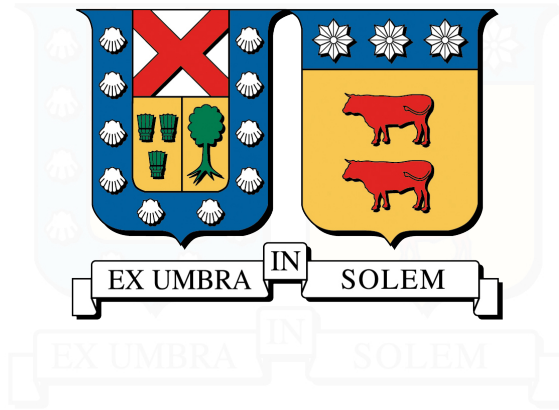


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS
VALPARAÍSO - CHILE



**ESTIMACIÓN DE FACTORES DE FORMA MEDIANTE APRENDIZAJE
SUPERVISADO: UN ENFOQUE BASADO EN MACHINE LEARNING Y
DEEP LEARNING**

MAURO MACIEL MERCADO FLEITER

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA INDUSTRIAL
PROFESOR GUÍA : PROF. ANDRÉS FUENTES.
PROFESOR CORREFERENTE : PROF. RODRIGO DEMARCO BULL.
PROFESOR CORREFERENTE EXTERNO : PROF. FELIPE ESCUDERO B.

MARZO 2026



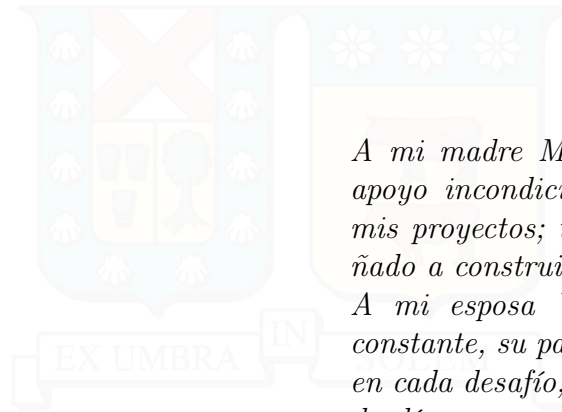
Índice de Contenidos

Nomenclatura	v
1. Introducción y Motivación	1
1.1. Introducción	1
1.2. Motivación	2
1.2.1. Evolución de incendios forestales	3
1.2.2. Incendios forestales en Chile	3
1.2.3. Incendios por Región	4
1.2.4. Zonas de peligro.	6
1.2.5. Alcance	9
1.2.6. Objetivo General	10
1.2.7. Objetivos Específicos	10
2. Estado del arte	11
2.1. Cálculo clásico del factor de forma	11
2.2. Métodos numéricos y aproximaciones estocásticas	11
2.3. Aprendizaje automático en ingeniería térmica	12
2.4. Brecha identificada	13
3. Marco Teórico	15
3.1. Mecanismos de transferencia de calor	15
3.1.1. Mecanismos de propagación del fuego	15
3.2. Aprendizaje automático: conceptos generales	17
3.2.1. Paradigmas de aprendizaje	17
3.2.2. Modelos predictivos utilizados en este trabajo	18
3.2.3. Consideraciones finales	20
3.3. Técnicas de optimización en machine learning	21
3.4. Fundamentos Matemáticos	23
3.4.1. Campos escalares y derivadas	23
3.4.2. Gradiente	23
3.4.3. Matriz Jacobiana	24
3.4.4. Matriz Hessiana	24
3.4.5. Descomposición en valores singulares	25
3.4.6. Optimización en Machine Learning	25
3.4.7. Descenso de gradiente	26

3.4.8.	Descenso de gradiente estocástico	26
3.4.9.	Método de Newton	27
3.4.10.	Optimizador Adam	27
3.4.11.	Curvas parametrizadas	28
3.4.12.	Integrales de línea	28
3.4.13.	Teorema de Stokes	28
3.5.	Factores de Forma	29
3.6.	Integrales de Línea y Superficie	29
3.7.	Factor de forma por integrales de superficies.	29
3.7.1.	Vectores Normales a las Superficies	29
3.7.2.	Expresión Final del Factor de Forma	30
3.8.	Factor de forma por Integrales de Línea (Stokes)	30
3.8.1.	Parametrización de los contornos	30
3.8.2.	Norma euclidiana de ambos vectores.	30
3.8.3.	Producto entre vectores tangentes	31
3.8.4.	Descomposición por segmentos.	31
3.8.5.	Expresión final del factor de forma	31
3.9.	Modelos de Machine Learning y Deep Learning	31
3.9.1.	Regresión lineal múltiple con interacciones	32
3.9.2.	Formulación del problema	33
3.10.	Estimación de parámetros vía mínimos cuadrados	33
3.10.1.	Crecimiento combinatorial en modelos polinomiales	35
3.11.	Support Vector Regression (SVR)	35
3.11.1.	Formulación del problema	36
3.11.2.	Función de pérdida ε -insensible	36
3.11.3.	Problema primal	37
3.11.4.	Formulación dual	37
3.11.5.	Extensión no lineal y kernel trick	38
3.11.6.	Ventajas y propiedades	38
3.12.	Perceptrón Multicapa (MLP)	38
3.12.1.	Propagación hacia adelante (Forward Propagation)	39
3.12.2.	Función de pérdida	39
3.12.3.	Funciones de activación en redes neuronales	40
3.12.4.	Retropropagación (Backpropagation)	41
3.12.5.	Actualización de parámetros	41
3.12.6.	Capacidad de aproximación	41
4.	Metodología	43
4.1.	Modelamiento físico de factores de forma	43
4.2.	Casos de estudio	44
4.3.	Generación del conjunto de datos	45
4.4.	Preprocesamiento de datos	45
4.5.	Modelos de Machine Learning	46
4.5.1.	Optimización de hiperparámetros	46
4.5.2.	Evaluación del desempeño	47

