14.2	4.20	15.0
7.0	7.85	9.20	7.20	4.10	15.0

El número de HRFs consideradas para cada nivel de Teslaje es el siguiente: 1.5 T: 882, 3.0 T: 3.623 y 7.0 T: 199. Estos valores representan la cantidad de registros incluidos en el análisis para cada intensidad de campo magnético.

Tabla 5.23: Shapiro-Wilk test for each teslaje

Teslaje	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Shapiro-Wilk W</i>					
1.5	0.0935	0.0905	0.0888	0.0867	0.0609
3.0	0.922	0.830	0.830	0.820	0.837
7.0	0.833	0.900	0.872	0.906	0.725
<i>Shapiro-Wilk p</i>					
1.5	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
3.0	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
7.0	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001

El test de Shapiro-Wilk indica que, para todos los teslajes analizados (1.5, 3.0 y 7.0), los valores de p son menores a 0.05 en todas las variables, lo que sugiere que se rechaza la hipótesis de

normalidad. Esto indica que la distribución de Altura, TimeToPeak, AnchoMitadAltura, Onset y AreaBajoCurva no sigue una distribución normal para ninguno de los teslajes evaluados.

Tabla 5.24: Resultados del test de Kruskal-Wallis para cada variable por teslaje

Variable	χ^2	df	p
Altura	89.8	2	l.t. .001
TimeToPeak	285.2	2	l.t. .001
AnchoMitadAltura	34.0	2	l.t. .001
Onset	858.4	2	l.t. .001
AreaBajoCurva	540.7	2	l.t. .001

El test de Kruskal-Wallis evidencia diferencias significativas entre los teslajes en todas las variables analizadas. Altura, TimeToPeak, AnchoMitadAltura, Onset y AreaBajoCurva presentan valores de p l.t. 0.001, indicando que la distribución de estas medidas difiere significativamente entre los distintos niveles de teslaje. Esto sugiere un efecto relevante del teslaje sobre las características evaluadas de la señal.

5.2.7. Comparación entre datasets

Tabla 5.25: Estadísticos descriptivos de los parámetros de la HRF según Experimento (Grupo 1)

Experimento	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
LANDMARK	0.0411	4.67	4.01	1.88	0.170
MOTOR	0.0151	8.37	6.57	3.70	0.0969
SELECTIVE	0.129	8.13	7.75	3.38	1.28
ds000102	0.0202	9.22	6.78	4.64	0.145
ds002422	4.48	3.47	3.08	1.42	15.0
<i>Mediana</i>					
LANDMARK	0.0235	4.65	3.75	2.00	0.0867
MOTOR	0.0113	8.70	6.70	3.80	0.0803
SELECTIVE	0.0408	9.30	7.50	3.80	0.264
ds000102	0.0181	9.30	7.40	4.20	0.133
ds002422	4.52	3.40	3.00	1.40	15.0
<i>Desviación estándar</i>					
LANDMARK	0.0435	2.25	1.50	1.20	0.226
MOTOR	0.0197	1.36	1.53	0.828	0.0953
SELECTIVE	0.323	2.51	2.63	1.46	3.46
ds000102	0.00802	0.691	1.29	1.25	0.0601
ds002422	0.343	0.569	0.723	0.146	0.00261
<i>Mínimo</i>					
LANDMARK	1.25e4	0.600	1.30	0.100	3.60e4
MOTOR	0.00376	2.00	2.40	0.100	0.0283
SELECTIVE	0.0179	1.70	2.60	0.300	0.115
ds000102	0.0108	8.10	4.70	3.60	0.0857
ds002422	1.32	3.40	3.00	1.00	15.0
<i>Máximo</i>					
LANDMARK	0.180	9.40	9.00	5.30	1.34
MOTOR	0.242	10.0	12.1	6.50	1.20
SELECTIVE	1.29	10.0	12.9	5.20	13.7
ds000102	0.0311	10.0	7.90	6.80	0.214
ds002422	4.52	9.90	10.6	3.20	15.0

Tabla 5.26: Estadísticos descriptivos de los parámetros de la HRF según Experimento (Grupo 2)

