

PREDICCIÓN PARAMÉTRICA DEL CAMPO DE TEMPERATURAS EN UN PROBLEMA DE CONDUCCIÓN DE CALOR CON FUENTE TÉRMICA MÓVIL, MEDIANTE REDES NEURONALES INFORMADAS POR FÍSICA SIN DATOS ROTULADOS

PARAMETRIC PREDICTION OF THE TEMPERATURE FIELD IN A HEAT CONDUCTION PROBLEM WITH A MOVING HEAT SOURCE, USING PHYSICS-INFORMED NEURAL NETWORKS WITHOUT LABELLED DATA

Benjamin A. Tourn^a, Juan C. Álvarez Hostos^{a,b} y Marcos U. Maillot^c

^a*Centro de Investigación y Transferencia (CIT), Universidad Nacional de Rafaela (UNRaf)/Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Bv. Roca 989, 2300 Rafaela, Argentina, citrafaela@conicet.gov.ar, <https://citrafaela.conicet.gov.ar/>*

^b*Departamento de Materiales, Facultad de Ingeniería Química, Universidad Nacional del Litoral (UNL), Santiago del Estero 2829, 3000 Santa Fe, Argentina, <https://www.fiq.unl.edu.ar/materiales/>*

^c*Departamento de Ingeniería Eléctrica, Universidad Tecnológica Nacional, Facultad Regional General Pacheco (UTN-FRGP), Av. Hipólito Yrigoyen 288, 1617 Gral. Pacheco, Argentina, <https://www.frgp.utn.edu.ar/>*

Palabras clave: PINN, Deep Learning, conducción de calor, solución paramétrica.

Resumen. En este trabajo se propone utilizar redes neuronales informadas por física (PINN, por sus siglas en inglés) para resolver paramétricamente las ecuaciones de gobierno del problema de conducción de calor bidimensional con fuente térmica móvil. Se computan soluciones generales no sólo en función de las variables independientes espacial y temporal, sino también de un parámetro geométrico que modula la intensidad de la fuente térmica denominado radio característico. Los resultados muestran un ajuste preciso frente a las soluciones de referencia obtenidas mediante el método de elementos finitos. No obstante, se reportan y analizan algunos modos de falla asociados con la satisfacción de las condiciones de borde, junto con aquellos encontrados a medida que la fuente térmica se torna más intensa.

Keywords: PINN, Deep Learning, heat conduction, parametric solution.

Abstract. In this work, we propose use physics-informed neural networks (PINN) to parametrically solve the governing equations of a two-dimensional heat conduction problem with a moving heat source. We obtain general solutions not only as a function of the space-and-time independent variables but also to a geometric parameter that controls the intensity of the thermal source, named the characteristic radius. The results show a precise fit with respect to the reference solutions obtained by the finite element method. Nevertheless, we report and analyze some failure modes linked with the accomplishment of the boundary conditions and also to those found as the heat source becomes more intense.

1. INTRODUCCIÓN

Los formidables avances en el desarrollo de herramientas de Machine Learning registrados en los últimos años han derivado en innumerables y sofisticadas aplicaciones de propósito general. El ámbito científico-tecnológico, aunque en menor medida, no ha sido ajeno a tales avances (Rabczuk y Bathe, 2023). En este contexto, las redes neuronales informadas por física (PINN, por sus siglas en inglés) constituyen un desarrollo novedoso y prometedor, con una tasa de evolución extremadamente veloz (Cuomo et al., 2022). En esencia, las PINN son esquemas de Deep Learning concebidos para resolver problemas directos e inversos descritos por ecuaciones en derivadas parciales (Raissi et al., 2019).

El enfoque planteado por los modelos PINN permite eliminar la dependencia de grandes volúmenes de datos rotulados (Big Data) que caracteriza a muchas de las herramientas de Machine Learning utilizadas como “modelos de caja negra” (Rabczuk y Bathe, 2023). Esto provocó un renovado interés en la comunidad científico-tecnológica debido al desafío práctico que significa configurar experimentos sofisticados en ambientes de trabajo desafiantes para obtener datos confiables en cantidad necesaria. Otra característica fundamental es la posibilidad de incorporar información de principios físicos elementales (Raissi et al., 2019; Cai et al., 2021), como el principio de conservación de la energía en el problema de conducción de calor aquí planteado. De esta manera, las soluciones obtenidas por los modelos PINN satisfacen por diseño tales principios, algo que los modelos de Deep Learning tradicionales de caja negra no son capaces de asegurar.

Los modelos PINN son particularmente útiles para tratar problemas inversos en presencia de datos escasos y/o ruidosos y para el descubrimiento de relaciones físicas ocultas (Karniadakis et al., 2021). En el caso de problemas directos, sin embargo, no son capaces de competir (aún) con estrategias de resolución numérica maduras y ampliamente aceptadas como el método de elementos finitos (MEF). No obstante, la modificación de cualquier parámetro presente en las ecuaciones de gobierno, ya sea propiedades materiales o características geométricas, implica resolver nuevamente el sistema de ecuaciones algebraicas construido por el método en cada paso de tiempo. Por este motivo, la obtención de soluciones paramétricas es una posibilidad que queda fuera de su alcance. En cambio, los esquemas PINN permiten incluir los parámetros que se desean variar como entradas adicionales en los modelos, por lo que sus predicciones resultan generales en relación con las variables independientes y con tales parámetros (Hosseini et al., 2023; Cuomo et al., 2022). Esta característica permitiría reducir drásticamente el costo computacional en problemas de optimización al emplearlos como modelos sustitutos (Rabczuk y Bathe, 2023).