4.5.3. Comparación con el modelo físico	47
5. Casos de estudio	49
5.1. Configuraciones geométricas consideradas	49
5.1.1. Placa rectangular y disco	49
5.1.2. Calefactor cilindrico y elemento diferencial	51
5.1.3. Cono calorímetro y elemento diferencial.	52
5.2. Modelos sustitutos	54
5.3. Caso 1: Placa rectangular y disco	54
5.3.1. Generación del dataset.	55
5.3.2. Construcción y entrenamiento de los modelos	55
5.4. Caso 2: Calentador cilíndrico y elemento diferencial	56
5.4.1. Generación del dataset.	56
5.4.2. Construcción y entrenamiento de los modelos	56
5.5. Caso 3: Cono calorímetro y elemento diferencial	57
5.5.1. Generación del conjunto de datos	57
5.5.2. Construcción y entrenamiento del modelo	58
6. Resultados	59
6.1. Caso 1: Placa rectangular y disco	59
6.1.1. Comparación de perfiles del factor de forma	59
6.1.2. Comparación del desempeño de los modelos predictivos	61
6.1.3. Comparación métricas de validación y prueba	62
6.1.4. Análisis de métricas de validación y prueba	62
6.1.5. Análisis de desempeño	64
6.1.6. Contraste entre valores reales y predicción.	65
6.1.7. Análisis de eficiencia computacional	65
6.2. Caso 2: Calentador cilíndrico y elemento diferencial	66
6.2.1. Comparación métricas de validación y prueba	66
6.2.2. Análisis métricas de validación y prueba	67
6.2.3. Análisis de desempeño	68
6.2.4. Comparación de perfiles del factor de forma	69
6.2.5. Análisis de eficiencia computacional	70
6.3. Caso 3: Cono calorímetro y elemento diferencial	70
6.3.1. Comparación métricas de validación y prueba	71
6.3.2. Análisis de desempeño	72
6.3.3. Contraste entre valores reales y predicción	73
6.3.4. Análisis de eficiencia computacional	74
7. Conclusiones	75
Bibliografía	77





A mi madre Marcia, por su amor y apoyo incondicional en cada uno de mis proyectos; tu fuerza me ha enseñado a construir la mía.

A mi esposa Valery, por su apoyo constante, su paciencia y su compañía en cada desafío; a tu lado aprendo cada día a ser una mejor persona.

Y a mi hija Amara, cuya alegría y ternura son la mayor inspiración de mi vida; en ti encuentro siempre el recordatorio de lo que realmente es importante.

AGRADECIMIENTOS

Quisiera expresar mi más sincero agradecimiento a las personas que me han acompañado y apoyado a lo largo de este proceso.

En primer lugar, quiero agradecer a mi esposa, Valery, por su comprensión, paciencia y apoyo durante todo este camino. Gracias por estar a mi lado en cada etapa, por motivarme a seguir adelante y por compartir conmigo los desafíos y logros de este proceso. Tu compañía, amor y presencia son un regalo invaluable para mí.

En segundo lugar, agradezco profundamente a mi madre, Marcia, por su amor incondicional, su apoyo constante y por haber creído siempre en mí. Su ejemplo de esfuerzo y perseverancia ha sido una inspiración permanente en mi vida. Verte luchar incansablemente día tras día me ha enseñado a enfrentar la vida con la misma determinación.

A mi hija, Amara, quien con su alegría y ternura ha sido una fuente constante de motivación. Su presencia ilumina cada uno de mis días y me recuerda siempre el verdadero sentido de seguir avanzando. Todo mi trabajo es también por ti.

Quisiera agradecer a mi amigo Pablo Pinto, de quien he tenido la oportunidad de aprender mucho a lo largo de estos años. Su amistad, sus conversaciones y sus conocimientos han sido una valiosa influencia en mi formación personal y académica.

También quiero agradecer a Ignacio Verdugo, un amigo que me ayudó a lo largo de este trabajo y de quien también aprendí mucho.

Por último, pero no menos importante, agradezco a todos los profesores que, de una u otra manera, han contribuido a mi formación a lo largo de este camino. Gracias por todos los conocimientos compartidos.

RESUMEN EJECUTIVO

Los factores de forma (VF) son fundamentales para modelar con precisión la transferencia de calor por radiación entre superficies, ya que gobiernan la distribución de los flujos radiativos y desempeñan un papel clave en la gestión térmica y en el análisis de seguridad frente a incendios. En este trabajo se presenta una metodología numérica para la estimación de factores de forma y la resolución de problemas inversos asociados mediante algoritmos de machine learning (ML) y deep learning (DL). El marco de trabajo propuesto, basado en datos, se compara con valores de referencia de factores de forma calculados mediante métodos de integrales de contorno derivados del teorema de Stokes. Se analizan tres configuraciones geométricas representativas con el fin de caracterizar el comportamiento espacial y geométrico del intercambio radiativo y evaluar la capacidad predictiva de los modelos sustitutos.

Entre los enfoques evaluados, el Perceptrón Multicapa (MLP) alcanzó de manera consistente la mayor precisión, capturando exitosamente las dependencias geométricas no lineales y reproduciendo las distribuciones espaciales del factor de forma con una concordancia casi perfecta respecto de las referencias analíticas. Los modelos entrenados ofrecen mejoras de varios órdenes de magnitud en la eficiencia computacional en comparación con la integración numérica clásica, lo que permite evaluaciones en tiempo real adecuadas para estudios paramétricos, tareas de optimización y problemas de diseño inverso.

Los resultados demuestran que los modelos sustitutos basados en ML y DL constituyen una alternativa robusta y escalable a las formulaciones tradicionales basadas en integrales, proporcionando estimaciones de factores de forma precisas, rápidas y físicamente consistentes en una amplia variedad de escenarios de transferencia radiativa.



Nomenclatura

Símbolos latinos

F	Factor de forma (—)
A	Área de superficie (m ²)
h	Parámetro espacial horizontal (mm)
k	Parámetro espacial vertical (mm)
L	Longitud característica (mm)
N	Número de muestras (—)

Símbolos griegos

θ	Coordenada angular (rad)
ϕ	Ángulo azimutal (rad)
ε	Métrica de error (—)
η	Tasa de aprendizaje (—)
σ	Desviación estándar (—)

Acrónimos

ML	Aprendizaje automático
DL	Aprendizaje profundo
MLP	Perceptrón multicapa
VF	Factor de forma
MSE	Error cuadrático medio
RMSE	Raíz del error cuadrático medio
APE	Error porcentual absoluto
MAPE	Error porcentual absoluto medio

1 | Introducción y Motivación

1.1. Introducción

Desde el inicio de la conciencia humana, el fuego ha sido reconocido y venerado. Este fenómeno, tan destructivo como fascinante, ha sido tanto aliado como adversario de la humanidad a lo largo de su historia. Esta manifestación volátil de la naturaleza, resultante de una reacción química exotérmica conocida comúnmente como combustión, se caracteriza por la liberación de luz y calor, entrelazándose con la existencia humana en una dualidad de creación y devastación. A través de los siglos, el fuego ha sido simultáneamente herramienta y amenaza, proporcionando calor, permitiendo la cocción de alimentos y siendo esencial en diversos procesos industriales, mientras que, al mismo tiempo, se erige como una potente fuerza capaz de destruir ecosistemas y construcciones humanas.

