

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y
APLICADAS

CENTRO DE INVESTIGACIÓN EN INGENIERÍA Y CIENCIAS
APLICADAS

MEJORA EN LA PREDICCIÓN DEL COP EN BOMBAS DE CALOR POR
ABSORCIÓN MEDIANTE EL USO DE NUEVAS FUNCIONES DE
TRANSFERENCIA DE ORDEN ENTERO, FRACCIONARIO Y
CONFORMABLE EN REDES NEURONALES ARTIFICIALES

TESIS PROFESIONAL PARA OBTENER EL GRADO DE:
MAESTRÍA EN INGENIERÍA Y CIENCIAS APLICADAS

PRESENTA:
GERARDO MANRÍQUEZ PARDO

DIRECTORA: DR. JOSÉ ALFREDO HERNÁNDEZ PÉREZ
CO-DIRECTOR: DR. JOSÉ FRANCISCO GÓMEZ AGUILAR

SINODALES:
DR. ALBERTO ARMANDO ÁLVAREZ GALLEGOS
DR. ARMANDO HUICOCHEA RODRÍGUEZ
DR. ABISAI ACEVEDO QUIROZ



Instituto de
Investigación en
Ciencias
Básicas y
Aplicadas

CUERNAVACA, MORELOS

SEPTIEMBRE, 2026

Dedicatoria

A mi mamá, que ha estado a mi lado desde que tengo memoria y nunca me ha soltado la mano.

*A mi papá, que siempre me ha apoyado y ha buscado incansablemente sacar la mejor versión de
mí.*

*A mi única hermana; aunque somos diferentes, siempre logras sorprenderme. Sé que tienes el
potencial para lograr cosas grandes e increíbles.*

"Si caminas solo llegarás más rápido,
si caminas acompañado
llegarás más lejos."
Proverbio africano.

Agradecimientos

A la Universidad Autónoma del Estado de Morelos (UAEM) y al Centro de Investigación en Ingeniería y Ciencias Aplicadas (CIICAp), por ser mi casa de estudios y brindarme las herramientas para mi formación profesional.

A la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI, anteriormente CONAHCYT), por el invaluable apoyo otorgado a través de la beca de posgrado, sin el cual no hubiese sido posible llevar a cabo este trabajo de investigación.

A mis directores de tesis, el Dr. José Alfredo Hernández Pérez y el Dr. José Francisco Gómez Aguilar por su amistad, enseñanzas, consejos y acompañamiento; gracias por la confianza depositada en mí y por permitirme trabajar a su lado.

Al Dr. Alberto Álvarez Gallegos, por su incondicional amistad, su constante apoyo, sus valiosos consejos y su acompañamiento a lo largo de mi formación.

Al Dr. Armando Huicochea Rodríguez, por brindarme su confianza, su amistad y sus sabios consejos. Gracias por ser una excelente persona conmigo, por tu acompañamiento y por buscar siempre sacar lo mejor de mí.

Al Dr. Abisaí Acevedo Quiroz, por sus enseñanzas, sus consejos, su amistad y su apoyo incondicional a lo largo de este proceso.

Al Dr. Jesús Emmanuel Solís Pérez, por su invaluable apoyo y orientación en el desarrollo de este trabajo.

Al Dr. Miguel Ángel Basurto Pensado, porque, aunque su nombre no aparezca formalmente en este trabajo, él es la razón principal por la cual decidí estudiar en esta institución. Ha sido un guía fundamental en los últimos años de mi vida y lo aprecio profundamente.

A todos los amigos y compañeros que conocí en este camino y que hicieron de esta etapa de mi vida una experiencia sumamente disfrutable. Ha sido un trayecto muy bonito y agradable, lleno de sabidurías, tropiezos, acompañamiento y confianza. Los quiero mucho a todos, gracias por formar parte de mi vida.

Contenido

1.	Introducción	18
2.	Marco teórico	21
2.1.	Energía	21
2.1.1.	Energía primaria y secundaria.....	22
2.1.2.	Energía renovable y no renovable	22
2.1.3	Energía Térmica.....	22
2.2.	Inteligencia Artificial	23
2.2.1.	Las Redes Neuronales Artificiales (RNA).....	25
2.2.2.	Funcionamiento teórico de las RNA	28
2.2.3.	Funciones de Transferencia	30
2.2.4.	Cálculo Fraccionario.....	34
2.2.5.	Cálculo Conformable	38
2.3.	Fundamentos de energía y sistemas térmicos de transformación.....	40
2.3.1	El desafío del calor residual industrial	41
2.3.2	Bombas de Calor por Absorción para calentamiento	41
2.3.3	Variables termodinámicas de funcionamiento para el sistema de bomba de calor por absorción	43
2.3.4	Descripción General del Ciclo de Absorción	45

2.3.5	COP (Coeficiente de Rendimiento)	46
2.4.	El reto del modelado termodinámico y el uso de Inteligencia Artificial	49
2.4.1.	Limitaciones operativas de los modelos analíticos y deterministas	50
2.4.2.	El paradigma de las Redes Neuronales Artificiales como aproximadores universales	51
2.5.	Complejidad física y operativa en sistemas de absorción	52
2.5.1.	Acoplamiento multivariable y dinámica transitoria del sistema	52
2.5.2.	Irreversibilidades térmicas y el riesgo crítico de cristalización	53
2.6.	Transición hacia el cálculo fraccionario y conformable	54
2.6.1.	Límite del orden entero y el costo de la memoria no local	55
2.6.2.	Eficiencia computacional y analítica del cálculo conformable	56
2.7	Avances contemporáneos en la aplicación de redes neuronales artificiales y cálculo fraccionario en el modelado de sistemas térmicos	57
3.	Antecedentes.	61
3.1.	Problemática	64
4.	Hipótesis	64
5.	Objetivo general	65
5.1	Objetivos específicos	65
6.	Metodología	65
7.	Aplicación de la metodología	66

7.1.	Arquitectura y topología de la red neuronal	68
7.1.1.	Topología de flujo unidireccional (Feed-Forward)	69
7.1.2.	Configuración 16:2:1	70
7.1.3.	Dinámica del cálculo sináptico.....	71
7.2.	Algoritmo de optimización	72
7.2.1.	Limitaciones de los métodos de primer y segundo orden puros.....	73
7.2.2.	Formulación híbrida: La interpolación de Levenberg y Marquardt	74
7.2.3.	Justificación técnica en el modelado de orden no entero	75
7.3.	Normalización de las variables	76
7.4.	Análisis de Sensibilidad de Variables de Entrada	77
7.4.1.	Metodología de Análisis.....	77
7.4.2.	Resultados Completos del Análisis.....	78
7.4.3.	Interpretación de resultados	80
7.5.	Métricas estadísticas	81
7.5.1.	Criterios de dispersión y tendencia direccional	83
7.5.2.	Validación paramétrica: Prueba de linealidad de pendiente e intercepto ...	84
7.6.	Funciones de transferencia a evaluar	85
8.	Resultados de la simulación y análisis comparativo	89
8.1	Modelos de orden entero	90
8.2	Modelos de orden fraccionario:.....	96

8.3 Modelos de orden conformable:	104
8.4 Codificación en Python.....	109
8.5 Análisis Comparativo del Flujo de Trabajo: MATLAB vs. Python	111
8.5.1 Normalización de las Variables	112
8.5.2 Segmentación de Datos y Nuevo Esquema (70/15/15)	113
8.5.3 Mantenimiento del Algoritmo de Optimización (Levenberg-Marquardt)	114
8.5.4 Conclusiones de la Migración de lenguaje de programación.....	114
8.6 Análisis comparativo de resultados previos y modelos propuestos	115
9. Discusión	119
9.1. Evolución de las arquitecturas neuronales en la predicción del COP para sistemas de absorción	119
9.2. Análisis de la normalización y preprocesamiento de datos	121
9.3. La inicialización acotada de parámetros	122
9.4. Evaluación de los modelos de orden entero	123
9.5. El cálculo fraccionario como herramienta de modelado con memoria	125
9.6. El cálculo conformable como alternativa local.....	126
9.7. Implicaciones de la migración a Python y herramientas de código abierto	128
9.8. Comparación con trabajos previos.....	129
9.9. Implicaciones para el modelado de sistemas térmicos complejos	130
9.10. Limitaciones y líneas de investigación futuras	131

10. Conclusiones	132
11. Limitaciones de los resultados	135
12. Trabajos Futuros	136
Referencias	138
ANEXOS	150
Anexo A. Reconocimiento grupal del 5.º Congreso CITCA 2025	153
Anexo B. Reconocimiento personal del 5.º Congreso CITCA 2025.....	155
Anexo C. Certificado del 15th ISTEf (España).....	157
Anexo D. Artículo en Extenso: Artificial Intelligence Applied To Absorption Heat Pumps.....	158
Anexo E. Artículo Corto: Evaluation of different neural networks for the thermodynamic	164

Tabla de Abreviaturas y Acrónimos

Abreviatura	Significado
<i>AHT</i>	Transformador de Calor por Absorción (Absorption Heat Transformer)
<i>IA</i>	Inteligencia Artificial
<i>RNA</i>	Redes Neuronales Artificiales
<i>DL</i>	Aprendizaje Profundo (Deep Learning)
<i>LiBr – H₂O</i>	Bromuro de litio y agua (Mezcla de trabajo)
<i>MTBE</i>	Metil-Terbutil Éter
<i>SHX</i>	Intercambiador de calor de soluciones (Solution Heat Exchanger)
<i>PLC</i>	Controlador Lógico Programable (Programmable Logic Controller)
<i>LOGSIG</i>	Función de transferencia sigmoide logística
<i>TANSIG</i>	Función de transferencia tangente hiperbólica
<i>ReLU</i>	Unidad Lineal Rectificada (Rectified Linear Unit)
<i>SiLU</i>	Unidad Lineal Sigmoide (Sigmoid Linear Unit / Swish)

<i>dSiLU</i>	Derivada de la Unidad Lineal Sigmoide
<i>FOPID</i>	Controlador Proporcional Integral Derivativo de Orden Fraccionario

Tabla de Nomenclatura

Símbolo	Descripción	Unidades
T	Temperatura	[°C] o [K]
P	Presión	[kPa] o [bar]
X	Concentración másica de la solución salina	[%] o [fracción]
$\dot{m}$	Flujo másico	[kg/s]
h	Entalpía específica	[kJ/kg]
Q	Capacidad térmica o calor transferido	[kW]
COP	Coefficiente de Rendimiento (Coefficient of Performance)	Adimensional
W	Peso sináptico de la Red Neuronal	Adimensional
b	Término de sesgo (bias) de la Red Neuronal	Adimensional
θ	Umbral de activación	Unidades
ϕ o φ	Función de activación o de transferencia	Adimensional
Φ^+	Matriz pseudoinversa de la dinámica generada	Adimensional

V	Pesos sinápticos de la capa de salida	Adimensional
Y	Vector definido por etiquetas, clases o salida normalizada	Adimensional
α	Orden de la derivada fraccionaria o conformable $(n - 1 < \alpha < n)$	Adimensional
t	Tiempo actual de evaluación	[s]
τ	Variable de integración (memoria histórica temporal)	[s]
Γ	Función matemática Gamma	Adimensional
D_t^α	Operador diferencial fraccionario de Caputo	-
T_α	Operador diferencial conformable local	-
I	Intercepto de la recta de regresión	Adimensional
S	Pendiente de la recta de regresión	Adimensional

Tabla de subíndices

Subíndice	Descripción
<i>in</i>	Entrada del flujo al componente térmico
<i>out</i>	Salida del flujo del componente térmico
<i>AB</i>	Absorbedor
<i>GE</i>	Generador
<i>CO</i>	Condensador
<i>EV</i>	Evaporador
<i>min</i>	Valor mínimo de la variable (utilizado en normalización)
<i>max</i>	Valor máximo de la variable (utilizado en normalización)
<i>n</i>	Identificador numérico de capa, iteración, peso o nodo (1, 2, 3...)
1,2...10	Puntos de estado termodinámico de flujo en el diagrama del sistema
<i>inf</i>	Límite inferior del intervalo de confianza
<i>sup</i>	Límite superior del intervalo de confianza

Resumen

La predicción precisa del Coeficiente de Rendimiento (COP) en los transformadores de calor por absorción sigue siendo un desafío crítico para el avance de la ingeniería energética y los sistemas térmicos sostenibles. En consecuencia, este estudio se enfoca en mejorar la exactitud y confiabilidad de la predicción del COP mediante un análisis comparativo exhaustivo de funciones de transferencia de orden entero, fraccionario y conformable dentro del marco de las Redes Neuronales Artificiales (RNA).

Para optimizar el proceso de estimación, se evaluó un conjunto diverso de funciones, incluyendo Logsig, Tansig, Swish, ELU, GELU y dSiLU, junto con sus respectivas variantes fraccionarias y conformables. Mientras que los intentos de modelado previos con arquitecturas estándar (las cuales incorporaban 16 variables de entrada y capas ocultas de 3 o más neuronas) fracasaban frecuentemente al intentar satisfacer criterios rigurosos de confiabilidad como la prueba de linealidad de pendiente e intercepto, esta investigación resuelve dichas limitaciones a través de un refinamiento arquitectónico y matemático.

La experimentación empleó una base de datos derivada de observaciones empíricas en sistemas de bombas de calor por absorción, con un entrenamiento optimizado en Python. Los hallazgos revelaron que las redes neuronales de orden no entero, operando con una arquitectura reducida de 16:2:1, superan significativamente a los modelos tradicionales. Específicamente, la función de transferencia fraccionaria dSiLU basada en la serie de Mittag-Leffler (dSiLU-ML) alcanzó una precisión predictiva excepcional, produciendo un coeficiente de correlación (R) de 0.99856 y un Error Cuadrático Medio (RMSE) de 0.00291. Por su parte, la variante conformable (cdSiLU) demostró un rendimiento altamente competitivo, obteniendo un R de 0.99857 y un RMSE de 0.00963.

Este estudio demuestra que la integración de funciones de transferencia fraccionarias y conformables logra un equilibrio perfecto entre la simplicidad del modelo y una estricta fiabilidad predictiva, estableciendo un nuevo y robusto estándar metodológico para el modelado de transformaciones térmicas complejas en aplicaciones energéticas.

Abstract

The precise prediction of the Coefficient of Performance (COP) in absorption heat transformers remains a critical challenge for the advancement of energy engineering and sustainable thermal systems. Consequently, this study focuses on improving the accuracy and reliability of COP prediction through a comprehensive comparative analysis of integer, fractional, and conformable-order transfer functions within the framework of Artificial Neural Networks (ANNs).

To optimize the estimation process, a diverse set of functions was evaluated, including Logsig, Tansig, Swish, ELU, GELU, and dSiLU, along with their respective fractional and conformable variants. While previous modeling attempts with standard architectures (which incorporated 16 input variables and hidden layers of 3 or more neurons) frequently failed to satisfy rigorous reliability criteria such as the slope-intercept linearity test, this research resolves these limitations through architectural and mathematical refinement.

The experimentation employed a dataset derived from empirical observations in absorption heat pump systems, with optimized training in Python. The findings revealed that non-integer order neural networks, operating with a reduced 16:2:1 architecture, significantly outperform traditional models. Specifically, the fractional transfer function dSiLU based on the Mittag-Leffler series (dSiLU-ML) achieved exceptional predictive precision, yielding a correlation coefficient (R) of 0.99856 and a Root Mean Square Error (RMSE) of 0.00291. Meanwhile, the conformable variant (cdSiLU) demonstrated highly competitive performance, obtaining an R of 0.99857 and an RMSE of 0.00963.

This study demonstrates that the integration of fractional and conformable transfer functions achieves a perfect balance between model simplicity and strict predictive reliability,

establishing a robust new methodological standard for modeling complex thermal transformations
in energy applications.

1. Introducción

El constante crecimiento de la demanda energética a nivel global y la imperante urgencia de mitigar los impactos ambientales derivados del uso de combustibles fósiles han posicionado a la eficiencia térmica como una prioridad en la ingeniería moderna. En el sector industrial, una cantidad masiva de energía se desperdicia diariamente en forma de calor residual de baja temperatura. Para contrarrestar esta pérdida y promover la sostenibilidad, las tecnologías de recuperación y reciclaje energético se han vuelto indispensables. Dentro de estas tecnologías, las bombas de calor por absorción (particularmente los transformadores de calor que operan con mezclas salinas como el bromuro de litio y agua, $\text{LiBr-H}_2\text{O}$) destacan por su capacidad para recibir calor residual y elevarlo a niveles de temperatura térmicamente útiles para otros procesos, minimizando drásticamente el consumo de energía eléctrica. El indicador fundamental para evaluar la eficiencia y viabilidad operativa de estos equipos es el Coeficiente de Rendimiento (COP), el cual cuantifica la relación entre el calor útil generado y la energía térmica suministrada al sistema.

A pesar de las ventajas operativas de las bombas de calor por absorción, predecir su COP en condiciones reales de funcionamiento representa un desafío analítico de altísima complejidad. El ciclo termodinámico de estos equipos está fuertemente acoplado; esto significa que cualquier perturbación transitoria en las temperaturas o presiones de entrada genera un efecto dominó altamente no lineal en los procesos de transferencia de masa y energía dentro del absorbedor, el generador, el evaporador y el condensador. Además, el sistema está restringido por límites operativos críticos, como el riesgo de cristalización de la sal de bromuro de litio ante caídas abruptas de temperatura, lo que puede provocar bloqueos mecánicos severos. Debido a estas irreversibilidades y a la naturaleza estocástica del calor residual, los modelos matemáticos

analíticos y deterministas resultan insuficientes. Al depender de sistemas de ecuaciones diferenciales rígidas y suposiciones de estado estacionario, estos métodos clásicos exigen un costo computacional masivo y tienden a divergir, volviéndose inviables para el monitoreo y control de la planta en tiempo real.

Como respuesta a las limitaciones del modelado determinista, el uso de la Inteligencia Artificial (IA), y específicamente de las Redes Neuronales Artificiales (RNA), ha emergido como un paradigma revolucionario en la ingeniería térmica. Basadas en el Teorema de Aproximación Universal, las arquitecturas de aprendizaje profundo de tipo Feed-Forward son capaces de actuar como aproximadores universales de "caja negra". Estas redes aprenden la fenomenología del equipo directamente a partir de vastos conjuntos de datos experimentales multivariantes (incluyendo temperaturas, presiones y concentraciones), evadiendo la necesidad de resolver ecuaciones diferenciales de forma explícita.

Sin embargo, tras casi dos décadas de desarrollo en el modelado neuronal de sistemas térmicos, la literatura científica ha evidenciado una barrera estadística persistente. Los modelos neuronales convencionales, contruidos sobre matemáticas de orden entero y equipados con funciones de transferencia clásicas como la sigmoide logística (LOGSIG), la tangente hiperbólica (TANSIG) o la Unidad Lineal Rectificada (ReLU), han alcanzado su límite de eficiencia. Aunque estas arquitecturas logran coeficientes de correlación aparentemente altos, sufren de fallas algorítmicas severas como el "desvanecimiento del gradiente" o la aparición de "neuronas muertas" cuando se enfrentan a entradas altamente ruidosas o negativas. En consecuencia, estos modelos de orden entero fallan sistemáticamente al someterse a pruebas rigurosas de linealidad (pendiente e intercepto) en los límites de operación del ciclo de absorción, lo que compromete su confiabilidad técnica para aplicaciones industriales.

Es en este contexto de necesidad algorítmica donde se inserta el núcleo científico y la principal aportación de la presente investigación. El presente trabajo de investigación tiene como objetivo fundamental mejorar drásticamente la predicción del Coeficiente de Rendimiento (COP) en un sistema de bomba de calor por absorción mediante la innovación de su núcleo matemático. Para lograrlo, se propone abandonar la rigidez de las funciones de activación clásicas e implementar nuevas funciones de transferencia fundamentadas en el cálculo de orden fraccionario y el cálculo conformable.

El cálculo fraccionario, a través de herramientas no locales como la función de generalización de Mittag-Leffler, introduce un efecto de "memoria histórica" en el algoritmo, permitiendo a la red asimilar las dinámicas transitorias y la inercia del ciclo de absorción sin saturar sus gradientes de aprendizaje. Paralelamente, la derivada local conformable dota a la arquitec

tura de una plasticidad analítica pura, garantizando el estricto cumplimiento de la regla de la cadena matemática. Esta compatibilidad algebraica es crucial, ya que permite que el algoritmo de optimización de segundo orden de Levenberg-Marquardt actualice los pesos sinápticos con una precisión y velocidad formidables.

Específicamente, este estudio evalúa la integración de la derivada de la Unidad Lineal Sigmoide (dSiLU) modificada con operadores de orden no entero (dSiLU-ML y cdSiLU). Esta fusión otorga a las neuronas un perfil de activación en forma de campana asimétrica, autoajutable y altamente sensible a los regímenes operativos seguros de la mezcla LiBr-H₂O. Además, para garantizar la escalabilidad técnica del proyecto y optimizar el consumo de recursos de hardware, la metodología incorpora una transición integral del entorno de entrenamiento hacia el lenguaje de programación Python, asegurando un procesamiento ágil y una segmentación estadística robusta de los datos.

En resumen, este trabajo pretende demostrar que la incorporación del cálculo fraccionario y conformable en arquitecturas neuronales ultraligeras (de solo dos neuronas ocultas) no solo reduce significativamente el Error Cuadrático Medio de las estimaciones, sino que consigue, por primera vez, superar las más estrictas pruebas estadísticas de linealidad. El éxito de estos algoritmos híbridos permitirá consolidar un sensor virtual predictivo altamente preciso, abriendo nuevas fronteras para la eficiencia energética y el control automatizado de procesos térmicos sostenibles.

2. Marco teórico

2.1. Energía

En el área de la tecnología y la economía, la energía se refiere a los recursos naturales que pueden ser explotados para generar trabajo o funcionar como intermediarios en procesos de transformación. Desde la perspectiva de la física y la termodinámica, la energía se define de manera más estricta como la capacidad que posee un cuerpo o sistema para realizar un trabajo, producir un cambio o transferir calor (Çengel & Boles, 2015).

En la sociedad moderna, la energía es el motor que sostiene desde las actividades cotidianas hasta los procesos industriales más complejos. De acuerdo con la primera ley de la termodinámica, la energía no se crea ni se destruye, únicamente se transforma. Para facilitar su estudio y aprovechamiento, los recursos energéticos naturales suelen clasificarse por su nivel de procesamiento y por su disponibilidad a largo plazo.

La energía tiene diversas clasificaciones, siendo las más comunes:

- a) Por su procedencia: renovables y no renovables.
- b) Por su procesamiento: primaria y secundaria.

2.1.1. Energía primaria y secundaria

Por su nivel de procesamiento, los recursos se dividen en dos grandes grupos.

Energía primaria

Esta clasificación de energía se refiere a aquella que es extraída directamente de la Tierra, considerando que los procesos por los que atraviesa sean solamente para limpiar impurezas. Entre la energía primaria se enlista: el petróleo crudo, gas natural, carbón, nuclear y renovables.

Energía secundaria

La energía secundaria es aquella que se obtiene por medio de procesamiento de la energía primaria, suele considerarse un producto de esta; por ejemplo: petrolíferos, gas seco y electricidad.

2.1.2. Energía renovable y no renovable

Es considerada como energía renovable toda aquella definida como inagotable a pesar de su creciente explotación.

La energía no renovable es aquella que se cuantifica en reservas, debido a que su explotación conlleva el agotamiento de esta.

2.1.3 Energía Térmica

La energía térmica es aquella que se almacena en un sistema debido al movimiento de sus partículas, y es esta vibración de átomos y moléculas la que se percibe como calor. La energía

térmica no solo es esencial para el funcionamiento de sistemas de calefacción y enfriamiento en hogares, sino que también es clave en la industria, donde se usa para realizar procesos de transformación y manufactura a través del calentamiento y enfriamiento de materiales. Se mide en unidades como el joule o la caloría y se puede transferir entre cuerpos de distintas maneras: por conducción, convección o radiación.

- **Conducción:** Es la forma en que el calor viaja a través de un sólido. Cuando una parte del material se calienta, sus partículas comienzan a moverse más rápido y a transmitir esa energía a las partículas cercanas.
- **Convección:** Ocurre en fluidos como líquidos o gases, donde el calor se transfiere al movimiento de las partículas del fluido mismo, que suben o bajan según su temperatura.
- **Radiación:** Esta transferencia ocurre sin necesidad de contacto entre cuerpos. Un ejemplo clásico es el calor del sol, que viaja a través del espacio vacío hasta llegar a la Tierra.

La energía térmica es muy valiosa en la industria moderna, donde se aprovecha para tareas que van desde la calefacción de edificios y el procesamiento de materiales, hasta la generación de electricidad en plantas termoeléctricas.

2.2. Inteligencia Artificial

La Inteligencia Artificial (IA) es una rama de la programación que busca dotar a las computadoras con la capacidad de imitar la inteligencia humana. Su propósito fundamental es mejorar y ampliar las habilidades que normalmente asociamos con los seres humanos. Este concepto tuvo sus inicios en una conferencia llevada a cabo en 1956 en la Universidad de Dartmouth, liderada por John McCarthy. Desde entonces, la IA ha evolucionado enormemente,

convirtiéndose en un campo de estudio dinámico y en constante crecimiento. (McCarthy, Minsky, Papert, & Shannon, 1956).

La esencia de la Inteligencia Artificial radica en entrenar sistemas computacionales para analizar conjuntos de datos y extraer patrones subyacentes. Mediante iteraciones y algoritmos de optimización, estos modelos logran encontrar relaciones no lineales y soluciones a problemas diversos, convirtiendo a la IA en una herramienta fundamental para abordar dinámicas de alta complejidad (Russell & Norvig, 2016).

Cuando hablamos de IA, nos enfrentamos a un vasto terreno que se subdivide en diferentes categorías o ramas. En este contexto, nos centraremos específicamente en una subcategoría crucial: el aprendizaje profundo (DL), también conocido como Deep Learning en inglés. Dentro de esta categoría, encontramos las Redes Neuronales Artificiales, un modelo inspirado en la estructura y funcionamiento de nuestro propio cerebro.

El aprendizaje profundo implica que las máquinas puedan aprender a través de capas de procesamiento, como si fueran capas de abstracción, permitiéndoles entender características cada vez más complejas de los datos. Las Redes Neuronales Artificiales, por su parte, imitan la forma en que operan las neuronas en nuestro cerebro para procesar información y tomar decisiones. (LeCun, Bengio, & Haffner, 1998).

La aplicación de estas tecnologías no solo busca replicar la inteligencia humana, sino también mejorar la eficiencia en la resolución de problemas. La IA, y en particular el aprendizaje profundo, se ha convertido en un campo de investigación crucial para abordar desafíos en áreas tan diversas como la medicina, la robótica, la industria y muchos más. (Chollet, 2018).

En resumen, la Inteligencia Artificial, con su enfoque en el aprendizaje profundo y las Redes Neuronales Artificiales, representa una frontera en la evolución de la tecnología. Al permitir

que las máquinas aprendan y comprendan de manera más sofisticada, estamos abriendo la puerta a soluciones innovadoras y a una revolución en la forma en que interactuamos con la tecnología y abordamos problemas complejos en la sociedad moderna.

2.2.1. Las Redes Neuronales Artificiales (RNA)

Las Redes Neuronales Artificiales (RNA) representan un modelo computacional y matemático diseñado para simular el funcionamiento de las redes neuronales biológicas, encontrando su inspiración en la forma en que operan estas últimas. La base de su diseño se centra en imitar el aprendizaje a partir de la experiencia, la capacidad de generalizar información de ejemplos previos a situaciones nuevas, y el análisis de las características fundamentales de conjuntos de datos (Basogain Olabe, 2020). La Figura 1 ilustra cómo el Aprendizaje Automático y el Aprendizaje Profundo se estructuran como subconjuntos de la Inteligencia Artificial.

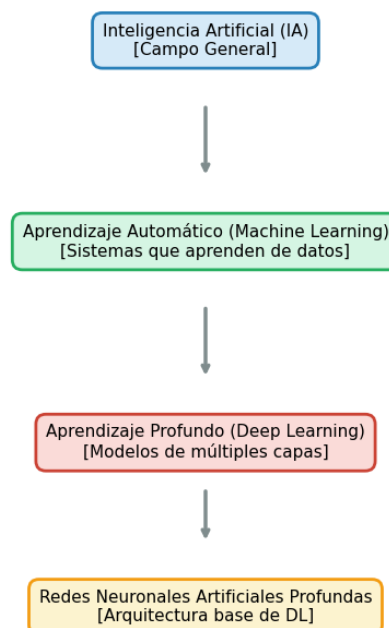


Figura 1. Clasificación de la inteligencia artificial

La estructura básica de las RNA sigue una metodología que emplea unidades fundamentales de procesamiento conocidas como "neuronas". Estas neuronas están interconectadas entre sí, y cada una recibe un dato de entrada. Cada entrada es asignada un "peso", que determina la influencia que tendrá en la salida del modelo (Parrales et al., 2018). Este proceso refleja la forma en que las conexiones sinápticas entre neuronas biológicas ajustan su fuerza en función de la relevancia de la información. (Tablada & Torres, 2020).

A pesar de que el concepto de RNA como modelo computacional tiene sus raíces en el año 1943, ha experimentado un aumento significativo en popularidad durante la última década (Abeliuk & Gutiérrez, 2020). Su aplicación se extiende a diversos campos, incluyendo la Ingeniería, Medicina, Administración y Leyes, demostrando su versatilidad y capacidad para abordar una amplia gama de problemas y desafíos en distintas disciplinas. (Franco Guzmán, 2023).

El auge de las RNA se debe en gran medida a su capacidad para aprender patrones complejos y adaptarse a situaciones diversas, convirtiéndolas en herramientas valiosas en la toma de decisiones y el análisis de datos. Su capacidad para procesar información de manera similar a las redes neuronales biológicas las hace particularmente efectivas en la identificación de tendencias, reconocimiento de patrones y toma de decisiones en entornos complejos. En resumen, las RNA han emergido como una herramienta esencial en la era de la inteligencia artificial, desempeñando un papel clave en la resolución de problemas y en la mejora de la eficiencia en una variedad de campos de aplicación.

Las redes neuronales artificiales intentan simular el cerebro humano; la unidad de cálculo en una red neuronal es la neurona, y esta puede utilizarse en diferentes tipos de arquitecturas. En la Figura 2 se muestra la arquitectura tipo perceptrón, la cual se describe en la ecuación (1):

$$y = \varphi \left(\sum_{n=1}^N w_n x_n + \theta \right), \quad (1)$$

$$\mathbf{V} = \Phi^+ \mathbf{Y},$$

donde X_n es la entrada de la red neuronal, W_n son los pesos sinápticos y θ son los umbrales, definidos de manera aleatoria para ampliar el co-dominio de las funciones de activación (φ). La Figura 2 representa el entrenamiento de los pesos sinápticos $\mathbf{V}$ de la capa de salida, donde Φ^+ es la matriz pseudoinversa de la dinámica generada por $\phi(W+\theta)$, y $\mathbf{Y}$ es un vector definido por las etiquetas o clases Y_n .

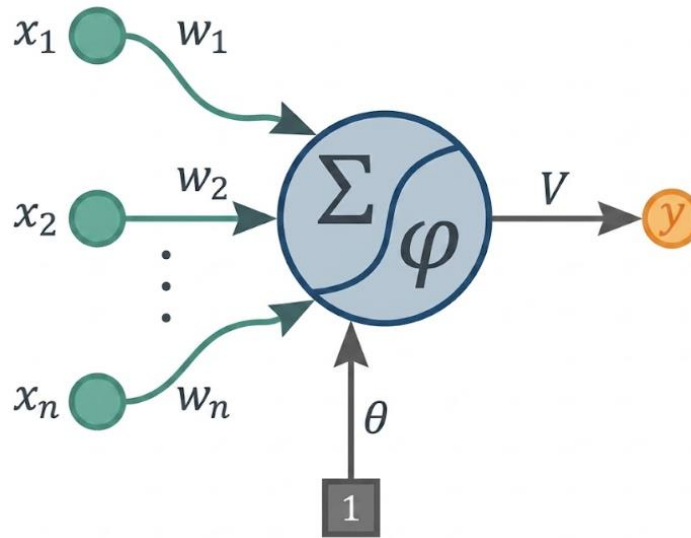


Figura 2. Diagrama de una neurona artificial con función de activación. Modificada a partir de Garrido

Alhama (2017).

2.2.2. Funcionamiento teórico de las RNA

De acuerdo con Tablada y Torres (2020), una RNA se divide en capas, tal como detalla la Figura 3:

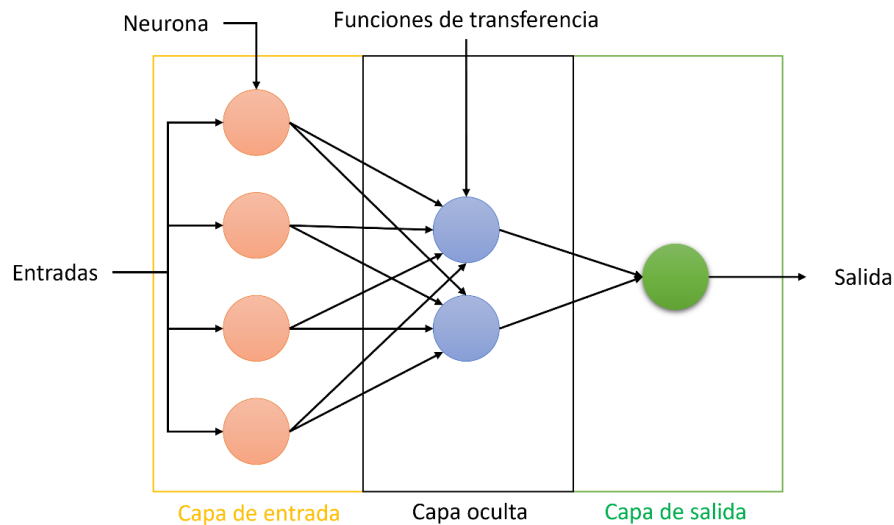


Figura 3. Arquitectura básica de una RNA. Modificada a partir de Cervantes-Bobadilla et al., (2021)

1. La primera capa se conoce como **“capa de entrada”** y en ella se encuentran los parámetros de entrada que tienen pertinencia en la salida o resultados que se espera obtener de la RNA.
2. La segunda capa, conocida como **“capa oculta”** será dónde se realizan los cálculos tomando las entradas.
3. La tercera y última capa, denominada **“capa de salida”**, representa el componente final de la arquitectura, cuya función es presentar los resultados del modelo."

El funcionamiento de las RNA es básicamente un cúmulo de iteraciones para balancear los pesos de las entradas y el término de sesgo, de modo que después de cada una de las iteraciones, el error entre los valores calculados y los valores modelados disminuya. (Haykin, 2009, pp. 59-60).

1. Se identifican e introducen los datos de entrada
2. Las variables de entrada deberán ser normalizadas para que el tiempo de cálculo disminuya, de acuerdo con la Ecuación (2) (Solís-Pérez et al., 2022).

$$Y_n = (Y_{max} - Y_{min}) * \frac{X - X_{min}}{X_{max} - X_{min}} + Y_{min} , \quad (2)$$

donde:

Y_n es la variable normalizada resultante.

X es el valor de la variable experimental de entrada a normalizar.

X_{min} y X_{max} representan los valores mínimo y máximo históricos registrados para dicha variable, respectivamente.

Y_{min} y Y_{max} definen el intervalo de normalización establecido por el usuario para optimizar el entrenamiento de la red.