Experimento	Altura	TimeToPeak	AnchoMitadAltura	Onset	AreaBajoCurva
<i>Media</i>					
ds005576	3.90	4.83	3.95	2.13	15.0
ds005613	3.41	2.28	3.34	0.553	9.54
ds005700	3.35	5.06	5.52	1.65	13.8
ds006092	3.13	3.94	5.48	0.925	11.4
ds006105	4.63	3.44	3.91	1.12	15.0
<i>Mediana</i>					
ds005576	3.89	4.60	3.60	2.10	15.0
ds005613	3.16	1.40	2.30	0.300	9.64
ds005700	4.74	3.10	3.00	1.20	15.0
ds006092	1.64	3.50	6.20	0.500	12.5
ds006105	5.46	2.30	2.60	0.800	15.0
<i>Desviación estándar</i>					
ds005576	1.24	1.78	1.21	0.860	4.91e5
ds005613	1.92	2.09	2.20	0.791	3.65
ds005700	1.74	2.63	3.22	0.723	2.70
ds006092	2.65	2.47	2.84	0.864	3.77
ds006105	1.57	2.35	2.63	0.718	0.00841
<i>Mínimo</i>					
ds005576	1.94	1.10	1.70	0.200	15.0
ds005613	0.240	0.900	1.90	0.100	2.03
ds005700	0.325	3.10	3.00	0.700	3.68
ds006092	0.220	1.00	1.70	0.100	2.31
ds006105	1.31	2.30	2.60	0.800	15.0
<i>Máximo</i>					
ds005576	7.85	9.20	7.20	4.10	15.0
ds005613	7.34	10.0	13.7	4.20	15.0
ds005700	4.74	10.0	13.2	3.70	15.0
ds006092	8.14	10.0	14.2	3.50	15.0
ds006105	5.46	10.0	10.6	4.00	15.0

La cantidad de HRFs varía significativamente entre los diferentes experimentos. Para el primer grupo, se registraron 72 HRFs en el experimento LANDMARK, 296 en MOTOR, 15 en SELECTIVE, 5 en ds000102 y 882 en ds002422. En el segundo grupo, la distribución fue de 199 HRFs en ds005576, 1.339 en ds005613, 810 en ds005700, 1.320 en ds006092 y 154 en ds006105. Estos valores reflejan la amplitud y el tamaño de los datos disponibles para cada experimento, lo que permite un análisis comparativo robusto de los parámetros de la HRF.

Tabla 5.27: Shapiro-Wilk test for each experiment

Experimento	Altura	TimeToPeak	AnchoMitadAltura	Onset	AUC
<i>Shapiro-Wilk W</i>					
LANDMARK	0.782	0.967	0.922	0.952	0.651
MOTOR	0.737	0.877	0.988	0.921	0.744
SELECTIVE	0.892	0.749	0.955	0.884	0.760
ds000102	0.970	0.868	0.864	0.791	0.862
ds002422	0.0935	0.0905	0.0888	0.0867	0.0609
ds005576	0.833	0.900	0.872	0.906	0.725
ds005613	0.962	0.587	0.669	0.550	0.947
ds005700	0.693	0.711	0.696	0.690	0.510
ds006092	0.796	0.894	0.911	0.737	0.854
ds006105	0.528	0.518	0.530	0.492	0.425
<i>Shapiro-Wilk p</i>					
LANDMARK	l.t. 0.001	0.059	l.t. 0.001	0.009	l.t. 0.001
MOTOR	l.t. 0.001	l.t. 0.001	0.019	l.t. 0.001	l.t. 0.001
SELECTIVE	0.086	0.001	0.649	0.066	0.002
ds000102	0.878	0.259	0.241	0.068	0.236
ds002422	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
ds005576	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
ds005613	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
ds005700	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
ds006092	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001
ds006105	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001	l.t. 0.001

El test de Shapiro-Wilk muestra que la mayoría de los experimentos no cumplen con la normalidad en varias variables, dado que los valores de p son menores a 0.05. En particular, experimentos como LANDMARK y MOTOR presentan normalidad parcial en algunas variables (por ejemplo, TimeToPeak en LANDMARK con $p = 0.059$), mientras que experimentos como ds002422, ds005576 y ds005613 no cumplen la normalidad en ninguna variable (p l.t. 0.001 en todas). Esto indica que para la comparación de variables entre experimentos es más apropiado usar tests no paramétricos, como el Kruskal-Wallis, para evitar suposiciones de normalidad.