En el planeta Tierra, el registro fósil del fuego aparece en el período Ordovícico medio, hace aproximadamente 470 millones de años, asociado al establecimiento de la flora terrestre, lo que propició la acumulación de oxígeno en la atmósfera hasta niveles no alcanzados previamente. Cuando el porcentaje de oxígeno en la atmósfera superó aproximadamente el 13 % de concentración, se generaron las condiciones necesarias para la aparición de incendios no deseados [1]. Entre estos emergen los incendios forestales, eventos que se propagan sin control en terrenos rurales, afectando combustibles vegetales, flora y fauna. Estos se distinguen de otros tipos de incendios por su amplia extensión, la velocidad con la que pueden propagarse desde su lugar de origen, su potencial para cambiar de dirección inesperadamente y su capacidad para superar diversos obstáculos naturales.

Sin duda, uno de los fenómenos más catastróficos asociados al fuego son los incendios forestales. Estos eventos ocurren en diferentes partes del mundo y generan consecuencias significativas para los ecosistemas, la vida silvestre y las comunidades humanas. Los incendios pueden surgir por diversas razones, ya sean accidentales o intencionales, y pueden tener efectos perjudiciales en múltiples dimensiones ambientales y sociales. En general, ocurren cuando el fuego se propaga sin control en áreas boscosas o en zonas de vegetación densa.

Una de las principales causas de los incendios forestales es la combinación de condiciones climáticas adversas, tales como altas temperaturas, baja humedad relativa y vientos fuertes. Dichas condiciones crean un entorno propicio para que el fuego se propague rápidamente y resulte difícil de controlar. No obstante, también existen factores humanos que contribuyen significativamente a la aparición de estos

fenómenos. Entre los más habituales se encuentran actividades negligentes, como el uso irresponsable del fuego en áreas forestales, la quema agrícola descontrolada o el abandono de cigarrillos encendidos, entre otras acciones atribuibles a la actividad humana.

Además de lo mencionado anteriormente, el aumento de la urbanización en zonas aledañas a los bosques ha incrementado el riesgo de incendios forestales debido a la mayor presencia de fuentes de ignición de origen antrópico. Este tipo de incendios se conoce como incendios de interfaz urbano-forestal o WUI (por sus siglas en inglés, *Wildland-Urban Interface*) [1, 2].

1.2. Motivación

El fuego ha despertado el interés de la humanidad desde tiempos remotos debido a su enorme capacidad de transformación y a su impacto en los sistemas naturales y sociales. Si bien ha sido fundamental para el desarrollo de numerosas actividades humanas, también representa una amenaza cuando ocurre de manera descontrolada. En este contexto, comprender los mecanismos que gobiernan los procesos de combustión y propagación del fuego resulta fundamental para reducir sus efectos negativos sobre los ecosistemas y las comunidades humanas.

Más allá de su manifestación visible, el fuego constituye un fenómeno físico y químico complejo que involucra múltiples procesos de transferencia de calor, reacciones químicas y dinámicas de fluidos. El estudio de estos procesos permite comprender de mejor manera el comportamiento del fuego en distintos entornos y desarrollar estrategias más efectivas para su prevención, detección y control.

En particular, el análisis científico del fuego y de los procesos de combustión abre oportunidades para el desarrollo de herramientas tecnológicas orientadas al monitoreo y modelación de incendios. La comprensión de los principios físicos y químicos involucrados en la combustión permite diseñar sistemas más eficientes de detección y extinción, así como mejorar las estrategias de manejo forestal orientadas a la reducción del riesgo de incendios.

En consecuencia, el estudio de los procesos de combustión y de los fenómenos asociados al fuego constituye un área de gran relevancia científica y tecnológica, especialmente en un contexto global donde la frecuencia e intensidad de los incendios forestales ha aumentado en las últimas décadas.

1.2.1. Evolución de incendios forestales

El incremento en la frecuencia e intensidad de los incendios forestales a nivel global se ha convertido en un tema crítico dentro de las discusiones ambientales y climatológicas contemporáneas. Este fenómeno representa una realidad compleja que ha sido abordada desde diversas perspectivas. No obstante, múltiples factores han contribuido a esta tendencia creciente, entre los cuales destacan el cambio climático, las prácticas de uso del suelo y diversos factores de origen antrópico, particularmente la intervención humana, que han favorecido el aumento de la ocurrencia de estos eventos [1].

1.2.2. Incendios forestales en Chile

En el caso particular de Chile, los incendios forestales son atribuibles casi en su totalidad a causas humanas. De acuerdo con cifras oficiales entregadas por la Corporación Nacional Forestal (CONAF), aproximadamente el 99,7 % de los incendios registrados tienen origen en la acción humana, ya sea por descuidos o por negligencia [3].

Chile presenta una geografía altamente diversa, que abarca desde zonas desérticas en el norte hasta bosques templados en el sur. Durante la temporada de incendios, que generalmente ocurre en el verano austral, las altas temperaturas, la baja humedad relativa y los vientos intensos generan condiciones particularmente favorables para la ignición y propagación de incendios forestales. En muchas regiones del país, los principales combustibles corresponden a plantaciones de especies como pino y eucalipto, árboles que han sido introducidos y que presentan características altamente inflamables [4, 5].

1.2.3. Incendios por Región

A lo largo de nuestra historia reciente, con el fin de contar con antecedentes, se han recopilado datos de incendios por región que se han producido en los últimos años en nuestro país, en la tabla 1.1 se presentan las estadísticas correspondientes al período 2019-2024.

Región	N° de incendios	Superficie afectada (ha)
	2019–2024	2019–2024
Arica y Parinacota	40	209,2
Tarapacá	29	388,8
Antofagasta	31	110,8
Atacama	102	166,0
Coquimbo	336	3.128,5
Valparaíso	2.412	41.178,7
Metropolitana	2.017	28.693,5
O’Higgins	1.685	32.846,3
Maule	4.456	71.945,4
Ñuble	2.498	71.922,0
Biobío	11.540	232.866,4
La Araucanía	7.744	261.431,8
Los Ríos	608	8.193,6
Los Lagos	1.386	5.531,5
Aysén	154	1.918,6
Magallanes y de la Antártica Chilena	77	5.656,2
Total	35.115	766.187,2

Tabla 1.1: Número de incendios forestales y superficie afectada por región durante el período 2019–2024.