Existen diversos antecedentes en la resolución de problemas de conducción de calor mediante PINN (Cai et al., 2021), y específicamente aquellos donde se considera una fuente térmica móvil. Hosseini et al. (2023) calcularon los perfiles de temperatura y las dimensiones del baño de fusión que evolucionan durante el proceso de fusión de lecho de polvo láser para cualquier conjunto dado de propiedades térmicas del material y condiciones de proceso a un costo computacional prácticamente nulo. Zhu et al. (2021) aplicaron modelos de aprendizaje profundo basados en física para el modelado tridimensional de procesos manufactura aditiva. Xie et al. (2022) obtuvieron predicciones tridimensionales del campo de temperatura en metales durante procesos de deposición directa de energía monocapa y multicapa.

En este trabajo se propone la construcción de modelos PINN para predecir de manera paramétrica el campo de temperaturas en un problema de conducción de calor bidimensional en régimen transiente con una fuente térmica que se desplaza a velocidad constante sobre una pla-

ca plana. El modelo ofrece predicciones que dependen no solo de las variables independientes espacio-temporales de las ecuaciones de gobierno, sino también del “radio característico”, que modula la intensidad de la fuente. El caso emula rudimentariamente un proceso tecnológico de soldadura en el que la fuente térmica representa el aporte concentrado de calor utilizado para fundir el metal. La implementación busca explorar el potencial del modelado de problemas conducción de calor vía PINN, resaltar sus fortalezas y debilidades, y descubrir posibles estrategias de mejora.

2. PROBLEMA DE CONDUCCIÓN DE CALOR EN RÉGIMEN TRANSIENTE

Sea una placa plana rectangular sobre la que se aplica una fuente de calor Gaussiana móvil sobre el eje de simetría longitudinal, con velocidad constante. El material es homogéneo e isótropo y sus propiedades termofísicas son independientes de la temperatura. La superficie de la placa no presenta pérdidas por convección ni por radiación hacia el medio circundante. Bajo estas hipótesis, las ecuaciones de gobierno adoptadas en forma adimensional son:

$$\frac{\partial u(\xi, \eta, \tau)}{\partial \tau} = \Delta u(\xi, \eta, \tau) + \dot{Q}, \quad \forall(\xi, \eta) \in \Omega, \tau \in (0, \tau_f], \quad (1a)$$

$$\nabla u(\xi, \eta, \tau) = 0, \quad \forall(\xi, \eta) \in \partial\Omega, \tau \in (0, \tau_f], \quad (1b)$$

$$u(\xi, \eta, 0) = u_0, \quad \forall(\xi, \eta) \in \Omega, \tau = 0. \quad (1c)$$

donde

$$\xi = \frac{x}{L}, \quad \eta = \frac{y}{L}, \quad \rho_0 = \frac{r_0}{L}, \quad \tau = \frac{k}{\rho c L^2} t \equiv Fo, \quad u = \frac{T}{T_0},$$

son las variables definidas en base a las coordenadas espaciales (x, y) , al tiempo t , a la temperatura T en el problema original (con dimensiones) y a diversos parámetros cuyos símbolos y sus valores correspondientes se encuentran listados en la tabla 1. Asimismo, Ω representa el dominio rectangular sin dimensiones definido como $\Omega = [0, 1] \times [-H/2L, H/2L]$, $\partial\Omega$ su perímetro y \dot{Q} el flujo de calor proporcionado por la fuente térmica.

TABLA 1: Parámetros físicos y geométricos del problema.

Descripción	Valor
Densidad ρ (kg mm ⁻³)	7,6e-6
Conductividad térmica k (W mm ⁻¹ K ⁻¹)	0,025
Capacidad térmica específica c (J kg ⁻¹ K ⁻¹)	658
Longitud del dominio L (mm)	100
Ancho del dominio H (mm)	50
Tiempo total del análisis transiente t_f (s)	50
Temperatura inicial T_0 (K)	273
Potencia total de la fuente \dot{Q}_T (W)	62,83185
Velocidad de la fuente v (mm s ⁻¹)	2
Radio característico r_0 (mm)	1,0 a 10,0

El descrito es un caso idéntico al presentado en el trabajo de [Álvarez Hostos et al. \(2022\)](#). Allí se muestran detalles de cómo la fuente térmica \dot{Q} puede ser expresada en términos tanto de las variables independientes adimensionales ξ , η y τ como del radio característico ρ_0 :

$$\dot{Q}(\xi, \eta, \tau, \rho_0) = \frac{\dot{Q}_T^*}{\pi \rho_0^2} e^{-(\xi - v^* \tau)^2 / \rho_0^2} e^{-\eta^2 / \rho_0^2}, \quad \text{con} \quad \dot{Q}_T^* = \frac{\dot{Q}_T}{k T_0} \quad \text{y} \quad v^* = \frac{\rho c L}{k} v,$$

donde $v^* \tau$ indica la posición de la fuente sobre el eje ξ en el instante τ .