3. Se proponen pesos para cada una de las entradas y un término de sesgo (bias), de acuerdo con la influencia que tienen sobre la(s) variable(s) de salida; todo esto ocurre dentro de la capa de entrada.
4. Cuando las variables de entrada han sido normalizadas y tienen pesos asignados, pasarán a la capa oculta, formando una especie de suma ponderada, como se muestra en la Ecuación (2) y después serán ingresadas a una función de transferencia (también conocida como función de activación). En la ecuación (3) se muestra cómo se realiza la suma ponderada de pesos, variables y bias.

$$N = W_1X_1 + W_2X_2 + \dots W_nX_n + b , \quad (3)$$

donde N es la suma ponderada o entrada neta a la función de transferencia, W_n = Peso de la variable, X_n = Variable de entrada, b = Sesgo o bias

5. Posteriormente, la RNA otorga una o más variables de salida de acuerdo con los cálculos que ha realizado.
6. Finalmente, se mide el error y con base en este y en la base de datos experimental, se entrena a la RNA por medio de una regla de entrenamiento hasta encontrar la precisión esperada.

2.2.3. Funciones de Transferencia

Derivado de que la suma ponderada que conforma los pesos, variables de entrada y sesgo es una función lineal, no es posible realizar un cálculo preciso debido a que la suma de dos o más rectas numéricas da como resultado otra recta. Para mitigar este problema, la suma ponderada pasa por una función de transferencia que le otorgará no-linealidad. (Haykin, 2009, pp. 122-125).

La esencia de la Inteligencia Artificial radica en entrenar sistemas computacionales para analizar conjuntos de datos y extraer patrones subyacentes. Mediante iteraciones y algoritmos de optimización, estos modelos logran encontrar relaciones no lineales y soluciones a problemas diversos, convirtiendo a la IA en una herramienta fundamental para abordar dinámicas de alta complejidad (Russell & Norvig, 2016).

Históricamente, durante las primeras etapas del desarrollo del modelado neuronal aplicado a sistemas térmicos, el estándar se basaba en el uso de funciones de transferencia simétricas y monótonas crecientes, siendo la función sigmoide logística (LOGSIG) y la tangente hiperbólica

(TANSIG) las más prominentes. Ambas funciones tienen la característica de "aplastar" el espacio de entrada infinito en un dominio de salida acotado, típicamente $(0, 1)$ o $(-1, 1)$, lo cual resultaba intuitivamente útil para imitar el potencial de acción binario de las neuronas biológicas.

No obstante, cuando estas arquitecturas clásicas se someten a la predicción de variables físicas altamente fluctuantes y ruidosas —como los gradientes de temperatura o los perfiles de concentración del bromuro de litio— emergen limitaciones algorítmicas severas. El principal obstáculo analítico es el fenómeno conocido como "desvanecimiento del gradiente" (vanishing gradient problem). Como lo demostraron formalmente Glorot y Bengio (2010), la derivada de las funciones TANSIG y LOGSIG alcanza su valor máximo únicamente en el origen (cuando la entrada es cercana a cero). A medida que el valor de entrada ponderado se aleja del centro, la función entra en sus regiones de saturación asintótica, provocando que la derivada local converja rápidamente a cero.

Dado que el algoritmo de retropropagación (backpropagation) calcula las actualizaciones de los pesos sinápticos multiplicando estas derivadas mediante la regla de la cadena, un gradiente cercano a cero paraliza la transferencia de información de error. Esto significa que, si una perturbación térmica genera una señal de entrada grande, la neurona se "satura" y detiene su aprendizaje, impidiendo que el modelo capture los picos transitorios en la operación del AHT.

Buscando una solución geométrica a la saturación asintótica, la comunidad científica introdujo la Unidad Lineal Rectificada (ReLU, por sus siglas en inglés), definida matemáticamente como $f(x) = \max(0, x)$. Al poseer una derivada constante de 1 para todos los valores positivos, ReLU resolvió el desvanecimiento del gradiente en el semiplano derecho, acelerando drásticamente los tiempos de convergencia matemática.

Sin embargo, su aplicación en la predicción de series de tiempo termodinámicas introdujo una falla estructural conocida como el problema de la "neurona muerta" (dying ReLU). De acuerdo con Lu et al. (2019), si durante el entrenamiento una neurona actualiza sus pesos de tal forma que su entrada ponderada se vuelve negativa, la función ReLU colapsa su salida y su gradiente a un valor exacto de cero. Al presentar un gradiente nulo, la neurona pierde la capacidad de actualizarse en las iteraciones futuras, apagándose permanentemente.

Este comportamiento representa un riesgo crítico en el modelado del Coeficiente de Rendimiento (COP). Dado que las temperaturas operativas del sistema de absorción pueden presentar caídas transitorias abruptas —generando entradas negativas tras el preprocesamiento de normalización— el uso exclusivo de la función ReLU resulta en redes topológicamente frágiles y sub-optimizadas para aplicaciones térmicas.

La necesidad de un operador matemático que combinara la no-saturación de ReLU con una región de transición suave motivó la búsqueda automatizada de nuevas topologías. Utilizando algoritmos de aprendizaje por refuerzo, los investigadores de Google Brain (Ramachandran et al., 2017) identificaron una nueva familia de funciones de transferencia que superaba consistentemente a las arquitecturas clásicas: la función Swish, también formalizada como la Unidad Lineal Sigmoide (SiLU).

La función SiLU se define como el producto algebraico entre la entrada de la neurona y la función sigmoide estándar: $f(x) = x \cdot \sigma(x)$. A diferencia de todas sus predecesoras históricas, SiLU introduce una característica geométrica revolucionaria: es una función suave, acotada inferiormente, pero no monótona. Al acercarse a los valores negativos, la función presenta una ligera depresión (bump o caída en valores negativos) antes de converger hacia cero (Elfwing et al., 2018). Esta concavidad no monótona actúa como un mecanismo de regularización matemática

intrínseca; permite que los gradientes negativos pequeños fluyan hacia atrás a través de la red, erradicando por completo el problema de las neuronas muertas y preservando la memoria a corto plazo de las variaciones térmicas sutiles (Misra, 2019).

En este contexto de innovación algorítmica es donde se justifica la selección de la función dSiLU (la derivada de la Unidad Lineal Sigmoide) como el núcleo de la presente investigación. Emplear dSiLU no como un paso intermedio de la retropropagación, sino como la función de transferencia principal, dota a la neurona de un perfil de activación en forma de campana suave pero asimétrica.

Esta geometría específica emula el comportamiento de un campo receptivo localizado. Para el modelado de un ciclo de absorción, esto significa que la neurona dSiLU presenta una sensibilidad extrema a una banda de operación específica de las variables (como el rango de concentración seguro de la mezcla LiBr-H₂O), atenuando suavemente las señales que representan condiciones anómalas sin llegar a saturar el gradiente abruptamente. Al fusionar la función dSiLU con los operadores del cálculo fraccionario (Mittag-Leffler) o conformable, la arquitectura neuronal obtiene una elasticidad analítica sin precedentes, siendo capaz de autoajustar la apertura y curvatura de su "campana" de activación para mimetizar la física no lineal con gran aproximación del transformador de calor. La Figura 4 evidencia estas diferencias entre los comportamientos de las distintas funciones de transferencia.

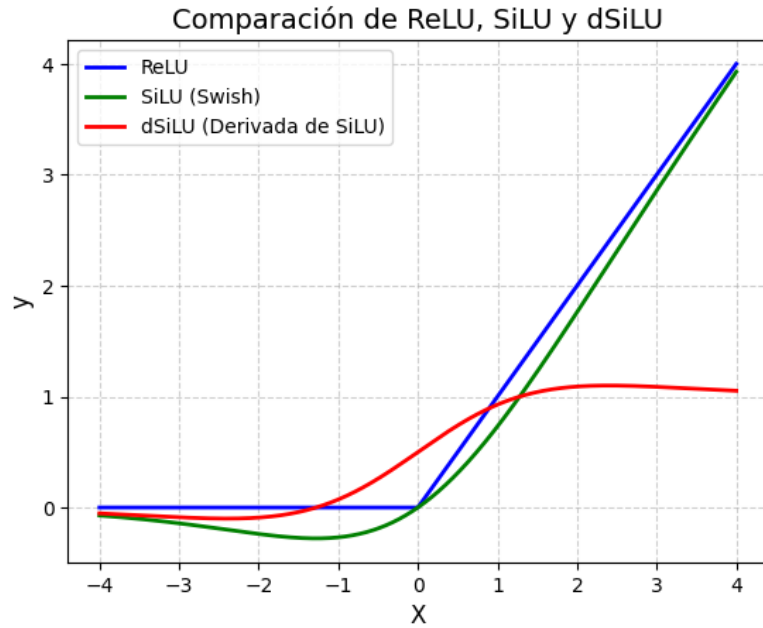


Figura 4. Comparación de comportamiento para las funciones ReLU, Silu, dSilu

El presente trabajo de investigación pretende introducir nuevas funciones de transferencia y comparar su desempeño con las funciones existentes, de modo que sean capaces de tener un mayor acercamiento al mínimo error posible.

La contribución científica esperada de esta tesis será precisamente disminuir el error de las RNA, para optimizar el procedimiento de cálculo.

2.2.4. Cálculo Fraccionario

El cálculo fraccionario, a diferencia del cálculo diferencial e integral de orden entero formulado por Newton y Leibniz, extiende el concepto del operador diferencial a órdenes arbitrarios no enteros (reales o complejos). En el contexto del modelado de sistemas físicos, las ecuaciones diferenciales de orden fraccionario han demostrado una capacidad excepcional para describir la dinámica de sistemas con "memoria" y procesos difusivos anómalos, características

intrínsecas de las irreversibilidades térmicas en mezclas como el bromuro de litio-agua (Baleanu et al., 2012).

El empleo del cálculo fraccionario en el diseño de algoritmos de Inteligencia Artificial ofrece una serie de ventajas y posibilidades que pueden mejorar significativamente la eficiencia y el rendimiento de los sistemas. En el contexto específico del diseño y simulación de un esquema de control fraccionario que involucra una red neuronal, la utilización de cálculo fraccional puede aportar beneficios clave, como puede ser el aumento de precisión en modelados de sistemas complejos, aumento de flexibilidad en modelos de Deep Learning, optimización de algoritmos de aprendizaje, entre otros. (Macías-Díaz, 2022).

En primer lugar, el cálculo fraccionario permite modelar de manera más precisa y flexible los comportamientos dinámicos de sistemas complejos, como los presentes en procesos energéticos (Tarasov, 2010; Baleanu et al., 2012). La inclusión de términos fraccionarios en las ecuaciones diferenciales proporciona una descripción más realista de fenómenos no lineales y memoria a largo plazo, lo cual es fundamental en el análisis y control de sistemas energéticos.

En el ámbito de la Inteligencia Artificial, donde la capacidad de aprendizaje y adaptación es esencial, la introducción de elementos fraccionarios en algoritmos de control puede mejorar la capacidad del sistema para ajustarse a patrones complejos y dinámicos en los datos. Esto se traduce en una mayor eficiencia en la toma de decisiones y una mejor adaptación a cambios en el entorno.

La capacidad de modelar de manera más precisa los comportamientos de los sistemas energéticos con términos fraccionarios puede conducir a estrategias de control más eficientes, reduciendo el consumo de energía y optimizando la operación del sistema.

En cuanto a la comparación con la eficiencia de la Inteligencia Artificial desarrollada con los algoritmos y matemáticas tradicionales, la inclusión de cálculo fraccionario brinda la

oportunidad de superar limitaciones y mejorar la capacidad predictiva y de adaptación del sistema. La validación y análisis de los resultados obtenidos permitirán evaluar de manera más precisa la contribución del cálculo fraccional en comparación con enfoques convencionales.

La combinación de técnicas de cálculo fraccionario con Inteligencia Artificial puede abrir nuevas perspectivas en el diseño de algoritmos avanzados, con aplicaciones significativas en diversos campos, desde la optimización de procesos industriales hasta el desarrollo de sistemas de control más eficientes y adaptables.

A lo largo de la historia, diversas formulaciones han buscado establecer una definición universal para la derivada fraccionaria. Las definiciones más rigurosas y ampliamente adoptadas en la literatura científica son las de Riemann-Liouville y Caputo. Ambas aproximaciones se fundamentan en el uso de la integral de convolución. Por ejemplo, la derivada fraccionaria de Caputo de orden α (donde $n - 1 < \alpha < n$) se define matemáticamente como (Eq. 4):

$$C_a D_t^\alpha f(t) = \frac{1}{\Gamma(n - \alpha)} \int_a^t \frac{f^{(n)}(\tau)}{(t - \tau)^{\alpha - n + 1}} d\tau. \quad (4)$$

Esta formulación integral es la responsable de la principal ventaja fenomenológica del cálculo fraccionario: su naturaleza no local (Laskin, 2000). A diferencia de una derivada estándar que evalúa la pendiente de una función en un punto infinitesimal exacto, la derivada de Caputo requiere integrar la historia completa de la variable desde un tiempo inicial a hasta el tiempo actual t . Al aplicar este concepto a las Redes Neuronales Artificiales mediante la función de generalización de Mittag-Leffler, se dota a las neuronas ocultas de un campo receptivo histórico, permitiéndoles "recordar" las perturbaciones previas del transformador de calor (Zeng et al., 2013).

Sin embargo, esta virtud teórica representa simultáneamente su mayor barrera tecnológica. La evaluación numérica de una integral de convolución en cada paso de actualización del algoritmo

de retropropagación exige un costo computacional masivo y un incremento exponencial en la memoria requerida (Diethelm, 2010). Para el desarrollo de sistemas de control predictivo que deben operar a nivel industrial mediante Controladores Lógicos Programables (PLC), la sobrecarga de evaluar series infinitas y funciones Gamma de manera recursiva resulta frecuentemente prohibitiva, restringiendo el cálculo fraccionario clásico a simulaciones computacionales fuera de línea (offline).

El cálculo fraccionario es una rama de las matemáticas que estudia operadores de diferenciación e integrales de orden arbitrario (Kilbas et al., 2006; Podlubny, 1999). El orden de los operadores fraccionarios no está restringido a valores enteros, sino que puede tomar valores reales o complejos. Asimismo, es posible considerar un orden variable, expresado como una función de la variable independiente, dando lugar a los operadores fraccionarios de orden variable (Almeida & Torres, 2013). La Tabla 1 resume diferentes definiciones de derivadas fraccionarias, abarcando desde las formulaciones clásicas hasta los operadores recientes de núcleo no singular, siendo estas algunas de las más estudiadas en la literatura actual (Gómez-Aguilar & Atangana, 2017).

Tabla 1. Derivadas fraccionarias

Definición	Ecuación
Riemann - Liouville	${}^{\text{RL}}D_t^\alpha f(t) = \frac{1}{\Gamma(n-m)} \frac{d^n}{dt^n} \int_0^t (t-\tau)^{n-m-1} f(\tau) d\tau, \alpha \in [n-1, n].$
Grünwald - Letnikov	${}^{\text{GL}}D_t^\alpha f(t) = \lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sum_{m=0}^{\frac{t-a}{h}} (-1)^m \binom{\alpha}{m} f(t-mh).$

Liouville - Caputo	${}^{LC}_0 D^\alpha f(x) = \frac{1}{\Gamma(1-\alpha)} \int_0^x \frac{f'(t)}{(x-t)^\alpha} dt.$
Caputo - Fabrizio	${}^{CF}_b D^\alpha_t f(t) = \frac{M(\alpha)}{1-\alpha} \int_b^t f'(\tau) \exp\left(-\frac{\alpha(t-\tau)}{1-\alpha}\right) d\tau.$
Atangana - Baleanu	${}^{ABC}_b D^\alpha_t f(t) = \frac{B(\alpha)}{1-\alpha} \int_b^t f'(\tau) E_\alpha\left(-\frac{\alpha(t-\tau)^\alpha}{1-\alpha}\right) d\tau.$

La formulación matemática de las derivadas fraccionarias se compone generalmente de dos operaciones sucesivas. La primera consiste en la convolución de la función objetivo con un núcleo integral, lo cual permite modelar el efecto de memoria del sistema en el dominio del tiempo, o bien, su fractalidad en el dominio espacial. La segunda operación corresponde a la diferenciación clásica de la función resultante de dicha convolución.

La diferencia entre las definiciones de Riemman-Liouville y Caputo-Fabrizio radica principalmente en el núcleo. La primera presenta una función de potencia en su núcleo, mientras que la segunda utiliza un núcleo exponencial que genera un efecto de memoria débil, presente en fenómenos naturales. Otra diferencia se encuentra en el sentido de la derivada, donde se puede encontrar en el sentido de Riemman y en el sentido de Liouville-Caputo. En el primero, el operador diferencial se encuentra dentro de la integral asociada, mientras que en el segundo está fuera de la integral. La definición de la derivada de Atangana-Baleanu utiliza un núcleo definido por la función de Mittag-Leffler, que permite representar comportamientos de núcleos de potencia o exponenciales según el orden de la derivada. (Kilbas et al., 2006).

2.2.5. Cálculo Conformable

Para resolver la dicotomía entre la necesidad de un comportamiento de orden no entero (para evitar la saturación neuronal) y la viabilidad computacional, la comunidad matemática

desarrolló un enfoque alternativo altamente eficiente: la derivada conformable. Propuesta formalmente por Khalil et al. (2014), esta aproximación redefinió el concepto diferencial introduciendo una extensión local y simplificada del límite clásico.

Dado un orden fraccionario $\alpha \in (0,1]$, la derivada conformable de una función $f: [0, \infty) \rightarrow \mathbb{R}$ en un punto $t > 0$ se define a través del siguiente límite fundamental (Eq. 5):

$$T_{\alpha}(f)(t) = \lim_{\epsilon \rightarrow 0} \frac{f(t + \epsilon t^{1-\alpha}) - f(t)}{\epsilon}. \quad (5)$$

A diferencia de las formulaciones de Riemann-Liouville o Caputo, el operador de Khalil no depende de integrales retrospectivas, sino que modifica localmente el diferencial de tiempo ponderándolo por una potencia fraccionaria del estado actual ($t^{1-\alpha}$). Esta formulación produce un operador estrictamente local que emula la flexibilidad geométrica del orden fraccionario sin arrastrar el lastre computacional de la memoria histórica (Abdeljawad, 2015). La eficacia de este operador local ha sido demostrada recientemente al modelar con alta precisión fenómenos termodinámicos clásicos, como la ley de enfriamiento de Newton, logrando ajustes analíticos superiores a los del cálculo de orden entero (Rosales-García et al., 2020).

La aportación más trascendental del cálculo conformable para la inteligencia artificial reside en su compatibilidad algebraica pura. Mientras que las derivadas fraccionarias tradicionales violan gran parte de los teoremas fundamentales del cálculo, el operador conformable satisface de manera exacta la regla del producto, la regla del cociente y, críticamente, la regla de la cadena (Chung, 2015).

Si f y g son funciones diferenciables, la regla de la cadena conformable establece (Eq. 6):

$$T_{\alpha}(f \circ g)(t) = t^{1-\alpha} g'(t) f'(g(t)). \quad (6)$$

Para el entrenamiento de una red neuronal, el cumplimiento de la regla de la cadena es un requisito innegociable (Tarasov, 2016). El algoritmo de retropropagación y los optimizadores de segundo orden (como Levenberg-Marquardt) calculan la matriz Jacobiana multiplicando sucesivamente las derivadas de las funciones de activación de cada capa. Al utilizar el operador conformable, es posible calcular estos gradientes algebraicamente de forma analítica y explícita, evitando las pesadas estimaciones numéricas aproximadas que paralizan a las redes de Mittag-Leffler.

2.3. Fundamentos de energía y sistemas térmicos de transformación

La comprensión de cualquier sistema físico, mecánico o térmico debe partir de su elemento más básico: la energía. En términos generales de la física y la termodinámica, la energía se define como la capacidad que posee un cuerpo o un sistema para realizar un trabajo, producir un cambio o transferir calor (Çengel & Boles, 2015).

En la sociedad moderna, la energía es el motor que sostiene desde las actividades cotidianas hasta los procesos industriales más complejos. No obstante, la energía no se crea ni se destruye, únicamente se transforma de una manifestación a otra, un principio dictado por la primera ley de la termodinámica. Para facilitar su estudio y aprovechamiento, los recursos energéticos naturales pueden clasificarse de distintas maneras, dependiendo de su origen, su capacidad de regeneración y el tratamiento que reciben antes de llegar al consumidor final.

2.3.1 El desafío del calor residual industrial

A pesar de que la energía térmica es indispensable, una proporción considerable de ella termina desperdiciándose. En la mayoría de los procesos industriales, una vez que el calor ha cumplido su propósito principal —como manufacturar un producto o mover una turbina— queda un remanente de energía térmica de baja temperatura. A este remanente se le conoce como calor residual, y debido a que resulta difícil de aprovechar, por lo general se libera directamente a la atmósfera o a cuerpos de agua cercanos, perdiéndose por completo (Dincer & Rosen, 2021).

Recuperar este calor residual representa un área de oportunidad muy importante para reducir el consumo de combustibles primarios en la industria. El principal obstáculo con este calor sobrante es que, al encontrarse a una temperatura baja, pierde su capacidad técnica para ser utilizado en otras partes del proceso. Por lo tanto, el reto actual no consiste únicamente en capturar ese calor, sino en encontrar un método eficiente para elevar su temperatura y volverlo útil nuevamente sin depender de un alto consumo eléctrico. Es justo para dar solución a esta problemática de reciclaje energético que entran en juego tecnologías térmicas especializadas, destacando entre ellas el uso de las bombas de calor.

2.3.2 Bombas de Calor por Absorción para calentamiento

Las bombas de calor por absorción para calentamiento son dispositivos que usan fuentes de energía térmica, en lugar de electricidad, para transferir calor desde una zona fría a una más caliente. La Figura 5 esquematiza el sistema de bomba de calor por absorción para calentamiento (Herold et al., 2016). Esto las hace ideales para situaciones en las que se dispone de calor residual o de una fuente de calor de bajo costo, como el gas natural o la energía solar. Una bomba de calor por absorción funciona a partir de un ciclo termodinámico en el cual el calor de entrada impulsa

una serie de reacciones y cambios de fase en un refrigerante y un absorbente, que son los fluidos que circulan dentro del sistema. Los pares de refrigerante-absorbente más comunes son el amoníaco-agua y el bromuro de litio-agua, cada uno con características específicas según la aplicación.

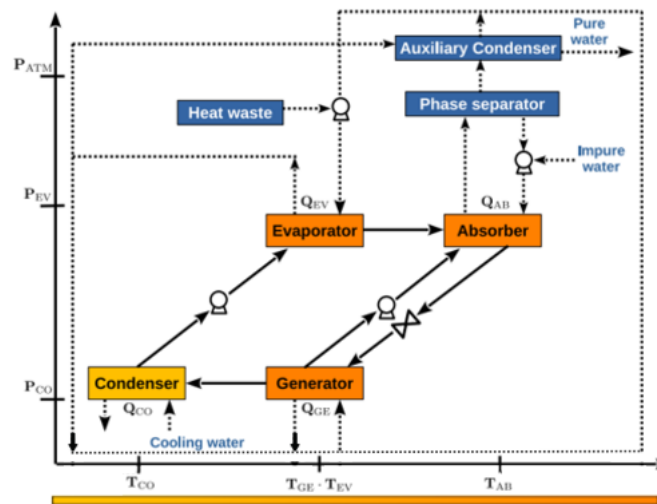


Figura 5. Transformador de calor por absorción con reciclaje de energía integrado al sistema de purificación de agua. Modificada de Solís-Pérez et al. (2022).

El ciclo de una bomba de calor por absorción se puede entender mejor al observar sus cuatro componentes principales (Universidad de Sevilla):

- Evaporador: Aquí el refrigerante absorbe calor de la fuente y se evapora, extrayendo así la energía térmica.
- Absorbedor: El refrigerante en estado de vapor se mezcla con el absorbente en una solución concentrada, que se encarga de transportar el calor hacia la siguiente fase.
- Generador: Esta mezcla es calentada, lo cual separa el refrigerante del absorbente, permitiendo que el vapor de refrigerante suba hacia el condensador.

- Condensador: El refrigerante vaporizado se condensa, liberando calor en el proceso, que puede usarse para calefacción o para algún proceso que requiera calor.

Ciclo de Absorción y Eficiencia del COP

El ciclo de absorción es un proceso termodinámico en el cual la energía térmica impulsa una serie de reacciones en el refrigerante y el absorbente, generando un flujo de calor que puede aprovecharse para calefacción o enfriamiento. Un aspecto crucial en las bombas de calor por absorción es el Coeficiente de Rendimiento (COP), que mide la eficiencia del sistema. Este coeficiente es la relación entre el calor útil que la bomba produce y el calor que consume. Cuanto más alto sea el COP, más eficiente es el sistema, ya que significa que está generando más calor útil con menor energía de entrada.

La eficiencia de las bombas de calor por absorción, y por tanto su COP, depende de varios factores como la temperatura de la fuente de calor, la concentración de la solución absorbente y las condiciones de funcionamiento (Herold et al., 2016).

2.3.3 Variables termodinámicas de funcionamiento para el sistema de bomba de calor por absorción

Las variables que influyen en el sistema son temperaturas, presiones y concentraciones, para hacer referencia a cada una de ellas en la programación de la RNA, la nomenclatura con el que se hará referencia a cada una de ellas es:

- Para temperaturas: $T_{in/out, "Elemento1-Elemento2"}$
- Para concentraciones: $X_{in/out, "Elemento"}$
- Para las presiones: $P_{Elemento}$

Para los elementos se utilizará una nomenclatura abreviada:

- Generador: GE

- Condensador: CO
- Evaporador: EV
- Absorbedor: AB

Temperaturas

$T_{in,GE-AB}$: Temperatura de entrada desde el generador al absorbedor. Controla la cantidad de calor transferido al absorbedor.

$T_{in,EV-AB}$: Temperatura de entrada desde el evaporador al absorbedor. Indica el nivel de calor absorbido en el evaporador que se transfiere al absorbedor.

$T_{out,AB-GE}$: Temperatura de salida del absorbedor hacia el generador. Influye en el calor que se dirige hacia el generador.

$T_{in,AB-GE}$: Temperatura de entrada del absorbedor al generador. Controla la energía que se transfiere del absorbedor al generador.

$T_{out,GE-AB}$: Temperatura de salida del generador al absorbedor. Afecta la eficiencia de la desorción del refrigerante en el generador.

$T_{out,GE-CO}$: Temperatura de salida del generador hacia el condensador. Determina la cantidad de calor transferido del generador al condensador.

$T_{in,CO}$: Temperatura de entrada en el condensador. Afecta el proceso de condensación del refrigerante.

$T_{out,CO}$: Temperatura de salida en el condensador. Influye en la cantidad de refrigerante condensado disponible en el sistema.

$T_{in,EV}$: Temperatura de entrada en el evaporador. Controla la energía que se transfiere en el evaporador.

$T_{out,EV-AB}$: Temperatura de salida del evaporador hacia el absorbedor. Determina el calor transferido desde el evaporador al absorbedor.

Concentraciones

$X_{in,AB}$: Concentración de absorbente en la entrada del absorbedor. Afecta la capacidad de absorción del refrigerante.

$X_{out,AB}$: Concentración de absorbente en la salida del absorbedor. Influye en el equilibrio de concentración en el absorbente al salir del absorbedor.

$X_{in,GE}$: Concentración de absorbente en la entrada del generador. Representa la cantidad de refrigerante en la mezcla en el generador.

$X_{out,GE}$: Concentración de absorbente en la salida del generador. Controla la concentración de refrigerante desorbido en el generador.

Presiones

P_{AB} : Presión en el absorbedor. Afecta la capacidad del sistema para absorber el refrigerante y la eficiencia del ciclo de absorción.

P_{GE} : Presión en el generador. Controla la eficiencia del proceso de desorción en el generador.

2.3.4 Descripción General del Ciclo de Absorción

Absorbedor (AB):

En el absorbedor, el refrigerante en forma de vapor ingresa y se mezcla con la solución absorbente, formando una solución rica en refrigerante. Este proceso libera calor, el cual puede ser disipado en un disipador.

- La concentración de entrada y salida en el absorbedor ($X_{in\ AB}$, $X_{out\ AB}$) afecta el equilibrio de la mezcla, permitiendo que el sistema regrese al generador con la concentración adecuada para que el ciclo se repita.
- La presión en el absorbedor (PAB) influye en la absorción efectiva del refrigerante.

Generador (GE):

En el generador, una fuente de calor (proveniente del Heat Source en el diagrama) proporciona la energía necesaria para desorbitar el refrigerante (agua) de la solución absorbente (generalmente una mezcla de agua y amoníaco o agua y bromuro de litio).

- La solución rica en refrigerante entra al generador con una concentración inicial ($X_{in,GE}$) y sale con una menor concentración ($X_{out,GE}$), ya que el refrigerante ha sido extraído.
- La temperatura y presión en el generador ($T_{in,AB-GE}$, $T_{out,GE-AB}$, P_{GE}) son factores clave para determinar la cantidad de refrigerante extraído de la solución.

Condensador (CO):

El vapor de refrigerante que sale del generador entra al condensador a alta presión (P_{GE}) y se enfría, cambiando de estado gaseoso a líquido.

- Las temperaturas de entrada y salida en el condensador ($T_{in,CO}$, $T_{out,CO}$) afectan la cantidad de refrigerante condensado y, por lo tanto, la cantidad de refrigerante disponible para el ciclo.

2.3.5 COP (Coeficiente de Rendimiento)

El COP, o Coefficient of Performance, es una medida de la eficiencia de la bomba de calor y se calcula en función del calor absorbido y del calor suministrado al sistema. Se define como la

relación entre el calor útil transferido (ya sea al evaporador o al absorbedor) y la energía térmica de entrada en el generador.

La Figura 6 especifica las referencias para los flujos másicos y entalpías que se presentarán en las ecuaciones que sirven para calcular el calor absorbido en cada componente.

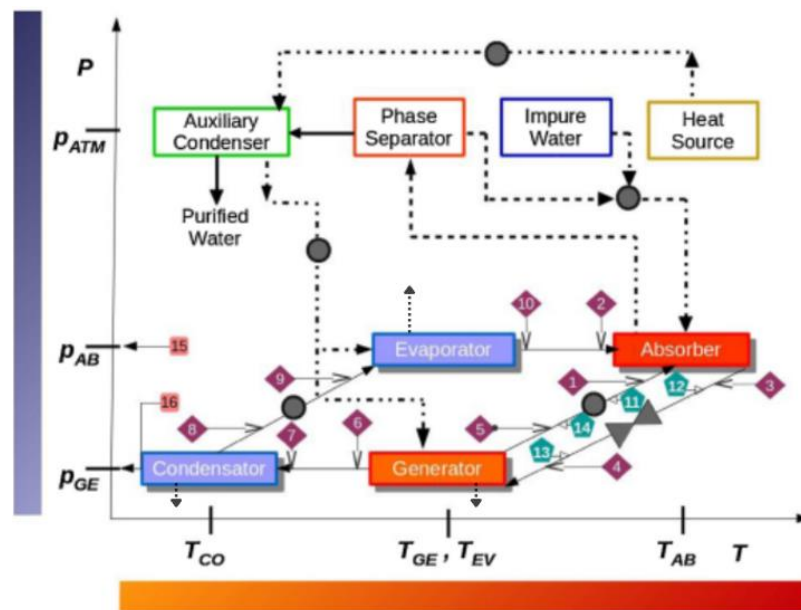


Figura 6. Esquema de un transformador de calor por absorción acoplado a un sistema de purificación de agua con reciclaje de energía. Modificada de Conde-Gutiérrez et al. (2018).

También, en la Figura 6 se muestra el sistema de bomba de calor por absorción para calentamiento, los números del 1 al 10 corresponden a las posiciones donde se ubican las temperaturas previamente mencionadas, los números del 11 al 14 hacen referencia a las concentraciones y tanto el 15 como el 16 son presiones. A continuación, se explica como calcular el calor producido en cada elemento, donde m hace referencia flujo másico, h a entalpía, y los números al costado hacen referencia a la posición de la temperatura asociada a dicha entalpía y flujo másico.

Calor en el Absorbedor (Q_{AB}):

➤ $Q_{AB} = m_2 \cdot h_2 + m_1 \cdot h_1 - m_3 \cdot h_3$

- Donde los flujos másicos y entalpías en diferentes puntos del absorbedor influyen en la cantidad de calor liberado en este proceso. La temperatura de entrada y salida y la concentración de la solución afectan los valores de entalpía y, en consecuencia, el calor liberado en el absorbedor.

Calor en el Generador (QGE):

- $QGE = m_6 \cdot h_6 + m_5 \cdot h_5 - m_4 \cdot h_4$
- Representa el calor necesario para desorbitar el refrigerante en el generador. Este es el "input" principal del sistema, y depende de la temperatura de entrada y salida del generador, así como de las concentraciones $X_{in,GE}$ y $X_{out,GE}$.

Calor en el Evaporador (QEV):

- $QEV = m_9 \cdot (h_{10} - h_9)$
- Determina el calor absorbido por el refrigerante al pasar por el evaporador. La temperatura de entrada y salida $T_{in,EV}$ y $T_{out,EV-AB}$ afectan este valor, ya que determinan el cambio de fase del refrigerante en el evaporador.

Calor en el Condensador (QCO):

- $QCO = m_7 \cdot (h_7 - h_8)$
- Es el calor liberado cuando el refrigerante se condensa en el condensador. Este valor depende de las temperaturas $T_{in,CO}$ y $T_{out,CO}$ y del flujo de refrigerante.

Fórmula del COP para la Bomba de Calor

El COP se calcula con la siguiente fórmula (Eq. (7)):

$$COP = \frac{Q_{AB}}{Q_{GE} + Q_{EV}}, \quad (7)$$

donde Q_{AB} es el calor total liberado en el absorbedor, que puede utilizarse para calentar el espacio o transferirse a otra carga térmica y Q_{GE} es el calor suministrado al generador para desorbitar el refrigerante.

Evaporador (EV):

El refrigerante líquido pasa del condensador al evaporador, donde se expande y reduce su presión. En el evaporador, el refrigerante absorbe calor del medio ambiente o de otra fuente (a la temperatura $T_{in,EV}$), cambiando de líquido a vapor.

- La temperatura y presión en el evaporador son claves para el intercambio de calor eficiente.

La salida del evaporador ($T_{out,EV-AB}$) y la concentración de refrigerante determinan la cantidad de calor absorbido.

2.4. El reto del modelado termodinámico y el uso de Inteligencia Artificial

En el panorama actual de la ingeniería térmica y los sistemas de energía sostenible, el modelado basado en datos ha pasado de ser un método estadístico auxiliar para convertirse en una herramienta esencial en la optimización y control de procesos físicos complejos (Kalogirou, 2001). A medida que la demanda de energía global sigue en aumento, a la par de la urgencia ineludible de mitigar impactos ambientales como las emisiones de carbono, la eficiencia de las tecnologías de transformación térmica se ha posicionado como un foco central de la investigación a nivel internacional (IEA, 2023). La descarbonización del sector industrial depende en gran medida de

la capacidad tecnológica para recuperar y elevar la calidad de la energía de bajo grado, comúnmente desperdiciada.

2.4.1. Limitaciones operativas de los modelos analíticos y deterministas

En la tarea de recuperación térmica, el modelado termodinámico tradicional suele encontrar obstáculos metodológicos severos debido a la naturaleza estocástica del calor residual industrial. Los modelos analíticos clásicos, aunque se fundamentan rígidamente en las leyes de conservación de masa, conservación de especies y la primera y segunda ley de la termodinámica, a menudo presentan limitaciones significativas al intentar describir los fenómenos no lineales de transferencia de calor y masa inherentes a los sistemas energéticos a escala industrial (Yang, 2008).