Tabla 5.28: Resultados del test de Kruskal-Wallis para cada variable por experimento

Variable	χ^2	df	p
Altura	1227	9	l.t. .001
TimeToPeak	1695	9	l.t. .001
AnchoMitadAltura	1096	9	l.t. .001
Onset	2407	9	l.t. .001
AreaBajoCurva	2337	9	l.t. .001

El test de Kruskal-Wallis evidencia diferencias significativas entre los experimentos en todas las variables analizadas. Altura, TimeToPeak, AnchoMitadAltura, Onset y AreaBajoCurva

presentan valores de p l.t. 0.001, lo que indica que la distribución de estas medidas difiere significativamente entre los distintos experimentos. Esto confirma un efecto relevante del tipo de experimento sobre las características evaluadas de la señal.

6 | Conclusión

El presente estudio permitió desarrollar un flujo de trabajo integral para el análisis de funciones de respuesta hemodinámica (HRF) a partir de múltiples *datasets*. Desde el preprocesamiento de los datos, se aplicaron procedimientos sistemáticos de limpieza y depuración, incluyendo la exclusión de valores anómalos o fisiológicamente poco plausibles, asegurando la fiabilidad y consistencia de la base de datos final. Esto permitió consolidar un total de 5.092 HRFs válidas, organizadas en archivos `.tsv` estructurados para su análisis estadístico.

La utilización de *Jamovi* [12] facilitó la obtención de estadísticas descriptivas de manera eficiente y reproducible, incluyendo medias, desviaciones estándar, valores mínimos y máximos, así como el número de observaciones por parámetro, lo que permitió caracterizar la variabilidad de las HRFs y establecer rangos fisiológicos esperables.

Es importante destacar que gran parte del desarrollo de este flujo de trabajo se basó en los códigos proporcionados por Danilo Lara. Gracias a ello, fue posible optimizar y parametrizar los scripts, evitando el *hardcodeo* y mejorando significativamente la reproducibilidad y escalabilidad del análisis. Este logro permitió adaptar los procedimientos a múltiples *datasets*, así perdiendo foco en acomodar los script para cada dataset y ajustándose el interés en el trabajo con las HRFs resultantes

El análisis de las HRFs reveló una variabilidad considerable entre sujetos y condiciones experimentales, lo que refleja la complejidad intrínseca de la respuesta hemodinámica en distintos contextos. Esta variabilidad se observa no solo entre *datasets* diferentes, sino también dentro de cada conjunto de datos, en hemisferios del cerebro, teslaje utilizado para las fMRI, por sexo, etc. Dada esta heterogeneidad, surge la necesidad de desarrollar métodos más robustos para la estimación de la HRF, capaces de capturar patrones generales sin perder sensibilidad ante diferencias individuales. La integración de bases de datos depuradas y estructuradas, junto con procedimientos automatizados y reproducibles, permite explorar estrategias avanzadas de modelado que minimicen el efecto de valores atípicos y mejoren la confiabilidad de los resultados.

Finalmente, se identificó que una dificultad clave en este tipo de análisis está en el manejo de los datos debido a la extensión y complejidad de las variables involucradas, tales como `pixel_log` y `ts_signal`. Esto resalta la importancia de buscar soluciones de optimización que mantengan la reproducibilidad y eficiencia del análisis sin comprometer la facilidad de acceso y manejo de los datos.

6.1. Trabajos futuros

Los resultados obtenidos en este estudio abren diversas líneas de investigación que podrían fortalecer y ampliar la comprensión de la respuesta hemodinámica (HRF). En primer lugar, sería relevante explorar nuevas formas de estimar la HRF, alternativas a las metodologías existentes como la propuesta por Shan, incluyendo revisiones exhaustivas de la literatura para identificar enfoques más precisos o adaptativos. Este tipo de estudios permitiría manejar de manera más efectiva la variabilidad observada entre sujetos, sesiones y condiciones experimentales.

El aumento de la cantidad de *datasets* y sujetos conlleva un incremento significativo en el costo computacional, particularmente debido a la gran cantidad de variables por HRF que deben procesarse y analizarse. Por ello, resulta esencial optimizar los procedimientos de preprocesamiento y análisis, buscando soluciones que mantengan la reproducibilidad y eficiencia sin comprometer la facilidad de acceso a los datos. Estrategias como la parametrización de scripts, la paralelización de procesos, la utilización de formatos estructurados y estándar (`.tsv`, BIDS) o técnicas de manejo eficiente de matrices y datos voluminosos podrían contribuir a lograr este objetivo.