Fuente: Elaboración propia a partir de datos de CONAF, reportados por la Biblioteca del Congreso Nacional de Chile en el documento “Incendios Forestales. Período 2019–2024”.

La Tabla 1.1 presenta el número de incendios forestales registrados y la superficie afectada por región en Chile durante el período 2019–2024. Los datos evidencian una mayor ocurrencia de incendios en las regiones del Biobío, La Araucanía y Maule, con 11.540, 7.744 y 4.456 eventos, respectivamente. En total, se registraron 35.115 incendios forestales a nivel nacional durante el período analizado.

En términos de superficie afectada, las regiones más impactadas fueron La Araucanía y Biobío, con 261.431,8 ha y 232.866,4 ha quemadas, respectivamente. Estas cifras muestran que la severidad de los incendios no solo depende del número de eventos registrados, sino también de la magnitud territorial alcanzada por cada siniestro. En conjunto, la superficie total afectada en el país durante el período 2019–2024 fue de 766.187,2 ha.

Por otra parte, regiones como Arica y Parinacota, Tarapacá, Antofagasta y Atacama presentan una menor frecuencia de incendios y una superficie afectada considerablemente inferior en comparación con las regiones centro-sur y sur del país. No obstante, la distribución geográfica de los eventos evidencia que el riesgo de incendios forestales se encuentra presente en gran parte del territorio nacional. Este escenario resalta la importancia de analizar los patrones de ocurrencia y daño territorial, con el fin de fortalecer las estrategias de prevención, planificación y asignación de recursos para el combate de incendios forestales.

Con el propósito de complementar el análisis descriptivo, se presentan dos gráficos de barras asociados a los incendios forestales registrados en Chile durante el período 2019–2024. La primera figura representa la cantidad de incendios por región, mientras que la segunda muestra la superficie afectada por dichos eventos, medida en hectáreas.

Distribución de incendios por región

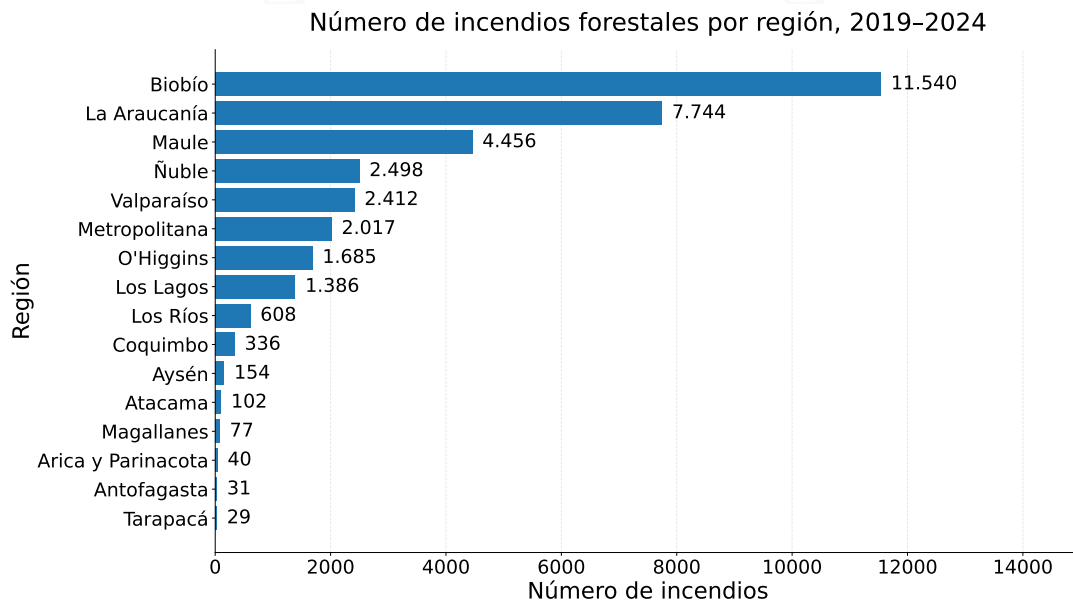


Figura 1.1: Número de incendios forestales por región durante el período 2019–2024.

Las Figuras 1.1 y 1.2 permiten visualizar gráficamente las diferencias regionales observadas en la tabla anterior. Estas representaciones facilitan la comparación entre regiones y permiten identificar con mayor claridad aquellas zonas donde la recurrencia de incendios y la superficie afectada presentan valores más elevados.

Distribución de área afectada por región

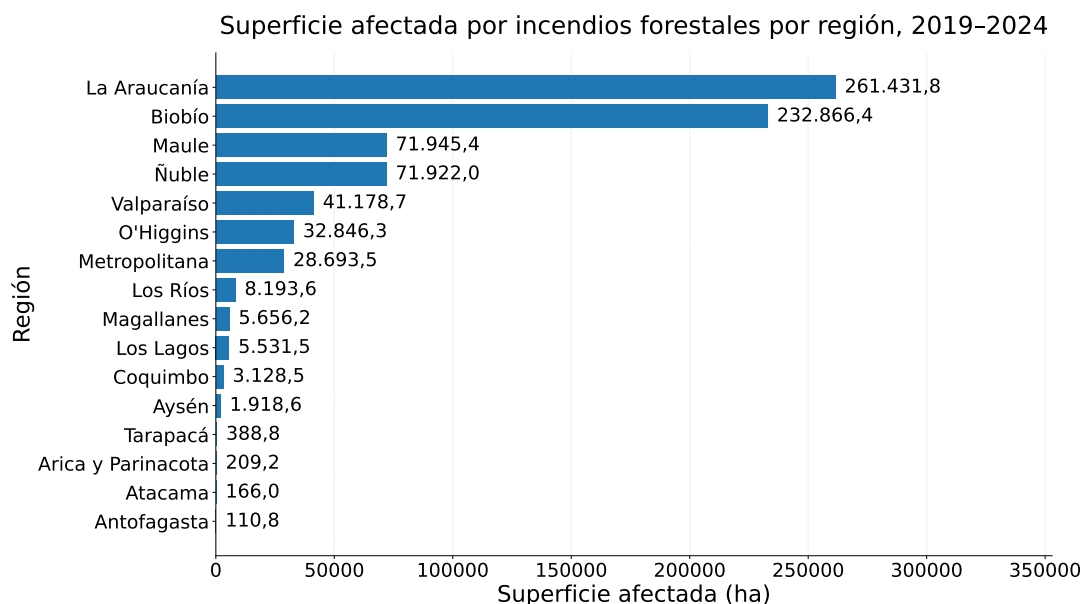


Figura 1.2: Superficie quemada por región durante el período 2019–2024.