3. MODELO PINN

De acuerdo con el enfoque original de Raissi et al. (2019), el objetivo de un modelo PINN para el problema de conducción de calor es aproximar el campo de temperaturas $u(\xi, \eta, \tau)$ que resuelve el sistema de ecuaciones (1) mediante una red neuronal de avance completamente conectada (FC-FFNN, por sus siglas en inglés) representada matemáticamente como $\hat{u} = f_{\theta}(\xi, \eta, \tau)$. Luego, se procede a escribir las ecuaciones de gobierno (1) como residuos:

$$r_{\text{PDE}}(\boldsymbol{\theta}) := \frac{\partial f_{\theta}(\xi, \eta, \tau)}{\partial \tau} - \Delta f_{\theta}(\xi, \eta, \tau) - \dot{Q}(\xi, \eta, \tau, \rho_0), \quad (2a)$$

$$r_{\text{BC}}(\boldsymbol{\theta}) := \nabla f_{\theta}(\xi, \eta, \tau), \quad (2b)$$

$$r_{\text{IC}}(\boldsymbol{\theta}) := f_{\theta}(\xi, \eta, 0) - u_0, \quad (2c)$$

donde los subíndices PDE, BC e IC refieren a las ecuaciones de balance de energía, condiciones de borde y condiciones iniciales, respectivamente. El cómputo de las derivadas de $f_{\theta}(\xi, \eta, \tau)$ respecto a las coordenadas espaciales y al tiempo se lleva a cabo mediante el algoritmo de diferenciación automática (Raissi et al., 2019) conocido como AD o *autodiff*, empleado de manera estándar en modelos de Deep Learning. En esencia, tal mecanismo permite obtener de manera computacionalmente eficiente la derivada analítica de la salida de una red neuronal con respecto a cualquier parámetro del modelo y a cualquiera de sus entradas (Goodfellow et al., 2016). Esto se consigue mediante la aplicación recursiva de la regla de la cadena a través de un grafo computacional, en virtud de que la red es una composición de funciones elementales cuyas derivadas analíticas son conocidas.

El ingrediente clave en los modelos PINN es la construcción de la función de pérdida:

$$\mathcal{L}(\boldsymbol{\theta}) = \int_{\Omega} r_{\text{PDE}}^2 d\Omega + \omega_{\text{BC}} \int_{\partial\Omega} r_{\text{BC}}^2 d\Gamma + \omega_{\text{IC}} \int_{\Omega} r_{\text{IC}}^2 d\Omega, \quad (3)$$

donde ω_{BC} y ω_{IC} son factores de ponderación para asignar pesos relativos a los términos que acompañan. En la práctica, las integrales deben ser reemplazadas por reglas de cuadratura adecuadas (De Ryck et al., 2023). Esto implica la definición de un conjunto de puntos de colocación en los dominios correspondientes donde se deben evaluar los residuos (2) que componen la función de pérdida. Se demuestra que los errores de generalización pueden ser arbitrariamente pequeños a medida que la cantidad de puntos de colocación se hace lo suficientemente grande (Mishra y Molinaro, 2023).

En el caso planteado en este trabajo, la única diferencia respecto a la formulación descrita al comienzo de esta sección es que la red neuronal cuenta con cuatro entradas en lugar de tres con el fin de incluir el parámetro adicional $\rho_0 \in [\rho_0^{\text{inf}}, \rho_0^{\text{sup}}]$, donde ρ_0^{inf} y ρ_0^{sup} refieren a los límites inferior y superior adimensionales del rango de variación adoptado para el radio característico (ver tabla 1). De esta manera, las predicciones que se obtienen con este modelo son paramétricas en relación a tal entrada. La forma matemática de este modelo paramétrico se expresa como $f_{\theta}(\xi, \eta, \tau, \rho_0)$.

3.1. Datasets para entrenamiento

La construcción de un modelo PINN descansa sobre la definición de tres *datasets* necesarios para realizar el entrenamiento, denominados como X_{PDE} , X_{BC} , X_{IC} . Matemáticamente,

$$X_{\text{PDE}} = \{P_i = (\xi_i, \eta_i, \tau_i, \rho_{0i}) / P_i \in \Omega \times (0, \tau_f] \times [\rho_0^{\text{inf}}, \rho_0^{\text{sup}}]; i = 1, \dots, N_{\text{PDE}}\}, \quad (4a)$$

$$X_{\text{BC}} = \{Q_i = (\xi_i, \eta_i, \tau_i, \rho_{0i}) / Q_i \in \partial\Omega \times (0, \tau_f] \times [\rho_0^{\text{inf}}, \rho_0^{\text{sup}}]; i = 1, \dots, N_{\text{BC}}\}, \quad (4b)$$

$$X_{\text{IC}} = \{R_i = (\xi_i, \eta_i, 0, \rho_{0i}) / R_i \in \Omega \cup \partial\Omega \times [\rho_0^{\text{inf}}, \rho_0^{\text{sup}}]; i = 1, \dots, N_{\text{IC}}\}. \quad (4c)$$

Debido a que la presencia de la fuente térmica genera gradientes térmicos elevados en las inmediaciones de su posición, resulta necesario considerar una estrategia de muestreo particular en el conjunto X_{PDE} similar a la propuesta por Hosseini et al. (2023). El objetivo es generar una mayor densidad de puntos en la zona térmicamente afectada por la fuente móvil. El descrito es un procedimiento que se asemeja a los empleados en la implementación de métodos numéricos clásicos, como por ejemplo las estrategias de refinamiento y mallado adaptativo en el MEF o la estrategia de superposición de nodos en el método Galerkin libre de elementos (Álvarez Hostos et al., 2022).