El problema fundamental radica en la formulación de los balances de energía. Para resolver el comportamiento de un equipo térmico mediante métodos numéricos deterministas, es necesario establecer sistemas de ecuaciones diferenciales ordinarias o parciales fuertemente acopladas. En la práctica, para que estos sistemas logren converger computacionalmente, los modelos convencionales dependen de suposiciones de estado estacionario, propiedades térmicas constantes y reglas de mezcla ideales que no reflejan la naturaleza cambiante de los entornos físicos reales. Estas limitaciones se vuelven evidentes en sistemas caracterizados por cambios de fase simultáneos, perfiles de concentración variables y condiciones de operación transitorias. Intentar resolver estas matrices de ecuaciones termodinámicas en tiempo real (por ejemplo, mediante algoritmos iterativos de Newton-Raphson), o ante condiciones de alimentación fluctuantes para tomar decisiones de control operativo de la planta, resulta inviable debido a los altos tiempos de procesamiento y al riesgo de divergencia matemática.

2.4.2. El paradigma de las Redes Neuronales Artificiales como aproximadores universales

Como respuesta a este vacío técnico en la simulación dinámica, las Redes Neuronales Artificiales (RNA) han surgido como una herramienta computacional robusta para modelar ciclos termodinámicos altamente no lineales (Kalogirou, 2001; Şencan et al., 2006). De acuerdo con los fundamentos matemáticos establecidos por el Teorema de Aproximación Universal (Hornik et al., 1989; Cybenko, 1989), una red neuronal prealimentada con al menos una capa oculta tiene la capacidad de mapear cualquier función continua, sin importar su grado de dimensionalidad geométrica. Al funcionar bajo este principio topológico, las RNA evitan el requerimiento de resolver ecuaciones diferenciales complejas de manera explícita.

En su lugar, estas arquitecturas matemáticas se especializan en asimilar relaciones directamente de conjuntos de datos experimentales, proporcionando estimaciones y predicciones de alta fidelidad. Este enfoque es sumamente valioso para el diseño, dimensionamiento y optimización operativa (Mohanraj et al., 2012; Labus et al., 2013). La adopción de esta metodología trasciende la programación tradicional basada en reglas predefinidas, destacándose por su capacidad para identificar patrones latentes dentro de vastas bases de datos multivariantes que los métodos numéricos clásicos pasarían por alto debido a la presencia de ruido instrumental o perturbaciones ambientales (Kalogirou, 2001; Yang, 2008). Este cambio de paradigma hacia el modelado de "caja negra" o "caja gris" permite una integración flexible de fuentes de calor residual en las redes industriales existentes, logrando anticipar el rendimiento del equipo de forma instantánea ante variables de entrada impredecibles (Jouhara et al., 2018).

2.5. Complejidad física y operativa en sistemas de absorción

El Transformador de Calor por Absorción (AHT, por sus siglas en inglés) representa una tecnología importante para la sostenibilidad industrial (Khamooshi et al., 2014), diseñado con el propósito específico de elevar el calor residual de baja calidad a niveles de temperatura térmicamente útiles —un proceso conocido formalmente como "impulso de temperatura"— sin requerir un trabajo mecánico de compresión significativo. Este proceso de reciclaje de energía térmica es fundamental para sostener aplicaciones industriales que demandan calor de grado medio y alto, tales como la purificación de agua, los sistemas de desalinización térmica y las redes de calefacción urbana a gran escala (Hernández et al., 2012; Herold et al., 2016; Jouhara et al., 2018).

2.5.1. Acoplamiento multivariable y dinámica transitoria del sistema

Dado que los principios de operación fundamentales del Transformador de Calor por Absorción (AHT) y las propiedades fisicoquímicas de la mezcla LiBr-H₂O han sido previamente establecidos, es imperativo analizar el sistema desde la perspectiva de su dinámica transitoria. La inmensa complejidad de modelar este equipo no radica en sus componentes individuales, sino en el fuerte acoplamiento termodinámico y multivariable que existe entre sus cuatro intercambiadores de calor principales durante la operación continua.

En un escenario teórico ideal, el comportamiento de la mezcla binaria se proyecta sobre un diagrama de Dühring asumiendo estados de equilibrio perfectos. Sin embargo, en un entorno de recuperación de calor industrial, los intercambiadores no operan como volúmenes de control aislados o estáticos. Cualquier perturbación externa, por mínima que sea, desencadena un efecto dominó altamente no lineal en todo el ciclo. Por ejemplo, una fluctuación en la temperatura de la fuente de calor residual suministrada al evaporador altera instantáneamente la presión de

vaporización del refrigerante. Esta variación de presión viaja hacia el absorbedor, modificando abruptamente el potencial químico y la cinética de absorción entre el vapor y la película de solución salina (Florides et al., 2003).

Esta interdependencia provoca que las propiedades termofísicas de la mezcla —tales como la entalpía, la viscosidad dinámica y la densidad— varíen de forma drástica y no proporcional. El absorbedor y el generador experimentan tasas de transferencia de calor y masa que son funciones exponenciales de la temperatura y la concentración simultáneamente. Cuando el generador recibe una carga térmica inestable, la concentración de la solución fuerte que retorna al absorbedor oscila, impidiendo que el sistema alcance un estado estacionario prolongado.

Desde el punto de vista del modelado analítico, este comportamiento interconectado transforma las ecuaciones rectoras del sistema en una matriz de ecuaciones diferenciales rígidas (stiff equations). Resolver el balance de energía acoplado a los gradientes de concentración en tiempo real requiere recursos computacionales masivos, ya que los métodos iterativos convencionales tienden a divergir o estancarse cuando se enfrentan a los fuertes transitorios térmicos de la mezcla salina (Chua et al., 2000). Es en este caos dinámico y multivariable donde los modelos deterministas revelan sus mayores debilidades empíricas, haciendo imperativa la transición hacia arquitecturas neuronales capaces de mapear estas complejas topologías de datos sin depender del equilibrio termodinámico ideal.

2.5.2. Irreversibilidades térmicas y el riesgo crítico de cristalización

Aunque el cálculo del Coeficiente de Rendimiento (COP) parece analíticamente directo mediante la primera ley de la termodinámica, predecir su valor en un entorno físico representa un desafío analítico sustancial. La operación en línea está intrínsecamente plagada de

irreversibilidades y generación de entropía. El equipo experimenta disipación de calor hacia los alrededores, pérdidas estructurales por factores de fricción en los circuitos de tuberías, caídas de presión en los distribuidores de líquido, y frecuentemente opera bajo estados prolongados de no-equilibrio termodinámico.

A este entorno de alta entropía se suma un límite operativo estrecho debido a la naturaleza química del LiBr-H₂O. La mezcla salina posee un límite de solubilidad estricto. A medida que la concentración másica de la sal aumenta y su temperatura disminuye, la solución cruza la línea de saturación, provocando la cristalización de la sal en forma sólida. Este fenómeno ocurre con mayor frecuencia en la línea de solución concentrada que viaja del generador al absorbedor, especialmente al atravesar el intercambiador de calor de soluciones (SHX), donde el fluido cede temperatura. Si la fuente de calor residual de la planta industrial presenta fluctuaciones transitorias —como una caída abrupta en la temperatura de suministro— los gradientes internos del equipo se desploman, lo que puede inducir la precipitación del bromuro de litio (Shirazi et al., 2016). La cristalización obstruye las líneas de flujo, bloqueando la operación del sistema y exigiendo protocolos de mantenimiento correctivo invasivos (Hernández et al., 2012). Estas vulnerabilidades mecánicas y las incertidumbres operativas (Hernández & Colorado, 2013) provocan que el COP experimental fluctúe de forma no lineal, justificando categóricamente la necesidad de emplear algoritmos de aprendizaje profundo que puedan monitorear y anticipar estas anomalías en fracciones de segundo.

2.6. Transición hacia el cálculo fraccionario y conformable

La introducción formal del cálculo fraccionario representa un avance matemático que modifica el núcleo de los modelos predictivos neuronales. A diferencia de los enfoques

convencionales que se apoyan estrictamente en derivadas e integrales de orden entero para propagar las señales a través de una red, las Redes Neuronales Artificiales Fraccionarias (FANNs) incorporan operadores no enteros que ofrecen la capacidad de registrar efectos de "memoria" a largo plazo y capturar las dinámicas de relajación lenta propias de los procesos térmicos (Podlubny, 1999). Al sustituir las funciones clásicas por expansiones en series matemáticas más amplias, como la función de Mittag-Leffler (Atangana & Baleanu, 2015), resulta posible optimizar la topología del algoritmo, dotando a las neuronas de mayores grados de libertad geométrica y paramétrica (Wang et al., 2017; Tarasov, 2010).

2.6.1. Límite del orden entero y el costo de la memoria no local

La validación para integrar un comportamiento fraccionario radica en los límites de flexibilidad que han encontrado los modelos neuronales en las últimas décadas. Las arquitecturas clásicas típicamente emplean funciones de activación sigmoidales, trigonométricas o logarítmicas (como TANSIG o LOGSIG). Si bien logran converger, estas funciones evidencian un sesgo asintótico severo, fallando en superar las pruebas de linealidad estricta entre los datos experimentales y los valores calculados. Esta deficiencia ocurre porque las funciones de orden entero clásicas sufren del fenómeno conocido como "desvanecimiento del gradiente" (vanishing gradient); al alcanzar los extremos de su dominio, la derivada de la función se aproxima a cero, paralizando la actualización de los pesos sinápticos e impidiendo que el algoritmo asimile las perturbaciones térmicas complejas del sistema.

Para superar este estancamiento, los estudios analíticos (Viera-Martin et al., 2022) demostraron que un orden fraccionario generalizado permite a la red alterar dinámicamente su propia velocidad de convergencia. Esta adaptación permite reducir el número total de neuronas

ocultas, minimizando la carga espacial en la memoria y previniendo el sobreajuste de los datos de entrenamiento (Solís-Pérez et al., 2022). No obstante, el principal reto técnico en el uso del cálculo fraccionario clásico —basado en las definiciones integrales de Riemann-Liouville o Caputo— es que su naturaleza no local implica resolver integrales de convolución desde el origen del tiempo hasta el estado actual (Machado et al., 2011). Este requerimiento retrospectivo exige el cálculo continuo de funciones Gamma y transformadas integrales complejas. Como resultado, el hardware experimenta un costo de procesamiento severo, lo que representa un obstáculo logístico prohibitivo para la programación de bucles de control predictivo dentro de controladores lógicos programables (PLC) industriales.

2.6.2. Eficiencia computacional y analítica del cálculo conformable

Para conciliar la necesidad de mayor flexibilidad matemática con la viabilidad del tiempo de cómputo, el cálculo conformable ha emergido como un paradigma analítico de alta eficiencia algorítmica. Propuesto formalmente en la definición de derivada local por Khalil et al. (2014), el operador conformable proporciona la escalabilidad geométrica y el ajuste fino del orden no entero α , pero lo hace mediante un cálculo estrictamente local, fundamentado en una extensión del límite diferencial clásico (Gómez-Aguilar & Atangana, 2017). La principal fortaleza algorítmica de esta transición reside en la compatibilidad algebraica. A diferencia de los operadores fraccionarios tradicionales, la derivada conformable respeta las propiedades axiomáticas del cálculo clásico elemental, cumpliendo de manera exacta con la regla de la cadena, la regla del producto y la regla del cociente. Esta característica es fundamental en la arquitectura de las redes neuronales, puesto que los algoritmos de retropropagación y de optimización de pesos dependientes del gradiente —tales como el método de Levenberg-Marquardt o los optimizadores

estocásticos como Adam (Kingma & Ba, 2014)— requieren calcular de forma explícita las matrices Jacobianas y Hessianas del sistema neuronal. Al poder derivar una función de activación conformable sin requerir expansiones de series infinitas ni estimaciones numéricas aproximadas paso a paso, el tiempo de iteración disminuye exponencialmente. Recientemente, se ha demostrado que las redes construidas bajo este principio logran equilibrar el balance entre precisión física y carga de procesamiento (Solís-Pérez et al., 2022). Refinar la neurona integrando el orden conformable en funciones del estado del arte, como SWISH, ELU o dSilu, suprime los gradientes saturados, permitiendo que una red sumamente compacta alcance el mismo rigor predictivo que modelos convencionales mucho más pesados, facilitando su eventual despliegue en sistemas de monitoreo de plantas térmicas en tiempo real.

2.7 Avances contemporáneos en la aplicación de redes neuronales artificiales y cálculo fraccionario en el modelado de sistemas térmicos

El desarrollo de modelos basados en redes neuronales artificiales (RNA) ha adquirido especial relevancia en el análisis de sistemas térmicos debido a su capacidad para representar relaciones no lineales entre múltiples variables de operación sin requerir necesariamente una formulación matemática explícita del fenómeno estudiado. En particular, las RNA han sido empleadas para el modelado, predicción y optimización de diferentes procesos relacionados con transferencia de calor, refrigeración, bombas de calor y sistemas de absorción. En este contexto, Mohanraj et al. (2012) realizaron una revisión de las aplicaciones de las RNA en sistemas de refrigeración, acondicionamiento de aire y bombas de calor, mientras que investigadores como Ledesma et al. (2015) analizaron su aplicación específica en el modelado termodinámico de componentes de refrigeración. Además, revisiones posteriores realizadas por los mismos autores

(Mohanraj et al., 2015) demostraron que las RNA son igualmente eficaces para el análisis térmico integral en intercambiadores de calor complejos. Estos antecedentes muestran que las RNA constituyen una herramienta consolidada para aproximar el comportamiento de sistemas térmicos complejos a partir de variables experimentales o de operación.

En los sistemas de absorción, las RNA han demostrado particular utilidad debido a la complejidad asociada con la determinación de las propiedades termodinámicas de los pares de trabajo y con la interacción entre las diferentes variables de operación. Şencan, et al. (2006) emplearon redes neuronales para determinar propiedades termodinámicas de soluciones LiBr–H₂O y LiCl–H₂O, pares ampliamente utilizados en sistemas de absorción. Los autores señalaron que el cálculo convencional de estas propiedades puede involucrar relaciones y procedimientos matemáticos complejos, mientras que las RNA permitieron obtener expresiones de cálculo más sencillas. Para la entalpía de las soluciones estudiadas se obtuvieron coeficientes de determinación de aproximadamente $R^2 = 0.999$, mostrando una elevada correspondencia entre los valores obtenidos mediante la RNA y los valores de referencia. Además, los modelos fueron utilizados para calcular el coeficiente de rendimiento (COP) de sistemas de absorción bajo diferentes condiciones de operación.

La aplicación de RNA se ha extendido también directamente a la predicción y optimización del comportamiento de transformadores de calor por absorción. Morales et al. desarrollaron un modelo neuronal para predecir el COP de un transformador de calor por absorción con una configuración de componentes compactos, considerando como variables de entrada temperaturas de operación, concentración de la solución LiBr–H₂O, presión y flujo másico. El modelo que presentó el mejor ajuste empleó una arquitectura 16–7–1, con una función de transferencia tangente hiperbólica en la capa oculta y una función lineal en la capa de salida. Los resultados

mostraron una elevada concordancia entre los valores experimentales y los calculados, con un coeficiente de determinación superior a 0.9969 y un error porcentual medio cercano al 3 %. A partir de este modelo directo, los autores desarrollaron además una RNA inversa para determinar condiciones de operación favorables y la combinaron con un algoritmo genético para optimizar el desempeño del sistema. De manera complementaria, Colorado et al. (2011) analizaron la propagación del error en la predicción del COP de un sistema de purificación de agua integrado a un transformador de calor por absorción, empleando una RNA con funciones de transferencia tangente hiperbólica y lineal y el algoritmo de Levenberg-Marquardt para el entrenamiento (previamente, la eficacia de esta metodología inversa para optimizar la transferencia de calor y masa ya había sido demostrada por Hernández (2009) en sistemas de secado termodinámico, sentando las bases para el uso de algoritmos predictivos en el procesamiento de variables energéticas). En dicho estudio se consideraron 16 variables asociadas con temperatura, presión y concentración, lo que evidencia la capacidad de las RNA para integrar simultáneamente diferentes variables termodinámicas y de operación en la estimación del desempeño de estos sistemas.

A pesar de estos resultados, el desarrollo de modelos neuronales convencionales se encuentra condicionado por la elección de la arquitectura, las funciones de activación, el número de neuronas y la cantidad y distribución de los datos empleados durante el entrenamiento. En este sentido, el cálculo fraccionario representa una alternativa de interés para introducir grados adicionales de flexibilidad en modelos matemáticos y computacionales destinados a representar fenómenos complejos. Una línea particularmente relevante consiste en incorporar operadores o funciones de orden no entero directamente en la arquitectura de una red neuronal. La viabilidad de esta integración ha sido demostrada con éxito en la literatura matemática; por ejemplo, Zúñiga-Aguilar et al. (2017) evidenciaron que las Redes Neuronales Artificiales pueden acoplarse con

operadores que involucran el núcleo de Mittag-Leffler para aproximar y resolver ecuaciones diferenciales fraccionarias complejas, dotando al algoritmo de una memoria histórica robusta. Adicionalmente a estos fundamentos, Viera-Martin et al. realizaron una revisión práctica de las aplicaciones de las RNA que involucran cálculo fraccionario, mostrando que esta combinación ha sido estudiada en diferentes problemas de modelado y procesamiento, lo que evidencia el interés creciente por integrar herramientas del cálculo fraccionario con técnicas de aprendizaje automático.

Un antecedente particularmente cercano a esta línea de investigación fue desarrollado por Solís-Pérez et al. (2022), quienes propusieron una función de transferencia conformable de orden no entero para redes neuronales multicapa de alimentación hacia adelante. La función propuesta combina la tangente hiperbólica con una función exponencial conformable basada en la definición de Khalil, incorporando parámetros adicionales que pueden determinarse durante el entrenamiento de la red. Los autores evaluaron la propuesta en tres casos experimentales, incluyendo la estimación del número de Nusselt en un evaporador de doble tubo helicoidal acoplado a un transformador de calor por absorción, además de aplicaciones en procesos electroquímicos y en un colector solar de canal parabólico. Los mejores modelos presentaron relaciones del 99 %, 97 % y 95 % con los datos experimentales, respectivamente, y los autores señalaron que la función de transferencia conformable permitió reducir el número de neuronas necesarias en la capa oculta. Este resultado es particularmente significativo para el desarrollo de modelos térmicos, ya que demuestra que la incorporación de una función de transferencia de orden no entero puede modificar favorablemente la configuración de una RNA y mantener una adecuada capacidad de aprendizaje.

En conjunto, los antecedentes revisados permiten identificar dos líneas de investigación que convergen en el presente trabajo. Por una parte, las redes neuronales artificiales han

demostrado su capacidad para representar propiedades termodinámicas, predecir el COP y optimizar condiciones de operación en sistemas de absorción, incluyendo configuraciones que emplean soluciones LiBr–H₂O. Por otra parte, la incorporación del cálculo fraccionario en las redes neuronales, particularmente mediante funciones de transferencia conformables, ha mostrado resultados favorables en el modelado de procesos térmicos y ha abierto la posibilidad de modificar la capacidad de adaptación de las redes mediante parámetros de orden no entero. Sin embargo, a partir de los antecedentes revisados se observa que ambas líneas no han sido exploradas con la misma profundidad en la predicción del comportamiento termodinámico de bombas de calor por absorción empleando una RNA con funciones de transferencia asociadas al cálculo fraccionario. Por ello, resulta pertinente estudiar una arquitectura neuronal que incorpore este tipo de funciones y evaluar su capacidad para modelar las variables de interés del sistema, comparando su desempeño con el de arquitecturas neuronales convencionales mediante indicadores estadísticos de ajuste y error.

3. Antecedentes.

La optimización y el control de los procesos térmicos complejos han dejado de depender exclusivamente de modelos analíticos tradicionales. En sistemas como el transformador de calor por absorción (AHT), las transiciones de fase simultáneas y las fluctuaciones operativas hacen que las ecuaciones diferenciales sean computacionalmente prohibitivas para un análisis en tiempo real. En este escenario, las Redes Neuronales Artificiales (RNA) se han posicionado como la herramienta predictiva por excelencia. Sin embargo, encontrar la arquitectura perfecta para este sistema específico ha sido un reto que ha tomado casi dos décadas de refinamiento técnico, marcando una clara transición desde los modelos de orden entero hasta la vanguardia del cálculo fraccionario

- 2008: El primer gran obstáculo estadístico (Hernández et al., 2008)

Los esfuerzos formales para modelar este sistema específico utilizando inteligencia artificial comenzaron con el trabajo de Hernández et al. En esta etapa temprana, los investigadores diseñaron una red neuronal alimentada por 16 variables de entrada y estructurada con una capa oculta de tres neuronas (arquitectura 16:3:1), utilizando la función de transferencia TANSIG. A primera vista, los resultados fueron un éxito rotundo: el modelo alcanzó un coeficiente de correlación (R) de 0.99 y un error cuadrático medio (RMSE) de 0.021. Sin embargo, al someter el modelo a rigurosas validaciones estadísticas, específicamente la prueba de linealidad de pendiente e intercepto, la red falló. Esta discrepancia demostró que, aunque la correlación global era altísima, el modelo presentaba sesgos significativos en ciertos rangos de operación térmica, volviéndolo poco confiable para una aplicación de ingeniería de alta precisión.

- 2013: La búsqueda de la simplicidad y el límite de las funciones enteras (Hernández & Colorado, 2013)

Conscientes de que los modelos debían ser no solo precisos sino también eficientes, Hernández y Colorado emprendieron un análisis exhaustivo de incertidumbre buscando simplificar la topología de la red. Redujeron la arquitectura a su mínima expresión funcional: solo dos neuronas en la capa oculta (16:2:1), alternando entre funciones LOGSIG y TANSIG. A pesar de lograr un coeficiente de correlación excepcional ($R = 0.9978$) con un RMSE de 0.104, el resultado de las pruebas de validación fue el mismo. Estos modelos parsimoniosos, contruidos sobre matemáticas de orden entero, seguían fracasando en la prueba de pendiente-intercepto al 99% de confianza. Quedó claro que las funciones de activación convencionales carecían de la flexibilidad matemática necesaria para mapear con total exactitud los intrincados gradientes térmicos del ciclo de absorción.

- 2015: La solución de "fuerza bruta" y el riesgo de sobreajuste (Morales et al., 2015)

Ante la incapacidad de las redes pequeñas para pasar los filtros estadísticos, Morales et al. optaron por incrementar significativamente la densidad computacional del modelo. Implementaron una arquitectura robusta de 16:7:1 (siete neuronas en la capa oculta) optimizada mediante algoritmos genéticos y utilizando nuevamente la función TANSIG. Esta estrategia finalmente rindió frutos: con un R de 0.9969 y un RMSE de apenas 0.0052, el modelo logró superar las estrictas pruebas de linealidad. Sin embargo, este triunfo estadístico trajo consigo un costo prohibitivo. El uso de una red tan densa incrementó radicalmente el esfuerzo computacional y encendió las alertas sobre un posible sobreajuste (overfitting). Se había resuelto el problema matemático, pero se había perdido la eficiencia práctica necesaria para su implementación en tiempo real.

- 2018: La era de la optimización multivariable (Conde-Gutiérrez et al., 2018)

Buscando un equilibrio entre precisión y aplicabilidad, Conde-Gutiérrez y su equipo introdujeron un esquema de Red Neuronal Artificial Inversa (ANNI) acoplado a un algoritmo genético. Su objetivo principal era determinar las condiciones operativas multivariadas óptimas para maximizar el reciclaje de energía. Regresaron a una estructura más manejable de 16:3:1 con TANSIG, logrando un R de 0.9935. Aunque este trabajo sentó un precedente invaluable para la resolución de problemas inversos complejos, volvió a evidenciar que los modelos de orden entero operaban justo en el límite de la viabilidad estadística. La altísima no linealidad del proceso térmico exigía un motor matemático distinto.

- 2022: El cambio de paradigma hacia la memoria sistémica (Viera-Martin et al., 2022 y Solís-Pérez et al., 2022)

La verdadera revolución metodológica llegó con la introducción del cálculo no entero. Viera-Martin et al. (2022) demostraron que el uso de funciones de transferencia de orden fraccionario podía cambiar las reglas del juego. Al implementar la función de Mittag-Leffler (TANSIG ML) en una estructura ultraligera de 16:2:1, lograron un R de 0.9969. Por primera vez, se demostró que los operadores no enteros permitían a la red "recordar" la dinámica del sistema y mapear comportamientos térmicos complejos sin necesidad de inflar la cantidad de neuronas. Paralelamente, investigadores como Solís-Pérez integraron el cálculo conformable de Khalil, que ofrecía la flexibilidad de las derivadas fraccionarias, pero con un enfoque local, evitando los altos costos computacionales de los operadores clásicos y sentando las bases para el uso de funciones más modernas como SWISH, ELU o Silu.

3.1. Problemática

Los modelos de RNA reportados para la predicción del COP en bombas de calor por absorción cuentan con un error medio, se busca disminuir este valor empleando modelos de orden fraccionario en las funciones de transferencia y compararlas con los métodos ya reportados.

4. Hipótesis

¿Es posible mejorar el rendimiento del cálculo de una Red Neurona Artificial para la predicción del COP en un sistema de bombas de calor por absorción para calentamiento empleando un modelo de orden fraccionario en la función transferencia?

La introducción de nuevas funciones de transferencia generalizadas mediante operadores de orden fraccionario podría mejorar el desempeño de las Redes Neuronales Artificiales.

5. Objetivo general

Introducir nuevas funciones de transferencia en modelos de inteligencia artificial y aplicarlas en la predicción del Coeficiente de Rendimiento en bombas de calor por absorción, con el fin de mejorar la eficiencia de los modelos existentes, considerando tanto el orden entero como el cálculo fraccionario.

5.1 Objetivos específicos

- Lograr construir una RNA para la predicción del COP en el sistema que se va a trabajar.
- Definir los operadores fraccionarios ideales para minimizar el error de la predicción.
- Implementar operadores fraccionarios a la función de transferencia.
- Comparar resultados obtenidos con los resultados publicados utilizando funciones de transferencia de orden entero y conformables.

6. Metodología

- Estudio del sistema que en el que se va a trabajar (sistema de bombas de calor por absorción, programación y entrenamiento de RNA, operadores de orden fraccionario que se aplicarán).
- Contar con una base de datos de pruebas realizadas en el sistema
- Programación y entrenamiento de la RNA (Cómputo)

- Modificación de las funciones de transferencia implementando técnicas de cálculo fraccionario
- Entrenamiento de la RNA considerando la base de datos.
- Validación y pruebas de la RNA.
- Análisis estadístico entre los resultados arrojados por las nuevas funciones y las funciones clásicas

7. Aplicación de la metodología

Lo primero que se realizó fue replicar el mejor modelo de RNA reportado en el sistema de bomba de calor por absorción del CIICAp, una vez comprendido cómo funcionan las redes neuronales artificiales, se procedió a realizar modelos sobre el mismo sistema modificando entre cada modelo las funciones de activación y la cantidad de neuronas en la capa oculta, esto con la finalidad de trabajar con nuevas funciones de transferencia y aplicarlas en este sistema para poder observar su comportamiento. Para verificar la calidad de cada modelo se calcularán métricas estadísticas, las cuales son:

- Correlación lineal de Pearson
- Error cuadrático medio
- Raíz de error cuadrático medio

La base de datos con la que se va a trabajar consta de 16 variables de entrada y 1 salida, a continuación, la Tabla 2 lista cada una de ellas:

Tabla 2. Variables de la base de datos del sistema de bomba de calor por absorción

Variable	Valores mínimos	Valores máximos
$T_{in,GE-AB}$ (°C)	71.84	91.61
$T_{in,EV-AB}$ (°C)	69.34	89.93
$T_{out,AB-GE}$ (°C)	74.46	98.27
$T_{in,AB-GE}$ (°C)	73.27	92.15
$T_{out,GE-AB}$ (°C)	71.84	91.61
$T_{out,GE-CO}$ (°C)	72.04	83.89
$T_{in,CO}$ (°C)	39.42	65.05
$T_{out,CO}$ (°C)	26.75	33.79
$T_{in,EV}$ (°C)	28.52	81.34
$T_{out,EV-AB}$ (°C)	69.34	89.93

$X_{in,AB}$ (%)	51.66	55.33
$X_{out,AB}$ (%)	50.75	54.36
$X_{in,GE}$ (%)	50.75	54.36
$X_{out,GE}$ (%)	53.16	56.07
P_{AB} (en.Hg)	7	11.5
P_{GE} (en.Hg)	20	21.1
COP	0.21	0.39

Se cuentan con 4702 muestras, se seleccionarán aleatoriamente el 50% de todos los datos para el entrenamiento de la red, 25 % de los datos para las pruebas de la red, y finalmente el 25 % restante para validación de la red.

7.1. Arquitectura y topología de la red neuronal

Para llevar a cabo el modelado predictivo del Coeficiente de Rendimiento (COP), se implementó una arquitectura de Red Neuronal Artificial de tipo Feed-Forward (red prealimentada de propagación hacia adelante). La elección de esta topología sobre alternativas recurrentes o

dinámicas se fundamenta en la naturaleza del fenómeno físico a predecir y en la estructura del mapeo matemático requerido.

7.1.1. Topología de flujo unidireccional (Feed-Forward)

En una red Feed-Forward, el flujo de información es estrictamente unidireccional. Las señales numéricas se propagan desde la capa de entrada, a través de las capas ocultas, hasta alcanzar la capa de salida, sin la existencia de ciclos o bucles de retroalimentación interna (Hagan et al., 2014). Como señalan Goodfellow et al. (2016) y Bishop (2006), las arquitecturas prealimentadas son los modelos de aprendizaje profundo por excelencia para aproximar funciones complejas, ya que proporcionan un mapeo algebraico directo y determinista desde el espacio de entrada hasta la salida, evitando inestabilidades. Esta característica topológica es altamente compatible con el análisis de estados termodinámicos en sistemas de absorción.

Dado que el vector de entrada proporciona una caracterización completa del estado del sistema en un instante específico (mediante las 16 variables de temperatura, presión y concentración), el objetivo del modelo es establecer una función de regresión estática y directa hacia el COP resultante. Modelos con memoria interna retroalimentada, como las redes neuronales recurrentes (RNN), suelen introducir inestabilidades dinámicas cuando se intentan mapear relaciones algebraicas de variables de estado que ya han sido estabilizadas temporalmente (Ledesma et al., 2015). Por lo tanto, la estructura Feed-Forward garantiza una convergencia matemática más estable y un costo de evaluación (inferencia) lo suficientemente bajo para ser implementado en controladores lógicos de uso industrial.

Para ilustrar de forma gráfica la topología base del algoritmo implementado, la Figura 7 presenta el diagrama de flujo de información para una estructura prealimentada simplificada de una sola neurona oculta.

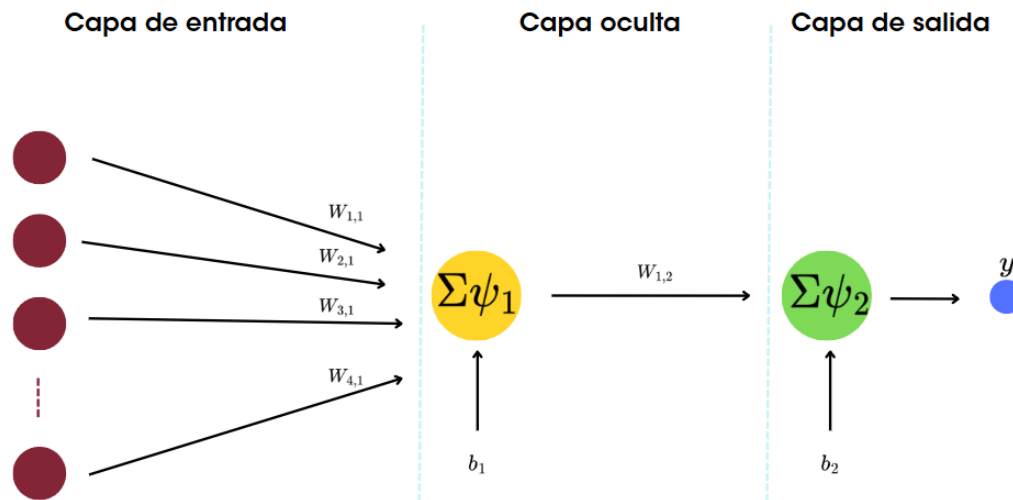


Figura 7. Estructura de una RNA de tipo Feed Forward

7.1.2. Configuración 16:2:1

El diseño arquitectónico de una red neuronal implica definir la dimensión de tres componentes fundamentales: la capa de entrada, la capa oculta y la capa de salida.

- Capa de entrada: Quedó rígidamente definida por 16 nodos receptores, correspondiendo a cada uno de los sensores físicos ubicados en el transformador de calor (previamente detallados en la sección de análisis de sensibilidad).
- Capa de salida: Se restringe a un único nodo, el cual tiene la tarea de entregar el valor escalar continuo correspondiente al Coeficiente de Rendimiento (COP).
- Capa oculta: La determinación de la cantidad de neuronas en esta capa representa el núcleo del diseño metodológico.

En la literatura especializada en modelado térmico, existe una tendencia a resolver la falta de precisión predictiva aumentando indiscriminadamente el número de neuronas ocultas. Trabajos previos en este mismo equipo experimental recurrieron a arquitecturas densas, como estructuras de 16:7:1, para lograr superar los filtros de error estadístico (Morales et al., 2015). Sin embargo, el aumento de neuronas conlleva un riesgo crítico conocido como sobreajuste (overfitting). Una red con demasiados nodos ocultos memoriza el ruido de los sensores y las fluctuaciones transitorias del equipo en lugar de generalizar la física subyacente del proceso de absorción (Prechelt, 1998).

Para evitar este fenómeno y garantizar la aplicabilidad del modelo, esta investigación adoptó el principio de parsimonia estructural. La metodología consistió en evaluar topologías sumamente compactas, analizando específicamente redes con una (16:1:1) y dos (16:2:1) neuronas en la capa oculta. El objetivo metodológico central de este trabajo es transferir la carga de la complejidad algorítmica en lugar de utilizar una red estructuralmente grande con funciones matemáticas simples, se propone utilizar una red estructuralmente pequeña basada en operadores matemáticos avanzados (cálculo fraccionario y conformable). Las pruebas preliminares demostraron que la configuración de dos neuronas (16:2:1) proporciona el equilibrio ideal entre capacidad de mapeo geométrico y ligereza computacional (Sivanandam & Deepa, 2006).

7.1.3. Dinámica del cálculo sináptico

Tal como se infiere del esquema topológico (Figura 7), el proceso de cálculo al interior de la red ocurre mediante operaciones algebraicas secuenciales. En la primera fase de la capa oculta, cada nodo realiza una combinación lineal de las 16 señales de entrada (x_i), multiplicándolas por sus respectivos pesos sinápticos ($W_{i,j}$) y añadiendo un término de sesgo o bias (b_1). Este sesgo

actúa como un umbral de desplazamiento que permite trasladar la función de activación a través del eje cartesiano, brindando mayor flexibilidad geométrica para ajustarse a los datos termodinámicos.

Este argumento de entrada ponderado ingresa al "motor" de la red: la función de activación(ψ_1). Es en este nodo específico donde recae la principal innovación metodológica de la tesis, ya que las funciones evaluadas(ψ) transicionan desde operadores sigmoides de orden entero hasta integrales de convolución de Mittag-Leffler y derivadas locales conformables.

Finalmente, la señal no lineal resultante de las neuronas ocultas se propaga hacia la capa de salida. En esta última etapa, las señales se combinan linealmente mediante un nuevo conjunto de pesos($W_{1,2}$) y un sesgo final(b_2), atravesando una función de transferencia de salida(ψ_2) — típicamente una función lineal (purelin) para problemas de regresión termodinámica— con el fin de proyectar el valor escalar definitivo del COP (y).