La Figura 1.2 complementa el análisis anterior al enfocarse en la magnitud territorial de los incendios forestales. A diferencia del número de eventos, esta variable permite observar el impacto espacial de los siniestros, evidenciando que algunas regiones concentran una proporción considerable de la superficie total afectada durante el período 2019–2024.

1.2.4. Zonas de peligro.

Las zonas de peligro de incendios forestales en Chile no se distribuyen de manera homogénea a lo largo del territorio nacional, sino que tienden a concentrarse principalmente en la zona centro-sur y sur del país. Esta concentración se relaciona con una combinación de factores climáticos, geográficos, vegetacionales y antrópicos, tales como la presencia de cobertura vegetal combustible, períodos prolongados de sequía, altas temperaturas, vientos intensos, expansión urbana hacia zonas de interfaz urbano-rural y actividades humanas que pueden actuar como fuentes de ignición.

A partir de los antecedentes presentados anteriormente, se observa que las regiones del Biobío, La Araucanía, Maule, Ñuble y Valparaíso presentan una mayor recurrencia de incendios forestales durante el período analizado. En particular, Biobío y La Araucanía destacan tanto por el número de eventos registrados como por la superficie afectada, lo que permite identificarlas como zonas de alta prioridad para el análisis del riesgo. Esta situación evidencia que el peligro no debe evaluarse únicamente en función de la cantidad de incendios, sino también considerando la magnitud territorial del daño ocasionado.

En este contexto, la identificación de zonas de peligro permite reconocer áreas donde existe una mayor probabilidad de ocurrencia y propagación de incendios fo-

restales. Esta información resulta relevante para apoyar la planificación territorial, la asignación de recursos preventivos, el diseño de estrategias de monitoreo y la toma de decisiones orientadas a reducir el impacto de estos eventos. Por ello, el análisis espacial de los incendios constituye una herramienta fundamental para comprender la vulnerabilidad del territorio chileno frente a este tipo de amenazas.

A continuación, se presentan dos mapas de calor correspondientes al territorio nacional. El primer recuadro muestra la distribución espacial asociada a la cantidad de incendios registrados, mientras que el segundo representa la superficie quemada. Visualmente, ambos mapas evidencian patrones espaciales similares, lo que sugiere una posible asociación entre la frecuencia de incendios y la magnitud de la superficie afectada.

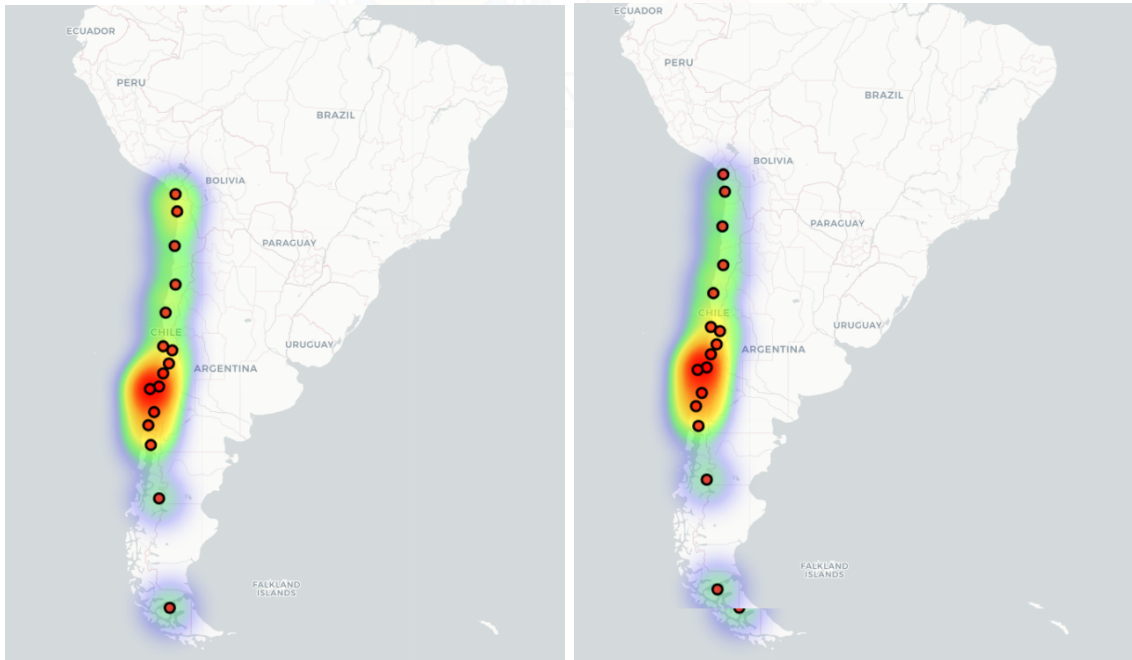


Figura 1.3: Mapas de calor asociados a incendios forestales en Chile. El panel izquierdo muestra la distribución espacial de la cantidad de incendios registrados, mientras que el panel derecho representa la distribución espacial de la superficie quemada. La similitud entre ambos patrones sugiere una posible asociación espacial entre la frecuencia de incendios y la magnitud del área afectada.

Los patrones espaciales observados en la ocurrencia de incendios y en la superficie quemada permiten identificar zonas del territorio con mayor exposición al riesgo. Sin embargo, para complementar este análisis territorial resulta necesario comprender también los mecanismos físicos que gobiernan la propagación del fuego a escala local. En este contexto, la transferencia de calor por radiación cumple un papel fundamental, ya que permite que una fuente térmica, como una llama o una superficie caliente, transfiera energía hacia combustibles cercanos incluso en ausencia de contacto directo.

Desde el punto de vista del modelamiento térmico, una variable clave para cuantificar este intercambio radiativo es el factor de forma, el cual representa la fracción de

energía emitida por una superficie que alcanza directamente a otra. Esta magnitud depende exclusivamente de la geometría relativa entre las superficies involucradas y permite estimar el flujo de calor radiativo entre llamas, receptores, calentadores, paredes, vegetación u otros elementos expuestos a una fuente térmica. Por esta razón, el cálculo de factores de forma resulta relevante en problemas asociados a incendios, evaluación de exposición térmica, diseño experimental y simulación de escenarios de propagación.

