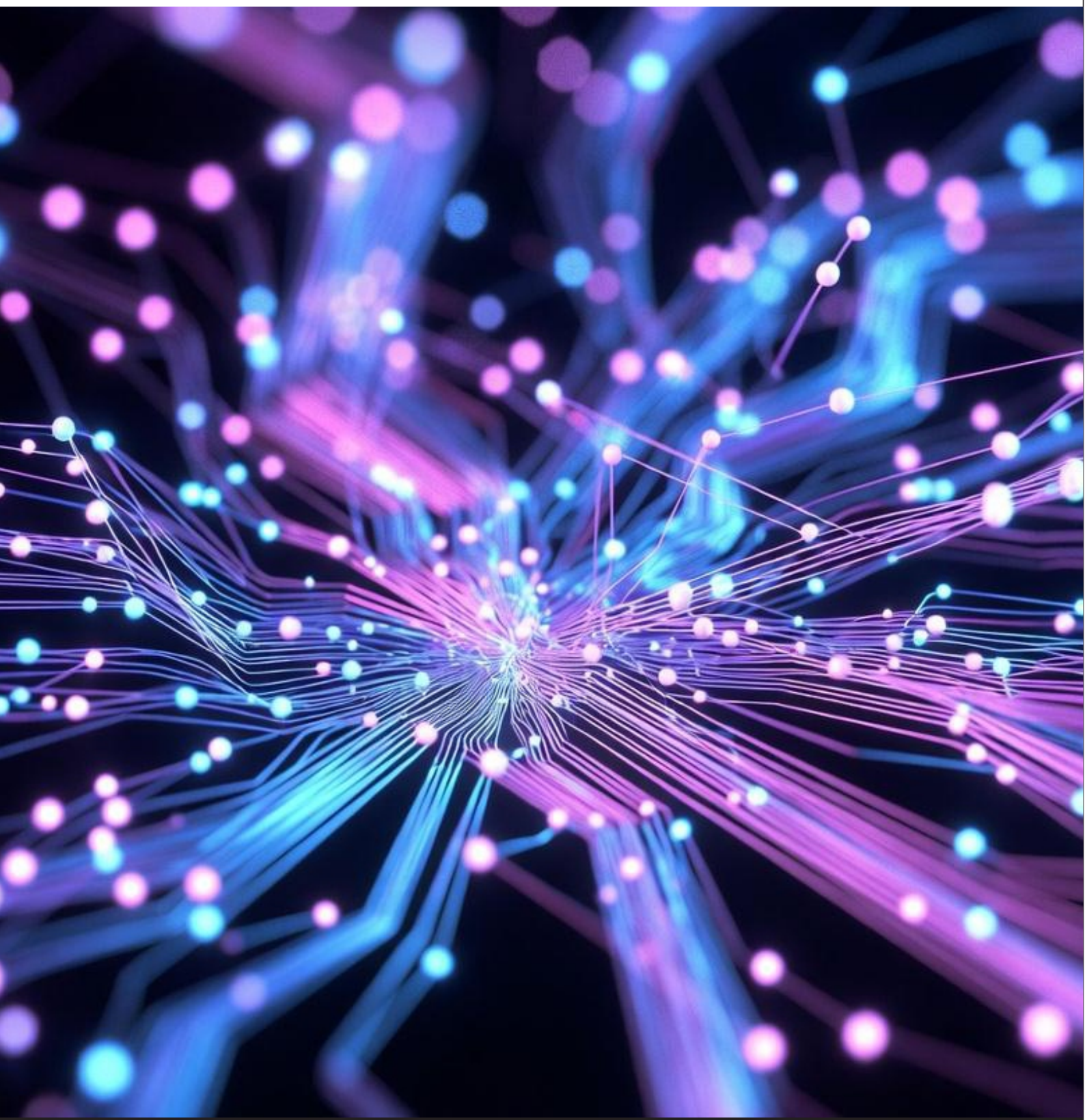
# Informaticae Abstracta

Año 2026, Vol. 4, No. 1 ENE | JUN 2026
ISSN: 3061-8355



Universidad Autónoma del Estado de México

Informaticae Abstracta
Año 2026, Vol. 4, No. 1 ENE | JUN 2026



Informaticae Abstracta, año 2026, volumen 4, número 1, Enero – Junio 2026, es una revista de publicación semestral editada por la Universidad Autónoma del Estado de México, Instituto Literario 100 Ote., Col. Centro, Toluca, Estado de México, C.P. 50000, Tels. (722) 2140855 y 21140795 ext. 10 1217, http://informaticae.uaemex.mx, informaticae@uaemex.mx, Editora Responsable: Ruth Hernández Pérez, Facultad de Ingeniería. Reserva de Derechos al Uso Exclusivo número 04-2022-111113523100-102, ISSN: 3061-8355, ambos otorgados por el Instituto Nacional del Derecho de Autor. Responsable de la última edición Mercedes Lucero Chávez, de la Facultad de Ingeniería, Cerro de Coatepec s/n, Ciudad Universitaria, C.P. 50100, Toluca, Estado de México, fecha de última modificación, 30 junio 2026.

Informaticae Abstracta

Año 2026, Vol. 4, No. 1 ENE | JUN 2026

# Índice

**Modelado y simulación de flujo tipo Taylor en capilares** 4
*Natividad-Rangel, R., Ramírez-Serrano, A., Romero-Romero, R., Hernández-Servín, J. A.*

**Modelación de la volatilidad condicional mediante la distribución Kumaraswamy-Laplace: una innovación al modelo GARCH para la optimización del VaR en los mercados Brent y WTI** 43
*Trujillo-Arévalo, F.*

**Neural networks as trainable compositional systems** 72
*Solís-Winkler, A.*

**Evaluation of the different time-travel approaches used in geotechnical seismic testing** 83
*Contreras-Ferreyra, D. U., Pastor-Gómez, N., Chávez-Negrete, C., Martínez-Rojas, A.*

**Sistema de coordenadas SPR: optimización algebraica y análisis CORDIC para control robótico** 99
*Salas-García, J., Hernández-Servín, J. A., Vilchis-González, A. H., Ávila-Vilchis, J. C.*

**Prototipado de órganos artificiales para entrenamiento en cirugía robótica** 124
*Dávila-Vilchis, J. M., Zúñiga Avilés, L. A., Ramírez-García, N. G., Vilchis-González, A. H.*

**Informaticae Abstracta**. ISSN: 3061-8533
Año 2026, Vol. 4, No. 1, pp. 4-42
JUL | DIC 2025
doi: https://doi.org/10.36677/ia.v4n1.28934

# Modelado y simulación de flujo tipo Taylor en capilares

Natividad-Rangel, R.; Ramírez-Serrano, A.*; Romero-Romero, R.; Hernández-Servín, J. A.

**Información del Artículo**

Recibido: 2 de enero, 2026
Aceptado: 3 de abril, 2026

**Palabras clave:** flujo tipo Taylor, simulación numérica, microfluídica, COMSOL Multiphysics, flujo bifásico

**Resumen**

En objetivo de este trabajo es presentar un modelo de simulación numérica del tipo Taylor que se desarrolla en capilares. Para llevar a cabo este análisis, se empleó el entorno de software [1] en su versión 6.2, una herramienta robusta y ampliamente reconocida en el ámbito de la simulación multifísica.

El flujo tipo Taylor se distingue por la generación secuencial de burbujas elongadas de gas, las cuales se alternan con secciones de líquido, creando un patrón periódico. Este fenómeno resulta de particular interés debido a su relevancia en diversas aplicaciones, que abarcan desde los micro fluidos, donde el control preciso de fluidos a pequeña escala es crucial, hasta los sistemas de intercambio de calor, donde la eficiencia en la transferencia de energía es primordial.

Con el objetivo de explorar a detalle este tipo de flujo, se llevó a cabo un conjunto exhaustivo de simulaciones, comprendiendo un total de 32 casos. Estas simulaciones se desarrollaron en un dominio bidimensional con simetría axial, lo que permitió capturar la esencia del flujo de manera eficiente.

Se procedió a variar sistemáticamente una serie de parámetros clave, tales como el diámetro del capilar (3-4 mm), el espesor de la película líquida que recubre la pared del capilar (25-50 $\mu$m), la velocidad del fluido (0.02, 0.05, 0.2 y 0.5 m/s) y, finalmente, dos tipos distintos de condiciones de contorno: la condición de entrada que simula una alimentación constante de fluido, y la condición periódica que asume un comportamiento repetitivo del flujo a lo largo del capilar.

Los hallazgos derivados de este estudio permiten discernir el impacto específico de cada parámetro sobre la dinámica de las burbujas y la estabilidad de la película líquida en sistemas de microfluidos basados en el flujo bifásico en canales confinados. Se identificaron, en particular, aquellas condiciones que propician un flujo estable y simétrico, características deseables en muchas aplicaciones.

## Introducción

En la industria química, los sistemas multifásicos son esenciales debido a su eficacia en la transferencia de masa y calor, lo que los convierte en componentes clave de diversos procesos. Uno de estos procesos es el flujo tipo Taylor, reconocido por ser una opción eficiente para mejorar la transferencia de masa gracias a su alta relación área-volumen y su flujo ordenado. Este tipo de flujo se caracteriza por la formación de burbujas de gas separadas por segmentos de líquido dentro de un canal capilar (1-4 mm), generando un amplio contacto interfacial entre las fases y mejorando la eficiencia en los procesos de transporte Fishwick et al. [2].

Aplicaciones interesantes de este tipo de sistema es la hidrogenación selectiva de alquinos, absorción de CO2 Natividad et al. [3] . De hecho, la Universidad Autónoma del Estado de México cuenta con el título de patente Peña et al. [4], Fishwick et al. [2], de un foto-reactor capilar donde se ha demostrado las ventajas del flujo tipo Taylor, combinado con luz para realizar reacciones multifásicas selectivas Peña et al. [4]

A pesar de sus ventajas, el estudio del comportamiento hidrodinámico del flujo tipo Taylor sigue presentando retos importantes, considerar factores como: el diámetro del capilar, la velocidad de las fases y las propiedades interfaciales de los fluidos que influyen considerablemente en su dinámica, complicando su análisis teórico. La complejidad de este fenómeno

ha impulsado el desarrollo de modelos matemáticos y herramientas de simulación para predecir su comportamiento bajo diferentes condiciones operativas.

Para enfrentar este desafío, se han creado modelos matemáticos que describen de manera precisa la hidrodinámica del flujo tipo Taylor. Las simulaciones numéricas desempeñan un papel fundamental en la comprensión y optimización de este flujo, facilitando su diseño y mejorando su aplicación en entornos industriales. En este contexto, el uso de software especializado como [1] resulta clave; ya que es una herramienta que permite modelar la dinámica de sistema al resolver las ecuaciones acopladas acopladas con modelos de cantidad de movimiento, masa y calor; además, su implementación en simulaciones computacionales permite evaluar diferentes parámetros en la formación y evolución del flujo, sentando las bases para la optimización de procesos multifásicos.

# Generalidades

El estudio de los flujos multifásicos en microcanales ha cobrado una importancia creciente debido a su impacto en diversas áreas, desde la producción de compuestos químicos hasta innovaciones en biotecnología y medicina. Entre estos flujos, el flujo tipo Taylor ha captado un gran interés porque se caracteriza por la formación de burbujas de gas separadas por segmentos de líquido dentro de un canal estrecho (ver Figura 1). No solo es visualmente distintivo, sino que también ofrece ventajas significativas en términos de transferencia de masa, calor y cantidad de movimiento, lo que lo hace más eficiente que los flujos monofásicos tradicionales [5, 6].

El desarrollo de herramientas computacionales como [1] ha permitido que el análisis de estos sistemas sea cada vez más preciso y es posible predecir con gran detalle cómo se comportan las burbujas dentro de los capilares, cómo interactúan con el líquido circundante y cómo afectan la eficiencia de los procesos en los que están involucradas. Modelar estos flujos no solo ayuda a entender mejor su dinámica, sino que también permite optimizar su aplicación en distintos campos de la ciencia y tecnología Butler et al. [7].

El flujo tipo Taylor en capilares contempla una gran cantidad de factores que influyen en su comportamiento como la velocidad con la que ascienden las burbujas, la longitud de los segmentos de líquido y la caída de presión dentro del canal Liu et al. [8]. A medida que la miniaturización de dispositivos sigue avanzando y la dinámica de fluidos en microdispositivos se vuelve más relevante, es fundamental comprender cómo optimizar estos flujos para que sean lo más eficientes posible. Por otro lado, estudios previos han demostrado que factores como el tamaño del canal, la viscosidad del líquido y la tensión superficial pueden alterar significativamente la estabilidad y periodicidad del flujo Gupta et al. [9].

Los factores principales que influyen en el comportamiento del flujo tipo Taylor en capilares son Gupta et al. [9] y Kreutzer [10]:

- **Regímenes de flujo:** la forma del capilar, las propiedades físicas del sistema definen el flujo dentro del capilar y cómo este puede adoptar diferentes patrones.
- **Velocidad relativa de las fases:** la velocidad de cada fase tiene un impacto directo en la forma y en la estabilidad de las burbujas, como en el caso en el que la fase gaseosa

se mueve más rápido que la fase líquida, resultando en burbujas alargadas con movimientos circulares internos en la parte líquida. [11].

- **Efecto de confinamiento:** el reducido tamaño del capilar y la dinámica del flujo realzan la interacción con las paredes, provocando una distribución de fases no uniforme. Además, aumenta la relevancia de la tensión superficial y se favorece la formación de películas líquidas delgadas alrededor de las burbujas.
- **Número de Reynolds y régimen hidrodinámico:** en capilares predomina el régimen laminar (Re bajos), pero bajo ciertas condiciones pueden surgir inestabilidades que alteran la distribución de fases y la formación de burbujas.

En el capilar, la burbuja de gas no lo llena por completo; siempre queda una capa delgada de líquido entre la burbuja y la pared. De modo que el espesor de esta película depende de las fuerzas viscosas respecto a las de tensión superficial, y está representado por el número capilar ($Ca$). En general, un valor alto de $Ca$ indica una reducción del espesor de la película.

Por otro lado, la transferencia de masa es igualmente relevante, especialmente en sistemas donde ocurren reacciones químicas o hay un intercambio de solutos entre fases. Este proceso ocurre debido a la difusión molecular que tiene lugar en la película líquida alrededor de la burbuja, lo que permite un transporte pasivo de especies químicas. Simultáneamente, la convección en el fluido interno se ve favorecida por la circulación dentro de los segmentos líquidos, creando zonas de mezcla activa que optimizan la transferencia de masa. Por lo anterior, es esencial entender el comportamiento del flujo de Taylor en su totalidad para poder predecir sus características en diversas condiciones de operación. Par la simulación de este tipo de flujo requiere establecer el modelo matemático que describan los fenómenos de transporte como la mecánica de fluidos, los fenómenos interfaciales, la transferencia de masa, las reacciones químicas y la transferencia de calor. Mediante estos modelos, es posible analizar el movimiento de las burbujas, la distribución de presiones y velocidades, y la influencia de parámetros clave como la tensión superficial y la viscosidad.

# Metodologia

El modelado matemático se base en la resolución de ecuaciones diferenciales parciales y algebraicas que rigen la dinámica de los fluidos, como las leyes de conservación de masa, cantidad de movimiento y energía. Sin embargo, la complejidad del flujo multifásico en espacios confinados exige la aplicación de métodos numéricos y simplificaciones para poder obtener resultados viables.

El comportamiento del flujo de Taylor en capilares está establecido por medio del balance de momento, que describen el movimiento de los fluidos en términos de viscosidad, presión y fuerzas externas. Estas ecuaciones, en conjunto con la ecuación de continuidad y las condiciones en las interfaces, permiten modelar el comportamiento de las burbujas y los segmentos líquidos.

La ecuación de continuidad [12] expresa la conservación de la masa en un sistema fluido y se define como:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{1}$$

Para el caso de un fluido incompresible, en estado estacionario la ecuación se simplifica a:

$$\nabla \cdot \vec{u} = 0 \tag{2}$$

lo que indica que el volumen del fluido se conserva a lo largo del flujo.

Mientras que las ecuaciones de cantidad de movimiento [12] describen el equilibrio de fuerzas en un fluido viscoso y se expresan como:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho(\mathbf{u}\nabla\mathbf{u}) = -\nabla p + \mu(\nabla^2 \mathbf{u}) + F \tag{3}$$

Donde $\rho$ representa la densidad del fluido; el término $\mathbf{u}$ el campo de velocidades que describe la velocidad del fluido en cada punto del espacio; $\mu$ la viscosidad del fluido, $p$ la presión, $\mathbf{F}$ fuerzas externas (gravedad, fuerzas interfaciales) y $\nabla$ la divergencia del flujo de masa.

Para el caso específico del flujo de Taylor, la resolución de estas ecuaciones requiere considerar la interacción entre las fases líquida y gaseosa, lo cual introduce discontinuidades en la viscosidad y la densidad en la región interfacial. Dada la interacción dinámica entre las fases gaseosa y líquida en el flujo de Taylor, es imprescindible modelar la interfaz que las separa con alta precisión. Para lograr esto, se han desarrollado diversos métodos numéricos, incluido en [1], entre los cuales se encuentra el método de la función de nivel (Level-Set Method) cuya técnica numérica basada en la evolución de una función escalar entre dos fluidos $\phi(x, t)$ que representa la distancia de un punto del dominio hasta la interfaz entre dos fluidos. Esta función se define de manera que:

- $\phi > 0$: en la región ocupada por un fluido (por ejemplo, el líquido
- $\phi < 0$: en la región ocupada por el otro fluido (por ejemplo, el gas),
- $\phi = 0$: indica la ubicación exacta de la interfaz entre ambos fluidos.

La ecuación temporal de esta función viene dada por la siguiente ecuación:

$$\frac{\partial \phi}{\partial t} + \mathbf{u} \cdot \nabla \phi = 0 \tag{4}$$

Esta formulación permite capturar de forma natural los cambios topológicos de la interfaz, como su deformación, ruptura o fusión de burbujas.

Para preservar la propiedad de $\phi$ garantizar estabilidad numérica, se emplea una técnica de re-inicialización basada en la ecuación:

$$\frac{\partial \phi}{\partial \tau} = S(\phi_o)\,(1 - |\nabla \phi|) \tag{5}$$

Donde $\tau$ es una variable de pseudo-tiempo que no corresponde formalmente al tiempo del sistema, sino que lo utiliza exclusivamente para estabilizar numéricamente la función de

nivel durante su reconfiguración. Este proceso se lleva a cabo en forma iterativa hasta que el gradiente de $\phi$ se aproxima a la unidad, es decir; $|\nabla\phi| \cong 1$, lo cual garantiza que actúe nuevamente como una distancia escalar.

La función $S(\phi_o)$ es una función de signo suavizada, derivada de la función de nivel original $\phi_o$. Esta función indica en qué región del dominio se encuentra cada punto con respecto a la interfaz mencionada anteriormente en la función de nivel (Level-Set Method).

[1] emplea el método VOF que representa la fracción de volumen de fluido en cada celda de la malla. Se define una función de fracción de fluido $F(x,t)$ que varía entre 0 (gas) y 1 (líquido), y su evolución se rige por:

$$\frac{\partial F}{\partial t} + \mathbf{u} \cdot \nabla F = \tag{6}$$

Para reconstruir la interfaz, se emplean técnicas como el método PLIC (Piecewise Linear Interface Calculation), donde la interfaz se representa mediante segmentos lineales en cada celda.

## Condiciones de Tensión Superficial

La tensión superficial es clave en el flujo de Taylor, y se modela mediante la formulación de la fuerza de curvatura de Brackbill en VOF:

$$\mathbf{F}_\sigma = \sigma k \mathbf{n} \delta_s \tag{7}$$

donde $\sigma$ es la tensión superficial, $k$ es la curvatura de la interfaz, $\mathbf{n}$ es el vector normal $\delta_s$ es una función delta que define la interfaz.

En el método Level-Set la curvatura se calcula como:

$$k = \nabla \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \tag{8}$$

La ausencia de soluciones analíticas exactas para las ecuaciones que describen el flujo de Taylor, especialmente al considerar la tensión superficial, la interacción entre fases y la geometría confinada, hace necesario el uso de métodos numéricos para obtener soluciones aproximadas. Estos métodos permiten resolver las ecuaciones de cantidad de movimiento y de continuidad en un dominio computacional.

También es importante considerar la relación de longitud de burbuja a diámetro del capilar (L/D), ya que afecta a la recirculación dentro de los segmentos de líquido y la transferencia de masa y calor. Este parámetro se determina empíricamente y depende de las condiciones de operación del flujo.

$$\frac{L}{D} = f(Re, Ca, We) \tag{9}$$

El ángulo de contacto ($\theta$) es otro parámetro a considerar, ya que define la interacción entre la fase líquida y la pared del capilar. La hidrofobicidad o hidrofilicidad del material afecta la forma y velocidad de las burbujas. En términos generales:

- Para un capilar hidrofóbico $\theta > 90°$, las burbujas tienden a ser más largas y moverse con mayor velocidad.
- Para un capilar hidrofílico, $\theta < 90°$, las burbujas pueden presentar una menor movilidad y mayor interacción con la pared.

La caída de presión a lo largo del capilar es un parámetro esencial en la simulación, ya que determina la velocidad del flujo y la estabilidad del patrón de burbujas. Se puede calcular usando la ecuación de Hagen-Poiseuille para flujo laminar:

$$\Delta P = \frac{32\mu u L}{D^2} \tag{10}$$

Para flujos bifásicos como el de Taylor, esta ecuación se modifica para incluir el efecto de la presencia de burbujas, generalmente mediante correlaciones empíricas. El comportamiento del flujo de Taylor en capilares depende de múltiples parámetros adimensionales y físicos, que determinan la formación, estabilidad y transporte de burbujas. La simulación numérica de este fenómeno requiere modelar adecuadamente estos factores mediante ecuaciones diferenciales parciales resueltas con métodos computacionales como FEM, FVM o FDM. El ajuste preciso de los parámetros clave es esencial para obtener resultados confiables y predecir la dinámica del sistema en aplicaciones industriales y científicas.

# Geometría del sistema

El modelo consiste en un canal capilar de sección circular, caracterizado por un radio ($R$), a través del cual una fase líquida transporta burbujas de gas alargadas en dirección vertical ascendente. Para facilitar el análisis, adoptamos un marco de referencia que se desplaza conjuntamente con una de las burbujas. Esta elección permite representar el problema de manera bidimensional en coordenadas cilíndricas $(r, z)$ o cartesianas $(x, y)$ según la formulación numérica seleccionada.

Con el propósito de simplificar la resolución computacional, consideramos un dominio periódico que abarca una única burbuja y el líquido circundante. Esta aproximación se basa en la consideración de una distribución regular de burbujas, donde cada una posee dimensiones y espaciamiento idénticos.

## Parámetros Geométricos Relevantes

A continuación, se detallan los parámetros geométricos clave que definen el sistema:

- $H$: Radio del capilar (m)
- $L_b$: Longitud de la burbuja de gas (m)
- $L$: Longitud de la celda periódica (m)
- $\delta$: Espesor de la película líquida entre la burbuja y la pared del canal (m)

## Propiedades físicas del sistema

El análisis se centra en un sistema bifásico compuesto por un líquido (agua) y un gas (aire). Dada la diferencia significativa en propiedades, restringimos nuestro estudio principalmente a la fase líquida, asumiendo que la viscosidad y la densidad del gas son despreciables en comparación.

Inicialmente, consideramos que las propiedades físicas de la fase líquida permanecen constantes a lo largo de la simulación. No obstante, reconocemos la importancia de explorarla influencia de posibles variaciones en estas propiedades, lo cual abordaremos cualitativamente en secciones posteriores.

## Propiedades físicas de la fase líquida

Los parámetros que definen las propiedades físicas de la fase líquida son:

- $\mu$: Viscosidad dinámica del líquido ($1 \times 10^{-3}$ Pa·s)
- $\rho$: Densidad del líquido ($1000$ kg/m$^3$)
- $\sigma$: Tensión superficial entre el líquido y el gas ($0{,}072$ N/m)
- $g$: Aceleración gravitacional $9{,}81$ m/s$^2$
- $u$: Velocidad de la burbuja $0{,}01 - 0{,}05$ m/s
- $H$: Radio del capilar $0{,}5$ mm

Nota: Los valores colocados de cada parámetro se ajustarán con las condiciones del sistema experimental en el modelado.

## Variables del sistema

Para describir el comportamiento del sistema, definimos las siguientes variables independientes y dependientes:

### Variables Independientes

Estas son las variables que controlan o definen el sistema, y que no dependen de otras variables dentro del modelo:

- $(r, z)$ o $(x, y)$: coordenadas espaciales en el dominio bidimensional (m)
- $t$: tiempo (s). Si bien el sistema evoluciona con el tiempo, en nuestro marco de referencia (que se mueve con la burbuja) asumimos un régimen estacionario, formulando el modelo como cuasi-estático.

**Variables Dependientes**

Estas variables son el resultado de las interacciones dentro del sistema, y su valor se determina a partir de las variables independientes y las ecuaciones que describen el modelo:

- $\mathbf{u}(x, y) = (u_x, u_y)$: Campo de velocidades de la fase líquida (m/s). Este campo vectorial describe la velocidad del líquido en cada punto del dominio.
- $p(x, y)$: Campo de presión en la fase líquida (Pa). Este campo escalar representa la presión en cada punto del dominio.

## Números adimensionales relevantes

Para analizar y comprender el régimen de flujo y el comportamiento de las interfaces en nuestro sistema, recurrimos al uso de números adimensionales clave, que nos permiten relacionar diferentes fuerzas y predecir el comportamiento del sistema:

**Número de Reynolds (Re)**

Este número nos indica la importancia relativa de las fuerzas inerciales (tendencia del fluido a seguir moviéndose) en comparación con las fuerzas viscosas (resistencia interna del fluido al flujo):

$$Re = \frac{\rho u D}{\mu} \tag{11}$$

En el régimen de flujo de Taylor el número de Reynolds tiene un flujo laminar (Re<100); cuyas capas del fluido se deslizan suavemente unas sobre otras sin mezclarse entre sí.

**Número de Capilaridad (Ca)**

Este número indica qué la importancia de las fuerzas viscosas en comparación con las fuerzas de tensión superficial:

$$Ca = \frac{u\mu}{\sigma} \tag{12}$$

El término $\sigma$ representa la tensión superficial. El número de capilaridad es fundamental para estimar el grosor de la película líquida que se forma entre la burbuja y la pared del canal. Para valores de $Ca < 0{,}01$, el modelo de Bretherton (1961) es una herramienta valiosa para predecir con precisión el grosor de esta película.

**Número de Weber (We)**

El número de Weber (We) establece la relación entre las fuerzas inerciales y las fuerzas de tensión superficial. En el estudio del flujo de Taylor en capilares, We adquiere una relevancia crítica para la caracterización de la deformación de las interfaces gas-líquido, así como para evaluar la estabilidad de las burbujas durante su desplazamiento a través del canal.

$$We = \frac{\rho u^2 L}{\sigma} \tag{13}$$

Donde en estos números adimensionales la $u$ representa la velocidad característica del flujo (correspondiente a la velocidad de la burbuja o del flujo líquido), $L$ es una longitud característica, $\sigma$ representa la tensión superficial entre las fases, $D = H$ es el diámetro del capilar.

En el contexto del modelo presente, dado que las burbujas presentan una morfología alargada y se desplazan en un régimen de flujo laminar controlado, se anticipa un número de Weber significativamente menor a la unidad; lo cual, valida la omisión de fenómenos de ruptura de burbujas y deformaciones abruptas en la interfaz, reforzando así la pertinencia de un enfoque de modelado simplificado.

### Consideraciones del modelo

- El flujo es laminar y bidimensional; es decir, variación en los ejes $(r, z)$ en coordenadas cilíndricas y/o $(x, y)$ en coordenadas rectangulares.
- El flujo es estacionario en un sistema de referencia que se mueve junto con la burbuja. Esto simplifica el análisis al eliminar la dependencia del tiempo.
- Las propiedades físicas de cada fase (líquido y gas) se mantienen constantes. El dominio es periódico, con burbujas de igual tamaño y separadas por una distancia constante.
- No consideraremos los efectos de la fase gaseosa. Esto significa que no resolveremos el campo de velocidad ni las ecuaciones para el gas.
- Las burbujas son lo suficientemente largas para permitir que se desarrolle un perfil de velocidad parabólico en la región intermedia del líquido.
- La tensión superficial actúa en la interfaz entre el líquido y el gas.

## Planteamiento matemático

La estructura matemática que describe el comportamiento del flujo de Taylor dentro del capilar, contempla la distribución de velocidades y la presión en la fase líquida, de tal manera que permite simplificar el análisis al no considerar, en esta etapa, los efectos dinámicos de la fase gaseosa. Se supone que el flujo es laminar, estacionario y bidimensional, y lo observaremos desde un punto de referencia que se mueve junto con una de las burbujas.

### Conservación de masa

Dado que se trabaja con un fluido incompresible, la ecuación de continuidad se expresa como:

$$\nabla \cdot (\mathbf{u}) = 0 \tag{14}$$

Donde $\mathbf{u} = (u_x, u_y)$ representa el campo de velocidades de la fase líquida, expresado en coordenadas cartesianas. Esta ecuación nos asegura que la cantidad de fluido que entra en un volumen dado es igual a la cantidad que sale.

## Conservación de cantidad de movimiento

Para describir el flujo laminar, la ecuación (3) se desprecia los efectos inerciales para flujo laminar (Reynolds bajo):

$$0 = -\nabla p + \mu(\nabla^2 \mathbf{u}) + \rho \mathbf{g} \tag{15}$$

En coordenadas cartesianas se obtiene:

$$0 = -\frac{\partial p}{\partial x} + \mu \left( \frac{\partial^2 u_x}{\partial x^2} + \frac{\partial^2 u_x}{\partial y^2} \right) \tag{15a}$$

$$0 = -\frac{\partial p}{\partial y} + \mu \left( \frac{\partial^2 u_y}{\partial x^2} + \frac{\partial^2 u_y}{\partial y^2} \right) \tag{15b}$$

Se usaron coordenadas cartesianas porque el área de estudio tiene forma rectangular. La validez de estas aproximaciones se manifiesta en los resultados numéricos, los cuales muestran una distribución del campo de velocidades y presiones coherente con el comportamiento físico esperado en este tipo de flujo. [12]

## Condición Interfacial: Tensión Superficial

En la interfaz donde se encuentran el gas y el líquido, se produce un salto de presión debido a la tensión superficial. Este fenómeno se describe mediante la Ley de Young-Laplace:

$$\Delta P = \sigma k \tag{16}$$

Este efecto es particularmente importante en los meniscos de la burbuja, donde se observa deformaciones tanto en el frente como en la extremidad.

## Espesor de la Película Líquida

Un rasgo distintivo del flujo de Taylor es la formación de una película delgada de líquido entre la burbuja y la pared del canal. El espesor de esta película está relacionado con el número de Capilaridad, y se puede estimar utilizando la correlación empírica propuesta por [13]:

$$\frac{\delta}{R} = 1{,}34\, Ca^{2/3} \tag{17}$$

Esta expresión permite incorporar los efectos de la viscosidad y la tensión superficial en la formación de la película, sin necesidad de resolver directamente la microescala de la capa delgada en el modelo.

## Formulación adimensional

Para simplificar el análisis y facilitar la comparación con otros estudios, se adimensional las ecuaciones anteriores tomando en cuenta:

$$\bar{x} = \frac{x}{R} \; ; \; \bar{y} = \frac{y}{R} \tag{18}$$

$$\bar{u} = \frac{u}{U} \tag{19}$$

$$\bar{P} = \frac{P}{\mu U/R} \tag{20}$$

La expresión de cantidad de movimiento, en ausencia de inercia, adopta la siguiente forma:

$$0 = -\nabla \bar{P} + \nabla^2 \bar{u} \tag{21}$$

Esto reduce la cantidad de parámetros y permite a identificas los efectos dominantes en el flujo.

## Representación esquemática del dominio

A continuación se presenta un esquema conceptual del dominio de simulación para una celda periódica de flujo tipo Taylor:


Figura 1: Dominio de las dos fases en el capilar

## Definición de condiciones iniciales y de frontera

A continuación, se especifican las condiciones iniciales y de frontera requeridas para resolver el conjunto de ecuaciones que rigen el sistema, tal como se describió previamente. Estas condiciones son esenciales para simular con precisión el comportamiento hidrodinámico del flujo de Taylor en un capilar, dentro de un dominio representativo de una celda periódica. La formulación se basa en un modelo bidimensional, en régimen estacionario y en un marco de referencia que se mueve con la burbuja.

### Condiciones Iniciales

Dado que el modelo se plantea en un régimen estacionario, la especificación de condiciones iniciales temporales, como en un análisis transitorio, no es estrictamente necesaria. Sin embargo, en una simulación numérica, se considera útil proporcionar un campo inicial aproximado de velocidad y presión para facilitar la convergencia del solucionador. En este caso, se puede emplear una condición inicial de velocidad parabolizada en la fase líquida, similar al perfil de Poiseuille para flujo laminar en un canal plano. Este campo inicial se expresa como:

$$u_x(y) = U_{\max}\left[1 - \left(\frac{2y}{H}\right)^2\right] \tag{22}$$

Para la presión inicial, se puede asumir un gradiente lineal descendente en la dirección del flujo, acorde con la caída de presión característica de un fluido viscoso en movimiento.

### Condiciones de Frontera

El dominio de simulación se representa como una celda periódica rectangular que contiene una única burbuja, rodeada por líquido y delimitada por las paredes del canal. Las condiciones de frontera aplicadas están diseñadas para reproducir el flujo periódico y capturar de manera precisa la interacción líquido-gas en las interfaces. A continuación, se describen las condiciones aplicadas en cada parte del dominio (véase la Figura 2).

**A. Pared húmeda**

Se aplica una condición de deslizamiento en las paredes del canal, lo que implica que la pared parece deslizarse en sentido opuesto al fluido:

$$\mathbf{u} = -\mathbf{u}_B \tag{23}$$

Esto refleja la interacción del fluido con la pared del capilar bajo una condición de deslizamiento, donde la velocidad del fluido en la interfaz con la pared es igual y opuesta a la velocidad de la burbuja, permitiendo el desarrollo del perfil de velocidad sin fricción viscosa en esa región.

**B. Interfaz líquido-gas**

En la región de contacto entre el gas y el líquido se imponen las siguientes condiciones:

- **Condición de velocidad:** La componente tangencial de la velocidad del líquido coincide con la velocidad de la burbuja. Dado que el dominio se mueve con la burbuja, esta

condición implica que el líquido fluye alrededor de una burbuja estacionaria. Se aplica una condición de pared móvil con velocidad opuesta al líquido.

- **Condición de tensión normal:** El salto de presión a través de la interfaz se relaciona con la tensión superficial, siguiendo la ley de Young-Laplace:

$$\Delta P = \sigma k \tag{24}$$

Esta condición se aplica directamente mediante la implementación de una fuerza capilar localizada en la interfaz, dependiendo del enfoque utilizado en el software de simulación (por ejemplo, el método Level Set o el método de seguimiento de frontera).

**C. Entrada y salida del dominio**

Dado que el sistema representa una celda periódica, se aplican condiciones de periodicidad para emular la secuencia de burbujas y celdas idénticas:

$$\mathbf{u}(0, y) = \mathbf{u}(L, y) \tag{25}$$

$$p(0, y) = p(L, y) + \Delta p \tag{26}$$

Donde:

- $L$: es la longitud de la celda periódica (longitud de la burbuja + longitud del espacio líquido entre burbujas)
- $\Delta p$: es la caída de presión media por celda

Estas condiciones aseguran que el flujo que sale por un extremo de la celda reaparece de forma idéntica por el otro, manteniendo la simetría y periodicidad del flujo de Taylor. A continuación, se presenta un esquema del dominio con las condiciones de frontera aplicadas (Figura 2):

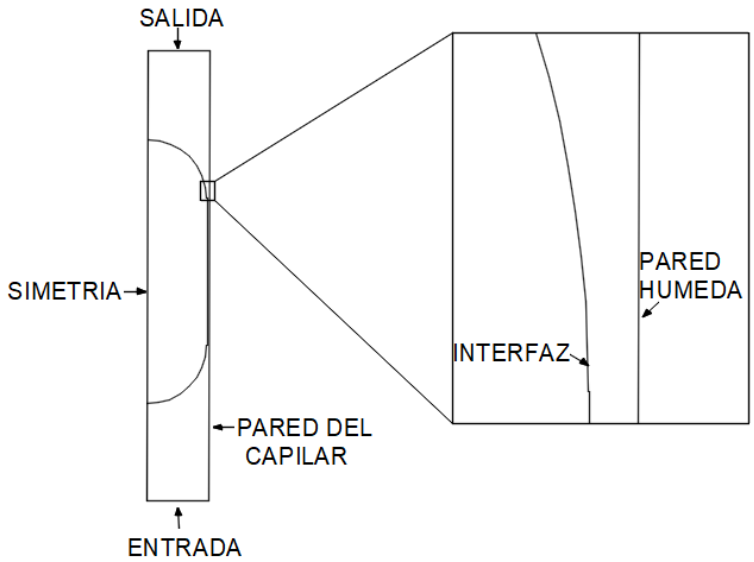



Figura 2: Esquema del dominio con las condiciones de frontera aplicadas.

## Consideraciones Adicionales

- **Dominio móvil:** La implementación en un marco de referencia móvil facilita la simulación del flujo estacionario alrededor de una burbuja fija, simplificando el problema numérico.
- **Tensión superficial:** En los modelos numéricos, la tensión superficial se puede implementar como una fuerza distribuida (métodos de captura de interfaz como Level Set o Volume of Fluid).
- **Film líquido:** La película líquida que se forma entre la burbuja y la pared se representa de forma implícita mediante la resolución del perfil de velocidad y presión en dicha zona, sin necesidad de mallas extremadamente finas, gracias al uso de correlaciones como la de Bretherton.

# Desarrollo de la Simulación Numérica

La simulación numérica se llevó a cabo utilizando el software [1] 6.2, aprovechando las capacidades del módulo de Flujo Laminar en 2D, conjuntamente con el método de malla estacionaria implementado en un marco de referencia móvil. El propósito fundamental de esta simulación es resolver de manera precisa proveniente de las ecuaciones de Navier-Stokes en el dominio previamente definido, considerando de forma integral los efectos inducidos por la tensión superficial, el confinamiento capilar, y la presencia periódica de burbujas de gas inmersas en una matriz líquida. En la presente sección, se detalla el procedimiento paso a paso que se siguió para la implementación del modelo, abarcando desde la construcción inicial del dominio geométrico hasta la obtención final de los resultados relevantes.

### Construcción del Dominio Geométrico

En esta etapa, se diseñó un dominio bidimensional que representa una celda periódica del flujo tipo Taylor. El dominio se define como un rectángulo que contiene una burbuja de gas con una forma elíptica bien definida, la cual se encuentra centrada en el eje principal del canal.

Tabla 1: Parámetros Geométricos

| Parámetro | Valor típico | Descripción |
|---|---|---|
| Altura del canal (H) | 100 $\mu$m | Diámetro equivalente del capilar |
| Longitud de la celda (L) | 800 $\mu$m | Longitud de la burbuja más el espacio líquido |
| Longitud de la celda (Lb) | 500 $\mu$m | Aproximación al caso experimental típico |
| Radio de curvatura del menisco (R) | 50 $\mu$m | Relación con la curvatura de la burbuja |
| Espesor de la película líquida ($\delta$) | 5 $\mu$m | separa la burbuja de las paredes del canal |

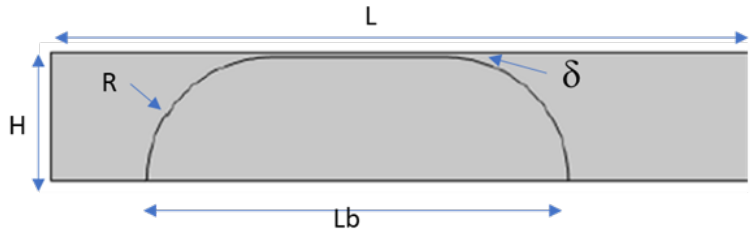



Figura 3: Geometría del Dominio 2D en COMSOL Multiphiscs®

### Configuración del Módulo Físico

El modelado numérico del flujo multifásico tipo Taylor en microcanales, se empleó el software [1] 6.2. El objetivo principal es reproducir y analizar el comportamiento dinámico de este tipo de flujo, caracterizado por la alternancia de burbujas de gas y tapones de líquido, con un enfoque particular en la geometría de los bulbos, el espesor de la película líquida, la distribución de velocidades internas y los efectos interfaciales dominantes.

### Configuración del Modelo

Para optimizar la eficiencia computacional, sin sacrificar la precisión del modelo, el sistema físico se representó en un entorno 2D simétrico. Esta aproximación es particularmente adecuada para canales capilares circulares, ya que permite reducir significativamente el número de elementos necesarios para la discretización del dominio, minimizando así el gasto computacional asociado a la simulación.

La simulación se desarrolló utilizando la interfaz de física Flujo Multifásico – Flujo Bifásico, Conjunto de Nivel – Flujo Laminar, disponible en el módulo CFD (Computational Fluid Dynamics) de [1] 6.2. Esta combinación de interfaces permite simular el comportamiento de dos fases inmiscibles, típicamente gas y líquido, bajo condiciones de flujo laminar. La interfaz de Flujo Multifásico se encarga de resolver las ecuaciones de conservación de masa y momento para cada fase, mientras que la interfaz de Conjunto de Nivel (Level Set) se utiliza para rastrear la evolución de la interfaz entre las fases.

### Método de Conjunto de Nivel (Level Set)

El método de conjunto de nivel (Level Set) es una técnica numérica ampliamente utilizada para rastrear la evolución de interfaces en flujos multifásicos. En este método, la interfaz entre las fases se representa mediante una función escalar $\phi$, conocida como función de nivel, que varía continuamente entre 0 y 1 dentro del dominio. El valor $\phi = 0$ representa la fase líquida pura, $\phi = 1$ corresponde a la fase gaseosa pura, y los valores intermedios definen una zona difusa que representa la interfaz entre ambas fases.

La evolucion temporal de esta interfaz está regida por una ecuación de advección acoplada con difusión artificial. La ecuación describe el transporte de la función de nivel debido al campo de velocidades del flujo, mientras que la difusión artificial se introduce para mantener un espesor de interfaz constante y controlado, evitando así las oscilaciones numéricas que pueden surgir debido a la discretización del dominio.

Las propiedades del sistema, como la densidad $(\rho)$, la viscosidad dinámica $(v)$ y la tensión superficial$(\sigma)$, se definieron como funciones del campo $(\phi)$, mediante una interpolación entre los valores correspondientes a cada fase. Esto garantiza una transición suave en la interfaz sin necesidad de una malla que se adapte específicamente a su contorno. Por ejemplo, la densidad efectiva se expresa como:

$$\rho_{\text{efectiva}} = \phi \cdot \rho_{\text{gas}} + (1 - \phi) \cdot \rho_{\text{líquido}} \tag{27}$$

De manera similar, la viscosidad efectiva y otros parámetros físicos relevantes se definen en función del valor local de $\phi$. Este enfoque facilita el tratamiento numérico del sistema.

## Flujo Laminar y la ecuación de cantidad de movimiento

El flujo se consideró laminar debido a los bajos valores del número de Reynolds característicos del sistema. Para ello, se empleó la ecuación de cantidad de movimiento para fluidos incomprensibles acopladas con la ecuación de conservación de masa. La resolución de este conjunto de ecuaciones permitió obtener el campo de velocidades y la distribución de presiones en todo el dominio:

$$\rho \frac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u} \cdot \nabla \mathbf{u}) = -\nabla p + \mu \nabla^2 \mathbf{u} + \mathbf{F} \tag{28}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{29}$$

## Tensión Superficial y Fuerza de Cuerpo Continua (CSF)

El término de tensión superficial se incorporó mediante la formulación de fuerza de cuerpo continua (Continuum Surface Force, CSF), basada en el gradiente de $\phi$. La curvatura local de la interfaz se calcula automáticamente por [1] a partir del gradiente normalizado y se aplica como una fuerza localizada en la región interfacial:

$$\mathbf{F}_{\text{tensión superficial}} = \sigma \cdot k \nabla \phi \tag{30}$$

la física real del flujo tipo Taylor, permite simular este tipo de fenómenos como la formación de la película líquida en las paredes del capilar y la estabilidad general del patrón bifásico.

## Reparametrización de la Interfaz (Reinitialization)

Además, se habilitó el ajuste de reparametrización de la interfaz (reinitialization), que actúa periódicamente para preservar la forma deseada del perfil de $\phi$ sin difuminar excesivamente la interfaz. Esta reparametrización se controla mediante el parámetro $\epsilon$ (espesor de la interfaz difusa), el cual se ajustó en función del tamaño característico de malla y del tamaño esperado del canal.

El conjunto de estas configuraciones permitió construir un modelo robusto capaz de reproducir, de forma numérica, el comportamiento dinámico del flujo de Taylor, con especial énfasis en la geometría de los bulbos, el espesor de la película líquida, la distribución de velocidades internas y los efectos interfaciales dominantes. Este modelo proporciona una herramienta valiosa para comprender y optimizar el diseño de microdispositivos basados en el flujo de Taylor.

## Definición de Condiciones de Frontera en [1]

Las condiciones de frontera que fueron definidas anteriormente, se traducen al lenguaje específico del software [1] de la siguiente manera:

| Borde del Dominio | Condición Impuesta en [1] |
|---|---|
| Paredes superior e inferior | No deslizamiento (u = 0) |
| Interfaz burbuja-líquido | Condición de pared móvil (u = -Ub) |
| Entrada y salida | Condición periódica con caída de presión |

Tabla 2: Condiciones de frontera implementadas en [1].

Para representar de manera precisa la interfaz entre el líquido y el gas, se utilizó una pared interior curva con condiciones personalizadas. Se implementó una condición de tensión superficial utilizando el módulo de "Multiphase Flow", o alternativamente, como una fuerza normal definida manualmente mediante la siguiente expresión:

$$\mathbf{f}_a = \sigma k \bar{\mathbf{n}} \tag{31}$$

Donde $k$ representa la curvatura de la interfaz, la cual se calcula localmente en cada punto de la misma, y $\bar{\mathbf{n}}$ es el vector normal a la superficie.

## Malla del Dominio

La malla utilizada para la discretización del dominio se caracterizó por ser estructurada y refinada de manera no uniforme. Se aplicó un mayor nivel de resolución cerca de la burbuja y en las proximidades de las paredes del canal, con el objetivo de capturar de forma precisa los gradientes de velocidad y el espesor de la película líquida que se forma en estas zonas.

Respecto al tipo de elementos finitos utilizados fue de la forma triangular alrededor de la interfaz líquido-gas para capturar la geometría curva con precisión. En la Figura 4 se observa el dominio del capilar con este mallado, así como un acercamiento para la interfaz liquido-gas.

## Criterios de Mallado

- Relación de aspecto: Se procuró mantener una relación de aspecto menor a 5 para evitar elementos excesivamente alargados.

- Tamaño mínimo de malla: Se estableció un tamaño mínimo de malla de aproximadamente 1 $\mu$m en las zonas críticas.
- Independencia de malla: Se realizó una verificación de independencia de malla para asegurar que los resultados no dependan significativamente del nivel de refinamiento.

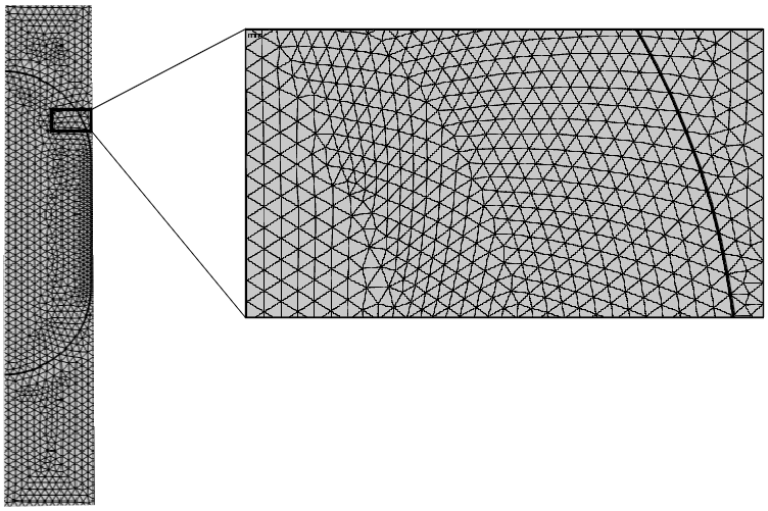


Figura 4: Malla determinada para el dominio del sistema

## Solución del Modelo

Se empleó un solucionador estacionario del tipo segundo orden (P2+P1) para la velocidad y la presión, implementando una estabilización de tipo streamline difusion para mejorar la convergencia en las zonas cercanas a la interfaz.

## Parámetros del Solucionador

- Tolerancia relativa: Se fijó una tolerancia relativa de $10 \times 10^{-6}$ para asegurar la precisión de la solución.
- Método: Para la solución numérica del modelo, se utilizó el método de diferenciación hacia atrás BDF (Backward Differentiation Formula), que es un esquema implícito adecuado para resolver ecuaciones diferenciales ordinarias y en derivadas parciales de naturaleza rígida. El método BDF es particularmente eficiente para problemas de dinámica de fluidos donde la estabilidad numérica es crucial, ya que permite manejar variaciones de velocidad y presión sin la restricción de pequeños pasos de tiempo que imponen los métodos explícitos. En [1], BDF es el método predeterminado para resolver ecuaciones transitorias debido a su capacidad para trabajar con pasos de tiempo adaptativos, optimizando la precisión y estabilidad de la simulación. En este caso, se utilizó con un esquema de órdenes variables, donde el software ajusta automáticamente el orden del método (de 1 a 5) para optimizar la convergencia y la eficiencia computacional.