7.2. Algoritmo de optimización

El entrenamiento de una Red Neuronal Artificial no es un simple procedimiento de ajuste de curvas, sino un problema matemático complejo de optimización no lineal multidimensional. El objetivo central de este proceso es navegar a través de un hiperespacio de pesos sinápticos y sesgos para encontrar el mínimo global de una función de costo; en este caso, la minimización del Error Cuadrático Medio (MSE) entre el COP empírico medido en el equipo de absorción y el COP estimado por la arquitectura matemática.

El algoritmo de optimización juega un papel muy importante en la velocidad de convergencia para los ajustes de los parámetros de una RNA. Se trabajará el algoritmo de

Levenberg–Marquardt (Eq. 8) ya que es el algoritmo más utilizado en redes de tipo Feed Forward, su ecuación es:

$$\mathbf{H} = \mathbf{J}^T \mathbf{J}, \quad (8)$$

donde:

H: es la Matriz Hessiana aproximada

J: es la Matriz Jacobiana que Contiene las derivadas parciales de los errores de la red respecto a los parámetros (pesos y sesgos).

Derivado de esto (Eq. 9):

$$\nabla f = \mathbf{J}^T \mathbf{e}, \quad (9)$$

donde:

∇f : es el gradiente del error.

7.2.1. Limitaciones de los métodos de primer y segundo orden puros

El enfoque más elemental para resolver esta minimización es el algoritmo de Descenso de Gradiente, un método de primer orden que actualiza los pesos en la dirección opuesta al vector gradiente de la superficie de error. Aunque garantiza matemáticamente una reducción del costo en cada iteración, el descenso de gradiente puro converge de manera extremadamente lenta, especialmente en "valles" estrechos de la superficie de error. Esta topología de valles prolongados es sumamente común cuando se modelan fenómenos fuertemente acoplados, como la interdependencia entre la temperatura del generador y la presión del absorbedor en el AHT.

Para acelerar la búsqueda, la teoría de optimización propone métodos de segundo orden basados en la serie de Taylor, como el algoritmo de Newton o Gauss-Newton (Nocedal & Wright, 2006). De acuerdo con la literatura fundamental en aprendizaje profundo (Goodfellow et al., 2016), aunque los métodos puros de segundo orden ofrecen tasas de convergencia cuadrática teóricamente superiores, su alta susceptibilidad a matrices Hessianas mal condicionadas los vuelve inestables frente a datos experimentales ruidosos. Estos métodos evalúan la curvatura del espacio de error calculando la matriz Hessiana, lo que les permite realizar pasos gigantescos hacia el mínimo, logrando una convergencia cuadrática sumamente veloz. Sin embargo, en el modelado de datos termodinámicos experimentales —los cuales contienen inherentemente ruido de instrumentación y perturbaciones estocásticas— la matriz Hessiana tiende a volverse mal condicionada o no definida positiva. Cuando esto ocurre, el algoritmo de Gauss-Newton se desestabiliza por completo, provocando que la red diverja hacia errores computacionalmente infinitos y paralizando el entrenamiento de los pesos fraccionarios o conformables.

7.2.2. Formulación híbrida: La interpolación de Levenberg y Marquardt

Para resolver este dilema entre estabilidad y velocidad, el entrenamiento de todas las arquitecturas evaluadas en este trabajo fue impulsado por el algoritmo Levenberg-Marquardt (LM). Desarrollado independientemente por Kenneth Levenberg y Donald Marquardt, este algoritmo representa una genialidad matemática: funciona como una interpolación adaptativa continua entre la robustez del descenso de gradiente y la agresiva velocidad de convergencia del método de Gauss-Newton (Levenberg, 1944; Marquardt, 1963).

La regla de actualización de los pesos en el algoritmo Levenberg-Marquardt se expresa mediante la siguiente ecuación (Eq. 10) matricial fundamental:

$$\Delta W = -[J^T J + \mu I]^{-1} J^T e, \quad (10)$$

donde:

ΔW : es el vector de actualización de los pesos sinápticos.

J : representa la matriz Jacobiana, la cual contiene las primeras derivadas parciales de los errores de la red con respecto a cada peso y sesgo de la arquitectura.

e : es el vector de los errores predictivos en la iteración actual.

I : es la matriz identidad.

μ : es el factor de amortiguamiento (damping factor), un escalar clave que altera dinámicamente el comportamiento de la ecuación.

La principal ventaja de esta formulación para el pronóstico del rendimiento térmico reside en el comportamiento del escalar μ (Hagan & Menhaj, 1994). Cuando el valor de μ es grande, el término μI domina la matriz, y el algoritmo se comporta geométricamente como un descenso de gradiente con una tasa de aprendizaje pequeña. Esto garantiza que la red neuronal explore las zonas caóticas del espacio de error de manera segura y estable. Por el contrario, a medida que la red se acerca al mínimo global y el error predictivo del COP comienza a disminuir, el algoritmo reduce automáticamente el valor de μ hacia cero. Al hacerlo, el término μI desaparece de la ecuación y el sistema se transforma fluidamente en el método de Gauss-Newton ($[J^T J]^{-1} J^T e$), acelerando hacia el fondo del "valle" con una precisión cuadrática devastadora (Yu & Wilamowski, 2011).

7.2.3. Justificación técnica en el modelado de orden no entero

La decisión de emplear Levenberg-Marquardt frente a optimizadores modernos de orden estocástico (como Adam o RMSprop) está firmemente respaldada por la naturaleza del problema físico abordado. Aunque optimizadores como Adam son el estándar para bases de datos masivas

(como millones de imágenes), son ineficientes para alcanzar un margen de error cercano a cero en redes neuronales estructuralmente pequeñas pero matemáticamente densas (arquitecturas 16-2-1). El algoritmo LM es el estándar de oro (gold standard) de la industria para problemas de regresión no lineal donde la precisión numérica absoluta es prioritaria y el número de pesos sinápticos a optimizar está en el rango de unos cientos o menos. Su mayor ventaja computacional es que aproxima la matriz Hessiana ($H \approx J^T J$) utilizando únicamente el producto interno de la matriz Jacobiana. Al no requerir el cálculo explícito de segundas derivadas cruzadas, LM logra integrar de manera armónica la regla de la cadena clásica con los nuevos operadores conformables y de Mittag-Leffler implementados en esta investigación. Esta compatibilidad permite que el motor de ajuste de la red evalúe con máxima precisión la sensibilidad del COP ante fluctuaciones infinitesimales en el ciclo de absorción, garantizando la estabilización del error (RMSE) a los niveles requeridos por las estrictas pruebas estadísticas de linealidad y confiabilidad presentadas en los resultados.

7.3. Normalización de las variables

Es necesario realizar una normalización en las variables de la base de datos porque muchas de las funciones de activación solo funcionan de manera óptima en un rango de operación específico, lo cual también es ideal para tener un mayor control en el análisis de datos. Para llevar a cabo este proceso, se empleó la normalización descrita previamente en la Ecuación (2).

El software que se utilizó para realizar el entrenamiento en busca de los mejores parámetros para todos los modelos de RNA será Matlab, una vez que fueron encontrados los mejores parámetros ajustados a cada modelo, estos son guardados en un archivo de tipo .txt. Posteriormente fue realizado un código donde se cargan los mejores parámetros ajustados y se realiza la

simulación que el usuario desee, este solo debe especificar la función a utilizar y si será para una o dos neuronas ocultas, este código se anexa en el documento.

7.4. Análisis de Sensibilidad de Variables de Entrada

Se realizó un análisis de sensibilidad con el objetivo de cuantificar la influencia relativa de cada una de las 16 variables de entrada sobre la variable objetivo, el Coeficiente de Performance (COP). Para asegurar la robustez de las conclusiones, se emplearon dos metodologías analíticas distintas que se complementan entre sí: un análisis de Correlación de Pearson y un análisis de Importancia de Variables basado en el modelo Random Forest.

7.4.1. Metodología de Análisis

A. Correlación de Pearson (Análisis Lineal)

El coeficiente de correlación de Pearson (r) es una medida estadística que cuantifica la fuerza de la relación lineal entre dos variables. El coeficiente varía entre -1 y +1:

Un valor cercano a +1 indica una fuerte correlación positiva (si una variable aumenta, la otra también tiende a aumentar).

Un valor cercano a -1 indica una fuerte correlación negativa (si una variable aumenta, la otra tiende a disminuir).

Un valor cercano a 0 indica una ausencia de relación lineal.

Este método es fundamental para identificar tendencias directas y fuertes entre una entrada y la salida.

B. Importancia de Variables (Análisis No Lineal con Random Forest)

Este es un método más avanzado basado en machine learning. Un modelo Random Forest (Bosque Aleatorio) se compone de un gran número de árboles de decisión. El modelo se entrena para predecir el COP a partir de las 16 variables de entrada.

La "importancia" de una variable se determina midiendo cuánto empeora la precisión del modelo (aumento en el error) cuando los valores de esa única variable se permutan aleatoriamente. Si al "desordenar" una variable, el modelo falla estrepitosamente en sus predicciones, se considera que dicha variable es muy importante. Este método tiene la ventaja de capturar relaciones no lineales y de interacción que el análisis de Pearson no puede detectar.

7.4.2. Resultados Completos del Análisis

Las Tablas 3 y 4 comparan los resultados de ambos métodos para cada variable de entrada.

Tabla 3. Análisis de Pearson

Nombre de la variable	Variable	Coefficiente de Correlación
$T_{in,EV}$ (°C)	9	0.664
$T_{in,GE-AB}$ (°C)	1	0.49
$T_{out,GE-AB}$ (°C)	5	0.489
$T_{in,CO}$ (°C)	7	0.428
$T_{out,CO}$ (°C)	8	0.419
$T_{out,AB-GE}$ (°C)	3	0.38
$T_{out,EV-AB}$ (°C)	10	-0.286
$T_{in,EV-AB}$ (°C)	2	-0.285
$T_{in,AB-GE}$ (°C)	4	0.207

$T_{\text{out,GE-CO}}$ (°C)	6	0.165
$X_{\text{in,AB}}$ (%)	11	-0.106
P_{GE} (en.Hg)	16	-0.098
$X_{\text{out,GE}}$ (%)	14	-0.097
$X_{\text{out,AB}}$ (%)	12	-0.091
$X_{\text{in,GE}}$ (%)	13	-0.091
P_{AB} (en.Hg)	15	-0.08

Tabla 4. Análisis de Random Forest

Nombre de la Variable	Número de variable	Puntuación de Importancia
$T_{\text{in,EV}}$ (°C)	9	0.457 (45.7%)
$T_{\text{in,GE-AB}}$ (°C)	1	0.126 (12.6%)
$T_{\text{out,GE-AB}}$ (°C)	5	0.118 (11.8%)
$T_{\text{in,CO}}$ (°C)	7	0.054 (5.4%)
$T_{\text{out,CO}}$ (°C)	8	0.051 (5.1%)
$T_{\text{out,AB-GE}}$ (°C)	3	0.043 (4.3%)
$T_{\text{out,GE-CO}}$ (°C)	6	0.027 (2.7%)
$T_{\text{in,AB-GE}}$ (°C)	4	0.024 (2.4%)
$T_{\text{in,EV-AB}}$ (°C)	2	0.023 (2.3%)
$T_{\text{out,EV-AB}}$ (°C)	10	0.022 (2.2%)

$X_{in,AB}$ (%)	11	0.019 (1.9%)
P_{AB} (en.Hg)	15	0.011 (1.1%)
$X_{in,GE}$ (%)	13	0.009 (0.9%)
$X_{out,AB}$ (%)	12	0.009 (0.9%)
$X_{out,GE}$ (%)	14	0.004 (0.4%)
P_{GE} (en.Hg)	16	0.004 (0.4%)

7.4.3. Interpretación de resultados

Variables con mayor influencia

La conclusión más significativa del análisis es la fuerte coincidencia de ambos métodos en la identificación de las variables con mayor impacto en la predicción del COP.

$T_{in,EV}$ es, indiscutiblemente, el parámetro más influyente. Lidera el análisis de Random Forest (45.7% de la importancia) y presenta la correlación lineal más fuerte ($r = 0.664$).

$T_{in,GE-AB}$ y $T_{out,GE-AB}$ se posicionan claramente como la segunda y tercera variables más importantes en ambos análisis, con puntuaciones de importancia significativas (12.6% y 11.8%) y las siguientes correlaciones más altas.

El hecho de que los tres parámetros principales sean los mismos en ambos análisis valida robustamente que el rendimiento del sistema está primordialmente gobernado por estas tres temperaturas.

Variables de Baja Influencia

De igual manera, ambos métodos coinciden en las variables con el impacto más bajo. Todas las variables de Concentración y Presión se agrupan al final de la lista. Sus puntuaciones de

importancia son mínimas (todas por debajo del 2%) y sus coeficientes de correlación son muy cercanos a cero, indicando una influencia despreciable en el COP dentro de los rangos experimentales estudiados.

Variables de Influencia Media

Se observan ligeras variaciones en el orden de las variables de influencia media. Por ejemplo, $T_{in,EV-AB}$ y $T_{out,EV-AB}$ tienen una correlación lineal negativa más fuerte que $T_{out,GE-CO}$, pero el modelo Random Forest las clasifica por debajo. Esto es esperable y resalta la ventaja de usar ambos métodos: el análisis de Pearson puede haber sido "engañado" por una tendencia lineal moderada, mientras que el Random Forest determinó que, a pesar de esa tendencia, la variable no era tan crítica para la predicción general debido a patrones no lineales.

Sin embargo, no es necesario centrarse en estas variaciones menores en la parte media de la tabla. El hallazgo fundamental es que ambos métodos separaron clara y consistentemente el mismo grupo de variables en un "Top 3" de alta influencia.

7.5. Métricas estadísticas

La evaluación de modelos de Redes Neuronales Artificiales (RNA) requiere de un análisis estadístico para la validación de su eficacia y confiabilidad.

Los cálculos estadísticos son fundamentales para la evaluación de cada modelo de RNA. Estos datos nos ayudan a poder determinar si el modelo es adecuado o no lo es, esto depende del sistema físico que se está estudiando y del criterio del usuario. La Tabla 5 muestra las métricas estadísticas que usaré para determinar la calidad de cada uno de los modelos de RNA. Es importante señalar que la "calidad" de un modelo no es un concepto absoluto, sino relativo: un mismo RNA podría considerarse óptimo para aplicaciones en tiempo real que priorizan velocidad

de inferencia, pero insuficiente en entornos científicos que exigen precisión microscópica. Esta dualidad refuerza la necesidad de que el usuario interprete las métricas desde una perspectiva crítica, ponderando siempre los resultados de exactitud, eficiencia computacional y adaptabilidad contextual. Eq. (11), (12) y (13) muestran las métricas estadísticas que serán contempladas para este estudio.

Tabla 5. Ecuaciones estadísticas para evaluar la eficiencia de los modelos de RNA

Nombre de la ecuación	Ecuación
Correlación lineal de Pearson (R)	$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} . \quad (11)$
Error Cuadrático Medio (MSE)	$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 . \quad (12)$
Raíz del Error Cuadrático Medio (RMSE)	$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} . \quad (13)$

--	--

7.5.1. Criterios de dispersión y tendencia direccional

La evaluación de arquitecturas neuronales orientadas a la predicción termodinámica requiere métricas que penalicen distintos tipos de desviaciones. El Coeficiente de Correlación de Pearson (R) proporciona una medida estandarizada sobre la sincronía direccional entre los valores observados y las predicciones del modelo. Un valor de R cercano a la unidad confirma que la red ha logrado asimilar la tendencia general del ciclo de absorción; sin embargo, en la validación de modelos físicos, esta métrica es insuficiente de forma aislada, ya que no detecta sesgos sistemáticos (Willmott & Matsuura, 2005). Por ejemplo, un modelo que prediga temperaturas consistentemente superiores a las reales por un factor constante podría registrar un coeficiente R elevado, a pesar de ser físicamente inexacto.

Para contrarrestar esta limitación, el desempeño direccional se complementa con la Raíz del Error Cuadrático Medio ($RMSE$). A diferencia del MSE , que cuantifica la varianza del error en magnitudes al cuadrado, el $RMSE$ devuelve la penalización a las unidades dimensionales originales del parámetro objetivo (adimensional en el caso del COP). La formulación cuadrática intrínseca del $RMSE$ le otorga una alta sensibilidad ante valores atípicos (outliers), obligando a la red neuronal a minimizar los errores residuales en los escenarios de operación transitoria más severos del transformador de calor (Chai & Draxler, 2014).

7.5.2. Validación paramétrica: Prueba de linealidad de pendiente e intercepto

Dado que el objetivo central del modelo es sustituir mediciones físicas complejas, la literatura establece que un modelo neuronal solo adquiere grado de confiabilidad ingenieril si demuestra equivalencia estadística uno a uno con el fenómeno físico (Montgomery & Runger, 2010). Para demostrar formalmente este grado de adaptación, todas las arquitecturas desarrolladas en esta investigación fueron sometidas a una prueba de linealidad mediante análisis de regresión ordinaria por mínimos cuadrados.

La prueba asume que la relación ideal entre los datos experimentales(y_{exp}) y las predicciones del modelo(y_{sim}) debe obedecer a una función lineal estricta descrita por la ecuación (Eq. 14) de la recta:

$$y_{sim} = S \cdot y_{exp} + I, \quad (14)$$

donde S representa la pendiente de la recta de regresión e I representa el coeficiente de intersección con el eje de las ordenadas (intercepto). En un modelo con capacidad predictiva teóricamente exacta, la pendiente adquiriría un valor de $S = 1$ y el intercepto cruzaría el origen en $I = 0$.

No obstante, debido a las incertidumbres operacionales inherentes al equipo físico y al ruido estocástico de los sensores, se requiere construir intervalos de confianza estadísticos. Para este estudio, se definió un nivel de confianza del 99. Los intervalos superior e inferior para ambos parámetros se determinaron aplicando la distribución t de Student, según las siguientes inecuaciones:

$$S_{lower} \leq S \leq S_{upper},$$

$$I_{lower} \leq I \leq I_{upper}.$$

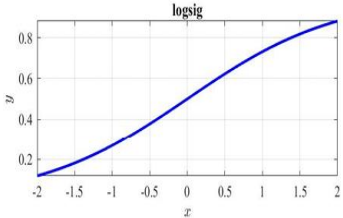
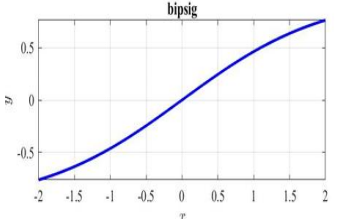
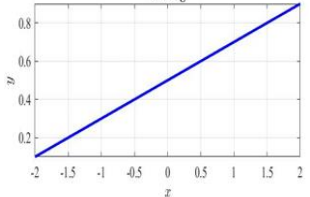
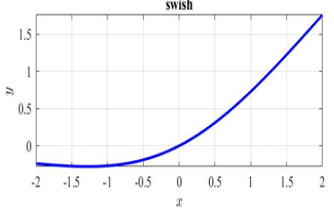
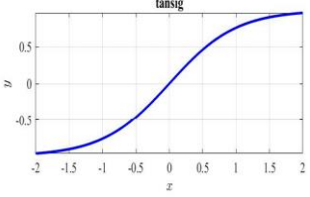
Para que una configuración neuronal (como la estructura 16:2:1) se considere robusta y carente de sesgos sistemáticos locales, la unidad (1) debe estar estrictamente contenida dentro del intervalo de confianza de la pendiente, y el cero (0) debe estar contenido dentro del intervalo de confianza del intercepto (Piñeiro et al., 2008).

Esta validación paramétrica constituye el filtro metodológico más riguroso del presente trabajo. Como se documenta en la literatura previa sobre modelado de ciclos térmicos, es común que las funciones de transferencia de orden entero logren minimizar el *RMSE*, pero presenten desviaciones finas en los límites operativos del ciclo de absorción (por ejemplo, en las zonas cercanas al límite de cristalización de la sal), fallando en satisfacer estas inecuaciones. Evaluar los modelos fraccionarios y conformables bajo esta métrica asegura que cualquier mejora reportada corresponda a una verdadera asimilación física del proceso y no a un ajuste numérico superficial.

7.6. Funciones de transferencia a evaluar

La evolución constante en el campo de las redes neuronales artificiales (RNA) ha impulsado el desarrollo de numerosas funciones de activación. La evaluación de distintas funciones de activación es crítica porque cada una introduce propiedades matemáticas únicas que condicionan la capacidad de la red para aprender patrones complejos y adaptarse a las particularidades del sistema en estudio. En este trabajo serán evaluadas la mayor cantidad de funciones posibles, entre las funciones a evaluar se incluirán no solo opciones ampliamente utilizadas, como la sigmoide o la ReLU, sino también alternativas recientes como Swish o GELU. Esto con el fin de determinar cual de ellas tiene el mejor ajuste al sistema con el que estará trabajando. Las Tablas 6, 7 y 8 muestran la ecuación y gráfica del comportamiento de estas funciones.

Tabla 6. Funciones de activación sigmoid, bipolar sigmoid, hard sigmoid, swish, tansig, hard tansig, lisht

Nombre de la función	Ecuación	Gráfica
Sigmoid	$f(x) = \frac{1}{1 + e^{-x}} \cdot (15)$	
Bipolar Sigmoid	$f(x) = -1 + \frac{2}{1 + e^{-x}} \cdot (16)$	
Hard Sigmoid	$f(x) = \begin{cases} 0 & \text{if } x < -2.5, \\ 0.2x + 0.5 & \text{if } -2.5 \leq x \leq 2.5, \\ 1 & \text{if } x > 2.5. \end{cases} (17)$	
Swish	$f(x) = x \cdot \frac{1}{1 + e^{-x}} \cdot (18)$	
Tansig	$f(x) = \frac{2}{1 + e^{-2x}} - 1 \cdot (19)$	

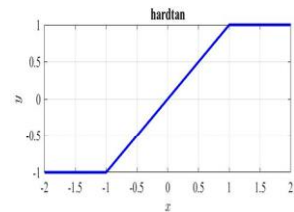
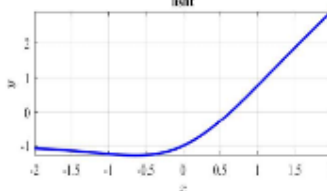
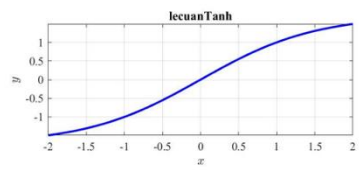
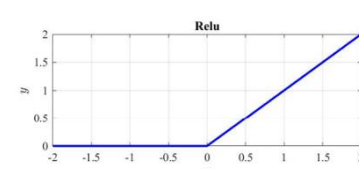
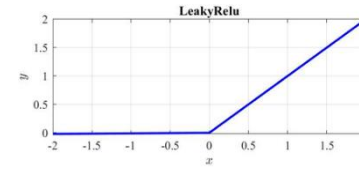
Hard Tansig	$f(x) = \begin{cases} -1 & \text{if } x < -1, \\ x & \text{if } -1 \leq x \leq 1, \\ 1 & \text{if } x > 1. \end{cases} \quad (20)$	
LISHT	$f(x) = x \cdot \left(\frac{2}{1 + e^{-2x}} \right) - 1. \quad (21)$	

Tabla 7. Funciones de activación lecuTanh, Relu, Leaky Rely, Parametric Relu, Randomized Relu, Elo, Parametric Elu

Nombre de la función	Ecuación	Gráfica
LeCun Tanh	$f(x) = 1.7159 \cdot \left(\frac{e^x - e^{-x}}{e^x + e^{-x}} \right) \cdot \left(\frac{2}{3}x \right). \quad (22)$	
ReLU	$f(x) = \max(0, x). \quad (23)$	
Leaky ReLU	$f(x) = \max(0.01x, x). \quad (24)$	

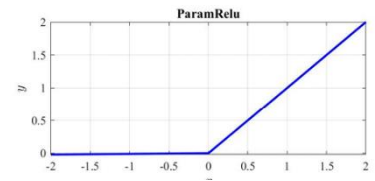
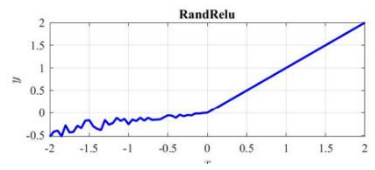
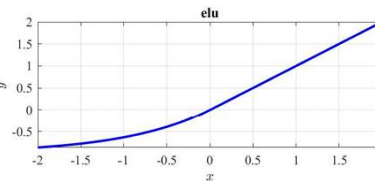
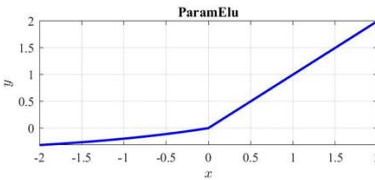
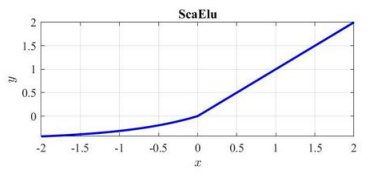
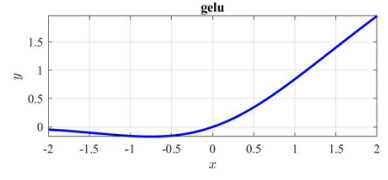
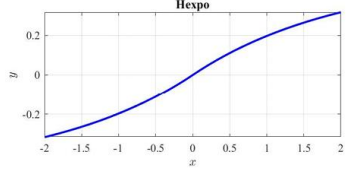
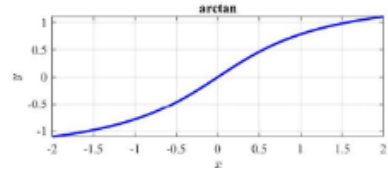
Parametric ReLU	$f(x) = \max(\alpha x, x). \quad (25)$	
Randomized ReLU	$f(x) = \max(Rx, x). \quad (26)$	
ELU	$f(x) = \begin{cases} x & \text{if } x > 0, \\ b(e^x - 1) & \text{if } x \leq 0. \end{cases} \quad (27)$	
Parametric ELU	$f(x) = \begin{cases} dx & \text{if } x > 0, \\ b(e^{x/c} - 1) & \text{if } x \leq 0. \end{cases} \quad (28)$	

Tabla 8. Funciones de activación Scaled Elu, Gelu, Hexpo, Arctan

Nombre de la función	Ecuación	Gráfica
Scaled ELU	$f(x) = \mu \begin{cases} x & \text{if } x > 0, \\ h(e^x - 1) & \text{if } x \leq 0. \end{cases} \quad (29)$	

GELU	$f(x) = 0.5x \left(1 + \tanh \left(\sqrt{2/\pi} (x + 0.044715x^3) \right) \right). \quad (30)$	
Hexpo	$f(x) = \begin{cases} -a \left(e^{-\frac{x}{b}} - 1 \right) & \text{if } x \geq 0, \\ c \left(e^{\frac{x}{d}} - 1 \right) & \text{if } x < 0. \end{cases} \quad (31)$	
ArcTan	$f(x) = \arctan(x). \quad (32)$	

8. Resultados de la simulación y análisis comparativo

Una vez definida la arquitectura base de la red neuronal y tras evidenciar que las funciones de transferencia clásicas de orden entero resultan insuficientes para superar de manera estricta las pruebas de linealidad en los regímenes operativos del equipo, este capítulo expone la evaluación predictiva de los modelos matemáticos avanzados. A continuación, se presentan los resultados de las simulaciones desarrolladas mediante la integración de operadores de cálculo fraccionario y conformable. Asimismo, se detalla la reestructuración del entorno de entrenamiento hacia el lenguaje de programación Python —implementada para optimizar el manejo de datos y reducir la carga computacional—, culminando con un análisis estadístico integral que contrasta la superioridad de estas nuevas arquitecturas frente a los modelos históricos reportados en el estado del arte.

8.1 Modelos de orden entero

Estas funciones fueron evaluadas en modelos de RNA para 1 y 2 neuronas ocultas y de manera general, los resultados obtenidos fueron muy bueno. La Tabla 9 muestra los resultados obtenidos para cada uno de estos modelos de orden entero.

Tabla 9. Tabla comparativa de resultados obtenidos con 1 neurona oculta y con 2 neuronas ocultas

Función	Arquitectur a	R	MSE	RMSE	Intercept ($I \pm 0$)		Slope ($S \pm 1$)	
					I _{Upper}	I _{Lower}	S _{Upper}	S _{Lower}
TANSIG	16:2:1	0.9984858	0.000008	0.0029	0.00391	0.00083		0.99462
		9	5	3	8	9	0.9988	6
LOGSIG	16:2:1	0.99853	0.00001	0.0029		0.00035		
				3	0.0018	1	0.999	0.9949
ARCTAN	16:2:1	0.99855	0.00001	0.0029				
				2	0.0018	0.00036	0.989	0.9949
LISHT	16:2:1	0.99855	0.00001	0.0029				
				2	0.0018	0.00355	0.999	0.9949
GELU	16:2:1	0.99854	0.00001	0.0029				
				2	0.0018	0.00035	0.989	0.9949
SCALED ELU	16:2:1	0.9985	0.00001	0.0029		0.00046		
				7	0.0019	9	0.986	0.9945

HEXPO	16:2:1	0.99854	0.00001	0.0029 3	0.0018	0.00034 9	0.999	0.9949
dSILU	16:2:1	0.998524	0.00001	0.0029 3	0.00324	0.00030 3	0.9995	0.99547
BIPSIG	16:2:1	0.99853	0.00001	0.0029 4	0.0018	0.00039 6	0.9988	0.9948
SWISH	16:2:1	0.99727	0.00004	0.0059 7	-0.0247	-0.0268	0.0719	0.066
LECUN TANH	16:2:1	0.98364	0.0001	0.0099 7	0.0104	0.0057	0.9775	0.9641
HARD TANSIG	16:2:1	0.9984	0.00001	0.0030 7	0.0021	0.00059 2	0.9984	0.9941
ELU	16:2:1	0.9984	0.0002	0.0140 2	0.0253	0.021	0.9116	0.8994
HARD SIGMOID	16:2:1	0.99839	0.00001	0.0030 7	0.0021	0.00059 2	0.9987	0.9944
LEAKY RELU	16:2:1	0.99839	0.00001	0.0030 7	0.0019	0.00044 1	0.9987	0.9944
RELU	16:2:1	0.99839	0.00001	0.0030 7	0.0019	0.00044 2	0.9987 1	0.9944
RAND RELU	16:2:1	0.98409	0.00009	0.0096 3	0.0131	0.0085	0.9754	0.9622
PARAM ELU	16:2:1	0.99844	0.00001	0.0030 3	0.0018	0.00029 7	0.9991	0.9949

PARAM RELU	16:2:1	0.99839	0.00001	0.0030 7	0.002	0.00046 5	0.9986	0.9944
ELLIOT	16:2:1	0.9985	0.00001	0.0029 6	0.0018	0.00035 5	0.9989	0.9948
SIGMOID								
SOFT MAX	16:2:1	0.98797	0.00007	0.0083 8	0.0102	0.0061	0.9823	0.9708
SQSINC	16:2:1	0.99852	0.00001	0.0029 5	0.0019	0.00041 8	0.9988	0.9947
SOFTPLUS	16:2:1	0.99849	0.00001	0.0029 8	0.00189	0.00042 2	0.9988	0.9946

En la Tabla 9 se puede observar el nombre de las funciones de activación que fueron evaluadas, así como sus valores correspondientes en las métricas estadísticas “R”, pendiente inferior, pendiente superior, intercepto inferior, intercepto superior, MSE, RMSE, tanto para una como para 2 neuronas ocultas.

En las pruebas realizadas se lograron obtener resultados muy eficientes, sin embargo, ningún modelo de 1 neurona oculta logró superar a lo ya reportado en la literatura, en algunos de los modelos de 2 neuronas ocultas se logró obtener mejores resultados en las métricas estadísticas, sin embargo, todavía no se logra superar la prueba de linealidad.

Las funciones de activación relacionadas a los mejores modelos de RNA de 2 neuronas ocultas fueron:

- Tansig.
- Logsig.
- Bipolar sigmoid.

- Arctan.
- liSHT.
- Gelu.
- Hexpo.
- Dsilu.
- Sqsinc.

Todas ellas superaron a las métricas previamente reportadas en este sistema, donde emplearon funciones de orden entero y funciones fraccionarias.