Finalmente, futuras investigaciones podrían enfocarse en desarrollar modelos estadísticos y algoritmos robustos capaces de integrar grandes volúmenes de datos depurados, capturando patrones generales de la HRF sin perder sensibilidad ante diferencias individuales. Esto permitirá no solo mejorar la estimación de la HRF, sino también facilitar su aplicación en estudios clínicos y experimentales a gran escala.

7 | Anexos

7.1. Análisis General

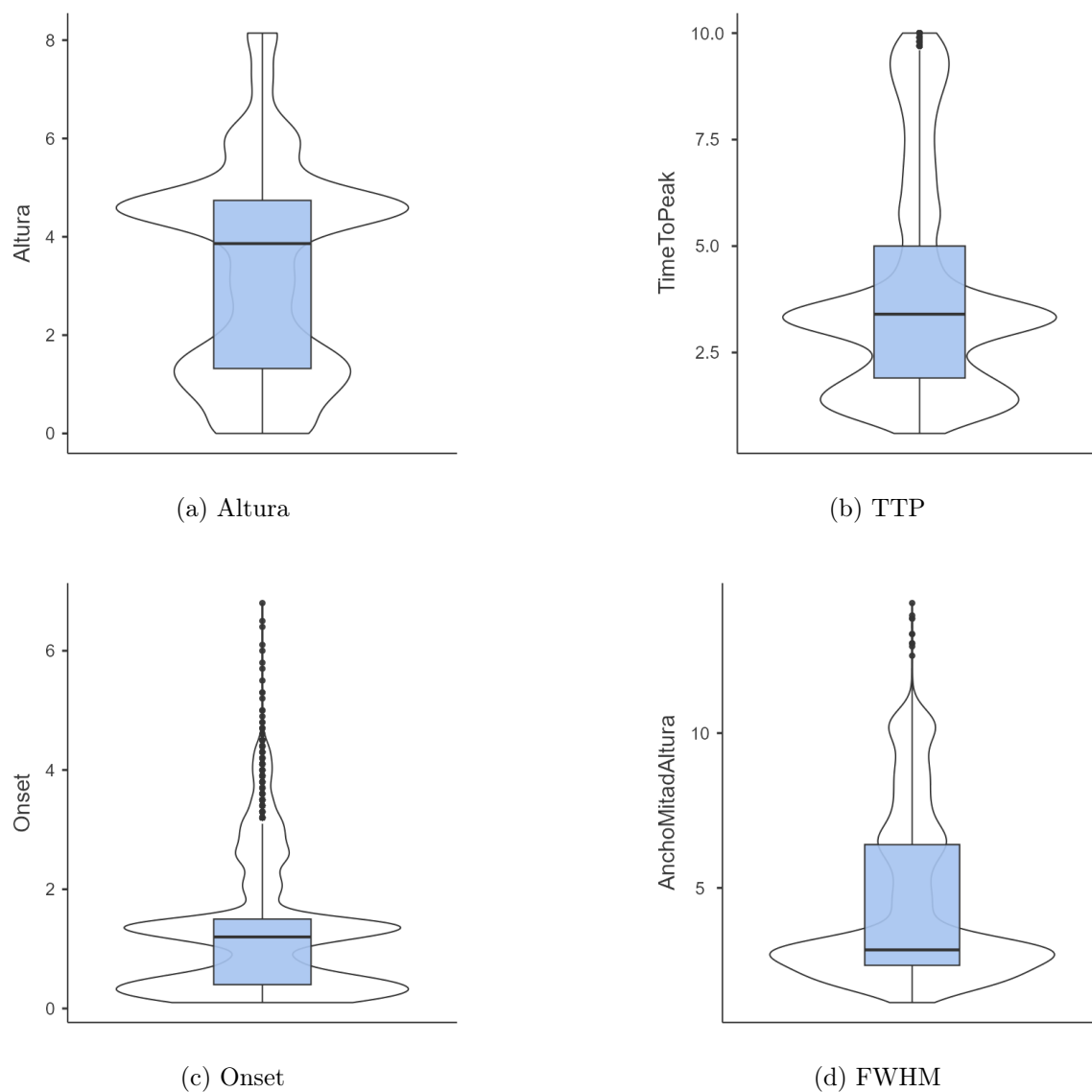


Figura 7.1: Boxplot que demuestran la distribuciones de 5092 HRFs

7.2. Tablas estadísticas

Tabla 7.1: Descriptivos y pruebas de normalidad de las métricas HRF no válidas

Motivo	Altura	TimeToPeak	PWHM	Onset	AUC
Descriptives					
Altura l.t 0.005 con mala forma	−0,530	1,13	1,12	1,86	0,459
TimeToPeak l.t 1 con altura m.t 4	−1,53	0,384	1,64	4,17	−2,32
TimeToPeak m.t 10 Área = 0	−0,571 3,87	0,913 1,20	−0,783 0,0829	1,85 0,936	−0,808 NaN
Kurtosis					
Altura l.t 0.005 con mala forma	−1,14	−0,207	0,0707	3,59	0,719
TimeToPeak l.t 1 con altura m.t 4	1,48	−1,35	1,69	15,7	3,97
TimeToPeak m.t 10 Área = 0	−0,257 15,0	2,16 1,52	3,17 −0,506	11,5 2,00	−0,156 NaN
Shapiro–Wilk W					
Altura l.t 0.005 con mala forma	0,900	0,781	0,836	0,774	0,966
TimeToPeak l.t 1 con altura m.t 4	0,747	0,796	0,712	0,227	0,474
TimeToPeak m.t 10 Área = 0	0,956 0,284	0,939 0,851	0,932 0,972	0,889 0,862	0,906 NaN
Shapiro–Wilk p					
Altura l.t 0.005 con mala forma	0,049	l.t.001	0,004	l.t.001	0,697
TimeToPeak l.t 1 con altura m.t 4	l.t.001	l.t.001	l.t.001	l.t.001	l.t.001
TimeToPeak m.t 10 Área = 0	l.t.001 l.t.001	l.t.001 0,018	l.t.001 0,882	l.t.001 0,026	l.t.001 NaN

Bibliografía

- [1] Z. Y. Shan, M. J. Wright, P. M. Thompson, K. L. McMahon, G. G. A. M. Blokland, G. I. de Zubicaray, N. G. Martin, A. A. E. Vinkhuyzen, and D. C. Reutens, “Modeling of the hemodynamic responses in block design fmri studies,” *Journal of Cerebral Blood Flow & Metabolism*, vol. 34, no. 2, pp. 316–324, 2014. PMID: 24252847.