No obstante, en aplicaciones donde se requiere evaluar múltiples configuraciones geométricas, como análisis paramétricos, simulaciones térmicas, estudios de sensibilidad u optimización, el cálculo analítico o numérico del factor de forma puede convertirse en una etapa computacionalmente costosa. Bajo esta perspectiva, el uso de modelos sustitutos basados en aprendizaje supervisado permite aproximar de manera eficiente dichos factores, reduciendo el tiempo de cálculo sin perder la relación funcional entre la geometría del sistema y el intercambio radiativo. En consecuencia, la estimación eficiente de factores de forma constituye una herramienta útil para apoyar el análisis térmico de sistemas expuestos a radiación, especialmente en contextos donde se requiere explorar una gran cantidad de escenarios geométricos [6, 7].

1.2.5. Alcance

El presente trabajo tiene como alcance el desarrollo, implementación y validación de modelos sustitutos basados en técnicas de *Machine Learning* y *Deep Learning* para la estimación del factor de forma (*View Factor*, VF) en problemas de transferencia de calor por radiación.

La investigación considera tres configuraciones geométricas representativas, seleccionadas por su relevancia en problemas de transferencia de calor por radiación y por su conexión con aplicaciones térmicas donde la estimación del factor de forma resulta necesaria. En particular, las geometrías estudiadas permiten representar distintos escenarios de intercambio radiativo: superficies receptoras expuestas a fuentes térmicas, sistemas cilíndricos de calentamiento y configuraciones asociadas a fuentes radiantes de mayor complejidad geométrica, tales como llamas o dispositivos experimentales de calorimetría.

Las tres configuraciones consideradas son:

- **Placa rectangular y disco:** representa una configuración fundamental de intercambio radiativo entre superficies planas y circulares, útil como caso base para analizar la dependencia geométrica del factor de forma.
- **Calefactor cilíndrico y elemento diferencial:** permite estudiar el intercambio radiativo producido por una fuente cilíndrica de calentamiento hacia un elemento receptor, configuración asociada a dispositivos experimentales de transferencia de calor.
- **Cono calorímetro y elemento diferencial:** representa una configuración ampliamente utilizada en ensayos de caracterización térmica, donde resulta relevante estimar el flujo radiativo incidente sobre una muestra o superficie receptora.

En cada caso, los valores de referencia del factor de forma son obtenidos mediante integración de contornos (integrales de línea) basada en el teorema de Stokes, lo que garantiza alta precisión numérica. A partir de estos datos, se construyen modelos predictivos capaces de aproximar el factor de forma sin necesidad de resolver explícitamente las integrales dobles asociadas al problema físico.

El estudio se enfoca en:

- La generación de bases de datos sintéticos de alta fidelidad.
- La construcción de modelos de regresión mediante:
 - Modelo de regresión polinómica con términos de interacción.
 - Modelo de regresión polinómica con términos de interacción y regularización.
 - Regresión mediante *Support Vector Regression* (SVR).
 - Regresión mediante redes neuronales multicapa completamente conectadas (*Multilayer Perceptron*, MLP).

- La optimización de hiperparámetros mediante técnicas de búsqueda sistemática y probabilística.
- La evaluación comparativa de desempeño en términos de precisión y eficiencia computacional.

No se abordan fenómenos radiativos dependientes de propiedades espectrales de materiales, medios participantes o efectos transitorios, limitándose el análisis a configuraciones geométricas y factores de forma en régimen estacionario.

1.2.6. Objetivo General

Desarrollar y validar modelos sustitutos basados en aprendizaje automático capaces de estimar con alta precisión y eficiencia computacional el factor de forma en distintas configuraciones geométricas de transferencia radiativa, reemplazando la evaluación directa de integrales de contorno.

1.2.7. Objetivos Específicos

1. Formular matemáticamente el cálculo del factor de forma para tres configuraciones geométricas mediante integración de contornos basada en el teorema de Stokes, y generar datasets sintéticos de alta fidelidad mediante la variación sistemática de los parámetros geométricos relevantes en cada caso de estudio.
2. Implementar modelos sustitutos para la predicción del factor de forma, incluyendo regresión polinómica clásica y regularizada, *Regresión por vectores de soporte* (SVR) y redes neuronales multicapa (MLP).
3. Optimizar los hiperparámetros de los modelos mediante técnicas de búsqueda y selección, tales como *Grid Search*, optimización bayesiana y criterios de información (AIC/BIC).
4. Evaluar el desempeño predictivo y la capacidad de generalización de los modelos frente a configuraciones geométricas de distinta complejidad, utilizando métricas estadísticas de validación.
5. Comparar la eficiencia computacional de los modelos sustitutos respecto del método integral tradicional, cuantificando el factor de aceleración (*speed-up*).

2 | Estado del arte

2.1. Cálculo clásico del factor de forma

El factor de forma (*View Factor*, VF) constituye un parámetro fundamental en la modelación de la transferencia de calor por radiación, ya que determina la fracción de energía radiativa intercambiada entre superficies. Desde un punto de vista matemático, su formulación general involucra una integral doble sobre las superficies emisora y receptora, lo que implica un elevado costo computacional, especialmente en configuraciones geométricas complejas [8, 9].

Con el fin de reducir la dimensionalidad del problema, Sparrow propuso la transformación de las integrales de superficie en integrales de contorno mediante la aplicación del teorema de Stokes [10]. Esta formulación permite expresar el factor de forma como una integral doble de línea evaluada sobre los contornos de las superficies involucradas. En muchas configuraciones geométricas, esta transformación conduce a una reducción significativa del esfuerzo computacional sin comprometer la precisión del cálculo.

Asimismo, diversos textos clásicos en transferencia radiativa han desarrollado expresiones analíticas cerradas para configuraciones geométricas específicas [8]. Sin embargo, dichas soluciones suelen estar restringidas a geometrías idealizadas, lo que limita su aplicabilidad en escenarios más complejos presentes en problemas reales de ingeniería, tales como hornos industriales, cámaras de combustión, sistemas de calentamiento radiativo, dispositivos de caracterización térmica de materiales.

2.2. Métodos numéricos y aproximaciones estocásticas

Cuando se consideran geometrías arbitrarias o configuraciones más complejas, los métodos analíticos resultan difíciles de aplicar. En estos casos, los métodos numéricos y estocásticos constituyen una alternativa ampliamente utilizada.

Entre estos enfoques, los métodos basados en trazado de rayos Monte Carlo han sido empleados extensamente para la estimación de factores de forma en dominios tridimensionales complejos [8]. Estas técnicas permiten modelar interacciones radiativas en geometrías arbitrarias y capturar efectos geométricos sofisticados. No obstante, presentan una limitación importante asociada a su costo computacional,

ya que requieren un elevado número de muestras para reducir la varianza estadística de las estimaciones.