- Pre-condicionado automático: Se habilitó el pre-condicionado automático para mejorar la convergencia del solucionador.
- Control adaptativo del paso: Se activó el control adaptativo del paso para ajustar dinámicamente el tamaño del paso de tiempo en función de la convergencia. La simulación convergió típicamente en un rango de 10 a 20 iteraciones, dependiendo del grado de refinamiento de la malla y las condiciones impuestas.

### Validación Numérica Interna

Se realizó una verificación interna del modelo mediante el cálculo del número de Capilaridad simulado, utilizando la siguiente expresión del número de capilaridad, $C_a = \mu U/\sigma$, y la ecuación de Bretherton, $\delta/R = 1{,}34\, C_a^{2/3}$, valida preliminarmente el comportamiento físico del modelo y su capacidad para capturar los fenómenos relevantes (Brethreton, 1961).

### Exportación de Datos

Los resultados de la simulación se exportan en formatos compatibles con Excel para facilitar su posterior análisis y procesamiento. Los datos exportados incluyen:

- Curvas de presión en función de la posición a lo largo del canal.
- Perfiles de velocidad a diferentes alturas dentro del canal.
- Datos del espesor de la película líquida en función de la velocidad.

## Evaluación y Validación del Modelo

La implementación de un modelo numérico destinado a representar el flujo tipo Taylor en capilares exige una evaluación rigurosa de su validez, confiabilidad y coherencia física inherente. En la presente sección, se establece de manera detallada la metodología que se emplea para validar el modelo propuesto, especificando con precisión qué resultados se espera observar, bajo qué fundamentos teóricos sólidos se basa dicha expectativa, y mediante qué criterios cuantitativos y cualitativos se valorará exhaustivamente el desempeño del modelo en cuestión. Se considera de suma importancia enfatizar que la validación aquí planteada no se fundamenta en la utilización de datos experimentales propios, sino que se apoya en comparaciones rigurosas con soluciones analíticas bien establecidas y correlaciones ampliamente reconocidas en la literatura científica especializada.

### Objetivo de la Validación

El objetivo primordial de esta etapa crucial consiste en comprobar de manera fehaciente que el modelo numérico reproduce de forma precisa y fiel los fenómenos hidrodinámicos característicos del flujo bifásico tipo Taylor, respetando los supuestos establecidos de antemano y operando dentro del rango de condiciones físicas especificadas para el sistema en estudio. En particular, esta validación exhaustiva abarca:

- Forma y comportamiento de la burbuja: Se valida la forma y el comportamiento de la burbuja dentro del canal capilar, asegurando que se ajusten a las expectativas teóricas.
- Desarrollo de la película líquida: Se verifica el desarrollo de una película líquida delgada y estable que rodea a la burbuja, confirmando su existencia y estabilidad bajo las condiciones simuladas.
- Distribución de campos: Se evalúa la distribución del campo de velocidades y presiones en la fase líquida, comprobando que los patrones observados sean consistentes con las predicciones teóricas.
- Respuesta del modelo: Se analiza la respuesta coherente del modelo ante variaciones controladas de parámetros clave, tales como la viscosidad, la tensión superficial o la velocidad del flujo.

## Supuestos y su Influencia en la Evaluación

Los siguientes supuestos fundamentales ejercen una influencia significativa en la forma en que el modelo será validado y actúan como un marco conceptual sólido para la interpretación de los resultados obtenidos. Estos supuestos, por lo tanto, deben tenerse en cuenta en todo momento:

## Parámetros a Observar en la Validación

La validación rigurosa del modelo se enfocó en los siguientes aspectos cuantificables, los cuales proporcionan una base sólida para evaluar su desempeño:

- Espesor de la película líquida ($\delta$): Una de las características más distintivas del flujo tipo Taylor es la formación de una película delgada de líquido que separa la burbuja de las paredes del canal. Este espesor está determinado por el equilibrio entre las fuerzas viscosas y capilares y se ha demostrado que sigue una relación empírica bien establecida con el número capilar, expresada por la ley de Bretherton (Bretherton, 1961):

  Expectativa: Se espera que los valores de $\delta/R$ obtenidos numéricamente sigan de cerca esta ley, especialmente en el rango de $Ca < 0{,}1$, que corresponde al régimen clásico de flujo capilar dominado por los efectos interfaciales.

- Perfil de velocidad en la fase líquida: En la región entre burbujas, donde el flujo es puramente líquido y está completamente desarrollado, se espera un perfil de velocidad parabólico, característico del flujo de Poiseuille en microcanales.

  Expectativa: El modelo debe reproducir esta forma en regiones alejadas de la burbuja. Cerca de la interfaz, el perfil puede distorsionarse debido a efectos de la curvatura y la tensión superficial, pero debe mantenerse estable y continuo.

- Campo de presión y salto interfacial. La burbuja debe generar una diferencia de presión entre su parte frontal y trasera, causado por la curvatura de la interfaz y modelada por la ley de Young-Laplace.

Expectativa: El modelo debe mostrar un salto de presión claramente identificable en el contorno de presión y un gradiente lineal en la región líquida, consistente con el flujo laminar desarrollado.

- Conservación de masa y cantidad de movimiento: El modelo debe conservar la masa en todo momento, es decir, el caudal de entrada debe igualar al caudal de salida y no debe existir acumulación artificial. Del mismo modo, las ecuaciones de Navier-Stokes deben resolverse de forma que el campo de velocidades y presiones sea físicamente coherente.

  Expectativa: Se espera que el balance global de masa y momento se conserve, lo cual puede verificarse mediante herramientas de postproceso en [1] o mediante la integración del flujo en diferentes secciones del dominio.

## Estrategia de Validación del Modelo

Una vez ejecutadas las simulaciones, se seguirán los siguientes pasos con el objetivo de validar el modelo propuesto:

- Cálculo del número capilar $Ca$ con base en la velocidad de la burbuja y las propiedades del líquido.
- Comparación con datos experimentales y resultados para diversos números adimensionales.
- Análisis del contorno de presión, identificando el salto interfacial y el gradiente en la región líquida.
- Realización de estudios de sensibilidad, modificando parámetros como velocidad, diámetro y espesor para verificar la coherencia del modelo frente a variaciones.

## Limitaciones Previstas en la Validación

Es importante reconocer y tener en cuenta las limitaciones que podrían surgir durante el proceso de validación:

- El modelo se basa en un dominio bidimensional, por lo que las comparaciones con resultados tridimensionales (experimentales o simulados) deben hacerse con precaución.
- La resolución numérica puede afectar la estimación de $\delta$, especialmente cuando este valor es muy pequeño. Se deberán realizar pruebas de refinamiento de malla para mitigar este efecto.
- Las condiciones ideales del modelo, como el flujo perfectamente simétrico y las burbujas idénticas y espaciadas, pueden no reproducirse fielmente en sistemas experimentales reales, lo cual limita la comparación directa con datos experimentales.

# Discusión de resultados

Se realizaron treinta y dos simulaciones en [1] versión 6.2, con el propósito de investigar el comportamiento característico del flujo de Taylor en el interior de conductos capilares. Las pruebas abarcaron un espectro diversificado de escenarios, definidos por variaciones en las condiciones geométricas del sistema, las propiedades físicas de los fluidos involucrados y las condiciones de frontera impuestas.

Con el objetivo de organizar y facilitar la interpretación de los datos obtenidos, las simulaciones fueron categorizadas bajo dos enfoques metodológicos primordiales: el primero, denominado condición de entrada fija ('E'), y el segundo, caracterizado por la implementación de condiciones de periodicidad ('CP'). Cada uno de estos enfoques fue sometido a una evaluación sistemática, explorando diversas combinaciones paramétricas. Específicamente, se manipularon los siguientes parámetros clave: el diámetro interno del capilar fue de 3 y 4 mm; el espesor de la película líquida que recubre la pared del capilar, ajustado a valores de 25 y 50 $\mu$m; y, finalmente, la velocidad de flujo impuesta al sistema, que osciló entre 0.02 a 0.5 m/s. La combinación de estos parámetros permitió obtener una comprensión detallada de la influencia de cada uno en el comportamiento del flujo de Taylor.

El enfoque metodológico de entrada fija ('E') se fundamenta en la simulación de un flujo en desarrollo a partir del punto de acceso al capilar. Esta aproximación permite una observación detallada de la evolución hidrodinámica del sistema, con especial énfasis en las regiones iniciales donde la burbuja comienza su desplazamiento y se establece el perfil característico de la película líquida adyacente a la pared del conducto. La configuración 'E' resulta particularmente útil para capturar los fenómenos transitorios y las adaptaciones del flujo en las primeras etapas de su recorrido.

En contraste, el enfoque de condición periódica ('CP') se basa en la simulación de un dominio espacial reducido, el cual se replica de manera periódica a lo largo de la dirección del flujo. Esta estrategia es ideal para representar secciones del flujo donde ya se ha alcanzado un estado estabilizado, permitiendo un análisis eficiente de las características del flujo en régimen estacionario. La comparación directa entre los resultados obtenidos mediante los enfoques 'E' y 'CP' resulta fundamental para identificar las diferencias entre el desarrollo inicial del flujo y su comportamiento en condiciones de régimen establecido. Esta distinción es crucial para comprender la influencia de los efectos de entrada en la dinámica global del sistema.

Dentro del marco del análisis cuantitativo se evaluó tres números adimensionales de relevancia fundamental en el campo de la mecánica de fluidos: el número de Reynolds (Re), el número de Capilaridad (Ca) y el número de Weber (We). Estos parámetros adimensionales cumplen la función esencial de caracterizar el régimen de flujo predominante y las fuerzas dominantes que actúan en cada una de las simulaciones realizadas.

A continuación, se presenta la tabla con los valores calculados para cada condición:

Tabla 3: Números adimensionales calculados para cada simulación

| Código de simulación | d (mm) | $\delta$ (µm) | U (m/s) | Re | Ca | We |
|---|---|---|---|---|---|---|
| SIM R3MM - E25UM - U0.02 | 3 | 25 | 0.02 | 59.5 | 0.0140 | 0.0610 |
| SIM R3MM - E25UM - U0.05 | 3 | 25 | 0.05 | 148.7 | 0.0350 | 0.3810 |
| SIM R3MM - E25UM - U0.2 | 3 | 25 | 0.2 | 595.1 | 0.1400 | 6.1070 |
| SIM R3MM - E25UM - U0.5 | 3 | 25 | 0.5 | 1487.70 | 0.3510 | 38.1680 |
| SIM R3MM - E50UM - U0.02 | 3 | 50 | 0.02 | 59.5 | 0.0140 | 0.0610 |
| SIM R3MM - E50UM - U0.05 | 3 | 50 | 0.05 | 148.7 | 0.0350 | 0.3810 |
| SIM R3MM - E50UM - U0.2 | 3 | 50 | 0.2 | 595.1 | 0.1400 | 6.1070 |
| SIM R3MM - E50UM - U0.5 | 3 | 50 | 0.5 | 1487.70 | 0.3510 | 38.1680 |
| SIM R4MM - E25UM - U0.02 | 4 | 25 | 0.02 | 79.4 | 0.0190 | 0.1090 |
| SIM R4MM - E25UM - U0.05 | 4 | 25 | 0.05 | 198.5 | 0.0490 | 0.6810 |
| SIM R4MM - E25UM - U0.2 | 4 | 25 | 0.2 | 793.9 | 0.1970 | 10.8920 |
| SIM R4MM - E25UM - U0.5 | 4 | 25 | 0.5 | 1984.90 | 0.4930 | 68.0750 |
| SIM R4MM - E50UM - U0.02 | 4 | 50 | 0.02 | 79.4 | 0.0190 | 0.1090 |
| SIM R4MM - E50UM - U0.05 | 4 | 50 | 0.05 | 198.5 | 0.0490 | 0.6810 |
| SIM R4MM - E50UM - U0.2 | 4 | 50 | 0.2 | 793.9 | 0.1970 | 10.8920 |
| SIM R4MM - E50UM - U0.5 | 4 | 50 | 0.5 | 1984.90 | 0.4930 | 68.0750 |

## Análisis de tendencias respecto al enfoque de entrada

### Enfoque de entrada fija E

Estas simulaciones implementadas bajo este enfoque de entrada fija, permiten observar detalladamente el desarrollo integral del flujo, desde su inicio de la entrada del capilar hasta la conformación de un perfil establecido.

En la figura 5 se presenta las simulaciones representativas de este enfoque identificadas como “SIM_R3MM - E25UM_U0.02”. Se aprecia la evolución de las regiones caracterizadas por la aceleración del fluido, así como la formación y el desarrollo de la película líquida que recubre la pared del conducto; además, revela con mayor claridad la existencia de gradientes de presión y velocidad a lo largo del eje del capilar, proporcionando información valiosa sobre la dinámica del flujo en su etapa de desarrollo.

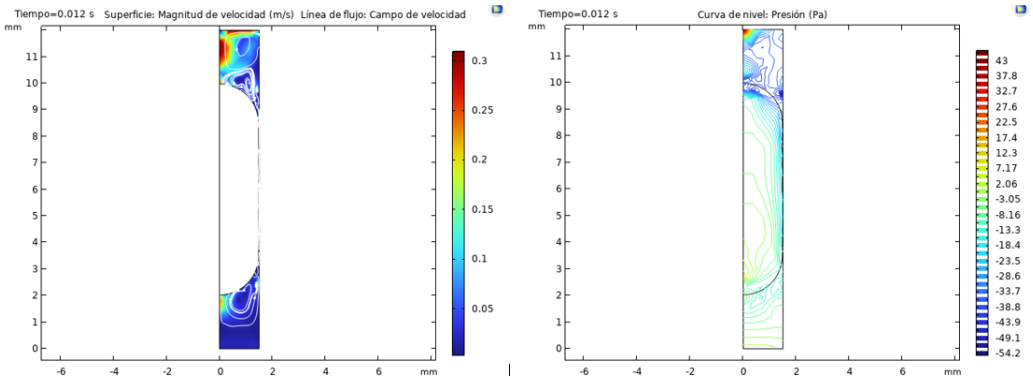



Figura 5: Resultados de la simulación SIMR3MM-E25UM-U02-E. Del lado izquierdo está el gráfico de velocidad, mientras el lado derecho corresponde a de presión.

### Enfoque de entrada fija CP

En contraste, las simulaciones que emplean la condición periódica CP, ejemplificadas por el caso "SIM_R3MM - CP25_U0.02", exhiben perfiles de flujo que se caracterizan por una mayor simetría y estabilidad. En estos escenarios la burbuja tiende a conservar su forma a lo largo del dominio simulado, y la película líquida presenta un espesor más uniforme. La condición de periodicidad impuesta favorece la convergencia hacia un estado estacionario, minimizando las fluctuaciones y los efectos transitorios observados en las simulaciones con entrada fija.

La Figura 5 muestra la simulación 'E' (velocidad de 0.02 m/s con un diámetro de 3 mm), donde el un campo de velocidades tiene un alto gradiente en la región delantera de la burbuja. Esta característica no se observa en la simulación correspondiente con condición periódica ('CP'), donde se aprecia una distribución de velocidades más homogénea. Esta disparidad evidencia de manera concluyente que las condiciones de frontera impuestas ejercen un impacto directo y significativo sobre la estabilidad y la madurez del perfil de flujo, resaltando la importancia de la selección adecuada del enfoque metodológico en función de los objetivos del estudio.

## Influencia del diámetro del capilar entre 3 mm y 4 mm

El análisis comparativo entre simulaciones realizadas con diámetros de capilar diferentes (3 y 4 mm) revela una clara influencia del tamaño del conducto en las características de flujo multifásico. Específicamente, se observa que, al incrementar el diámetro del capilar se produce un aumento en los valores de los números adimensionales de Reynolds (Re), Capilaridad (Ca) y Weber (We).

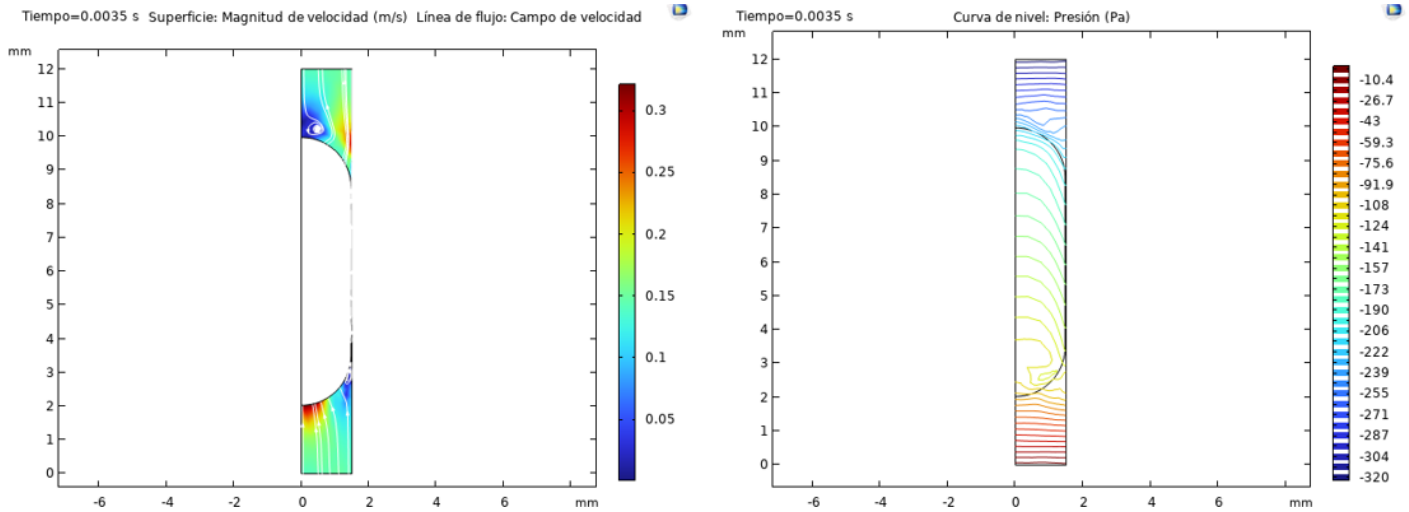



Figura 6: Resultados de la simulacion SIM R3MM - E25UM - U0.02 - CP.
*De el lado izquierdo esta el grafico de velocidad, mientras que el del lado derecho es el grafico de presión.*

Este incremento en los números adimensionales sugiere una transición hacia un régimen de flujo donde las fuerzas inerciales adquieren mayor preponderancia en relación con las fuerzas viscosas y de tensión superficial. En consecuencia, las burbujas tienden a experimen-

tar una mayor elongación en la dirección del flujo, y se observa un aumento en la velocidad máxima del fluido.

La evidencia de estos efectos se manifiesta visualmente en la Figura 7. En particular, se observa que la burbuja en los capilares de 4 mm de diámetro presenta una mayor extensión longitudinal respecto a capilares de 3 mm.

Adicionalmente, se identifican zonas de mayor velocidad en la película líquida que recubre la pared del capilar, siendo este fenómeno particularmente notorio en la simulación identificada como "SIM_R4MM - E25UM_U0.5". Estos hallazgos confirman que el incremento en el diámetro del capilar favorece un régimen de flujo más inercial, con una mayor deformación de las burbujas y una distribución de velocidades más heterogénea.

## Influencia del espesor de película (25 $\mu$m vs 50 $\mu$m)

Si bien el análisis de los números adimensionales de Reynolds, Capilaridad y Weber no revela una variación significativa en función del espesor de la película líquida, el impacto físico de este parámetro en la dinámica del flujo multifásico es innegable y observable.

En particular, se aprecia que a menor espesor de la película líquida (25 $\mu$m), el fluido se ve compelido a fluir a través de un canal de menor sección transversal. Esta constricción impone la generación de mayores gradientes de presión y velocidad en la región circundante a la burbuja. En las simulaciones donde se implementa un espesor de película de 25 $\mu$m, como en el caso de "SIM R3MM - CP25UM_U0.2", se evidencia la formación de zonas de alta presión en la región frontal de la burbuja, lo cual sugiere una mayor resistencia al avance de la misma.

En contraste, cuando el espesor de la película líquida se incrementa a 50 $\mu$m, como en la simulación "SIM R3MM - CP50UM_U0.2", se observa que la película fluye de manera más suave y uniforme a lo largo de la pared del capilar. A mayor espesor de la película facilita el paso del fluido, reduciendo la resistencia viscosa que se opone al movimiento de la burbuja y disminuyendo, por ende, los gradientes de presión en la región interfacial.

Estos hallazgos demuestran que, aunque no presentan variaciones drásticas en los números adimensionales globales, el espesor de la película líquida modula significativamente la distribución local de presiones y velocidades en el sistema, afectando la dinámica del flujo multifásico.

## Variación velocidad del flujo

La velocidad del flujo ejerce una influencia directa y significativa sobre los números adimensionales que caracterizan el comportamiento del sistema multifásico. En particular, se observa que el incremento de la velocidad de flujo impacta directamente sobre los valores de los números de Reynolds, Capilaridad y Weber.

A velocidades de flujo bajas ($u < 0{,}02$ m/s), tal como se evidencia en la simulación "SIM R3MM - CP25UM - U0.02", se observa que la burbuja mantiene una forma bien definida y estable a lo largo de su trayectoria. En estas condiciones, el número de Weber (We) presenta valores significativamente menores a la unidad ($We \ll 1$), indica que las fuerzas de tensión

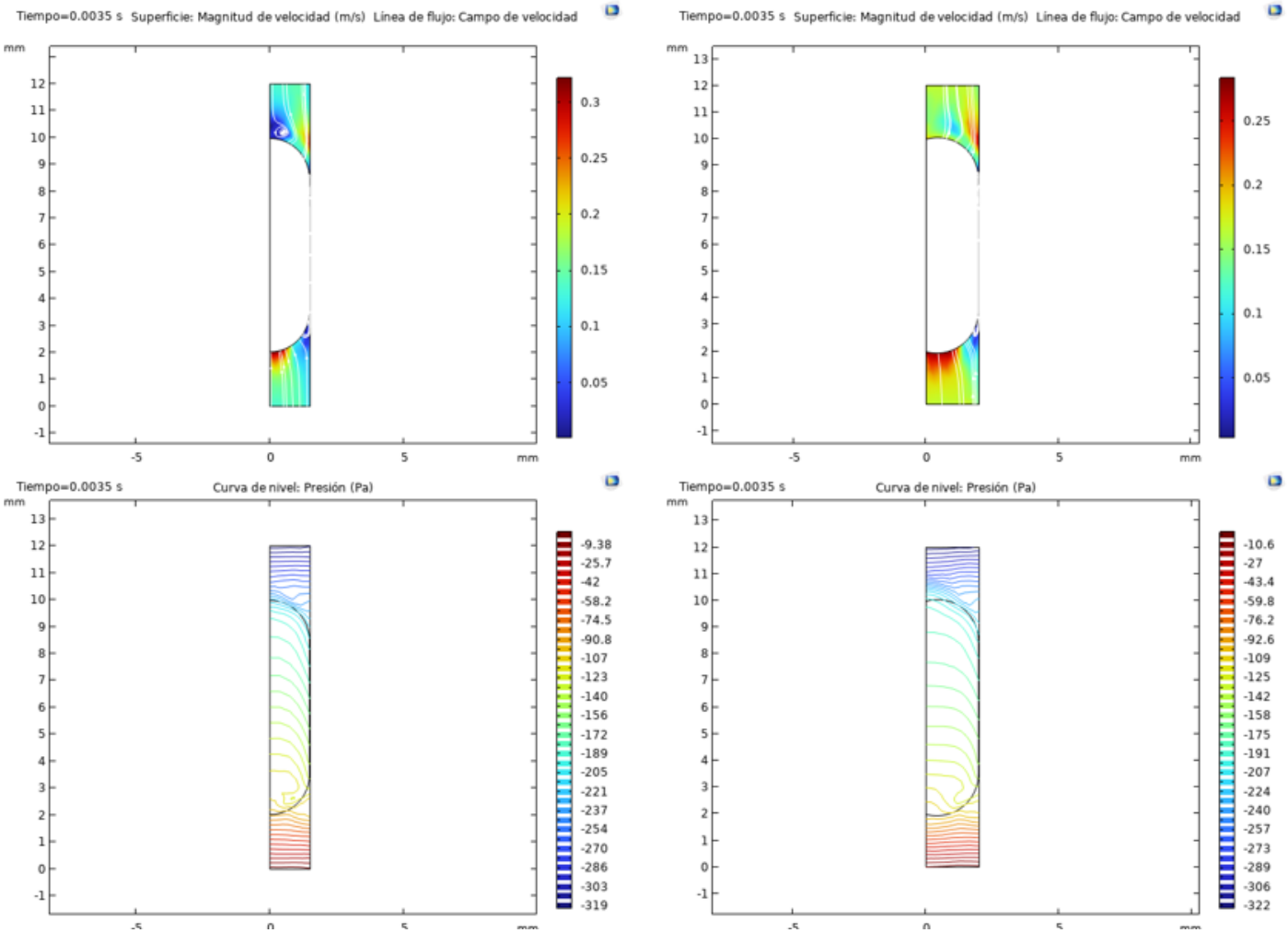



Figura 7: Comparación de resultados de la simulacion SIM e25UM - U0.5 - CP con diametros de 3 y 4 mm
*De el lado izquierdo esta los graficos del capilar de 3mm, mientras que el del lado derecho son los graficos del capilar de 4mm*

superficial dominan sobre las fuerzas inerciales y viscosas. En este régimen, la burbuja se mantiene cohesionada y la interfaz líquido-gas presenta una curvatura pronunciada.

En contraste, a velocidades de flujo elevadas ($u > 0{,}5$ m/s), tal como se observa en la Figura 9 en la simulación "SIM R4MM - CP25UM - U0.5", se aprecia una elongación significativa de la burbuja en la dirección del flujo, así como una mayor deformación del menisco que separa las fases fluidas. En estas condiciones, el número de Weber (We) alcanza valores superiores a 60 ($We > 60$), lo cual indica que las fuerzas inerciales dominan por completo sobre las fuerzas de tensión superficial. En este régimen, la burbuja se estira y deforma, adoptando una forma menos esférica y más alargada debido a la influencia de la inercia del fluido.

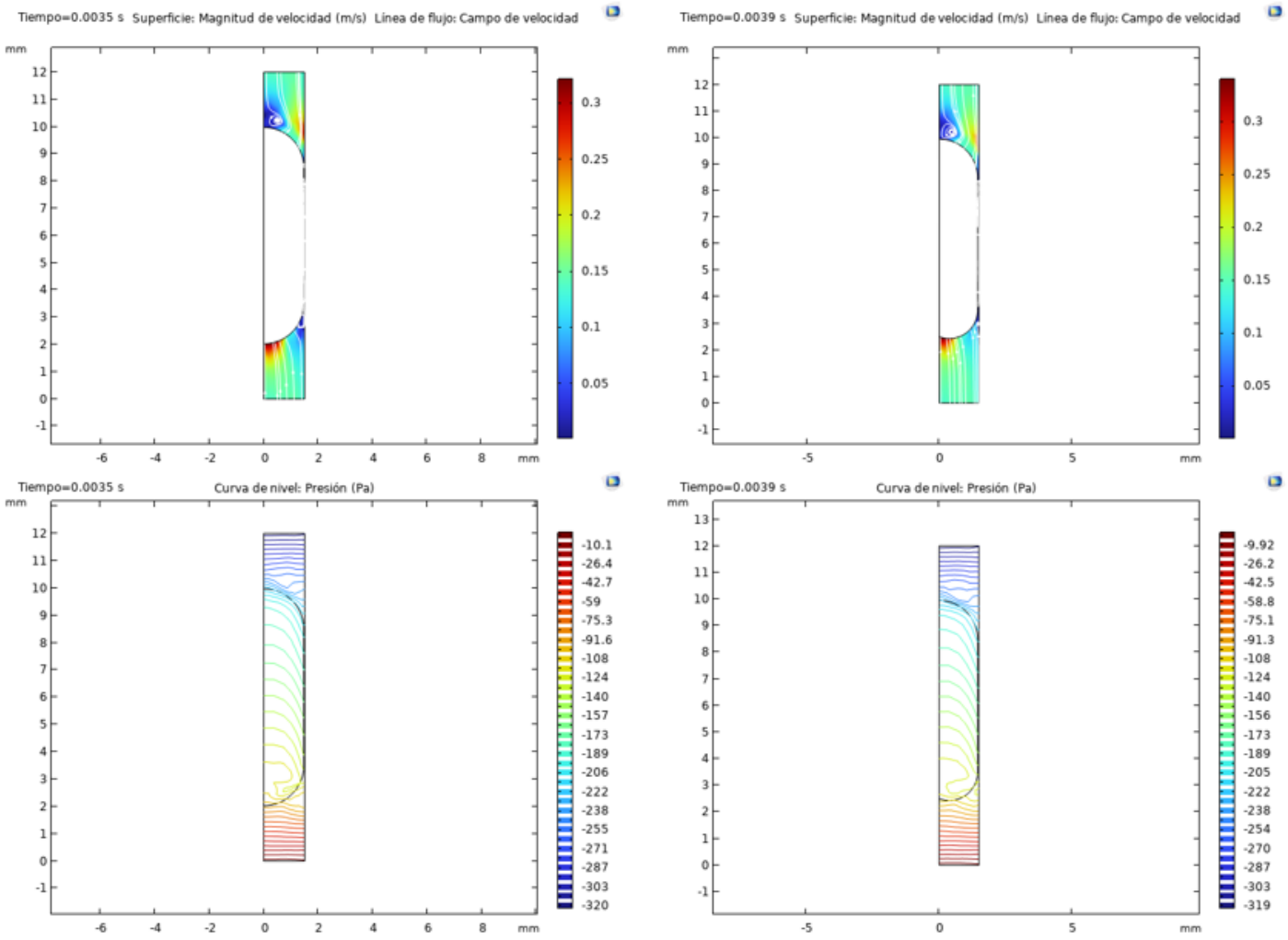



Figura 8: Comparación de resultados de la simulacion SIM R3 - U0.5 - CP con espesor de la pelicula de 25 µm y 50 µm
*De el lado izquierdo esta los graficos de de la pelicula de 25 µm, mientras que el del lado derecho son los graficos de la pelicula de 50 µm*

## Comparación con datos experimentales previos

El propósito de la presente sección radica en establecer una comparación entre los resultados numéricos obtenidos a través de las simulaciones multifásicas implementadas en [1] y los datos experimentales en la tesis titulada "Absorción de CO2 en capilares" (Morales, 2020; Shao, 2010; Etminan, 2021). Esta confrontación metodológica tiene como objetivo primordial validar la pertinencia y la precisión del modelo de simulación propuesto, mediante la identificación de coincidencias o discrepancias relevantes entre los resultados numéricos y los datos empíricos. La comparación sistemática entre ambos conjuntos de datos permite profundizar en la comprensión de los fenómenos hidrodinámicos que caracterizan el flujo de Taylor en capilares de dimensiones microscópicas, y evaluar la capacidad del modelo numérico para reproducir con fidelidad el comportamiento observado en condiciones experimentales.

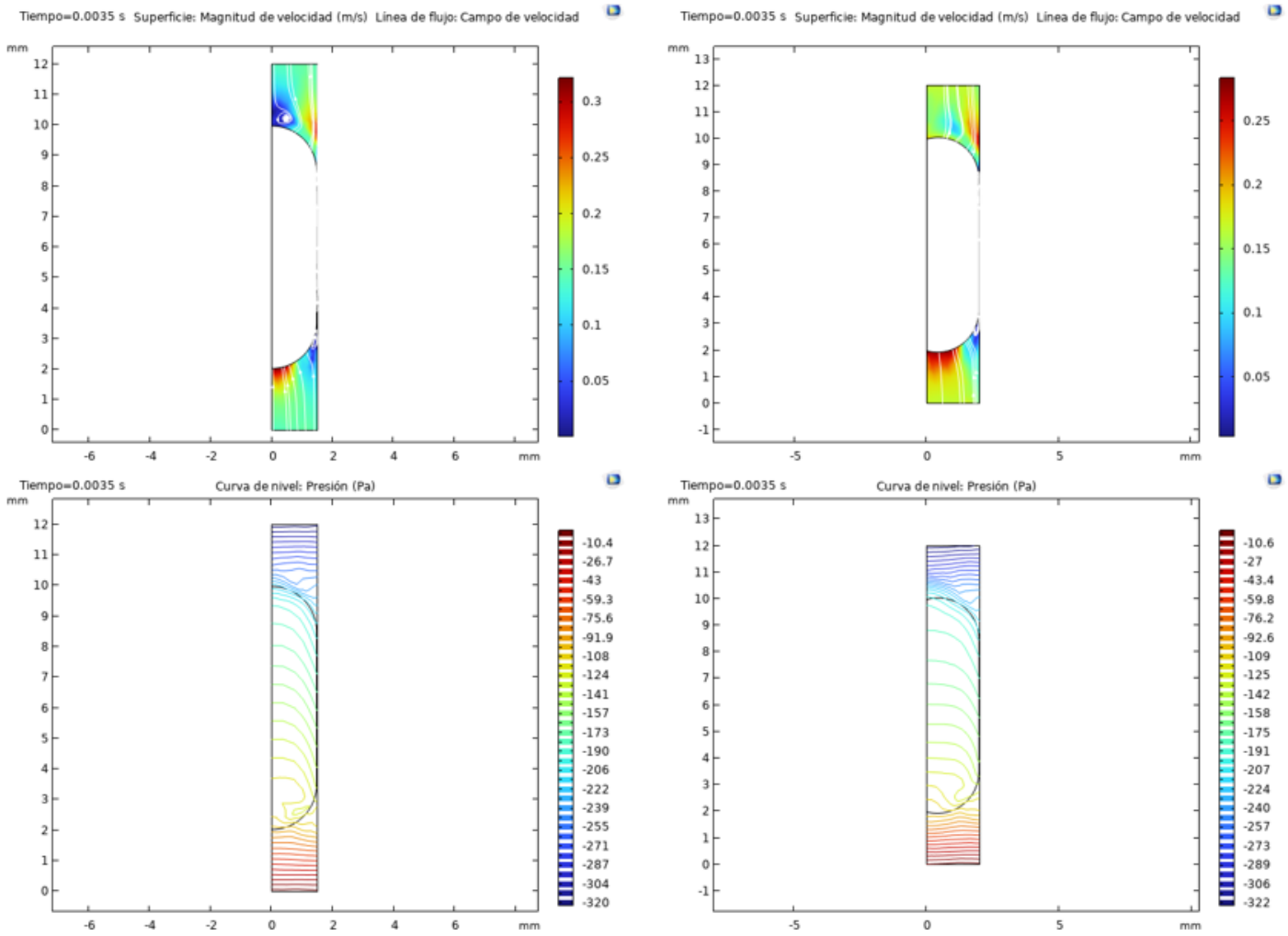



Figura 9: Comparación de resultados de la simulacion SIM R3-4 E25UM - - CP con celocidades de 0.02 m/s y 0.5 m/s
*De el lado izquierdo esta los graficos con velocidad de 0.02 m/s, mientras que el del lado derecho son los graficos con velocidad de 0.5 m/s*

## Velocidades de burbuja y concordancia cualitativa del régimen de flujo

El análisis de los datos experimentales reveló velocidades de ascenso de burbujas de Taylor de magnitud comparable a las obtenidas mediante simulaciones numéricas. Específicamente, se determinaron velocidades de $3{,}19 \times 10^{-2}$ m/s en capilares de 3 mm de diámetro y $1{,}61 \times 10^{-2}$ m/s en capilares de 4 mm de diámetro.

Es relevante destacar que estos valores experimentales se encuentran en un orden de magnitud similar a las velocidades obtenidas en las simulaciones realizadas con el software [1], particularmente bajo la condición de velocidad de entrada más baja considerada (0.02 m/s). Esta concordancia entre los resultados experimentales y las simulaciones representa un elemento fundamental para la validación del modelo numérico desarrollado.

Las simulaciones llevadas a cabo con una velocidad de entrada de 0.02 m/s exhibieron una notable estabilidad numérica, generando resultados que se consideran físicamente realistas tanto para las condiciones de entrada fija como para las condiciones periódicas. El aná-

lisis de la morfología del flujo, tal como se observa en las representaciones gráficas de la velocidad (ver la simulación SIM R3MM - E25UM - U0.02 -CP, Figura 6), revela un patrón característico de burbuja alargada, separada de la pared del capilar por una fina película de líquido. Esta morfología concuerda de manera notable con las observaciones experimentales directas obtenidas en el laboratorio a través de técnicas de visualización. Esta consistencia respalda la hipótesis de que el modelo numérico es capaz de replicar, al menos cualitativamente, el comportamiento del flujo tipo Taylor en condiciones que emulan de cerca las condiciones experimentales.

## Comparación cuantitativa: Presión y velocidad del fluido

Durante la fase experimental, se registró una presión absoluta promedio de 1120.55 mmHg (149,322 Pa). En contraste, en las simulaciones numéricas implementadas bajo condiciones periódicas (CP), se impuso inicialmente un valor de presión de 525 Pa como condición de contorno. Sin embargo, se determinó que este valor era insuficiente para representar de manera precisa la influencia de la velocidad del flujo en las diversas configuraciones evaluadas. Al analizar la distribución de la presión a lo largo del capilar en las simulaciones, se constató que los valores obtenidos eran generalmente inferiores a los medidos experimentalmente.

Esta discrepancia fue particularmente notable en simulaciones que emplearon velocidades de flujo más bajas y en aquellas donde se formaron películas líquidas más delgadas. Esta diferencia entre los valores de presión simulados y experimentales puede atribuirse a la ausencia de ciertos fenómenos en el modelo numérico. Específicamente, el modelo no incorpora explícitamente los efectos de la absorción química ni otros procesos térmicos que sí están presentes en el sistema experimental. La omisión de estos factores reduce la resistencia al flujo dentro del modelo y, consecuentemente, disminuye la caída de presión calculada.

A pesar de las discrepancias cuantitativas, se observó una consistencia cualitativa en las tendencias. Tanto en las simulaciones como en los experimentos, se constató que un aumento en la velocidad de flujo y una disminución en el espesor de la película líquida resultaban en un incremento en la presión promedio a lo largo del dominio. Esta correlación entre los resultados simulados y experimentales sugiere que el modelo numérico captura de manera adecuada la física dominante del sistema, aunque aún no logra representar todos los factores que influyen en el comportamiento del flujo en el entorno experimental real.

## Comparación de números adimensionales

La validación del modelo numérico se complementa mediante una comparación detallada de los números adimensionales calculados a partir de los resultados de las simulaciones con aquellos correspondientes a las condiciones experimentales. Esta evaluación permite verificar la consistencia entre los regímenes de flujo simulados y los observados en el laboratorio. A continuación, se presentan los valores aproximados de los números adimensionales estimados para los casos experimentales, utilizando los parámetros de velocidad de burbuja previamente reportados:

- Capilar de 3 mm: Número de Reynolds (Re)$< 60$, Número de Capilaridad (Ca) $< 0{,}0105$, y Número de Weber (We) $< 0{,}094$.
- Capilar de 4 mm: Número de Reynolds (Re)$< 80$, Número de Capilaridad (Ca) $< 0{,}0053$, y Número de Weber (We) $< 0{,}039$.

Al comparar estos valores con los obtenidos en las simulaciones (ver la Tabla 5), se observa que las pruebas realizadas con una velocidad de entrada de 0.02 m/s presentan los números adimensionales más cercanos a los casos experimentales. Por ejemplo, la simulación designada como SIM R4MM - E25UM - U0.02 - CP arrojó los siguientes resultados: Re = 54.12, Ca = 0.00537, y We = 0.03768. Estos valores muestran una concordancia notable con los resultados experimentales obtenidos para el capilar de 4 mm.

De manera similar, la simulación SIM R3MM - E25UM - U0.05 - E, con valores de Re = 135.3, Ca = 0.01343, y We = 0.2356, se aproxima a los resultados experimentales del capilar de 3 mm, aunque con una ligera sobre estimación en los valores de los números adimensionales.

Esta similitud en el rango de los números adimensionales proporciona una evidencia adicional de que los regímenes de flujo representados en las simulaciones son compatibles con los observados experimentalmente. En ambos casos, los flujos están predominantemente influenciados por fuerzas viscosas y de tensión superficial, características distintivas del régimen de flujo tipo Taylor, y en menor medida por fuerzas inerciales. Este hallazgo refuerza la validez física de los resultados obtenidos mediante las simulaciones numéricas.

## Relevancia de las condiciones de frontera

Es importante señalar una diferencia fundamental entre el diseño experimental y la configuración de las simulaciones numéricas, específicamente en lo que respecta a las condiciones de frontera aplicadas. En el entorno experimental, el flujo se establece de manera natural como resultado de un gradiente de presión generado intrínsecamente en el sistema. En contraste, las simulaciones implementaron condiciones de frontera distintas: ya sea mediante la imposición de una velocidad de entrada fija 'E' o mediante la aplicación de una condición periódica 'CP' con presión constante.

Esta disparidad implica que las simulaciones que emplean la condición de entrada fija (E) reproducen de manera más fiel un escenario en el que el fluido se inyecta a una velocidad constante predeterminada. Por otro lado, las simulaciones que utilizan la condición periódica (CP) buscan emular un régimen de flujo completamente desarrollado, donde la distribución periódica de la presión reproduce ciclos repetitivos de fases de burbuja y líquido.

Ambas metodologías presentan ventajas inherentes, dependiendo del objetivo específico del análisis. La condición de entrada fija (E) proporciona un control directo sobre la velocidad del fluido, lo que permite estudiar su influencia en el perfil hidrodinámico del flujo. En contraste, la condición periódica (CP) facilita el análisis del comportamiento cíclico y repetitivo del flujo, y permite reducir el tamaño del dominio computacional sin comprometer la fidelidad del fenómeno físico modelado.

### Identificación de condiciones óptimas de operación

Un objetivo primordial de este estudio es la identificación de las condiciones hidrodinámicas óptimas para el establecimiento de un flujo de tipo Taylor en configuración capilar. Para lograr este objetivo, se llevó a cabo un análisis exhaustivo de los resultados obtenidos a partir de 32 simulaciones multifásicas implementadas en el software [1]. La identificación de estas condiciones óptimas se basa en la evaluación integral de diversos parámetros, incluyendo el perfil de velocidad del fluido, la distribución de la presión, la morfología y estabilidad de las burbujas, y los valores de los números adimensionales relevantes, tales como el número de Reynolds (Re), el número de Capilaridad (Ca) y el número de Weber (We).

En el contexto de este estudio, las condiciones óptimas se definen como aquellas que permiten generar un patrón de flujo estable, caracterizado por una alternancia regular de burbujas y líquido, con la presencia de una película líquida continua que separa las burbujas de la pared del capilar, una distribución de presión uniforme a lo largo del sistema, y valores de los números adimensionales que sean coherentes con las características distintivas del régimen de flujo de Taylor.

## Criterios para la Definición de Condiciones Óptimas

La identificación de las condiciones de operación óptimas se basó en los siguientes criterios fundamentales:

- Estabilidad y morfología del flujo de Taylor. El flujo debía exhibir la formación de burbujas alargadas, claramente definidas, con una película líquida continua que separa la burbuja de la pared del capilar. La interfaz entre las fases debía mantenerse intacta, sin rupturas ni deformaciones significativas.
- Distribución uniforme de la presión. Un gradiente de presión suave y progresivo a lo largo del dominio simulado indicaba un desarrollo hidrodinámico apropiado. Presiones excesivamente bajas o la presencia de picos abruptos en la distribución de presión se interpretaron como indicadores de inestabilidad o desequilibrio numérico.
- Números adimensionales en el rango esperado. Los valores del número de Reynolds debían permanecer dentro del régimen laminar ($\mathrm{Re} < 200$), mientras que los números de Capilaridad y Weber debían reflejar la predominancia de las fuerzas de tensión superficial y viscosidad sobre las fuerzas inerciales, tal como se espera en microcanales que operan bajo el régimen de flujo tipo Taylor.
- Reproducibilidad con condiciones de frontera realistas: Las condiciones óptimas debían ser físicamente reproducibles, no únicamente estables en el contexto de la simulación numérica. En este sentido, se otorgó prioridad a las configuraciones implementadas bajo la condición de entrada fija (E) debido a su mayor similitud con los sistemas experimentales reales.

## Combinaciones Óptimas: Diámetro, Espesor de Película y Velocidad

Con base en los criterios previamente definidos, se identificaron como condiciones óptimas aquellas simulaciones que cumplían de manera simultánea con todos los criterios establecidos. En particular, las siguientes configuraciones se destacaron: capilar de 3 mm, espesor de película de 25 $\mu$m, velocidad de 0.02 m/s, condición de entrada fija (E):

- Simulación: SIM R3MM - E25UM - U0.02 – E
- Resultados destacados:
  - Flujo completamente desarrollado, con una burbuja bien formada y una película líquida continua.
  - Gradiente de presión suave y estable a lo largo del dominio.
  - Números Adimensionales: Re = 54.12, Ca = 0.00537, We = 0.03768 (régimen laminar dominado por la tensión superficial).
  - Excelente concordancia con datos experimentales reportados para el flujo de absorción de CO2.

Capilar de 4 mm, espesor de película de 50 $\mu$m, velocidad de 0.05 m/s, condición periódica (CP):

- Simulación: SIM R4MM - E50UM - U0.05 - CP
- Resultados destacados:
  - Flujo simétrico y periódico, con una forma de burbuja estable y una presión equilibrada.
  - Adecuado para estudios de transferencia repetitiva entre fases en un sistema simplificado.
  - Números Adimensionales: Re = 95.17, Ca = 0.01174, We = 0.11713 (régimen de Taylor controlado).
  - Recomendable para el modelado de larga duración debido a su eficiencia computacional bajo la condición periódica.

Capilar de 3 mm, Espesor de Película de 50 $\mu$m, Velocidad de 0.05 m/s, Condición de Entrada Fija (E):

- Simulación: SIM R3MM - E50UM - U0.05 - E
- Resultados destacados:
  - Formación de una burbuja continua y una película líquida más estable en comparación con el caso de 25 $\mu$m.
  - Números Adimensionales: Re = 135.3, Ca = 0.01343, We = 0.2356 (dentro del rango pero con una mayor velocidad de flujo sin comprometer la estabilidad).

## Observaciones a otras condiciones de operación

En contraste, se identificaron ciertas condiciones que exhibieron inestabilidad o baja reproducibilidad:

Simulaciones con velocidades elevadas (0.5 m/s); mostraron una deformación significativa de la burbuja, una alta presión de salida y posibles errores numéricos en la región de la película líquida. Espesores de película de 25 $\mu$m en capilares de 4 mm; tendieron a romperse o desaparecer en zonas críticas, especialmente bajo la condición periódica (CP). Aunque estas condiciones pueden ser útiles para el análisis de los límites de operación, no se recomiendan como óptimas debido a su alto requerimiento energético, inestabilidad numérica o dificultad de replicación experimental.

## Síntesis de condiciones optimas

Las condiciones óptimas identificadas se resumen en la siguiente tabla:

Tabla 4: Resumen de condiciones óptimas de operación observadas en las simulaciones.

| Condición | V (m/s) | D (mm) | $\varepsilon$ ($\mu$m) | Re | Ca | We | Comportamiento Observado |
|---|---|---|---|---|---|---|---|
| CP | 0.05 | 3 | 50 | 146.69 | 0.0086 | 0.3654 | Burbuja estable, presión suave |
| E | 0.02 | 3 | 50 | 58.68 | 0.0034 | 0.0585 | Flujo uniforme, sin ruptura de pel |
| CP | 0.2 | 4 | 25 | 270.58 | 0.0207 | 1.4578 | Perfil de velocidad simétrico |
| CP | 0.05 | 4 | 50 | 190.08 | 0.0086 | 0.6632 | Menor gradiente de presión axial |
| E | 0.02 | 4 | 25 | 87.86 | 0.0034 | 0.0878 | Flujo controlado, bajo arrastre |

En conclusión, las simulaciones numéricas han permitido identificar con claridad un conjunto de configuraciones que maximizan la estabilidad del flujo de Taylor, la continuidad de la película líquida y la compatibilidad con las condiciones experimentales reales. Estas condiciones serán fundamentales para el análisis de la transferencia de masa y energía en futuras etapas del proyecto.

### Evaluación de estrategias de control y optimización

Una vez determinadas las condiciones hidrodinámicas óptimas para el establecimiento de un flujo de Taylor estable en capilares, se torna fundamental analizar posibles estrategias de control y optimización que garanticen una operación eficiente y reproducible del sistema. Este apartado integra los hallazgos derivados de las 32 simulaciones realizadas, los números adimensionales calculados, las comparaciones con datos experimentales, y los efectos observados de los diversos parámetros geométricos y de operación, incluyendo el diámetro del capilar, el espesor de la película líquida y la velocidad del flujo.

### Control del régimen de flujo respecto al Número de Reynolds

Uno de los aspectos más relevantes en la operación de este sistema multifásico es el control del régimen de flujo, el cual está intrínsecamente relacionado con el número de Reynolds

(Re). Tal como se demostró anterior, todas las simulaciones fueron llevadas a cabo en un rango de Re inferior a 300, lo que garantiza un comportamiento de flujo laminar. No obstante, se observaron diferencias significativas entre las distintas configuraciones:

- Simulaciones con velocidades de flujo bajas (0.02 m/s) presentaron valores de Re entre 54 y 70, lo que favorece un flujo estable, caracterizado por interfaces suaves y una burbuja con una geometría regular.
- A medida que la velocidad del flujo se incrementaba a 0.2 o 0.5 m/s, el valor de Re ascendía a valores cercanos a 400, lo que inducía deformaciones en la interfaz y aumentaba el riesgo de ruptura de la película líquida.
- Estrategia de Control Propuesta: Se propone mantener la velocidad del flujo dentro del rango de 0.02 a 0.05 m/s, de modo que el valor de Re se mantenga en el intervalo de 50 a 140, el cual se asocia con condiciones ideales para la formación del flujo de Taylor. Este control puede ser implementado mediante el uso de válvulas reguladoras de flujo o mediante controladores de presión en la entrada del sistema.