A continuación, las Figuras 8 a 16 exponen las gráficas correspondientes al cálculo del coeficiente de correlación de los mejores modelos evaluados:

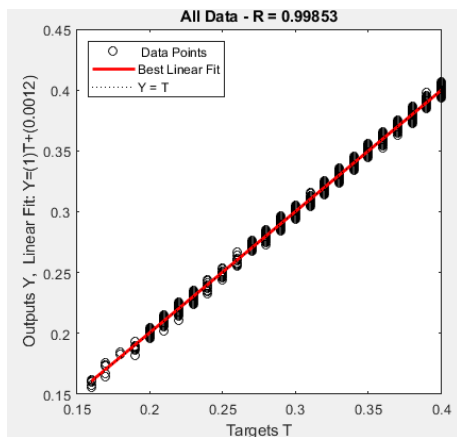


Figura 8. Coeficiente Correlación lineal tansig 16-2-1

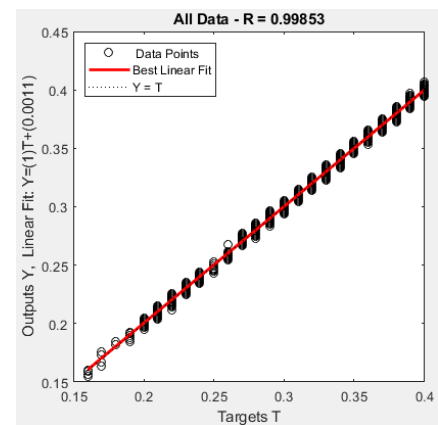


Figura 9. Coeficiente Correlación lineal logsig 16-2-1

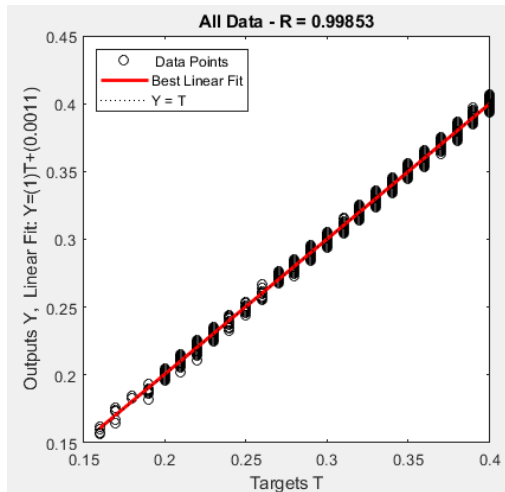


Figura 10. Coeficiente Correlación lineal bipsig 16-2-1

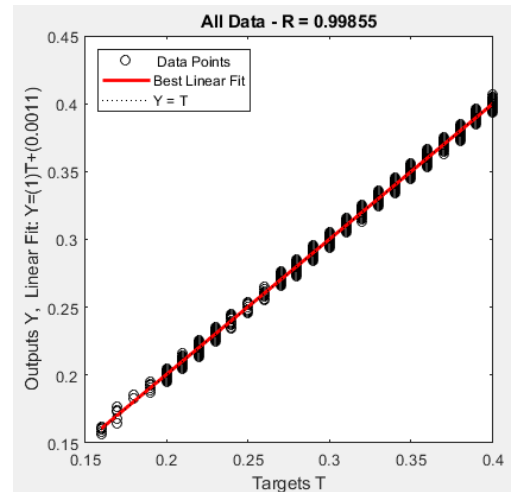


Figura 11. Coeficiente Correlación lineal arctan 16-2-1

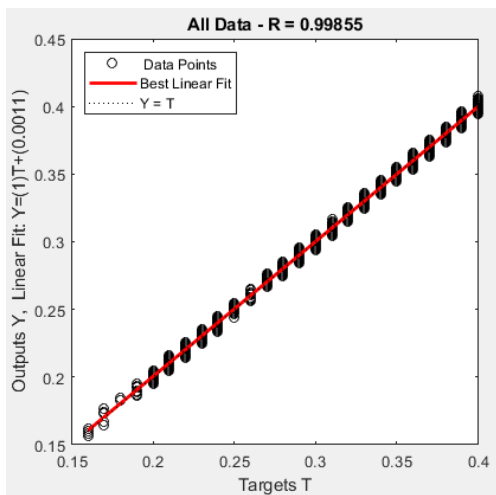


Figura 12. Coeficiente Correlación lineal liSHT 16-2-1

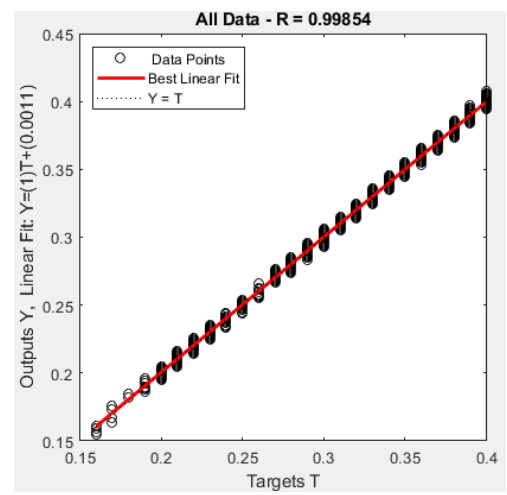


Figura 13. Coeficiente Correlación lineal Hexpo 16-2-1

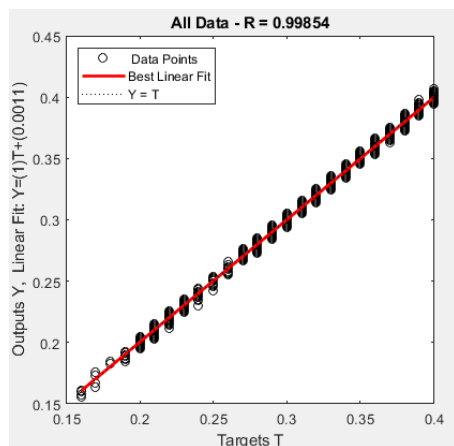


Figura 14. Coeficiente Correlación lineal Gelu 16-2-1

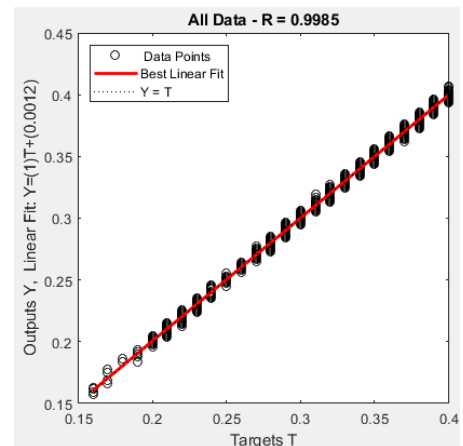


Figura 15. Coeficiente Correlación lineal ScaElu 16-2-1

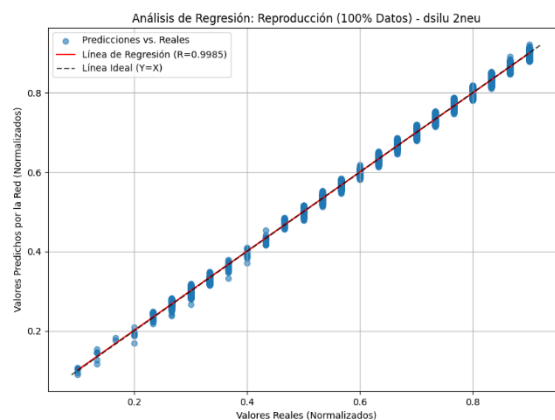


Figura 16. Coeficiente Correlación lineal dSilu 16-2-1

Al observar las gráficas de regresión correspondientes a los mejores modelos de orden entero —como los desarrollados con las funciones tansig, logsig, bipsig, arctan y dsilu—, el primer aspecto que destaca es la alta correspondencia entre los valores reales y las predicciones de la red neuronal. Visualmente, los puntos de datos se agrupan de manera muy estrecha sobre la línea de ajuste ideal, aquella donde el valor predicho es exactamente igual al valor experimental. Esto

indica que el modelo logró aprender la dinámica general de la bomba de calor y es capaz de estimar el Coeficiente de Rendimiento (COP) con un margen de error bastante reducido en la mayor parte del rango de operación.

Sin embargo, en el modelado de sistemas térmicos complejos, la apreciación visual y el coeficiente de correlación no son suficientes para validar por completo una herramienta predictiva. Aunque los diagramas muestran valores de correlación muy favorables, cercanos a 0.998, el análisis estadístico detallado revela una limitación importante. Al someter estas predicciones a la prueba de linealidad para evaluar la pendiente y el intercepto con un nivel de confianza del 99%, los modelos no logran cumplir con los criterios de aceptación.

En términos prácticos, esto significa que, aunque la línea de ajuste del modelo parece coincidir con la línea ideal, existen desviaciones sistemáticas muy finas. Es común que en los extremos de operación del sistema —cuando las temperaturas y presiones alcanzan sus valores máximos o mínimos— la red neuronal de orden entero pierda un poco de capacidad de adaptación y genere un ligero sesgo. Estas pequeñas variaciones, que son casi imperceptibles a simple vista en las gráficas de dispersión, son detectadas por la rigurosidad de la prueba estadística.

Esta situación demuestra que las redes de orden entero hicieron el mejor trabajo posible dentro de sus limitaciones matemáticas. Es justamente esta pequeña brecha estadística, la cual impide validar el modelo al cien por ciento, la que justifica la necesidad de integrar las funciones fraccionarias avanzadas y el cálculo conformable en la siguiente fase de la investigación.

A continuación, se plantean los modelos con funciones de orden no entero (Eq. (33) – (37)):

8.2 Modelos de orden fraccionario:

Tansig ML
$$\psi^\gamma(z) = \frac{2}{1 + E_\gamma(-2z)} - 1, \quad -1 < \psi^\gamma(z) \leq 1, \quad (33)$$

Logsig ML
$$\psi^\gamma(z) = \frac{1}{1 + E_\gamma(-z)}, \quad 0 < \psi^\gamma(z) \leq 1, \quad (34)$$

Gelu ML
$$\psi^\gamma(z) \approx 0.5z \left(1 + \tanh \left(\sqrt{\frac{2}{\pi}} \cdot (z + 0.044715 z^3) \right) \right), \quad (35)$$

Dsilu ML
$$f(x) = \frac{1 + E_{\{\alpha,\beta\}}^\gamma(-x) + x \left(E_{\{\alpha,\beta\}}^\gamma(-x) \right)}{\left(1 + E_{\{\alpha,\beta\}}^\gamma(-x) \right)^2}, \quad (36)$$

donde E_γ es la función de Mittag-Leffler de orden γ .

Las Figuras 17 a 20 muestran el comportamiento de estas ecuaciones en distintos órdenes:

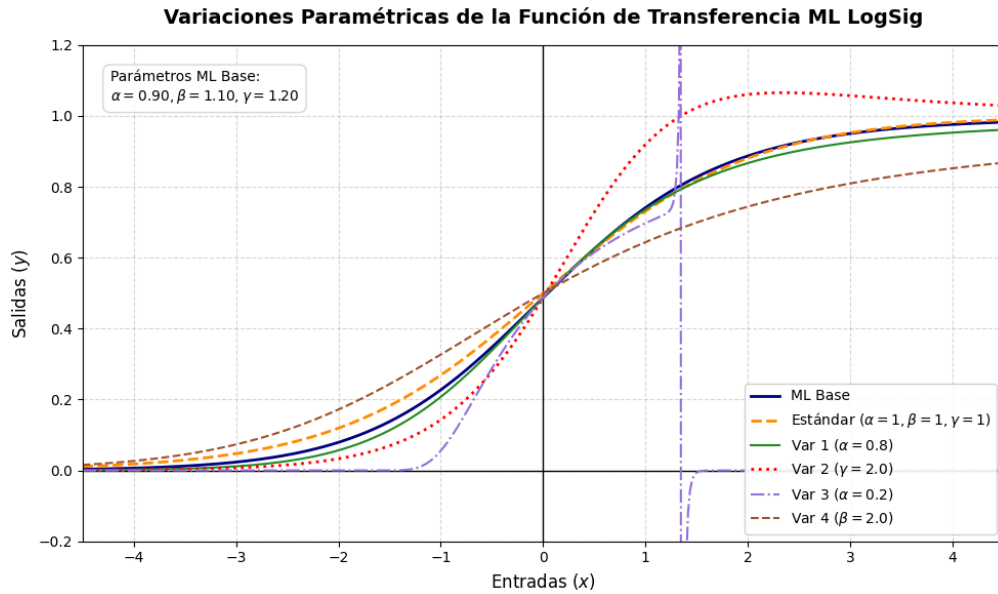


Figura 17. Comportamiento de la función Logsig ML con diferentes ordenes

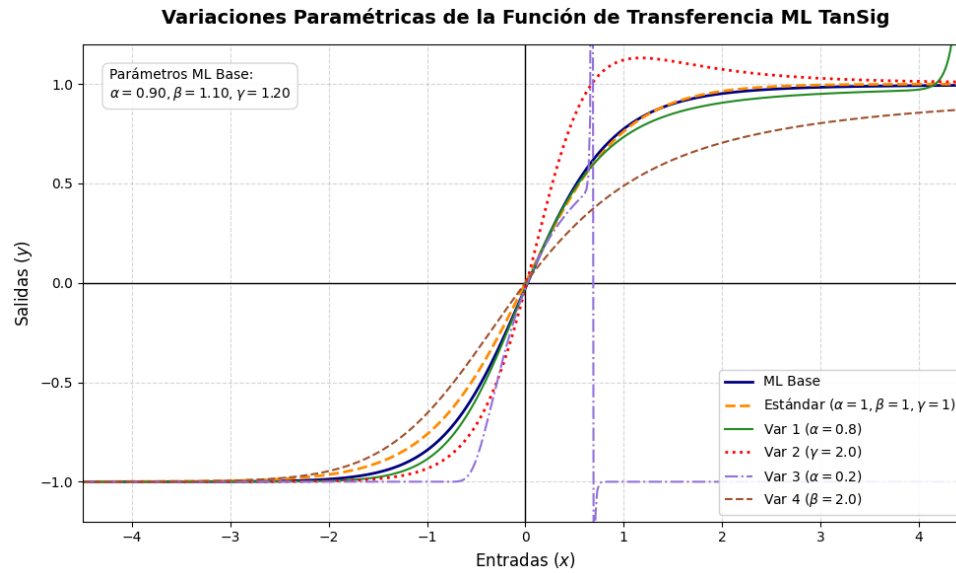


Figura 18. Comportamiento de la función Tansig ML con diferentes ordenes

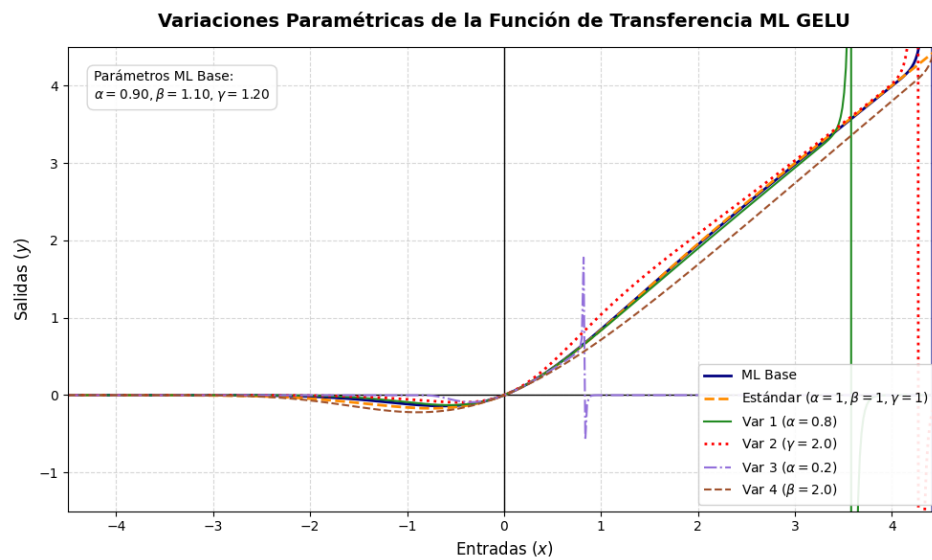


Figura 19. Comportamiento de la función Gelu ML con diferentes ordenes

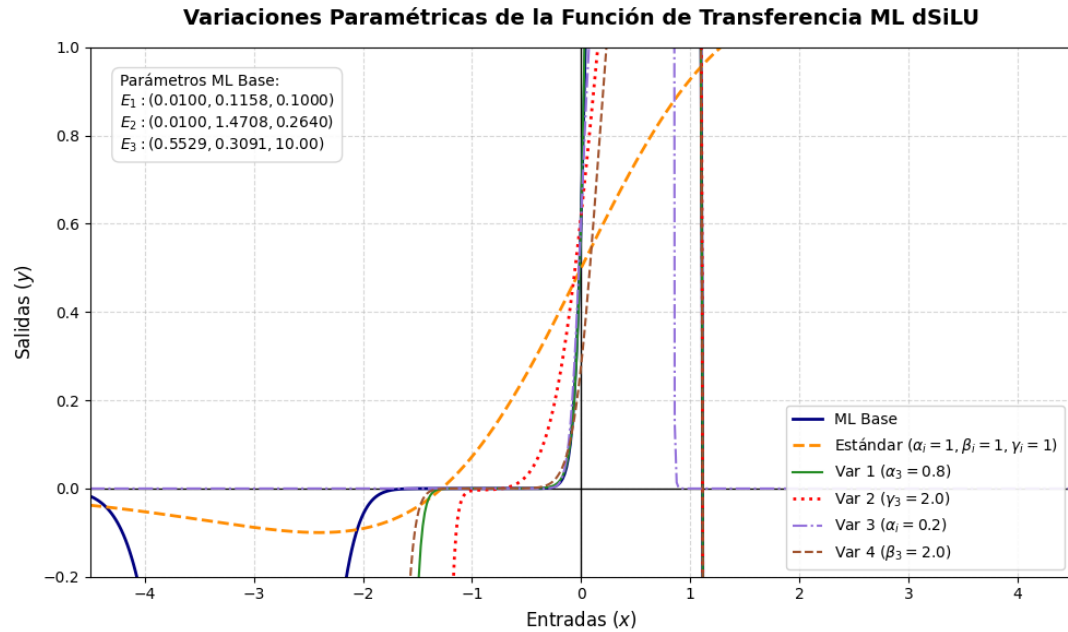


Figura 20. Comportamiento de la función dSilu ML

Se realizaron entrenamientos con estas 4 funciones generalizadoras empleando la función de Mittag-Leffler; la Tabla 10 concentra estos resultados a continuación:

Tabla 10. Tabla comparativa de resultados obtenidos con las ecuaciones de orden fraccionario con estructura 16-2-1

Función 2 neuronas ocultas	R	Inter. inf	Inter. sup	Pen. Inf.	Pen. Sup.
Tansig ML	0.998478	0.00040529	0.0018586	0.994647	0.99878477
Logsig ML	0.998398	0.0024789	0.003284	0.993264	0.997189
Gelu ML	0.998445	0.001370	0.004488	0.993811	0.997992
Dsilu ML	0.99856	-0.00035	0.00171	0.9994	1.00048

Tras evaluar las nuevas opciones matemáticas, el modelo entrenado con la función fraccionaria Dsilu ML destacó claramente por sus resultados. A diferencia de las redes anteriores que conservaban ligeros sesgos, esta arquitectura no solo mantuvo un coeficiente de correlación alto ($R = 0.99856$), sino que logró corregir el problema de adaptación en todo el rango operativo. Al revisar a detalle sus métricas, se observa que los intervalos de confianza finalmente cumplen con la exigencia estadística: el intercepto abarca desde -0.00035 hasta 0.00171 , y la pendiente va de 0.9994 a 1.00048 .

Con estos rangos, el modelo dSilu ML logra pasar satisfactoriamente la estricta prueba de linealidad al 99% de confianza, rompiendo por fin la barrera que limitaba a las funciones de orden entero. Para ilustrar de forma visual cómo se comporta este ajuste en la práctica y la precisión alcanzada, a continuación, se muestra la gráfica de regresión, en la cual se contrastan directamente los valores reales del sistema contra las predicciones arrojadas por la red.

Para ilustrar de forma visual cómo se comporta este ajuste en la práctica y la precisión alcanzada, la Figura 21 contrasta directamente los valores reales del sistema contra las predicciones arrojadas por la red neuronal mediante el análisis de regresión lineal del modelo 16:2:1 dSilu ML.

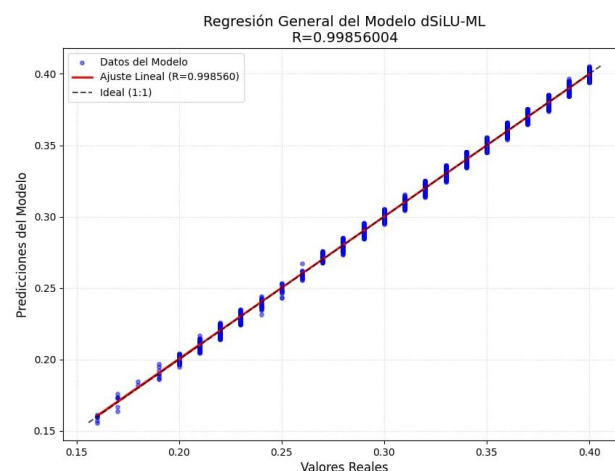


Figura 21. Análisis de regresión lineal dSilu ML

Como punto culminante del análisis predictivo, el desempeño de la arquitectura dSiLU-ML se constata mediante el análisis de regresión lineal. La Figura 21 contrasta visualmente las estimaciones del modelo frente a los valores reales del sistema, revelando el nivel de convergencia alcanzado. Al combinar la estructura base de la función dSiLU con la serie de Mittag-Leffler, el modelo alcanza un coeficiente de correlación de $R = 0.99856$. Este valor marca una diferencia clara, ya que logra superar estadísticamente el límite de 0.99855, un punto donde los modelos con funciones de orden entero se habían estancado.

La Figura 21 muestra una alineación muy precisa entre la línea de ajuste y el comportamiento ideal del sistema. Físicamente, esta reducción en la dispersión en los extremos operativos (valores de COP máximos y mínimos) es un hallazgo crítico. En modelos clásicos de orden entero reportados previamente, las desviaciones en estos extremos sugerían que la red no lograba capturar las fuertes no linealidades cuando la mezcla salina se acercaba a su límite de cristalización o cuando la fuente de calor residual fluctuaba abruptamente. El ajuste ceñido a la línea ideal demuestra que la arquitectura dSiLU-ML actúa como un amortiguador matemático; su naturaleza no local (memoria fraccionaria) asimila los gradientes térmicos abruptos sin saturar la neurona, garantizando predicciones estadísticamente seguras y físicamente viables para el monitoreo del equipo de absorción en tiempo real.

Este avance respalda la idea central del estudio: la estructura propia de la función dSiLU funciona como un buen mecanismo estabilizador. Esto le permite al operador fraccionario ajustar los datos de manera más fina, aprovechando su efecto de "memoria", pero evitando la saturación que limitó el desempeño de otras funciones como Tansig-ML o Logsig-ML.

Una vez validada la superioridad estadística de la arquitectura dSiLU-ML frente a sus contrapartes, el siguiente paso fundamental de esta investigación consiste en consolidar el

algoritmo entrenado como una herramienta predictiva práctica y reproducible. Para lograr esto, es necesario extraer y formalizar la estructura matemática explícita que rige la estimación del Coeficiente de Rendimiento (COP), revelando así el funcionamiento interno de la "caja negra" neuronal.

La estructura matemática general para la estimación de orden fraccionario desarrollada en este trabajo se define mediante la siguiente ecuación (Eq. (37)):

$$COP = \left(\sum_{n=1}^2 [W_{o(1,n)} \cdot dSiLU_{ML}(\eta_n, \alpha_i, \beta_i, \gamma_i)] \right) + b, \quad (37)$$

donde cada término representa un componente físico-matemático de la red:

COP : Es el Coeficiente de Rendimiento estimado del sistema térmico.

n : Representa la neurona oculta específica dentro de la topología 16:2:1 empleada.

$W_{o(1,n)}$: Son los pesos de salida que escalan la contribución individual de cada neurona oculta hacia el resultado final.

$dSiLU_{ML}$: Corresponde a la función de transferencia de orden fraccionario, fundamentada en la serie de Mittag-Leffler de tres parámetros.

η_n : Representa el argumento de entrada ponderado (la suma lineal de variables) que alimenta a la n -ésima neurona oculta.

$\alpha_i, \beta_i, \gamma_i$: Son los parámetros fraccionarios compartidos que gobiernan la forma no local y la "memoria" de la función de transferencia.

b : Es el sesgo (bias) correspondiente a la capa de salida.

Al sustituir los pesos sinápticos optimizados y el sesgo final obtenidos tras la fase de entrenamiento, el modelo matemático explícito queda formulado de la siguiente manera (Eq. 38):

$$COP = -0.0898 \cdot dSiLU_{ML}(\eta_1, \alpha_i, \beta_i, \gamma_i) + 1.3919 \cdot dSiLU_{ML}(\eta_2, \alpha_i, \beta_i, \gamma_i) - 0.9287. \quad (38)$$

La lógica interna de la capa oculta está regida por la función fraccionaria, la cual se define mediante la siguiente relación no lineal que involucra la serie de Mittag-Leffler de tres parámetros (Eq. (39)):

$$dSiLU_{ML}(\eta_n, \alpha_i, \beta_i, \gamma_i) = \frac{1 + E_{\alpha_1, \beta_1}^{\gamma_1}(-\eta_n) + \eta_n \cdot E_{\alpha_2, \beta_2}^{\gamma_2}(-\eta_n)}{\left(1 + E_{\alpha_3, \beta_3}^{\gamma_3}(-\eta_n)\right)^2}, \quad (39)$$

donde $dSiLU_{ML}$: Es la función de transferencia de orden fraccionario que combina las propiedades geométricas de la Unidad Lineal Sigmoidea (Sigmoid Linear Unit) con la expansión en series de Mittag-Leffler.

η_n : Representa el argumento de entrada ponderado para la n -ésima neurona oculta.

$E_{\alpha_i, \beta_i}^{\gamma_i}$: Representa la función Mittag-Leffler de tres parámetros para el componente i -ésimo del mapeo de activación.

$\alpha_i, \beta_i, \gamma_i$: Son los parámetros fraccionarios optimizados que definen de manera precisa la curvatura y la adaptación no local de la función de transferencia frente a los datos térmicos.

Durante la fase de entrenamiento, el algoritmo optimizó y ajustó estos parámetros de activación compartidos. Para garantizar la estabilidad numérica y evitar el sobreajuste, los valores fueron acotados según las restricciones del sistema, obteniendo las siguientes magnitudes:

Para E_1 :

$$\alpha_1 = 0.01, \beta_1 = 0.1158, \gamma_1 = 0.1000,$$

Para E_2 :

$$\alpha_2 = 0.0100, \beta_2 = 1.4708, \gamma_2 = 0.2640,$$

Para E_3 :

$$\alpha_3 = 0.5529, \beta_3 = 0.3091, \gamma_3 = 10.00.$$

Finalmente, las variables intermedias ponderadas (η_1 y η_2), que concentran la información de entrada proveniente de los 16 sensores físicos del transformador de calor, siguen la forma general de sumatoria previamente definida en la metodología. Las expansiones numéricas específicas que alimentan a este modelo de orden fraccionario se detallan en las siguientes ecuaciones (Eq. (40) y (41)):

$$\begin{aligned}\eta_1 = & -0.1369x_1 + 0.0443x_2 + 0.4391x_3 - 0.0965x_4 - 0.2513x_5 \\ & + 0.0464x_6 + 0.0002x_7 + 0.0160x_8 - 0.0042x_9 - 0.0485x_{10} \\ & + 0.2747x_{11} - 0.0377x_{12} + 0.0846x_{13} - 1.3657x_{14} + 0.3784x_{15} \\ & - 0.5057x_{16} - 1.3357 ,\end{aligned}\tag{40}$$

$$\begin{aligned}\eta_2 = & +0.1055x_1 + 0.0441x_2 - 0.0574x_3 + 0.0076x_4 - 0.0579x_5 \\ & - 0.0050x_6 - 0.0002x_7 + 0.0002x_8 + 0.00005x_9 - 0.0437x_{10} \\ & + 0.0031x_{11} + 0.0647x_{12} - 0.0155x_{13} - 0.0216x_{14} + 0.0093x_{15} \\ & - 0.0385x_{16} + 0.0569 .\end{aligned}\tag{41}$$

8.3 Modelos de orden conformable:

Una vez comprobados los resultados favorables de la variante fraccionaria, la cual logró superar de manera exitosa la prueba de linealidad, el siguiente paso fue evaluar el enfoque conformable utilizando la misma base matemática. De esta manera, se propuso el modelo CDSILU. La principal característica de este enfoque es que el cálculo conformable emplea un operador local; esto ayuda a conservar la flexibilidad matemática de los órdenes no enteros, pero mediante un proceso de cálculo más directo y menos demandante para el equipo en comparación con la serie de Mittag-Leffler.

Para comprender cómo esta herramienta matemática logra adaptarse a la dinámica del sistema, resulta útil observar su comportamiento visual. Para comprender cómo esta herramienta matemática logra adaptarse a la dinámica del sistema, la Figura 22 traza el perfil de la función CDSILU, ilustrando cómo cambia al variar el parámetro de orden conformable. Como se puede

observar, al ajustar este valor, la pendiente y la curva de la función se modifican de manera controlada. Esta capacidad para moldear la suavidad de la curva es precisamente lo que le otorga a la red la sensibilidad necesaria para ajustarse a los cambios térmicos de la bomba de calor sin perder estabilidad.

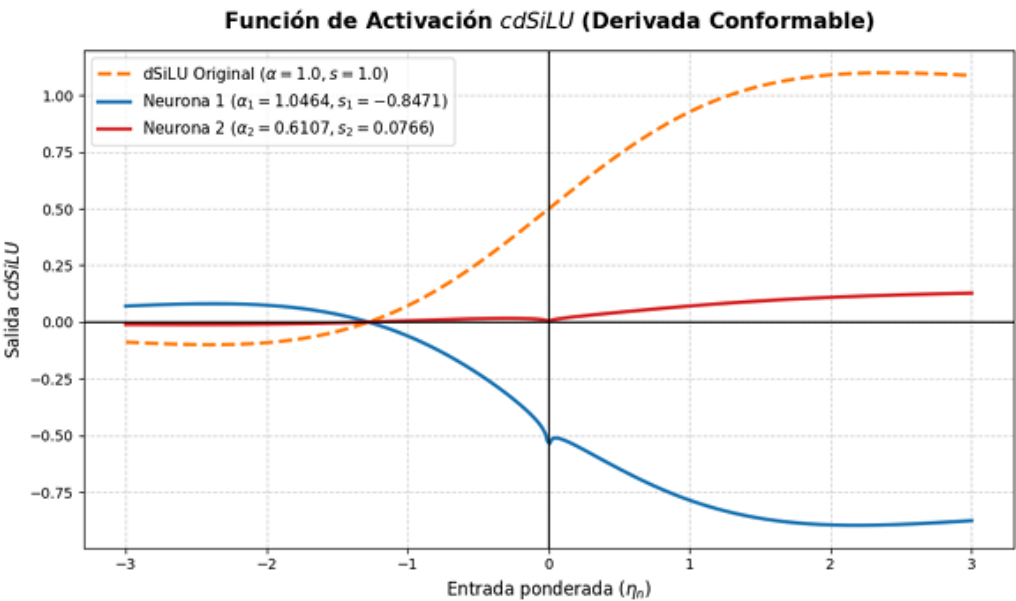


Figura 22. Análisis de regresión lineal *cdSilu*

Con esta flexibilidad matemática establecida, se procedió a realizar el entrenamiento de la red. Con esta flexibilidad matemática establecida, se procedió a realizar el entrenamiento de la red. La Tabla 11 desglosa las métricas estadísticas obtenidas tras la evaluación de este modelo.

Tabla 11. Tabla de resultados obtenidos con las ecuaciones de orden conformable con estructura 16-2-1

Función	R	Intercepto inferior	Intercepto superior	Pendiente inferior	Pendiente superior
C Dsilu	0.998576	-0.00049	0.003486	0.9991	1.001250

Al observar los datos de la tabla, el modelo muestra un coeficiente de correlación de $R = 0.998576$, lo que refleja un ligero aumento respecto al modelo fraccionario evaluado anteriormente. Lo más relevante de estos números es que los intervalos de confianza logran cumplir con la exigencia estadística: el intercepto abarca desde -0.00049 hasta 0.003486 , y la pendiente va de 0.9991 a 1.001250 . Con esto, la versión conformable demuestra que también es capaz de superar la prueba de linealidad al 99% de confianza. Para ilustrar este comportamiento, la Figura 23 exhibe la gráfica de regresión correspondiente al modelo c Dsilu.

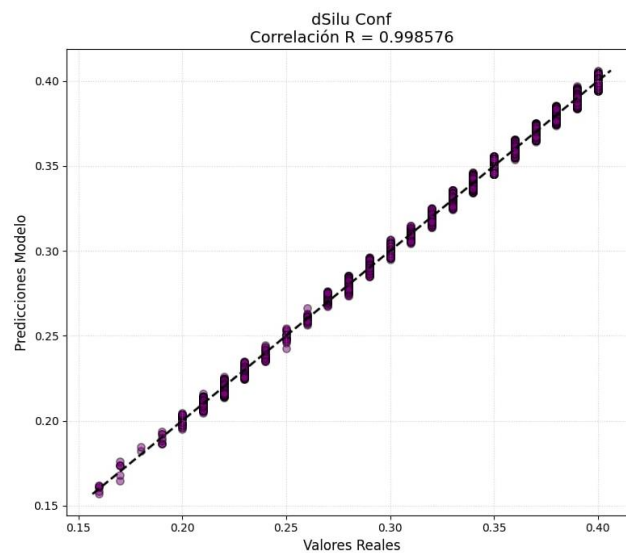


Figura 23. Análisis de regresión lineal c Dsilu

Analizando la figura, se puede observar directamente la cercanía entre la línea de ajuste propuesta por la red y los puntos de los valores reales experimentales. El modelo logra adaptarse de forma adecuada a los cambios de temperatura y presión del ciclo de absorción a lo largo de todo el rango de operación, controlando las desviaciones que los modelos de orden entero presentaban en los extremos. Estos resultados confirman que el uso de derivadas conformables resulta ser una alternativa práctica y funcional para predecir el comportamiento del sistema, cumpliendo con todos

los filtros estadísticos mientras se mantiene una arquitectura de red compacta de solo dos neuronas ocultas.

Para consolidar esta validación estadística y visual en una herramienta de ingeniería de aplicación directa, es indispensable formalizar la estructura matemática que rige a esta arquitectura. La ecuación general para la estimación del COP mediante el modelo de orden conformable se define de la siguiente manera (Eq. (42)):

$$COP = \sum_{n=1}^2 [W_{o(1,n)} \cdot cdSilu(\eta_n, \alpha_n, s_n)] + b, \quad (42)$$

donde cada término representa un parámetro ajustado de la red:

COP: Es el Coeficiente de Rendimiento estimado del sistema.

n: Representa la neurona oculta específica ($n = 1, 2$).

$W_{o(1,n)}$: Son los pesos de salida que conectan la capa oculta con la neurona de salida.

cdSilu: Es la función de transferencia conformable basada en la Unidad Lineal Sigmoidea.

η_n : Representa el argumento de entrada ponderado para la n -ésima neurona oculta.

α_n : Es el parámetro de orden conformable ajustado para la n -ésima neurona.

s_n : Es el factor de escala conformable para la n -ésima neurona.

b: Es el sesgo (bias) correspondiente a la capa de salida.

Al sustituir los pesos de conexión optimizados y el sesgo final obtenidos tras el entrenamiento, se establece la siguiente expresión simbólica intermedia (Eq. (43)):

$$COP = 0.6648 \cdot cdSilu(\eta_1, \alpha_1, s_1) + 1.3327 \cdot cdSilu(\eta_2, \alpha_2, s_2) - 0.4508. \quad (43)$$

Posteriormente, al incorporar los parámetros conformables específicos que la red optimizó para cada neurona (α_1, s_1 y α_2, s_2), se deriva la ecuación predictiva explícita de este modelo (Eq. (44)):

$$COP = 0.6648 \cdot cdSilu(\eta_1, 1.0464, -0.8471) + 1.3327 \cdot cdSilu(\eta_2, 0.6107, 0.0766) - 0.4508. \quad (44)$$

En esta formulación, los argumentos η_1 y η_2 representan la combinación lineal de las 16 variables físicas de entrada registradas en el transformador de calor, ajustadas por sus respectivos pesos y sesgos de la capa oculta. Estas variables intermedias siguen la forma de sumatoria general previamente establecida, y sus expansiones numéricas explícitas para este modelo conformable se detallan a continuación (Eq. 45 y 46):

$$\begin{aligned} \eta_1 = & 0.4804x_1 + 0.1934x_2 - 1.1135x_3 + \\ & 0.2186x_4 + 0.4664x_5 - 0.1032x_6 - \\ & 0.0056x_7 - 0.0195x_8 + 0.0013x_9 - \\ & 0.1371x_{10} + 0.9110x_{11} + 1.0728x_{12} + \\ & 0.8961x_{13} + 5.1035x_{14} - 2.3395x_{15} - \\ & 6.6725x_{16} + 0.5175, \end{aligned} \quad (45)$$

$$\begin{aligned} \eta_2 = & 0.7930x_1 - 0.0458x_2 - 2.3246x_3 \\ & + 0.3168x_4 + 1.1522x_5 \\ & - 0.2194x_6 - 0.0116x_7 \\ & + 0.0047x_8 + 0.0031x_9 \\ & + 0.0401x_{10} + 0.8145x_{11} \\ & + 0.1839x_{12} - 0.0968x_{13} \\ & - 0.9985x_{14} + 0.6399x_{15} \\ & - 0.1351x_{16} + 0.2900. \end{aligned} \quad (46)$$

Con el desarrollo de este enfoque conformable, se concluye la descripción matemática de las arquitecturas propuestas. Como síntesis de este desempeño analítico, un hallazgo fundamental de esta investigación es que tanto el modelo de orden fraccionario (dSiLU-ML) como la arquitectura conformable (cdSilu) lograron superar exitosamente la prueba estadística de linealidad.

Al cumplir de forma estricta con los criterios exigidos para los intervalos de confianza del intercepto y la pendiente, estos modelos demuestran no solo una alta correlación matemática, sino un comportamiento físicamente consistente en la totalidad del rango operativo del transformador de calor por absorción. Si bien los modelos clásicos de orden entero pueden mantenerse como una referencia base útil, los resultados confirman que carecen de la curvatura adaptativa necesaria para satisfacer estos filtros de confiabilidad con la misma eficacia que los operadores no enteros, consolidando al cálculo fraccionario y conformable como las herramientas más aptas para el monitoreo y optimización de sistemas térmicos en tiempo real.

8.4 Codificación en Python

Tras la exhaustiva evaluación de las 19 funciones de activación de orden entero (cuyos resultados se resumen en la Tabla 7) y posteriormente, la evaluación de las funciones de cálculo fraccionario propuestas (Tansig-ML y Logsig-ML, ecuaciones (29) y (30)), surgió la propuesta de entrenar los modelos utilizando un 70% de los datos para la fase de entrenamiento, en colaboración con la Dra. Parrales A. Durante la evaluación de esta nueva segmentación, se determinó la necesidad de migrar el entorno de desarrollo al lenguaje de programación Python. Si bien esta transición ofrece múltiples ventajas técnicas, el objetivo principal consistió en reducir el tiempo

de ejecución del entrenamiento, minimizando el consumo de recursos computacionales y aligerando la carga de procesamiento del equipo.