La cantidad de puntos de colocación utilizados en el conjunto de entrenamiento para el interior del dominio, las fronteras y la condición inicial es $N_{\text{PDE}} = 40\,000$, $N_{\text{BC}} = 10\,000$ (2500 por cada lado) y $N_{\text{IC}} = 10\,000$, respectivamente. La figura 1a muestra en simultáneo las coordenadas dimensionales (x, y) de los puntos de colocación pertenecientes a los conjuntos definidos en el interior del dominio y sobre la frontera, X_{PDE} y X_{BC} respectivamente. La figura 1b muestra lo propio, pero respecto al conjunto X_{IC} . Se aprecia claramente la mayor densidad de puntos del conjunto X_{PDE} en la región por donde pasa la fuente térmica. Por otro lado, si bien la distribución de puntos perteneciente al conjunto X_{IC} es más homogénea en el dominio, también incluye una región de mayor densidad centrada en la posición inicial de la fuente térmica, el punto $(0, 0)$.

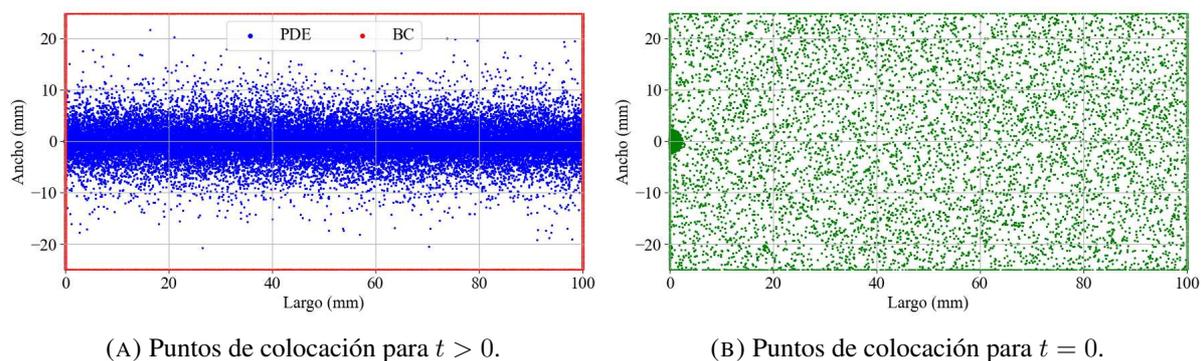


FIGURA 1: Subconjuntos de datos para entrenamiento.

3.2. Arquitectura e hiperparámetros

La red neuronal tipo FC-FFNN cuenta con cuatro entradas: la posición ξ, η sobre el dominio, el tiempo τ y el radio característico ρ_0 ; y una salida que corresponde a la temperatura \hat{u} . La figura 2 muestra un esquema del modelo PINN adoptado en el que se aprecia la mencionada red, las funciones de pérdida expresadas en términos del error cuadrático medio de los residuos (2) y el lazo de retroalimentación. Este último permite actualizar iterativamente los parámetros θ del modelo mediante el algoritmo de aprendizaje basado en el gradiente $\nabla_{\theta} \mathcal{L}(\theta)$ de la función de pérdida global, hasta satisfacer una tolerancia deseada ϵ .

Se ensayaron 5 experimentos cuyos modelos asociados se denotan como M1, M2, M3, M4 y M5. La tabla 2 muestra los detalles de los hiperparámetros adoptados en cada caso. Los acrónimos NCO y NNxC refieren al número de capas ocultas y al número de neuronas por capa, respectivamente. En todos los casos, se utilizaron funciones de activación tipo tangente hiperbólica en todas las unidades de las capas ocultas. Asimismo, el proceso de minimización de la función de pérdida se lleva a cabo con el algoritmo iterativo cuasi-Newton de memoria limitada

L-BFGS (Nocedal y Wright, 1999). En particular, en M2 se utiliza una estrategia secuencial que combina el optimizador conocido como Adam (Goodfellow et al., 2016) con L-BFGS. lr es la tasa de aprendizaje asignada al algoritmo de optimización que controla el paso de avance en la dirección del gradiente computado, y el número de épocas refiere a la cantidad máxima de iteraciones en el bucle de entrenamiento.

TABLA 2: Hiperparámetros definidos en cada experimento.

Modelo	Arquitectura		Tipo	Optimizador	
	NCO	NNxC		lr	Épocas
M1	5	50	BFGS	1,0	15 000
M2	5	50	Adam+BFGS	$5e-4 1,0$	$3\,500 + 10\,000$
M3	5	100	BFGS	1,0	15 000
M4	6	150	BFGS	1,0	15 000
M5	6	200	BFGS	1,0	15 000

Por último, cabe mencionar que el comportamiento físico esperado como consecuencia del aporte de calor por parte de la fuente térmica es un incremento de la temperatura en el dominio desde el valor dado u_0 en la condición inicial. Para asegurar que las predicciones del modelo de PINN sean consistentes con esta condición, se utiliza una función de activación tipo *softplus* en la salida de la neurona correspondiente a la última capa del modelo. Tal función, cuya expresión es $f(x) = \beta^{-1} \log(1 + e^{\beta x})$, representa una aproximación a la función rampa cuyo grado de suavidad depende del parámetro β .