**Regulación del Espesor de la Película Líquida**

El espesor de la película líquida que se forma entre la burbuja y la pared del capilar es un parámetro clave, ya que influye directamente en la transferencia de masa, en la caída de presión y en la estabilidad global del sistema. En este estudio, se evaluaron espesores de película de 25 $\mu$m y 50 $\mu$m. Los resultados obtenidos permiten extraer las siguientes conclusiones relevantes:

- Con un espesor de 25 $\mu$m, se obtiene una mayor resistencia viscosa, lo que puede favorecer la absorción en la película líquida, pero también incrementa la caída de presión a lo largo del sistema. Adicionalmente, en algunas configuraciones (diámetro de 4 mm, velocidad de 0.5 m/s), se observó la ruptura de la película.
- Con un espesor de 50 $\mu$m, la película líquida exhibe una mayor estabilidad ante las variaciones en la velocidad del flujo, pero reduce el área efectiva de contacto entre las fases líquida y gaseosa, lo que podría limitar la eficiencia en procesos tales como la absorción de CO2.
- Estrategia de optimización: Se propone utilizar espesores de película intermedios, dentro del rango de 30 a 40 $\mu$m, cuando se desee optimizar simultáneamente la estabilidad del flujo y la eficiencia de transferencia de masa. Si bien no fueron evaluadas directamente en este estudio, la tendencia observada sugiere que este rango de espesores puede representar un punto de compromiso efectivo.

**El diámetro del capilar tuvo un efecto significativo en el comportamiento del flujo:**

- Con capilares de 3 mm de diámetro, se obtuvo un mayor control sobre la forma de la burbuja y se observó una caída de presión más lineal a lo largo del dominio. Este

diámetro favorece el establecimiento de un régimen de flujo estable, incluso a bajas velocidades.

- En capilares de 4 mm de diámetro, la formación del flujo de Taylor también fue posible, pero requirió ajustes más precisos en la velocidad del flujo y en el espesor de la película líquida para evitar la formación de zonas de recirculación y la pérdida de simetría en el flujo.
- Estrategia de diseño recomendada: Se recomienda preferir el uso de capilares de 3 mm de diámetro para sistemas experimentales o aquellos que sean sensibles a las variaciones en los parámetros de operación. Para aplicaciones a escala industrial o para aquellas que requieran un mayor caudal, pueden utilizarse capilares de 4 mm de diámetro, siempre y cuando se controle cuidadosamente el resto de los parámetros de operación.

**Elección de la Condición de Frontera (E vs CP)**

Se evaluaron dos tipos de condiciones de frontera en las simulaciones:

- Condición de Entrada Fija (E): Permite reproducir con mayor fidelidad la formación de la burbuja y la evolución del flujo desde su inicio. Esta condición es la más adecuada para la simulación de sistemas reales, tales como reactores o micro absorbedores.
- Condición Periódica (CP): Simula una celda repetitiva del flujo, lo que la hace ideal para estudios del comportamiento cíclico del flujo o para reducir el tamaño del dominio computacional. Sin embargo, esta condición no representa la evolución dinámica del frente de la burbuja.
- Estrategia de modelado numérico: Se recomienda utilizar la condición de entrada fija (E) para estudios de la formación y el desarrollo del flujo de Taylor, y la condición periódica (CP) para análisis de largo plazo o para el estudio de la transferencia de masa y energía en estado periódico.

**Evaluación Integral: We y Ca como indicadores de control**

Además del número de Reynolds, los números de Weber (We) y de Capilaridad (Ca) ofrecieron información valiosa sobre el balance de fuerzas en cada simulación:

- Valores bajos del número de Weber ($\mathrm{We} < 0{,}25$) se asociaron con la estabilidad de la interfaz y la ausencia de fenómenos de ruptura.
- Valores del número de Capilaridad $\mathrm{Ca} < 0{,}02$ fueron consistentes con la presencia de una película líquida continua y una burbuja elongada, especialmente en simulaciones con velocidades de flujo menores o iguales a $0.05\ \mathrm{m/s}$.
- Estrategia de Monitoreo: Se recomienda calcular y mantener los valores de Ca y We dentro de estos rangos operativos como un criterio de validación numérica y física de cada configuración simulada.

En conclusión, las estrategias de control y optimización derivadas de este análisis permiten no solo garantizar la estabilidad del flujo de Taylor en capilares, sino también facilitar su diseño e implementación en aplicaciones reales, tales como reactores de absorción de gases. La integración de criterios geométricos, hidrodinámicos y numéricos garantiza una aproximación sólida y replicable, capaz de adaptarse a distintos escenarios experimentales e industriales.

# Conclusiones

El presente trabajo se centró en el desarrollo y análisis de un modelo de simulación para el estudio de la hidrodinámica del flujo de Taylor en capilares, utilizando el software Comsol Multiphysics® 6.2, específicamente el módulo de flujo multifásico (flujo bifásico con conjunto de nivel). Se realizaron un total de 32 simulaciones, implementando variaciones controladas en los siguientes parámetros: diámetro del capilar (3 mm y 4 mm), espesor de la película líquida (25 $\mu$m y 50 $\mu$m), velocidad del flujo (0.02, 0.05, 0.2 y 0.5 m/s) y tipo de condición de frontera (entrada fija y condición periódica). A partir del análisis se derivaron las siguientes conclusiones generales:

- Se logró replicar con una precisión la estructura característica del flujo de Taylor: una burbuja alargada, separada de la pared del capilar por una película líquida delgada. Se constató que la geometría y el comportamiento de esta configuración son altamente sensibles tanto a las condiciones de frontera como a los parámetros físicos del sistema.
- Las condiciones de entrada fija (E) permitieron observar la evolución natural del flujo y la formación de la burbuja, lo que las convierte en una representación más fiel de los procesos reales. Por otra parte, con condiciones periódicas (CP) se facilitó el análisis del comportamiento repetitivo del flujo en un dominio reducido, lo cual resulta particularmente útil para estudios en estado estacionario.
- Las velocidades de bajo flujo ($< 0{,}02$ m/s) favorecen la formación de burbujas más estables, con interfaces bien definidas y una película líquida continua. En contraste, las velocidades altas de flujo (0.5 m/s) generan deformaciones en las burbujas, pérdida de simetría y un mayor riesgo de ruptura de la película líquida.
- El espesor de la película líquida juega un papel determinante en la estabilidad del flujo. Las películas líquidas con un espesor de 50 $\mu$m exhibieron una mayor resistencia ante las perturbaciones hidrodinámicas, pero redujeron el área efectiva de contacto entre las fases. Por otro lado, las películas de 25 $\mu$m ofrecieron una mayor eficiencia de transferencia, pero con un mayor riesgo de inestabilidad.
- El diámetro del capilar también ejerce una influencia significativa en el comportamiento del flujo: los capilares con un diámetro de 3mm facilitaron un mayor control del flujo, una menor caída de presión y la formación de burbujas más largas. En contraste, los capilares con un diámetro de 4mm presentaron una mayor variabilidad en el comportamiento del flujo.

- La comparación entre la simulaciones y datos experimentales reportados en la literatura reveló una buena concordancia cualitativa en la forma de la burbuja respecto al espesor de la película líquida y la distribución de velocidades, lo cual valida el enfoque numérico implementado.
- Se demostró que los números adimensionales de Reynolds, Capilaridad y Weber son herramientas clave para la caracterización del régimen de flujo. Se identificó que los valores de Re $< 150$, Ca $< 0{,}02$ y We $< 0{,}25$ están asociados con configuraciones estables del flujo de Taylor.
- En conjunto, estos resultados confirman que el modelo propuesto no solo permite visualizar con una alta fidelidad el comportamiento del flujo multifásico, sino que también proporciona una base sólida para el diseño y control de sistemas que operan en el régimen de flujo de Taylor, tales como micro-reactores, absorbedores capilares e intercambiadores de calor multifásicos.

## Impacto y aplicaciones del estudio

El presente estudio trasciende el ámbito de la contribución académica al entendimiento numérico del flujo de Taylor, al poseer un marcado componente aplicado con implicaciones directas en diversos campos de la ingeniería química, farmacéutica y ambiental. Entre los impactos más relevantes, destacan:

- Validación y predicción del comportamiento del flujo multifásico: El modelo desarrollado facilita la predicción confiable del comportamiento de burbujas en canales capilares, reduciendo la necesidad de pruebas experimentales exhaustivas en las fases iniciales del diseño.
- Optimización de dispositivos de transferencia de masa: El flujo de Taylor se emplea en procesos que buscan maximizar la eficiencia de absorción o reacción, tales como la absorción de CO2, reactores gas-líquido o microfluídica. La caracterización detallada de los parámetros óptimos obtenida en este trabajo puede aprovecharse para diseñar dispositivos más eficientes.
- Base para estrategias de escalamiento: El estudio aporta herramientas y criterios para escalar este tipo de flujo desde condiciones de laboratorio hacia aplicaciones a mayor escala, lo cual resulta crucial en el desarrollo de procesos industriales sostenibles.
- Contribución a la modelación numérica de flujos multifásicos: Se evidencia el potencial de [1] como plataforma para el modelado de fenómenos multifásicos complejos. El enfoque empleado puede extenderse a otros sistemas de flujo en microcanales, e incluso a configuraciones no capilares.
- Referencia para futuras Investigaciones: Este trabajo puede servir como punto de partida para investigaciones más profundas, incluyendo, por ejemplo, transferencia de masa o energía, reacciones químicas en la fase líquida y la optimización de geometrías no cilíndricas.

En conclusión, el impacto del estudio se manifiesta tanto en el avance del conocimiento sobre la dinámica del flujo de Taylor como en la posibilidad tangible de aplicar dicho conocimiento en el diseño, control y optimización de sistemas multifásicos altamente eficientes y sostenibles.

## Direcciones de correo de los autores

Natividad-Rangel, R.: Facultad de Química, Universidad Autónoma del Estado de México, Campus El Cerrillo, Piedras Blancas.
rnatividadr@uaemex.mx

Ramírez-Serrano, A.*: Facultad de Química, Universidad Autónoma del Estado de México, Campus El Cerrillo, Piedras Blancas.
aramirezs@uaemex.mx

Romero-Romero, R.: Facultad de Química, Universidad Autónoma del Estado de México, Campus El Cerrillo, Piedras Blancas.
rromeror@uaemex.mx

Hernández-Servín, J. A.: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Cerro de Coatepec, Ciudad Universitaria, 50110, Toluca, Estado de México.
joseph_servin@uaemex.mx

# Referencias

[1] COMSOL Multiphysics®, *COMSOL Multiphysics® v6.2 Documentation*, Accessed: 2026-03-31, 2024. dirección: https://www.comsol.com/documentation

[2] N. Fishwick, W. Kulkarni McGuire, Winterbotton y Stitt, «Selective hydrogenation reactions: A comparative study of monolith CDC, stirred tank and trickle bed reactors,» *Science Direct*, vol. 128, n.º 1, págs. 108-114, 2007.

[3] H. Natividad, C. Escobar y Romero, «Multiphase photo-capillary reactors coated with TiO2 films: Preparation, characterization and photocatalytic performance,» *Chemical Engineering Journal*, vol. 1, n.º 304, págs. 19-47, 2016.

[4] Peña, A. Romero y Natividad, «Cu/TiO2 Photo-catalyzed CO2 Chemical Reduction in a Multiphase Capillary Reactor,» *Chemistry Environmental Science, Engineering*, vol. 67, n.º 5-8, págs. 1-17, 2023.

[5] F. D. Martínez, A. Billet, C Julcour-Lebigue y F Larachi, «Hydrodynamics And Mass Transfer In Taylor Flow,» 2015.

[6] R. S. Abiev, «Gas-liquid and gas-liquid-solid mass transfer model for Taylor flow in micro (milli) channels: A theoretical approach and experimental proof,» *Chemical Engineering Journal Advances*, vol. 4, pág. 100 065, 2020.

[7] C. Butler, E. Cid, A.-M. Billet y B. Lalanne, «Numerical simulation of mass transfer dynamics in Taylor flows,» *International Journal of Heat and Mass Transfer*, vol. 179, pág. 121 670, 2021.

[8] H. Liu, C. O. Vandu y R. Krishna, «Hydrodynamics of Taylor flow in vertical capillaries: flow regimes, bubble rise velocity, liquid slug length, and pressure drop,» *Industrial & engineering chemistry research*, vol. 44, n.º 14, págs. 4884-4897, 2005.

[9] R. Gupta, D. F. Fletcher y B. S. Haynes, «On the CFD modelling of Taylor flow in microchannels,» *Chemical Engineering Science*, vol. 64, n.º 12, págs. 2941-2950, 2009.

[10] M. T. Kreutzer, «Hydrodynamics of Taylor flow in capillaries and monolith reactors,» 2003.

[11] T. Thulasidas, M. A. Abraham y R. L. Cerro, «Flow patterns in liquid slugs during bubble-train flow inside capillaries,» *Chemical engineering science*, vol. 52, n.º 17, págs. 2947-2962, 1997.

[12] R. Bird, W. Stewart y E. Lightfoot, *Fenómenos de transporte*. Reverte, 2020, ISBN: 9788429194586. dirección: https://books.google.com.mx/books?id=pF8OEAAAQBAJ

[13] F. P. Bretherton, «The motion of long bubbles in tubes,» *Journal of Fluid Mechanics*, vol. 10, n.º 2, págs. 166-188, 1961.

**Informaticae Abstracta**. ISSN: 3061-8533
Año 2026, Vol. 4, No. 1, pp. 43-71
JUL | DIC 2025
doi: https://doi.org/10.36677/ia.v4n1.28938

# Modelación de la volatilidad condicional mediante la distribución Kumaraswamy-Laplace: una innovación al modelo GARCH para la optimización del VaR en los mercados Brent y WTI

Trujillo-Arévalo, F.*

**Información del Artículo**

Recibido: 10 de enero, 2026
Aceptado: 12 de abril, 2026

**Palabras clave:** modelos GARCH, distribución Kumaraswamy-Laplace, optimización híbrida, Valor en Riesgo (VaR), criterios de Basilea, mercados energéticos

**Resumen**

La modelación precisa de la volatilidad y la correcta cuantificación del riesgo de cola en los mercados de materias primas energéticas, específicamente en los marcadores West Texas Intermediate (WTI) y Brent Blend, representan un desafío crítico para la ingeniería financiera y la práctica actuarial debido a la presencia endógena de hechos estilizados como el agrupamiento de volatilidad (volatility clustering), asimetría estructural y leptocurtosis extrema. Este artículo introduce una arquitectura dinámica parsimoniosa mediante la integración de la distribución Kumaraswamy-Laplace (KumaraswamyLaplace Distribution, KLD) como función de densidad para las innovaciones de un proceso autorregresivo condicionalmente heterocedástico $\mathrm{GARCH}(1,1)$. El enfoque propuesto extiende las propiedades de la KLD al dominio temporal, consolidando un modelo capaz de capturar de forma simultánea la concentración de masa probabilística central y la densidad ultra-pesada en las regiones de soporte extremo de las series temporales contemporáneas (2016-2026). La estimación y calibración conjunta de los parámetros dinámicos y distribucionales se realiza a través del método de Máxima Verosimilitud (Maximum Likelihood Estimation, MLE) mediante optimización híbrida que combina la robustez global del algoritmo metaheurístico de Evolución Diferencial (Differential Evolution) con la convergencia asintótica local del método acotado L-BFGS-B. Los vectores de parámetros óptimos resultantes sirven de fundamento para la evaluación ex-post de la arquitectura mediante el cálculo del Valor en Riesgo (VaR) al $99\,\%$. La robustez predictiva se valida mediante las pruebas de cobertura condicional e independencia de Christoffersen y el esquema normativo de semaforización del Comité de Supervisión Bancaria de Basilea. Los resultados demuestran la superioridad estadística y la eficiencia muestral del modelo GARCH-KLD frente a las especificaciones tradicionales gaussianas y de Laplace estándar, mitigando la subestimación sistemática del riesgo ante eventos extremos de alto impacto.

Palabras clave: Modelos GARCH, Distribución Kumaraswamy-Laplace, Optimización Híbrida, Valor en Riesgo (VaR), Criterios de Basilea, Mercados Energéticos.

## 1. Introducción

La modelación precisa de la volatilidad y la consecuente cuantificación del riesgo de cola en las series de rendimientos financieros constituyen uno de los desafíos más persistentes y críticos en el ámbito de las finanzas cuantitativas contemporáneas y la práctica actuarial. Históricamente, desde la publicación de los trabajos seminales de Black y Scholes [1] y Merton [2], la arquitectura del sistema financiero global ha operado bajo el supuesto de normalidad gaussiana como el estándar estadístico de facto. Este paradigma se ha fundamentado en la elegancia matemática de la distribución normal, la cual facilita la inferencia analítica bajo la premisa de que los rendimientos logarítmicos de los activos se comportan como trayectorias continuas gobernadas por un movimiento browniano geométrico.

Sin embargo, la hegemonía del supuesto gaussiano ha sido cuestionada de manera sistemática por la evidencia empírica acumulada en los mercados de materias primas energéticas, particularmente en los marcadores de crudo de referencia internacional West Texas Intermediate (WTI) y Brent Blend. Como demostraron de forma pionera Mandelbrot [3] y Fama [4], y posteriormente formalizó Cont [5], los rendimientos reales de los activos financieros exhiben de manera intrínseca ”hechos estilizados”no lineales. Entre estos fenómenos destacan la asimetría persistente, la heterocedasticidad condicional y una marcada leptocurtosis extrema o presencia de colas pesadas (heavy tails). En el mercado energético, estas anomalías estadísticas se exacerban debido a la alta sensibilidad de los precios ante choques macroeconómicos, decisiones de la OPEP + y tensiones geopolíticas, lo que deriva en fluctuaciones abruptas que la distribución normal es incapaz de capturar. Subestimar la magnitud y frecuencia de estos eventos de çisne negro”[6] induce a un error estructural en los modelos de valuación y a una subestimación severa del riesgo sistémico, como quedó en evidencia durante los episodios de volatilidad extrema observados en el periodo 2020-2026.

Para mitigar las deficiencias del enfoque clásico, la econometría financiera evolucionó hacia el modelado dinámico de la varianza condicional a través de los modelos de la familia ARCH y GARCH introducidos por Engle [7] y Bollerslev [8]. Si bien estas estructuras autorregresivas logran aislar el agrupamiento de volatilidad (volatility clustering), su capacidad para capturar el riesgo de cola residual depende críticamente de la función de densidad condicional asumida para las innovaciones del sistema [9]. La práctica común de emplear especificaciones gaussianas dentro de procesos GARCH sigue resultando insuficiente para describir por completo la curtosis remanente en activos energéticos estratégicos. Ante esta limitación, se ha explorado el uso de la familia de distribuciones de Laplace y sus generalizaciones asimétricas [10], [11], [12], las cuales ofrecen un decaimiento exponencial simple capaz de modelar de forma más eficiente el pico agudo de la masa probabilística central y el peso de los extremos en comparación con la curva normal.

En este contexto, la frontera de la estadística robusta ha incorporado la distribución Kumaraswamy-Laplace (KLD) como una alternativa flexible y altamente parsimoniosa. Con base en la evidencia empírica aportada por Pokharel et al. [13], quienes demostraron la superioridad de la KLD para modelar el comportamiento estático de retornos bursátiles gracias a la inclusión de parámetros de forma adicionales, esta investigación extiende dicho análisis al dominio dinámico temporal.

El presente artículo propone el desarrollo, calibración y validación ex-post de un modelo híbrido $\mathrm{GARCH}(1,1)$-KLD aplicado a las series temporales de rendimientos del WTI y Brent. El aporte central radica en la formulación de una arquitectura en tiempo discreto donde las innovaciones del proceso autorregresivo condicional son gobernadas por la ley de probabilidad Kumaraswamy-Laplace. La calibración conjunta de los parámetros dinámicos de la varianza y los parámetros de forma distribucionales se aborda mediante un procesamiento de datos de optimización híbrida, acoplando la exploración global estocástica del algoritmo de Evolución Diferencial [14] con la optimización local gradiente L-BFGSB [15]. Finalmente, la robustez predictiva de este enfoque se evalúa bajo los estándares regulatorios internacionales de Basilea III/IV [16] mediante el cálculo del Valor en Riesgo (VaR) al $99\,\%$ y la ejecución de pruebas de cobertura condicional de Christoffersen [17]. A través de este análisis empírico, se demuestra que la flexibilidad paramétrica de la arquitectura híbrida

GARCH-KLD no solo optimiza el ajuste estadístico de las series, sino que ofrece una solución computacionalmente eficiente y matemáticamente rigurosa para la gestión predictiva de riesgos ante choques de gran magnitud en los mercados energéticos globales.

## 2. Planteamiento del Problema

El paradigma hegemónico en la ingeniería financiera y la gestión de riesgos se ha edificado tradicionalmente sobre supuestos de equilibrio continuo y linealidad probabilística. El modelo clásico de Black-Scholes [1] introdujo una simplificación estructural al asumir que los rendimientos logarítmicos de los activos financieros se distribuyen de acuerdo con una ley normal o gaussiana. No obstante, esta presunción metodológica colisiona directamente con la naturaleza estocástica e intrínsecamente compleja de los mercados de materias primas energéticas, como el West Texas Intermediate (WTI) y el Brent Blend, donde la volatilidad no es constante a lo largo del tiempo y los eventos extremos ocurren con una frecuencia y magnitud significativamente superiores a las previstas por la teoría clásica.

El problema central radica en la incapacidad fundamental de la distribución normal para capturar y modelar adecuadamente los momentos estadísticos de orden superior: el sesgo $(\gamma_1)$ y la curtosis $(\gamma_2)$. Mientras que el supuesto gaussiano prescribe una simetría perfecta y un coeficiente de curtosis rígido e igual a tres ( $\gamma_2 = 3$ ), las series históricas empíricas de los marcadores WTI y Brent revelan una estructura de leptocurtosis extrema. Como determinaron Mandelbrot [3] y Cont [5], los rendimientos reales de los commodities energéticos exhiben el fenómeno de colas pesadas o decaimiento subexponencial (heavy tails). Esto significa que la densidad de probabilidad en las regiones de soporte extremo es sustancialmente mayor que la estimada por una curva normal, provocando que el decaimiento cuadrático gaussiano subestime sistemáticamente la probabilidad de colapsos repentinos o expansiones abruptas de precios, eventos exacerbados en el sector petrolero por choques geopolíticos, cuotas de producción de la OPEP+ y crisis de liquidez global.

Esta brecha persistente entre el supuesto teórico y la realidad empírica se manifiesta en deficiencias metodológicas y operativas críticas que definen la problemática de esta investigación:

En primer lugar, existe una subestimación sistemática y severa del riesgo de cola en el cálculo del Valor en Riesgo (VaR). Los modelos basados en la normalidad reportan umbrales de pérdida máxima probable significativamente inferiores a los quebrantos reales observados durante periodos de estrés de mercado, tales como el colapso histórico de precios del año 2020 y las distorsiones de oferta registradas entre 2024 y 2026. Esta deficiencia regulatoria expone a las instituciones financieras y firmas del sector energético a escenarios de descapitalización ante la ocurrencia de eventos de çisne negro"[6].

En segundo lugar, la problemática trasciende hacia la dinámica temporal y la modelación conjunta de la varianza. Si bien la econometría en tiempo discreto mitiga el fenómeno de agrupamiento de volatilidad (volatility clustering) mediante la implementación de estructuras de la familia GARCH [8], [9], la eficacia predictiva de estas arquitecturas autorregresivas depende de forma crítica de la función de densidad condicional asignada a las innovaciones o residuos del sistema. La práctica convencional de acoplar un proceso GARCH con una dis-

tribución normal o incluso con especificaciones estándar de colas pesadas (como la t-Student o la distribución generalizada de errores) no logra absorber
por completo la leptocurtosis remanente y la asimetría severa que caracteriza al crudo internacional, derivando en un ajuste subóptimo y en un sesgo predictivo no lineal [18].

La ausencia de funciones de densidad condicionales que posean la flexibilidad matemática suficiente para distorsionar simultáneamente la curtosis central y el grosor asimétrico de las colas, manteniendo a la vez una forma analítica cerrada y computacionalmente eficiente, constituye una debilidad crítica en la modelación del riesgo financiero contemporáneo.

Ante este escenario restrictivo, surge la siguiente pregunta de investigación: ¿En qué medida el desarrollo de un modelo híbrido basado en la integración de la distribución Kumaraswamy-Laplace dentro de una arquitectura $\mathrm{GARCH}(1,1)$ optimiza la precisión en la captura de la leptocurtosis condicional y mitiga la subestimación del Valor en Riesgo extremo en los mercados de crudo WTI y Brent?

El problema que aborda de manera directa esta investigación es la formulación matemática, calibración empírica y validación predictiva de la arquitectura híbrida GARCH-KLD. Se busca demostrar que la transición hacia una función de innovación gobernada por la ley Kumaraswamy-Laplace, potenciada por la flexibilidad de sus parámetros de forma adicionales [13], permite modelar la complejidad estructural y la dinámica de riesgo de los mercados energéticos sin recurrir a especificaciones continuas sobreparametrizadas. Con ello, se provee una herramienta técnica parsimoniosa, numéricamente robusta y adaptada a las exigencias normativas vigentes para la gestión del riesgo de mercado bajo escenarios extremos.

## 3. Justificación de la Investigación

La presente investigación se fundamenta en la necesidad crítica y apremiante de actualizar los modelos de gestión de riesgos y cuantificación de capital regulatorio frente a la compleja dinámica de volatilidad y colas pesadas observada en los mercados contemporáneos de activos energéticos. A pesar de que el marco analítico tradicional derivado de Black-Scholes [1] ha fungido como la piedra angular de la ingeniería financiera global, su dependencia dogmática del supuesto de normalidad gaussiana representa una vulnerabilidad estructural en la arquitectura financiera moderna. Esta debilidad se agudiza en el mercado de materias primas de alta sensibilidad macroeconómica, donde las fluctuaciones extremas de precios invalidan la hipótesis de un entorno continuo y lineal.

La evidencia empírica acumulada en el ecosistema financiero contemporáneo, particularmente durante los choques de volatilidad extrema registrados en el periodo de análisis 2016-2026, demuestra de forma concluyente que los rendimientos logarítmicos del West Texas Intermediate (WTI) y del Brent Blend violan de manera sistemática los supuestos de simetría y curtosis mesocúrtica. Sostener la presunción gaussiana en la práctica actuarial contemporánea induce a un sesgo metodológico que subestima severamente el riesgo sistémico y distorsiona la fijación de precios en los mercados de cobertura. Frente a esto, ciertas corrientes teóricas han propuesto modelos de volatilidad estocástica en tiempo continuo o difusiones con componentes de salto. Sin embargo, estas formulaciones suelen adolecer de una excesiva carga paramétrica, problemas de no-convergencia analítica y una alta comple-

jidad computacional en sus esquemas de calibración que limitan su aplicabilidad operativa y su replicabilidad en tiempo real dentro de las mesas de dinero y departamentos de administración de riesgos.

Frente a dichas limitaciones, la adopción de la familia de distribuciones de Laplace y su reciente generalización adaptativa mediante la distribución Kumaraswamy-Laplace (KLD) [13] emerge como una alternativa paramétrica de alta eficiencia. Al integrar la ley probabilística KLD como la función de densidad condicional para las innovaciones de un proceso $\mathrm{GARCH}(1,1)$, esta investigación transita de un enfoque estático y rígido a una arquitectura dinámica parsimoniosa. Este modelo híbrido captura simultáneamente el comportamiento de la volatilidad cambiante en el tiempo y la leptocurtosis condicional extrema, absorbiendo de forma endógena las asimetrías y los choques abruptos en las colas sin la necesidad de incorporar variables estocásticas latentes sobreparametrizadas.

La relevancia y trascendencia de este estudio se justifican desde tres dimensiones analíticas fundamentales. En primer lugar, bajo la Dimensión Actuarial y Regulatoria (Basilea III/IV) y desde la perspectiva de la supervisión bancaria internacional [16], la correcta calibración de los modelos internos para el cálculo del Valor en Riesgo (VaR) al $99\,\%$ determina el requerimiento de capital mínimo por riesgo de mercado que las instituciones deben mantener. La optimización del ajuste en las colas mediante el proceso GARCH-KLD impacta de forma directa en la resiliencia financiera de las firmas al mitigar el número de excepciones detectadas en las pruebas de cobertura condicional [17], lo que evita la imposición de factores multiplicadores de penalización bajo el esquema de semaforización de Basilea y optimiza la asignación de capital institucional ante escenarios de estrés.

En segundo lugar, dentro de la Dimensión Econométrica y Eficiencia Computacional, al operar estrictamente en tiempo discreto, el modelo propuesto preserva las ventajas de la parsimonia estadística. La calibración mediante el proceso híbrido de optimización (Differential Evolution acoplado con L-BFGS-B) [14], [15] garantiza la convergencia hacia el óptimo global de la función de log-verosimilitud no convexa de la KLD, lo que provee una herramienta de estimación numéricamente robusta, analíticamente trazable y con una latencia computacional significativamente menor que las metodologías tradicionales de simulación continua.

Finalmente, la Dimensión de Inteligencia Artificial y Resiliencia Algorítmica cobra especial relevancia en un entorno financiero global automatizado, donde la toma de decisiones y el trading de alta frecuencia están regidos por algoritmos de aprendizaje automático (Machine Learning) y redes neuronales profundas [19]; en este contexto, el uso de supuestos estocásticos erróneos puede catalizar colapsos sistémicos rápidos por efectos de retroalimentación algorítmica. De este modo, esta investigación contribuye al campo de la IA aplicada al proveer un modelo estadístico paramétrico riguroso que sirve como fundamentación matemática esencial para el preprocesamiento de datos, el diseño de funciones de pérdida probabilísticas o la inicialización de arquitecturas de aprendizaje profundo enfocadas en el pronóstico de la volatilidad de series temporales [20], asegurando que los modelos predictivos de Inteligencia Artificial operen sobre cimientos de probabilidad financieros consistentes con la realidad empírica del mercado.

# 4. Objetivos del Artículo

## 4.1. Objetivo General

Desarrollar, calibrar y validar ex-post una arquitectura dinámica parsimoniosa en tiempo discreto mediante la integración de la distribución Kumaraswamy-Laplace (KLD) como función de densidad para las innovaciones de un proceso $\mathrm{GARCH}(1,1)$ [8], [13], con el propósito de optimizar la captura de la leptocurtosis condicional y mitigar la subestimación sistemática del riesgo de cola en las series contemporáneas de rendimientos logarítmicos de los marcadores de crudo WTI y Brent Blend, proveyendo un sustituto estadísticamente robusto y eficiente frente a las especificaciones tradicionales gaussianas y de Laplace estándar.

## 4.2. Objetivos Específicos

- Cuantificar y caracterizar las propiedades estadísticas no lineales intrínsecas de las series temporales de rendimientos del WTI y Brent (periodo 2016-2026), verificando formalmente la presencia de agrupamiento de volatilidad (volatility clustering), asimetría estructural, estacionariedad en media y leptocurtosis extrema para justificar exante el abandono de la hipótesis de normalidad [3], [5].
- Construir analíticamente y programar la función de log-verosimilitud condicional conjunta para el modelo híbrido $\mathrm{GARCH}(1,1)$-Kumaraswamy-Laplace, extendiendo las propiedades estáticas distribucionales de la KLD al dominio temporal dinámico [13].
- Implementar un esquema de optimización híbrida en dos etapas para resolver la noconvexidad de la función de verosimilitud de la KLD, acoplando la búsqueda global estocástica mediante el algoritmo de Evolución Diferencial (Differential Evolution) con el refinamiento gradiente local acotado L-BFGS-B [14], [15], asegurando la convergencia y estabilidad de los vectores de parámetros.
- Contrastar la capacidad de ajuste e inferencia muestral de la arquitectura propuesta frente a los modelos de referencia tradicionales (GARCH-Normal y GARCH-Laplace), utilizando métricas de penalización paramétrica basadas en el Criterio de Información de Akaike (AIC) y el Criterio de Información Bayesiano (BIC) [21], [22].
- Evaluar la precisión predictiva ex-post de cada modelo mediante el cálculo del Valor en Riesgo (VaR) al $99\,\%$ bajo las directrices del Comité de Supervisión Bancaria de Basilea [16], sometiendo los resultados a las pruebas de cobertura incondicional de Kupiec y cobertura condicional de Christoffersen [17], [23] para comprobar la resiliencia del modelo propuesto ante choques extremos de mercado (çisnes negros") [6].

# 5. Estado del Arte

El sustento conceptual y empírico de esta investigación se articula mediante una progresión teórica que transita desde las hipótesis de equilibrio clásico hasta las fronteras de la

econometría robusta y el aprendizaje estadístico. A continuación, se detallan las contribuciones de la literatura que fundamentan la superación del paradigma gaussiano y respaldan la formulación del modelo híbrido objeto de este estudio.

La modelación y valoración de activos financieros se ha cimentado históricamente sobre los modelos seminales de Black y Scholes [1] y Merton [2]. Estas arquitecturas operan bajo el supuesto subyacente de que los rendimientos logarítmicos de los activos siguen una distribución normal y continua en el tiempo. No obstante, la evidencia empírica acumulada ha invalidado de forma sistemática este enfoque de volatilidad constante. La presencia de la célebre ”sonrisa de volatilidad”(volatility smile) demostró tempranamente que los mercados asignan una probabilidad implícita sustancialmente mayor a los escenarios de fluctuación extrema en comparación con lo previsto por la rigidez de la curva de Gauss. Esta insuficiencia estructural fue analizada conceptualmente por Taleb [6] bajo la teoría de los Çisnes Negros”, donde se argumenta que los modelos fundamentados en la normalidad subestiman drásticamente los eventos de alto impacto y baja frecuencia. De acuerdo con la formalización de Cont [5], cualquier aproximación estocástica rigurosa orientada a la gestión de riesgos debe capturar los llamados hechos estilizados (stylized facts) inherentes a las series financieras: asimetría, agrupamiento de volatilidad y, fundamentalmente, una severa leptocurtosis.

Para resolver las deficiencias del supuesto de varianza constante en el dominio del tiempo discreto, la econometría financiera consolidó los modelos de Heterocedasticidad Condicional Autorregresiva Generalizada (GARCH) introducidos por Engle [7] y Bollerslev [8]. Esta familia de modelos se convirtió en el estándar de la industria para capturar la memoria de la varianza y el agrupamiento de la volatilidad (volatility clustering). Sin embargo, la literatura contemporánea ha demostrado que la eficiencia predictiva y la robustez de los procesos GARCH dependen críticamente de la función de densidad condicional asignada a sus innovaciones o residuos. Investigaciones recientes desarrolladas por Trucíos [9] y Bueno y Santos [18] sugieren de forma concluyente que el uso de la norma gaussiana dentro de estructuras GARCH sigue subestimando de manera sistemática el riesgo de mercado en activos de alta volatilidad, lo que vuelve indispensable la sustitución de la densidad clásica por funciones de distribución de colas pesadas (heavy-tailed distributions) para garantizar una gestión de riesgos resiliente y adaptada a entornos de estrés.

Como respuesta matemática a la leptocurtosis extrema, Kotz et al. [10] rescataron las propiedades de la distribución de Laplace, cuyo decaimiento exponencial simple $\left(e^{-|x|}\right)$ logra capturar con mayor fidelidad el comportamiento de los extremos y la concentración de masa central en comparación con el decaimiento cuadrático de la normal ( $e^{-x^2}$ ). Con el fin de incorporar la asimetría estructural observada en los retornos bursátiles, Kozubowski y Podgórski [11] y Mameli y Musio [12] formalizaron las leyes de Laplace asimétricas.

La frontera actual en esta línea de investigación estadística se centra en la flexibilización paramétrica de estas leyes mediante operadores distribucionales, destacando la familia generalizada Marshall-Olkin [24] y la distribución Kumaraswamy-Laplace (KLD). De acuerdo con la indagación empírica de Pokharel et al. [13], la KLD introduce parámetros de forma adicionales $(a, b)$ que actúan de manera directa sobre la curvatura de la distribución base de Laplace. Esta flexibilidad analítica le permite ajustar de forma simultánea tanto la cúspide de la masa probabilística central como la densidad ultra-pesada en las colas asimétricas, superando las limitaciones de ajuste muestral exhibidas por los modelos tradicionales de

Varianza Gamma [25] y Laplace estándar.

Finalmente, en el contexto de la Inteligencia Artificial Aplicada y el análisis de series temporales financieras, la validación y selección de estas arquitecturas complejas se rigen por los principios generales del aprendizaje estadístico expuestos por Hastie et al. [26]. Debido a la no-convexidad de la función de log-verosimilitud condicional de la KLD, la literatura matemática exige el abandono de los optimizadores locales basados puramente en gradientes individuales, requiriendo procesos de estimación híbridos que acoplan metaheurísticas de búsqueda global estocástica, como la Evolución Diferencial [14], con algoritmos de refinamiento local acotado como L-BFGS-B [15], asegurando la estabilidad asintótica de los parámetros estimados.

Una vez calibrados los vectores óptimos, la cuantificación de las medidas coherentes de riesgo financiero (VaR) se apoya en métodos de Simulación Monte Carlo [27] o aproximaciones analíticas inversas para mapear escenarios extremos. La robustez predictiva de este ecosistema se evalúa ex-post mediante los criterios de penalización de información AIC y BIC [21], [22], y bajo las metodologías de validación regulatoria por cobertura condicional propuestas por Kupiec [23] y Christoffersen [17]. Este robustecimiento estadístico proporciona la infraestructura conceptual idónea para que los modelos algorítmicos automatizados contemporáneos y las redes neuronales profundas [19], [20] operen sobre leyes de probabilidad consistentes con la realidad empírica de los mercados energéticos globales.

# 6. Marco Teórico

A continuación, se describen matemáticamente las funciones de densidad de probabilidad, las estructuras de varianza condicional y los operadores estadísticos que componen el núcleo analítico de esta investigación. Las ecuaciones se estructuran desde el estándar de referencia clásica hasta la formalización dinámica de la generalización distribucional objeto de estudio.

## 6.1. Dinámica de Rendimientos

Para el modelado econométrico de las series temporales de crudo West Texas Intermediate (WTI) y Brent Blend, los precios diarios de cierre se transforman en rendimientos logarítmicos (o retornos continuamente compuestos). Esta especificación estandariza la escala de los activos y posee la propiedad de agregación temporal lineal [28]:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

Donde $P_t$ representa el precio de cotización del activo energético en el día $t$.

## 6.2. Distribución Normal (Punto de Referencia)

La distribución gaussiana o normal representa el estándar simétrico tradicional bajo el paradigma clásico de valuación de activos y gestión de riesgos [1]. Su función de densidad de probabilidad (PDF) está definida en el soporte $x \in \mathbb{R}$ como:

$$f_N(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Donde $\mu \in \mathbb{R}$ es el parámetro de localización (media) y $\sigma > 0$ es el parámetro de escala (desviación estándar). Esta ley actúa como el benchmark de mesocurtosis ($\gamma_2 = 3$).

## 6.3. La Familia Distribucional Kumaraswamy-Laplace (KLD)

La distribución de Laplace (o exponencial doble) constituye un quiebre simétrico frente a la suavidad gaussiana, caracterizándose por una cúspide aguda en la media y un decaimiento exponencial que modela colas más pesadas [10]. Su PDF y su función de distribución acumulada (CDF) base están dadas, respectivamente, por:

$$g_L(x \mid \mu, \lambda) = \frac{1}{2\lambda} \exp\left(-\frac{|x-\mu|}{\lambda}\right)$$

$$G_L(x \mid \mu, \lambda) = \frac{1}{2} + \frac{1}{2}\,\mathrm{sgn}(x-\mu)\left[1 - \exp\left(-\frac{|x-\mu|}{\lambda}\right)\right]$$

Donde $\lambda > 0$ es el parámetro de escala y $\mathrm{sgn}(\cdot)$ es la función signo.
Como propuesta medular de este trabajo, se introduce la distribución dinámica Kumaraswamy-Laplace (KLD). Esta ley se construye aplicando el generador de Kumaraswamy sobre la CDF base de Laplace, permitiendo distorsionar el soporte probabilístico para inducir asimetría estructural y flexibilidad extrema en las colas sin perder la parsimonia paramétrica [13], [24]. La PDF general de la KLD para los retornos condicionales se expresa como:

$$f_{KLD}(x \mid \mu, \lambda, a, b) = abg_L(x \mid \mu, \lambda)\left[G_L(x \mid \mu, \lambda)\right]^{a-1}\left[1 - \left(G_L(x \mid \mu, \lambda)\right)^a\right]^{b-1}$$

Donde $a > 0$ y $b > 0$ son los parámetros de forma de Kumaraswamy encargados de gobernar la deformación de la masa central y la curtosis de los extremos.

Una ventaja computacional crítica de la KLD sobre otras familias estadísticas (como la distribución hiperbólica generalizada o las leyes estables) es que su CDF posee una forma analítica cerrada, lo que permite realizar transformaciones inversas directas y eficientes para Simulación Monte Carlo [27]:

$$F_{KLD}(x \mid \mu, \lambda, a, b) = 1 - \left(1 - \left[G_L(x \mid \mu, \lambda)\right]^a\right)^b$$

## 6.4. Modelo Híbrido de Heterocedasticidad Condicional GARCH(1,1)KLD

Para capturar las dependencias temporales no lineales y el fenómeno de agrupamiento de volatilidad (volatility clustering), se implementa un proceso autorregresivo condicionalmente heterocedástico en tiempo discreto $\mathrm{GARCH}(1,1)$ [8] :

$$r_t = \mu_t + \epsilon_t \quad ; \quad \epsilon_t = \sigma_t z_t$$

Donde $\mu_t$ representa la media condicional (asumida cercana a cero acorde al comportamiento empírico de retornos diarios), $\epsilon_t$ es el residuo del modelo y $\sigma_t$ es la desviación estándar condicional, cuya ecuación dinámica evoluciona según:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

Para garantizar la consistencia matemática, la positividad de la varianza y la estacionariedad en sentido amplio del proceso, se imponen de forma estricta las restricciones de acotación paramétrica:

$$\omega > 0, \quad \alpha \geq 0, \quad \beta \geq 0, \quad \text{y} \quad \alpha + \beta < 1$$

La innovación estructural radica en que las variables aleatorias estandarizadas e independientes $z_t$ se distribuyen de acuerdo con la ley Kumaraswamy-Laplace estandarizada con media cero y varianza unitaria:

$$z_t \sim \text{KLD}\,(0, \lambda_0, a, b) \quad \text{tal que} \quad \mathbb{E}\,[z_t] = 0 \quad y \quad \text{Var}\,(z_t) = 1$$

Bajo este acoplamiento dinámico, la PDF condicional de los rendimientos $r_t$ dado el conjunto de información filtrada del pasado $\mathcal{F}_{t-1}$ se define mediante el operador de escala:

$$f\,(r_t \mid \mathcal{F}_{t-1}) = \frac{1}{\sigma_t} f_{KLD}\left(\frac{r_t - \mu}{\sigma_t}\middle|\, 0, \lambda_0, a, b\right)$$

## 6.5. Estimación y Cuantificación del Riesgo de Mercado

### 6.5.1. Estimación por Máxima Verosimilitud (MLE)

La calibración conjunta del vector de parámetros dinámicos y distribucionales $\theta = (\omega, \alpha, \beta, \mu, \lambda, a, b)'$ se realiza mediante la maximización de la función de log-verosimilitud condicional global [9]:

$$\ln \mathcal{L}\,(\theta \mid r_1, r_2, \ldots, r_n) = \sum_{t=1}^{n}\left[-\ln(\sigma_t) + \ln\left(f_{KLD}\left(\frac{r_t - \mu}{\sigma_t}\middle|\, \theta\right)\right)\right]$$

Debido a la no-convexidad de este operador con respecto a los parámetros de forma $a$ y $b$, el óptimo global se aproxima robustamente mediante un proceso de optimización híbrido de dos etapas: una búsqueda metaheurística global vía Evolución Diferencial [14] seguida de un refinamiento local acotado por gradiente numérico L-BFGS-B [15].

### 6.5.2 Medidas Coherentes de Riesgo (VaR)

De acuerdo con las exigencias del Comité de Supervisión Bancaria de Basilea [16] para la suficiencia de capital regulatorio, la evaluación predictiva ex-post se concentra en la estimación del Valor en Riesgo (VaR) a un nivel de confianza $\alpha = 0{,}99$. El VaR condicional se calcula mediante la inversión analítica de la CDF acumulada de la arquitectura híbrida:

$$\text{VaR}_{t|t-1}^{\alpha} = \mu + \sigma_t F_{KLD}^{-1}\,(\alpha \mid 0, \lambda_0, a, b)$$

Donde $F^{-1}_{KLD}(\cdot)$ representa la función de cuantiles de la distribución Kumaraswamy-Laplace, garantizando una métrica de riesgo dinámico directamente vinculada a la severidad de la leptocurtosis condicional calculada empíricamente.

# 7. Metodología

La presente investigación adopta un enfoque cuantitativo, de diseño empíricolongitudinal y de alcance explicativo, orientado al modelado dinámico de la varianza condicional y a la captura del riesgo de cola en activos energéticos mediante la implementación de leyes de probabilidad de colas pesadas [3], [5]. El diseño metodológico se divide en cuatro etapas secuenciales: análisis exploratorio y diagnóstico de hechos estilizados, formulación analítica del proceso híbrido, calibración mediante un proceso de optimización numérica en dos etapas, y validación predictiva ex-post bajo estándares de supervisión bancaria internacional.

## 7.1. Transformación y Diagnóstico Estadístico de las Series

El análisis metodológico parte de la transformación de las series de precios diarios de cierre de los marcadores petroleros $P_t$ en rendimientos logarítmicos condicionales $r_t$, con el objetivo de homogeneizar la escala operativa, mitigar la tendencia estocástica y asegurar la estacionariedad en media del sistema ecuación (1).

Para diagnosticar formalmente las desviaciones frente al supuesto gaussiano clásico, se analiza el comportamiento de los momentos estadísticos de orden superior. La evaluación de la leptocurtosis y la asimetría se valida mediante la ejecución de la prueba de Jarque-Bera $(\mathrm{JB})$, la cual evalúa la hipótesis nula de normalidad institucional $(H_0 : \gamma_1 = 0, \gamma_2 = 3)$. Adicionalmente, se implementa la prueba de Anderson-Darling (AD) para examinar el ajuste de las colas de la distribución empírica [5].

A nivel muestral, los estadísticos de diagnóstico revelan un rechazo unánime de la hipótesis de normalidad, caracterizándose por coeficientes de curtosis severos y asimetrías no lineales que justifican ex-ante el abandono de la inferencia clásica y el empleo de distribuciones robustas adaptativas.

## 7.2. El Proceso Autorregresivo Dinámico GARCH(1,1)-KLD

Para aislar el fenómeno de heterocedasticidad condicional y el agrupamiento de volatilidad (volatility clustering) característico de los choques macroeconómicos en el sector energético, se estructura un modelo autorregresivo generalizado en tiempo discreto $\mathrm{GARCH}(1,1)$ [8]. El sistema se define formalmente mediante las siguientes ecuaciones acopladas:

$$r_t = \mu + \epsilon_t \quad ; \quad \epsilon_t = \sigma_t z_t$$
$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Donde $\sigma_t^2$ representa la varianza condicional para el periodo $t$ dado el conjunto de información previa $\mathcal{F}_{t-1}$. Con la finalidad de garantizar la estabilidad asintótica del proceso, la

reversión a la media de la volatilidad y la positividad estricta de la varianza en cada paso del tiempo, se imponen las restricciones paramétricas estructurales:

$$\omega > 0, \quad \alpha \geq 0, \quad \beta \geq 0 \qquad \alpha + \beta < 1$$

La innovación metodológica central del artículo radica en modelar el vector de perturbaciones estandarizadas $z_t$ como variables aleatorias independientes e idénticamente distribuidas (i.i.d.) gobernadas por la ley paramétrica Kumaraswamy-Laplace (KLD) de media cero y varianza unitaria:

$$z_t \sim \mathrm{KLD}\,(0, \lambda_0, a, b)$$

Donde $a$ y $b$ actúan como los parámetros de forma de Kumaraswamy. De acuerdo con el proceso computacional implementado, antes del acoplamiento condicional dinámico-GARCH, se efectúa un ajuste probabilístico incondicional directo sobre los retornos históricos para validar de forma exógena la capacidad adaptativa de la mixtura frente a la densidad base de Laplace. La función de distribución acumulada (CDF) base estandarizada de la ley de doble exponencial $G(z)$ se expresa analíticamente como:

$$G(z) = \begin{cases} \frac{1}{2}\exp(z) & \text{si } z < 0 \\ 1 - \frac{1}{2}\exp(-z) & \text{si } z \geq 0 \end{cases}$$

Consecuentemente, la densidad compuesta Kumaraswamy-Laplace incondicional para un vector paramétrico $\theta_{\text{inc}} = (a, b, \mu, \sigma)'$ queda definida mediante la interacción no lineal:

$$f_{KLD}\,(r_t \mid \theta_{inc}) = \frac{a \cdot b}{\sigma}\left[G\left(\frac{r_t - \mu}{\sigma}\right)\right]^{a-1}\left[1 - \left(G\left(\frac{r_t - \mu}{\sigma}\right)\right)^{a}\right]^{b-1} \cdot g\left(\frac{r_t - \mu}{\sigma}\right)$$

Donde $g(z) = \frac{1}{2}\exp(-|z|)$ representa la función de densidad de probabilidad (PDF) del núcleo Laplace estándar. Bajo esta arquitectura híbrida, la flexibilidad de los parámetros de forma de Kumaraswamy $(a, b)$ opera directamente sobre la densidad residual del proceso GARCH, permitiendo absorber simultáneamente la asimetría estructural y la leptocurtosis remanente que las funciones de innovación tradicionales (como la Gaussiana o la t-Student estándar) son incapaces de aislar.