- [2] W. Chen, *Hemodynamic Response Function Modeling*. PhD thesis, University of North Carolina at Chapel Hill, 2019.
- [3] D. F. Lara Vergara, “Caracterización de la función de respuesta hemodinámica (hrf) en una muestra amplia de sujetos sanos,” 2024.
- [4] H. Girouard and C. Iadecola, “Neurovascular coupling in the normal brain and in hypertension, stroke, and alzheimer disease,” *Journal of applied physiology*, vol. 100, no. 1, pp. 328–335, 2006.
- [5] X. Han and T. Li, “A mini review of functional magnetic resonance imaging,” *OAText*, 2018.
- [6] C. Gauthier and A. Fan, “Bold signal physiology: Models and applications,” *NeuroImager*, vol. 187, pp. 116–127, 2019.
- [7] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-B. Poline, C. D. Frith, and R. S. J. Frackowiak, “Analysis of fmri time-series revisited,” *NeuroImage*, vol. 2, no. 1, pp. 45–53, 1995.
- [8] O. Josephs, R. Turner, and K. Friston, “Event-related f mri,” *Human Brain Mapping*, vol. 5, no. 4, pp. 243–248, 1997.
- [9] A. Veloz, C. Moraga, A. Weinstein, L. Hernández-García, S. Chabert, R. Salas, R. Riveros, C. Bennett, and H. Allende, “Fuzzy general linear modeling for functional magnetic resonance imaging analysis,” *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 1, pp. 100–111, 2020.
- [10] K. Worsley and K. Friston, “Analysis of fmri time-series revisited—again,” *NeuroImage*, vol. 2, no. 3, pp. 173–181, 1995.
- [11] Wellcome Trust Centre for Neuroimaging, *SPM12 Manual*. University College London, London, UK, 2014. Disponible en: <https://www.fil.ion.ucl.ac.uk/spm/doc/manual.pdf>.
- [12] The Jamovi Project, “Jamovi project (version 2.5) [computer software],” 2023. Accedido: Sep. 2025.
- [13] “Talairach daemon.” <http://www.talairach.org/>, —. Accedido: 2 de octubre de 2025.