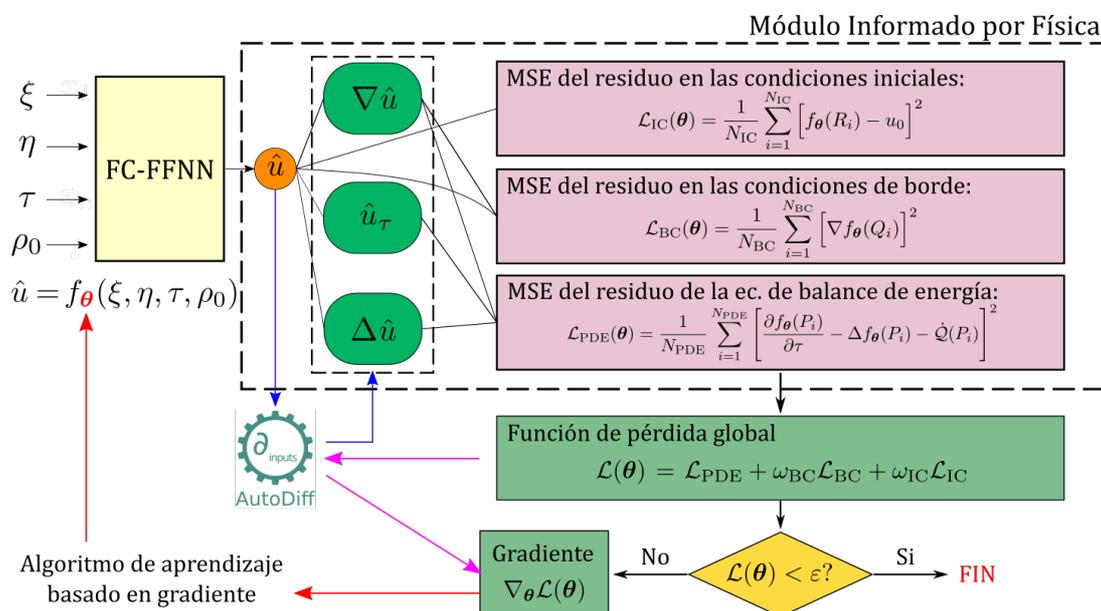


FIGURA 2: Arquitectura del modelo PINN.

3.3. Implementación computacional

Todas las rutinas desarrolladas para los distintos casos fueron escritas en Python. Se utilizó principalmente la biblioteca PyTorch para permitir el desarrollo de las rutinas asociadas con los modelos de aprendizaje profundo, en combinación con PyTorch Lightning. Los códigos fuente

de la implementación pueden encontrarse en el repositorio público disponible en internet bajo la dirección https://github.com/btourn/PINN_HeatConduction2D.

4. RESULTADOS

Las gráficas de la figura 3 muestran la evolución de la función de pérdida global y la influencia de sus distintos términos, para cada uno de los modelos definidos. Se reportan los resultados solamente hasta las 10 000 épocas de entrenamiento.

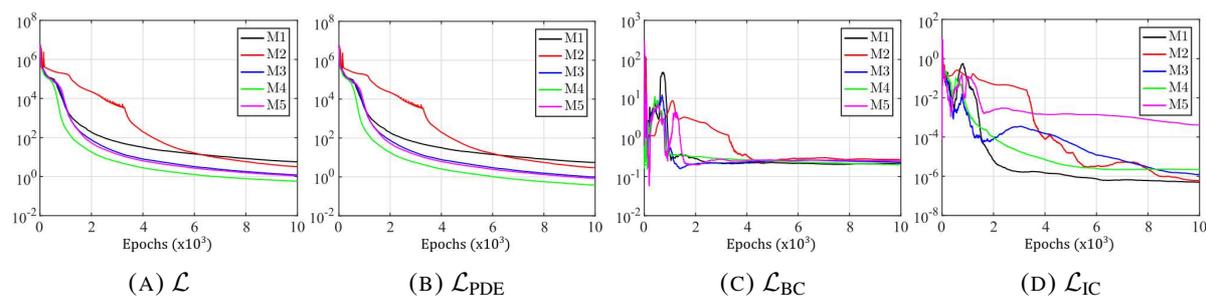


FIGURA 3: Evolución de las funciones de pérdida.