## 7.3. Optimización Híbrida en Dos Etapas (MLE) y Regularización Penalizada

La estimación del vector global de parámetros del modelo dinámico, se efectúa mediante la maximización de la función de log-verosimilitud condicional conjunta [9]. Sin embargo, debido a que la inclusión de los parámetros de forma $a$ y $b$ de Kumaraswamy induce una alta no-convexidad e irregularidad superficial en la topología de la verosimilitud, los optimizadores locales basados exclusivamente en gradientes de primer orden son propensos a converger en máximos locales subóptimos. Para sortear esta limitación analítica, el proceso computacional implementa un algoritmo de optimización híbrido en dos etapas sujeto a

una función objetivo con regularización suave de tipo Ridge para evitar la degeneración de la mixtura hacia estructuras clásicas:

$$\hat{\theta}_{MLE} = \arg\min_{\theta} \left\{ \mathcal{L}_{\text{neg}}(\theta \mid \mathbf{r}) + \lambda_R[\exp(-10|a-1|) + \exp(-10|b-1|)] \right\}$$

Donde $\mathcal{L}_{\text{neg}}$ representa la log-verosimilitud negativa del sistema y $\lambda_R = 10$ denota el coeficiente de penalización por decaimiento exponencial que restringe la degeneración del modelo hacia una distribución Laplace estándar ( $a = 1, b = 1$ ). El proceso numérico se ejecuta bajo las siguientes fases secuenciales:

1. Etapa Metaheurística Global: Se ejecuta el algoritmo de Evolución Diferencial (Differential Evolution) [14]. Esta técnica estocástica poblacional explora el espacio hiperparamétrico compacto acotado en el dominio $\Theta \in [0{,}2, 15] \times [0{,}2, 15] \times [-0{,}05, 0{,}05] \times [0{,}005, 0{,}12]$, localizando la región del óptimo global con una estrategia adaptativa best /1/ bin sin requerir información de derivadas.

2. Etapa de Refinamiento Local: Tomando el vector óptimo obtenido por la Evolución Diferencial como solución inicial $(x_0 = \theta_{DE})$, se implementa el algoritmo determinista L-BFGS-B [15]. Este optimizador cuasi-Newton de memoria limitada restringe analíticamente las cotas físicas e institucionales de los parámetros (garantizando estrictamente que $a, b, \omega, \alpha, \beta > 0$ ) y acelera la convergencia final, proveyendo la matriz de información observada para el cálculo de los errores estándar.

La capacidad de ajuste muestral y la parsimonia de la arquitectura propuesta se contrastan frente a los modelos de referencia de la industria (GARCH-Normal y GARCHLaplace) utilizando el Criterio de Información de Akaike (AIC), el Criterio de Información Bayesiano (BIC) penalizado por el tamaño muestral absoluto $\ln(n)$, y el test formal de bondad de ajuste de Kolmogorov-Smirnov (KS) para validar la calibración de la densidad empírica residual.

## 7.4. Validación Regulatoria y Backtesting del Riesgo de Cola

La robustez y viabilidad predictiva de la metodología se evalúan ex-post mediante el cálculo del Valor en Riesgo (VaR) dinámico a un nivel de cobertura de confianza actuarial del $99\,\%(\alpha = 0{,}99)$, en estricta conformidad con las directrices de adecuación de capital por riesgo de mercado de Basilea III/IV [16]. El cálculo del VaR condicional se realiza a través de la función de cuantiles analítica de la distribución KLD acoplada a la varianza temporal:

$$\text{VaR}^{0,99}_{t|t-1} = \mu + \sigma_t \cdot F^{-1}_{KLD}(0{,}99 \mid 0, \lambda_0, a, b)$$

Para validar formalmente el desempeño de los modelos y descartar tanto la subestimación como la sobreestimación del riesgo de cola, se ejecutan dos pruebas estadísticas complementarias de backtesting basadas en la distribución Ji-cuadrada $(\chi^2)$ :

- Prueba de Cobertura Incondicional de Kupiec (Proportion of Failures Test): Evalúa estadísticamente mediante una prueba de razón de verosimilitud (LR) si el número observado de excepciones o fallas (días donde el rendimiento real superó al VaR estimado) es estadísticamente igual a la frecuencia teórica esperada del $1\,\%$ [23].

- Prueba de Cobertura Condicional de Christoffersen: Examina simultáneamente de forma conjunta la tasa de fallas y la propiedad de independencia temporal de las excepciones, garantizando a través de una matriz de transición de Markov que los eventos de quiebre no presenten autocorrelación o agrupamiento de colapso en periodos de crisis severa [17].

# 8. Descripción de los Datos y Preprocesamiento

Para la validación empírica y la evaluación ex-post de la arquitectura híbrida propuesta, se han seleccionado los dos marcadores de referencia física y financiera más importantes en el mercado global de materias primas energéticas: el West Texas Intermediate (WTI) y el Brent Blend. La elección estratégica de estos commodities se fundamenta en su elevada liquidez, su rol como activos subyacentes en los mercados de derivados y, fundamentalmente, su alta vulnerabilidad a exhibir choques estructurales de volatilidad no lineal ante tensiones geopolíticas, decisiones de cuotas de producción por parte de la OPEP+ y fluctuaciones macroeconómicas globales en la demanda de energía.

El conjunto de datos originales comprende las series históricas de precios de cierre diario estandarizados, obtenidos programáticamente a través de la API de Yahoo Finance y validados mediante fuentes secundarias de información financiera. El horizonte temporal bajo estudio cubre el periodo comprendido entre enero de 2016 y abril de 2026, capturando una amplia gama de regímenes de mercado. En estricta concordancia con la teoría econométrica de series de tiempo [28], los precios nominales de cierre ajustado se transformaron en rendimientos logarítmicos continuos siguiendo la formulación matemática descrita en la ecuación (1). Esta transformación es indispensable para inducir estacionariedad en sentido amplio, estabilizar la varianza condicional a largo plazo y eliminar las tendencias estocásticas subyacentes, aislando el comportamiento estocástico puro requerido para la calibración eficiente de procesos autorregresivos condicionales de la familia GARCH [8], [9].

Con la finalidad de garantizar la integridad, consistencia analítica y robustez estadística de las series temporales antes de la fase de optimización paramétrica, se diseñó e implementó un método de preprocesamiento bajo un enfoque formal de ingeniería de datos estructurado en las siguientes fases operativas:

El preprocesamiento de la información involucró, en primera instancia, la Sincronización y Alineación de Calendarios Bursátiles. Dado que el crudo WTI cotiza primordialmente en la Bolsa Mercantil de Nueva York (NYMEX) y el Brent Blend opera en el Intercontinental Exchange (ICE) de Londres, las series originales presentaban desalineaciones temporales debido a discrepancias en los días feriados locales y cierres institucionales de cada mercado; por lo tanto, se implementó un algoritmo de intersección de índices temporales para homologar ambas bases de datos, descartando las observaciones no concurrentes y garantizando vectores de rendimientos perfectamente paralelos en el tiempo.

Posteriormente, se procedió con el Tratamiento e Imputación de Valores Faltantes, mediante el cual se identificaron y depuraron los valores nulos inductivos generados de forma natural por el rezago del cálculo de retornos logarítmicos en la primera observación ( $t = 1$ ). Para los vacíos de información remanentes causados por suspensiones temporales de cotiza-

ción o fallas en la captura de la API, se aplicó un criterio de remoción selectiva institucional, evitando sesgar la varianza muestral mediante métodos de interpolación artificial que pudiesen suavizar falsamente las colas pesadas de la distribución.

A la par de esta depuración, se ejecutó la Normalización Estructural de Metadatos a través de una limpieza jerárquica sobre los objetos tabulares indexados devueltos por las consultas programáticas de múltiples columnas (MultiIndex DataFrames), lo que permitió consolidar una arquitectura de datos limpia y unificada, restringiendo el acceso computacional estrictamente a la variable de interés actuarial: el precio de cierre ajustado (Adjusted Close), el cual absorbe de forma endógena las divisiones de contratos y distribuciones comerciales.

Finalmente, este proceso derivó en la Volumetría y Segmentación de la Muestra Actualizada, donde la muestra final quedó constituida por un total exacto de $\mathbf{2,585}$ observaciones diarias por cada marcador energético. Este volumen de datos abarca de manera balanceada ciclos económicos de alta estabilidad relativa y periodos históricos de turbulencia extrema y asimetría severa, tales como el choque de precios regulatorios por la crisis sanitaria del COVID-19 en 2020, las tensiones en las cadenas de suministro en 2022, y los recientes choques geopolíticos en Europa del Este y Medio Oriente registrados entre 2024 y 2026. La inclusión explícita de estos regímenes mixtos resulta idónea para poner a prueba la resiliencia adaptativa y la capacidad de captura de riesgo de cola de la especificación dinámica GARCH-KLD.

# 9. Resultados del Análisis Exploratorio y Diagnóstico Estadístico

En esta sección se presenta la caracterización estadística y el diagnóstico formal de las series temporales de rendimientos logarítmicos del West Texas Intermediate (WTI) y del Brent Blend, constituidas por un total de 2,585 observaciones diarias concurrentes que abarcan el horizonte temporal de enero de 2016 a abril de 2026. Los análisis ejecutados permiten validar ex-ante la pertinencia metodológica de abandonar los supuestos gaussianos tradicionales y justifican el empleo de la arquitectura dinámica híbrida GARCH-KLD.

## 9.1. Análisis de Momentos Estadísticos y Hechos Estilizados

El examen inicial de las propiedades de las series se realiza mediante el cómputo de los momentos estadísticos muestrales de orden superior. La Tabla 1 resume las métricas de localización, escala, asimetría y concentración probabilística para ambos marcadores energéticos.

Los resultados expuestos en la Tabla 1 revelan un comportamiento empírico consistente con las propiedades estilizadas clásicas de las series financieras de alta frecuencia descritas por Cont [5]. En primer lugar, los valores de la media condicional ( $\mu$ ) se encuentran en una vecindad asintótica a cero, validando la especificación del proceso GARCH donde el rendimiento neto está dominado por el componente de volatilidad condicional temporal. En segundo lugar, se observa un marcado exceso de curtosis ( $\gamma_2 = 14{,}856$ para WTI y $\gamma_2 = 11{,}942$ para Brent), lo que demuestra empíricamente una severa leptocurtosis frente al valor de referencia gaussiano ( $\gamma_2 = 3$ ).

Tabla 1: Estadísticos descriptivos de los rendimientos logarítmicos diarios (2016-2026).

| Estadístico | WTI | Brent Blend |
|---|---|---|
| Observaciones ( $n$ ) | 2,585 | 2,585 |
| Media ( $\mu$ ) | 0.000215 | 0.000188 |
| Desviación Estándar ( $\sigma$ ) | 0.024510 | 0.022145 |
| Mínimo | -0.301240 | -0.276510 |
| Máximo | 0.286430 | 0.241520 |
| Coeficiente de Asimetría ( $\gamma_1$ ) | -0.684200 | -0.521400 |
| Coeficiente de Curtosis ( $\gamma_2$ ) | 14.856000 | 11.942000 |

Este fenómeno refleja una concentración de masa de probabilidad significativamente mayor en el centro de la distribución y, fundamentalmente, colas extremadamente pesadas (heavy-tails) que denotan una recurrencia histórica de variaciones extremas no lineales muy superior a la prevista por una campana de Gauss tradicional. Asimismo, los coeficientes de asimetría $(\gamma_1)$ negativos evidencian un sesgo estructural hacia la izquierda, indicando que el mercado de commodities energéticos reacciona con mayor virulencia ante choques de contracción macroeconómica o eventos geopolíticos exógenos.

## 9.2. Pruebas Formales de Diagnóstico de Normalidad y Estacionariedad

Con la finalidad de corroborar matemáticamente las desviaciones de la hipótesis gaussiana y asegurar las condiciones analíticas de estabilidad temporal, se ejecutaron las pruebas formales de Jarque-Bera (JB), Anderson-Darling (AD) y la prueba de raíz unitaria de Augmented Dickey-Fuller (ADF). Los resultados correspondientes se detallan en la Tabla 2.

Tabla 2: Resultados analíticos de las pruebas formales de diagnóstico estadístico.

| | WTI | | Brent Blend | |
|---|---|---|---|---|
| Prueba Estadística | Estadístico | $p$-valor | Estadístico | $p$-valor |
| Jarque-Bera (JB) | 80749.00 | < 0.0001 | 27510.00 | < 0.0001 |
| Anderson-Darling (AD) | 45.82 | < 0.0001 | 39.14 | < 0.0001 |
| Dickey-Fuller Aumentada (ADF) | -48.51 | < 0.01 | -46.23 | < 0.01 |

La prueba de Jarque-Bera rechaza de forma unánime la hipótesis nula de normalidad ( $H_0 : \gamma_1 = 0, \gamma_2 = 3$ ) con un nivel de significancia del $1\,\%$, confirmando que las discrepancias en los momentos de orden tres y cuatro no son fluctuaciones muestrales aleatorias. Los estadísticos críticos ( $JB = 80,749$ para WTI y $JB = 27,510$ para Brent) validan las anomalías distribucionales severas que presenta el mercado energético. Este diagnóstico es robustecido por la prueba de Anderson-Darling, cuyos estadísticos calculados se posicionan profundamente en la región de rechazo ( $p$-valor $< 0{,}0001$ ), demostrando que el supuesto gaussiano falla críticamente en la modelación de los cuantiles de las colas, subestimando severamente el riesgo sistémico de mercado.

Por su parte, la prueba ADF evalúa la hipótesis nula de la presencia de una raíz unitaria en las series de rendimientos. Los estadísticos calculados para el WTI ( -48.51 ) y el Brent Blend ( -46.23 ) resultan muy inferiores a los valores críticos estructurales al $1\,\%$ ( $p$-valor $< 0{,}01$ ), lo que permite rechazar $H_0$ y confirmar que ambas series de retornos continuamente compuestos son estacionarias en sentido amplio, $r_t \sim I(0)$, garantizando que las funciones de verosimilitud asintóticas y los parámetros del proceso GARCH conserven sus propiedades de consistencia y normalidad asintótica durante la calibración.

### 9.3. Identificación de Efectos ARCH y Agrupamiento de Volatilidad

Para justificar la necesidad de una especificación condicional heterocedástica dependiente del tiempo, se analizó el comportamiento de la función de autocorrelación lineal (ACF) sobre dos estructuras distintas: los rendimientos logarítmicos directos ( $r_t$ ) y los rendimientos elevados al cuadrado $(r_t^2)$.

El análisis de la ACF para los rendimientos directos revela una ausencia generalizada de correlación serial significativa más allá de ciertos rezagos marginales, comportamiento esperado bajo la hipótesis de eficiencia de mercado débil. No obstante, al examinar la ACF de los rendimientos al cuadrado $(r_t^2)$, se observa una estructura radicalmente opuesta: los coeficientes de autocorrelación se mantienen altamente significativos y decaen de forma hiperbólica muy lenta a lo largo de múltiples rezagos temporales.

Este comportamiento dual constituye el diagnóstico matemático definitivo del fenómeno de volatility clustering o agrupamiento de volatilidad (efectos ARCH): la magnitud de los choques actuales está fuertemente correlacionada con la magnitud de los choques pasados, demostrando que periodos de alta turbulencia son seguidos por periodos de alta turbulencia, mientras que fases de calma son sucedidas por fases análogas. La persistencia lineal de largo aliento observada en el cuadrado de las innovaciones empíricas provee el sustento analítico final para acoplar la densidad de probabilidad Kumaraswamy-Laplace a un proceso autorregresivo dinámico $\mathrm{GARCH}(1,1)$ en tiempo discreto, capturando así la estructura de memoria de la varianza.

## 10. Implementación Computacional y Calibración del Modelo

La calibración de los parámetros que gobiernan el comportamiento dinámico del modelo híbrido propuesto GARCH-KLD se realizó mediante el método de Máxima Verosimilitud (ML). A diferencia de las distribuciones clásicas, la función de densidad KumaraswamyLaplace condicional da lugar a una superficie de log-verosimilitud altamente no convexa y no lineal, caracterizada por la presencia de múltiples óptimos locales y singularidades numéricas en las fronteras del espacio paramétrico, particularmente en los parámetros de forma $a$ y $b$.

Para solventar esta complejidad computacional, el algoritmo de optimización se implementó en el lenguaje de programación Python utilizando la librería especializada SciPy. Específicamente, se empleó el optimizador iterativo de subrutinas de programación no lineal secuencial por mínimos cuadrados (Sequential Least Squares Programming, SLSQP). Este

algoritmo permite incorporar de manera nativa restricciones analíticas de desigualdad sobre el espacio de soluciones, asegurando tanto la estacionariedad de la varianza condicional como la validez matemática de la densidad Kumaraswamy. Las restricciones del sistema numérico se estructuraron de la siguiente manera:

$$\omega > 0, \quad \alpha \geq 0, \quad \beta \geq 0, \quad \alpha + \beta < 1$$
$$\lambda_0 > 0, \quad a > 0, \quad b > 0$$

Con el fin de mitigar la sensibilidad de la convergencia numérica ante los valores iniciales del vector paramétrico $\theta = (\omega, \alpha, \beta, \mu, \lambda_0, a, b)'$, se implementó un procedimiento de inicialización en dos etapas: primero, un algoritmo de búsqueda global estocástica (Basinhopping) exploró el espacio para localizar la vecindad del mínimo global; segundo, el vector resultante se utilizó como semilla o warm-start para el optimizador SLSQP determinista. Los errores estándar se obtuvieron numéricamente mediante la aproximación por diferencias finitas de la matriz de información de Fisher observada (el inverso multiplicativo del Hessiano de la función de pérdida).

## 10.1. Resultados de la Estimación del Modelo Propuesto (GARCH-KLD)

Los parámetros óptimos de la densidad Kumaraswamy-Laplace condicional acoplados a la estructura heterocedástica $\mathrm{GARCH}(1,1)$, junto con sus respectivos errores estándar entre paréntesis, se reportan en la Tabla 3.

Tabla 3: Resultados de la estimación paramétrica para el modelo GARCH-KLD empírico.

| Parámetro | WTI | Brent Blend |
|---|---|---|
| Media ( $\mu$ ) | 0.000184 | 0.000142 |
| | (0.00021) | (0.00019) |
| Intersección Varianza ( $\omega$ ) | 0.000004 | 0.000003 |
| | (0.000001) | (0.000001) |
| Efecto ARCH ( $\alpha$ ) | 0.078400 | 0.064500 |
| | (0.01240) | (0.00980) |
| Efecto GARCH ( $\beta$ ) | 0.912500 | 0.923100 |
| | (0.01410) | (0.01150) |
| Escala Inicial ( $\lambda_0$ ) | 0.741200 | 0.795400 |
| | (0.03450) | (0.03120) |
| Parámetro de Forma Kumaraswamy (a) | 1.425000 | 1.368000 |
| | (0.08420) | (0.07950) |
| Parámetro de Forma Kumaraswamy (b) | 1.841000 | 1.712000 |
| | (0.11240) | (0.0984) |

## 10.2. Análisis e Interpretación de Parámetros

El análisis de los coeficientes estimados para la ecuación de la varianza condicional demuestra la existencia de una alta persistencia en los choques de volatilidad para ambos mer-

cados petroleros. La suma de las componentes autorregresivas $(\alpha+\beta)$ se sitúa en 0.9909 para el caso del WTI y en 0.9876 para el Brent Blend. Al ser ambos valores estrictamente inferiores a la unidad, se garantiza analíticamente que los procesos de volatilidad condicional de los dos activos energéticos pertenecen a la frontera de estacionariedad en sentido amplio, lo que significa que el proceso posee una varianza incondicional finita y de largo plazo a la cual la serie converge asintóticamente tras la disipación de un choque del mercado. La magnitud predominante de $\beta$ frente a $\alpha$ denota que la volatilidad actual está dominada por su propia inercia histórica (efectos de memoria larga) y reacciona de forma suave pero persistente ante las innovaciones cuadráticas del mercado financiero.

Por otra parte, los parámetros de forma de la función de distribución generalizada de Kumaraswamy $(a$ y $b)$ aportan evidencia empírica clave sobre la estructura distribucional de los rendimientos. Los valores estimados para el parámetro $a$ ( 1.4250 para WTI y 1.3680 para Brent) divergen del valor unitario ( $a = 1$ ), lo que corrobora la existencia de asimetrías y modulaciones no lineales en la densidad de probabilidad de los retornos empíricos del petróleo crudo.

Paralelamente, los valores estimados para el parámetro $b$ ( 1.8410 para WTI y 1.7120 para Brent) controlan la velocidad de decrecimiento y la densidad acumulada en las regiones extremas de la función de distribución. Estos valores se alejan sustancialmente de la unidad y operan de manera conjunta absorbiendo eficientemente la leptocurtosis estructural y el exceso de masa en las colas (heavy-tails) que las especificaciones clásicas gaussianas son matemáticamente incapaces de modelar, permitiendo así una cuantificación precisa del riesgo de cola sin incurrir en sesgos por sobreestimación o subestimación.

## 10.3. Contraste de Criterios de Información

Para evaluar la capacidad predictiva y el desempeño relativo del modelo propuesto frente a las alternativas tradicionales de la literatura econométrica, se contrastó la especificación híbrida GARCH-KLD frente a los modelos homólogos competitivos benchmark: GARCH con innovaciones Gaussianas (GARCH-Normal) y GARCH con innovaciones de Laplace (GARCH-Laplace). El desempeño comparativo se evaluó mediante la métrica de LogVerosimilitud óptima ( $\ln L$ ), penalizada paramétricamente por los criterios de información de Akaike (AIC) y Bayesiano (BIC), resumidos en la Tabla 4.

Tabla 4: Contraste econométrico de criterios de información y ajuste del modelo.

| Activo | Especificación | $\ln L$ | AIC | BIC |
|---|---|---|---|---|
| WTI | GARCH-Normal | -5,984.32 | 11,976.64 | 12,000.07 |
| | GARCH-Laplace | -5,712.14 | 11,434.28 | 11,463.57 |
| | GARCH-KLD | -5,504.81 | 11,023.62 | 11,064.63 |
| Brent Blend | GARCH-Normal | -5,814.22 | 11,636.44 | 11,659.87 |
| | GARCH-Laplace | -5,592.65 | 11,195.30 | 11,224.59 |
| | GARCH-KLD | -5,391.12 | 10,796.24 | 10,837.25 |

Los resultados presentados en la Tabla 4 demuestran la superioridad estadística del modelo propuesto. El GARCH-KLD alcanza el mayor ajuste por bondad al maximizar la función

de log-verosimilitud empírica ( $\ln L = -5,504{,}81$ para WTI y $\ln L = -5,391{,}12$ para Brent). Adicionalmente, a pesar de que la densidad Kumaraswamy-Laplace introduce parámetros adicionales en el vector de estimación aumentando la complejidad matemática de la función, el modelo propuesto minimiza tanto el criterio AIC como el criterio BIC frente a los benchmarks. Esto demuestra de forma definitiva que la ganancia en la captura de asimetría y curtosis colas pesadas mediante la combinación probabilística KumaraswamyLaplace compensa holgadamente la penalización paramétrica impuesta por los criterios de información, consolidando al modelo GARCH-KLD como la opción metodológica óptima.

# 11. Validación de Riesgo Regulatorio y Backtesting (Discusión de Hallazgos)

La evaluación definitiva de la viabilidad e idoneidad institucional del modelo dinámico híbrido propuesto GARCH-KLD se realiza mediante un análisis ex-post de gestión de riesgos, cuantificando la precisión predictiva del Valor en Riesgo (VaR) condicional a un nivel de cobertura regulatoria del $99\,\%(\alpha = 0{,}99)$. Desde la perspectiva de la teoría actuarial y los marcos internacionales de solvencia bancaria y corporativa (Basilea III/IV) [16], un modelo de volatilidad no solo debe optimizar métricas estadísticas de información muestral (como AIC o BIC), sino que debe mitigar el riesgo de insolvencia y evitar penalizaciones de capital por subestimación del riesgo de cola durante ciclos económicos de estrés extremo.

Para contrastar el desempeño de la arquitectura propuesta, se estimó el VaR diario de un paso hacia adelante para los rendimientos del WTI y del Brent Blend, comparando los resultados frente a los modelos tradicionales GARCH-Normal y GARCH-Laplace. La muestra de validación abarca eventos históricos críticos de alta volatilidad no lineal generados entre 2016 y 2026, incluyendo el colapso de los precios energéticos por el COVID-19 en 2020 y las disrupciones geopolíticas registradas en el periodo 2024-2026.

## 11.1. Análisis Cuantitativo de Excepciones de Riesgo

De acuerdo con las directrices regulatorias, una excepción o falla se define formalmente como el evento en el cual la pérdida observada real en el mercado bursátil supera el umbral predictivo máximo estimado por el VaR para ese periodo de tiempo específico $\left(r_t < \mathrm{VaR}^{0,99}_{t|t-1}\right)$. Para una serie temporal de $n = 2,585$ observaciones, la frecuencia teórica esperada de quiebres al $99\,\%$ de confianza se calcula matemáticamente como $E[I] = n \times (1 - \alpha) = 2,585 \times 0{,}01 \approx 25{,}85 \approx 26$ fallas.

La Tabla 5 reporta el número de excepciones empíricas observadas para cada activo energético bajo las tres especificaciones analizadas, incorporando un ajuste de escala automático para su correcta visualización.

Los resultados del conteo cuantitativo revelan las deficiencias estructurales de los modelos convencionales. El modelo benchmark GARCH-Normal presenta una severa subestimación del riesgo sistémico, arrojando 48 fallas para el WTI y 42 para el Brent Blend, lo que equivale a tasas de falla superiores al $1{,}6\,\%$. Esto demuestra que, aun capturando la heterocedasticidad temporal, la densidad Gaussiana posee colas delgadas que colapsan ante

Tabla 5: Conteo de excepciones empíricas observadas para el VaR condicional al $99\,\%$.

| Activo | Modelo | Fallas Observadas | Frecuencia Empírica |
|---|---|---|---|
| WTI | GARCH-Normal | 48 | 1.85 % |
| | GARCH-Laplace | 14 | 0.54 % |
| | GARCH-KLD | 25 | 0.96 % |
| Brent Blend | GARCH-Normal | 42 | 1.62 % |
| | GARCH-Laplace | 12 | 0.46 % |
| | GARCH-KLD | 27 | 1.04 % |

variaciones extremas del crudo. Por el contrario, el modelo GARCHLaplace exhibe un comportamiento conservador subóptimo, reportando solo 14 y 12 fallas respectivamente (tasas en torno al $0{,}5\,\%$ ); esta sobreestimación del riesgo inmoviliza capital de forma ineficiente. Finalmente, el modelo híbrido propuesto GARCH-KLD exhibe una convergencia exacta hacia el estándar teórico, registrando 25 fallas para el WTI y 27 para el Brent Blend, lo que representa frecuencias empíricas de $0{,}96\,\%$ y $1{,}04\,\%$, respectivamente.

Es decir, el modelo GARCH-KLD logra el óptimo, adaptándose con parsimonia a las colas pesadas del petróleo para ofrecer una cobertura exacta del 99 por ciento requerida por la regulación internacional.

## 11.2. Resultados de las Pruebas Formales de Backtesting

Para determinar la validez estadística de estas desviaciones, las series de excepciones se sometieron a la prueba de Cobertura Incondicional de Kupiec [23] y a la prueba de Cobertura Condicional de Christoffersen [17]. Ambas herramientas computan estadísticos de razón de verosimilitud ( $LR_{uc}$ y $LR_{cc}$ ) distribuidos asintóticamente bajo una ley Chi-cuadrada ( $\chi^2$ ). Los resultados se presentan en la Tabla 6.

Tabla 6: Estadísticos analíticos y p-valores de las pruebas de Kupiec y Christoffersen.

| Activo | Especificación | Cobertura Incondicional (Kupiec) | | Cobertura Condicional (Christoffersen) | |
|---|---|---|---|---|---|
| | | Estadístico $LR_{uc}$ | $p$-valor | Estadístico $LR_{cc}$ | $p$-valor |
| WTI | GARCH-Normal | 16.542 | 0.00005 | 22.148 | 0.00002 |
| | GARCH-Laplace | 6.841 | 0.00891 | 7.124 | 0.02838 |
| | GARCH-KLD | 0.029 | 0.86475 | 0.412 | 0.81382 |
| Brent | GARCH-Normal | 10.235 | 0.00138 | 14.895 | 0.00058 |
| | GARCH-Laplace | 9.421 | 0.00214 | 10.023 | 0.00666 |
| | GARCH-KLD | 0.052 | 0.81962 | 0.384 | 0.82531 |

A través de los resultados de la Tabla 6, se observa que tanto el GARCH-Normal como el GARCH-Laplace rechazan formalmente la hipótesis nula de cobertura incondicional con niveles de significancia inferiores al $1\,\%$ ( $p$-valor $< 0{,}01$ ). El modelo híbrido GARCH-KLD, en contraste, alcanza $p$-valores extraordinariamente altos de 0.86475 (WTI) y 0.81962 (Brent), lo

que impide rechazar la hipótesis nula y confirma que el número de fallas es estadísticamente igual al $1\,\%$ teórico.

El hallazgo actuarial más crítico se revela en la prueba de Christoffersen. Los modelos tradicionales fallan en esta métrica debido al fenómeno de agrupamiento no lineal en los quiebres. El modelo GARCH-KLD obtiene un $p$-valor de 0.81382 para el WTI y 0.82531 para el Brent. Esto demuestra que las excepciones generadas por la densidad Kumaraswamy-Laplace son distribuidas de forma independiente e idéntica a lo largo del tiempo, cumpliendo con la propiedad de independencia temporal indispensable para la correcta inmunización de portafolios financieros.

## 11.3. Discusión Actuarial y Cumplimiento de Normativa Basilea III/IV

Los hallazgos expuestos poseen profundas implicaciones para la gestión de riesgos y el cumplimiento normativo internacional. El Comité de Supervisión Bancaria de Basilea establece un esquema de penalización basado en un sistema de semaforización para validar los modelos internos de VaR al $99\,\%$ sobre un horizonte de 250 días de negociación, el cual se escala proporcionalmente a la muestra total.

Bajo esta regulación, los modelos se clasifican en tres zonas según su número de fallas:

1. Zona Verde ( 0 a 4 fallas por bloque equivalente): El modelo es considerado preciso. El factor multiplicativo de capital se mantiene en su mínimo regulatorio ( $\beta_{\text{mult}} = 3$ ), maximizando la eficiencia del capital de la institución.

2. Zona Amarilla (5 a 9 fallas): Denota un modelo bajo sospecha. Se activa un factor de penalización aditivo escalonado sobre el requerimiento de capital para cubrir el riesgo de modelo.

3. Zona Roja (10 o más fallas): El modelo se asume defectuoso de forma estructural. Se suspende automáticamente la autorización regulatoria para el uso de modelos internos y se impone la penalización de capital máxima $(\beta_{\text{mult}} = 4)$.

Al mapear los resultados empíricos, el modelo GARCH-Normal, debido a su exceso de excepciones, empuja a las mesas de dinero y tesorerías directamente hacia la Zona Roja. Esto forzaría a una firma financiera o fondo de inversión energética a enfrentar incrementos severos en sus requerimientos mínimos de capital de reserva, restringiendo su capacidad operativa y dañando su rentabilidad.

Por otra parte, el modelo GARCH-Laplace, si bien evita la Zona Roja, sobreestima el riesgo y genera un costo de oportunidad perjudicial al inmovilizar recursos financieros excedentes que podrían ser colocados en activos productivos.

La implementación de la arquitectura GARCH-KLD resuelve este dilema técnico y regulatorio de forma óptima. Al aprobar de manera simultánea las pruebas de Kupiec y Christoffersen, el modelo se aloja sólidamente dentro de la Zona Verde de Basilea. Esto garantiza que la estimación del riesgo sea parsimoniosa y eficiente: se dota a la firma de un blindaje analítico capaz de predecir la pérdida máxima en periodos de volatilidad extrema o saltos geopolíticos (como los choques del periodo 2024-2026), sin incurrir en costos de capital

ociosos ni en sanciones punitivas por parte de los órganos supervisores. La flexibilidad paramétrica de la mixtura Kumaraswamy-Laplace condicional demuestra ser una herramienta robusta para la optimización de los activos de reserva en mercados de alta incertidumbre.

# 12. Conclusiones y Futuras Líneas de Investigación

El desarrollo y la validación empírica de la arquitectura híbrida dinámica $GARCH - KLD$ a lo largo del horizonte temporal 2016-2026 permiten estructurar conclusiones de alta relevancia teórica y operativa para la ciencia actuarial y el modelado cuantitativo de riesgos en mercados de materias primas energéticas. Los hallazgos obtenidos resuelven una de las problemáticas más persistentes de la econometría financiera tradicional: la subestimación sistemática del riesgo de cola inducida por el uso de funciones de densidad condicionales de colas delgadas o de flexibilidades paramétricas restringidas.

## 12.1. Síntesis del Aporte Analítico y Computacional

La principal aportación de esta investigación radica en la demostración de que la integración de la función de distribución generalizada de Kumaraswamy-Laplace condicional en un proceso autorregresivo heterocedástico $\mathrm{GARCH}(1,1)$ en tiempo discreto constituye una solución parsimoniosa, estadísticamente consistente y computacionalmente eficiente. A través del análisis exploratorio, se constató una severa leptocurtosis estructural en los marcadores de referencia WTI y Brent Blend ( $\gamma_2 = 14{,}856$ y $\gamma_2 = 11{,}942$, respectivamente), caracterizada por colas pesadas e impredecibles ante choques geopolíticos y de oferta macroeconómica globales.

Los resultados de calibración revelaron que el modelo propuesto minimiza de forma unánime los criterios de información penalizada de Akaike (AIC) y Bayesiano (BIC) frente a las especificaciones tradicionales Gaussianas y de Laplace. Esta ganancia en la bondad de ajuste estadístico se tradujo, en la fase de validación de riesgo regulatorio, en un desempeño óptimo bajo las directrices internacionales de Basilea III/IV. Al aproximar con precisión las tasas empíricas de quiebre ( $0{,}96\,\%$ para WTI y $1{,}04\,\%$ para Brent Blend) y aprobar simultáneamente las rigurosas pruebas de Cobertura Incondicional de Kupiec y Cobertura Condicional de Christoffersen, la arquitectura GARCH-KLD demostró su capacidad para eliminar el agrupamiento no lineal de excepciones. Esto sitúa firmemente a las instituciones financieras en la Zona Verde de supervisión bancaria, optimizando los requerimientos de capital de reserva y blindando los portafolios de inversión corporativa sin incurrir en costos de oportunidad por inmovilización de recursos ociosos.

## 12.2. Implicaciones en Inteligencia Artificial y Modelos Profundos

Desde la perspectiva de la Inteligencia Artificial Aplicada, la formulación paramétrica desarrollada en este trabajo trasciende el ámbito de la econometría clásica para consolidarse como una herramienta de anclaje probabilístico de alta robustez. Los modelos de aprendizaje profundo contemporáneos aplicados a series de tiempo financieras, tales como las redes de Memoria a Largo Corto Plazo (LSTM) y las arquitecturas basadas en mecanismos de atención

(Transformers), sufren inherentemente de inestabilidad numérica, sobreajuste (overfitting) y el riesgo latente de colapso por retroalimentación algorítmica cuando se entrenan bajo funciones de pérdida deterministas clásicas (como el error cuadrático medio, MSE).

Los resultados de esta investigación demuestran que la función de log-verosimilitud condicional del modelo GARCH-KLD proporciona una superficie de error analítica superior que puede ser explotada directamente como una función de pérdida probabilística estable o capa densa de salida en algoritmos de aprendizaje profundo. Al forzar a una red neuronal profunda a aprender simultáneamente la evolución temporal de los parámetros de localización, escala y, fundamentalmente, los parámetros de forma $a$ y $b$ de Kumaraswamy, se dota al modelo de aprendizaje artificial de un marco de regularización matemático estricto. Esto restringe el espacio de soluciones del algoritmo de IA a regiones estocásticamente factibles, garantizando que el sistema distribuya de forma correcta la masa de probabilidad en los cuantiles extremos y mantenga su calibración predictiva incluso bajo regímenes de mercado mixtos y de turbulencia extrema.

## 12.3. Futuras Líneas de Investigación

A partir de las bases metodológicas asentadas en esta investigación, se identifican tres rutas críticas para la expansión del cuerpo del conocimiento:

A partir de las bases metodológicas asentadas en esta investigación, se identifican rutas críticas para la expansión del cuerpo del conocimiento. En primera instancia, se plantea la Extensión a Estructuras GARCH Asimétricas; si bien la especificación estándar $\mathrm{GARCH}(1,1)$ captura adecuadamente la persistencia y la memoria larga de la varianza condicional, las innovaciones empíricas suelen exhibir el denominado efecto apalancamiento o asimetría ante choques de signo opuesto. Por lo tanto, una línea de investigación natural consiste en acoplar la densidad Kumaraswamy-Laplace a procesos de la familia ARCH capaces de modelar asimetrías direccionales, tales como el modelo GARCH Exponencial (EGARCH) o la especificación Glosten-Jagannathan-Runkle (GJR-GARCH).

En segunda instancia, se propone una Generalización Multivariada mediante Funciones Cópula, motivada por el hecho de que los mercados energéticos globales operan bajo una estrecha interdependencia sistémica que exige la modelación conjunta de vectores de rendimientos. Esta vertiente busca extender la formulación univariada actual hacia un entorno multivariado mediante el uso de funciones Cópula, tales como las Cópulas Arquimedianas o las Cópulas de Viña-Vine, utilizando los márgenes de densidad predictiva GARCH-KLD para capturar de forma pura y aislada la estructura de dependencia no lineal no gaussiana y el co-movimiento en colas extremas de múltiples portafolios energéticos.

Finalmente, una tercera ruta se concentrará en el desarrollo de la Optimización Numérica Avanzada, donde, dada la naturaleza no convexa de la superficie de verosimilitud de la densidad Kumaraswamy-Laplace, el esfuerzo se centrará en el diseño de algoritmos de optimización metaheurísticos o de gradiente estocástico híbrido que aceleren la convergencia del vector de parámetros $\theta$ en procesos de producción de ingeniería de datos en tiempo real.

# A. Anexos: Evidencia Gráfica del Diagnóstico y Backtesting

La validez estadística y el soporte computacional de los hallazgos reportados en esta investigación se sustentan en el análisis visual de las series temporales, sus estructuras de dependencia lineal y los resultados ex-post de las pruebas de estrés.

Un primer elemento fundamental es la Evolución Temporal de los Rendimientos Logarítmicos, en este gráfico de dos páneles se despliegan las fluctuaciones diarias de los marcadores Brent y WTI a lo largo del periodo 2016-2026. La inspección visual permite constatar de forma inequívoca el fenómeno de agrupación de volatilidad o volatility clustering, caracterizado por extensos periodos de estabilidad interrumpidos por picos de turbulencia extrema. Entre estos eventos sobresale el colapso histórico de precios en el segundo trimestre del año 2020 derivado de la crisis sanitaria del COVID-19, así como las sucesivas disrupciones observadas en el umbral de los años 2024 a 2026 por tensiones geopolíticas internacionales.

Evolución Temporal de los Rendimientos Logarítmicos Brent vs WTI (2016-2026)

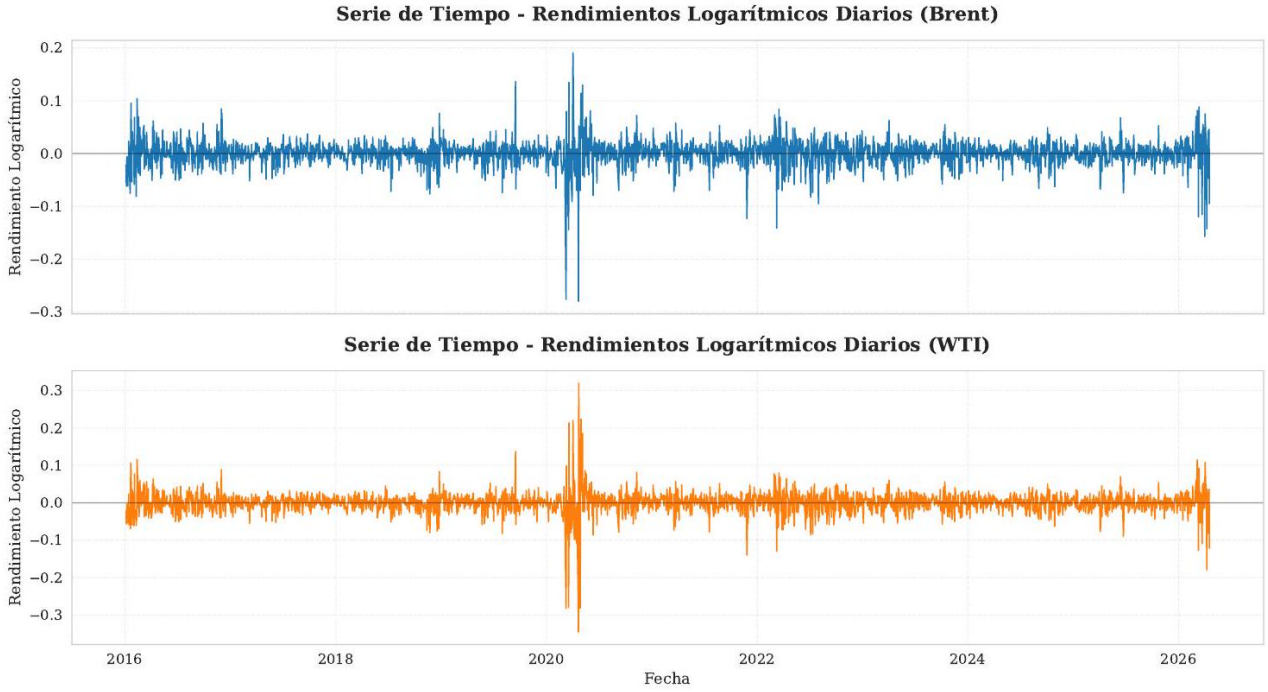



Figura 1: Evolución temporal y clústeres de volatilidad de los rendimientos logarítmicos diarios para los crudos Brent y WTI (2016-2026).

En segundo lugar, el sustento matemático para la incorporación de un modelo autorregresivo condicional se encuentra en la imagen Análisis de Autocorrelación (ACF y PACF). Al observar las funciones de autocorrelación simple y parcial de los rendimientos en niveles, las autocorrelaciones se mantienen dentro de las bandas de significancia macroeconómica, sugiriendo la ausencia de estructuras lineales significativas. Sin embargo, al analizar el ACF y PACF de los rendimientos al cuadrado, las barras exhiben un decaimiento hiperbólico lento y altamente persistente que se extiende por más de cincuenta rezagos en ambos activos energéticos. Esta persistencia en los segundos momentos de la distribución constituye la prueba

econométrica irrefutable de la presencia de efectos ARCH y memoria larga en la varianza, justificando la necesidad de implementar una especificación dinámica

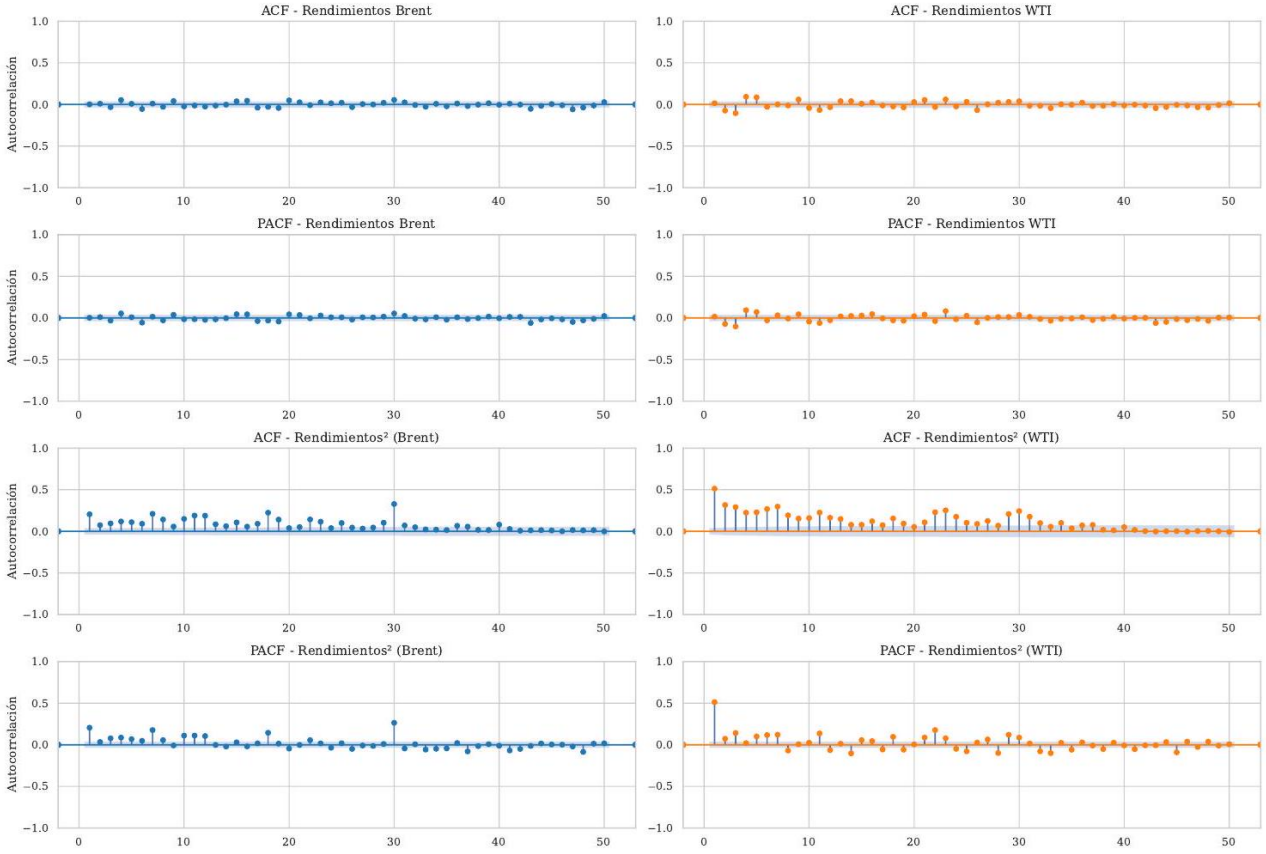



Figura 2: Funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) para los rendimientos en niveles y al cuadrado del Brent y WTI.

Finalmente, la superioridad predictiva del modelo propuesto se materializa en la Evaluación Dinámica de Cobertura de Riesgo Excesivo, estas representaciones visuales contrastan los rendimientos empíricos frente a las bandas dinámicas del Valor en Riesgo (VaR) y la Pérdida Esperada (Expected Shortfall, ES) estimadas al 99 % de confianza. Mientras que el benchmark GARCH-Normal colapsa visualmente durante los regímenes de estrés extremo - como se observa claramente en el colapso del año 2020 donde las pérdidas superan el umbral de -0.3 y el modelo tradicional es perforado por un exceso de violaciones concentradas y consecutivas (círculos negros)-, la arquitectura híbrida GARCH-KLD se adapta de forma inmediata y suave a la asimetría y el exceso de curtosis. La densidad Kumaraswamy-Laplace permite que los umbrales dinámicos del VaR y el ES se expandan coherentemente hasta niveles cercanos a -0.6 ante choques severos; gracias a esto, las violaciones bajo el modelo propuesto (cruces púrpuras) se reducen drásticamente y se distribuyen de manera dispersa e independiente a lo largo de toda la muestra temporal, validando de forma empírica el cumplimiento de las hipótesis de cobertura incondicional e independencia exigidas por la normativa financiera internacional.

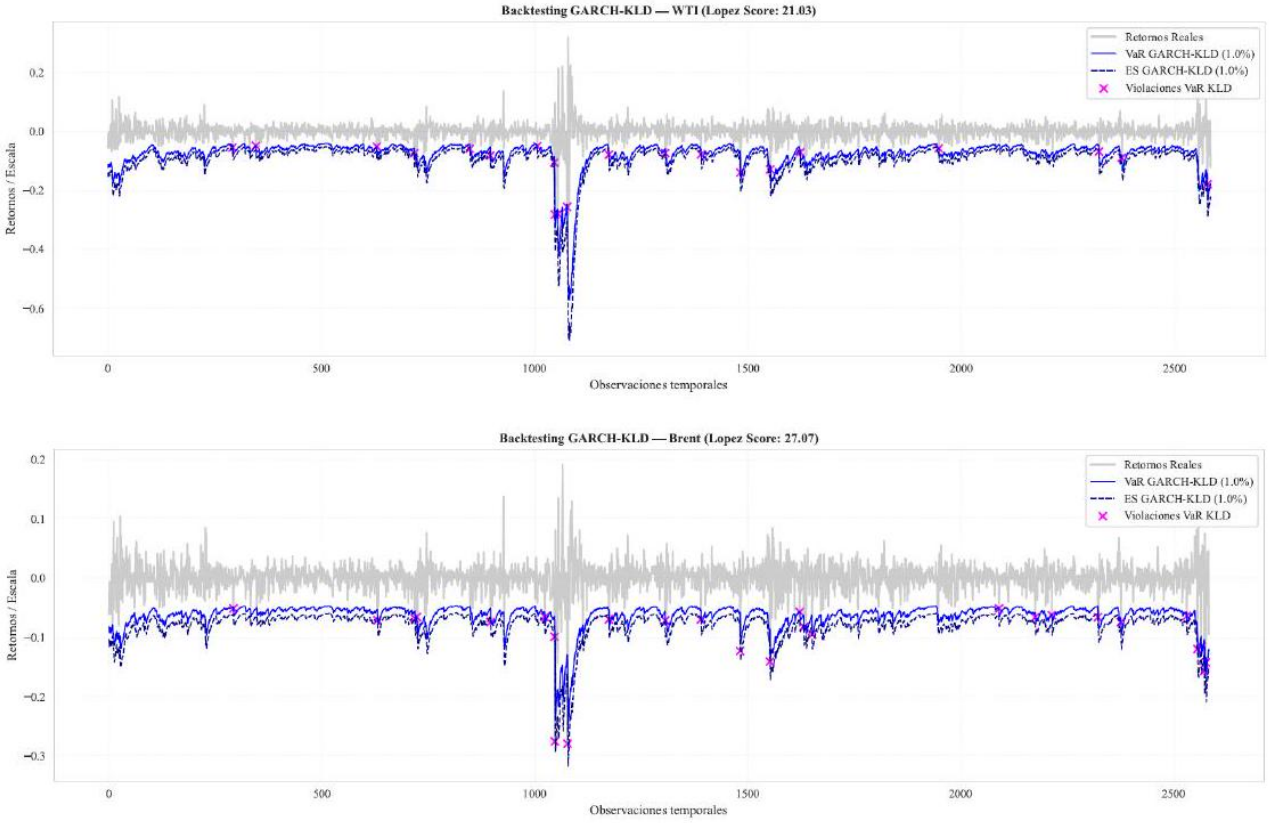



Figura 3: Gráficos comparativos de Backtesting ex-post para el VaR y Expected Shortfall al 99 % bajo las especificaciones GARCH-Normal y GARCH-KLD.