En el contexto de simulaciones térmicas avanzadas, donde el factor de forma debe evaluarse repetidamente dentro de procesos iterativos de diseño, simulación u optimización, esta carga computacional puede convertirse en un cuello de botella. Esta situación ha motivado el desarrollo de estrategias orientadas a acelerar el cálculo del factor de forma mediante aproximaciones eficientes o modelos sustitutos.

2.3. Aprendizaje automático en ingeniería térmica

En los últimos años, las técnicas de *Machine Learning* han emergido como herramientas prometedoras para la construcción de modelos sustitutos (*surrogate models*) en diversas áreas de la ingeniería computacional, incluyendo dinámica de fluidos, transferencia de calor y modelación de fenómenos físicos complejos [11, 12].

De manera general, un modelo sustituto puede definirse como una aproximación matemática de bajo costo computacional diseñada para emular la respuesta de un modelo físico, numérico o experimental de alta fidelidad, preservando con suficiente precisión la relación entre variables de entrada y salida en el dominio de interés [11, 12]. Este tipo de enfoque resulta especialmente útil cuando la evaluación repetida del modelo original es costosa, por ejemplo, en el cálculo de factores de forma mediante integrales de contorno.

En este contexto, las redes neuronales artificiales han demostrado ser herramientas eficaces para aproximar relaciones no lineales presentes en diversos problemas de ingeniería [13, 14]. En particular, el término *profundo* en (*Deep Learning*) se utiliza para describir arquitecturas neuronales con múltiples capas ocultas, lo que permite construir representaciones jerárquicas de creciente nivel de abstracción a partir de las variables de entrada [13, 14].

Desde un punto de vista teórico, el Teorema de Aproximación Universal establece que una red neuronal *feedforward* con al menos una capa oculta y funciones de activación no lineales adecuadas puede aproximar, con precisión arbitraria, funciones continuas definidas sobre conjuntos compactos, bajo ciertas condiciones [15, 16]. Aunque este resultado no garantiza por sí solo eficiencia de entrenamiento, generalización ni superioridad frente a otros modelos, sí proporciona una base teórica para emplear redes neuronales como aproximadores funcionales en problemas donde la variable de salida depende de manera continua y no lineal de los parámetros geométricos. En el caso del factor de forma, esta interpretación es razonable siempre que se trabaje dentro del dominio geométrico considerado y bajo configuraciones físicamente admisibles.

2.4. Brecha identificada

A pesar del amplio desarrollo de métodos analíticos y numéricos para el cálculo del factor de forma, así como del creciente uso de modelos de aprendizaje automático en diversas áreas de la ingeniería, la literatura muestra una integración limitada entre formulaciones físicas de alta fidelidad basadas en integrales de contorno y modelos supervisados entrenados específicamente para sustituir la evaluación integral del VF en múltiples configuraciones geométricas.

En este contexto, el presente trabajo propone una metodología que combina datos generados mediante integrales de línea o de contornos de alta precisión con modelos supervisados de aprendizaje automático, incluyendo regresión polinómica, *Support Vector Regression* (SVR) y redes neuronales multicapa (MLP). Estos modelos son evaluados en términos de precisión, capacidad de generalización y eficiencia computacional, con el objetivo de determinar su potencial como aproximaciones sustitutas para el cálculo del factor de forma en problemas de ingeniería térmica.



3 | Marco Teórico

3.1. Mecanismos de transferencia de calor

Un incendio, involucra una dinámica crucial de transferencia de calor, la cual es imperativa tanto en la expansión como en la intensidad del fuego. El calor no sólo induce daño material directo, sino que también actúa como un multiplicador de la severidad del incendio, ampliando su alcance, impulsando su propagación y complicando los esfuerzos para sofocarlo. Además, la velocidad con la que el fuego se propaga está inextricablemente ligada al tipo y cantidad de calor que se transfiere en el entorno incendiado, mientras que la combustión misma es dependiente de una aportación constante y sostenida de calor.

La transferencia de calor durante un incendio se materializa a través de tres mecanismos primordiales: conducción, convección y radiación. Cada uno de estos mecanismos se desencadena en simultaneidad durante el desarrollo de un incendio, coexistiendo e interactuando en un complejo entramado que determina la evolución del fuego.

La relevancia proporcional de cada uno de estos mecanismos varía significativamente, en función de factores como la intensidad del fuego, la magnitud del incidente y las particularidades del entorno en el que se desata el incendio.

3.1.1. Mecanismos de propagación del fuego

En el contexto de los incendios, la transferencia de calor no solo se interpreta mediante los mecanismos clásicos de conducción, convección y radiación, sino también a través de procesos propios de la propagación del fuego, tales como el contacto directo de la llama, el transporte convectivo de gases calientes, la radiación térmica y la generación de pavesas. Estos mecanismos pueden actuar de manera simultánea y su importancia relativa depende de las condiciones ambientales, la disponibilidad de combustible, la geometría del entorno y la intensidad del incendio [5].

Contacto directo

El contacto directo ocurre cuando la llama alcanza físicamente un combustible cercano, transfiriendo energía de manera intensa y localizada. Este mecanismo puede favorecer la ignición cuando el material expuesto alcanza condiciones térmicas suficientes para iniciar procesos de degradación, pirólisis o combustión. En incendios

de vegetación, su relevancia depende de la continuidad del combustible, la distancia entre elementos combustibles y la estructura espacial de la vegetación [4].

Convección

La convección corresponde al transporte de energía mediante el movimiento de gases calientes generados por la combustión. Estos gases pueden precalentar combustibles ubicados en la dirección del flujo, facilitando su ignición posterior. En incendios abiertos, la convección está fuertemente influenciada por el viento, la pendiente del terreno y la intensidad de la llama, factores que pueden modificar la velocidad y dirección de propagación del frente de fuego [6].

Radiación térmica

La radiación térmica permite transferir energía desde la llama o superficies calientes hacia combustibles cercanos sin necesidad de contacto directo. Este mecanismo es especialmente relevante cuando existe separación espacial entre la fuente térmica y el material expuesto, ya que puede contribuir al precalentamiento e ignición del combustible. En este contexto, el factor de forma adquiere importancia, pues permite cuantificar qué fracción de la energía radiativa emitida por una superficie alcanza a otra, dependiendo exclusivamente de la geometría relativa entre ambas [6, 7].