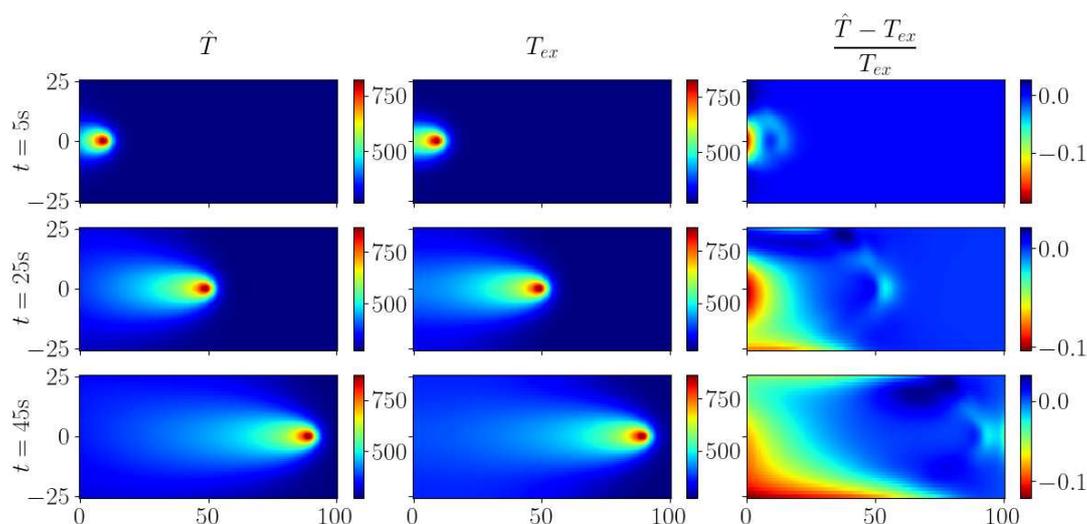
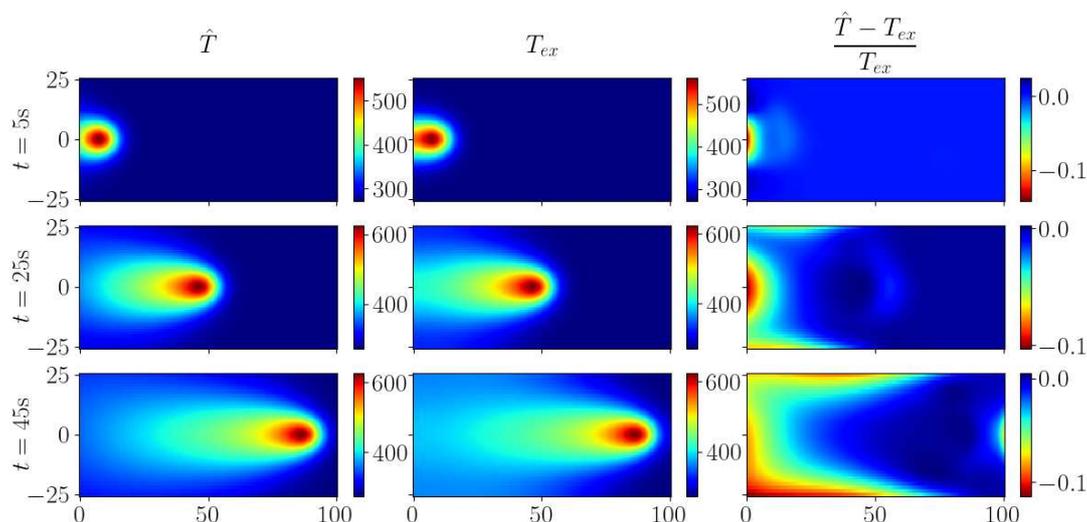
El modelo M4 es el que presenta el mejor desempeño en relación con el valor de la función de pérdida global \mathcal{L} . Por ello, es el que se emplea a continuación para mostrar y analizar las predicciones proporcionadas en diversos casos. Cabe resaltar que, una vez entrenado hasta un umbral de error aceptable, el modelo es capaz de brindar inferencias de temperatura \hat{u} para cualquier punto (ξ, η) del dominio considerado, instante de tiempo τ y valor del radio característico ρ_0 , a costo computacional prácticamente nulo. Tales inferencias tienen sentido solo si los valores de los integrantes de la tétrada $(\xi, \eta, \tau, \rho_0)$ se encuentran cada uno de ellos dentro del rango de valores para los que fue entrenado el modelo.

Las figuras 4, 5 y 6 muestran los campos de temperatura \hat{T} obtenidos por el modelo M4 sobre un conjunto de predicción que consiste en una grilla suficientemente fina de puntos en los que se computa las inferencias. Tales predicciones se comparan con soluciones obtenidas por el MEF y consideradas como *ground-truth* (denotadas como T_{ex}) sobre una malla de cuadriláteros definida a partir de la grilla de puntos previamente mencionada. Por último, se muestra la diferencia proporcional punto a punto entre ambas distribuciones.