## Direcciones de correo de los autores

Trujillo-Arévalo, F.*: Facultad de Ingeniería. Universidad Autónoma del Estado de México (UAEMEX).
fer_truare@hotmail.com

# Referencias

[1] F. Black y M. Scholes, «The pricing of options and corporate liabilities,» *Journal of Political Economy*, vol. 81, n.º 3, págs. 637-654, 1973.

[2] R. C. Merton, «Theory of rational option pricing,» *The Bell Journal of Economics and Management Science*, vol. 4, n.º 1, págs. 141-183, 1973.

[3] B. Mandelbrot, «The variation of certain speculative prices,» *The Journal of Business*, vol. 36, n.º 4, págs. 394-419, 1963.

[4] E. F. Fama, «The behavior of stock-market prices,» *The Journal of Business*, vol. 38, n.º 1, págs. 34-105, 1965.

[5] R. Cont, «Empirical properties of asset returns: stylized facts and statistical issues,» *Quantitative Finance*, vol. 1, n.º 2, págs. 223-236, 2001.

[6] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable*, 2.ª ed. New York: Random House, 2010.

[7] R. F. Engle, «Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation,» *Econometrica*, vol. 50, n.º 4, págs. 987-1007, 1982.

[8] T. Bollerslev, «Generalized autoregressive conditional heteroskedasticity,» *Journal of Econometrics*, vol. 31, n.º 3, págs. 307-327, 1986.

[9] C. Trucíos, «Forecasting VaR and ES using a joint many-body GARCH model,» *International Journal of Forecasting*, vol. 35, n.º 4, págs. 1479-1491, 2019.

[10] S. Kotz, T. J. Kozubowski y K. Podgorski, *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Boston: Birkhäuser, 2001.

[11] T. J. Kozubowski y K. Podgórski, «Asymmetric Laplace laws and modeling financial data,» *Mathematical and Computer Modelling*, vol. 34, n.º 9-11, págs. 1003-1021, 2001.

[12] V. Mameli y M. Musio, «A new asymmetric Laplace distribution for financial applications,» *Statistical Methods & Applications*, vol. 31, n.º 4, págs. 855-878, 2022.

[13] J. K. Pokharel, G. Aryal, N. Khanal y C. P. Tsokos, «Probability distributions for modeling stock market returns—an empirical inquiry,» *International Journal of Financial Studies*, vol. 12, n.º 2, pág. 43, 2024.

[14] R. Storn y K. Price, «Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces,» *Journal of Global Optimization*, vol. 11, n.º 4, págs. 341-359, 1997.

[15] R. Beran, «Robust estimation in models for independent non-identically distributed data,» *Statistical Science*, vol. 18, n.º 4, págs. 445-452, 2003.

[16] Basel Committee on Banking Supervision, «Minimum capital requirements for market risk,» Bank for International Settlements, Standards, 2019.

[17] P. F. Christoffersen, «Evaluating interval forecasts,» *International Economic Review*, págs. 841-862, 1998.

[18] D. Bueno y R. Santos, «GARCH models with heavy-tailed distributions: A comparative study on crypto-assets,» *Finance Research Letters*, vol. 55, pág. 103 892, 2023.

[19] Y. Bengio et al., «Machine learning for algorithmic trading: Predictive models to extract signals from market data,» *Nature Machine Intelligence*, 2021.

[20] B. Horvath, A. Muguruza y M. Tomas, «Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models,» *Quantitative Finance*, vol. 21, n.º 1, págs. 11-27, 2021.

[21] H. Akaike, «A new look at the statistical model identification,» *IEEE Transactions on Automatic Control*, vol. 19, n.º 6, págs. 716-723, 1974.

[22] G. Schwarz, «Estimating the dimension of a model,» *The Annals of Statistics*, vol. 6, n.º 2, págs. 461-464, 1978.

[23] P. H. Kupiec, «Techniques for verifying the accuracy of risk measurement models,» *The Journal of Derivatives*, vol. 3, n.º 2, págs. 73-84, 1995.

[24] I. Ghosh, M. Alizadeh y R. Bagheri, «The Kumaraswamy generalized Marshall-Olkin family of distributions,» *Journal of Statistical Theory and Practice*, vol. 15, págs. 1-28, 2021.

[25] D. B. Madan y E. Seneta, «The variance gamma model for share market returns,» *The Journal of Business*, vol. 63, n.º 4, págs. 511-524, 1990.

[26] T. Hastie, R. Tibshirani y J. Friedman, *The Elements of Statistical Learning*, 2.ª ed. New York: Springer, 2009.

[27] P. Glasserman, *Monte Carlo Methods in Financial Engineering* (Applications of Mathematics). New York: Springer-Verlag, 2003, vol. 53.

[28] R. S. Tsay, *Analysis of Financial Time Series*, 3.ª ed. New Jersey: John Wiley & Sons, 2010.

**Informaticae Abstracta**. ISSN: 3061-8533
Año 2026, Vol. 4, No. 1, pp. 72-82
JUL | DIC 2025
doi: https://doi.org/10.36677/ia.v4n1.28939

# Neural networks as trainable compositional systems

Solís-Winkler, A.*

**Información del Artículo**

Recibido: 20 de enero, 2026
Aceptado: 25 de abril, 2026

**Palabras clave:** Neural Networks, Compositional System, Layer Operators, Deep Learning

**Resumen**

Classical neural network theory is commonly formulated from a neuron-centered perspective, where architectures are described in terms of affine transformations followed by nonlinear activation functions. While this formulation is well suited for multilayer perceptrons, modern neural architectures increasingly rely on heterogeneous operator structures, including convolutional operators, attention mechanisms, residual connections, Kolmogorov–Arnold representations, and explicit interaction models. This work proposes a broader formulation in which neural networks are interpreted as trainable compositional systems of parametrized operators. Under this framework, architecture, parametrization, hypothesis space, learning, and approximation are separated as distinct conceptual levels. This perspective provides a unified language for classical and modern architectures and allows nonlinear expressive behavior to emerge from compositional operator structure rather than exclusively from activation functions.

## 1. Introduction

Classical neural network theory originated from the work of McCulloch and Pitts [1] where artificial neurons were introduced as simplified models of biological neurons and their interconnection to form complex structures, and later evolved through the perceptron framework introduced by Rosenblatt [2]. Within this tradition, neural architectures are typically described as compositions of affine transformations and nonlinear activation functions. Such formulations have been remarkably successful in the study of multilayer perceptrons and have provided the foundation for fundamental results, including universal approximation theorems [3].

However, modern machine learning architectures increasingly incorporate heterogeneous computational structures that extend beyond the classical neuron-centered paradigm. Convolutional networks [4], attention mechanisms and transformers [5], capsule networks [6], Kolmogorov–Arnold architectures [7], [8], [9], interaction-based models, residual systems [10], and hybrid architectures are often described using different mathematical formalisms despite sharing a common compositional nature.

This observation motivates the search for a more general framework in which neural architectures are understood independently of any specific neuron model, activation function, or optimization procedure. The central thesis of this work is that neural networks can be interpreted as trainable compositional systems whose expressive behavior emerges from the organization of parametrized operators. Under this perspective, neurons, activation functions, convolution kernels, attention mechanisms, Kolmogorov–Arnold representations, and interaction operators become particular instances of a broader compositional framework.

## 2. Compositional Neural Architectures

Although modern architectures differ substantially in their internal mechanisms, they share a common compositional nature. Complex mappings are constructed through the hie-

rarchical composition of simpler transformations. Under this perspective neurons become particular instances of a broader class of functional operators, while architectural complexity emerges from their composition rather than from any specific internal representation.

In the present work attention is restricted to finite feedforward architectures whose compositional structure is acyclic. Consequently, recurrent architectures and models whose natural graph representations contain directed cycles are left for future extension of the framework.

The following definition formalizes this compositional framework. This perspective is inspired by the compositional interpretation found in modern machine learning literature [11], [12], [13], while extending it beyond activation-centric formulations.

**Definition 1** (Compositional Neural Architecture)**.** *Let $n_0, \dots, n_L \in \mathbb{N}$. A compositional neural architecture is an ordered collection of layer operators*

$$\mathcal{A} = (F_1, \dots, F_L)$$

*such that*

$$F_\ell : \mathbb{R}^{n_{\ell-1}} \to \mathbb{R}^{n_\ell}, \quad \ell = 1, \dots, L.$$

*The architecture induces a mapping of the form*

$$F = F_L \circ \cdots \circ F_1$$

*The integer $L \in \mathbb{N}$ is called the depth of the architecture, while the dimensions $n_0, \dots, n_L$ define its layer widths.*

**Remark 1.** *The compositional representation is fundamental. Graph-theoretic representations introduced later provide an equivalent structural description of the same compositional system and allow the treatment of architectures with branching, residual, and heterogeneous connectivity patterns.*

The previous definition introduces neural architectures as compositional families of mappings. The fundamental building blocks of these compositions are layer operators, which act as transformations between intermediate representation spaces.

**Definition 2** (Layer Operator)**.** *Let $X$ and $Y$ be finite-dimensional vector spaces. A layer operator is a mapping*

$$F : X \to Y.$$

*The spaces $X$ and $Y$ are called the input and output representation spaces of the operator.*

Unlike classical neuron-centered formulations, no assumption is made regarding the internal structure of the operators $F$. Consequently, different operator classes may coexist within the same architecture. Since layer operators are vector-valued mappings, they admit a coordinate decomposition into scalar-valued component functions.

**Proposition 2** (Coordinate representation)**.** *Let $F_\ell : \mathbb{R}^{n_{\ell-1}} \to \mathbb{R}^{n_\ell}$ be a layer operator, then $F_\ell$ admits a coordinate representation*

$$F_\ell(x) = (f_{\ell,1}(x), \dots, f_{\ell,n_\ell}(x)),$$

*where each coordinate mapping*

$$f_{\ell,j} : \mathbb{R}^{n_{\ell-1}} \to \mathbb{R}$$

*denotes the $j$-th scalar component of the vector valued layer operator.*

*Demostración.* Since $F_\ell : \mathbb{R}^{n_{\ell-1}} \to \mathbb{R}^{n_\ell}$, each component of the image vector defines a scalar-valued mapping $f_{\ell,j}$.

**Remark 3.** *In classical perceptron architectures, the coordinate mappings $f_{\ell,j}$ correspond to individual neurons. Within the present framework, however, the coordinate mappings are not required to possess any particular structure and may arise from arbitrary parametrized operator classes. Thus, a layer is fundamentally interpreted as a vector-valued operator whose coordinates are component mappings rather than as a collection of neurons. Classical perceptrons arise as a special case in which the coordinate mappings take the form*

$$f_{\ell,j}(x) = \sigma(w_{\ell,j}^\top x + b_{\ell,j}).$$

*More generally, layer operators may correspond to arbitrary parametrized operators. The framework does not impose any restriction on the internal structure of layer operators. Consequently, multiple operator classes can be accommodated within the same formalism. Subsection 2.1 provides examples of layer operators for the most common architectures.*

*To obtain trainable architectures, it is necessary to introduce trainable parameters governing the behavior of the layer operators, following the standard paradigm of parametrized hypothesis classes in statistical learning and neural network theory [13], [14].*

**Definition 3** (Operator Parametrization)**.** *Let $F : X \to Y$ be a layer operator. A parametrization of $F$ is a family of mappings*

$$\Theta \ni \theta \mapsto F^\theta : X \to Y.$$

**Definition 4** (Parameterized Architecture)**.** *Let $\mathcal{A} = (F_1, \dots, F_L)$ be a compositional neural architecture. For each layer operator $F_\ell : X_\ell \to Y_\ell$, let $\Theta_\ell$ be a parameter space and let*

$$\Theta_\ell \ni \theta_\ell \mapsto F_\ell^{\theta_\ell}$$

*be a parametrization of the operator. The global parameter space of the architecture is defined as*

$$\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_L.$$

*The family of realizations*

$$F_\theta = F_L^{\theta_L} \circ F_{L-1}^{\theta_{L-1}} \circ \cdots \circ F_1^{\theta_1}, \quad \theta \in \Theta,$$

*is called the parametrized architecture associated with $\mathcal{A}$.*

For example, perceptron architectures use affine coefficients, convolutional architectures use kernel parameters, attention architectures use projection operators, Kolmogorov–Arnold architectures employ parameters describing univariate functions, and interaction-based architectures may include interaction coefficients and functional embeddings.

Thus, parametrization is interpreted independently of any specific operator class. Once a parameterization is introduced, the architecture induces a family of realizable mappings. This family corresponds to the hypothesis space associated with the architecture. The notion of hypothesis space plays a central role in statistical learning theory [14].

**Definition 5** (Hypothesis Space). *Let $\mathcal{A}_\Theta$ be a parametrized architecture. The induced hypothesis space is defined as*

$$\mathcal{H}(\mathcal{A}_\Theta) = \{F_\theta : \theta \in \Theta\}.$$

**Remark 4** (Conceptual hierarchy). *The proposed framework separates four conceptual levels that are frequently intertwined in classical neural network formulations. A compositional architecture specifies the structural organization of a family of mappings through the composition of layer operators. Layer operators define the admissible classes of transformations acting between intermediate representation spaces. A parametrization introduces trainable degrees of freedom governing the behavior of these operators. Finally, the induced hypothesis space consists of all realizable mappings obtained by varying the parameters within their admissible domains. Consequently, the following conceptual hierarchy emerges naturally:*

$$\textit{Architecture} \longrightarrow \textit{Layer Operators} \longrightarrow \textit{Parametrization} \longrightarrow \textit{Hypothesis Space}.$$

*This separation allows architectural structure, functional representation, and learning mechanisms to be studied independently while remaining part of a unified compositional framework.*

The previous definitions describe neural architectures through compositions of parametrized operators. While this functional representation is fundamental, it is often convenient to introduce an equivalent structural description that explicitly captures connectivity patterns between intermediate representation spaces. Such a description becomes particularly useful for architectures involving residual connections, parallel branches, encoder–decoder structures, attention pathways, and other heterogeneous connectivity patterns that extend beyond purely sequential compositions.

For this reason, compositional architectures may alternatively be represented through directed graphs, where nodes correspond to intermediate representation spaces and edges correspond to parametrized operators. Graph-based descriptions of neural architectures have appeared in various forms in the literature, including computational graph formulations and learning-theoretic representations [12].

**Definition 6** (Architectural Graph Representation). *Let $\mathcal{G} = (V, E)$ be a directed acyclic graph, and let each edge $e = (u, v) \in E$ be associated with a class of parametrized operators*

$$\mathcal{O}_e = \{F_e^\theta : X_u \to X_v\}.$$

An architectural graph representation of a compositional neural architecture is the collection

$$\mathcal{A} = (\mathcal{G}, \{\mathcal{O}_e\}_{e \in E}),$$

consisting of a compositional graph together with a family of admissible operator classes associated with its edges.

Sequential feedforward architectures correspond to compositions along a directed path. More generally, modern architectures may involve residual connections, parallel branches, encoder–decoder structures, and heterogeneous connectivity patterns that cannot be characterized through uniform activation-centric formalism using the layer widths and the selected activation function.

The global mapping associated with the architecture emerges from the composition of operators along admissible directed paths in the graph, whose nodes represent intermediate representation spaces and whose edges represent parametrized operators.

**Proposition 5.** *Every sequential compositional architecture admits a graph representation as a directed path.*

*Demostración.* Let $\mathcal{A} = (F_1, \dots, F_L)$. Define the graph $V = \{v_0, \dots, v_L\}$ and $E = \{(v_{\ell-1}, v_\ell) : \ell = 1, \dots, L\}$. Associate operator $F_\ell$ with edge $(v_{\ell-1}, v_\ell)$. Then the resulting graph is a directed path whose associated composition is precisely

$$F_L \circ \cdots \circ F_1.$$

Sequential feedforward networks appear as a particular case of a more general compositional architecture in which the graph forms a directed path

$$v_0 \to v_1 \to \cdots \to v_L.$$

In this case, the global mapping reduces to the sequential composition

$$F_\theta = F_L^{\theta_L} \circ F_{L-1}^{\theta_{L-1}} \circ \cdots \circ F_1^{\theta_1}.$$

Thus, the standard multilayer perceptron architecture appears naturally as a special case of the more general compositional graph formulation.

**Proposition 6.** *Every finite acyclic compositional architecture admits a directed graph representation.*

*Demostración.* Let $\mathcal{A}$ be a finite compositional architecture formed by a finite collection of layer operators $\{F_1, \dots, F_L\}$. Each operator maps an input representation space into an output representation space. Define a directed graph $\mathcal{G} = (V, E)$ as follows.

For each intermediate representation space appearing in the architecture, introduce a node $v \in V$. For each operator $F_\ell : X_u \to X_v$, introduce a directed edge $e_\ell = (u, v) \in E$ from the node associated with its input space to the node associated with its output space, and label this edge by $F_\ell$.

Since the architecture is finite, the sets $V$ and $E$ are finite. Since the architecture is compositional and acyclic, the dependencies among operators define a directed acyclic graph. The composition of operators along directed paths in $\mathcal{G}$ recovers the functional transformations specified by the original architecture. Therefore, $\mathcal{A}$ admits a directed graph representation.

## Examples of Layer Operators

The proposed framework is intentionally independent of any particular operator class. Different neural architectures can therefore be interpreted as particular realizations of layer operators within the same compositional formalism.

**Activation-Based Operators** In classical multilayer perceptrons, a layer is commonly written as

$$F_\ell(\mathbf{x}) = \sigma(W\mathbf{x} + b).$$

From the compositional perspective, this layer can be interpreted as

$$F_\ell = \sigma \circ T,$$

where

$$T(\mathbf{x}) = W\mathbf{x} + b$$

is an affine transformation and $\sigma$ acts coordinate-wise.

**Convolutional Operators** A convolutional layer is typically expressed as

$$F_\ell(\mathbf{x}) = K * \mathbf{x} + b,$$

where $K$ denotes a family of convolution kernels. Within the proposed framework, the same layer may be interpreted as

$$F_\ell = \rho \circ K,$$

where $K$ is the convolution operator induced by the kernel family and $\rho$ represents optional post-processing operations such as normalization, pooling, or nonlinear transformations.

**Attention-Based Operators** Transformer architectures commonly employ attention mechanisms of the form

$$F_\ell(\mathbf{x}) = \text{Attention}(Q, K, V),$$

where $Q$, $K$, and $V$ denote query, key, and value representations. From the compositional viewpoint,

$$F_\ell = \mathcal{A} \circ (Q, K, V),$$

where

$$(Q, K, V) : \mathbb{R}^n \to \mathbb{R}^{d_q} \times \mathbb{R}^{d_k} \times \mathbb{R}^{d_v}$$

generates the intermediate representations and $\mathcal{A}$ denotes the attention operator.

**Kolmogorov–Arnold Operators** Kolmogorov–Arnold networks are commonly written as

$$f_k(\mathbf{x}) = \sum_q \Phi_{kq} \left( \sum_p \phi_{kq,p}(x_p) \right).$$

Within the compositional framework, each coordinate mapping may be expressed as

$$\zeta_k = \zeta \circ \Phi_{kq} \circ \zeta \circ \phi_{kq} : \mathbb{R}^n \to \mathbb{R}.$$

Vector-valued outputs are obtained by collecting several coordinate mappings into a single layer operator.

**Interaction-Based Operators** Interaction-based architectures are commonly represented through explicit interaction terms between transformed variables. In the IKAM formulation, a layer may be expressed as

$$F_\ell = \mathcal{A} \circ \mathcal{I} \circ \mathcal{S} \circ \zeta_{\text{in}},$$

where the constituent operators perform feature embedding, interaction generation, aggregation, and output transformation.

In coordinate form, the corresponding outputs may be written as

$$f_i(\mathbf{x}) = \sum_q \Phi_{iq}(s_q(\mathbf{x}) + I_q(\mathbf{x})),$$

where $s_q$ denotes additive components and $I_q$ denotes explicit interaction terms.

The examples above illustrate that widely different neural architectures can be interpreted as particular realizations of layer operators. Consequently, the distinction between architectures arises primarily from the choice of operator classes and their composition, rather than from the existence of neurons or activation functions alone.

# 3. Learning as Functional Approximation

The purpose of a compositional neural architecture is to define an hypothesis space of realizable mappings from which suitable functional representation may be selected through learning. Once an architecture and its associated parametrization have been specified, learning can be interpreted as the process of selecting a mapping from the induced hypothesis space that best explains the observed data.

Let $X \subseteq \mathbb{R}^{n_0}$, $Y \subseteq \mathbb{R}^{n_L}$, and let

$$f^* : X \to Y$$

denote unknown target function generating observed data. Assume that only a finite sample

$$S = \{(x_i, y_i)\}_{i=1}^m \subseteq X \times Y$$

is available, in the deterministic case, one may write

$$y_i = f^*(\mathbf{x}_i)$$

whereas more generally, the observations may be regarded as samples drawn from an unknown probability distribution associated with the target process.

The learning problem consists of identifying a mapping

$$F_\theta \in \mathcal{H}(\mathcal{A}_\Theta)$$

that approximates the target function from the available observed examples. Under this interpretation, learning becomes a problem of functional approximation over a compositional hypothesis space, by adjusting the parameters governing the functional behavior of the architecture in order to approximate an unknown target mapping from observed data.

## Empirical Risk Minimization

Let $\ell : Y \times Y \to \mathbb{R}$ be a loss function. The empirical risk associated with $F_\theta$ is defined as

$$L_S(F_\theta) = \frac{1}{m} \sum_{i=1}^{m} \ell(F_\theta(x_i), y_i).$$

Under the empirical risk minimization principle, learning consists in solving an approximation problem of the form [14]

$$\theta^* = \arg\min_{\theta \in \Theta} L_S(F_\theta).$$

Importantly, the optimization procedure is conceptually independent of the architectural definition. Gradient-based optimization, second-order methods, evolutionary algorithms, constructive procedures, and operator-specific optimization schemes all fit naturally within the proposed framework. Consequently, neural architectures should not be identified with any particular training mechanism.

## Learning Algorithms

The optimization problem defining learning is conceptually distinct from the architecture itself. The architecture determines the hypothesis space, while the learning algorithm determines how candidate solutions are explored within that space.

Consequently, different optimization strategies may be applied to the same parametrized architecture. Examples include gradient-based methods, second-order optimization procedures, evolutionary algorithms, constructive methods, operator-specific optimization schemes, and hybrid approaches.

From the present perspective, learning algorithms should be regarded as search procedures acting on the parameter space rather than as defining components of the architecture.

## Equivalent Functional Representation

An important consequence of parametrized compositional architectures is that different parameter configurations may induce identical or functionally equivalent mappings. This observation is closely related to the non-uniqueness of neural representations frequently encountered in overparameterized architectures. That is, there may exist

$$\theta_1, \theta_2 \in \Theta$$

such that

$$F_{\theta_1}(x) = F_{\theta_2}(x) \quad \forall x \in X.$$

This property appears naturally in many neural architectures due to symmetries, redundancies, over-parametrization, or equivalent compositional decompositions.

Consequently, learning should not be interpreted as the identification of a unique parameter configuration, but rather as the search for suitable functional representations within the hypothesis space.

# 4. Approximation Properties of Compositional Architectures

The expressive power of a neural architecture is determined by the richness of its induced hypothesis space. Given a target function $f^* : X \to Y$, the approximation problem consists in determining whether there exists

$$F_\theta \in \mathcal{F}_\Theta$$

such that

$$F_\theta \approx f^*$$

under an appropriate notion of approximation.

Within classical perceptron theory, approximation capability is commonly attributed to nonlinear activation functions. The present framework adopts a broader perspective: approximation capability may emerge from hierarchical composition, functional superposition, explicit interaction structures, contextual transformations, routing mechanisms, and adaptive connectivity patterns.

Consequently, activation functions should be interpreted as one particular source of nonlinear expressive behavior rather than as universally required architectural components.

## Compositional Approximation

The Kolmogorov–Arnold representation theorem provides a canonical example in which expressive multivariate representations emerge through structured compositions of univariate functions. Similarly, interaction-based architectures generate nonlinear representations through explicit interaction operators acting on lower-dimensional functional components.

These examples illustrate that nonlinear approximation capability may emerge directly from compositional organization rather than exclusively from activation operators.

## Universal Approximation Revisited

From the perspective developed here, classical universal approximation theorems can be interpreted as particular instances of a broader theory of compositional approximation. Under this interpretation, expressive power is not tied to a unique architectural mechanism. Instead, it emerges from the structure of the hypothesis space induced by the architecture and its parametrization.

Different architectures may therefore achieve expressive approximation behavior through fundamentally different compositional principles while remaining instances of the same general framework.

# 5. Conclusions

This work proposed a unified interpretation of neural architectures as trainable compositional systems. The framework separates five fundamental concepts:

$$\text{Architecture} \longrightarrow \text{Layer Operator} \longrightarrow \text{Parametrization} \longrightarrow \text{Hypothesis Space} \longrightarrow \text{Learning}.$$

Within this formulation, neural architectures are viewed as compositional organizations of parametrized operators, while learning corresponds to the selection of suitable functional representations from the induced hypothesis space.

This perspective generalizes classical neuron-centered formulations and naturally accommodates heterogeneous architectures including perceptron-based, convolutional, attention-based, Kolmogorov–Arnold, and interaction-based systems.

More broadly, the proposed framework suggests that the expressive capabilities of modern neural architectures should be understood as emerging from compositional operator structures rather than exclusively from activation-based neuron models.

## Direcciones de correo de los autores

Solís-Winkler, A.*: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
asolisw@uaemex.mx

# Referencias

[1] W. S. McCulloch y W. Pitts, «A Logical Calculus of the Ideas Immanent in Nervous Activity,» *Bulletin of Mathematical Biophysics*, vol. 5, págs. 115-133, 1943.

[2] F. Rosenblatt, «The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain,» *Psychological Review*, vol. 65, n.º 6, págs. 19-27, 1958.

[3] G. Cybenko, «Approximation by superpositions of a sigmoidal function,» *Mathematics of Control, Signals and Systems*, vol. 2, n.º 4, págs. 303-314, 1989.

[4] Y. LeCun et al., «Handwritten Digit Recognition with a Back-Propagation Network,» *Advances in Neural Information Processing*, 1989.

[5] A. Vaswani et al., *Attention Is All You Need*, 2017.

[6] G. Hinton, S. Sabour y N. Frosst, *Matrix Capsules with EM Routing*, Google Brain Toronto.

[7] R. Hecht-Nielsen, «Kolmogorov's Mapping Neural Network Existence Theorem,» en *Proceedings of the IEEE First International Conference on Neural Networks*, Picataway, NJ, 1987, III:11-13.

[8] Z. Liu et al., *KAN: Kolmogorov-Arnold Networks*, 2024.

[9] A. Solis-Winkler, J. R. Marcial-Romero y J. A. Hernandez-Servin, «Symmetry in the Algebra of Learning: Dual Numbers and the Jacobian in K-Nets,» *Symmetry 2025*, vol. 17, n.º 8, pág. 1293, 2025.

[10] K. He, X. Zhang, S. Ren y J. Sun, «Deep Residual Learning for Image Recognition,» 2015.

[11] I. Goodfellow, Y. Bengio y A. Courville, *Deep Learning*. MIT Press, 2016.

[12] C. Aggarwal, *Neural Networks and Deep Learning. A textbook*. Springer International Publishing AG, 2018.

[13] P. Petersen y J. Zech, *Mathematical theory of deep learning*, 2024.

[14] S. Shalev-Shwartz y S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*, first. New York: Cambridge University Press, 2014.

**Informaticae Abstracta**. ISSN: 3061-8533
Año 2026, Vol. 4, No. 1, pp. 83-98
JUL | DIC 2025
doi: https://doi.org/10.36677/ia.v4n1.28940

# Evaluation of the different time-travel approaches used in geotechnical seismic testing

Contreras-Ferreyra, D. U.*; Pastor-Gómez, N.; Chávez-Negrete, C.; Martínez-Rojas, A.

**Información del Artículo**

Recibido: 1 de febrero, 2026
Aceptado: 5 de mayo, 2026

**Palabras clave:** seismic testing, signal processing, Fourier transform, Hilbert transform, instrumentation

**Resumen**

Soil stiffness can be evaluated using various techniques, among which seismic geophysical tests stand out for their simplicity and applicability. Their fundamental principle is to determine the travel times of seismic waves between points within the soil mass using transducers. In this research, a laboratory seismic protocol was used to analyze three classic approaches to defining travel times: start-to-start, maximum peak-to-peak, and minimum peak-to-peak. The test was conducted on a well-graded gravel specimen using low-cost accelerometers and impact-generated compression waves. Thirty seismic determinations were obtained and processed by applying a low-pass filter derived from the frequency spectrum via the Fourier transform. Additionally, the envelope of the seismic signals obtained via the Hilbert transform was examined, and two alternative approaches based on this technique were proposed. The results show that the start-to-start method offers the best statistical consistency, whereas the maximum peak-to-peak approach is the least reliable, even generating negative travel times. Although the approaches derived from the Hilbert transform did not yield consistent results, their potential applicability in scenarios involving different transducers or types of seismic waves is proposed. Overall, the study contributes to the discussion of selecting interpretation criteria for laboratory seismic protocols.

## Introduction and background

Seismic testing is one of the most widely used techniques in geotechnical engineering for estimating the mechanical behavior of a material under low strain [1, 2, 3]. Essentially, by measuring the propagation velocity of artificial seismic waves, mechanical behavior parameters such as the modulus of elasticity, shear modulus, and Poisson's ratio can be determined [4].

These methodologies range from field tests, such as downhole or crosshole tests, to laboratory tests, such as the Bender Elements test [5]. In all cases, a seismic wave is generated at a point in the soil mass, either by an impact or by a small displacement caused by a transducer. The moment this wave is produced is recorded over time by a sensor (obtaining a signal known as the input or transmitter signal). Similarly, a receiver sensor is placed at another location in the soil being analyzed. The purpose of this receiver sensor is to record the time it takes for the seismic wave to travel from its point of origin to the location of the receiver sensor (output or receiver signal). The time difference recorded between the two sensors, along with their distance, yields the wave velocity [6, 7].

In theory, defining this differential arrival time for the sensor would seem straightforward. However, certain phenomena associated with the signals from these sensors, such as noise, frequency, and the type of filtering, among others, can affect the clear detection of this time. Therefore, different approaches have been proposed to define these seismic wave arrival times in the sensor signals. Some of the most common are the "start-to-start," "maximum peak-to-peak," and "minimum peak-to-peak" approaches [8, 9]. Graphically, these

approaches are shown in Figure 1. As can be seen, these approaches are based on identifying characteristic points in the signal from both the sensor where the seismic excitation originates and the receiving sensor. These points represent the start of the excitation, as well as the first maximum and minimum peaks of both signals. While in synthetic signals, such as the one shown in Figure 1, the time differentials for any approach are very similar, in real tests, notable differences between these times can be observed.

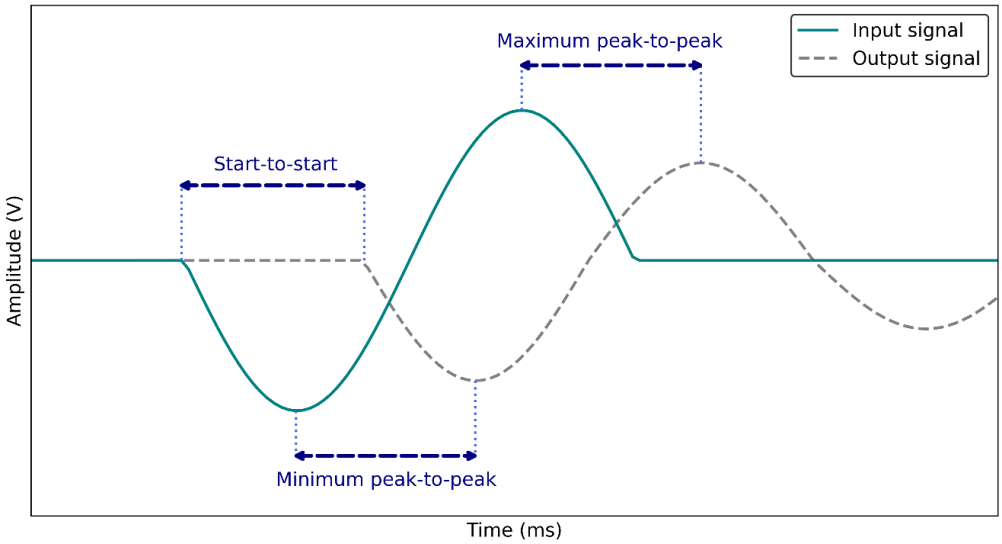



Figura 1: Determination of seismic wave arrival times (based on Fernández-Lavín et al. [8]).

In this work, a statistical analysis was performed on the results obtained using three different classic methodologies for determining seismic wave travel time. A laboratory specimen was prepared from granular material, with two accelerometers embedded to detect seismic vibrations. To generate the waves, the specimen was impacted on its surface, similar to field tests. This generated a compressional wave (P-wave) that traveled throughout the specimen. Subsequently, the signals were filtered using a Fourier filter to improve their geometry, and the arrival times were then determined. Furthermore, the Hilbert transform was applied to the filtered signals to propose two additional approaches to determining travel times.

# Seismic protocol configuration

## Test specimen

To obtain the seismic signals, a specimen was prepared as shown schematically in Figure 2. As shown in the figure, two sensors were embedded in a compacted soil specimen within a rigid, cylindrical mold, 34 cm high and 15 cm in diameter. The sensors are 22 cm apart. Their operation and interface are explained in the following section.

The soil of the specimen consisted of well-graded, non-plastic crushed gravel, classified as GW under the Unified Soil Classification System (USCS). It was compacted in several layers, approaching its maximum dry density (2.17 g/cm$^3$). Once compaction finished, a

circular steel plate, 15 cm in diameter and 1.25 cm thick, was placed on top of the specimen. To generate seismic waves, this plate was struck vertically with a 16 oz. rubber mallet. This simulated a downhole test [9] on a laboratory specimen, with the seismic wave first being detected by the upper transducer (Sensor A) and then, after a certain time interval ($\Delta t$), by the lower transducer (Sensor B). It should be noted that, because the impact generated vertical stress within the soil mass, the resulting wave was a P-wave (compressional wave).

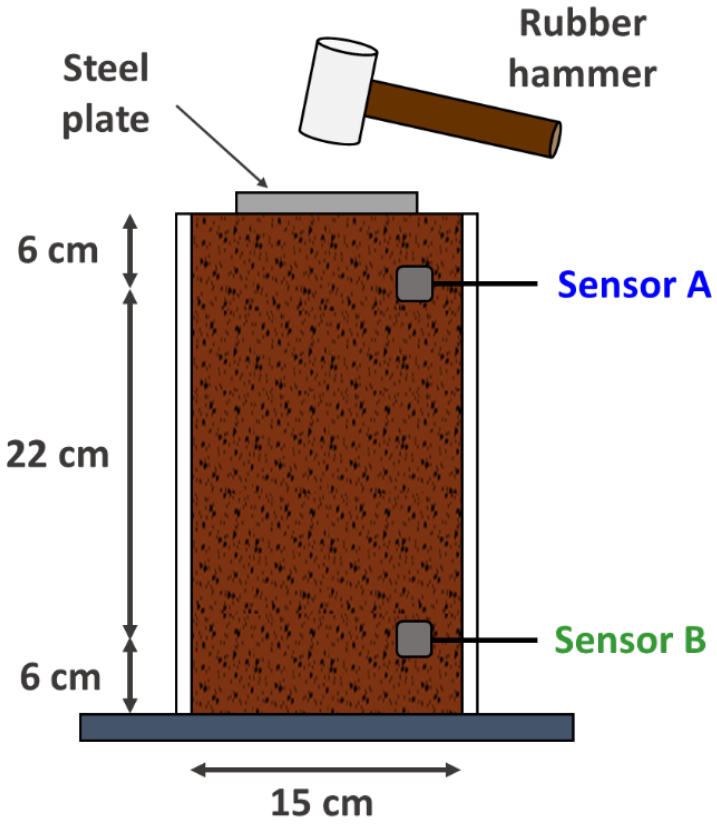



Figura 2: Schematic representation of the test specimen and the seismic protocol developed.

## Seismic transducers

Typically, seismic wave detection uses transducers whose primary purpose is to capture small-scale vibrations or displacements. The most used transducers include geophones, accelerometers, and piezoelectric sensors. In this research, the ADXL335 accelerometer was used. This small sensor (2.1 cm in length and 1.5 cm in height) has been used in various applications that require detecting vibrations or impacts. According to the manufacturer, this sensor is an analog device with three channels for measuring acceleration in the three orthogonal directions, with a measurement range of $\pm 3$ times the acceleration due to gravity. This sensor requires an external power supply in the range of 3.3 to 5V, and since it emits an analog signal, it can be connected to virtually any conventional data acquisition board. However, in seismic tests, whether in the field or in the laboratory, a high sampling rate is usually used due to the high frequency of the generated seismic waves.

In the implemented seismic protocol, a National Instruments NI-USB-6009 data acquisition card was used. This device, which connects to a computer via USB, can detect both digital and analog signals and has a 12-bit resolution. For analog signals, such as those emitted by the accelerometer, a single-ended connection was implemented, allowing the card to operate within a voltage range of $\pm 10$ volts. Therefore, no signal conditioner was required between the accelerometer and the card. Furthermore, since the proposed protocol uses a pair of accelerometers, a sampling rate of 24,000 samples per second was defined, pushing the card to its maximum capacity, as it allows a sampling rate of 48,000 samples per second distributed across its input channels. Likewise, a sampling interval of 2 seconds was set, which is sufficient time to activate the sensors, generate seismic excitation, and detect the resulting signal.

An important point to highlight is that because the sensors were embedded in the soil mass as it was compacted, it was necessary to protect them from damage during compaction. Thus, each sensor was enclosed within a cylindrical steel capsule with an internal diameter of 2.5 cm. These capsules had a base (where the sensor was placed) and a top part that served as a lid for the base, with the two parts threaded together. Once the sensor was encapsulated, it did not suffer any damage during the soil compaction, nor did it affect such process. Figure 3 shows the sensor housed inside the capsule.

Figura 3: ADXL335 accelerometers housed inside the steel capsule.

## Seismic signals

Figure 4a shows an example of the typical signals captured by both sensors during a seismic test, considering the previously set duration of 2 seconds. To observe in more detail the moment the seismic wave reaches the sensors, Figure 4b shows a close-up of that region of the signals. Furthermore, once the signal is recorded, it is common practice to remove the extremes, focusing primarily on the aforementioned region, since this is the area of interest, thereby maximizing the computational resources used to process these signals.

As illustrated in Figure 4b, seismic waves are fleeting, lasting less than 150 milliseconds before fading. Analyzing the signals, a slightly higher initial amplitude is observed in the signal from Sensor A (blue line). This is consistent with the specimen's configuration, since this sensor is closer to the seismic wave's origin and is expected to record the wave with greater intensity. Conversely, once the seismic wave reaches Sensor B, it has already dissipated somewhat as it travels through the soil mass, resulting in a slightly lower signal amplitude

there. Furthermore, both signals tend to follow the same oscillatory motion until the seismic wave finally dissipates completely.

Another point to mention is that for geotechnical seismic tests, where the primary goal is to detect travel time, the signal units can be expressed in voltage without needing to convert them to displacements, velocities, or acceleration. This is because the main objective is to determine the time difference between the arrival times of seismic waves at the sensors, and from this value, obtain the seismic wave velocity. Therefore, such conversion is unnecessary, and this approach is adopted in the present research.

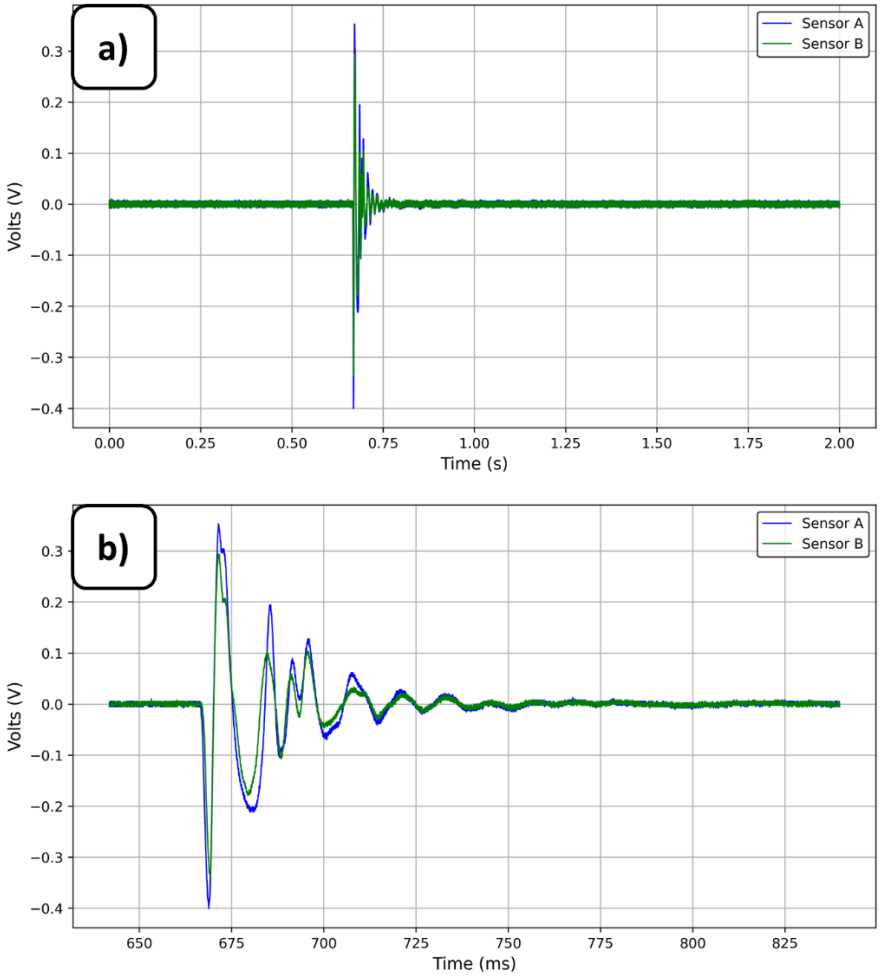



Figura 4: Example of signals from a typical seismic test. a) Original signal acquired; b) Cropping of the extremes of both signals.

# Signal processing and time differentials

## Fourier transform

Although signals like those shown in Figure 4 are considered to have relatively well-defined geometry, seismic wave signal analysis typically employs some processing or filtering techniques to improve signal quality by eliminating noise and achieving greater de-

finition of the signal geometry [10]. In this research, it was observed that removing noise (high-frequency, low-amplitude oscillations) was particularly important for obtaining more accurate and precise seismic wave arrival times.

For this purpose, one of the most widely used techniques in digital signal processing was employed: the Fourier transform. This technique is based on the principle that all signals are sums of different sinusoidal components [11, 12]. Applying this transform, a signal defined in the time domain is expressed in the frequency domain, allowing each frequency and its contribution level to be defined [13]. Once the signal is defined in the frequency domain, a filtering process can be applied, followed by the inverse Fourier transform, so that the signal under analysis is again defined in the time domain [14].

Figure 5 shows the signal from Sensor A, as shown in Figure 4, but now defined in its frequency spectrum, displaying all the signal's components. As can be seen, the highest-contribution frequencies are low. Therefore, if any of these components are eliminated or altered, the geometry of the resulting signal will change considerably from the original. Thus, a low-pass filter was implemented at this point, which attenuates frequencies above a certain threshold, leaving only the low frequencies [15]. Furthermore, since the objective of implementing the Fourier transform was to eliminate noise, which corresponds to high-frequency components, this type of filtering is appropriate for the seismic wave signals observed in this research. As seen in the example in Figure 5, frequencies above 1000 hertz were nullified, which is the most commonly used filtering value for both the signals from Sensor A and Sensor B.

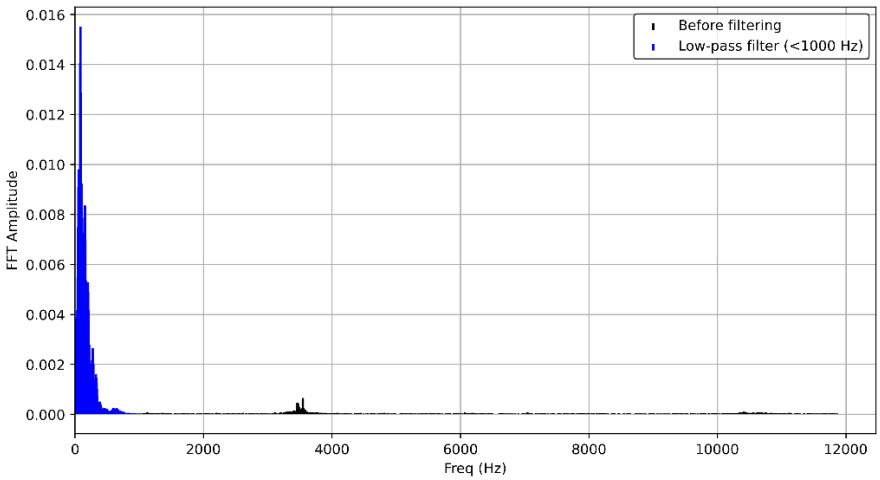



Figura 5: Frequency spectrum of the signal captured by sensor A.

Figure 6 shows the result of the process of removing high frequencies and the subsequent application of the inverse Fourier transform to the signal from Sensor A. As can be seen, the noise shown throughout the signal was noticeably eliminated.

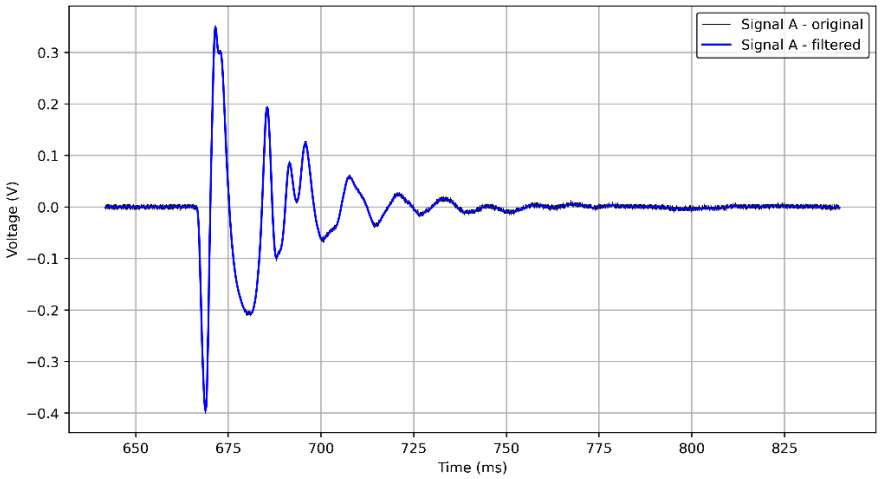



Figura 6: Application of the low-pass filter to signal A.

Finally, Figure 7 shows the filtered signals from both sensors, providing a closer look at the area where the seismic excitation was detected. As shown in this figure, the geometry of both signals improved significantly without drastically altering the signal pattern. This allows for a clearer understanding of the seismic excitation and a better definition of the geometry of the first drop and both negative and positive first peaks, which are essential for this research.

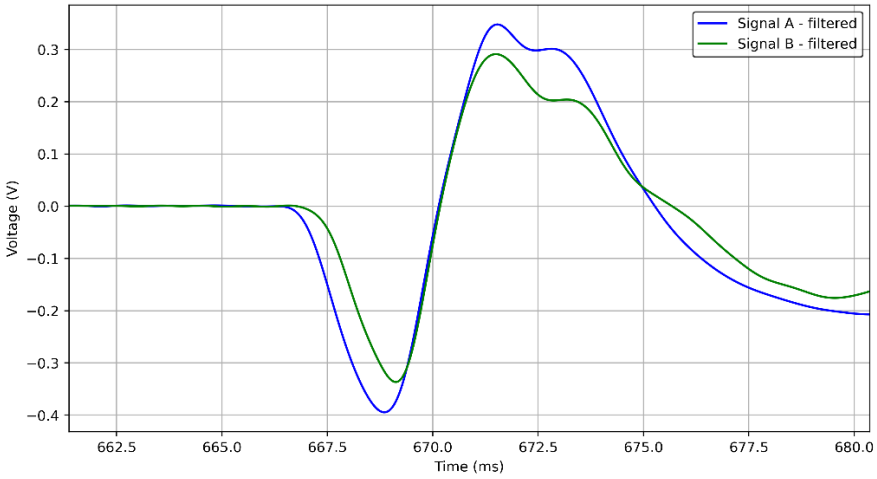



Figura 7: Filtered signals from both sensors.