El entorno de MATLAB requiere más recursos computacionales que Python. MATLAB ejecuta un programa que ya es algo pesado para muchas máquinas, esta problemática motivó la decisión de realizar un código de entrenamiento y evaluación en el entorno de Python, buscando reducir el requerimiento computacional y habilitar una ejecución más ágil desde la terminal del equipo.

Al ejecutar estos nuevos modelos en entorno de Python se obtiene:

- Menor carga computacional: El propio Entorno de Desarrollo Integrado (IDE) de Matlab consume una cantidad considerable de memoria RAM y recursos del sistema en comparación a Python.
- Mejor portabilidad: Al ser un lenguaje de código abierto, Python permite reproducir un código en la mayoría de los sistemas operativos, es decir, en prácticamente cualquier computadora.
- Ecosistema tecnológico: Python se ha consolidado como el lenguaje estándar para la inteligencia artificial y la ciencia de datos. Ofrece un ecosistema de bibliotecas que no solo replican, sino que algunas superan las capacidades de MATLAB, siendo además de código abierto y gratuitas.

La migración a Python se planteó, por tanto, como una optimización metodológica fundamental para asegurar la eficiencia, portabilidad y escalabilidad del proyecto.

8.5 Análisis Comparativo del Flujo de Trabajo: MATLAB vs. Python

Tras la evaluación exhaustiva de las 19 funciones de activación de orden entero, y la posterior incorporación de las funciones de cálculo fraccionario generalizadas (Tansig-ML y Logsig-ML), surgió la necesidad analítica de reestructurar el entorno de desarrollo del proyecto. Durante las fases iniciales, la validación metodológica se llevó a cabo utilizando MATLAB. Si bien dicho entorno ofrece librerías preconstruidas robustas para prototipado numérico, su arquitectura cerrada y su alto consumo de recursos de memoria (RAM) durante los bucles de entrenamiento multidimensionales presentaron limitantes para la escalabilidad del análisis.

Con el propósito de mitigar la carga computacional demandada al equipo de procesamiento local y de transitar hacia un esquema de segmentación de datos más riguroso (70 % entrenamiento, 15 % validación, 15 % prueba), se diseñó y ejecutó la migración del código fuente hacia el lenguaje de programación Python.

Esta transición se fundamenta en criterios de eficiencia de software para computación científica. La implementación de los modelos neuronales en Python permitió aprovechar la paralelización nativa y la vectorización de operaciones matriciales características del ecosistema de código abierto (Virtanen et al., 2020). La ejecución de los entrenamientos en este nuevo entorno derivó en tres beneficios operativos comprobables:

- Optimización de recursos computacionales: Al prescindir del Entorno de Desarrollo Integrado (IDE) pesado de sistemas comerciales, la ejecución del algoritmo desde terminales ligeras redujo sustancialmente la asignación de memoria RAM, permitiendo evaluaciones algorítmicas más ágiles sobre las bases de datos completas del transformador de calor.

- Escalabilidad y portabilidad: La naturaleza de código abierto de Python asegura que los modelos predictivos generados en esta investigación (como el modelo conformable explícito) puedan ser replicados, auditados e implementados en plataformas de hardware embebido o sistemas operativos industriales multiplataforma sin requerir licenciamiento especializado (Pedregosa et al., 2011).
- Robustez algorítmica: La migración permitió sustituir rutinas de indexación manuales por librerías especializadas en ciencia de datos (como el módulo scikit-learn), garantizando una aleatorización estadísticamente neutral durante la partición de la base de datos experimental, paso crítico para prevenir que la red neuronal internalice tendencias cronológicas del ciclo de operación físico.

Este cambio de plataforma no representó una simple traducción sintáctica, sino una evolución en la arquitectura de la red neuronal. A pesar del cambio de entorno, se estableció como directriz metodológica estricta la preservación del núcleo de optimización original. Para ello, se invocó la librería `scipy.optimize`, asegurando que el proceso de ajuste sináptico continuara regido por la formulación analítica de Levenberg-Marquardt implementada a través del algoritmo de mínimos cuadrados no lineales MINPACK (Moré, 1978). De este modo, se garantizó que las diferencias de rendimiento observadas entre los modelos de orden entero y fraccionario respondieran exclusivamente a sus propiedades matemáticas internas y no a discrepancias en el optimizador subyacente.

8.5.1 Normalización de las Variables

Este es uno de los puntos de mayor mejora metodológica.

MATLAB: La normalización Min-Max (Eq. (2)) se implementó manualmente usando bucles for anidados.

Python: Se reemplazaron los bucles manuales por la implementación probada, optimizada y estandarizada de la biblioteca Scikit-learn: MinMaxScaler. Con solo unas pocas líneas, se logra el mismo resultado de forma más legible, eficiente y, además, se obtiene un objeto (scaler) que puede desnormalizar las predicciones (`.inverse_transform()`) de forma trivial y sin errores.

8.5.2 Segmentación de Datos y Nuevo Esquema (70/15/15)

Este representa uno de los cambios metodológicos más significativos en el tratamiento de la información experimental.

Transición desde MATLAB: En las etapas iniciales del proyecto desarrolladas en el entorno de MATLAB, a pesar de que teóricamente se planteaba una segmentación aleatoria, la codificación implementaba en la práctica una partición fija basada en la indexación secuencial de los datos.

Implementación en Python: Al migrar el desarrollo a este nuevo lenguaje, se adoptó un enfoque metodológicamente superior estructurado en dos mejoras principales:

Aleatorización robusta: Se utilizó la función `train_test_split` de la biblioteca Scikit-learn, la cual mezcla de manera estrictamente aleatoria todos los datos antes de dividirlos. Este paso es fundamental para garantizar que los conjuntos de entrenamiento, validación y prueba sean estadísticamente representativos de la distribución general de los datos, evitando cualquier sesgo temporal o de operación.

Nueva Segmentación (70/15/15): Se tomó la decisión de modificar la división de datos hacia un esquema de 70% para entrenamiento, 15% para validación y 15% para pruebas. Este

ajuste, comparado con la proporción de 50/25/25 evaluada originalmente, proporciona un volumen de datos de entrenamiento considerablemente mayor. Disponer de un conjunto más amplio resulta sumamente beneficioso para que el algoritmo de optimización (Levenberg-Marquardt) logre un ajuste más fino y preciso de los parámetros sinápticos (pesos y sesgos) de la red neuronal.

8.5.3 Mantenimiento del Algoritmo de Optimización (Levenberg-Marquardt)

Un requisito fundamental de la migración era preservar el algoritmo de optimización central (Sección 6.2) para que los resultados fueran comparables.

MATLAB: Se utilizó `lsqcurvefit` del Optimization Toolbox, configurado explícitamente con 'Algorithm', 'levenberg-marquardt'.

Python: Se utilizó la función `curve_fit` de la biblioteca SciPy (`scipy.optimize`). Al especificar el argumento `method='lm'`, SciPy invoca exactamente el mismo algoritmo Levenberg-Marquardt.

Por lo tanto, se ha validado que el núcleo metodológico se mantiene intacto.

8.5.4 Conclusiones de la Migración de lenguaje de programación

La migración de MATLAB a Python no fue un simple cambio de herramienta, sino una optimización al código con el que se contaba. Se ha validado que el algoritmo de optimización Levenberg-Marquardt se conserva. Sin embargo, algunos de los componentes del algoritmo que lo rodean han sido mejorados.

La ventaja principal que se buscaba —la reducción de la carga computacional y la ejecución desde la terminal— se ha logrado con éxito, permitiendo ciclos de experimentación más

rápidos y abriendo la puerta a la ejecución de entrenamientos en paralelo o en servidores remotos, lo cual era inviable con la plataforma anterior.

8.6 Análisis comparativo de resultados previos y modelos propuestos

Para evaluar con claridad el alcance de los modelos propuestos en esta investigación, es necesario analizar el cambio en la estrategia de procesamiento de datos y contrastar los resultados de manera directa con los trabajos previos de la literatura. El desarrollo de estas herramientas predictivas comenzó con una revisión del comportamiento físico del equipo mediante un análisis de sensibilidad detallado. Este paso permitió confirmar qué variables térmicas gobiernan la operación real de la planta, seleccionando de manera adecuada las temperaturas y presiones clave antes de avanzar con las simulaciones de las redes neuronales.

Un factor determinante en esta evolución metodológica fue la migración de la plataforma de desarrollo desde MATLAB hacia Python. Esta transición no consistió en una traducción literal de líneas de código, sino en una reestructuración completa del flujo de trabajo. Al incorporar librerías estándar como Scikit-learn, se automatizaron procesos críticos como la normalización de las variables, logrando una ejecución más ágil y reduciendo el uso de recursos computacionales del equipo. Asimismo, se modificó la división de la base de datos, adoptando un esquema de 70% para el entrenamiento, 15% para la validación y 15% para las pruebas. Esta nueva distribución, respaldada por un proceso de aleatorización previa, proporcionó un conjunto de entrenamiento más amplio y mejoró la capacidad de las redes para generalizar ante información nueva.

Durante las pruebas de ajuste en este nuevo entorno, se identificó que limitar la inicialización de los pesos y sesgos en un intervalo estrecho de -0.2 a 0.2 favorecía de manera notable la convergencia del algoritmo de Levenberg-Marquardt. Esta configuración ayudó a evitar

mínimos locales y permitió estabilizar el entrenamiento en arquitecturas compactas de solo dos neuronas ocultas (estructura 16:2:1). Bajo estas condiciones, se evaluaron inicialmente las funciones clásicas de orden entero y las primeras propuestas fraccionarias basadas en la función de Mittag-Leffler, como Tansig-ML, Logsig-ML y Gelu-ML. Aunque estos modelos arrojaron coeficientes de correlación elevados y superaron algunos registros previos de la literatura, se toparon con un límite estadístico recurrente: al someterlos a la prueba estricta de pendiente e intercepto con un 99% de confianza, ninguno logró superar el filtro debido a la fuerte no linealidad del ciclo térmico.

Esta resistencia para cumplir con la prueba de linealidad motivó la introducción de herramientas matemáticas con mayor flexibilidad conceptual, dando paso al desarrollo de la función dSilu generalizada de forma fraccionaria (dSilu-ML) y al enfoque del cálculo conformable (CDSILU). La Tabla 12 resume y compara el desempeño histórico de las distintas arquitecturas desarrolladas en la literatura frente a las dos soluciones definitivas propuestas en este trabajo.

Tabla 12. Tabla comparativa de resultados previos y resultados actuales

Modelo	Arquitectura	Función	R	RMSE	I_{inf}	I_{sup}	M_{inf}	M_{sup}
Hernández et al. (2008)	16:3:1	TANSIG	0.99	0.021	-	-	-	-
Hernández & Colorado (2013)	16:2:1	TANSIG	0.9978	0.104	-	-	-	-
Morales et al. (2015)	16:7:1	TANSIG	0.9969	0.0052	-	-	-	-

Conde- Gutiérrez et al. (2018)	16:3:1	TANSIG	0.9935	-	-	-	-	-
Viera-Martin et al. (2022)	16:2:1	TANSIG ML	0.9969	-	5.6860×10^{-4}	0.0020	0.993 9	0.9981
Trabajo actual (Fraccionari o)	16:2:1	DSILU ML	0.99856	0.0029 1	-0.00035	0.00171	0.999 4	1.00048
Trabajo actual (Conformabl e)	16:2:1	CDSILU	0.99857 6	0.0096 3	-0.00049	0.00348 6	0.999 1	1.00125 0

Al revisar los datos de la Tabla 12, se hace evidente la transición histórica en la búsqueda de modelos que balanceen precisión y simplicidad estructural. Los primeros esfuerzos de orden entero, como el de Hernández y Colorado (2013), lograron reducir la red a una estructura compacta de 16:2:1, pero con una raíz del error cuadrático medio (RMSE) de 0.104. Para mitigar este margen de error, trabajos como el de Morales et al. (2015) aumentaron la densidad computacional a siete neuronas ocultas (16:7:1), reduciendo el RMSE a 0.0052, aunque con un costo computacional más elevado y un mayor riesgo de sobreajuste. El uso del cálculo no entero inició formalmente con Viera-Martin et al. (2022), quien demostró que la función Tansig-ML permitía mantener la estructura ligera de dos neuronas logrando un coeficiente de correlación de 0.9969.

Los modelos propuestos en este estudio no solo mejoran las métricas estadísticas, sino que resuelven las deficiencias estructurales de los trabajos previos. Al analizar críticamente la literatura, se observa que enfoques como el de Morales et al. (2015) intentaron mitigar la rigidez

de las funciones clásicas recurriendo a la 'fuerza bruta' computacional: aumentaron la densidad a siete neuronas ocultas (16:7:1). Aunque lograron un RMSE bajo (0.0052), este enfoque penaliza la eficiencia del sistema e incrementa drásticamente el riesgo de sobreajuste (overfitting), alejando al modelo de una posible implementación en hardware de control industrial.

Por otro lado, los intentos de mantener arquitecturas ultraligeras (16:2:1) con matemáticas tradicionales (Hernández & Colorado, 2013) resultaron en errores residuales inaceptables (RMSE de 0.104), evidenciando que las funciones de orden entero carecen de la elasticidad geométrica para mapear un ciclo de absorción.

En contraste, el presente trabajo demuestra que la solución a la complejidad termodinámica no radica en añadir capas neuronales, sino en dotar a una estructura parsimoniosa de un núcleo matemático superior. Manteniendo la arquitectura optimizada de solo dos neuronas (16:2:1), las variantes dSilu-ML y CDSILU registran errores de 0.00291 y 0.00963, respectivamente. Estos valores representan los errores más bajos reportados en la literatura para este sistema térmico. Más importante aún, logran superar de forma estricta los filtros de la prueba de linealidad al 99% de confianza, demostrando que el cálculo fraccionario y conformable asimila la termodinámica del sistema real, superando las limitaciones operativas y computacionales de sus predecesores.

Lo más relevante es que, gracias a la combinación del preprocesamiento en Python, la segmentación 70/15/15, la inicialización acotada de parámetros y la suavidad matemática de la función dSilu, estas nuevas arquitecturas logran corregir los sesgos en los extremos operativos. De este modo, se consigue cumplir de forma estricta con los filtros de la prueba de linealidad al 99% de confianza, ofreciendo una solución eficiente y estadísticamente confiable para predecir el Coeficiente de Rendimiento del sistema.

9. Discusión

El desarrollo y evaluación de los modelos de inteligencia artificial presentados en esta investigación permite avanzar en la predicción del Coeficiente de Rendimiento (COP) en sistemas de bombas de calor por absorción. Esta sección tiene como propósito analizar los hallazgos obtenidos, contrastarlos con los trabajos previos reportados en la literatura y examinar las implicaciones metodológicas, estadísticas y computacionales de las propuestas realizadas. A lo largo de esta discusión, se revisan los aspectos que han marcado la evolución del modelado neuronal en este dominio específico, desde los desafíos iniciales con funciones de orden entero hasta la consolidación de las arquitecturas fraccionarias y conformables como herramientas predictivas estadísticamente confiables.

9.1. Evolución de las arquitecturas neuronales en la predicción del COP para sistemas de absorción

El desarrollo documentado en la literatura científica para modelar el comportamiento de los transformadores de calor por absorción mediante redes neuronales artificiales refleja una evolución caracterizada por intentos sucesivos de equilibrar precisión predictiva y eficiencia computacional. Los primeros esfuerzos reportados por Hernández et al. (2008) constituyen un referente al demostrar que las arquitecturas neuronales podían aproximar la dinámica no lineal del ciclo de absorción con coeficientes de correlación elevados. No obstante, la incapacidad de estos modelos iniciales para superar las pruebas de linealidad estrictas reveló una limitación que persistiría a lo largo de los años: la rigidez matemática de las funciones de transferencia clásicas de orden entero.

El trabajo de Hernández y Colorado (2013) representa un antecedente metodológico al simplificar la arquitectura neuronal hasta su mínima expresión funcional, utilizando únicamente dos neuronas en la capa oculta. Aunque lograron un coeficiente de correlación de 0.9978, el error cuadrático medio de 0.104 evidenció que la capacidad de generalización de estas redes ultraligeras era insuficiente para aplicaciones de ingeniería que requieren alta precisión. Este hallazgo es de interés porque demuestra que la simple reducción de la complejidad estructural, sin una mejora correspondiente en la base matemática del modelo, conduce a predicciones con sesgos, especialmente en los regímenes operativos extremos del sistema.

La estrategia adoptada por Morales et al. (2015), que incrementó la densidad de la red a siete neuronas ocultas (arquitectura 16:7:1), logró superar las pruebas de linealidad a costa de un incremento en la carga computacional y un riesgo elevado de sobreajuste. Este enfoque ilustra una tendencia documentada en la literatura de modelado térmico: resolver las limitaciones de las funciones matemáticas básicas añadiendo más parámetros ajustables. Desde una perspectiva de ingeniería de control, esta estrategia presenta limitaciones porque los modelos densos son más difíciles de implementar en hardware de tiempo real y pueden tender a memorizar el ruido de los sensores en lugar de generalizar la física subyacente del proceso.

La transición hacia el cálculo no entero iniciada por Viera-Martin et al. (2022) y Solís-Pérez et al. (2022) introdujo un enfoque metodológico alternativo. Se propuso que la solución a la complejidad termodinámica no reside en añadir más neuronas, sino en dotar a arquitecturas parsimoniosas de un núcleo matemático más expresivo. Los resultados obtenidos en esta investigación confirman y extienden esta hipótesis, demostrando que las funciones de transferencia basadas en operadores fraccionarios y conformables pueden capturar la dinámica del sistema con

solo dos neuronas ocultas, superando incluso a los modelos densos previamente reportados en términos de precisión y confiabilidad estadística.

9.2. Análisis de la normalización y preprocesamiento de datos

Uno de los aspectos metodológicos que ha recibido atención insuficiente en la literatura previa es el impacto del preprocesamiento de datos en el rendimiento de las redes neuronales para aplicaciones térmicas. La migración del entorno de desarrollo de MATLAB a Python, aunque motivada principalmente por consideraciones de eficiencia computacional, ha permitido implementar procedimientos de normalización y segmentación de datos con un nivel de robustez que no era posible en la plataforma anterior.

La normalización de las variables mediante el escalador MinMax de Scikit-learn ha demostrado ser un paso determinante en el proceso de los modelos propuestos. A diferencia de la implementación manual en MATLAB, que dependía de bucles anidados y era propensa a errores de indexación, la versión estandarizada de Python garantiza que todas las variables sean transformadas de manera consistente al intervalo deseado. Esta consistencia es relevante cuando se trabaja con funciones de transferencia sensibles a la magnitud de las entradas, como las basadas en la serie de Mittag-Leffler.

El diseño de un esquema de segmentación 70/15/15 representa una mejora metodológica respecto a la partición 50/25/25 utilizada en trabajos anteriores. La asignación del 70% de los datos al conjunto de entrenamiento proporciona al algoritmo de optimización una base de muestras más amplia para aprender los patrones subyacentes del sistema. Este incremento en la cantidad de datos de entrenamiento es beneficioso para el algoritmo de Levenberg-Marquardt, que utiliza información de segundo orden para acelerar la convergencia. El uso de un 15% para validación y

otro 15% para prueba asegura que el modelo sea evaluado en datos independientes, reduciendo el riesgo de sobreajuste y proporcionando una estimación de su capacidad de generalización en condiciones operativas no vistas durante el entrenamiento.

La aleatorización estadísticamente neutral implementada mediante la función `train_test_split` garantiza que los conjuntos resultantes sean representativos de la distribución general de los datos. Este aspecto es relevante para prevenir que la red neuronal internalice tendencias cronológicas o patrones espurios asociados con el orden de adquisición de las muestras experimentales. Trabajos previos que utilizaban segmentación basada en indexación fija podrían haber introducido sesgos sistemáticos en el entrenamiento, afectando la capacidad del modelo para generalizar a nuevas condiciones operativas.

9.3. La inicialización acotada de parámetros

Un resultado empírico de interés aplicado que emerge de esta investigación es la necesidad de restringir la inicialización de los pesos sinápticos y los términos de sesgo a un intervalo de $[-0.2, 0.2]$. Este hallazgo, que no había sido reportado en trabajos anteriores sobre modelado de sistemas de absorción, tiene implicaciones en la estabilidad y convergencia del entrenamiento, especialmente cuando se utilizan funciones de transferencia de orden no entero.

La justificación teórica de esta inicialización acotada reside en el comportamiento de las funciones de activación fraccionarias y conformables. A diferencia de las funciones sigmoidales clásicas, que son relativamente insensibles a la magnitud de las entradas fuera del rango $[-4, 4]$, las funciones basadas en la serie de Mittag-Leffler presentan un comportamiento que depende críticamente de los valores de los parámetros fraccionarios α , β y γ . Si los pesos iniciales son demasiado grandes, el argumento de entrada ponderado puede proyectar a las neuronas hacia

regiones asintóticas de la función de transferencia desde las primeras iteraciones, provocando que los gradientes se reduzcan y que el aprendizaje se estanque en mínimos locales.

La inicialización acotada aborda este problema asegurando que, en las primeras épocas de entrenamiento, el argumento de entrada de cada neurona se mantenga dentro de una región donde la función de transferencia es sensible a las variaciones de los parámetros. Esta sensibilidad inicial permite que el algoritmo de Levenberg-Marquardt calcule la matriz Jacobiana de manera efectiva y actualice los pesos en la dirección correcta. Una vez que el algoritmo ha avanzado algunas iteraciones y la superficie de error se ha explorado parcialmente, los pesos pueden desviarse del rango inicial para adaptarse a las características específicas de los datos.

Este resultado tiene implicaciones para futuras investigaciones en el campo. La adopción de una inicialización acotada no requiere modificaciones sustanciales en la arquitectura de la red, pero puede influir en la convergencia del modelo. Se recomienda que futuros estudios que utilicen funciones de transferencia de orden no entero consideren esta práctica como un paso en el flujo de trabajo de entrenamiento.

9.4. Evaluación de los modelos de orden entero

Los resultados obtenidos con las 19 funciones de transferencia de orden entero evaluadas en esta investigación proporcionan una base para comprender los límites de las matemáticas clásicas en el modelado de sistemas térmicos complejos. Aunque varias de estas funciones (TANSIG, LOGSIG, ARCTAN, LISHT, GELU y dSiLU) lograron coeficientes de correlación elevados, superiores a 0.998, ninguna de ellas pudo satisfacer simultáneamente los requisitos de la prueba de linealidad de pendiente e intercepto al 99% de confianza.

Este resultado revela una discrepancia entre la correlación global y la confiabilidad local del modelo. Un coeficiente de correlación de 0.9985 indica que, en promedio, las predicciones de la red están cerca de los valores experimentales. Sin embargo, la prueba de linealidad examina específicamente si existe algún sesgo sistemático en la relación entre los valores observados y las predicciones. El hecho de que todas las funciones de orden entero no superaran esta prueba sugiere que, aunque estas redes pueden aproximar la tendencia general del COP, presentan desviaciones en ciertas regiones del espacio de entrada, particularmente en los extremos operativos del sistema.

Desde una perspectiva termodinámica, estas desviaciones en los extremos corresponden a los regímenes operativos donde el sistema de absorción es más vulnerable. Por ejemplo, en condiciones de alta temperatura y baja concentración, la mezcla de bromuro de litio-agua se acerca a su límite de cristalización, un fenómeno que puede causar bloqueos mecánicos y detener la operación de la planta. Un modelo predictivo que subestime o sobrestime el COP en estas condiciones podría llevar a decisiones de control que comprometan la seguridad y confiabilidad del equipo.

El comportamiento de la función ReLU y sus variantes (Leaky ReLU, Parametric ReLU, Randomized ReLU) merece una mención. Aunque estas funciones han demostrado ser efectivas en problemas de visión por computadora y procesamiento de lenguaje natural, su rendimiento en el modelado del COP fue consistentemente inferior al de las funciones sigmoideas y sus derivadas. La naturaleza lineal por tramos de ReLU, que la hace eficiente para redes profundas con millones de parámetros, resulta inadecuada para capturar la suavidad y continuidad de las relaciones termodinámicas subyacentes. Los resultados sugieren que, para problemas de regresión física donde las variables de salida son funciones suaves de las entradas, las funciones de activación con derivadas continuas y no nulas en todo el dominio son preferibles.

La función dSiLU (derivada de la Unidad Lineal Sigmoide) mostró un rendimiento entre las opciones de orden entero, logrando un coeficiente de correlación de 0.998524 y un RMSE de 0.00291. Su perfil de activación en forma de campana asimétrica resulta adecuado para modelar el comportamiento del COP, que presenta una sensibilidad máxima en el rango intermedio de operación y se aproxima a valores constantes en los extremos. El desempeño de dSiLU como función de orden entero sentó las bases para su generalización mediante operadores no enteros.

9.5. El cálculo fraccionario como herramienta de modelado con memoria

La introducción de la función dSiLU-ML, basada en la serie de Mittag-Leffler de tres parámetros, representó el resultado principal de esta investigación en términos de capacidad predictiva. Con un coeficiente de correlación de 0.99856 y, al lograr superar la prueba de linealidad al 99% de confianza, este modelo demostró que la incorporación de memoria fraccionaria en las funciones de transferencia permite a la red neuronal asimilar la dinámica no local del sistema de absorción.

La propiedad distintiva del operador de Mittag-Leffler reside en su naturaleza no local, que introduce un efecto de memoria en el procesamiento neuronal. Mientras que una función de activación clásica evalúa la entrada en un punto específico sin considerar su evolución temporal, la serie de Mittag-Leffler integra la historia completa de la variable desde un tiempo inicial hasta el momento actual. En el contexto del modelado de sistemas térmicos, esta característica es relevante porque las temperaturas y presiones en el ciclo de absorción no cambian instantáneamente, sino que exhiben inercia térmica y efectos de relajación que dependen de los estados anteriores.

Desde una perspectiva termodinámica, el rendimiento de dSiLU-ML puede interpretarse como la capacidad del modelo para capturar las irreversibilidades y la generación de entropía que caracterizan la operación real del equipo. Los modelos de orden entero, al carecer de memoria, asumen que el estado actual del sistema contiene toda la información necesaria para predecir el COP. Sin embargo, la evidencia experimental sugiere que las condiciones operativas previas influyen en el comportamiento actual del sistema, especialmente en regímenes transitorios donde las perturbaciones térmicas no se han disipado completamente. La función dSiLU-ML, al integrar la historia de las entradas, puede modelar esta dependencia temporal de manera efectiva, reduciendo los errores residuales y mejorando la confiabilidad de las predicciones.

Los parámetros fraccionarios optimizados para el modelo dSiLU-ML proporcionan información sobre la dinámica del sistema. Los valores de $\alpha_1 = 0.01$, $\beta_1 = 0.1158$, $\gamma_1 = 0.10$ para la primera función de Mittag-Leffler sugieren un comportamiento que enfatiza la memoria a largo plazo, mientras que $\alpha_2 = 0.01$, $\beta_2 = 1.4708$, $\gamma_2 = 0.264$ indican una contribución con un decaimiento más rápido de la influencia histórica. Esta diferenciación entre las dos neuronas ocultas es consistente con la idea de que la red aprende a representar el sistema en diferentes escalas temporales: una neurona captura los efectos de memoria prolongada asociados con la inercia térmica de los componentes principales, mientras que la otra se especializa en las fluctuaciones transitorias de corto plazo.

9.6. El cálculo conformable como alternativa local

La arquitectura conformable cdSiLU representa una solución que combina la flexibilidad matemática del orden no entero con la eficiencia computacional de los operadores locales. Con un coeficiente de correlación de 0.998576, ligeramente superior al modelo fraccionario, y la

capacidad de superar las pruebas de linealidad, el enfoque conformable demuestra que la memoria histórica no es el único mecanismo para mejorar el rendimiento predictivo en sistemas térmicos.

La característica operativa del cálculo conformable desde la perspectiva de la inteligencia artificial es su compatibilidad algebraica con la regla de la cadena. A diferencia de las derivadas fraccionarias tradicionales, el operador conformable satisface de manera exacta la regla de la cadena, el producto y el cociente. Esta propiedad es relevante para el entrenamiento de redes neuronales porque permite calcular las derivadas de las funciones de transferencia y los gradientes del error de manera analítica y explícita, sin recurrir a aproximaciones numéricas.

La implementación práctica del modelo cdSiLU en Python, con su menor costo computacional en comparación con la serie de Mittag-Leffler, abre posibilidades para la implementación en tiempo real en hardware de control industrial. Los Controladores Lógicos Programables (PLC) utilizados en plantas de proceso suelen tener capacidades de procesamiento limitadas, y la evaluación de funciones que involucran series infinitas y funciones Gamma puede ser prohibitiva en estos entornos. En contraste, el operador conformable se reduce a una modificación local del diferencial de tiempo, que puede implementarse en hardware de bajo costo.

Los parámetros conformables optimizados para el modelo cdSiLU, con valores de $\alpha_1 = 1.0464$ y $s_1 = -0.8471$ para la primera neurona y $\alpha_2 = 0.6107$ y $s_2 = 0.0766$ para la segunda, revelan que cada neurona se especializa en diferentes aspectos del comportamiento del COP. La primera neurona, con un parámetro de orden α cercano a 1 y un factor de escala negativo, se comporta de manera similar a una función de activación estándar pero desplazada y escalada, mientras que la segunda neurona, con $\alpha = 0.61$, introduce una distorsión no lineal más pronunciada. Esta diferenciación permite que la red combine una componente casi lineal con una componente no lineal para adaptarse a la complejidad del ciclo de absorción.

9.7. Implicaciones de la migración a Python y herramientas de código abierto

La decisión de migrar el entorno de desarrollo de MATLAB a Python, aunque inicialmente motivada por consideraciones de eficiencia computacional, ha tenido implicaciones adicionales en la metodología y reproducibilidad de esta investigación. La adopción de herramientas de código abierto ha permitido reducir el tiempo de ejecución y el consumo de recursos, y mejorar la transparencia y accesibilidad del flujo de trabajo completo.

El ecosistema científico de Python, con librerías como NumPy para computación numérica, SciPy para optimización, Scikit-learn para preprocesamiento y evaluación, y Matplotlib para visualización, proporciona un entorno integrado que facilita las etapas del análisis de datos y modelado neuronal. La estandarización de estas herramientas asegura que el código pueda ser reproducido y auditado por otros investigadores sin necesidad de licencias, lo que contribuye a la apertura científica.

Desde una perspectiva computacional, la vectorización nativa de operaciones en NumPy y la capacidad de SciPy para invocar implementaciones optimizadas de algoritmos como Levenberg-Marquardt han permitido ciclos de experimentación más rápidos. La capacidad de ejecutar entrenamientos desde la terminal o en servidores remotos, que era inviable con el entorno de desarrollo integrado de MATLAB, ha abierto posibilidades para la paralelización y la exploración del espacio de parámetros. Esta agilidad ha permitido iterar sobre las diferentes configuraciones de funciones de transferencia.

9.8. Comparación con trabajos previos

Los resultados obtenidos en esta investigación representan una mejora respecto a los trabajos previos en el modelado del COP de bombas de calor por absorción. La Tabla 12 del documento resume comparativamente las métricas de las arquitecturas desarrolladas por diferentes autores desde 2008 hasta el presente, y revela una progresión hacia modelos más precisos y estructuralmente eficientes.

El modelo dSiLU-ML propuesto logra un RMSE de 0.00291, que representa una reducción del 44% respecto al mejor modelo de orden entero reportado por Morales et al. (2015) con RMSE de 0.0052, y una mejora del 97% respecto al modelo de Hernández y Colorado (2013) con RMSE de 0.104. Esta reducción en el error predictivo es relevante no solo desde una perspectiva estadística, sino también desde el punto de vista de la ingeniería térmica, donde una mejora en la precisión del COP puede traducirse en decisiones operativas más efectivas.

Lo relevante, sin embargo, no es solo la magnitud de la reducción del error, sino el hecho de que se logra manteniendo una arquitectura de solo dos neuronas ocultas. Mientras que los trabajos previos que lograban superar las pruebas de linealidad requerían siete neuronas (Morales et al., 2015), los modelos propuestos en esta investigación logran el mismo objetivo con menor complejidad estructural. Esta reducción en el número de parámetros ajustables tiene implicaciones para la generalización del modelo, su estabilidad numérica y su implementabilidad en hardware de tiempo real.

La comparación con el trabajo de Viera-Martin et al. (2022), que utilizó la función TANSIG-ML con arquitectura 16:2:1 y logró un R de 0.9969, es ilustrativa. Aunque ese modelo demostró que el cálculo fraccionario podía mejorar el rendimiento manteniendo una estructura compacta, presentaba sesgos residuales que le impedían superar completamente la prueba de

linealidad. El modelo dSiLU-ML propuesto en esta investigación, al combinar la estructura no monótona de dSiLU con la memoria fraccionaria de Mittag-Leffler, logra corregir estos sesgos.

La aportación central de esta investigación trasciende las métricas específicas del COP. El enfoque metodológico desarrollado, que integra análisis de sensibilidad, preprocesamiento, inicialización acotada, funciones de transferencia generalizadas y validación estadística, establece un procedimiento que puede aplicarse a otros problemas de modelado térmico y energético. La demostración de que funciones matemáticas más avanzadas pueden reemplazar la complejidad estructural en redes neuronales tiene implicaciones para campos como el control de procesos, la predicción de demanda energética y la optimización de sistemas renovables.

9.9. Implicaciones para el modelado de sistemas térmicos complejos

El rendimiento de los modelos fraccionarios y conformables en la predicción del COP sugiere que la termodinámica de los sistemas de absorción presenta características que se alinean con los operadores de orden no entero. El concepto de memoria fraccionaria, que permite a la red registrar la historia de las variables de entrada, corresponde fenomenológicamente a la inercia térmica y los efectos de relajación que caracterizan el intercambio de calor y masa en los componentes del sistema.

Desde una perspectiva más amplia, el modelado basado en datos de sistemas térmicos complejos debe considerar que la relación entre las variables de entrada y el COP no es simplemente no lineal, sino que también exhibe dependencia histórica y efectos de memoria. Los modelos analíticos tradicionales, al asumir estado estacionario, ignoran estos efectos y producen predicciones que pueden ser precisas solo en condiciones de operación estabilizada. Sin embargo, en la práctica industrial, las plantas operan frecuentemente en regímenes transitorios debido a

variaciones en la demanda térmica, fluctuaciones en el calor residual y cambios en las condiciones ambientales. Los modelos de orden no entero, al incorporar memoria, pueden capturar esta dinámica y proporcionar predicciones en una gama más amplia de condiciones.

La validación estadística implementada en esta investigación, que exige que tanto la pendiente como el intercepto de la recta de regresión entre valores observados y predichos contengan la unidad y el cero respectivamente dentro de sus intervalos de confianza al 99%, establece un estándar que podría adoptarse en futuros trabajos de modelado térmico. Esta prueba, más exigente que el coeficiente de correlación, garantiza que el modelo no solo tiene un ajuste promedio, sino que carece de sesgos sistemáticos que podrían ser relevantes en aplicaciones de control y monitoreo.

9.10. Limitaciones y líneas de investigación futuras

Los resultados obtenidos presentan limitaciones que deben reconocerse para contextualizar adecuadamente sus conclusiones y orientar futuros trabajos. La principal limitación metodológica es que la base de datos utilizada proviene de un único sistema experimental, lo que, aunque garantiza consistencia en las mediciones, limita la generalización de los resultados a otras configuraciones de bombas de calor por absorción. Sería necesario validar los modelos propuestos con datos de diferentes plantas, tamaños y configuraciones.