5. DISCUSIÓN

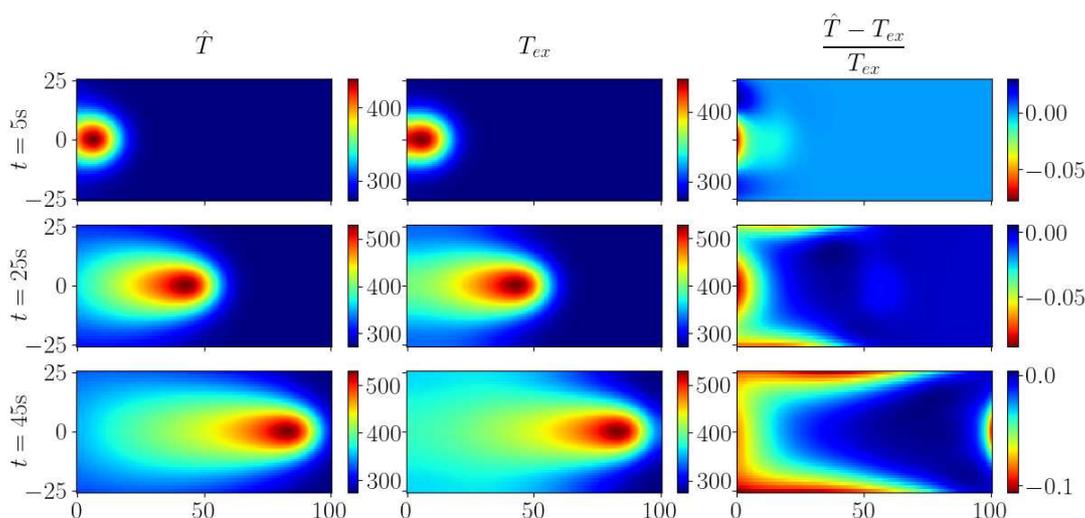
Los modelos entrenados desde el inicio con el algoritmo L-BFGS presentaron una convergencia significativamente más rápida y estable que el modelo M2 entrenado hasta la época 3500 con *Adam*. Esta observación se encuentra en línea con los resultados de Hosseini et al. (2023), que emplearon exclusivamente L-BFGS en sus modelos. Resulta evidente que el precondicionamiento de la dirección de descenso con información de la curvatura de la función de pérdida proporcionada por tal método favorece el proceso de entrenamiento.

En todos los casos, la tasa de convergencia se reduce notoriamente a medida que progresa el entrenamiento al punto que, a partir de las 10 000, no se registran prácticamente cambios favorables. La componente \mathcal{L}_{PDE} presenta la máxima contribución a la pérdida total (ver figura 3b). Por su parte, la satisfacción de la condición inicial, cuantificada por \mathcal{L}_{IC} , es más que adecuada (ver figura 3d). La reducción en la tasa de convergencia se debe posiblemente a la influencia que posee el término \mathcal{L}_{BC} sobre la pérdida global (ver figura 3c). \mathcal{L}_{BC} permanece estancada en

FIGURA 4: $r_0 = 2$ mm.FIGURA 5: $r_0 = 5$ mm.

un valor relativamente alto alrededor de 0,2 en todos los modelos. Esto quiere decir que las condiciones de borde Neumann sobre los contornos del dominio no se satisfacen adecuadamente, y que en consecuencia se ve comprometido el desempeño del modelo. Esta observación es habitual en el modelado por PINN y constituye uno de los principales modos de falla (Krishnapriyan et al., 2021; Wang et al., 2021). Sin embargo, las estrategias específicas para mitigar tales efectos indeseables se encuentran fuera del alcance de este trabajo.

De las figuras 4, 5 y 6 se aprecia que la temperatura en la zona delante de la posición de la fuente térmica móvil queda prácticamente inalterada respecto al estado inicial, mientras que la zona posterior se encuentra claramente afectada por su paso. Asimismo, se observa que las predicciones son aceptablemente simétricas respecto al eje de simetría longitudinal. Vale destacar que esta última característica fue alcanzada sin imponer ninguna condición particular sobre el modelo. En todos los casos se aprecia un ajuste óptimo en prácticamente todo el dominio. Sin embargo, la zona próxima a la frontera en $x = 0$ es la que presenta las mayores discrepancias, como así también aquellas porciones de las fronteras en $y = -25$ y $y = 25$ afectadas térmica-

FIGURA 6: $r_0 = 8$ mm.

mente luego del avance de la fuente. Si bien tales discrepancias no superan en ningún caso el 12 % de error respecto a la solución exacta, no puede afirmarse que la predicción del modelo sea adecuada en dichas regiones. Es importante resaltar que no ha sido posible mejorar el ajuste sobre tales regiones mediante técnicas tradicionales. En consecuencia, resulta evidente la necesidad de comprender el mecanismo que genera tales defectos en presencia de condiciones de borde tipo Neumann, de manera de elaborar una estrategia que provea un ajuste adecuado del término de pérdida asociado.

5.1. Limitaciones del modelo

Pese a la gran cantidad de experimentos preliminares en los que se modificaron parámetros de la arquitectura, optimizadores, cantidad y distribución de puntos de colocación, entre otras modificaciones, no fue posible construir un modelo PINN que permita resolver paramétricamente las ecuaciones (1) cuando los *datasets* (4) contemplaron rangos de variación del radio característico $\rho_0 \in [\rho_0^{\text{inf}}, \rho_0^{\text{sup}}]$ correspondientes con $r_0^{\text{inf}} < r_0^{\text{sup}} < 1,0$ mm.

Si bien este comportamiento no ha podido ser vinculado a una explicación cuantitativa satisfactoria, es posible asociarlo con algunos aspectos de la naturaleza física del problema en estudio. Las figuras reportadas por [Álvarez Hostos et al. \(2022\)](#) para $r_0 = 0,2$ mm y $r_0 = 0,02$ mm muestran que la solución presenta prácticamente una singularidad en el campo de temperaturas en la ubicación instantánea de la fuente. Esto indica que los gradientes de temperatura en dicha zona son excesivamente pronunciados. Es evidente que esta característica ocasiona dificultades tanto en el modelo PINN propiamente dicho (a través del cómputo de las derivadas de la temperatura respecto a las entradas mediante el mecanismo de autodiferenciación) como en el proceso de optimización, dado que la función de pérdida también se ve afectada de una manera que no resulta fácilmente predecible.