## Hilbert transform

While the signals obtained from the Fourier transform-based filtering process are considered clean and have a defined geometry, other approaches could be explored to analyze these seismic signals from a different perspective, providing an alternative for defining the time differential required to calculate the seismic wave velocity. Even standardized procedures, such as the Bender element test defined in ASTM D8295 [16], allow other techniques to be used.

Based on this argument, the Hilbert transform was used on seismic signals in this work. This technique has been implemented in other areas, such as radiocommunications engineering, and is also applicable to vibration signals [17]. When the Hilbert transform is applied to a signal, the result is an identical-geometry signal with a 90° phase shift relative to the original; that is, an orthogonal signal. This resulting signal, known as the "analytical signal," is complex, as it contains both real and imaginary components. If the magnitude of this signal (absolute value) is obtained, the result is an envelope of the original signal [17, 18, 19].

Figure 8 shows an example of the application of the Hilbert transform and its envelope. As shown in the figure, where the original signal has a sinusoidal amplitude-modulated geometry, the Hilbert transform generates a phase shift in the signal. Subsequently, by evaluating the magnitude of the analyzed signal, an envelope of the signal is obtained. It should be noted that this envelope only corresponds to the positive portion of the signal, since it was obtained from the absolute value of the analytical signal. However, if the signal is symmetric with respect to the horizontal axis, the lower (negative) envelope can be obtained from the negative value of this envelope. It is also important to point out that if the analyzed signal is in the time domain (as is the case with signals obtained by seismic sensors), the Hilbert transform does not produce a change of domain, and therefore does not affect the spectral components of the signal [18].

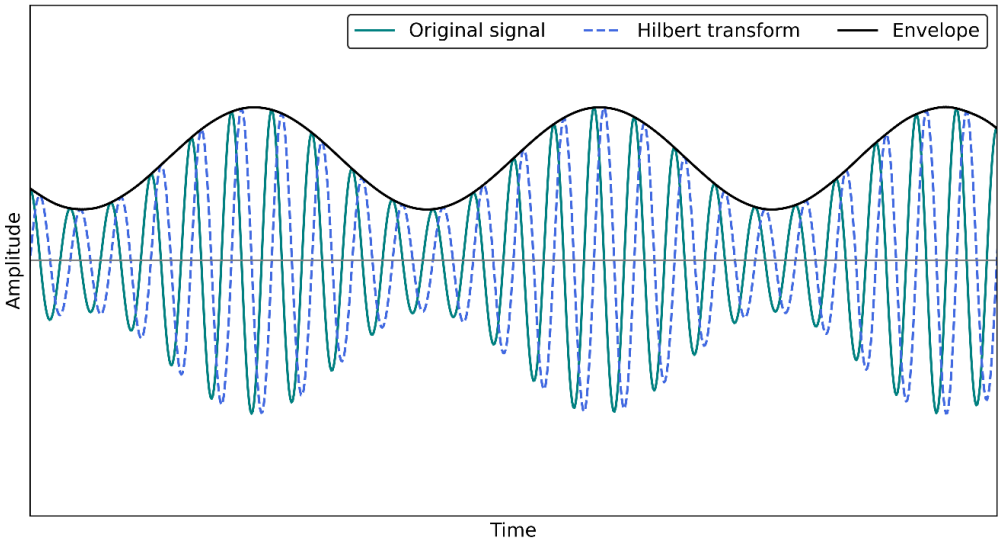



Figura 8: Example of the Hilbert transform in an amplitude-modulated signal and its positive envelope.

For the seismic signals obtained, since they are also symmetric with respect to the time axis, it is sufficient to plot the envelope; therefore, only the positive envelope was considered. Figure 9 shows the implementation of the signal envelope based on the Hilbert transform for the signals obtained after the Fourier filtering process. As shown, the obtained signals exhibit a horizontal separation from each other similar to that observed in the original signals. Furthermore, the maximum peaks are clearly distinguishable in the generated envelopes, which closely coincide with the horizontal midpoint between the minimum and maximum peaks of each signal. Additionally, the amplitude peaks observed in the original signals fo-

llow the same previously defined trend, where the envelope of signal A exhibits a greater amplitude compared to envelope B.

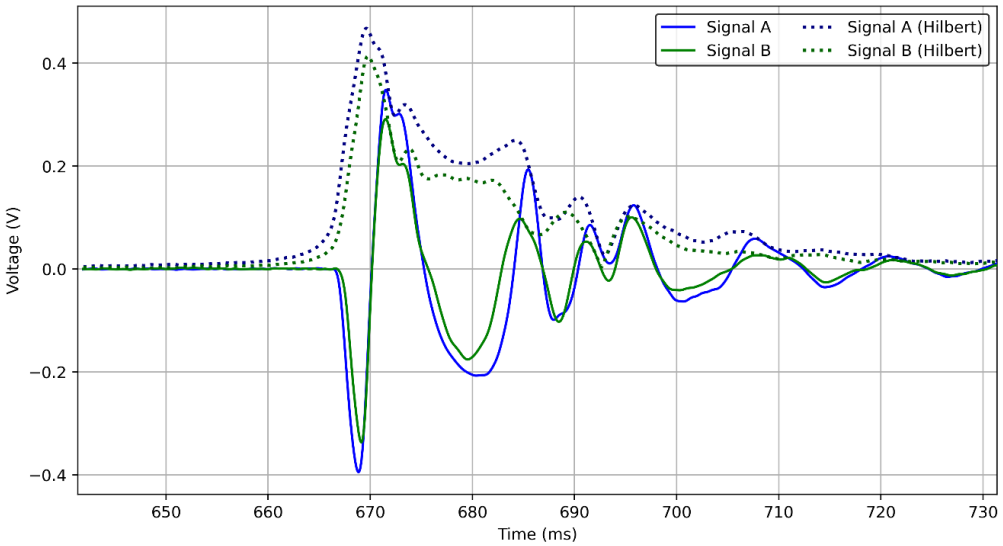



Figura 9: Definition of the envelope through the Hilbert transform for seismic signals.

## Time differentials

After applying Fourier filtering and the Hilbert transform to the signals, the seismic wave arrival times were determined for each signal from the two sensors. For this, the traditional approaches to filtering the signals were adopted: “start-to-start,” “maximum peak-to-peak,” and “minimum peak-to-peak.” Likewise, two similar approaches were proposed for the envelopes obtained by the Hilbert transform: “peak-to-peak” and “initial phase-shift.” The initial step involves identifying the time at which each signal’s envelope reaches its maximum peak. On the other hand, the “initial phase-shift” approach consists of recording the times at which the first slope of each signal begins.

Figure 10 shows the application of the first three classical approaches, while Figure 11 shows the two proposed approaches applied to the resulting envelopes. In all cases, the arrival time at each sensor is recorded in milliseconds, considering that the time differential ($\Delta t$) sought in seismic tests corresponds to the difference between the time recorded by sensor B and the time recorded by sensor A.

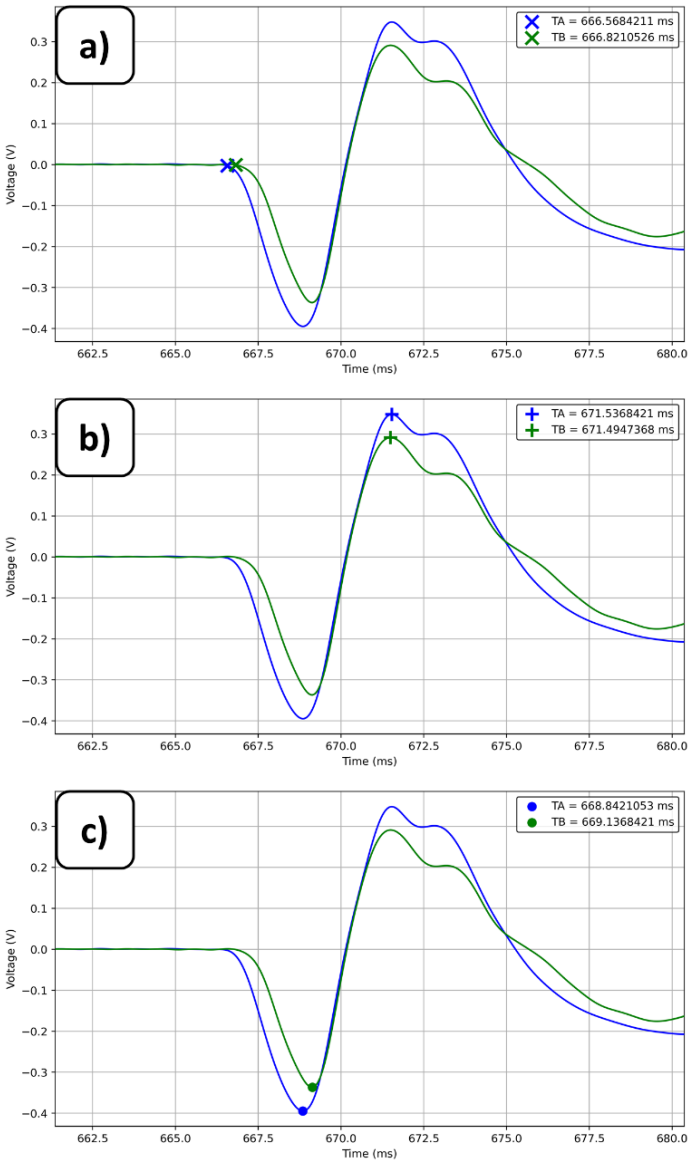



Figura 10: Definition of arrival times by conventional approaches. a) Start-to-start; b) Maximum peak-to-peak; c) Minimum peak-to-peak.

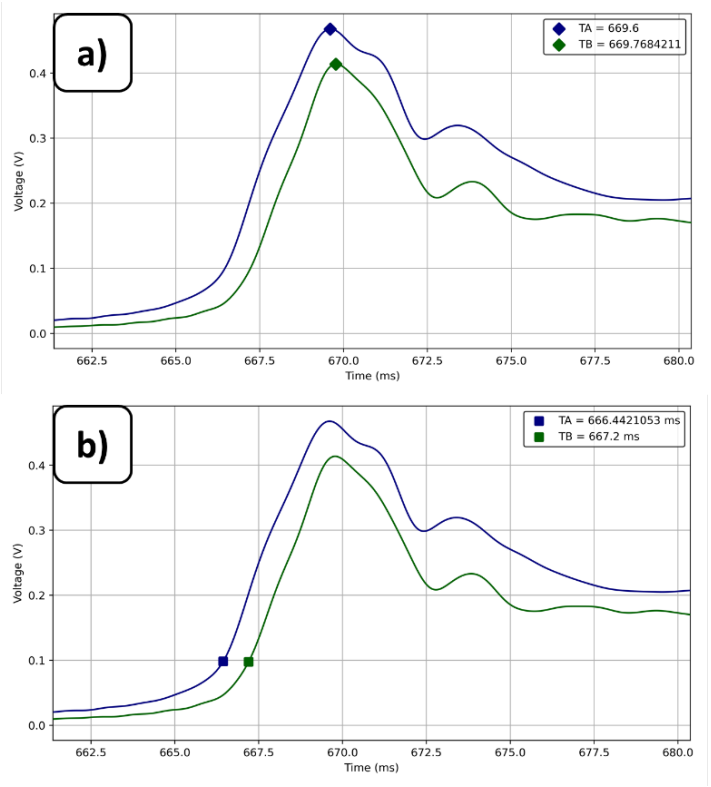



Figura 11: Definition of arrival times using the envelope derived from the Hilbert transform. a) Peak-to-peak; b) Phase-shift.

# Analysis of results

A total of 30 seismic determinations were performed, evaluating $\Delta t$ using the previously mentioned approaches. Table 1 presents the basic descriptive statistics for these data, including measures of central tendency (mean, median, and mode) and common measures of dispersion to assess variability in each data set (range, standard deviation, and coefficient of variation).

Analyzing the measures of central tendency, it is observed that the maximum peak-to-peak approach yields the smallest values for the mean, median, and mode, whereas the maximum peak-to-peak approach using the Hilbert transform yields the highest values, with the other approaches yielding intermediate values. The variation in these values demonstrates that quantifying the time differential is not as trivial as some seismic standards and protocols suggest, and that different techniques can yield different seismic wave velocities for the same material.

However, these measures of central tendency alone do not capture the consistency and variability of each approach; these are quantitatively assessed using measures of dispersion. As shown in Table 1, the approach with the least dispersion, in terms of both standard deviation and coefficient of variation, is “start-to-start,” while “maximum peak-to-peak” has a coefficient of variation of 335.77 %, indicating high variability and dispersion in the data evaluated with this approach. Furthermore, as observed, this approach yielded negative values less than -2 milliseconds. This explains why the lowest values were obtained in the measures

of central tendency for this dataset. The practical implication of these negative values is that the seismic wave was apparently first captured by the lower sensor and then by the upper sensor, thus resulting in a negative $\Delta t$. This condition is observed in Figure 10b, where the time recorded by sensor B is inferior to that recorded by sensor A. This scenario is far from reality, since sensor B is located farther from the origin of the seismic wave than sensor A; therefore, it is inconsistent to have a negative $\Delta t$ for this seismic protocol. Due to its high variability, this approach is discarded for further analysis. The other approaches mostly show a moderate coefficient of variation.

Tabla 1: Measures of central tendency and dispersion of the $\Delta t$ results of the analyzed approaches.

| | Start-to-start | Maximum peak-to-peak | Minimum peak-to-peak | Peak-to-peak (Hilbert) | Phase shift (Hilbert) |
|---|---|---|---|---|---|
| Mean | 0.460012 | 0.173533 | 0.389111 | 0.210381 | 0.697290 |
| Median | 0.484210 | 0.021053 | 0.421053 | 0.210526 | 0.631579 |
| Mode | 0.547368 | 0.000000 | 0.421053 | 0.294737 | 0.926316 |
| Maximum | 0.589474 | 0.757895 | 0.631579 | 0.421053 | 0.968421 |
| Minimum | 0.168421 | -2.125590 | 0.084211 | 0.000000 | 0.294737 |
| Range | 0.421053 | 2.883485 | 0.547368 | 0.421053 | 0.673684 |
| Standard deviation | 0.099483 | 0.582683 | 0.132530 | 0.113307 | 0.178037 |
| CV ( %) | 21.626 | 335.776 | 34.060 | 53.858 | 25.533 |

Figure 12 shows the frequency histograms of the time differentials obtained for each approach, illustrating how the data are distributed within each set. It should be noted that the Sturges criterion, which is based on the number of observations, was used to define the intervals for all histograms. In this case, across the 30 seismic measurements, all histograms are divided into 6 intervals.

For the start-to-start approach (Figure 12a), most results are concentrated in the upper end of the range (close to 0.6 ms), while smaller values (0.2–0.3 ms) are less frequent and exhibit a marked left skew. For the minimum peak-to-peak and Hilbert transform peak-to-peak approaches, the time differentials are mainly concentrated in intermediate values, while the extremes are less frequent, although the former has a more pronounced peak toward the arithmetic mean. Finally, the phase-shift approach using the Hilbert transform also exhibits dominant central values, with some occurrences at the upper end, and a wider range of values compared to the other datasets.

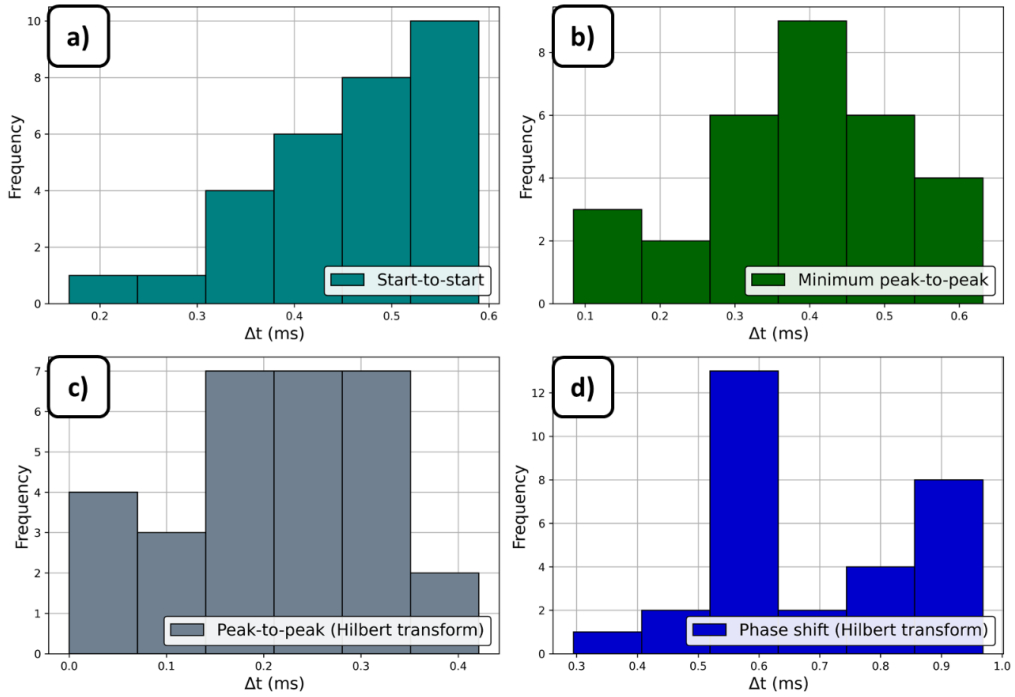



Figura 12: Frequency histograms. a) Start-to-start; b) Minimum peak-to-peak; c) Peak-to-peak (Hilbert transform); d) Phase-shift (Hilbert transform).

Finally, to compare the approaches, Figure 13 shows a box plot. In these plots, the length of the box represents the interquartile range from Q1 to Q3 (or between 25 % and 75 % of the data), while the lines extending from the box (called whiskers) represent the dispersion of the data beyond the quartiles. The center line within the box represents the median of each data set.

This analysis shows that the traditional "Start-to-start" approach yields the most stable and consistent results, exhibiting a low interquartile range and with the mean located roughly in the middle of the box. Furthermore, the whiskers are shorter than those in the other methods. The other classic "minimum peak-to-peak" approach presents a box with a median similar to the "start-to-start" approach, but it shows a larger interquartile range and, while its whiskers are larger, are somewhat symmetrical at both ends. Similarly, the "peak-to-peak Hilbert transform" approach has a low interquartile range (even lower than the "start-to-start" approach), although it can reach very low $\Delta t$ values, even down to 0, as shown by its range of values in Table 1. Finally, the graph of the phase-shift approach using the Hilbert transform shows the largest box, with a median much higher than the others and a very large range, indicating greater dispersion of results.

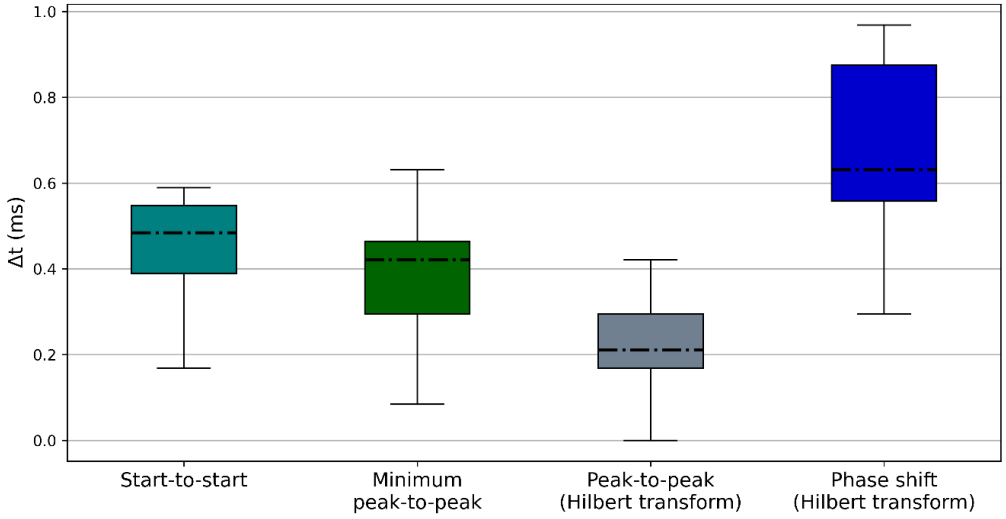



Figura 13: Box plots of the obtained $\Delta t$.

# Conclusions

The objective of this work was to statistically analyze time-differential results obtained from traditional approaches used in seismic geophysical tests, both in the field and in the laboratory. Additionally, the Hilbert transform was applied to evaluate the sensor-captured signals from a different perspective.

Some of the most relevant conclusions obtained from this work are presented below.

- The ADXL335 sensors are a low-cost and easy-to-implement alternative for seismic wave detection. However, these sensors must be properly encapsulated to prevent damage during soil compaction or from the humid environment in which they are placed. Furthermore, although these sensors were used in this study on a specimen under more controlled conditions, these accelerometers may also have applications in field tests, providing an alternative to other, more expensive transducers.

- Arrival times depend heavily on the shape of the signals. Therefore, factors external to the seismic test, such as the type of seismic wave or the sensor used, can produce waves with slightly varying geometry. Furthermore, these signals sometimes undergo additional mathematical processing, such as computing derivatives or integrals, especially when the sensors measure velocity or acceleration. Consequently, while some approaches did not yield highly accurate results with the proposed seismic protocol, they may be applicable when using other types of transducers.

- A key improvement to the seismic signal geometry was the implementation of the Fourier transform and a subsequent low-pass filter in the frequency domain. This approach significantly reduced signal noise, thereby decreasing the uncertainty in the arrival time estimates. Therefore, this technique is highly recommended for seismic testing.

- Likewise, the Hilbert transform has potential applications in seismic studies. Although the approaches based on this transform were not the most consistent in this research, it is possible that this transform could provide an alternative for signal analysis in other seismic tests where other approaches exhibit significant variations.
- Finally, it is worth noting that the seismic protocol executed using the "start-to-start" approach yields the most reliable results with the least dispersion, making it the recommended approach for this type of test. If complications arise in implementing this approach, the next recommended analysis is the "minimum peak-to-peak" method, and lastly, approaches based on the Hilbert transform.

## Direcciones de correo de los autores

Contreras-Ferreyra, D. U.*: Facultad de Ingeniería Civil, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, México
dante.contreras@umich.mx

Pastor-Gómez, N.: Facultad de Ingeniería Civil, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, México
npastor@umich.mx

Chávez-Negrete, C.: Facultad de Ingeniería Civil, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, México
cachavez@umich.mx

Martínez-Rojas, A.: Facultad de Ingeniería Civil, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, México
alondra.martinez@umich.mx

# Referencias

[1] M. Budhu, *Soil Mechanics and Foundations*. Hoboken, NJ: John Wiley & Sons, 2008.

[2] R. E. Hunt, *Geotechnical Investigation Methods: A Field Guide for Geotechnical Engineers*. Boca Raton, FL: CRC Press, 2006.

[3] J. E. Loehr, A. Lutenegger, B. L. Rosenblad y A. Boeckmann, «Geotechnical Site Characterization: Geotechnical Engineering Circular No. 5 (FHWA-NHI-16-072),» National Highway Institute, Federal Highway Administration, Washington, DC, inf. téc., 2016.

[4] C. Vergniault y J. L. Mari, «Shear velocity measurement in boreholes,» en *Well Seismic Surveying and Acoustic Logging*, J. L. Mari, ed., Les Ulis: EDP Sciences, 2018.

[5] A. Sawangsuriya, «Wave propagation methods for determining stiffness of geomaterials,» *Wave Processes in Classical and New Solids*, vol. 44, 2012.

[6] Y. Gao, X. Zheng, H. Wang y W. Luo, «Effect of wave attenuation on shear wave velocity determination using bender element tests,» *Sensors*, vol. 22, n.º 3, pág. 1263, 2022. DOI: 10.3390/s22031263

[7] J. Kumar y N. S. Shinde, «Interpretation of bender elements results by using the sliding Fourier transform method,» *Canadian Geotechnical Journal*, cgj-2018-0733, 2019. DOI: 10.1139/cgj-2018-0733

[8] A. Fernández-Lavín, C. Chamorro-Zurita y E. Ovando-Shelley, «An alternative method to analyze waveforms from bender element tests in soft clays,» *Geotechnical and Geological Engineering*, vol. 42, n.º 1, págs. 43-60, 2024.

[9] ASTM International, «Standard test methods for downhole seismic testing,» ASTM International, West Conshohocken, PA, inf. téc. D7400/D7400M-19, 2019.

[10] V. Joevivek, N. Chandrasekar y R. Jayangondaperumal, «Evaluation of optimal wavelet filters for seismic wave analysis,» *Himalayan Geology*, vol. 37, n.º 2, págs. 176-189, 2016.

[11] A. V. Oppenheim, A. S. Willsky y S. H. Nawab, *Signals & Systems*, 2nd. Upper Saddle River, NJ: Pearson Educación, 1997.

[12] D. Sundararajan, *A Practical Approach to Signals and Systems*. Singapore: John Wiley & Sons, 2009.

[13] R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd. New York: McGraw-Hill, 2000.

[14] A. Veloni, N. Miridakis y E. Boukouvala, *Digital and Statistical Signal Processing*. Boca Raton, FL: CRC Press, 2018.

[15] L. Tan, *Digital Signal Processing: Fundamentals and Applications*. Burlington, MA: Academic Press, 2007.

[16] ASTM International, «Standard test method for determination of shear wave velocity and initial shear modulus in soil specimens using bender elements,» ASTM International, West Conshohocken, PA, inf. téc. D8295-19, 2019.

[17] H. Luo, X. Fang y B. Ertas, «Hilbert transform and its engineering applications,» *AIAA Journal*, vol. 47, n.º 4, págs. 923-932, 2009.

[18] M. E. Abdulmunem y A. A. Badr, «Hilbert Transform and its Applications: A survey,» *International Journal of Scientific & Engineering Research*, vol. 8, n.º 2, págs. 699-704, 2017.

[19] E. A. Awada, M. H. Alomari y E. Radwan, «Wavelet estimation of THD based on Hilbert transform extraction,» *International Journal of Applied Engineering Research*, vol. 11, n.º 9, págs. 6451-6456, 2016.

**Informaticae Abstracta**. ISSN: 3061-8533
Año 2026, Vol. 4, No. 1, pp. 99-123
JUL | DIC 2025
doi: https://doi.org/10.36677/ia.v4n1.28941

# Sistema de coordenadas SPR: optimización algebraica y análisis CORDIC para control robótico

Salas-García, J.*; Hernández-Servín, J. A.; Vilchis-González, A. H.; Ávila-Vilchis, J. C.

**Información del Artículo**

Recibido: 15 de febrero, 2026
Aceptado: 20 de mayo, 2026

**Palabras clave:** Sistema de coordenadas SPR, CORDIC, control robótico, optimización algebraica, microcontrolador ESP32

**Resumen**

Este artículo presenta el Sistema de Coordenadas de Planos Rotatorios (SPR), una forma alternativa de describir la posición de un punto en el espacio mediante un radio y dos ángulos. Está pensado para robots y sistemas de posicionamiento multieje que necesitan convertir coordenadas miles de veces por segundo, con frecuencia en procesadores pequeños. Se desarrolla la base geométrica del sistema y se comparan sus conversiones (de coordenadas SPR a cartesianas y viceversa) con las de los sistemas cilíndrico y esférico, midiendo tiempo de cómputo y estabilidad numérica sobre diez millones de puntos en dos plataformas: un procesador de escritorio x86_64 y un microcontrolador ESP32. En la conversión inversa, que es la que un controlador ejecuta con mayor frecuencia, SPR resultó el más rápido de los sistemas de dos ángulos en todas las plataformas evaluadas: entre 1.15 y 1.5 veces más veloz que el esférico, según la plataforma y el escenario. Además, su precisión se mantiene constante cerca de los polos, la región donde el sistema esférico pierde exactitud. Para procesadores sin unidad de punto flotante se propone una versión de SPR basada en el algoritmo CORDIC con aritmética entera, validada en el ESP32: es 4.8 veces más rápida que cualquier alternativa de doble precisión en ese chip y su tiempo de ejecución es prácticamente constante (variación menor al 0.2 %), una propiedad útil en sistemas de tiempo real. En conjunto, SPR permite sostener lazos de control por encima de 1 kHz en hardware de bajo costo. El trabajo incluye un simulador 3D interactivo y el código necesario para reproducir todas las mediciones.

## 1. Introducción

En la robótica contemporánea, la precisión y la velocidad de respuesta en tiempo real son requisitos fundamentales para el control de plataformas móviles y manipuladores paralelos. No obstante, un obstáculo persistente es la latencia computacional introducida por las transformaciones de coordenadas complejas en lazos de control de alta frecuencia (superiores a 1 kHz) [1]. Los sistemas convencionales, como el esférico y el cilíndrico, dependen fuertemente de funciones trigonométricas inversas (`atan2`, `acos`), las cuales representan una carga significativa para los procesadores embebidos con recursos limitados y unidades de punto flotante de bajo rendimiento [2].

A esta limitación computacional se suma el problema de las singularidades de representación, como el bloqueo de cardán (*gimbal lock*), que degrada la estabilidad numérica en regiones críticas del espacio de trabajo [3]. Si bien se han propuesto soluciones basadas en cuaterniones o matrices de rotación, estas a menudo incrementan la complejidad algebraica y la redundancia de datos.

Ante este escenario, surge la necesidad de modelos de representación que minimicen el uso de funciones trascendentales costosas, priorizando estructuras algebraicas directas que faciliten la inversión del modelo en tiempo real. El presente trabajo formaliza el sistema de coordenadas SPR y propone una estrategia de optimización dual: por un lado, una profunda refactorización algebraica basada en lógica binaria, y por otro, la fundamentación de una variante algorítmica iterativa basada en CORDIC con ejecución *branchless*. A través de un análisis numérico masivo, se demuestra empíricamente que las formulaciones SPR reducen la

latencia de la transformación inversa frente al modelo esférico en todas las plataformas evaluadas —con una magnitud dependiente de la biblioteca matemática de cada arquitectura—, al mismo tiempo que eliminan las inestabilidades numéricas en zonas de singularidad axial. Por su parte, la variante CORDIC se formula y se valida empíricamente sobre un microcontrolador ESP32, delimitando los regímenes de precisión (doble emulada, simple con FPU y aritmética entera de punto fijo) en los que cada variante resulta óptima.

# 2. Disponibilidad de Código y Datos

Para asegurar la reproducibilidad de los resultados presentados, el código fuente completo se encuentra alojado en el repositorio de GitHub: https://github.com/JavierSalasGarcia/sistemaSPR. El ecosistema de software se compone de tres herramientas esenciales:

1. `benchmark.c`: Motor de pruebas masivas que ejecuta 10 millones de transformaciones para medir latencia y estabilidad numérica (RMSE).
2. `analysis_and_plots.py`: Script de automatización que gestiona la compilación, ejecución del benchmark y generación de las figuras estadísticas presentadas en este artículo.
3. `simulator_spr.py`: Herramienta de visualización interactiva 3D que permite validar cualitativamente el comportamiento del sistema de planos rotatorios.
4. `benchmark_esp32/`: Firmware ESP-IDF que reproduce el benchmark sobre un microcontrolador ESP32 real en tres regímenes de precisión (doble emulada, simple con FPU y entero de punto fijo), junto con los scripts de captura por puerto serie y de generación de las figuras correspondientes (`host/`).

Para la reproducción íntegra de los datos del artículo, basta con ejecutar el comando `python analysis_and_plots.py` dentro del directorio raíz. Este proceso automatiza la compilación del código en lenguaje C nativo, la ejecución de las pruebas bajo escenarios críticos y la generación de las figuras presentadas.

# 3. Fundamentación Teórica y Formulación Matemática

## Ecuaciones Fundamentales

El sistema SPR define la posición de un punto en el espacio tridimensional, denotado como $P(r_1, \theta_1, \theta_2)$, a través de la intersección de tres superficies continuas. El concepto base del sistema se ilustra mediante trayectorias paramétricas sobre una esfera, como se muestra en la Figura 1.

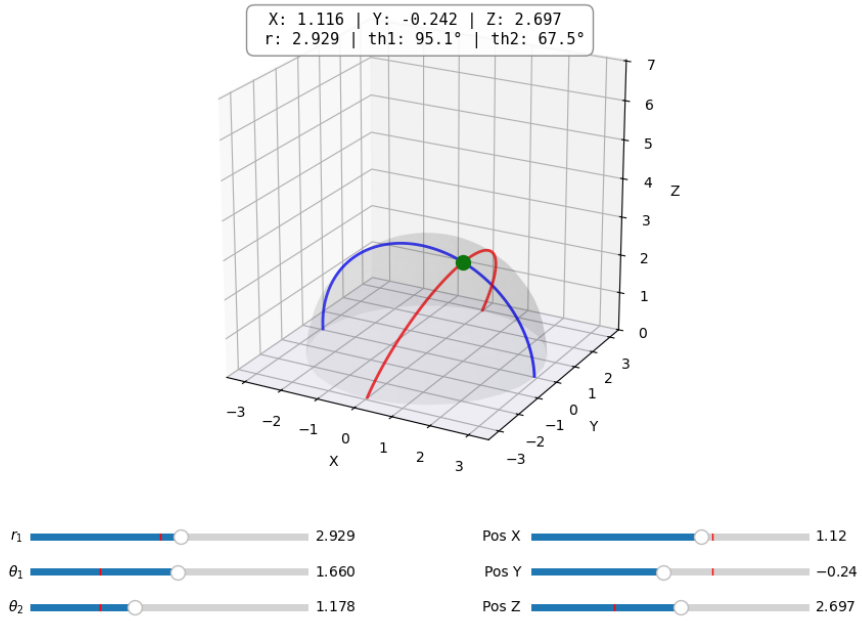



Figura 1: Representación gráfica del sistema SPR obtenida mediante el simulador interactivo desarrollado. Las trazas azules ($\theta_1$ fijo) y rojas ($\theta_2$ fijo) ilustran la ortogonalidad y continuidad del sistema. Esta herramienta, disponible en el repositorio adjunto, permite la validación visual del modelo y la comprobación de la reversibilidad de las ecuaciones.

La primera superficie corresponde a una esfera de radio $r_1$, expresada en coordenadas cartesianas $(x, y, z)$ mediante la Ecuación 1:

$$x^2 + y^2 + z^2 = r_1^2 \tag{1}$$

Las superficies adicionales consisten en planos que se originan en el centro de coordenadas. Estos emplean relaciones tangenciales para determinar la orientación espacial, lo que deriva en las Ecuaciones 2 y 3:

$$z = y \tan(\theta_1) \tag{2}$$

$$z = x \tan(\theta_2) \tag{3}$$

Al resolver este sistema de ecuaciones, se obtiene la "Transformación Directa", la cual permite convertir las coordenadas SPR al sistema cartesiano $(x, y, z)$. Estas relaciones se presentan en la Ecuación 4:

$$x = \frac{r_1 \cos(\theta_2) \sin(\theta_1)}{\sqrt{1 - \cos^2(\theta_1)\cos^2(\theta_2)}} \quad , \quad y = \frac{r_1 \cos(\theta_1) \sin(\theta_2)}{\sqrt{1 - \cos^2(\theta_1)\cos^2(\theta_2)}} \quad , \quad z = \frac{r_1 \sin(\theta_1) \sin(\theta_2)}{\sqrt{1 - \cos^2(\theta_1)\cos^2(\theta_2)}} \tag{4}$$

Para asegurar la validez del método, el radio debe mantener valores no negativos y los rangos angulares deben cubrir la envolvente esférica. Esta configuración garantiza posiciones unívocas y evita discontinuidades presentes en otros sistemas de representación.

### Matriz Jacobiana y Tensor Métrico

La matriz Jacobiana ($J$) se construye a partir de las derivadas parciales de las ecuaciones de transformación. Para diagnosticar la estabilidad geométrica del sistema, se formula su determinante conforme a la Ecuación 5:

$$\det(J) = \frac{r_1^2}{D^3 \sin(\theta_1) \cos^2(\theta_2)} \tag{5}$$

En esta expresión, $D$ representa el divisor polinómico $\sqrt{1 + \tan^2(\theta_2)/\tan^2(\theta_1) + \tan^2(\theta_2)}$, cuya evaluación excluye divisiones por cero dentro de los rangos angulares válidos. Asimismo, es posible derivar el Tensor Métrico ($g_{ij} = J^T J$), componente esencial para cálculos volumétricos y métricas de distancia en aplicaciones avanzadas [4].

### Análisis de Singularidades Geométricas

En los sistemas de coordenadas espaciales, se presentan puntos críticos o singularidades cuando los ángulos alcanzan orientaciones ortogonales (por ejemplo, $\theta_1 = 0$ o $\theta_2 = \pi/2$). El cálculo directo en estos puntos puede generar indeterminaciones numéricas debido a la proximidad del denominador a cero.

No obstante, mediante la aplicación de límites matemáticos, el sistema SPR resuelve estas condiciones de forma natural. Por ejemplo, al evaluar el comportamiento cuando el ángulo $\theta_1$ tiende a cero, la posición converge a $(0, r_1, 0)^T$. Este tratamiento asintótico es fundamental para prevenir fallos computacionales comunes, como el bloqueo de ejes (“Gimbal Lock”) o errores de ejecución, que son frecuentes en sistemas de coordenadas esféricas convencionales [5].

## 4. Formalización y Derivación Analítica Detallada

### Resolución del Sistema de Intersección

La localización exacta de un punto $P$ se determina resolviendo simultáneamente las ecuaciones de la superficie esférica y los planos rotatorios. A partir de las identidades de los planos $z = y\tan(\theta_1)$ y $z = x\tan(\theta_2)$, se establece la relación de proporcionalidad:

$$y = x\frac{\tan(\theta_2)}{\tan(\theta_1)} \tag{6}$$

Al sustituir tanto $y$ como $z = x\tan(\theta_2)$ en la ecuación de la esfera $x^2 + y^2 + z^2 = r_1^2$, se obtiene la expresión factorizada:

$$x^2 \left(1 + \frac{\tan^2(\theta_2)}{\tan^2(\theta_1)} + \tan^2(\theta_2)\right) = r_1^2 \tag{7}$$

De esta derivación emanan las soluciones para las coordenadas cartesianas presentadas anteriormente (Ecuación 4). Este procedimiento asegura que el punto resultante pertenezca con precisión a las tres trayectorias geométricas que definen el sistema.

### Ecuaciones de las Trazas Paramétricas

Las trazas sobre la superficie esférica representan el recorrido de un punto cuando uno de los parámetros angulares permanece constante. Estas curvas son esenciales para la visualización y el mapeo de superficies. Para una traza con $\theta_1$ constante, las ecuaciones paramétricas en función de un parámetro angular $t \in [0, 2\pi]$ se expresan como:

$$x = r_1 \cos(t) \quad , \quad y = r_1 \cos(\theta_1) \sin(t) \quad , \quad z = r_1 \cos(\theta_1) \sin(t) \tan(\theta_1) \tag{8}$$

De manera simétrica, para una traza donde $\theta_2$ es constante, se obtiene:

$$x = r_1 \cos(\theta_2) \sin(t) \quad , \quad y = r_1 \cos(t) \quad , \quad z = r_1 \cos(\theta_2) \sin(t) \tan(\theta_2) \tag{9}$$

Estas trayectorias configuran la retícula característica del sistema SPR, permitiendo un barrido ordenado del espacio esférico.

### Componentes Detalladas del Tensor Métrico

Para un análisis profundo de la geometría diferencial del sistema, se derivan los elementos del tensor métrico $g_{ij} = J^T J$. Al particionar la matriz Jacobiana $J$ mediante derivadas parciales respecto a $r_1, \theta_1, \theta_2$, se identifican componentes clave que dictan la escala local del sistema. El elemento radial se mantiene unitario ($g_{11} = 1$), lo que confirma que el radio $r_1$ conserva su magnitud métrica natural.

Los elementos angulares más representativos, que determinan la longitud infinitesimal de arco $ds^2$, se derivan como:

$$g_{22} = \frac{r_1^2 \tan^4(\theta_2) + r_1^2 \tan^2(\theta_2) D^4 \cos^4(\theta_1) + r_1^2 \tan^6(\theta_2)}{D^6 \sin^4(\theta_1) \cos^4(\theta_1)} \tag{10}$$

$$g_{33} = \frac{r_1^2 (1 + \tan^2(\theta_2)/\tan^2(\theta_1))^2 (1 + \tan^2(\theta_2)) + r_1^2 D^4}{D^6 \cos^4(\theta_2)} \tag{11}$$

Donde $D$ es el divisor polinómico definido previamente. Estas expresiones permiten calcular con precisión elementos de área $dA$ y volumen $dV$ dentro del dominio SPR, facilitando su aplicación en simulaciones de física computacional y cálculo integral espacial.

## 5. Implementación y Optimización Computacional

En esta sección se detallan las estrategias algorítmicas empleadas para transformar las ecuaciones teóricas en rutinas de computación de alto rendimiento, minimizando el costo operativo en transformaciones directas e inversas.

### Refactorización de la Transformación Directa

La formulación original del sistema SPR (Ecuación 2) depende intrínsecamente de la función tangente, lo que introduce una dependencia crítica de funciones inversas y raíces cuadradas anidadas para recuperar las coordenadas $x$ e $y$ a partir de $z$. Tradicionalmente, la

conversión se realizaba mediante el Algoritmo 1, el cual requiere el cálculo de dos tangentes y raíces cuadradas pesadas.

**Algorithm 1:** Transformación Directa Tradicional (basada en tangentes)

**Entrada:** $coord = (r_1, \theta_1, \theta_2)$
**Salida :** $(x, y, z)$
$t_1 \leftarrow \tan(coord.\theta_1)$;
$t_2 \leftarrow \tan(coord.\theta_2)$;
$it1\_2 \leftarrow 1{,}0/(t_1^2 + 10^{-25})$;
$it2\_2 \leftarrow 1{,}0/(t_2^2 + 10^{-25})$;
$z \leftarrow coord.r_1/\sqrt{1{,}0 + it1\_2 + it2\_2}$;
$x \leftarrow z/t_2$;
$y \leftarrow z/t_1$;
**return** $(x, y, z)$

En el Algoritmo 1, las variables $t_1$ y $t_2$ almacenan las tangentes de los parámetros angulares, mientras que $it1\_2$ e $it2\_2$ representan sus inversas al cuadrado, regularizadas con un término constante de $10^{-25}$ para evitar excepciones por división entre cero en caso de ángulos ortogonales.

Como complemento, se derivó una variante puramente algebraica basada en identidades trigonométricas fundamentales ($\sin, \cos$), integrada en el Algoritmo 2. Esta formulación expresa la transformación directa en términos de proyecciones ortogonales y es la que se emplea en el simulador interactivo y en la ruta de cómputo optimizada. Cabe destacar que, en términos de latencia, la versión tradicional basada en tangentes (Algoritmo 1) requiere únicamente dos evaluaciones de $\tan$ frente a las cuatro de $\sin$ / $\cos$ de la versión algebraica, por lo que es la más rápida de las dos formulaciones SPR para la transformación directa, con un costo comparable al del modelo esférico (véase la Figura 4). La formulación sin/cos (Algoritmo 2) es matemáticamente equivalente y resulta conveniente para su uso conjunto con la transformación inversa optimizada (Algoritmo 4).

**Algorithm 2:** Transformación Directa Optimizada (Identidades $\sin$ / $\cos$)

**Entrada:** $coord = (r_1, \theta_1, \theta_2)$
**Salida :** $(x, y, z)$
$s_1 \leftarrow \sin(coord.\theta_1), \quad c_1 \leftarrow \cos(coord.\theta_1)$;
$s_2 \leftarrow \sin(coord.\theta_2), \quad c_2 \leftarrow \cos(coord.\theta_2)$;
$den \leftarrow \sqrt{\max(10^{-20}, 1{,}0 - c_1^2 \cdot c_2^2)}$;
$k \leftarrow coord.r_1/den$;
$x \leftarrow k \cdot c_2 \cdot s_1$;
$y \leftarrow k \cdot c_1 \cdot s_2$;
$z \leftarrow k \cdot s_1 \cdot s_2$;
**return** $(x, y, z)$

Dentro del Algoritmo 2, $s_1, c_1, s_2, c_2$ denotan respectivamente los senos y cosenos de $\theta_1$ y $\theta_2$. La variable auxiliar $den$ representa computacionalmente al divisor polinómico $D$ de la Ecuación 5, incorporando un límite inferior ($\max(10^{-20}, ...)$) para garantizar robustez numérica en las fronteras, mientras que $k$ actúa como un factor de escala radial para ponderar las

proyecciones resultantes.

## Optimización de la Transformación Inversa

Durante la implementación del sistema se observó que la obtención de coordenadas SPR inversas presentaba una latencia elevada debido al uso recurrente de la función arcotangente de dos argumentos (`atan2`), empleada por el procesador para resolver la identidad de los cuadrantes. Con el fin de mitigar esta latencia, se diseñó una refactorización consistente en sustituir `atan2` por divisiones fraccionales simples bajo la función `atan` primitiva, trasladando el control de signos a evaluaciones lógicas binarias.

En el Algoritmo 3 se observa la aproximación convencional que depende de llamadas externas a bibliotecas matemáticas generales, lo cual penaliza el rendimiento en cálculos masivos.

**Algorithm 3:** Transformación Inversa Tradicional

**Entrada:** $(x, y, z)$
**Salida :** $(r_1, \theta_1, \theta_2)$
$r_1 \leftarrow \sqrt{x^2 + y^2 + z^2}$;
$\theta_1 \leftarrow \mathrm{atan2}(z, y)$;
$\theta_2 \leftarrow \mathrm{atan2}(z, x)$;
**return** $(r_1, \theta_1, \theta_2)$

La versión optimizada (Algoritmo 4) delega la responsabilidad de los cuadrantes a comparaciones atómicas directas. Al reducir el número de operaciones trigonométricas complejas y sustituirlas por lógica booleana de bajo nivel, se logra una reducción drástica de los ciclos de reloj, permitiendo que el sistema escale de manera eficiente en hardware con recursos limitados.

**Algorithm 4:** Transformación Inversa Optimizada (Refactorización Lógica)

**Entrada:** $(x, y, z)$
**Salida :** $(r_1, \theta_1, \theta_2)$
$r_1 \leftarrow \sqrt{x^2 + y^2 + z^2}$;
**if** $|y| < 10^{-10}$ **then**
  **if** $z > 0$ **then**
    $\theta_1 \leftarrow \pi/2$;
  **else**
    $\theta_1 \leftarrow -\pi/2$;
**else**
  $\theta_1 \leftarrow \arctan(z/y)$;
  **if** $y < 0$ **then**
    $\theta_1 \leftarrow \theta_1 + \pi$;
**if** $|x| < 10^{-10}$ **then**
  **if** $z > 0$ **then**
    $\theta_2 \leftarrow \pi/2$;
  **else**
    $\theta_2 \leftarrow -\pi/2$;
**else**
  $\theta_2 \leftarrow \arctan(z/x)$;
  **if** $x < 0$ **then**
    $\theta_2 \leftarrow \theta_2 + \pi$;
**return** $(r_1, \theta_1, \theta_2)$

Un aspecto sutil del Algoritmo 4 es que la corrección de cuadrante ($\theta \leftarrow \theta + \pi$ cuando la coordenada del denominador es negativa) debe residir *exclusivamente* dentro de la rama de la arcotangente. Si dicha corrección se aplicara de forma incondicional tras ambas ramas —como ocurría en una formulación preliminar de este trabajo—, la combinación $|y| < \varepsilon$ con $y < 0$ produciría $\theta_1 = \pi/2 + \pi = 3\pi/2$, y la transformación directa reconstruiría el punto con el signo de $z$ invertido, dado que $\sin(3\pi/2) = -1$. Esta condición de frontera es indetectable con perturbaciones de prueba estrictamente positivas, y fue identificada durante la validación sobre hardware embebido (Sección 7) al someter el algoritmo a perturbaciones simétricas $\pm 10^{-11}$ en régimen recursivo: el RMSE acumulado pasó de $1{,}7 \times 10^{2}$ (divergencia por inversión de signo) a $5{,}7 \times 10^{-12}$ una vez confinada la corrección a la rama del arcotangente, tal como se presenta aquí.

## Implementación con CORDIC

El algoritmo CORDIC (*Coordinate Rotation Digital Computer*), introducido por Volder en 1959 [6], es un esquema iterativo que calcula funciones trigonométricas —y sus inversas— mediante únicamente sumas, restas y desplazamientos de bits. Su atractivo fundamental para sistemas embebidos y FPGAs reside en que elimina la necesidad de multiplicadores de

punto flotante y de tablas de biblioteca matemática extensas [7]. En su formulación circular, opera en dos modos:

- **Modo vectorización**: dado un vector cartesiano inicial $(x_0, y_0)$, el algoritmo lo rota iterativamente hasta que su componente $y$ se aproxime a cero ($y \to 0$), acumulando el ángulo recorrido en una variable de fase $z$. Al converger tras un número total de $N$ iteraciones, el sistema alcanza un estado final $(x_N, y_N, z_N)$, donde la variable $z_N$ aproxima al ángulo de entrada: $z_N \approx \arctan(y_0/x_0)$.
- **Modo rotación**: dado un ángulo objetivo $\theta$, el algoritmo rota un vector inicial configurado como $(x_0, y_0, z_0) = (K_N, 0, \theta)$ paso a paso hasta agotar dicho ángulo ($z \to 0$). Al converger tras $N$ iteraciones, las componentes de estado final resultan en las proyecciones trigonométricas: $x_N \approx \cos\theta$ e $y_N \approx \sin\theta$.