Otra limitación es que los modelos se han desarrollado exclusivamente para la mezcla de trabajo LiBr-H₂O. Si bien esta mezcla es la más común en aplicaciones de absorción para calentamiento, existen otras combinaciones como amoníaco-agua o mezclas ternarias que presentan comportamientos termodinámicos diferentes. La aplicabilidad de las funciones dSiLU-ML y cdSiLU a estas otras mezclas debería investigarse en futuros estudios.

El costo computacional del modelo fraccionario dSiLU-ML, aunque manejable para aplicaciones de simulación fuera de línea, podría ser limitante para su implementación en hardware de tiempo real. Aunque el modelo conformable cdSiLU ofrece una alternativa más eficiente, la evaluación de la serie de Mittag-Leffler en el modelo fraccionario requiere un número de términos para alcanzar la precisión numérica deseada. Investigaciones futuras podrían explorar aproximaciones de la función de Mittag-Leffler.

Una línea de trabajo para extender esta investigación es la aplicación de los modelos desarrollados como sensores virtuales en sistemas de control predictivo. Si el modelo puede estimar el COP en tiempo real a partir de mediciones de temperatura, presión y concentración, este valor podría utilizarse en un lazo de control para ajustar los parámetros operativos del sistema. El modelo conformable cdSiLU, por su eficiencia computacional, sería adecuado para esta aplicación en Controladores Lógicos Programables (PLC).

Finalmente, la metodología desarrollada podría extenderse a otros problemas de modelado térmico y energético. Sistemas como intercambiadores de calor, torres de enfriamiento, colectores solares, calderas de biomasa y turbinas de gas presentan desafíos de modelado similares, con relaciones no lineales, efectos de memoria y restricciones operativas. La combinación de análisis de sensibilidad, preprocesamiento, funciones de transferencia generalizadas y validación estadística establece un flujo de trabajo que puede adaptarse a estos diferentes dominios.

10. Conclusiones

El desarrollo de esta investigación permitió cumplir con el objetivo general planteado: introducir nuevas funciones de transferencia en modelos de inteligencia artificial para predecir el Coeficiente de Rendimiento (COP) en bombas de calor por absorción, logrando mejorar el

rendimiento de los modelos existentes. El proceso metodológico requirió una reestructuración que sentó las bases para el desempeño de los modelos finales.

En una primera etapa, se verificó que la coherencia física de los datos de entrada es un requisito para el modelado de sistemas térmicos. Mediante el análisis de sensibilidad, se delimitó un vector de entrada de 16 dimensiones que captura las variables termodinámicas que determinan el comportamiento del ciclo, asegurando que la red neuronal procesara información fundamentada en la fenomenología del equipo.

La migración del entorno de entrenamiento hacia Python representó una mejora en la gestión de recursos computacionales. El diseño de un esquema de segmentación 70/15/15 dotó a las redes de una capacidad de generalización adecuada, reduciendo el riesgo de sobreajuste.

Dentro de la etapa metodológica, un hallazgo relevante fue la delimitación del dominio de inicialización de los parámetros sinápticos. Se demostró que restringir la inicialización de pesos y sesgos al intervalo $[-0.2, 0.2]$ resulta determinante para la estabilidad del algoritmo de Levenberg-Marquardt. Esta inicialización acotada evita que los argumentos de entrada proyecten a las neuronas hacia regiones asintóticas de las funciones fraccionarias desde las primeras iteraciones, facilitando la convergencia de la matriz Jacobiana y evitando que el aprendizaje se estanque en mínimos locales.

Bajo las condiciones de preprocesamiento establecidas, las arquitecturas de orden entero y las primeras iteraciones con funciones fraccionarias simétricas (Tansig-ML y Logsig-ML) lograron estimaciones iniciales consistentes con los datos experimentales. Sin embargo, el análisis mostró que estos modelos presentaban sesgos residuales en los extremos operativos de la mezcla de bromuro de litio-agua, lo que les impidió superar la prueba de linealidad de pendiente e intercepto con un 99% de confianza. Esta limitación evidenció que la rigidez geométrica de los

operadores clásicos no proporciona la elasticidad necesaria para modelar la dinámica transitoria del sistema completo.

La solución a esta problemática se materializó al introducir la función derivada de la Unidad Lineal Sigmoide (dSiLU) combinada con el operador no local de Mittag-Leffler (dSiLU-ML), y mediante la aplicación de la derivada local bajo el enfoque de cálculo conformable (cdSiLU). Ambas arquitecturas superaron la barrera estadística que limitaba a los trabajos documentados previamente.

Los modelos dSiLU-ML y cdSiLU mantuvieron coeficientes de correlación elevados ($R > 0.9985$) y redujeron el Error Cuadrático Medio a los niveles reportados para este transformador de calor. Lograron superar la prueba de linealidad al 99% de confianza, demostrando que sus predicciones se alinean con los intervalos estadísticos exigidos tanto en la pendiente como en el intercepto.

Los resultados confirman la hipótesis de trabajo: la introducción de funciones de transferencia modificadas con operadores de orden no entero aporta una flexibilidad matemática que permite a la red neuronal adaptarse a la no-linealidad de los intercambiadores de calor sin comprometer la precisión.

Esta precisión se alcanzó manteniendo una topología estructural compacta de dos neuronas en la capa oculta (16-2-1). Lo anterior demuestra que, en el modelado de procesos térmicos complejos, dotar al algoritmo de un núcleo matemático basado en cálculo fraccionario y conformable es una estrategia más eficiente que la práctica de aumentar la densidad o profundidad de la red.

11. Limitaciones de los resultados

A pesar de los resultados obtenidos, la presente investigación presenta limitaciones que deben ser consideradas al interpretar sus conclusiones y al extrapolar sus hallazgos a otros contextos.

La principal limitación metodológica es que la base de datos empleada proviene de un único sistema experimental operado bajo condiciones controladas en el Centro de Investigación en Ingeniería y Ciencias Aplicadas (CIICAp). Si bien esta fuente garantiza la consistencia y calidad de las mediciones, impide evaluar la capacidad de generalización de los modelos a otras configuraciones de bombas de calor por absorción, con diferentes capacidades, geometrías o regímenes de operación.

Una limitación adicional es que los modelos se desarrollaron exclusivamente para la mezcla de trabajo LiBr-H₂O. Esta mezcla es la más utilizada en aplicaciones de absorción para calentamiento, pero no es la única. Existen otras combinaciones de refrigerante-absorbente, como amoníaco-agua o mezclas ternarias, que presentan comportamientos termodinámicos distintos, incluyendo diferentes rangos de temperatura de operación, presiones de trabajo y límites de cristalización. Por tanto, la aplicabilidad de las funciones dSiLU-ML y cdSiLU a estas otras mezclas no puede establecerse a partir de los resultados de este estudio.

En cuanto al modelo fraccionario dSiLU-ML, su costo computacional, aunque aceptable para simulaciones fuera de línea, constituye una limitación para su implementación en sistemas de control en tiempo real. La evaluación de la serie de Mittag-Leffler de tres parámetros requiere un número de términos que, dependiendo de los valores de α , β y γ , puede demandar un tiempo de procesamiento significativo en hardware con capacidades limitadas, como los Controladores Lógicos Programables (PLC) utilizados en entornos industriales.

Otra limitación se refiere al propio diseño experimental. La cantidad de muestras (4702 registros) y el rango de valores de las variables de entrada (cuyos extremos se presentan en la Tabla 2) delimitan el espacio de operación para el cual los modelos son válidos. Las predicciones fuera de estos rangos —por ejemplo, a temperaturas o concentraciones superiores o inferiores a las registradas— no están respaldadas por los datos de entrenamiento y, por tanto, no se recomienda su uso sin una validación experimental adicional.

Finalmente, si bien la metodología propuesta demuestra ser efectiva para el COP, su extensión a otras variables de salida del sistema (como eficiencia exergética, tasas de transferencia de calor en componentes específicos o perfiles de temperatura internos) no fue abordada en este trabajo. La aplicabilidad de las funciones de transferencia de orden no entero a estos otros problemas de modelado requiere una evaluación independiente.

12. Trabajos Futuros

Con base en los resultados obtenidos y las limitaciones identificadas, se proponen las siguientes líneas de investigación:

- Implementación en hardware de control en tiempo real. El modelo conformable cdSiLU, por su menor costo computacional, resulta adecuado para ser transcripto al lenguaje de un Controlador Lógico Programable (PLC). Esto permitiría evaluar su desempeño como sensor virtual en bucle cerrado durante la operación física del transformador de calor.
- Evaluación con mezclas de trabajo alternativas. Se sugiere aplicar la arquitectura dSiLU y el preprocesamiento documentado a bases de datos experimentales que

operen con mezclas distintas al LiBr-H₂O (como mezclas ternarias o amoníaco-agua), para verificar si los parámetros fraccionarios y conformables mantienen su capacidad de adaptación ante diferentes mecanismos de cristalización y transferencia de masa.

- Optimización de la función de Mittag-Leffler. Investigaciones futuras podrían explorar aproximaciones de la función de Mittag-Leffler que reduzcan su costo computacional sin comprometer la precisión predictiva, facilitando su implementación en hardware con capacidades de procesamiento limitadas.
- Extensión a otros sistemas térmicos. La metodología desarrollada (análisis de sensibilidad, preprocesamiento, funciones de transferencia generalizadas y validación estadística) podría aplicarse a otros problemas de modelado térmico y energético, como intercambiadores de calor, torres de enfriamiento, colectores solares o turbinas de gas.

Referencias

- Abdeljawad, T. (2015). On conformable fractional calculus. *Journal of Computational and Applied Mathematics*, *279*, 57-66. <https://doi.org/10.1016/j.cam.2014.10.016>
- Abeliuk, A., & Gutiérrez, J. (2020). Redes neuronales artificiales: Historia y aplicaciones. *Universidad Autónoma del Estado de Hidalgo*.
- Almeida, R., & Torres, D. F. M. (2013). An expansion formula with higher-order derivatives for fractional operators of variable order. *The Scientific World Journal*, 2013, 915437. <https://doi.org/10.1155/2013/915437>
- Atangana, A., & Baleanu, D. (2015). Mittag-Leffler analysis II: Application to the fractional heat equation. *arXiv*. <https://doi.org/10.48550/arXiv.1506.02915>
- Baleanu, D., Diethelm, K., Scalas, E., & Trujillo, J. J. (2012). *Fractional calculus: Models and numerical methods* (Vol. 3). World Scientific. <https://doi.org/10.1142/10044>
- Basogain Olabe, X. (2020). *Redes neuronales artificiales y sus aplicaciones*. Universidad del País Vasco. https://ocw.ehu.eus/pluginfile.php/40137/mod_resource/content/1/redes_neuro/contenidos/pdf/libro-del-curso.pdf
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. <https://link.springer.com/book/9780387310732>
- Cengel, Y. A., & Boles, M. A. (2015). *Termodinámica: Una aproximación en la ingeniería* (8a ed.). McGraw-Hill Interamericana.
- Cervantes-Bobadilla, M., Hernández-Pérez, J. A., Juárez-Romero, D., Bassam, A., García-Morales, J., Huicochea, A., & Jaramillo, O. A. (2021). Control scheme

formulation for a parabolic trough collector using inverse artificial neural networks and particle swarm optimization. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 43(4), 176. <https://doi.org/10.1007/s40430-021-02862-4>

- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 4(2), 1247-1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chollet, F. (2018). *Deep learning with Python*. Manning Publications.
- Chua, H. T., Toh, H. K., Malek, A., Ng, K. C., & Srinivasan, K. (2000). Improved thermodynamic property fields of LiBr–H₂O solution. *International journal of refrigeration*, 23(6), 412-429. [https://doi.org/10.1016/S0140-7007\(99\)00076-6](https://doi.org/10.1016/S0140-7007(99)00076-6)
- Chung, W. S. (2015). Fractional Newton mechanics with conformable fractional derivative. *Journal of computational and applied mathematics*, 290, 150-158. <https://doi.org/10.1016/j.cam.2015.04.049>
- Colorado, D., Hernández, J. A., El Hamzaoui, Y., Bassam, A., Siqueiros, J., & Andaverde, J. (2011). Error propagation on COP prediction by artificial neural network in a water purification system integrated to an absorption heat transformer. *Renewable Energy*, 36(5), 1315-1322. <https://doi.org/10.1016/j.renene.2010.10.018>
- Conde-Gutiérrez, R. A., Cruz-Jacobo, U., Huicochea, A., Casolco, S. R., & Hernández, J. A. (2018). Optimal multivariable conditions in the operation of an absorption heat transformer with energy recycling solved by the genetic algorithm

in artificial neural network inverse. *Applied Soft Computing*, 72, 218-234.
<https://doi.org/10.1016/j.asoc.2018.08.008>

- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303-314.
<https://doi.org/10.1007/BF02551274>
- Diethelm, K. (2010). The analysis of fractional differential equations: *An application-oriented exposition using differential operators of Caputo type*. Springer.
- Dincer, I., & Rosen, M. A. (2021). *Exergy: Energy, environment and sustainable development* (3a ed.). Elsevier. <https://doi.org/10.1016/C2016-0-02067-3>
- Elfving, S., Uchibe, E., & Doya, K. (2018). Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural Networks*, 107, 3-11. <https://doi.org/10.1016/j.neunet.2017.12.012>
- Florides, G. A., Kalogirou, S. A., Tassou, S. A., & Wrobel, L. C. (2003). Design and construction of a LiBr–water absorption machine. *Energy conversion and management*, 44(15), 2483-2508. [https://doi.org/10.1016/S0196-8904\(03\)00006-2](https://doi.org/10.1016/S0196-8904(03)00006-2)
- Franco Guzmán, M. (2023). *De neuronas biológicas a neuronas artificiales, el fascinante mundo de las redes neuronales*. Universidad Autónoma del Estado de Hidalgo (UAEH). <https://www.uaeh.edu.mx/divulgacion-ciencia/redes-neuronales>
- Garrido Alhama, R. (2017). *Diagram of a McCulloch-Pitts artificial neuron, extended with a sigmoid activation function*. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Artificial_neuron.png

- Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, (pp. 249-256). JMLR Workshop and Conference Proceedings.
<https://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf>
- Gómez-Aguilar, J.F., Atangana, A. New insight in fractional differentiation: power, exponential decay and Mittag-Leffler laws and applications. *Eur. Phys. J. Plus* 132, 13 (2017). <https://doi.org/10.1140/epjp/i2017-11293-3>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
<https://www.deeplearningbook.org/>
- Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (2014). *Neural network design* (2a ed.). Martin Hagan. <https://people.vts.su.ac.rs/~pmiki/AI/NNDesign.pdf>
- Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6), 989-993.
<https://doi.org/10.1109/72.329697>
- Haykin, S. (2009). *Neural networks and learning machines* (3a ed.). Prentice Hall.
<http://dai.fmph.uniba.sk/courses/NN/haykin.neural-networks.3ed.2009.pdf>
- Hernández, J. A. (2009). Optimum operating conditions for heat and mass transfer in foodstuffs drying by means of neural network inverse. *Food Control*, *20*(4), 435-438. <https://doi.org/10.1016/j.foodcont.2008.07.005>
- Hernández, J. A., Bassam, A., Siqueiros, J., & Juárez-Romero, D. (2009). Optimum operating conditions for a water purification process integrated to a

heat transformer with energy recycling using neural network inverse. *Renewable Energy*, 34(4), 1084-1091. <https://doi.org/10.1016/j.renene.2008.07.004>

- Hernández, J. A., & Colorado, D. (2013). Uncertainty analysis of COP prediction in a water purification system integrated into a heat transformer using several artificial neural networks. *Desalination and Water Treatment*, 51(7-9), 1443-1456. <https://doi.org/10.1080/19443994.2012.693658>
- Hernández, J. A., Juárez-Romero, D., Morales, L. I., & Siqueiros, J. (2008). COP prediction for the integration of a water purification process in a heat transformer: With and without energy recycling. *Desalination*, 219(1-3), 66-80. <https://doi.org/10.1016/j.desal.2007.05.008>
- Herold, K. E., Radermacher, R., & Klein, S. A. (2016). *Absorption chillers and heat pumps* (2a ed.). CRC Press.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359-366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- Incropera, F. P., DeWitt, D. P., Bergman, T. L., & Lavine, A. S. (2011). *Fundamentals of heat and mass transfer* (7a ed.). John Wiley & Sons.
- International Energy Agency (IEA). (2023). *World Energy Outlook 2023*. <https://www.iea.org/reports/world-energy-outlook-2023>
- Jouhara, H., Khniefeh, A. J., Almahmoud, T., Delpech, B., Chauhan, A., & Tassou, S. A. (2018). Waste heat recovery technologies and applications. *Thermal Science and Engineering Progress*, 6, 268-289. <https://doi.org/10.1016/j.tsep.2018.04.017>

- Kalogirou, S. A. (2001). Artificial neural networks in renewable energy systems applications: A review. *Renewable and Sustainable Energy Reviews*, 5(4), 373-435. [https://doi.org/10.1016/S1364-0321\(01\)00006-5](https://doi.org/10.1016/S1364-0321(01)00006-5)
- Khalil, R., Al Horani, M., Yousef, A., & Sababheh, M. (2014). A new definition of fractional derivative. *Journal of Computational and Applied Mathematics*, 264, 65-70. <https://doi.org/10.1016/j.cam.2014.01.002>
- Khamooshi, M., Parham, K., Yari, M., Egelioglu, F., Salati, H., & Babadi, S. (2014). Thermodynamic analysis and optimization of a high temperature triple absorption heat transformer. *The Scientific World Journal*, 2014(1), 980452. <https://doi.org/10.1155/2014/980452>
- Kilbas, A. A., Srivastava, H. M., & Trujillo, J. J. (2006). *Theory and applications of fractional differential equations*. Elsevier Academic Press.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint*, arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>
- Labus, J., Bruno, J. C., & Coronas, A. (2013). Performance analysis of small capacity absorption chillers by using different modeling methods. *Applied Thermal Engineering*, 58(1-2), 305-313. <https://doi.org/10.1016/j.applthermaleng.2013.04.032>
- Laskin, N. (2000). Fractional quantum mechanics and Lévy path integrals. *Physics Letters A*, 268(4-6), 298-305. [https://doi.org/10.1016/S0375-9601\(00\)00201-2](https://doi.org/10.1016/S0375-9601(00)00201-2)
- LeCun, Y., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324. <https://doi.org/10.1109/5.726791>

- Ledesma, S., Belman-Flores, J. M., & Barroso-Maldonado, J. M. (2015). Analysis and modeling of a variable speed reciprocating compressor using ANN. *International Journal of Refrigeration*, 59, 190-197. <https://doi.org/10.1016/j.ijrefrig.2015.08.009>
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2(2), 164-168. <https://doi.org/10.1090/qam/10666>
- Lu, L., Shin, Y., Su, Y., & Karniadakis, G. E. (2019). Dying ReLU and initialization: Theory and numerical examples. *Communications in Computational Physics*, 28(5), 1671-1706. <https://doi.org/10.4208/cicp.OA-2020-0165>
- Machado, J. T., Kiryakova, V., & Mainardi, F. (2011). Recent history of fractional calculus. *Communications in Nonlinear Science and Numerical Simulation*, 16(3), 1140-1153. <https://doi.org/10.1016/j.cnsns.2010.05.027>
- Macías Díaz, J. E. (Ed.). (2022). *Fractional calculus - Theory and applications*. MDPI.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431-441. <https://doi.org/10.1137/0111030>
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence: August 31, 1955. *AI magazine*, 27(4), 12-14. <https://doi.org/10.1609/aimag.v27i4.1904>
- Misra, D. (2019). Mish: A self-regularized non-monotonic activation function. *arXiv preprint*, arXiv:1908.08681. <https://arxiv.org/abs/1908.08681>

- Mohanraj, M., Jayaraj, S., & Muraleedharan, C. (2012). Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systems—A review. *Renewable and Sustainable Energy Reviews*, 16(2), 1340-1358. <https://doi.org/10.1016/j.rser.2011.10.015>
- Mohanraj, M., Jayaraj, S., & Muraleedharan, C. (2015). Applications of artificial neural networks for thermal analysis of heat exchangers—A review. *International Journal of Thermal Sciences*, 90, 150-172. <https://doi.org/10.1016/j.ijthermalsci.2014.11.030>
- Montgomery, D. C., & Runger, G. C. (2010). *Applied statistics and probability for engineers*. John Wiley & Sons.
- Morales, L. I., Conde-Gutiérrez, R. A., Hernández, J. A., Huicochea, A., Juárez-Romero, D., & Siqueiros, J. (2015). Optimization of an absorption heat transformer with two-duplex components using inverse neural network and solved by genetic algorithm. *Applied Thermal Engineering*, *85*, 322-333. <https://doi.org/10.1016/j.applthermaleng.2015.04.018>
- Moran, M. J., Shapiro, H. N., Boettner, D. D., & Bailey, M. B. (2014). *Fundamentals of engineering thermodynamics* (8a ed.). John Wiley & Sons.
- Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. *In Numerical analysis* (pp. 105-116). Springer.
- Nocedal, J., & Wright, S. J. (2006). *Numerical optimization* (2a ed.). Springer.
- Parrales, A., Colorado, D., Díaz Gómez, J., Huicochea, A., Álvarez, A., & Hernández, J. (2018). New void fraction equations for two-phase flow in helical

heat exchangers using artificial neural networks. *Applied Thermal Engineering*, 130, 149-160. <https://doi.org/10.1016/j.applthermaleng.2017.10.139>

- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Piñeiro, G., Perelman, S., Guerschman, J. P., & Paruelo, J. M. (2008). How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecological Modelling*, 216(3-4), 316-322. <https://doi.org/10.1016/j.ecolmodel.2008.05.006>
- Podlubny, I. (1999). *Fractional differential equations: An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*. Academic Press.
- Prechelt, R. (1998). Automatic early stopping using cross validation: Quantifying the criteria. *Neural Networks*, 11(4), 761-767. [https://doi.org/10.1016/S0893-6080\(98\)00010-0](https://doi.org/10.1016/S0893-6080(98)00010-0)
- Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. *arXiv preprint*, arXiv:1710.05941.
- Rosales-García, J., Andrade-Lucio, J. A., & Shulika, O. (2020). Conformable derivative applied to experimental Newton's law of cooling. *Revista Mexicana de Física*, 66(2), 224-227. <https://doi.org/10.31349/RevMexFis.66.224>
- Russell, S., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3a ed.). Pearson.

- Şencan, A., Yakut, K. A., & Kalogirou, S. A. (2006). Thermodynamic analysis of absorption systems using artificial neural network. *Renewable Energy*, 31(1), 29-43. <https://doi.org/10.1016/j.renene.2005.03.011>
- Shirazi, A., Taylor, R. A., White, S. D., & Morrison, G. L. (2016). A systematic parametric study and feasibility assessment of solar-assisted single-effect, double-effect, and triple-effect absorption chillers for heating and cooling applications. *Energy conversion and management*, 114, 258-277. <https://doi.org/10.1016/j.enconman.2016.01.070>
- Sivanandam, S. N., & Deepa, S. N. (2006). *Introduction to neural networks using Matlab 6.0*. Tata McGraw-Hill Education.
- Solís-Pérez, J. E., Hernández, J. A., Parrales, A., Gómez-Aguilar, J. F., & Huicochea, A. (2022). Artificial neural networks with conformable transfer function for improving the performance in thermal and environmental processes. *Neural Networks*, 152, 44-56. <https://doi.org/10.1016/j.neunet.2022.04.016>
- Tablada, C. J., & Torres, G. A. (2020). Redes neuronales artificiales 1: Introducción. *FaMAF*. <https://www.famaf.unc.edu.ar/~revm/digital24-3/redes>
- Tarasov, V. E. (2010). *Fractional dynamics: Applications of fractional calculus to dynamics of particles, fields and media*. Springer.
- Tarasov, V. E. (2016). Local fractional derivatives of differentiable functions are integer-order derivatives or zero. *International Journal of Applied and Computational Mathematics*, 2(2), 195-201. <https://doi.org/10.1007/s40819-015-0054-6>

- Universidad de Sevilla. (s.f.). *Máquinas de absorción*.
<https://biblus.us.es/bibing/proyectos/abreproy/5070/fichero/CAPITULO+3%252FCAP%C3%8DTULO+3.0+MAQUINAS+DE+ABSORCION.pdf>
- Viera-Martin, E., Gómez-Aguilar, J. F., Solís-Pérez, J. E., Hernández-Pérez, J. A., & Escobar-Jiménez, R. F. (2022). Artificial neural networks: a practical review of applications involving fractional calculus. *The European Physical Journal Special Topics*, 231(10), 2059-2095. <https://doi.org/10.1140/epjs/s11734-022-00455-3>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & Van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261-272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wang, Z., Wang, X., Li, Y., & Huang, X. (2017). Stability and Hopf bifurcation of fractional-order complex-valued single neuron model with time delay. *International Journal of Bifurcation and Chaos*, 27(13), 1750209. <https://doi.org/10.1142/S0218127417502091>
- Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79-82. <https://doi.org/10.3354/cr030079>
- Yang, K. T. (2008). Artificial neural networks (ANNs): a new paradigm for thermal science and engineering. <https://doi.org/10.1115/1.2944238>
- Yu, H., & Wilamowski, B. M. (2011). Levenberg-Marquardt training. *Industrial Electronics Handbook*, 5(12), 1-15.

- Zubair, S. M., & Khedher, N. B. (2019). Performance analysis of absorption heat pump systems. *Energy Reports*, *5*, 123-130.
<https://doi.org/10.1016/j.egyr.2019.01.009>
- Zúñiga-Aguilar, C. J., Romero-Ugalde, H. M., Gómez-Aguilar, J. F., Escobar-Jiménez, R. F., & Valtierra-Rodríguez, M. (2017). Solving fractional differential equations of variable-order involving operators with Mittag-Leffler kernel using artificial neural networks. *Chaos, Solitons & Fractals*, 103, 382-403.
<https://doi.org/10.1016/j.chaos.2017.06.030>

ANEXOS

El desarrollo de la presente investigación de maestría no se limitó a la redacción del documento recepcional, sino que generó una serie de productos científicos, participaciones en foros internacionales y aplicaciones tecnológicas. Estas actividades representan el cumplimiento del compromiso de retribución académica y social, demostrando la viabilidad de transferir las metodologías basadas en Inteligencia Artificial hacia distintos ejes estratégicos de las energías limpias. A continuación, se detallan los logros obtenidos durante el periodo de posgrado.

I. Producción Científica y Publicaciones

La validación de las arquitecturas neuronales y los operadores no enteros requirió someter los hallazgos al escrutinio de la comunidad científica mediante la publicación de artículos formales.

Artículo en Revista Científica (En preparación): Derivado de los hallazgos principales de esta tesis, se tiene un manuscrito en etapa avanzada de redacción para su inminente postulación en una revista científica internacional indizada (JCR). Este documento expone el impacto del cálculo fraccionario y conformable en la superación de las pruebas de linealidad para el modelado del Coeficiente de Rendimiento (COP).

Artículo en Memorias de Congreso In Extenso: Se publicó el artículo titulado "Artificial Intelligence Applied To Absorption Heat Pumps" en las memorias del 5.º Congreso Internacional de Tecnología y Ciencia Aplicada (CITCA 2025). En este trabajo, se evaluó el desempeño de diversas funciones de transferencia de orden entero (como Hard Sigmoid y ReLU), sentando las bases estadísticas y metodológicas que posteriormente dieron origen a los modelos fraccionarios desarrollados en esta tesis (el documento íntegro se incluye en el Anexo D).

II. Participación en Congresos y Foros Internacionales

La difusión del conocimiento y la retroalimentación con pares investigadores se llevó a cabo mediante ponencias en foros de alto nivel, tanto dentro de la República Mexicana como en el extranjero.

5.º Congreso Internacional de Tecnología y Ciencia Aplicada (CITCA 2025): Se participó como ponente en este congreso celebrado del 26 al 28 de noviembre de 2025 en las instalaciones del Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) en Cuernavaca, Morelos. La ponencia abordó directamente la problemática de la predicción termodinámica en bombas de calor por absorción mediante redes neuronales, obteniendo el reconocimiento oficial grupal e individual por la destacada contribución (ver Anexo A y Anexo B).

15th International Seminar on Thermodynamic Engineering of Fluids (15ISTEF): Se participó en este seminario internacional organizado por la Universitat Rovira i Virgili en Tarragona, España, los días 22 y 23 de julio de 2026. En este foro, se presentó la investigación "Evaluation of different neural networks for the thermodynamic prediction of metal hydride alloys". Esta participación internacional evidencia la capacidad de colaboración con grupos de investigación externos (como el grupo CREVER en España) y proyecta el trabajo del autor a escala global (ver Anexo C).

III. Retribución Académica y Transferencia de Metodologías

Uno de los mayores impactos de un programa de maestría es la capacidad del investigador para extrapolar las habilidades adquiridas a nuevos campos de estudio. La investigación presentada en el 15ISTEF en España es la prueba directa de esta retribución.

Extrapolación del Modelo Neuronal a Nuevas Tecnologías: La rigurosa metodología de inteligencia artificial desarrollada en esta tesis —basada en la selección crítica de funciones de transferencia y la exigencia de superar pruebas de regresión lineal al 99% de confianza— fue

transferida exitosamente a un proyecto paralelo sobre el almacenamiento seguro de hidrógeno. Se utilizaron redes neuronales profundas (ResNets) para resolver el problema de diseño inverso y predecir propiedades termodinámicas complejas (entalpía y entropía de formación) en aleaciones de hidruros metálicos. Demostrar que los algoritmos creados para las bombas de calor pueden adaptarse para descubrir nuevos materiales de almacenamiento de hidrógeno subraya la transversalidad y la solidez de la formación recibida (el artículo correspondiente a esta investigación se anexa íntegramente en el Anexo E).

IV. Contribución a la Sostenibilidad

Todas las líneas de investigación abordadas por el autor durante la maestría responden directamente a los Programas Nacionales Estratégicos impulsados por las instituciones de ciencia en México.

Impacto en Energías Limpias: Tanto el perfeccionamiento de las bombas de calor por absorción para el reciclaje de calor residual, como la predicción termodinámica de aleaciones para sistemas de cero emisiones basados en hidrógeno, representan aportaciones directas a la transición energética y la sostenibilidad industrial.

Anexo A. Reconocimiento grupal del 5.º Congreso CITCA 2025



Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO

EL TECNOLÓGICO NACIONAL DE MÉXICO
A TRAVÉS DEL CENTRO NACIONAL DE INVESTIGACIÓN
Y DESARROLLO TECNOLÓGICO

OTORGA EL PRESENTE

RECONOCIMIENTO

A

**Gerardo Manríquez Pardo | José Alfredo Hernández Pérez | José
Francisco Gómez Aguilar | Armando Huicochea Rodríguez | Arianna
Parrales Bahena | Jesús Emmanuel Solís Pérez**

Por su destacada participación con la Contribución: Artículo Congreso (4-6 hojas máximo),
en el área: Inteligencia Artificial, con el Título:

Artificial Intelligence Applied To Absorption Heat Pumps

presentada en el marco del 5.º Congreso Internacional de Tecnología y Ciencia
Aplicada (CITCA), celebrado en el Centro Nacional de Investigación y Desarrollo
Tecnológico del 26 al 28 de Noviembre de 2025.

Cuernavaca, Morelos, 28 de noviembre de 2025

LORENZO O. HERNÁNDEZ
Director del CENIDET

cenidet
Centro Nacional de Investigación
y Desarrollo Tecnológico

MANUEL ADAM MEDINA
Coordinador General
5.º CITCA

5.º CITCA
Congreso Internacional de
Tecnología y Ciencia Aplicada
del 26 al 28 de noviembre, 2025
Cuernavaca, Morelos



Sello Digital:

i7T14EflvJ3XiwpXEdpwmnnlFeLFBFITILSdFDEF6uHcmGwV/CU6C84zpVai50rdGfqeE2kTfU6adW6nt/vDsCZJrlogl
QvOR0eu7qwm/aQ1o/n4W9EEyaA2H0p89A5JCopYaWoCQJr5C468UcMTqUoKo7zC3RlZKps33hunLtHMHSURd
RkXxcznpsALN7NzcOZNPvse45j25CrPQ2Do6gn+RYO15gSY7C0UxzUzfV4WtunrzZQ0aC0CDcJMy/tCtgFHB06Sl
pMvjtcnVFLc/YgJllgC30KjaWr3VWfRfHYGsCQOzBzkr4Z1rzeevdT79EgrzFaOwMO41JWkXNuAM9JGEPv77ersct4
gkMZGNiilbaagl7VTX3pu9vJg7uok9AwOWL3x7WDrSNxV1r2Eui5OI09J2eTQX2m6s0oielljmuTPxESFihjL9tE5Au1
xq5dvbniWjwFAzj0JtDD+sKyK06p9graUtk5ngxcX9pcxSCImUI9RIUESXyeEKp3yg0/XN1eFsQUfMqiKHeLqBmTDL
1RHcL3feE34UKbPY8miEmdfi8s5LbfbrLakmR52WUkUnMTZPdkiZd4e/77m+MyIVJ7VmFulDZ+hMcdz9EWPQTh
Q7utiWp8Ldl3JWLQEFJrNMCCtcjQfhVxo1UqyD1JA5y9s+VyPw43r4MtHA=

La firma electrónica que sustituye a la firma autógrafa del firmante, garantiza la integridad de la constancia y producirá los mismos efectos que las leyes otorgan a los documentos con firma autógrafa

Anexo B. Reconocimiento personal del 5.º Congreso CITCA

2025



Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO

EL TECNOLÓGICO NACIONAL DE MÉXICO
A TRAVÉS DEL CENTRO NACIONAL DE INVESTIGACIÓN
Y DESARROLLO TECNOLÓGICO

OTORGA EL PRESENTE

RECONOCIMIENTO

A

Gerardo Manríquez Pardo

Por su destacada participación con la Contribución: Artículo Congreso (4-6 hojas máximo),
en el área: Inteligencia Artificial, con el Título:

Artificial Intelligence Applied To Absorption Heat Pumps

presentada en el marco del 5.º Congreso Internacional de Tecnología y Ciencia
Aplicada (CITCA), celebrado en el Centro Nacional de Investigación y Desarrollo
Tecnológico del 26 al 28 de Noviembre de 2025.