Si bien el estudio de los motivos que originan el comportamiento descrito, así como las medidas necesarias para mitigarlo adecuadamente, quedan fuera de los alcances de este trabajo, la obtención de modelos de PINN para casos más generales en presencia de gradientes muy pronunciados constituye otro modo de falla típico que ha tratado de ser abordado con mayor o menor éxito en la literatura ([Krishnapriyan et al., 2021](#); [Wang et al., 2021](#); [Deguchi y Asai, 2023](#); [McClenny y Braga-Neto, 2023](#)). Los autores reconocen esta como una oportunidad para

orientar esfuerzos futuros con el fin de realizar contribuciones significativas al modelado via PINN.

6. CONCLUSIONES

El caso de estudio aquí presentado evidenció la capacidad de los modelos PINN entrenados de manera no supervisada para ofrecer soluciones generales paramétricas al problema bidimensional de conducción de calor con fuente térmica móvil, en relación al radio característico ρ_0 . Sin embargo, las predicciones sufrieron discrepancias de alrededor de un 10 % respecto a la solución de referencia en zonas próximas a las fronteras físicas detrás de la posición de la fuente, asociadas al incumplimiento de las condiciones de borde impuestas en tal región, lo que sugiere la necesidad de tratar específicamente este problema en futuros trabajos. Asimismo, el modelo presentó limitaciones en presencia de gradientes de temperatura pronunciados provocados por valores de $r_0 < 1$ mm. Se deja también para trabajos futuros el análisis de los motivos que ocasionan el desempeño sub-óptimo bajo tales condiciones y sus medidas paliativas. Un factor clave a destacar es la estrategia de muestreo de los puntos de colocación, que se encuentra íntimamente relacionada con la física del problema. Por último, en relación con el modelado y simulación de aplicaciones tecnológicas realistas como procesos de soldadura vía PINN, existen varias líneas de investigación posibles como la inclusión de propiedades termofísicas dependientes de la temperatura o la posibilidad de contemplar un movimiento arbitrario de la fuente térmica con velocidad variable, por mencionar algunos.

REFERENCIAS

- Cai S., Wang Z., Wang S., Perdikaris P., y Karniadakis G.E. Physics-informed neural networks for heat transfer problems. *Journal of Heat Transfer*, 143(6):060801, 2021.
- Cuomo S., Di Cola V.S., Giampaolo F., Rozza G., Raissi M., y Piccialli F. Scientific machine learning through physics-informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.
- De Ryck T., Bonnet F., Mishra S., y de Bézenac E. An operator preconditioning perspective on training in physics-informed machine learning. *arXiv preprint arXiv:2310.05801*, 2023.
- Deguchi S. y Asai M. Dynamic & norm-based weights to normalize imbalance in back-propagated gradients of physics-informed neural networks. *Journal of Physics Communications*, 7(7):075005, 2023.
- Goodfellow I., Bengio Y., y Courville A. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- Hosseini E., Ghanbari P.G., Müller O., Molinaro R., y Mishra S. Single-track thermal analysis of laser powder bed fusion process: Parametric solution through physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, 410:116019, 2023.
- Karniadakis G.E., Kevrekidis I.G., Lu L., Perdikaris P., Wang S., y Yang L. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.
- Krishnapriyan A., Gholami A., Zhe S., Kirby R., y Mahoney M.W. Characterizing possible failure modes in physics-informed neural networks. *Advances in neural information processing systems*, 34:26548–26560, 2021.
- McClenny L.D. y Braga-Neto U.M. Self-adaptive physics-informed neural networks. *Journal of Computational Physics*, 474:111722, 2023.
- Mishra S. y Molinaro R. Estimates on the generalization error of physics-informed neural networks for approximating pdes. *IMA Journal of Numerical Analysis*, 43(1):1–43, 2023.
- Nocedal J. y Wright S.J. *Numerical optimization*. Springer, 1999.
- Rabczuk T. y Bathe K.J. *Machine Learning in Modeling and Simulation: Methods and Applications*. Springer, 2023. Part of the book series: Computational Methods in Engineering & the Sciences (CMES).
- Raissi M., Perdikaris P., y Karniadakis G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.
- Wang S., Teng Y., y Perdikaris P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5):A3055–A3081, 2021.

- Xie J., Chai Z., Xu L., Ren X., Liu S., y Chen X. 3d temperature field prediction in direct energy deposition of metals using physics informed neural network. *The International Journal of Advanced Manufacturing Technology*, 119(5):3449–3468, 2022.
- Zhu Q., Liu Z., y Yan J. Machine learning for metal additive manufacturing: predicting temperature and melt pool fluid dynamics using physics-informed neural networks. *Computational Mechanics*, 67:619–635, 2021.
- Álvarez Hostos J.C., Storti B., Tourn B.A., y Fachinotti V.D. Solving heat conduction problems with a moving heat source in arc welding processes via an overlapping nodes scheme based on the improved element-free galerkin method. *International Journal of Heat and Mass Transfer*, 192:122940, 2022.