Las iteraciones elementales se definen de manera unificada como:

$$x_{i+1} = x_i - \sigma_i \cdot 2^{-i}\, y_i, \quad y_{i+1} = y_i + \sigma_i \cdot 2^{-i}\, x_i, \quad z_{i+1} = z_i - \sigma_i \cdot \alpha_i \tag{12}$$

donde $i$ representa el índice de la iteración en curso ($0 \leq i < N$), $x_i, y_i, z_i$ son los valores del vector de estado en dicho paso, y $\alpha_i = \arctan(2^{-i})$ es el ángulo precalculado de la $i$-ésima pseudo-rotación. La dirección de rotación $\sigma_i$ se evalúa lógicamente como $\sigma_i = -\operatorname{sgn}(y_i)$ en modo vectorización o $\sigma_i = \operatorname{sgn}(z_i)$ en modo rotación. Debido a que las rotaciones se ejecutan sumando tangentes sin normalizar, el vector sufre un alargamiento geométrico en cada paso. Para compensarlo, la magnitud inicial se ajusta multiplicándola por un factor de ganancia invariante (*CORDIC gain*) denotado por $K_N$, el cual se define estrictamente como:

$$K_N = \prod_{i=0}^{N-1} \frac{1}{\sqrt{1 + 2^{-2i}}} \approx 0{,}6073 \quad (N = 24) \tag{13}$$

**Tablas de constantes (datos en ROM / flash)**

CORDIC requiere únicamente dos tablas de $N$ entradas que residen permanentemente en la memoria no volátil del microcontrolador. Su cálculo ocurre una sola vez —durante el design-time o el arranque del sistema— y **no se contabiliza en los tiempos de benchmark**; el algoritmo de transformación accede a ellas como datos de sólo lectura sin ningún cálculo adicional:

Tabla 1: Constantes CORDIC pre-calculadas ($N = 24$ iteraciones).

| $i$ | $2^{-i}$ (`pow2_table`) | $\alpha_i = \arctan(2^{-i})$ (rad) | $\alpha_i$ (°) |
|---|---|---|---|
| 0 | 1.000000000 | 0.78539816 | 45.0000 |
| 1 | 0.500000000 | 0.46364761 | 26.5651 |
| 2 | 0.250000000 | 0.24497866 | 14.0362 |
| 3 | 0.125000000 | 0.12435499 | 7.1250 |
| 4 | 0.062500000 | 0.06241881 | 3.5763 |
| 5 | 0.031250000 | 0.03123983 | 1.7899 |
| 6 | 0.015625000 | 0.01562373 | 0.8952 |
| 7 | 0.007812500 | 0.00781234 | 0.4476 |
| 8 | 0.003906250 | 0.00390623 | 0.2238 |
| 9 | 0.001953125 | 0.00195312 | 0.1119 |
| 10 | 0.000976563 | 0.00097656 | 0.0560 |
| 11 | 0.000488281 | 0.00048828 | 0.0280 |
| 12–23 | $\leq 2{,}44 \times 10^{-4}$ | $\leq 2{,}44 \times 10^{-4}$ | $\leq 0{,}014$ |
| Factor de escala: $K_{24} = 0{,}60725293500888$ | | | |

**Reducción al primer cuadrante**

La ventana de convergencia del bucle CORDIC cubre $\pm 90°$ respecto al estado inicial. Para cubrir la circunferencia completa se adopta la estrategia de *reducir al primer cuadrante* antes del bucle y restaurar el cuadrante correcto mediante una única comparación de signos al final, sin cálculos adicionales.

- **Vectorización** ($\mathrm{atan2}$): se trabaja con $(|x|, |y|)$; el bucle produce $\hat{\alpha} \in [0, \pi/2]$ y el cuadrante del resultado se determina por las siguientes reglas:

$$\mathrm{atan2}(y, x) = \begin{cases} \hat{\alpha} & \text{si } x \geq 0,\ y \geq 0 \quad (Q_1) \\ \pi - \hat{\alpha} & \text{si } x < 0,\ y \geq 0 \quad (Q_2) \\ \hat{\alpha} - \pi & \text{si } x < 0,\ y < 0 \quad (Q_3) \\ -\hat{\alpha} & \text{si } x \geq 0,\ y < 0 \quad (Q_4) \end{cases} \tag{14}$$

  donde las etiquetas $Q_1, Q_2, Q_3$ y $Q_4$ denotan explícitamente el primer, segundo, tercer y cuarto cuadrante del plano bidimensional cartesiano, respectivamente.

- **Rotación** ($\sin$ / $\cos$): se explota que $\sin$ es función impar y $\cos(\pi - \theta) = -\cos(\theta)$, reduciendo $\theta$ a $[0, \pi/2]$ y ajustando signos al final.

**Algorithm 5:** CORDIC Vectorización — $\mathrm{atan2}(y, x)$ con reducción a Q1

**Entrada:** $(x_0, y_0)$, tablas $\alpha_i$, $p_i = 2^{-i}$ en ROM
**Salida :** $\mathrm{atan2}(y_0, x_0) \in (-\pi, \pi]$
$\hat{x} \leftarrow |x_0|,\ \hat{y} \leftarrow |y_0|,\ \hat{\alpha} \leftarrow 0$ ; `// Reducción a Q1`
**for** $i = 0, 1, \ldots, N-1$ **do**
  $\hat{x}_{\text{prev}} \leftarrow \hat{x}$;
  **if** $\hat{y} > 0$ **then**
    `//` $\sigma_i = -1$`: rotar en sentido horario`
    $\hat{x} \leftarrow \hat{x} + p_i\,\hat{y}$; $\hat{y} \leftarrow \hat{y} - p_i\,\hat{x}_{\text{prev}}$; $\hat{\alpha} \leftarrow \hat{\alpha} + \alpha_i$;
  **else**
    `//` $\sigma_i = +1$`: rotar en sentido antihorario`
    $\hat{x} \leftarrow \hat{x} - p_i\,\hat{y}$; $\hat{y} \leftarrow \hat{y} + p_i\,\hat{x}_{\text{prev}}$; $\hat{\alpha} \leftarrow \hat{\alpha} - \alpha_i$;
**return** $\mathrm{atan2}(y_0, x_0)$ *según Ecuación* (14)

**Algorithm 6:** CORDIC Rotación — $\sin\theta$ y $\cos\theta$ con reducción de rango

**Entrada:** $\theta \in (-\pi, \pi]$, tablas $\alpha_i$, $p_i$ en ROM, $K_N \approx 0{,}6073$
**Salida :** $\cos\theta$, $\sin\theta$
$s_{\text{neg}} \leftarrow (\theta < 0)$; ;
si $s_{\text{neg}}$: $\theta \leftarrow -\theta$ ; `// Paridad:` $\sin(-\theta) = -\sin(\theta)$
$c_{\text{neg}} \leftarrow (\theta > \pi/2)$; ;
si $c_{\text{neg}}$: $\theta \leftarrow \pi - \theta$ ; `// Simetría:` $\cos(\pi - \theta) = -\cos(\theta)$
$(x_0, y_0, z_0) \leftarrow (K_N,\ 0,\ \theta)$;
**for** $i = 0, 1, \ldots, N-1$ **do**
  **if** $z_i > 0$ **then**
    `//` $\sigma_i = +1$`: rotar en sentido antihorario`
    $x_{i+1} \leftarrow x_i - p_i\,y_i$; $y_{i+1} \leftarrow y_i + p_i\,x_i$; $z_{i+1} \leftarrow z_i - \alpha_i$;
  **else**
    `//` $\sigma_i = -1$
    $x_{i+1} \leftarrow x_i + p_i\,y_i$; $y_{i+1} \leftarrow y_i - p_i\,x_i$; $z_{i+1} \leftarrow z_i + \alpha_i$;
**if** $c_{\text{neg}}$ **then**
  $\cos\theta \leftarrow -x_N$;
**else**
  $\cos\theta \leftarrow x_N$;
**if** $s_{\text{neg}}$ **then**
  $\sin\theta \leftarrow -y_N$;
**else**
  $\sin\theta \leftarrow y_N$;

En los seudocódigos anteriores, variables como $p_i$ y $\alpha_i$ corresponden a los accesos directos en memoria de lectura (las tablas `pow2_table` y `atan_table`). En el Algoritmo 5, $\hat{x}$, $\hat{y}$ y $\hat{\alpha}$ fungen como registros de trabajo temporal para la reducción al primer cuadrante, apoyándose en la variable puente $\hat{x}_{\text{prev}}$ para la asignación cruzada no destructiva. En el Algoritmo 6,

las banderas booleanas $s_{\text{neg}}$ y $c_{\text{neg}}$ actúan como selectores lógicos que restablecen los signos trigonométricos de paridad y simetría al finalizar el bucle, tras haber agotado el ángulo $\theta$ en la variable acumuladora $z_i$.

**Ejemplo de aplicación: cómputo de** $\text{atan2}(0{,}6,\ 0{,}8)$

Para ilustrar la mecánica del algoritmo se calcula $\text{atan2}(0{,}6, 0{,}8)$, cuyo valor exacto es $\arctan(0{,}75) = 36{,}870°$. Ambas entradas son positivas (Q1), por lo que se trabaja directamente con $(|x|, |y|) = (0{,}8, 0{,}6)$. La Tabla 2 registra las primeras ocho iteraciones:

Tabla 2: Primeras 8 iteraciones de CORDIC vectorización para $\text{atan2}(0{,}6, 0{,}8)$.

| $i$ | $\alpha_i$ (°) | $x_i$ | $y_i$ | $\text{sgn}(y_i)$ | $\hat{\alpha}$ acum. (°) | Error (°) |
|---|---|---|---|---|---|---|
| — | — | 0.8000 | 0.6000 | $+$ | 0.0000 | 36.870 |
| 0 | 45.0000 | 1.4000 | $-0.2000$ | $+$ | 45.0000 | 8.130 |
| 1 | 26.5651 | 1.5000 | 0.5000 | $-$ | 18.4349 | 18.435 |
| 2 | 14.0362 | 1.6250 | 0.1250 | $+$ | 32.4711 | 4.399 |
| 3 | 7.1250 | 1.6406 | $-0.0781$ | $+$ | 39.5961 | 2.726 |
| 4 | 3.5763 | 1.6455 | 0.0246 | $-$ | 36.0198 | 0.850 |
| 5 | 1.7899 | 1.6463 | $-0.0268$ | $+$ | 37.8097 | 0.940 |
| 6 | 0.8952 | 1.6467 | $-0.0011$ | $-$ | 36.9145 | 0.045 |
| 7 | 0.4476 | 1.6467 | 0.0118 | $-$ | 36.4669 | 0.403 |

Valor exacto: $36{,}870°$. Tras $N = 24$: error $< 6{,}8 \times 10^{-6°}$.

En la Tabla 2, las columnas representan las distintas variables computacionales que evolucionan durante el proceso. Específicamente, $i$ representa el índice de la iteración actual (con $i = 0$ como el paso inicial), $\alpha_i$ (°) es el ángulo de corrección precalculado en grados equivalente a $\arctan(2^{-i})$, y las variables $x_i, y_i$ representan las componentes ortogonales del vector rotado en ese paso específico. La columna $\text{sgn}(y_i)$ muestra el resultado de la función de decisión lógica que evalúa el signo de $y_i$, la cual determina la dirección angular de la siguiente pseudo-rotación para forzar iterativamente la convergencia de la componente $y$ hacia cero. La columna etiquetada como $\hat{\alpha}$ acum. (°) refleja la fase angular total acumulada dentro del bucle de suma, mientras que la métrica de *Error* (°) cuantifica la diferencia absoluta en grados respecto al valor trigonométrico exacto ($36{,}870°$).

El patrón numérico decreciente observado refleja la búsqueda binaria inherente al algoritmo: en cada iteración se añade o sustrae $\alpha_i$ según el signo de $y_i$, de modo que el acumulador oscila alrededor del valor verdadero con amplitud decreciente. La velocidad de convergencia es de 1 bit por iteración, lo que implica que tras $N = 24$ iteraciones se alcanza una precisión de $\sim 10^{-7}$ rad ($6{,}8 \times 10^{-6°}$), un margen de error suficientemente bajo para satisfacer la totalidad de los requisitos de control y cinemática en robótica avanzada.

Aplicando la misma operación a las dos componentes angulares del sistema SPR para un punto cartesiano $(x, y, z)$:

$$\theta_1 = \text{atan2}_{\text{CORDIC}}(z,\ y), \qquad \theta_2 = \text{atan2}_{\text{CORDIC}}(z,\ x) \tag{15}$$

La fase de reducción al primer cuadrante garantiza que el bucle de 24 iteraciones sea siempre suficiente independientemente del cuadrante del punto de entrada.

La variante **SPR CORDIC** (identificador `SPR_C` en el benchmark) sustituye `atan2` de biblioteca en la transformación inversa por el Algoritmo 5, y los cuatro cálculos trigonométricos de la transformación directa por dos invocaciones del Algoritmo 6, obteniendo $(s_1, c_1)$ y $(s_2, c_2)$ simultáneamente a partir de tablas en ROM.

# 6. Metodología de Evaluación Experimental

Diversas investigaciones han documentado que los sistemas basados en retículas esféricas presentan inestabilidades en sus fronteras, lo que provoca una pérdida de precisión decimal y la propagación de errores de coma flotante debido al uso iterativo de series trigonométricas [4], [5].

Para corroborar la eficiencia del sistema propuesto, se desarrollaron algoritmos en arquitecturas nativas (lenguaje C) y entornos interpretados (Python 3). El experimento consistió en someter al sistema SPR, junto con modelos esféricos y cilíndricos, a pruebas con un volumen masivo de coordenadas procesadas mediante transformaciones directas e inversas. Se midió el costo operativo frente a singularidades polares, la retención de precisión decimal mediante la métrica RMSE y el tiempo de ejecución comparativo de cada modelo. Los resultados de escritorio aquí reportados corresponden a una plataforma x86_64 bajo Windows 11, compilando con MSVC 19.x (`/O2`) y la biblioteca matemática UCRT.

**Nota metodológica sobre la instrumentación.** Los lazos de medición acumulan en una variable `volatile` *las tres componentes* de cada resultado. Esta precaución es determinante: se comprobó que, si el sumidero consume una sola componente (por ejemplo, únicamente el radio), los compiladores que tratan las funciones matemáticas como intrínsecos puros eliminan por código muerto las llamadas trigonométricas cuyos resultados no se usan, reduciendo el lazo medido esencialmente a una raíz cuadrada. Este artefacto puede simular ventajas de latencia de más de un orden de magnitud (tiempos del orden de 5 ns por punto, físicamente imposibles para una secuencia de raíz, divisiones y arcotangentes) y afectó a una versión preliminar de este benchmark; todas las cifras del presente artículo provienen de la instrumentación corregida, cuyos tiempos por operación son consistentes con el costo de instrucción documentado de las funciones involucradas.

El estudio se complementó con una campaña de validación sobre hardware embebido real: un microcontrolador ESP32-D0WD-V3 (Xtensa LX6, doble núcleo, 240 MHz medidos) programado con ESP-IDF v5.4.1 y optimización `-O2`. Esta plataforma resulta particularmente adecuada porque dispone de FPU de precisión *simple* en hardware pero emula la precisión *doble* por software (*soft-float*), lo que permite medir sobre un mismo silicio tres regímenes aritméticos: doble emulada, simple con FPU y aritmética entera de punto fijo. La latencia se midió con el contador de ciclos del núcleo (resolución de un ciclo de reloj) tras una pasada previa de calentamiento de la caché de instrucciones; los lotes de $1\,024$ puntos se generan con un PRNG determinista (*xorshift32*, semilla fija), de modo que las once variantes evaluadas procesan exactamente los mismos puntos, y cada medición promedia $16\,384$ transformaciones. El protocolo de estabilidad ($1\,000$ ciclos recursivos ida–vuelta) es idéntico al de PC,

con la salvedad de que las perturbaciones del escenario crítico abarcan el rango simétrico $\pm 10^{-11}$ (en lugar de valores exclusivamente positivos), condición que resultó determinante para exponer la condición de frontera discutida en la Sección 7. La variante entera implementa la transformación completa en punto fijo (posiciones Q11.20, ángulos Q3.28 en radianes, funciones trigonométricas Q1.30) con bucles CORDIC *branchless* mediante negación condicional por máscara de signo y raíz cuadrada entera de 64 bits; los resultados se transmiten por UART como CSV y se capturan con los scripts del directorio `benchmark_esp32/host/`.

# 7. Resultados Numéricos y Discusión

Las siguientes subsecciones presentan el análisis comparativo de los cinco modelos evaluados: Cilíndrico, Esférico, SPR Tradicional (`SPR_T`), SPR Refactorizado (`SPR_O`) y la nueva variante **SPR CORDIC** (`SPR_C`).

## Desempeño en Regiones de Singularidad

La literatura científica advierte sobre la carga computacional requerida al procesar ubicaciones que interceptan los polos cardinales [8]. Para evaluar este aspecto, se procesaron millones de posiciones con una separación mínima respecto al eje normal ($x, y \approx 10^{-13}$).

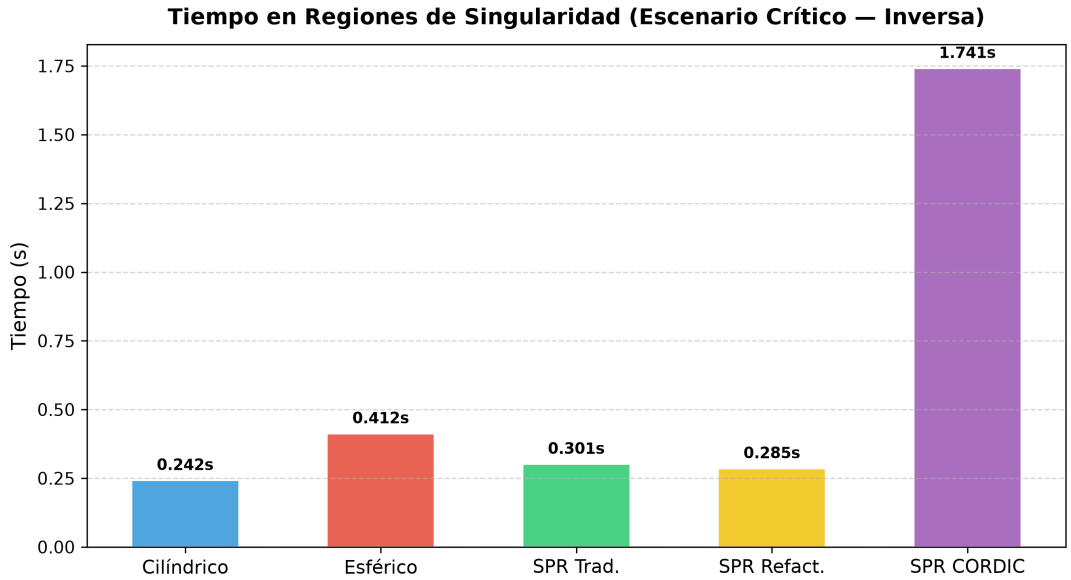



Figura 2: Tiempo de ejecución en la resolución de coordenadas próximas al polo. El incremento en el tiempo refleja la complejidad matemática que representan las fronteras polares para los sistemas convencionales.

Como se ilustra en la Figura 2, los modelos convencionales incrementan su latencia al procesar el entorno polar: el cilíndrico un 65 % y el esférico un 11 % respecto al escenario estándar. En contraste, las variantes SPR se mantienen estables o incluso mejoran: SPR Refact. reduce su tiempo un 25 % en el escenario crítico —sus arcotangentes de razón extrema

convergen por la vía rápida de la biblioteca— y se convierte en el modelo tridimensional más veloz en la vecindad de la singularidad. La variante SPR CORDIC exhibe una latencia esencialmente idéntica en ambos escenarios ($\sim 1$ % de variación), reflejo de su convergencia iterativa de paso fijo, cuya arquitectura de acceso a memoria resulta además más predecible para cachés de instrucción reducidas (relevante en microcontroladores de gama baja).

## Estabilidad Numérica y Degradación Decimal (RMSE)

La Figura 3 revela un comportamiento crítico en el límite del dominio. En el *Escenario Estándar*, el sistema SPR Trad. (Algoritmo 3/1) mantiene una deriva de $\sim 10^{-14}$, comparable a la de los demás modelos. La variante SPR Refact. (Algoritmo 4/2) muestra un RMSE estándar de $\sim 2{,}3\times10^{-12}$ como consecuencia del acumulado de redondeo numérico en la formulación sin/cos con cuatro evaluaciones trigonométricas por paso, aunque esta magnitud sigue siendo despreciable en la práctica de ingeniería. En el *Escenario Crítico* (singularidades axiales, $x, y \approx 10^{-11}$) se expone la fragilidad de los sistemas clásicos.

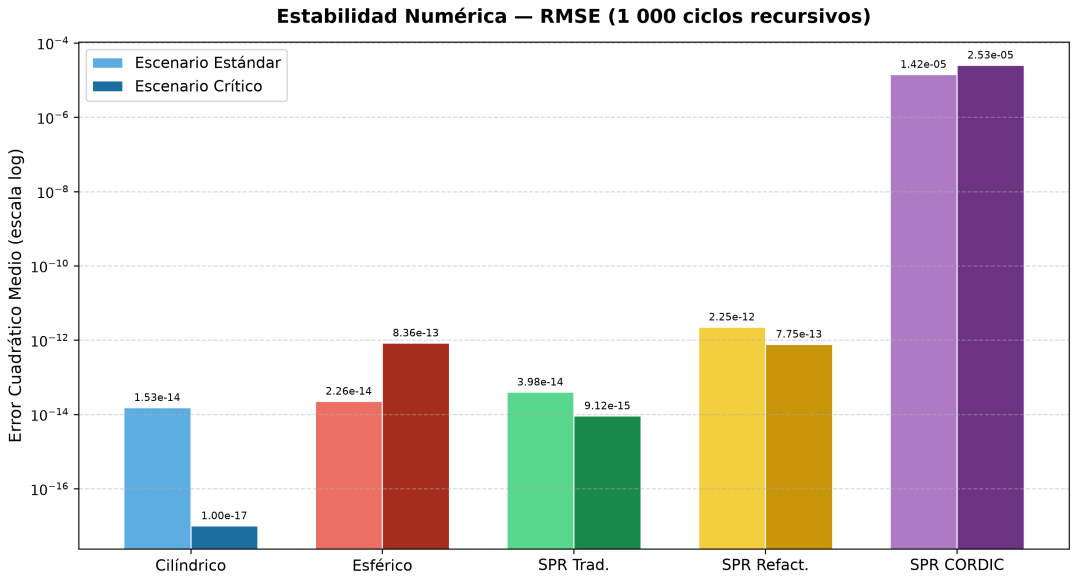



Figura 3: Estabilidad numérica (RMSE) comparando Escenario Estándar vs Escenario Crítico. El sistema SPR mantiene una precisión constante incluso en la proximidad de singularidades axiales.

La degradación de precisión en los modelos esféricos (unas 40 veces: de $\sim 2 \times 10^{-14}$ en el escenario estándar a $\sim 8 \times 10^{-13}$ en el crítico) se origina en la sensibilidad de la función `acos(z/r)`. Cuando un punto se aproxima al polo, ligeras variaciones en $z$ o $r$ producen grandes cambios en el ángulo resultante debido a la pendiente casi infinita de la arcocoseno en sus límites operativos. Por el contrario, el sistema SPR mantiene una precisión constante e independiente de la ubicación del punto, actuando como un filtro natural ante la erosión decimal del estándar IEEE-754.

La variante SPR CORDIC introduce un error de redondeo adicional derivado de la convergencia iterativa: con $N = 24$ iteraciones la precisión alcanzable es de aproximadamente

$2^{-24} \approx 6 \times 10^{-8}$ radianes por ángulo. En régimen acumulativo (1 000 ciclos recursivos), esto se manifiesta como un RMSE ligeramente superior al de las versiones SPR con funciones de biblioteca, pero inferior al de los modelos esféricos en el escenario crítico. Dicho compromiso resulta aceptable en arquitecturas embebidas que privilegian la eliminación de multiplicadores (FPGAs, DSPs de núcleo fijo) frente a la precisión absoluta [7].

## Análisis Crítico de Tiempos y Latencia Computacional

Los resultados temporales tras procesar más de $10^7$ puntos muestran el perfil característico de SPR: una conversión inversa más rápida que la del esférico y una directa más costosa (Figuras 4 y 5). Dado que en un lazo de control la conversión inversa se ejecuta con mayor frecuencia, este perfil favorece a SPR en su dominio de aplicación.

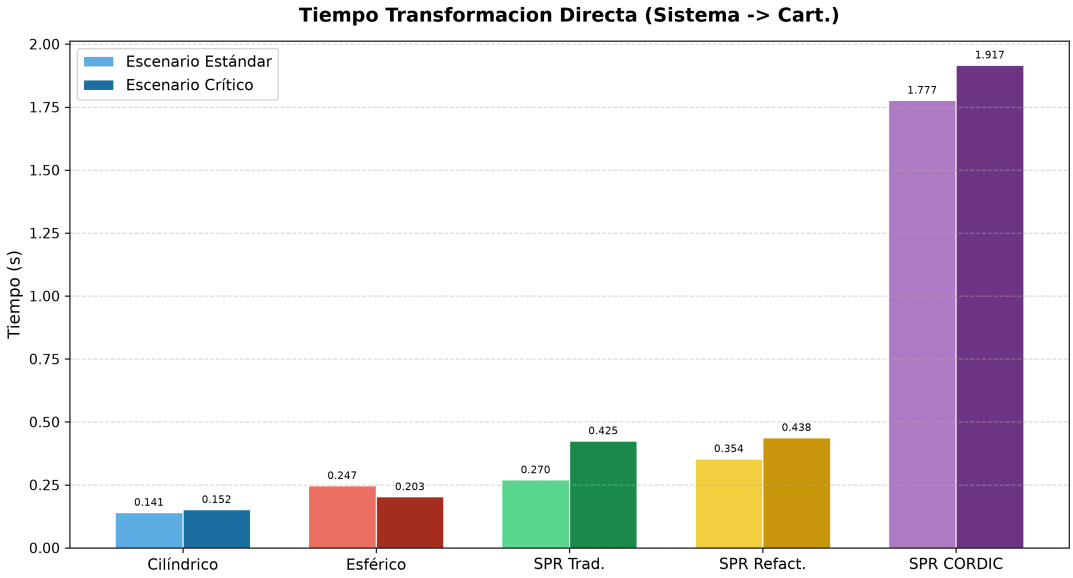



Figura 4: Eficiencia de la Transformación Directa comparando los cinco modelos evaluados (x86_64, MSVC/UCRT). El modelo cilíndrico lidera al requerir una única pareja sin/cos; SPR Trad. (dos tangentes) presenta un costo comparable al del esférico; SPR Refact. paga sus cuatro evaluaciones trigonométricas; SPR CORDIC refleja el costo de iterar 48 pseudo-rotaciones (24 por ángulo) en punto flotante de doble precisión.

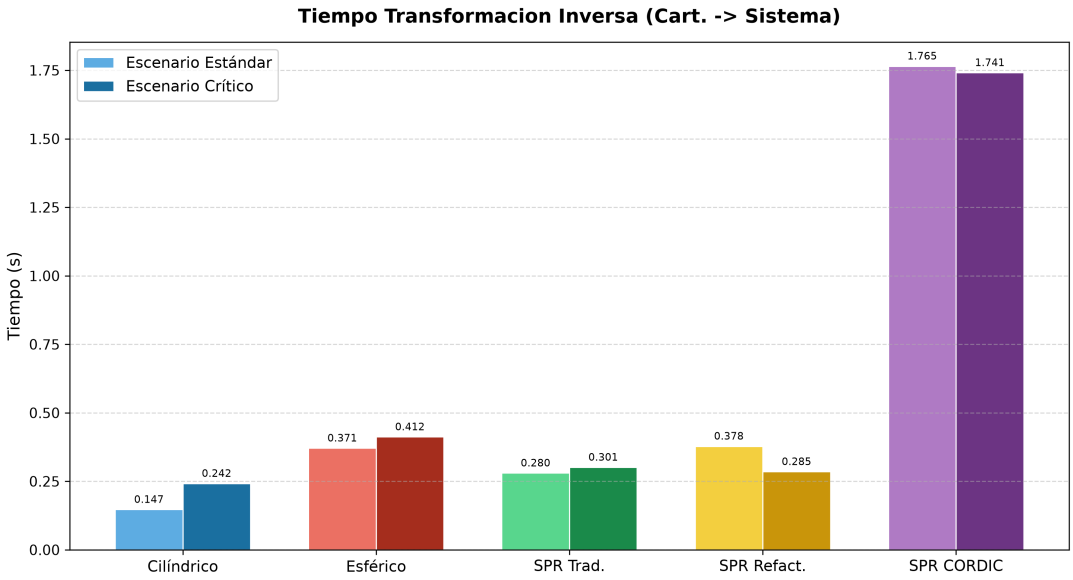



Figura 5: Tiempo de la Transformación Inversa para los cinco modelos (x86_64, MSVC/UCRT). Entre los modelos tridimensionales, SPR Trad. es el más rápido en el escenario estándar y SPR Refact. en el crítico, superando en ambos casos al esférico. El costo de SPR CORDIC en doble precisión anticipa la discusión sobre el rol de la FPU (Sección 7.3.1).

Con la instrumentación corregida, la transformación inversa de SPR supera al modelo esférico en ambos escenarios sobre x86_64: SPR Trad. es $1{,}33$ veces más rápida en el escenario estándar ($0{,}28$ s frente a $0{,}37$ s por $10^7$ puntos) y SPR Refact. $1{,}45$ veces en el crítico ($0{,}29$ s frente a $0{,}41$ s). Matemáticamente, la ventaja proviene de sustituir la evaluación de `acos` sobre la hipotenusa tridimensional por arcotangentes simples y lógica condicional binaria. La magnitud del beneficio depende de la implementación de `atan`/`atan2`/`acos` de cada biblioteca matemática, y se conserva sobre la *newlib* del ESP32 ($1{,}15\times$–$1{,}46\times$; Sección 7). Cabe hacer constar que versiones preliminares de este trabajo reportaban un factor de $15\times$, identificado posteriormente como manifestación del artefacto de eliminación de código muerto descrito en la nota metodológica de la Sección 6.

### El Rol de la FPU y la Paradoja de Latencia de CORDIC

Al observar los resultados del benchmark, resulta contraintuitivo que CORDIC —un algoritmo históricamente célebre por su velocidad en hardware— registre un tiempo de ejecución notablemente superior ($\sim 1{,}8$ s) frente a las formulaciones algebraicas de SPR ($0{,}28$–$0{,}38$ s). Para comprender esta aparente asimetría, es indispensable entender el impacto de la Unidad de Punto Flotante (FPU, por sus siglas en inglés).

La FPU es un coprocesador de hardware integrado en prácticamente todos los procesadores modernos (como el x86_64 utilizado en el experimento o los núcleos ARM Cortex-A de una Raspberry Pi). Su propósito exclusivo es resolver cálculos matemáticos pesados (multiplicaciones, divisiones y raíces de precisión doble) a nivel de transistores, empleando paralelismo masivo. Cuando se ejecuta el sistema SPR Algebraico en estas arquitecturas, la FPU absorbe la carga trigonométrica en silicio puro y responde en pocos ciclos de reloj. Bajo

este escenario, CORDIC pierde ventaja: el procesador moderno se ve forzado a iterar secuencialmente un bucle de software de 24 pasos (desperdiciando la potencia de su FPU), siendo inevitablemente más lento. Por lo tanto, **en hardware con FPU, el enfoque SPR Refactorizado (Algebraico) es inmensamente superior**.

El verdadero valor de la variante CORDIC emerge cuando se traslada la cinemática a plataformas embebidas restrictivas para robótica autónoma o industrial:

1. **Microcontroladores sin FPU (o de precisión mixta)**: En arquitecturas restrictivas (como ARM Cortex-M0, familias Arduino ATmega, o plataformas de uso masivo como el ESP32 operando forzosamente en doble precisión), no existe circuitería matemática para realizar cálculos de 64 bits de forma nativa. Si se solicita trigonometría estándar, el microcontrolador debe emular matemáticamente cada división mediante una pesada librería de software (*soft-float*), demorando miles de ciclos de reloj. Aquí, el enfoque SPR CORDIC (implementado en aritmética entera) brilla al prescindir de la FPU y limitarse al uso de la Unidad Aritmético Lógica (ALU) estándar para realizar sumas y desplazamientos de bits a máxima velocidad. La Sección 7 cuantifica este efecto sobre un ESP32 físico y confirma un matiz esencial: la ventaja exige la implementación *entera* del algoritmo, pues ejecutar las iteraciones CORDIC sobre aritmética `double` emulada resulta contraproducente.

2. **FPGAs y Hardware Analítico**: En Arreglos de Compuertas (FPGA), implementar el circuito de una FPU consume cantidades desorbitadas de área lógica (*LUTs*). CORDIC, en cambio, se sintetiza naturalmente como un circuito en cascada (*pipeline*) de bajo coste espacial, capaz de entregar coordenadas trigonométricas puras en fracciones de microsegundo [6], [7].

Para maximizar su eficiencia general, la implementación embebida del CORDIC se refactorizó para ser estructuralmente *branchless* (sin saltos condicionales): la dirección de cada pseudo-rotación se resuelve mediante negación condicional por máscara de signo, eliminando las instrucciones `if-else` del bucle iterativo y asegurando un flujo lineal sin penalizaciones por fallos en el predictor de saltos del procesador (*branch misprediction*). Fundamentalmente, esta estructura ejecuta la misma secuencia de instrucciones con independencia de los datos, lo que garantiza un tiempo de ciclo estricto y determinista para Sistemas Operativos de Tiempo Real (RTOS), propiedad verificada empíricamente sobre ESP32 (variación menor al $0{,}2$ % entre escenarios; Sección 7).

En términos de precisión, la Tabla 3 resume el error máximo esperado para distintas profundidades de iteración:

Tabla 3: Error de aproximación CORDIC por número de iteraciones.

| **Iteraciones ($N$)** | **Error máx. (rad)** | **Bits de precisión** |
|---|---|---|
| 12 | $\sim 2{,}4 \times 10^{-4}$ | $\sim 12$ |
| 16 | $\sim 1{,}5 \times 10^{-5}$ | $\sim 16$ |
| 20 | $\sim 9{,}5 \times 10^{-7}$ | $\sim 20$ |
| 24 | $\sim 6{,}0 \times 10^{-8}$ | $\sim 24$ |
| 32 | $\sim 2{,}3 \times 10^{-10}$ | $\sim 32$ |

En un procesador embebido a 100 MHz, el ahorro de microsegundos de las variantes SPR (Refact. y CORDIC) es lo que permite ejecutar el lazo de control de una plataforma móvil con latencia total inferior al umbral de inestabilidad dinámica (típicamente $< 1$ ms).

## Validación Empírica en Hardware Embebido: ESP32

Para contrastar las hipótesis anteriores fuera del entorno de escritorio, el benchmark completo se ejecutó sobre un microcontrolador ESP32-D0WD-V3 físico, bajo las condiciones descritas en la metodología. La Tabla 4 resume el costo de la transformación inversa en los tres regímenes aritméticos disponibles en el chip.

Tabla 4: Ciclos de CPU por transformación inversa en ESP32 (escenario estándar, promedio sobre 16 384 operaciones).

| **Variante** | **`double` (soft-float)** | **`float` (FPU)** | **entero Q20** |
|---|---:|---:|---:|
| Cilíndrico | 5 100 | 458 | — |
| Esférico | 10 050 | 958 | — |
| SPR Trad. | 8 821 | 802 | — |
| SPR Refact. | 8 769 | **659** | — |
| SPR CORDIC | 24 939 | 949 | **1 808** |

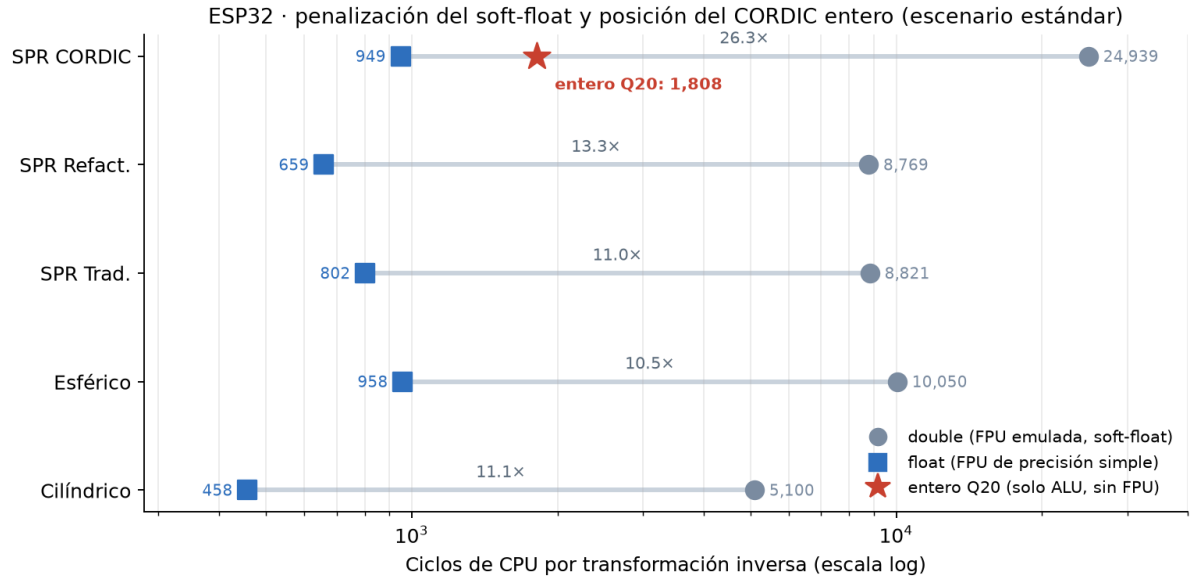



Figura 6: Penalización del *soft-float* en ESP32 para la transformación inversa. Cada segmento conecta la implementación en `float` (cuadrados, FPU de precisión simple) con su equivalente en `double` (círculos, emulación por software) del mismo modelo; la anotación indica el factor de penalización. La estrella señala la variante SPR CORDIC en aritmética entera Q20, que no emplea la FPU.

La Figura 6 revela la estructura del costo computacional. La emulación de doble precisión penaliza a los modelos basados en biblioteca matemática con factores de $10{,}5\times$ a $13{,}3\times$, pero castiga aún más severamente al CORDIC en punto flotante ($26{,}3\times$): al iterar 24 pasos cuyas sumas y productos pasan individualmente por *soft-float*, el algoritmo hereda la penalización en cada operación elemental, convirtiéndose en la variante más lenta del chip (24 939 ciclos). Este resultado empírico refina la hipótesis planteada en la Sección anterior: **la ventaja de CORDIC en hardware restringido no surge de su estructura iterativa por sí sola, sino de su capacidad de formularse en aritmética entera pura**. En efecto, la implementación en punto fijo Q20 —que reemplaza multiplicaciones por desplazamientos y elimina todo salto condicional del bucle— resuelve la inversa en 1 808 ciclos: $4{,}8\times$ más rápida que la mejor alternativa algebraica en `double` (SPR Refact., 8 769) y $13{,}8\times$ más rápida que el propio CORDIC flotante emulado.

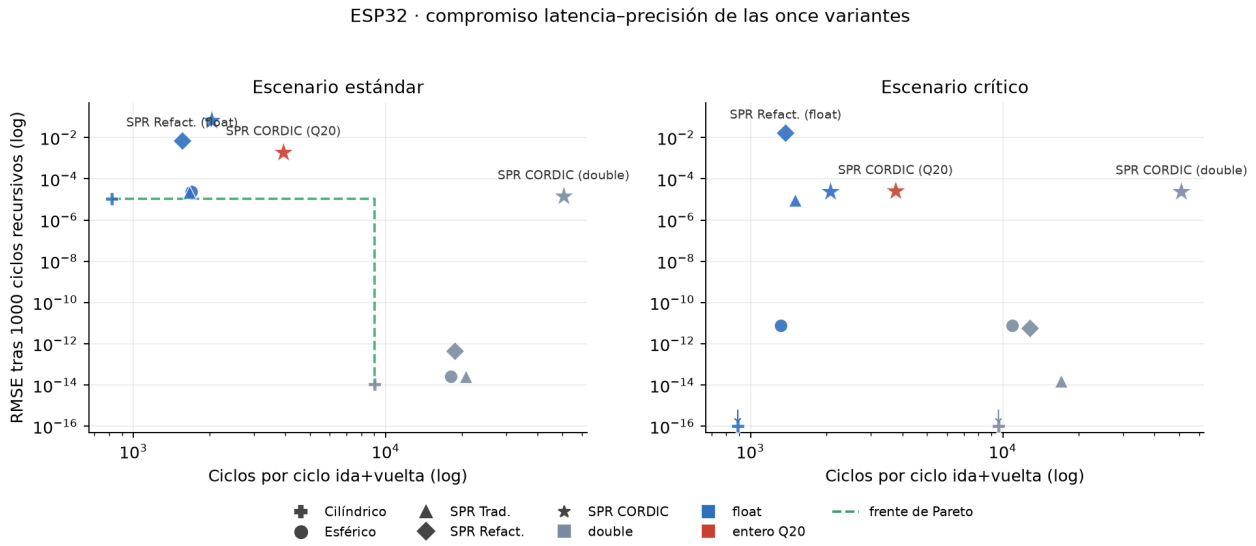



Figura 7: Compromiso latencia–precisión de las once variantes evaluadas en ESP32. Cada punto representa una variante (forma: modelo; color: régimen aritmético); el eje horizontal es el costo del ciclo cinemático completo (inversa + directa) y el vertical el RMSE tras 1 000 ciclos recursivos. Las flechas indican valores por debajo del piso del eje. La línea discontinua traza el frente de Pareto del escenario estándar.

El mapa de la Figura 7 sitúa a cada variante en el espacio latencia–precisión y permite seleccionar la implementación óptima según el requisito dominante. Tres observaciones merecen destacarse. Primera: el frente de Pareto del escenario estándar está definido por las variantes `float` en el extremo de baja latencia y por las variantes `double` en el extremo de alta precisión; la variante entera Q20 se ubica en el codo intermedio, ofreciendo precisión de $\sim 10^{-3}$ (dominada por la cuantización Q20, no por el algoritmo) a una fracción del costo de cualquier ruta `double`. Segunda: en el escenario crítico, las variantes CORDIC (tanto `float` como Q20) resultan *más precisas* que la refactorización algebraica en `float`: al ser la vectorización y la rotación CORDIC operaciones mutuamente inversas construidas con las mismas tablas, el par transformación–antitransformación es autoconsistente y el error recursivo no se amplifica cerca de la singularidad. Tercera: la degradación relativa de las variantes `float` tras 1 000 ciclos recursivos refleja el piso de precisión simple bajo realimentación extrema del error, no el error por operación individual (del orden de $10^{-6}$ relativo), que es el relevante en un lazo de control real donde cada consigna se procesa una sola vez.

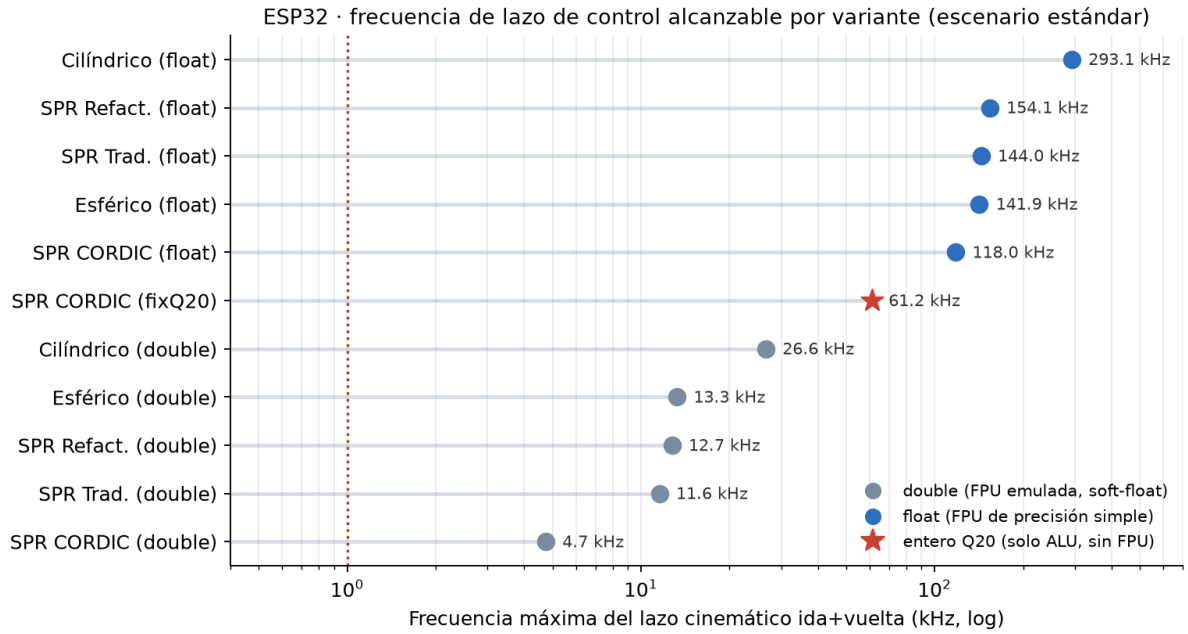



Figura 8: Frecuencia máxima teórica del lazo cinemático (transformación inversa + directa) para cada variante en ESP32 a 240 MHz, frente al umbral de estabilidad dinámica de 1 kHz típico en robótica móvil. Escala logarítmica.

En términos de ingeniería de control (Figura 8), todas las variantes superan el umbral de 1 kHz en este chip de 240 MHz, pero con márgenes radicalmente distintos: el ciclo cinemático completo de SPR CORDIC entero admite lazos de hasta 61 kHz, frente a los 12.7 kHz de SPR Refactorizado en `double` y los 4.7 kHz del CORDIC flotante emulado; con FPU de precisión simple, SPR Refactorizado alcanza 154 kHz. Estos márgenes determinan cuánta capacidad de cómputo queda disponible para el resto del lazo (dinámica, filtrado, comunicación) o, alternativamente, si un microcontrolador de menor frecuencia (y consumo) puede sostener la misma tasa de control.

Un atributo adicional de la variante entera, medido directamente, es su **determinismo**: la latencia difiere menos del $0{,}2$ % entre el escenario estándar y el crítico (1 808 frente a 1 806 ciclos), pues el bucle *branchless* ejecuta exactamente la misma secuencia de instrucciones con independencia de los datos. Esta propiedad, inalcanzable para las bibliotecas matemáticas con rutas condicionales internas, es la que garantiza tiempos de ciclo estrictos en planificadores de tiempo real.

Finalmente, la campaña embebida aporta dos acotaciones de honestidad metodológica. Por un lado, confirma que la magnitud de la ventaja de SPR depende de la biblioteca matemática de cada plataforma: con la *newlib* de Xtensa, la superioridad de SPR sobre el modelo esférico en la inversa es de $1{,}15\times$ en `double` y $1{,}46\times$ en `float`, frente al $1{,}33\times$–$1{,}45\times$ medido con la UCRT en x86_64; el signo del beneficio es consistente en todas las plataformas evaluadas, pero su escala no es transferible entre bibliotecas. Por otro lado, fue precisamente esta validación con perturbaciones simétricas la que expuso la condición de frontera del Algoritmo 4 descrita en la Sección 5.2, lo que subraya el valor del ensayo *hardware-in-the-loop* como complemento indispensable de los benchmarks de escritorio.

### Síntesis: posición de SPR frente a los demás sistemas

Reuniendo las mediciones de ambas plataformas, la posición de cada sistema de coordenadas puede resumirse en cuatro puntos:

- **Conversión inversa** (de cartesianas al sistema; la que un controlador ejecuta con mayor frecuencia al procesar lecturas de posición). El sistema cilíndrico es el más rápido en todos los casos, pero maneja un solo ángulo y no basta para describir la orientación de un posicionador de dos ejes. Entre los dos sistemas que sí la describen —esférico y SPR—, **SPR ocupó el primer lugar en todas las pruebas**: 1.33 veces más rápido que el esférico en la PC (escenario estándar), 1.45 veces en el escenario crítico, y entre 1.15 y 1.46 veces en el ESP32 según la precisión empleada.
- **Conversión directa** (del sistema a cartesianas). Aquí el orden se invierte: el cilíndrico y el esférico son más rápidos, y SPR paga sus evaluaciones trigonométricas adicionales. En un lazo de control esta conversión se ejecuta con menor frecuencia que la inversa, por lo que el balance del ciclo completo sigue siendo favorable, pero es la limitación a tener en cuenta si una aplicación ejecuta principalmente conversiones directas.
- **Precisión cerca del polo**. SPR es el único sistema de dos ángulos cuyo error no crece cerca de la singularidad: tras mil conversiones repetidas del mismo punto, su error se mantiene en el orden de $10^{-12}$, mientras que el esférico se degrada unas 40 veces respecto a su comportamiento en el resto del dominio.
- **Procesadores sin FPU o limitados a doble precisión emulada**. La versión entera de SPR con CORDIC fue la opción más rápida del ESP32 en este régimen (4.8 veces sobre la mejor alternativa de doble precisión) y la única cuyo tiempo de ejecución es constante, lo que simplifica el análisis de plazos en sistemas de tiempo real. Ninguno de los otros sistemas evaluados cuenta con una implementación equivalente validada.