Cuernavaca, Morelos, 28 de noviembre de 2025

LORENZO O. HERNÁNDEZ
Director del CENIDET

MANUEL ADAM MEDINA
Coordinador General
5.º CITCA

cenidet
Centro Nacional de Investigación
y Desarrollo Tecnológico

5 CITCA
Congreso Internacional de
Tecnología y Ciencia Aplicada
del 26 al 28 de noviembre, 2025
Cuernavaca, Morelos



Sello Digital:

i7T14EflvJ3XiwpXEdpwmnnlFeLFBFITILSdFDEF6uHcmGwV/CU6C84zpVai50rdGfqpe2kTfU6adW6nt/vDsCZJrlogl
QvOR0eu7qwm/aQ1o/n4W9EEyaA2H0p89A5JCopYaWoCOJr5C468UcMTqUoKo7zC3Rlzkps33hunLtHMHSURd
RkXxcznpsALN7NzcOZNPvse45j25CrPQ2Do6gn+RYO15gSY7C0UxzUzfV4WtunrzZQ0aC0CDcJMy/tCtgFHB06Sl
pMvjtcnVFLc/YgJllgC30KjaWr3VWfRfHYGsCQOzBzkr4Z1rzeevdT79EgrzFaOwMO41JWkXNuAM9JGEPv77ersct4
gkMZGNiilbaagl7VTX3pu9vJg7uok9AwOWL3x7WDrSNxV1r2Eui5OI09J2eTQX2m6s0oielljmuTPxESFihjL9tE5Au1
xq5dvbnIWjwFAzj0JtDD+sKyK06p9graUtk5ngxcX9pcxSCImIU9RIUESXyeEKp3yg0/XN1eFsQUfMqiKHeLqBmTDL
1RHcL3feE34UKbPY8miEmdfi8s5LbfbRlLakmR52WUkUnMTZPdkiZd4e/77m+MyIVJ7VmFulDZ+hMcdz9EWPQTh
Q7utiWp8Ldl3JWLQEFJrNMCCtcjQfhVxo1UqyD1JA5y9s+VyPw43r4MtHA=

La firma electrónica que sustituye a la firma autógrafa del firmante, garantiza la integridad de la constancia y producirá los mismos efectos que las leyes otorgan a los documentos con firma autógrafa

Anexo C. Certificado del 15th ISTEf (España)



UNIVERSITAT ROVIRA I VIRGILI

DOCTORAL PROGRAM ON THERMODYNAMIC ENGINEERING OF FLUIDS

CERTIFICATE OF PARTICIPATION

Gerardo Manríquez Pardo

Universidad Autónoma del Estado de Morelos, México

has participated at the

15th International Seminar on Thermodynamic Engineering of Fluids (15ISTEF)

**Evaluation of different neural networks
for the thermodynamic prediction of
metal hydride alloys**

Tarragona (Spain), July 22nd-23rd, 2026

Daniel Salavera
Member of the 15ISTEF organization

With the collaboration of



Anexo D. Artículo en Extenso: Artificial Intelligence Applied To Absorption Heat Pumps



Memorias del 5to Congreso Internacional de Tecnología y Ciencia Aplicada (5to CITCA 2025)
26 - 28 de Noviembre 2025, Cuernavaca, Morelos



Artificial Intelligence Applied To Absorption Heat Pumps

Manríquez Pardo, Gerardo*, Hernández Pérez, José A. **, Gómez Aguilar, José F.**, Huicochea Rodríguez, Armando**, Parrales Bahena, Arianna*, Solís Pérez, Jesús E.***

*POSGRADO - Centro de Investigación en Ingenierías y Ciencias aplicadas, MOR 62209
MEX (e-mail: gerardo.manriquez@uaem.edu.mx).

** Centro de Investigación en Ingenierías y Ciencias aplicadas, MOR 62209
MEX (e-mail: Alfredo.jose.ga_huicochea_arianna.parrales@uaem.mx).

*** Escuela Nacional de Estudios Superiores Unidad Juriquilla, QUE 76230
MEX (e-mail: jsolis@unam.mx).

Abstract: Artificial Neural Networks (ANNs) are a powerful tool for modeling complex physical systems like absorption heat pumps. However, previous models developed using a comprehensive experimental database (with 16 input variables, which include temperatures, concentrations, and pressures, to estimate the COP) have shown acceptable, but not optimal, results. Specifically, they struggled with linearity tests, indicating a lack of the precision and reliability required for robust engineering applications. The primary objective of this work is to improve upon these benchmarks. We introduce and evaluate four different integer-order transfer functions within ANN models to enhance the prediction of the Coefficient of Performance (COP). Our results show that two of these functions, Hard Sigmoid and Relu, yield models with significantly higher accuracy, surpassing the performance of previously published work. This study demonstrates that the careful selection of a transfer function is critical for creating highly precise and reliable predictive models.

Keywords: ANN (Artificial Neural Network), transfer function, artificial intelligence, absorption heat pumps for heating

1. INTRODUCTION

In recent years, data-driven modeling has become essential for understanding and optimizing complex physical systems. Due to the limitations of traditional analytical models, Artificial Neural Networks (ANNs) have emerged as a robust and effective approach for simulating energy systems, such as absorption heat pumps. ANNs, inspired by the structure of the human brain, are computational models capable of learning highly non-linear relationships directly from experimental data. Their ability to act as "universal function approximators" makes them an exceptionally powerful tool for creating accurate predictive models where underlying physical relationships are too complex to be described by simple equations.

1.1 Artificial Intelligence

To understand the tool, it is essential first to situate it within its parent field: Artificial Intelligence (AI). AI is an expansive branch of computer science that seeks to equip computers with capabilities that mimic human intelligence, such as learning and problem-solving. Its importance in modern engineering lies in its ability to tackle problems that cannot be solved with traditional, rule-based programming. Instead of telling a computer exactly how to model a system, the essence of AI lies in teaching machines to learn patterns from vast amounts of data, allowing them to find relationships and solutions to diverse problems on their own (Russell & Norvig, 2016). This

makes it the perfect approach for complex physical systems whose dynamics are not fully understood.

1.2 ANN

While AI is a vast field, this study focuses on a specific and powerful subfield known as Artificial Neural Networks (ANNs). ANNs are particularly well-suited for engineering problems because their architecture is designed to identify subtle patterns in numerical data, making them ideal for modeling physical phenomena. Inspired by the human brain.

The basic processing unit in a neural network is the neuron, which can be used in different types of architectures. Figure 2 shows the perceptron-type architecture, which is described by equations (1) and (2):

$$y = \phi \left(\sum_{n=1}^N w_n x_n + \theta \right), \quad (1)$$

$$V = \phi^{+Y} \quad (2)$$

Where x_n is the input to the neural network, w_n are the synaptic weights, and θ are the biases, randomly defined to expand the codomain of the activation functions (ϕ). The basic structure of ANNs follows a methodology that uses fundamental processing units (neurons). These neurons are interconnected, and each one receives an input data point. Each input is assigned a "weight," which determines its influence on the model's output (Parrales et al., 2018). According to

Tablada and Torres (2020), an ANN is divided into layers, as shown in **Figure 1**:

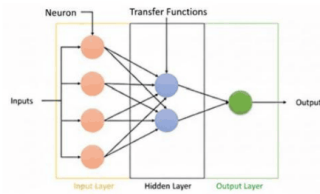


Figure 1. Structure of an ANN model. (Adapted from Cervantes-Bobadilla et al., 2021).

1. The first layer is known as the "input layer," and it contains the input parameters that are relevant to the output or results expected from the ANN.
2. The second layer, known as the "hidden layer," is where calculations are performed using the inputs.
3. The third and final layer, called the "output layer," is where the neurons that display the results are located.

1.3 Transfer Functions

Having established the layered structure of an ANN, it is crucial to understand the mathematical engine that operates within each neuron. Given that ANN models for predicting values are based on complex physical systems that a simple mathematical equation cannot represent, it is necessary to introduce equations that implement non-linearity. The transfer function takes the linear sum and transforms it into a non-linear output, unlocking the network's ability to approximate any complex function and accurately model real-world relationships (Haykin, 2009).

1.4 Absorption Heat Pumps and Coefficient of Performance (COP)

The physical system being modeled in this study is an absorption heat pump (**Figure 2**). Unlike conventional compression-based heat pumps, absorption systems use a heat source (such as waste heat from an industrial process) to drive a thermodynamic cycle that transfers thermal energy. The system consists of four main components: a generator, a condenser, an evaporator, and an absorber.

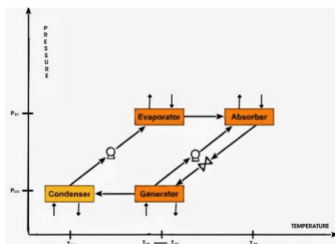


Figure 2. Schematic diagram of an absorption heat pump cycle. (Modified from Solís-Pérez et al., 2022).

The efficiency of such a system is measured by its Coefficient of Performance (COP). The COP is a dimensionless ratio that

quantifies the heat pump's efficiency. A higher COP indicates a more efficient system.

2. DATABASE

The database used in this work consists of 16 input variables (temperatures in °C, concentrations in %, and pressures in Hg) and one output variable, comprising a total of 4702 experimental tests. **Table 1** below shows the variables that will be used, along with their maximum and minimum obtained values:

Table 1. Absorption heat pump database.

Variable	Min Value	Max Value
$T_{in, GE-AB}$	71.84	91.61
$T_{in, EV-AB}$	69.34	89.93
$T_{out, AB-GE}$	74.46	98.27
$T_{in, AB-GE}$	73.27	92.15
$T_{out, GE-AB}$	71.84	91.61
$T_{out, GE-CO}$	72.04	83.89
$T_{in, CO}$	39.42	65.05
$T_{out, CO}$	26.75	33.79
$T_{in, EV}$	28.52	81.34
$T_{out, EV-AB}$	69.34	89.93
$X_{in, AB}$	51.66	55.33
$X_{out, AB}$	50.75	54.36
$X_{in, GE}$	50.75	54.36
$X_{out, GE}$	53.16	56.07
P_{AB}	7	11.5
P_{GE}	20	21.1
COP	0.21	0.39

For the development and evaluation of the model, this dataset was randomly partitioned into three subsets: 70% for the training set, 15% for the validation set, and 15% for the testing set.

Prior to model training, a sensitivity analysis was conducted to identify the input variables most critical to the target variable (COP). The analysis, utilizing both Pearson correlation and a Random Forest model, revealed that $T_{in, EV}$ the evaporator inlet temperature, has the most significant impact. The next most influential variables were identified as the generator-absorber inlet temperature, $T_{in, GE-AB}$ and the generator + absorber temperature, $T_{out, GE-AB}$.

3. BACKGROUND

With the database established, the next logical step is to understand the existing body of work related to it.

2.1 Best Reported Results with Integer-Order Functions

In 2018, Conde-Gutiérrez et al. (2018) developed ANN models using the same database. At that time, models with 1 to 4 hidden neurons were evaluated using the "Tansig" and "Logsig" transfer functions. The results of that work found that the model with the 16-3-1 structure using the "Tansig"

function obtained the best scores in statistical metrics, as shown in **Table 2**.

Table 2. Best results of ANN models with integer calculus (Conde-Gutiérrez et al., 2018)

Architecture ANN	Num. of Neurons	Epoch	RMSE	R ²
16-03-01	3	1000	0.0045	0.9935

2.2 Best Reported Results with Fractional-Order Functions

In 2023, the same database was used to create new models. In this case, fractional-order transfer functions were employed, which improved upon the results of the previous models (Viera Martin, 2023). **Table 3** shows the statistical metrics.

Table 3. Best results of ANN models with fractional calculus (Viera M, 2023)

Architecture	R ²	A.F.	Inf Inter	Sup Inter	Inf. Slope	Sup. Slope
16-02-01	0.9969	Tansig-ML	0.00056	0.002	0.9939	0.9981
16-02-01	0.9965	Logsig-ML	-0.002	-0.00054	0.9982	1.0023

Consequently, the primary objective of this investigation is to introduce and evaluate different new integer-order transfer functions within ANN models. By applying these new models to the absorption heat pump database, we aim to enhance the prediction of the Coefficient of Performance (COP). These functions were selected as they represent a novel application to this specific database, and their mathematical properties are hypothesized to be highly suitable for modeling this physical system. Success will be determined by comparing the statistical performance of these new models against established benchmarks in the literature.

4. DEVELOPMENT

To address the objective of creating more efficient and reliable models, a systematic development process was established. This process involves creating new ANN models using different transfer functions. The key components of this methodology are detailed in the following subsections.

4.1 Optimization Algorithm

In the training phase of an ANN, it is essential to have an optimization algorithm that adjusts the network's parameters (weights and biases) to find the optimal values for the model. The optimization algorithm plays a crucial role in determining the convergence speed when adjusting the parameters of an ANN.

In this case, the Levenberg-Marquardt algorithm was used, as it is the most used algorithm for Feed-Forward networks. Its equation is:

$$H = J^T J, \quad (3)$$

where:

H: is the approximate Hessian Matrix

J: is the Jacobian Matrix, which contains the partial derivatives of the network's errors with respect to the parameters (weights and biases).

4.2 Variable Normalization

It is necessary to perform normalization on the database variables. For this work, all inputs and outputs were normalized to the range of 0.1 to 0.9. This specific range is chosen to prevent the saturation of certain activation functions. Activation functions that saturate, like the S-shaped Tanh or the piecewise linear hardsigmoid, have boundaries where their output flattens out. If a neuron's output falls into these flat regions, its gradient becomes almost zero. This "vanishing gradient" problem can significantly slow down or completely stall the learning process. By mapping the data to a slightly narrower range, we ensure that the neuron activations remain in a more responsive, non-saturated region, which often leads to more stable and efficient training. The equation used is:

$$x_n = (v_{\max} - v_{\min}) \frac{x - x_{\min}}{x_{\max} - x_{\min}} + v_{\min}, \quad (4)$$

where x_n is the normalized variable, x is the value of the variable to be normalized, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values that the variable to be normalized can take, respectively; while $v_{\min}$ and $v_{\max}$ represent the minimum and maximum values that the variables can take within the normalization interval, which are defined by the user.

4.3 Functions for Evaluation

The equations and corresponding behavior graphs for the functions that will be used are shown below.

4.3.1 Lecun Tanh

$$f(x) = 1.7159 \cdot \left(\frac{e^x - e^{-x}}{e^x + e^{-x}} \right) \cdot \left(\frac{2}{3} x \right) \quad (5)$$

The Lecun Tanh function (Eq. 5) exhibits a classic "S-shaped" or sigmoidal curve, but it is scaled to have a steeper slope near the origin (Figure 4).

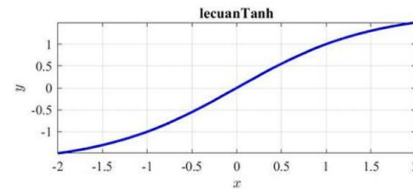


Figure 4. Behavior graph of the Lecun TanH function

4.3.2 Hard Sigmoid

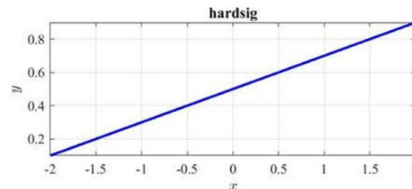


Figure 5. Behavior graph of the Hard Sigmoid function

$$f(x) \begin{cases} 0 & \text{if } x < -2.5 \\ 1 & \text{if } x > 2.5 \\ 0.2x + 0.5 & \text{if } -2.5 \leq x \leq 2.5 \end{cases} \quad (6)$$

The hardsigmoid function (Eq. 6) is a piecewise linear approximation of the sigmoid function (Figure 5). Its primary advantage is computational efficiency, as it avoids the costly exponential calculations of the logistic sigmoid function, replacing them with much faster linear operations.

4.3.3 Relu

$$f(x) = \max(0, x) \quad (7)$$

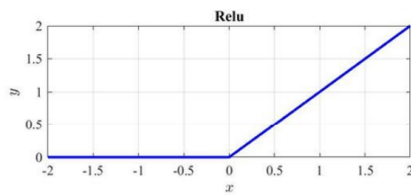


Figure 6. Behavior graph of the Relu function

The Rectified Linear Unit (Relu) function (Eq. 7) is fundamentally different. It is a piecewise linear function that is zero for all negative inputs and increases linearly for positive inputs (Figure 6). This simplicity makes it computationally very efficient.

4.3.4 Randomized Relu

$$f(x) = \max(Rx, x) \quad (8)$$

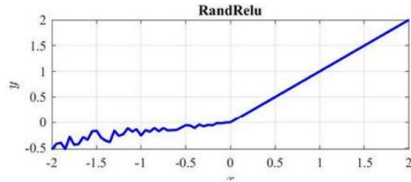


Figure 7. Behavior graph of the RandRelu function

This is a variation of Relu (Eq. 8) where the slope for negative inputs is a small, random value (R) rather than zero. (Figure 7).

4.4 Flowchart

Figure 8 provides a visual representation of the methodology used to develop, train, and evaluate each ANN model in this study.

A key component of this training methodology was the use of an early stopping criterion. The training process for each ANN was set to run for a maximum number of epochs. The patience for this criterion was set to 100,000 iterations (epochs) without improvement in the best-recorded R-value.

4.5 Statistical analysis

The evaluation of Artificial Neural Network (ANN) models requires statistical analysis to validate their effectiveness and reliability.

The statistical metrics calculated in the models studied will be the following:

Pearson's Linear Correlation (Eq. 9):

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (9)$$

Mean Squared Error (Eq. 10):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

Root Mean Squared Error (Eq. 11):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (11)$$

5. RESULTS

The following figures present the behavior graphs for the models configured with two hidden neurons. Figure 9 corresponds to the Rand Relu function, Figure 10 to Hard Sigmoid, Figure 11 to Lecun TanH, and Figure 12 to Relu.

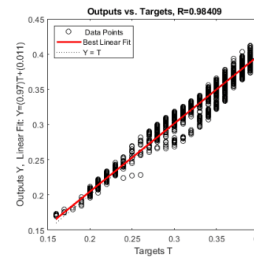


Figure 9. Linear correlation coefficient for Rand Relu, R=0.98409

The Randomized Relu function shows a good linear correlation. However, there is visible dispersion of the data points around the red line, particularly for COP values between 0.25 and 0.4.

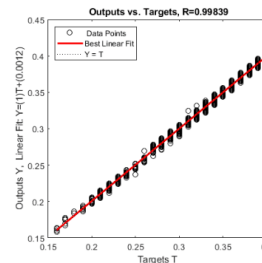


Figure 10. Linear correlation coefficient for Hard Sigmoid, R=0.99839

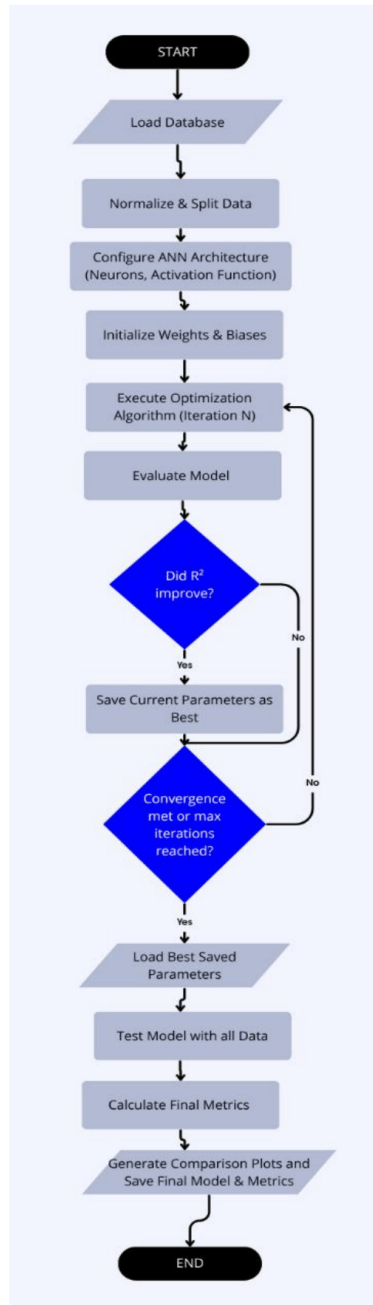


Figure 8. Flowchart from the ANN codes.

The Hard Sigmoid function yields a model that passes the linearity test. The data points form a very tight and consistent

band around the ideal line, with minimal deviation across all operating conditions.

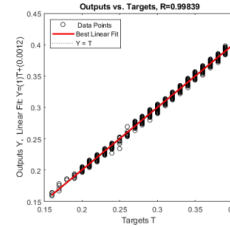


Figure 11. Linear correlation coefficient for Lecun TanH, $R=0.98364$

This model displays a significant issue with linearity. For low COP values (below 0.2), the predictions are quite accurate and close to the ideal line. However, as the target COP increases towards 0.4, the data points become increasingly dispersed and inaccurate.

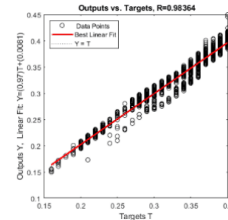


Figure 12. Linear correlation coefficient for Relu, $R=0.99839$

Similar to hardsigmoid, the Relu function demonstrates excellent linearity and achieves the exact same top-performing correlation coefficient of $R=0.99839$. The Relu model exhibits minimal dispersion, though its data point distribution differs slightly from hardsigmoid's in the central COP ranges.

ANN models with one hidden neuron (a 16-1-1 structure) were evaluated but proved non-competitive, failing to outperform established models from the literature.

Consequently, this section focuses only on the results from the more successful two-neuron configurations. **Table 4** summarizes the statistical results for the models with a 16-2-1 structure, evaluated using the entire dataset (100%).

Table 4. Evaluation of ANN models with 16-2-1 structure

Func	R^2	MS E	RM SE	Inf Inter	Sup Inter	Inf Slope	Sup Slope
Lecuh TanH	0.98364	0.00018	0.01339	0.0031	0.0095	0.9961	0.9843
Hard Sig	0.99839	0.00007	0.00815	-0.0019	0.002	0.9953	1.0066
Relu	0.99839	0.00007	0.00847	-0.0019	0.0022	0.9947	1.0064

Rand	0.984	0.00	0.01	0.0	0.0	0.9	0.9822
Relu	09	018	343	056	121	637	

Based on the strong linearity tests and superior statistical metrics, Hard Sigmoid and Relu are the clear winners. They achieved superior metrics (R^2 , errors) and, critically, the reliable, linear behavior that prior models failed to meet.

6. ANALYSIS OF RESULTS

Based on the results obtained, it can be concluded that we successfully improved upon the last reported model using this same data. Of the transfer functions used, two—hardsigmoid and Relu—were highly successful, achieving nearly identical top-tier performance. Both models produced a correlation coefficient of $R=0.99839$, showing better values than those reported in the literature and, most importantly, passing the critical linearity tests. However, for LecunTanh and Randomized Relu, the same cannot be said; their use is not recommended.

7. CONCLUSIONS

The objectives of this work were successfully met by developing and rigorously evaluating ANN models for COP prediction using four transfer functions not previously tested on this specific database. The investigation yielded two highly competitive and effective functions: Hard Sigmoid and Relu. Their competitiveness is demonstrated by producing results superior to published models and passing the linearity test where prior models failed. This success suggests the underlying physics of this system is best modeled by functions exhibiting strong linear behavior combined with critical nonlinearities. Conversely, models employing Lecun TanH and Randomized Relu are not recommended. Beyond predictive accuracy, the primary value of these robust models is their practical application in system optimization and real-time predictive control.

ACKNOWLEDGEMENTS

Jesús Emmanuel Solís Pérez thanks PAPIIT-DGAPA, UNAM, for the computational resources provided through projects IA102525 and TA100223.

REFERENCES

- Conde-Gutiérrez, R.A., Cruz-Jacobo, U., Huicochea, A., Casolco, S.R., and Hernández, J.A. (2018). Optimal multivariable conditions in the operation of an absorption heat transformer with energy recycling solved by the genetic algorithm in artificial neural network inverse.
- Viera Martin, E. (2023). Algoritmos de optimización aplicados a la solución de redes neuronales multivariadas. Tesis de Doctorado, Tecnológico Nacional de México/Centro Nacional de Investigación y Desarrollo Tecnológico, Cuernavaca, Morelos, México.
- Haykin, S. (2009). Neural Networks and Learning Machines (3rd ed.). Prentice Hall.
- Parrales, A., Colorado, D., Díaz Gómez, J., Huicochea, A., Álvarez, A., and Hernández, J. (2018). New void fraction equations for two-phase flow in helical heat exchangers using artificial neural networks. Applied Thermal Engineering, 149-160.
- Russell, S. and Norvig, P. (2016). Artificial Intelligence: A Modern Approach (3rd ed.). Pearson.
- Tablada, C. J. and Torres, G. A. (2020). Redes Neuronales Artificiales I. Introducción. FaMAF.
- Solís-Pérez, J.E., Hernández, J.A., Parrales, A., Gómez-Aguilar, J.F., and Huicochea, A. (2022). Artificial neural networks with conformable transfer function for improving the performance in thermal and environmental processes. Neural Networks, 152, 44–56.
- Cervantes-Bobadilla, M., Hernández-Pérez, J.A., Juárez-Romero, D., Bassam, A., García-Morales, J., Huicochea, A., and Jaramillo, A. (2021). Control scheme formulation for a parabolic trough collector using inverse artificial neural networks and particle swarm optimization. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 43:176.

Anexo E. Artículo Corto: Evaluation of different neural networks for the thermodynamic

*15th International Seminar on Thermodynamic Engineering of Fluids
Tarragona, 22-23 July 2026*

EVALUATION OF DIFFERENT NEURAL NETWORKS FOR THE THERMODYNAMIC PREDICTION OF METAL HYDRIDE ALLOYS

Gerardo Manríquez-Pardo^(a), Abdurahman Hasan Assagaf^(b), Daniel Salavera^(b), J. A. Hernández^(c), J. F. Gómez-Aguilar^(c)

(a) POSGRADO - Centro de Investigación en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Col. Chamilpa, C.P. 62209 Cuernavaca, Morelos, Mexico

(b) CREVER – Group of Applied Thermal Engineering, Department of Mechanical Engineering, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Spain

(c) Centro de Investigación en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Col. Chamilpa, C.P. 62209 Cuernavaca, Morelos, Mexico

ABSTRACT

The global transition toward a zero-emission energy system requires efficient hydrogen storage. Solid-state metal hydrides offer high volumetric hydrogen density and safety at near-ambient pressures. The performance of these reversible systems depends on formation enthalpy (ΔH) and entropy (ΔS), which dictate operating windows. Designing alloys with tailored thermodynamic properties remains a major challenge [1]. Historically, discovering alloys relied on empirical methods or expensive Density Functional Theory (DFT) calculations [2]. A key challenge is solving the inverse design problem: predicting alloy composition from target thermodynamic properties. Using the Metal Hydride Pairs database, we mapped four inputs (Temperature, Pressure, ΔH , ΔS) to 25 sparse elemental stoichiometric fractions (Mo, Zr, Mg, Ni, Al, La, Co, Ce, Sm, H, Si, B, V, Fe, Ti, Ca, Nd, Ge, Mn, Cr, Sn, Cu, Mn, Ga, Y).

To navigate the highly asymmetric input-output dimensionality, we evaluated several Artificial Neural Network (ANN) architectures, including deep residual networks (ResNets), to mitigate vanishing gradients and improve learning under sparse outputs. We tested various transfer functions (Sigmoid, Tanh, ReLU, GELU, dSiLU, RBF) [3] and optimizers (Levenberg-Marquardt, Adam, L-BFGS). To ensure physical consistency, all models were required to pass a linearity test at the 99% confidence level. [4]. Two top-performing architectures were identified: a deep Tansig ResNet (4-40-40-40-25) and a shallower Inverse Multiquadric (IMQ) ResNet (4-60-60-25). The Tansig model achieved the highest correlation ($R=0.998$) and nearly passed the linearity test (Figure 1). Conversely, the IMQ model ($R=0.9929$) provides a highly competitive alternative that significantly reduces the computational parameter count (Figure 2). Meeting this stringent statistical criterion suggests that the ANN successfully captured the underlying thermodynamic relationships, providing a reliable framework for accelerating metal hydride alloy discovery.

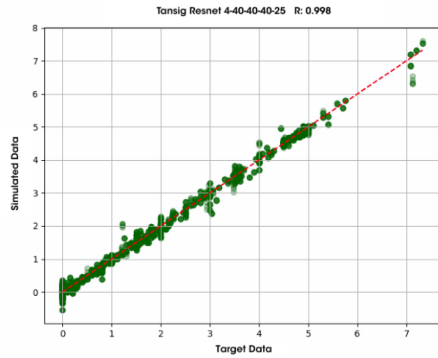


Figure 1 Linear regression analysis of Tansig Resnet 4-40-40-25

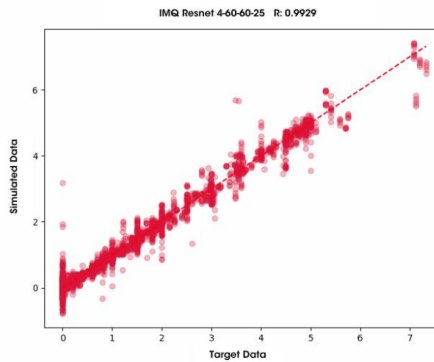


Figure 2 Linear regression analysis of IMQ Resnet 4-60-60-25

Acknowledges

The authors gratefully thank the support and Master's scholarship provided by the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECITI, formerly CONAHCYT), as well as the institutional backing from the Universidad Autónoma del Estado de Morelos (UAEM) and CIICAp.

References

- [1] Surname, Initial/s; Surname2, Initial/s. *Publish* (2017) 3, 4-6.
- [1] Schlapbach, L., & Züttel, A. (2001). Hydrogen-storage materials for mobile applications. *Nature*, 414(6861), 353-358.
- [2] Witman, M., Ling, S., Grant, D. M., Walker, G. S., Agarwal, S., Stavila, V., & Allendorf, M. D. (2020). Extracting an empirical metallic hydride design limit from machine learning. *The Journal of Physical Chemistry Letters*, 11(1), 40-47.
- [3] Ramprasad, R., Batra, R., Pilia, G., Mannodi-Kanakkithodi, A., & Kim, C. (2017). Ramprasad, R., Batra, R., Pilia, G. et al. Machine learning in materials informatics: recent applications and prospects. *npj Computational Materials*, 3(54), 1-13. <https://doi.org/10.1038/s41524-017-0056-5>
- [4] Agrawal, A., & Choudhary, A. (2016). Perspective: Materials informatics and big data: Realization of the fourth paradigm of science in materials science. *APL Materials*, 4(5), 053208.



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

JEFATURA DE POSGRADO EN INGENIERÍA Y CIENCIAS APLICADAS



Cuernavaca, Morelos, 09 de septiembre de 2026.

DR. ARTURO MOLINA OCAMPO
COORDINADOR DEL POSGRADO
EN INGENIERÍA Y CIENCIAS APLICADAS
P R E S E N T E

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada: MEJORA EN LA PREDICCIÓN DEL COP EN BOMBAS DE CALOR POR ABSORCIÓN MEDIANTE EL USO DE NUEVAS FUNCIONES DE TRANSFERENCIA DE ORDEN ENTERO, FRACCIONARIO Y CONFORMABLE EN REDES NEURONALES ARTIFICIALES que presenta el alumno **Gerardo Manríquez Pardo**, para obtener el título de la **Maestría en Ingeniería y Ciencias Aplicadas**.

Nos permitimos informarle que nuestro voto es **APROBATORIO**.

Atentamente
Por una humanidad culta

DR. ALBERTO ARMANDO ÁLVAREZ GALLEGOS

DR. ARMANDO HUICOCHEA RODRÍGUEZ

DR. ABISAI ACEVEDO QUIROZ

DR. JOSÉ FRANCISCO GÓMEZ AGUILAR

DR. JOSÉ ALFREDO HERNÁNDEZ PÉREZ

Se adiciona efirma UAEM



Av. Universidad 1001 Col. Chamilpa, Cuernavaca Morelos, México, 62209, Edificio 65,
Tel. 777 329 70 00 Ext. 6208 / ciicap.posgrado@uaem.mx

UAEM
RECTORÍA
2023-2029



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento con firma electrónica UAEM, soportada por el certificado vigente a la fecha de su elaboración y con efectos plenos de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS PUBLICADOS en el ÓRGANO INFORMATIVO UNIVERSITARIO "ADOLFO MENÉNDEZ SAMARÁ" número 117 de fecha 20 de abril de 2021.

Sello electrónico

JOSE FRANCISCO GOMEZ AGUILAR | Fecha:2026-09-14 09:35:55 | FIRMANTE

k1MORfDqhGaq45MAm+v4p02l9mwXWe2ihpOx9+Vd7zLsDWH2zRzWra/hAUDFn+9Zy2AT899HN1RGeYb1fH0AWSp9CpxByfKATeJl/670yhbJFD1yf78dqOIJgLqgMiB3Az+Rd01HtG4qzlitV317cLQY6vVxc3YtoVouNmvaJZwUwto8P+qjeXZTBdWC3bJdktH8P6cxT9+Gyy6nYJsgqcW7dJTJKg645dxjnxmndLN5v1Fa/4JDBlgmYvDd+kjhGqZV4lrSiIJdO4+uVxyO3uXOsPqEzXsGgUpvGYqjOdRo56JCSU3HaGB3hiyrTwpQgXVEMEFk4cWjV6Q==

ARMANDO HUICOCHEA RODRIGUEZ | Fecha:2026-09-14 09:40:07 | FIRMANTE

CoDmdOKN6Asvim47eFvPIZ+eQAms2y2yb9Dko8lkC6gQHVMY/Nq84Np8OcqRh35Km3cqJHolclXont6LT9GbPjsPv16biB21xIMfpr09/G2ZBY+10pBEAGolV9urnsCe30dW+mNc/SPOE82jTazR8jjcKG6D93Rd90Svqbs9EYV2DEVAmAC7qdZop4xVgDFBLvmhWA7lkLkjrUKJ24olrjfi5Eaj862j4R+DuRUO6QBf9leo7Hb7Ps/llOGKxyaHDD1XgUI+RwVmR4BGh1fRzdZlfi5UyDtnI45hRUm1cox+PXZz9ot1DA2Mw5ebS2kYczbiey2DpMupDM2uzxvg==

ABISAI ACEVEDO QUIROZ | Fecha:2026-09-14 09:50:03 | FIRMANTE

xAbCdpy7I2DIOsV1+n61YSyUGQXog5E1wrsPC05uJwCuoVC0USrEmQx4i/ZHPLH7BDMaU2SfMvNfbUeEr72/EHU5i1XIO75FqjazyCdG4USvRmOEuMww4pOFW9wJcG+mnrCxnwVj7pcGmOSqUBh0f5r7Q8f5bLnBJnhfzWaYgR1Pbda+9fedifX88UzyscLQcxNZQJXurpszyxrN201JTsH031cpF7AVr/MQ8QXJEiNUuglhtV4yfwksp8N78e/BEemUIY6PdM29D43sTz4SujMgcljOGKIsHGE9u4EafXtO1IWH3sDlz+umnYHCfqQ6mX5wzNVKbXzUdy2mw==

ALBERTO ARMANDO ALVAREZ GALLEGOS | Fecha:2026-09-14 10:20:24 | FIRMANTE

mYTtMukJ9O2m1ZBEz15DaJz0m4ygBcyqRBkij8NlbG0YEHLLUuDD2TuvQbilp/J80rIDBWHb7KoQaalz9p0gXONNkJFqyUlenS3T/AQGRCyK7mEvN82lbFsl8cZ72phfZ7ZoLuL7vHnnKD7Cxb3ZFexDqCB8zGPGiOcH5QubHl2ululaiSzUev9luVMDbzSiDd0tPbIKpyl0mpfpjQ99rJU5M2j3kvWbP4/NcJwWaxJL4NIKufSHeYFHyaA+seelkQERro8xYZYTZcAKdB05zyGxcUpqyl/MiMB8VpwXnnKqZzfG8lksvzDdGODkkigAAwLndFhSqkUUVLUobXtg==

JOSE ALFREDO HERNANDEZ PEREZ | Fecha:2026-09-14 11:49:29 | FIRMANTE

ochPODe2Nr32LNMHRH1/yNs2FsYvg12ANDkOL9wr8PRqElbFmdwozMCliPBCDpV5+VGUYZHNm/xOJzkhwTtDjITFX22DjZb5ws57XRb+P4ED/reQ8IH1bidhB42V0BgvMGygcdQDSPYQmLxgU2gu/cWX/NqN2Jl3k4pWFVxvUK7VAJDJ6z3lt70g2qXqsg8GPQOgO7Whpe0gjr/EkoXD3wTgC8wNU51ggMYdUsji13E3yBQsmeG7Nu8pNxbp6mkYdU+bemNpQ6y2KdCRulxhNwCH8oRQeCNy8AdOgepLuT5oly20lm71Qts/4lBmS+skTcdsyZ8Um6Ouw==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o
escaneando el código QR ingresando la siguiente clave:



EDdVknSGS

<https://firma.uaem.mx/noRepudio/KYeU46QkkWMGcgcmzkSr5JeundUU8V45>



UAEM
RECTORÍA
2023-2029