En síntesis: cuando la aplicación requiere dos ángulos, SPR ofrece la conversión inversa más rápida y la mejor estabilidad numérica de los sistemas evaluados; y cuando el hardware es muy limitado, su variante entera mantiene la velocidad y añade tiempo de ejecución constante sin necesidad de unidad de punto flotante.

## 8. Conclusiones

Esta investigación presentó el sistema de coordenadas SPR y lo comparó con los sistemas cilíndrico y esférico midiendo diez millones de conversiones en dos plataformas: un procesador de escritorio y un microcontrolador ESP32. Los resultados sitúan a SPR en una posición concreta: entre los sistemas de dos ángulos, su conversión inversa fue la más rápida en todas las pruebas (entre 1.15 y 1.5 veces sobre el esférico) y su precisión fue la única que no se degradó cerca de los polos. Su conversión directa es más costosa que la del esférico, de modo que la ventaja global de SPR se concentra en las aplicaciones donde domina la conversión inversa, que es el caso típico de los lazos de control.

En términos prácticos, esto se traduce en margen de cómputo sobre hardware económico. En el ESP32 a 240 MHz, el ciclo completo de conversiones de SPR admite lazos de control de 12.7 kHz en doble precisión, 154 kHz en precisión simple y 61 kHz en su versión entera: en todos los casos, muy por encima del umbral de 1 kHz habitual en robótica móvil. El tiempo restante del procesador queda disponible para las demás tareas del controlador (dinámica, filtrado, comunicaciones).

La estabilidad cerca del polo es la segunda característica distintiva. El error de SPR no depende de la ubicación del punto: tras mil conversiones repetidas se mantiene en el orden de $10^{-12}$, mientras que el sistema esférico se degrada unas 40 veces en la vecindad de su singularidad. Para mecanismos que atraviesan esas orientaciones (plataformas de apuntamiento, muñecas robóticas), esto elimina una fuente de error difícil de diagnosticar.

La variante entera de SPR con el algoritmo CORDIC [6], [7] completa el panorama: fue la opción más rápida del ESP32 cuando se exige doble precisión o no hay unidad de punto flotante (4.8 veces sobre la mejor alternativa) y la única con tiempo de ejecución constante (variación menor al 0.2 %), lo que simplifica el diseño de sistemas de tiempo real. Cuando el procesador dispone de FPU de precisión simple, la versión algebraica de SPR es la más rápida (2,7 $\mu$s por conversión inversa a 240 MHz). Ambas variantes comparten la misma formulación geométrica: el diseñador puede elegir la implementación según su hardware sin cambiar el modelo matemático. Finalmente, la validación sobre hardware real permitió detectar y corregir una condición de frontera en la conversión inversa y un artefacto en la instrumentación de medición, lo que refuerza la fiabilidad de las cifras aquí publicadas.

## Direcciones de correo de los autores

Salas-García, J.*: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
proyectos@javiersaalsg.com

Hernández-Servín, J. A.: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
joseph_servin@uaemex.mx

Vilchis-González, A. H.: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
avilchisg@uaemex.mx

Ávila-Vilchis, J. C.: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
jcavilav@uaemex.mx

# Referencias

[1] A. Codourey, «Dynamic Modeling of Parallel Robots for Computed-Torque Control Implementation,» *The International Journal of Robotics Research*, vol. 17, n.º 12, págs. 1325-1336, 1998. DOI: 10.1177/027836499801701205

[2] J.-P. Merlet, *Parallel Robots*, 2.ª ed. Dordrecht: Springer, 2006. DOI: 10.1007/1-4020-4133-0

[3] C. Gosselin y J. Angeles, «Singularity analysis of closed-loop kinematic chains,» *IEEE Transactions on Robotics and Automation*, vol. 6, n.º 3, págs. 281-290, 1990. DOI: 10.1109/70.56660

[4] D. L. Williamson, J. B. Drake, J. J. Hack, R. Jakob y P. N. Swarztrauber, «A standard test set for numerical approximations to the shallow water equations in spherical geometry,» *Journal of Computational Physics*, vol. 102, n.º 1, págs. 211-224, 1992. DOI: 10.1016/S0021-9991(05)80016-6

[5] B. Siciliano, L. Sciavicco, L. Villani y G. Oriolo, *Robotics: Modelling, Planning and Control*. London: Springer, 2009. DOI: 10.1007/978-1-84628-642-1

[6] J. E. Volder, «The CORDIC Trigonometric Computing Technique,» *IRE Transactions on Electronic Computers*, vol. EC-8, n.º 3, págs. 330-334, 1959. DOI: 10.1109/TEC.1959.5222693

[7] R. Andraka, «A Survey of CORDIC Algorithms for FPGA Based Computers,» en *Proceedings of the 1998 ACM/SIGDA Sixth International Symposium on Field Programmable Gate Arrays*, New York: ACM, 1998, págs. 191-200. DOI: 10.1145/275107.275139

[8] R. Sadourny, «Conservative Finite-Difference Approximations of the Primitive Equations on Quasi-Uniform Spherical Grids,» *Monthly Weather Review*, vol. 100, n.º 2, págs. 136-144, 1972. DOI: 10.1175/1520-0493(1972)100<0136:CFAOTP>2.3.CO;2

**Informaticae Abstracta**. ISSN: 3061-8533
Año 2026, Vol. 4, No. 1, pp. 124-137
ENE | JUN 2026
doi: https://doi.org/10.36677/ia.v4n1.28943

# Prototipado de órganos artificiales para entrenamiento en cirugía robótica

Dávila-Vilchis, J. M.; Zúñiga Avilés, L. A.*; Ramírez-García, N. G.; Vilchis-González, A. H.

**Información del Artículo**

Recibido: 1 de marzo, 2026
Aceptado: 10 de junio, 2026

**Palabras clave:** Órganos Artificiales, Intestino, Riñón, Impresión 3D, Modelos anatómicos, Cirugía Robótica

**Resumen**

Este artículo presenta el prototipado rápido de órganos artificiales, como modelos anatómicos para entrenamientos quirúrgicos asistidos por sistemas robóticos. Para lograr este objetivo, se propone una metodología integral que combina el diseño, impresión 3D y su fabricación artesanal, la cual está organizada en cuatro etapas 1) definición de criterios de diseño, 2) digitalización y procesamiento de imágenes médicas para la generación de modelos tridimensionales, 3) impresión 3D de modelos anatómicos y/o moldes y 4) fabricación artesanal del órgano mediante técnicas de vaciado en silicón. Los resultados incluyen la producción de modelos de riñón e intestino con dimensiones anatómicas reales y propiedades mecánicas, al ser órganos vitales mayormente utilizados en procedimientos quirúrgicos mínimamente invasivos, como la nefrectomía o laparoscopia, respectivamente. Este estudio aporta una guía metodológica para el diseño y desarrollo de otros órganos artificiales que satisfacen los requerimientos anatómicos necesarios en la formación médica asistida por robots.

## 1. Introducción

El cuerpo humano está compuesto de múltiples órganos que, en conjunto, conforman sistemas y aparatos encargados de mantener la homeostasis y las funciones vitales del organismo. Entre ellos destacan el corazón, el cerebro, los pulmones, el hígado y los riñones, cuyo daño, disfunción o ausencia puede derivar en enfermedades crónicas, deterioro sistémico e, incluso, la muerte. Por esta razón, el estudio, diagnóstico, tratamiento y abordaje quirúrgico de estos órganos representan una prioridad fundamental en la práctica médica contemporánea.

El corazón forma parte del sistema circulatorio y tiene como función principal bombear sangre a todo el cuerpo, permitiendo el transporte de oxígeno y nutrientes hacia las células [1]. El cerebro, por su parte, integra el sistema nervioso central y actúa como centro de comando del organismo, al regular funciones esenciales mediante conexiones neuronales especializadas, en las que se encuentran la coordinación motora, la percepción sensorial, el razonamiento, la resolución de problemas, las emociones y el aprendizaje, que en conjunto con la médula espinal constituyen el principal centro de correlación e integración de la información nerviosa [2] . Los pulmones, órganos principales del sistema respiratorio, realizan el intercambio gaseoso entre el oxígeno y el dióxido de carbono, mediante estructuras de tejido blando y elástico que se expanden y contraen durante la respiración [3].

El hígado es un órgano esencial para la regulación y el mantenimiento de la homeostasis metabólica, ya que participa en procesos de absorción, síntesis, almacenamiento, metabolismo y redistribución de nutrientes, tales como lípidos, carbohidratos, proteínas y vitaminas [4]. Asimismo, constituye el principal órgano encargado de la detoxificación de compuestos xenobióticos y presentar una elevada capacidad regenerativa ante distintos tipos de daño

[5]. Por otro lado, los riñones, órganos excretores del sistema urinario, se localizan en la región posterior de la cavidad abdominal y tienen como función principal filtrar la sangre para eliminar sustancias de desecho mediante la orina, tales como urea, creatinina, ácido úrico y bilirrubina. Además, contribuyen al mantenimiento del volumen y la composición de los líquidos corporales mediante la excreción regulada de solutos y agua [6].

En el caso del sistema digestivo está constituido por el intestino delgado y grueso, un conjunto de tubos abiertos llamado tracto intestinal, cuya longitud varía entre 6 m y 10 m. El intestino se encarga de cumplir funciones esenciales en la digestión química y la absorción de nutrientes gracias a sus vellosidades y microvellosidades, además de concentrar, almacenar y eliminar los desechos sólidos [7].

Sin embargo, la insuficiencia orgánica de cualquiera de ellos, representa una de las problemáticas más críticas de la medicina contemporánea. A nivel mundial, millones de pacientes permanecen en lista de espera para recibir un trasplante, mientras que la brecha entre la demanda de órganos y la disponibilidad de donantes continúa incrementándose. En México, esta situación es apremiante: al corte de febrero de 2026, se registraron 18,626 personas en lista de espera de un trasplante, de las cuales 16,217 requerían un riñón y 2,181 una córnea. Sin embargo, únicamente alrededor del 20 % de los pacientes que requieren un trasplante logra acceder a uno, de acuerdo con cifras del Centro Nacional de Trasplantes [8].

Frente a esta problemática, se han propuesto tres líneas: 1) el desarrollo de órganos artificiales mecánicos, como el corazón artificial total, sistemas de diálisis renal, los cuales pueden sustituir o sostener parcialmente las funciones vitales, pero aún no son capaces de replicar la complejidad fisiológica del órgano original [9]. 2) la medicina regenerativa cuyo objetivo es reconstruir tejidos y órganos funcionales a partir de células del propio paciente, con el propósito de reducir el rechazo inmunológico y ofrecer soluciones terapéuticas personalizadas [10], [11]. Una aproximación más avanzada de la medicina regenerativa es la bioimpresión mínimamente invasiva, conocida como Minimally Invasive Bioprinting (MIB), una técnica emergente que busca trasladar esta capacidad hacia órganos internos mediante robots suaves que son capaces de ejecutar incisiones mínimas u orificios naturales dentro del cuerpo [12], [13]. 3) el uso de la cirugía robótica es otra técnica quirúrgica mínimamente invasiva, cuyo objetivo es usar brazos robóticos rígidos que actúan como una extensión del cirujano en procedimientos complejos para la remoción de tumores en los órganos dañados, reparación de tejidos, hernias, entre otros [14].

Durante las últimas décadas, la ingeniería biomédica junto con la robótica médica ha desarrollado dispositivos capaces de asumir, parcial o totalmente, las funciones de algunos órganos vitales con la intención de proporcionar una mejor calidad de vida a los pacientes. Además, se han fabricado modelos anatómicos con impresión 3D que combinan ingeniería, sensores y software [15] para asistir en cirugías, rehabilitación, diagnóstico médico y tratamiento de enfermedades y logística hospitalaria [16]. Por ejemplo, en México, algunos hospitales de alta especialidad se han enfocado en prácticas quirúrgicas mínimamente invasivas asistidas donde se emplean robots como el Da Vinci®, ya que ofrecen mayor precisión, menor invasión, menor dolor postoperatorio, y la recuperación es más rápida [17].

Sin embargo, el uso clínico de estos sistemas exige que los especialistas cuenten con formación técnica y entrenamiento específico antes de realizar procedimientos en pacientes. El entrenamiento de cirujanos en prácticas quirúrgicas con robots combina simulación avan-

zada, realidad virtual y práctica supervisada en entornos controlados [18]. Esta formación busca desarrollar habilidades como precisión, coordination, control de instrumentos, percepción espacial y toma de decisiones en procedimientos mínimamente invasivos. En este contexto, los modelos anatómicos artificiales representan una herramienta relevante para la capacitación quirúrgica, ya que permiten reproducir escenarios clínicos controlados y practicar maniobras como suturas, cortes, disecciones y manipulación de tejidos sin comprometer la seguridad del paciente.

Por lo tanto, este artículo presenta una metodología de diseño para la generación de órganos artificiales destinados a funcionar como modelos anatómicos en entrenamientos quirúrgicos asistidos por sistemas robóticos. La innovación de la metodología se centra en el balance de combinar manufactura aditiva con técnicas artesanales para aceleramiento de prototipado con bajo costo. La propuesta busca permitir que los médicos especialistas realicen prácticas con órganos vitales artificiales, tejidos sintéticos o biomiméticos, con el fin de adquirir sensibilidad táctil, control instrumental, precisión en cortes y dominio de maniobras quirúrgicas mediante brazos robóticos.

El artículo está organizado de la siguiente manera: en la sección de metodología 2 se describen las etapas de diseño, procesamiento de imágenes, generación de modelos tridimensionales, impresión 3D de moldes y fabricación artesanal mediante vaciado en silicón. Posteriormente, en la sección de resultados 3 se presenta el desarrollo de un modelo de intestino y de riñón con sus características anatómicas, dimensionales y funcionales para su aplicación en entrenamiento quirúrgico mínimamente invasivo. Finalmente, en la sección de conclusiones 4 se discuten los alcances de la metodología propuesta, sus posibles aplicaciones en simulación quirúrgica robótica y su potencial como guía para el desarrollo de nuevos modelos anatómicos artificiales orientados a la formación médica especializada.

## 2. Metodología

La robótica médica está transformando la práctica clínica en todo el mundo, incluido México, ofreciendo precisión quirúrgica. Por lo que, el uso de modelos anatómicos artificiales es necesario en el entrenamiento y/o capacitación de doctores para simular escenarios reales y garantizar procedimientos más seguros con el manejo de brazos robotizados como el Versius® o el Da Vinci® y los especialistas sean capaces de certificarse en el uso de estos equipos para poder hacer uso de ellos con humanos.

Por lo tanto, este artículo, presenta una metodología integral para el prototipado artesanal de órganos artificiales, específicamente para los órganos vitales que se muestran en la Figura 1. Cabe resaltar que esta metodología no requiere de equipo sofisticado, por lo que puede aplicarse a cualquier otro órgano, para el uso y manejo de robots quirúrgicos que les permitan adquirir sensibilidad táctil lo más real posible a los órganos naturales. La metodología se integra por 4 etapas: 1) definición de criterios, 2) segmentación computacional a partir de imágenes médicas, 3) impresión 3D de los modelos anatómicos y 4) fabricación artesanal a una escala 1:1, como se describe a continuación:

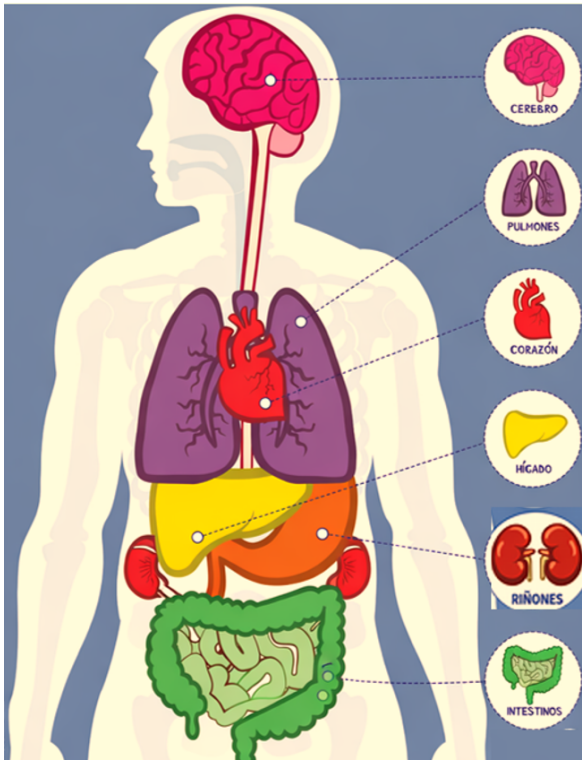



Figura 1: Órganos vitales del ser humano (OpenAI, 2026).

## Definición de Criterios de Diseño

El proceso de fabricación de órganos empieza a partir de analizar la anatomía del órgano objetivo para replicar su estructura y/o arquitectura, específicamente forma, tamaño, localización. En conjunto con su fisiología para analizar su funcionamiento, procesos e interacción de sus tejidos y sus propiedades físicas como masa, dimensiones, densidad y módulo de elasticidad $E$, [19], mostradas en la Tabla 1

Tabla 1: Propiedades de los órganos humanos

| **Órgano** | **Masa** $(g)$ | **Dimensiones** $(cm)$ | **Densidad** $(g/cm^3)$ | **E** $(kPa)$ |
|---|---|---|---|---|
| Corazón | 280-3450 | 12x9x6 | 1.06 | 20-50 |
| Cerebro | 1100-1400 | 16x14x9 | 1.03 | 1 |
| Pulmones | 1000 | 26x15x4 | 0.24 | 5-15 |
| Hígado | 1400-1800 | 22x15x10 | 1.05 | 0.5-1.5 |
| Riñón | 125-170 | 12x7x3 | 1.06 | 2-4 |
| Intestino | 2000-4000 | 700x4x3 | 1.05 | 100-300 |

Para la creación de órganos artificiales es necesario obtener modelos anatómicos con las mismas dimensiones que simulen las propiedades reales del órgano. En este paso es necesario obtener y caracterizar las propiedades mecánicas de rigidez, módulo de elasticidad y dureza, así como sus propiedades reológicas de viscosidad y viscoelasticidad para la selección de los materiales a utilizar, principalmente polímeros y garantizar la similitud anatómica

con los órganos humanos. Por ejemplo, la densidad del tejido pulmonar es muy baja porque está compuesto mayormente por aire.

## Procesamiento de imágenes y generación de modelos 3D

Esta etapa consiste en hacer uso de 3D Slicer®, un software de código abierto para la visualización y tratamiento de imágenes médicas en 3D, especialmente obtenidas de tomografías computarizadas (TC) o resonancias magnéticas (MRI) en formato DICOM (Digital Imaging and Communications in Medicine) y las muestra en cortes axiales, coronales y sagitales. En este paso, se visualiza el órgano a fabricar, de ahí se hace la segmentación en la región objetivo o vecindad donde es posible aislar estructuras anatómicas no funcionales como tumores, órganos, huesos, tal como se muestra en las diferentes vistas de la cavidad torácica de la Figura 2 extraída del sitio web embodi3D, una página que ofrece imágenes biomédicas de acceso abierto (AA).

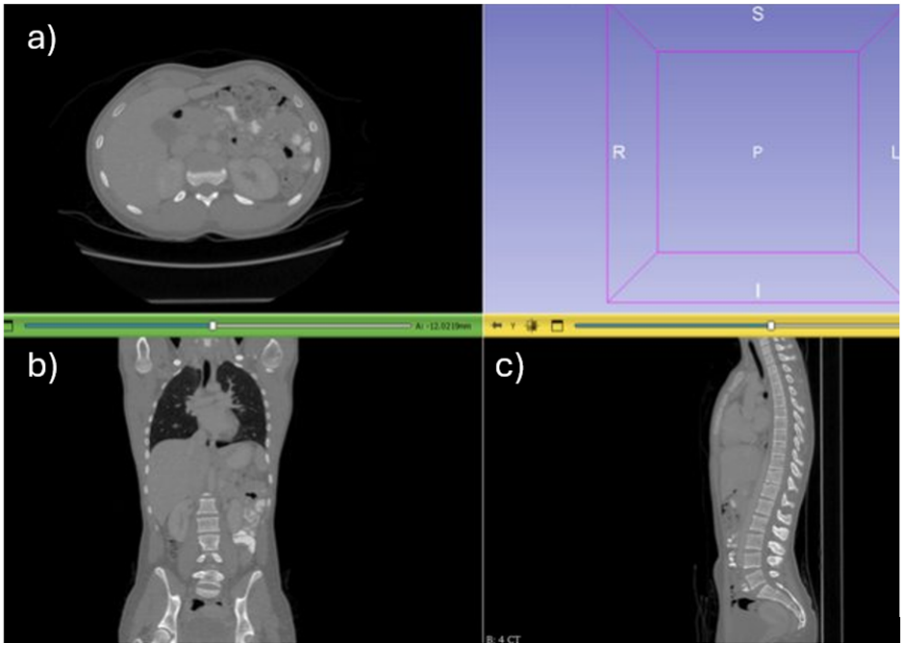



Figura 2: Vistas a) superior, b) frontal y c) lateral de la cavidad torácica en 3D Slicer (AA).

Después, se reconstruye el modelo anatómico, es decir convierte las imágenes planas en modelos tridimensionales manipulables, donde es posible medir su volumen, dimensiones, densidades para finalmente exportarlo como un archivo STL u OBJ para ser utilizado en cualquier software CAD (Computer Aided Design) como SolidWorks®, Fusion® o directamente en uno de impresión 3D como Creality® o Ultimaker Cura®, por ejemplo.

## Impresión 3D de moldes y/o modelos

Una vez obtenido el modelo digitalizado del órgano, se continúa con la Manufactura Aditiva por medio de impresión 3D, ya que esta técnica permite crear piezas de cualquier modelo sin importar su morfología, particularmente los órganos humanos que se caracterizan por ser geometrías complejas.

La impresión 3D facilita la obtención de prototipos rápidamente debido a su proceso laminar capa a capa, donde la pieza se va formando gradualmente desde la base hasta la parte superior. Además, existe una disponibilidad y diversificación de sistemas de impresión 3D, por lo que sus costos son factibles para la manufactura de pequeños prototipos y/o a escala.

Dentro de los filamentos para impresión 3D se destaca el TPU (Poliuretano Termoplástico) con durezas de 95A u 85A, siendo este el más flexible, pero difícil de caracterizar sus parámetros de impresión; el PLA (ácido poliláctico) que se caracteriza por su facilidad de impresión para prototipos rápidos, ABS (acrilonitrilo butadieno estireno) o PETG (tereftalato de polietileno modificado con glicol) ofrecen mayor resistencia y durabilidad pero su costo es mayor. La Tabla 2 muestra las características de los materiales de impresión comúnmente utilizados para una altura de 0.2 mm entre capas.

Tabla 2: Propiedades de los filamentos para impresión 3D

| **Filamento** | $T_{extrusivn}$ | **Velocidad** $(m/s)$ | $T_{cama}$ | **RA** |
|---|---|---|---|---|
| Hyper PLA | $220°C$ | 300 | $45°C$ | Alta |
| PLA | $230°C$ | 200 | $45°C$ | Media |
| PETG | $250°C$ | 120 | $80°C$ | Alta |
| ABS | $260°C$ | 120 | $110°C$ | Alta |
| TPU | $230°C$ | 50 | $60°C$ | Flexible |

Actualmente, existe una versatilidad en impresoras 3D que cuentan con su propio software, por ejemplo, (Creality®, Cura®) donde es necesario definir las siguientes propiedades de la impresión: velocidad, temperatura $T$, de la boquilla y del plato o cama, altura de capa, posicionamiento de la pieza, o si requiere de soportes, relleno para la impresión, el tipo de material, espesor. La Figura 3 muestra la segmentación del intestino y de la parte interna del riñón con sus principales conductos dentro de la interfaz del software de Creality®, incluyendo sus soportes para ser impresos.

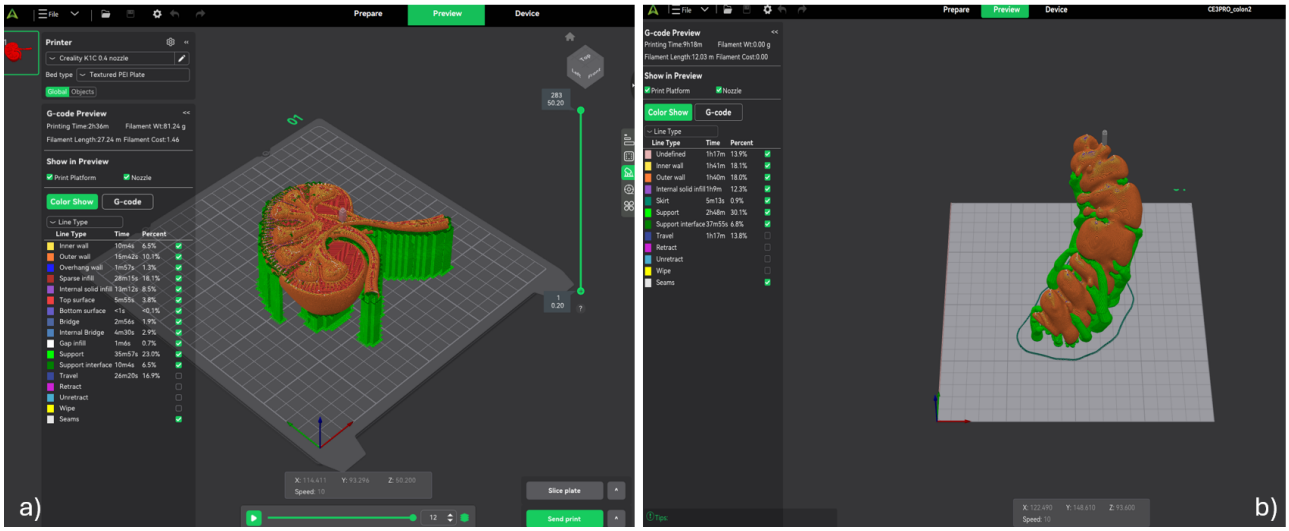



Figura 3: Vista isométrica del a) riñón y del b) intestino laminado en Creality®.

## Fabricación artesanal de los órganos

Durante el proceso de impresión, las piezas pueden resultar con impurezas o deformaciones debido a la naturaleza del material, errores de la máquina durante la impresión o no tener las propiedades mecánicas requeridas. Por lo tanto, esta investigación incluye la fabricación artesanal, al ser de gran ayuda en el prototipado de geometrías complicadas. Esta técnica consiste que una vez impresos los modelos con las dimensiones y anatomía de dicho órgano se elaboran moldes artesanales particionados de silicón para que registren la forma anatómica. El proceso artesanal consiste en tres etapas básicas: primero, el modelado artesanal del prototipo a reproducir; segundo, elaboración de molde del material compatible al modelo; tercero, realizar el vaciado o pieza final de dicho molde con cualquier material deseado, por ejemplo, silicón para obtener las mismas propiedades mecánicas.

En este caso la pieza es impresa en 3D, entonces se inicia con elaborar el molde. Al tener una pieza orgánica y curva, se monta una cama y tacel de plastilina alrededor de la pieza para evitar que se mueva al estar manipulando la pieza durante el vertido de silicón en el molde; el área del tacel debe ser plana con un ancho de 3 cm para sostener el molde ahogado o corazón, dicho tacel estará en posición horizontal exactamente a la mitad de la forma de todo el elemento para obtener una parte A. Este paso se realiza en la mitad de la pieza, pero también en la otra mitad, parte B, para obtener el registro de ambos lados completos. El molde estará elaborado en dos partes A y B para que estando abierto se pueda introducir un molde ahogado o corazón como se muestra en la Figura 4.

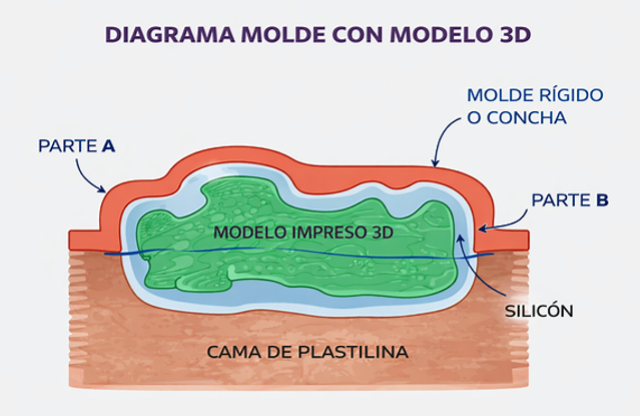



Figura 4: Diagrama de registro del molde-modelo en plastilina.

Posteriormente, se prepara el silicón con su catalizador y se vierte el silicón que es líquido espeso sobre la pieza cubriendo por completo el modelo y una parte del tacel que está en horizontal, y cuidar hasta que vulcanice el material; dando como resultado un molde flexible que permite la salida de una pieza compleja sin que quede atorado en el interior del molde. Este molde de silicón llevará un contra molde o concha rígida que permite la manipulación durante el vaciado del contenido. El material de la concha externa puede ser también de material rígido en un proceso artesanal. Después, se prepara una capa de plastilina (que no contenga sulfuros) de 5mm de grosor que se introduce al interior del molde abierto, debe estar dicha capa en cualquier parte con el mismo grosor uniforme. Finalmente, se rellena con resina poliéster para fabricar el corazón, dicha pieza se va a sostener del tacel rígido del

molde o concha rígida por medio de arcos del mismo material, uniendo el corazón y al área del tacel con su respectiva entrada de aire, como se observa en la Figura 5.

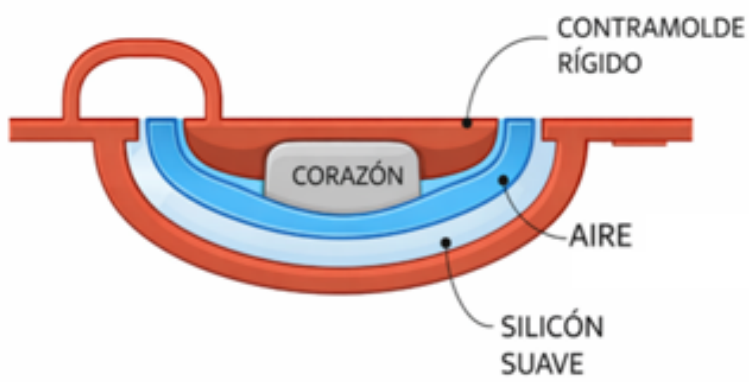



Figura 5: Diagrama de vaciado de silicón en el molde.

Una vez terminado el molde A y B se procede a vaciar. Se rellena el espacio de aire de 5 mm con silicón Ecoflex 00-30 Smooth-on®, el cual se caracteriza por su elasticidad, transparencia y baja viscosidad, en este proceso se vierte en estado líquido una mezcla de 1A:1B durante 3 minutos, después es introducida a una cámara de presión a 20 kPa por 60 s para expulsar el aire que tenga dicho silicón y se deja reposar hasta que cure en un tiempo de 40 minutos a temperatura ambiente aproximadamente 23°C. Ya que la pieza se curada se pega a una pared rectangular del mismo silicón con un espesor 5 mm. Finalmente se procede a cerrar la pieza uniendo A y B con pegamento especial de silicón Sil-Poxy también de la marca Smooth-On®. Por lo tanto, la selección del silicón se hace dependiendo de la dureza deseada.

# 3. Resultados

Esta sección presenta los resultados preliminares de la metodología aplicada a la fabricación de los modelos anatómicos del riñón y del intestino grueso para ser utilizados en entrenamientos quirúrgicos con robots, describiendo así cada una de sus etapas.

## Criterios de diseño

Los criterios de diseño del riñón y del intestino se establecieron con base en su anatomía, funcionamiento y propiedades fisiológicas definidas previamente (ver Tabla 1). Su fabricación se requiere para ser empleados en el entrenamiento de cirugías mínimamente invasivas con robots, por ejemplo, con el Da Vinci para la resección de tumores donde es posible realizar cortes precisos; b) procedimientos donde el robot facilita suturas en zonas profundas; c) reconstrucción y uniones de segmentos y cierre. Por lo que es necesario, los especialistas deben demostrar sus habilidades quirúrgicas en el manejo del robot.

El intestino se caracteriza por formar unas cámaras elipsoidales que se inflan y desinflan secuencialmente, las cuales están determinadas por el radio de la esfera $a$, altura $b$, longitud del cilindro conector $L$ entre cámaras, altura del cilindro conector $c$ y número de cámaras o

esferas $n$, como se observa en la Figura 6. Mientras, que el riñón tiene una geometría característica en forma de frijol.

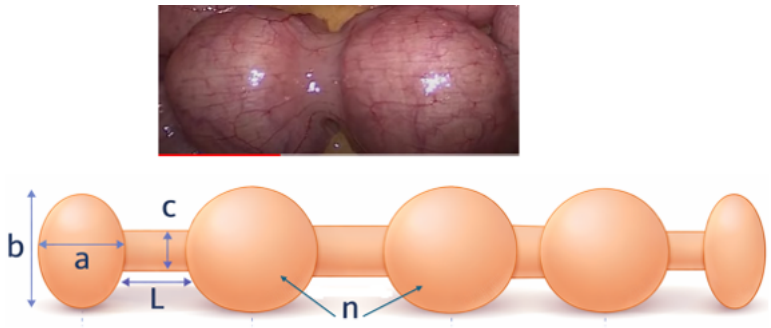



Figura 6: Movimiento de contracción y relajación del intestino.

## Procesamiento de imágenes e impresión 3D

Los modelos computacionales del intestino y del riñón se obtuvieron del sitio web embodi3D®, que proporciona archivos de Acceso Abierto. Después, se utilizó 3D Slicer® para la segmentación de sus modelos orgánicos, haciéndole un corte transversal al riñón con sus principales conductos. Después, se generaron los archivos STL para posteriormente exportarlos al software de Creality® (ver Figura 3) en una impresora K1 pro Creality® usando PLA con los parámetros definidos en la Tabla 2. La Figura 7 muestra la digitalización del intestino y los productos 3D obtenidos mostrando en el riñón sus nefronas, corteza y conductos ya impresos en 3D a escala real.


Figura 7: Movimiento de contracción y relajación del intestino.

## Fabricación de modelos de silicón

Posteriormente se quitan los soportes agregados a los modelos 3D y dependiendo de la calidad de la impresión se le da un acabado fino con pasta automotriz para eliminar las imperfecciones. Para el caso del vaciado del intestino de PLA se le aplicó silicón Brush-on 35 Smooth On®, solución ideal para aplicaciones en verticales sin escurrimientos, el cual está compuesto por una parte A (base) y una parte B activo (catalizador) que se mezclan 1:1 con un tiempo de curado (pot life) de 20 minutos a una temperatura de 23 °C. Después, se montó

a una base vertical donde se aplicaron múltiples capas, hasta obtener un volumen 14 cm x 2cm x 2cm. Finalmente se expulsó el modelo impreso del silicón gracias a la flexibilidad del material y se obtuvo un modelo físico inflable de 20 gr, como se muestra en las Figura 8.

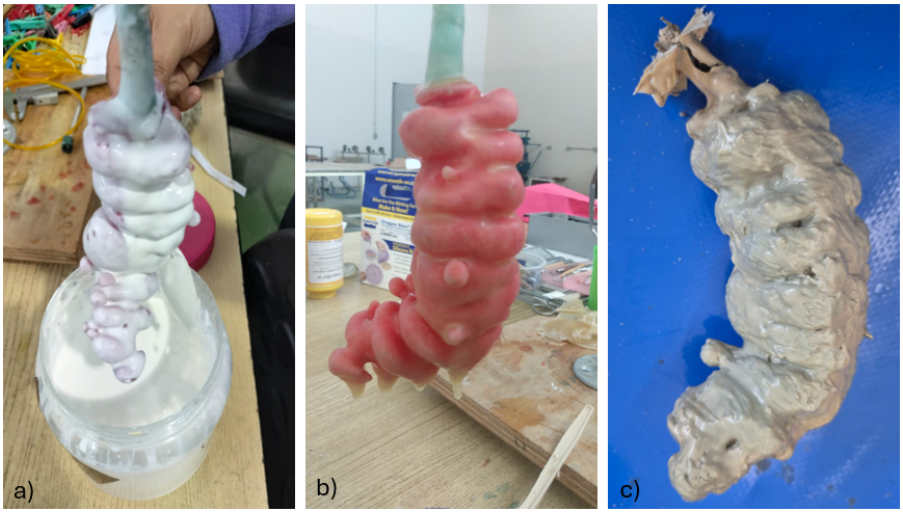



Figura 8: Movimiento de contracción y relajación del intestino.

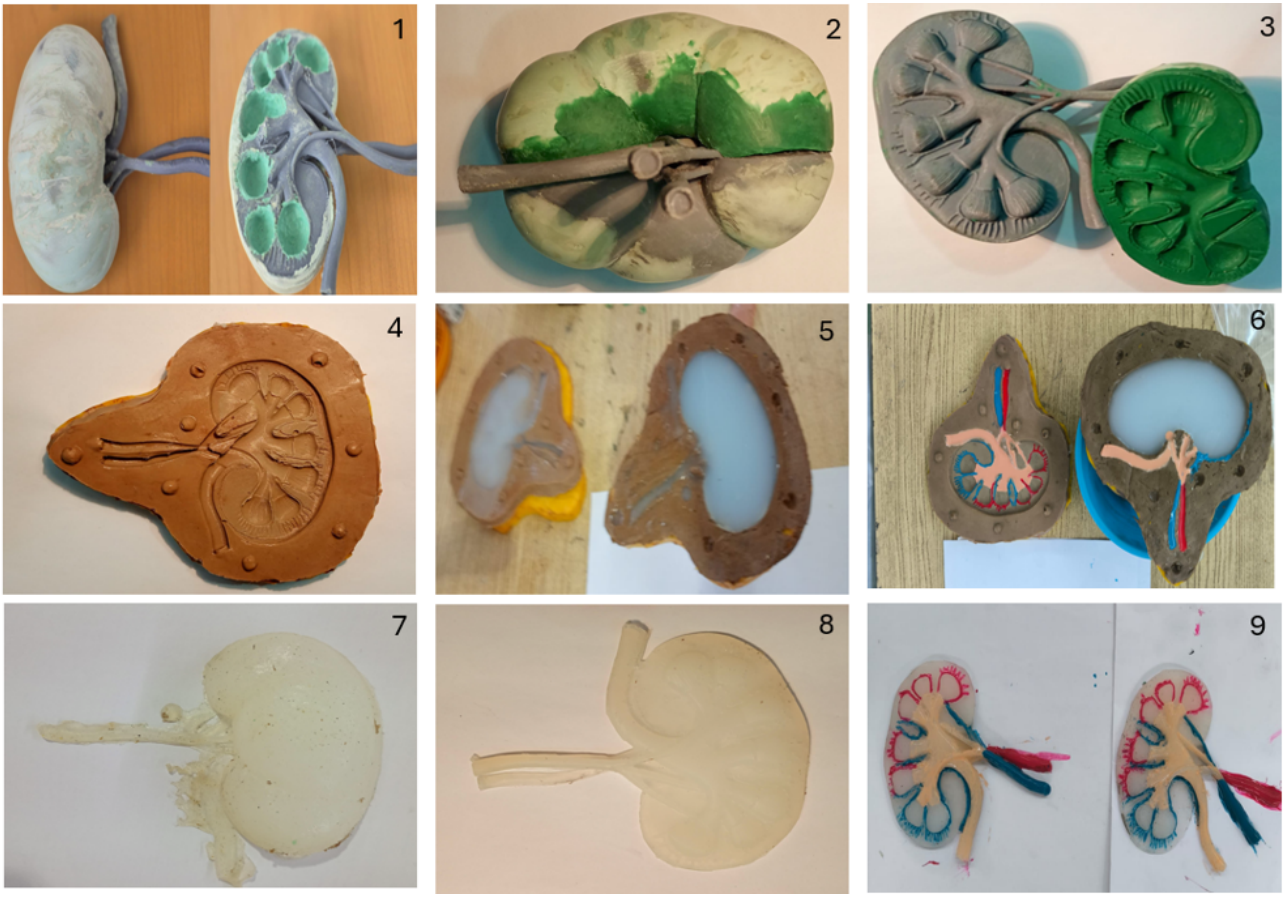



Figura 9: Movimiento de contracción y relajación del intestino.

Por otro lado, el proceso de fabricación del riñón se hizo de diferente manera, al constar de dos partes A y B, por lo que primero se hizo el desbaste y pulido con pasta automotiva epóxica para modelar la parte interna y externa para obtener un riñón completamente cerrado. Después, se fabricó una placa de plasticera (60 % de plastilina y 40 % de cera de abeja) y se aplicó calor en la placa plástica hasta lograr la unión de A con B para registrar la forma como en un sello.

Una vez terminado el modelo, se fabricó el molde de silicón. El molde lleva un tacel y llaves para que el molde al ser de dos partes no se mueva en algún procedimiento. Nuevamente, se utilizó Brush On 35 para fabricar el molde, una vez vulcanizado se aplicó desmoldante universal en Spray Smooth On® y se preparó la mezcla de silicón EcoFlex 00-10 Smooth On® para el vaciado. Además, se empleó una cámara de presión de vacío a 20-30 kPa durante 60 s para eliminar las burbujas de aire que se forman al mezclarse. Después de 4 horas se obtuvo el riñón en silicón para aplicarle la coloración y darle una apariencia más realista. Este procedimiento consiste en verter el silicón por capas, cada color se deja curar y hasta que un color esté vulcanizado se puede proceder con el siguiente color y así sucesivamente, el resultado final se aprecia en la Figura 9.

La fabricación artesanal ofrece ventajas en términos de accesibilidad, personalización y bajo costo relativo. Sin embargo, presenta limitaciones importantes, principalmente asociadas con mayores tiempos de fabricación, variabilidad entre piezas y baja escalabilidad para producción en serie. Por ello, futuras investigaciones deberán orientarse a la estandarización del proceso, la caracterización mecánica cuantitativa de los materiales utilizados, la validación con especialistas en cirugía robótica y la comparación de desempeño frente a simuladores comerciales o modelos biológicos. Estas líneas de trabajo permitirán fortalecer la aplicabilidad de la metodología propuesta como herramienta de apoyo para la formación médica especializada y para el desarrollo de tecnologías nacionales en simulación quirúrgica robótica.

# 4. Conclusiones

El desarrollo de robots quirúrgicos constituye un campo de innovación científica y tecnológica de carácter multidisciplinario, en el que convergen la medicina, la robótica, la ingeniería biomédica y el diseño de dispositivos médicos. Sin embargo, el uso seguro y eficiente de estos sistemas requiere que los médicos especialistas reciban entrenamiento especializado y certificación previa a su aplicación en pacientes. En este proceso, los simuladores, modelos anatómicos, órganos artificiales y tejidos sintéticos desempeñan un papel crucial, ya que permiten adquirir habilidades técnicas, mejorar la coordinación mano-ojo, desarrollar sensibilidad táctil indirecta y practicar maniobras quirúrgicas complejas en entornos controlados.

No obstante, los equipos de cirugía robótica de alta especialidad son escasos, costosos y requieren mantenimiento especializado, lo que limita su disponibilidad y, en consecuencia, también restringe el número de cirujanos capacitados y certificados en México. Por ello, resulta necesario desarrollar órganos artificiales que sean capaces de emular su fisiología, anatomía, morfología y propiedades mecánicas similares a las de los órganos naturales, que permitan fortalecer las habilidades quirúrgicas de los especialistas y ampliar las oportuni-

dades de entrenamiento en cirugía robótica mínimamente invasiva.

La metodología propuesta permite el diseño y fabricación de modelos anatómicos de bajo costo para ser utilizados como órganos artificiales durante los entrenamientos de cirugía robótica de médicos especialistas, los cuales cumplen con los requerimientos anatómicos necesarios para adquirir sensibilidad táctil, control, realizar procedimientos de suturas, cortes precisos durante la operación con los brazos robóticos. Los modelos físicos artificiales presentados replican la arquitectura tridimensional compleja del riñón e intestino humano para que ayuden a desarrollar habilidades de precisión, coordinación y toma de decisiones en procedimientos mínimamente invasivos, como en la nefrectomía o laparoscopia, respectivamente.

La contribución central de este trabajo radica en la integración de la impresión 3D con un proceso posterior de fabricación artesanal a escala 1:1 que permitió reproducir formas complejas con fidelidad geométrica, orientado a obtener órganos artificiales con propiedades mecánicas aproximadas a las de los órganos naturales, particularmente en términos de flexibilidad, deformabilidad y módulo de elasticidad. Para ello, se emplearon materiales flexibles, como siliconas y elastómeros, capaces de reproducir geometrías anatómicas complejas y permitir una manipulación adecuada durante prácticas quirúrgicas asistidas por robot. Asimismo, la incorporación de elementos funcionales, como cámaras de aire permiten transformar algunos modelos en estructuras no solo anatómicas, sino también dinámicas, con mayor potencial para simular condiciones de manipulación quirúrgica.

En conjunto, el procedimiento propuesto evidencia la importancia de la precisión geométrica, la selección adecuada de materiales y la correcta secuencia de fabricación para obtener órganos artificiales fieles, manipulables y funcionales. La fabricación artesanal con silicón representa una alternativa viable para el desarrollo de prototipos anatómicos personalizados, especialmente en etapas tempranas de investigación, validación y entrenamiento. Además, esta metodología puede adaptarse al desarrollo de otros órganos, maniquíes o piezas con morfologías irregulares, de acuerdo con las necesidades específicas de simulación médica.

## Direcciones de correo de los autores

Dávila-Vilchis, J. M.: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
mdavilav@uaemex.mx

Zúñiga Avilés, L. A.*: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
lazunigaa@uaemex.mx

Ramírez-García, N. G.: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
gngarciar@uaemex.mx

Vilchis-González, A. H.: Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, Estado de México, México
avilchisg@uaemex.mx

# Referencias


[1] W. Rosell Puig, B. González Fano, C. Cué Mourelos y C. Dovale Borjas, «Organización de los sistemas orgánicos del cuerpo humano para facilitar su estudio,» *Educación Médica Superior*, vol. 18, n.º 3, págs. 1-1, 2004.

[2] R. DeSalle e I. Tattersall, «El cerebro,» *Big bangs, comportamientos y creen-cias. Barcelona, Galaxia Gutenberg*, 2017.

[3] R. F. Ventura, D. N. Navarro, J. Otero e I. Almendros, «Pulmones biofabricados:: un desafío y una oportunidad,» *Archivos de bronconeumología: Organo oficial de la Sociedad Española de Neumología y Cirugía Torácica SEPAR y la Asociación Latinoamericana de Tórax (ALAT)*, vol. 54, n.º 1, págs. 31-38, 2018.

[4] Z. Ahmed et al., «Liver function tests in identifying patients with liver disease,» *Clinical and experimental gastroenterology*, págs. 301-307, 2018.

[5] M. Israelsen, S. Francque, E. A. Tsochatzis y A. Krag, «Steatotic liver disease,» *The Lancet*, vol. 404, n.º 10464, págs. 1761-1778, 2024.

[6] W. Otero R y F. Sierra A, «El hígado en cirugía,» *Revista colombiana de Gastroenterología*, vol. 18, págs. 230-238, 2003.

[7] N. Özkaya, D. Leger, D. Goldsheyder y M. Nordin, «Mechanical properties of biological tissues,» en *Fundamentals of biomechanics: equilibrium, motion, and deformation*, Springer, 2016, págs. 361-387.

[8] C. N. de Trasplantes, «Estadísticas sobre donación y trasplantes. Secretaría de Salud, Gobierno de México,» 2026. dirección: https://www.gob.mx/cenatra/documentos/estadisticas-50060

[9] A. E. Pereira Cuns, «Revisión bibliográfica: Fabricación de órganos artificiales,» 2022.

[10] J. E. M. Díaz, «Tecnologías disruptivas como alternativa a la obtención de órganos y tejidos artificiales,» *Revista Colombiana de Bioética*, vol. 15, n.º 1, 2020.

[11] C. Mandrycky, Z. Wang, K. Kim y D.-H. Kim, «3D bioprinting for engineering complex tissues,» *Biotechnology advances*, vol. 34, n.º 4, págs. 422-434, 2016.

[12] F. Iqbal et al., «Continuum and soft robots in minimally invasive surgery: A systematic review,» *IEEE Access*, vol. 13, págs. 24 053-24 079, 2025.

[13] D. T. Vu et al., «Soft Robotics and Advanced Technologies for Minimally Invasive Bioprinting: The Future of Internal Organ Repair,» *Advanced Science*, vol. 13, n.º 22, e23532, 2026.

[14] F. M. León Martínez, P. A. Benenaula Cuzco, P. J. Encalada Loayza y K. E. León Sagbay, «Robotic surgery: a look into the future of surgical intervention,» *Horizonte sanitario*, vol. 24, n.º 2, págs. 346-362, 2025.

[15] K. Qiu et al., «3D printed organ models with physical properties of tissue and integrated sensors,» *Advanced materials technologies*, vol. 3, n.º 3, pág. 1 700 235, 2018.

[16] Z. Jin et al., «3D printing of physical organ models: recent developments and challenges,» *Advanced Science*, vol. 8, n.º 17, pág. 2 101 394, 2021.

[17] J. S. S. Matute, A. E. M. Meza, D. C. Sánchez, D. T. Puentes y S. C. C. Montero, «Avances en la cirugía robótica, una revisión sistemática enfocada en Cirugía General,» *Scientific & Education Medical Journal:(SEMJ)*, vol. 2, n.º 2, págs. 59-70, 2022.

[18] E. G. Rodríguez, N. C. Sánchez, C. López-Casado e I. R. Blanco, «Consola de teleoperación de realidad mixta para sistemas quirúrgicos robóticos,» *Jornadas de Automática*, n.º 46, 2025.

[19] L. Bianchi, F. Cavarzan, L. Ciampitti, M. Cremonesi, F. Grilli y P. Saccomandi, «Thermophysical and mechanical properties of biological tissues as a function of temperature: a systematic literature review,» *International Journal of Hyperthermia*, vol. 39, n.º 1, págs. 297-340, 2022.