Pavesas

Las pavesas corresponden a partículas incandescentes o fragmentos de material en combustión que pueden ser transportados por el viento a distintas distancias del frente principal del incendio. Este mecanismo es particularmente relevante porque puede generar focos secundarios, incluso en zonas donde no existe contacto directo con la llama principal. Diversos estudios han analizado la ignición de combustibles forestales por pavesas idealizadas, destacando la influencia de propiedades como la humedad, el tipo de combustible y las condiciones térmicas de exposición [17, 18, 19].

3.2. Aprendizaje automático: conceptos generales

El aprendizaje automático (*Machine Learning*) corresponde a un área de la inteligencia artificial orientada al desarrollo de modelos capaces de identificar patrones, relaciones y regularidades a partir de datos, con el fin de realizar predicciones, clasificaciones o tomar decisiones sin necesidad de programar explícitamente cada regla del problema [13]. En términos generales, estos métodos buscan construir funciones para aproximar resultados a partir de información observada, de modo que puedan generalizar su comportamiento frente a nuevos datos no vistos durante el entrenamiento [20].

3.2.1. Paradigmas de aprendizaje

De acuerdo con la forma en que se dispone la información y el tipo de retroalimentación entregada al algoritmo, el aprendizaje automático suele clasificarse en tres grandes paradigmas:

- Aprendizaje supervisado.
- Aprendizaje no supervisado
- Aprendizaje por refuerzo

Aprendizaje supervisado

El aprendizaje supervisado se basa en el uso de un conjunto de datos etiquetado, es decir, un conjunto de ejemplos en el que cada vector de entrada $\mathbf{x}$ se encuentra asociado a una salida objetivo y [21, 22]. El propósito del modelo es aprender una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ o $f : \mathbb{R}^n \rightarrow \mathcal{C}$ que permita aproximar la relación entre entradas y salidas, de modo que pueda predecir correctamente la respuesta correspondiente a nuevos datos no observados.

Dependiendo de la naturaleza de la variable objetivo, el aprendizaje supervisado puede dividirse en problemas de regresión y clasificación [22]. En regresión, la salida corresponde a una variable continua, mientras que en clasificación la salida pertenece a un conjunto discreto de categorías. En el presente trabajo, el problema abordado corresponde a uno de regresión, ya que el objetivo es predecir valores continuos del factor de forma a partir de parámetros geométricos de entrada.

Aprendizaje no supervisado

El aprendizaje no supervisado se ocupa de conjuntos de datos que no poseen una variable objetivo explícita. En este caso, el propósito del algoritmo es identificar estructuras internas, agrupamientos, patrones latentes o representaciones compactas de los datos [13]. Entre las tareas más comunes dentro de este paradigma se encuentran el agrupamiento (*clustering*), la reducción de dimensionalidad y la detección de anomalías [22].

A diferencia del aprendizaje supervisado, aquí no se busca predecir directamente una salida conocida, sino explorar la organización subyacente del conjunto de datos [21]. Este paradigma resulta útil en problemas exploratorios o cuando no se dispone de etiquetas confiables o bien, no se disponen de datos etiquetados.

Aprendizaje por refuerzo

El aprendizaje por refuerzo se basa en la interacción de un agente con un entorno. En este esquema, el agente ejecuta acciones, observa las consecuencias de dichas acciones y recibe recompensas o penalizaciones en función de su desempeño [23]. El objetivo consiste en aprender una política de decisión que maximice la recompensa acumulada a lo largo del tiempo [13, 23].

Este paradigma se distingue de los anteriores porque no aprende a partir de etiquetas fijas ni de la sola estructura de los datos, sino a partir de un proceso secuencial de prueba y error. Aunque ha mostrado gran éxito en problemas de control, robótica y toma de decisiones, no constituye el enfoque central de este trabajo.

3.2.2. Modelos predictivos utilizados en este trabajo

Dentro del marco del aprendizaje supervisado, existen diversos modelos para abordar problemas de regresión. En esta investigación se emplearon métodos de distinta naturaleza, con el propósito de comparar su capacidad predictiva, su costo computacional y su comportamiento frente a relaciones no lineales entre las variables geométricas de entrada y el factor de forma.

Regresión polinómica

La regresión polinómica es una extensión de la regresión lineal que permite modelar relaciones no lineales entre las variables de entrada y la variable objetivo mediante la incorporación de términos polinomiales e interacciones [24]. Aunque el modelo sigue siendo lineal respecto de sus coeficientes, la transformación de las variables de entrada permite representar comportamientos curvos y dependencias más complejas [22].

Si $\mathbf{x} = (x_1, \dots, x_n)$ denota el vector de entrada, la regresión polinómica construye nuevas características a partir de potencias e interacciones entre sus componentes. Posteriormente, el modelo ajusta una combinación lineal de dichas características para aproximar la salida [24]. Esta estrategia resulta especialmente útil cuando la relación funcional es suave y puede representarse adecuadamente mediante expansiones polinomiales de grado finito [25].

finalmente, para encontrar el mejor conjunto de parámetros, se resuelve el siguiente problema de minimización:

$$\min_B \|AB - y\|_2^2$$

Donde A es la matriz que contiene los datos de entrada, y es el vector que contiene las respectivas salidas y β es el vector de parámetros que se busca encontrar en dicho

problema de optimización.

Modelos regularizados y no regularizados

En modelos de regresión, una distinción importante corresponde a la diferencia entre modelos no regularizados y modelos regularizados. Un modelo no regularizado ajusta sus parámetros minimizando únicamente el error sobre los datos de entrenamiento. Si bien esto puede conducir a un ajuste muy preciso sobre los datos observados, también puede aumentar el riesgo de sobreajuste, especialmente cuando el número de variables es elevado o cuando existe colinealidad entre ellas [22].

Por su parte, un modelo regularizado incorpora un término de penalización en la función objetivo, con el fin de restringir la magnitud de los coeficientes y favorecer soluciones más estables y generalizables [22]. La regularización no solo ayuda a reducir el sobreajuste, sino que además puede mejorar el condicionamiento numérico del problema de ajuste [22].

Regresión Ridge

La regresión Ridge corresponde a una versión regularizada de la regresión lineal, en la que se añade una penalización cuadrática sobre la magnitud de los coeficientes del modelo [13]. Si β representa el vector de parámetros, la estimación en Ridge se obtiene minimizando una función del tipo

$$\min_{\beta} \{ \|A\beta - y\|_2^2 + \lambda \|\beta\|_2^2 \}$$

donde $\lambda \geq 0$ es un hiperparámetro que controla la intensidad de la penalización [26]. Cuando $\lambda = 0$, se recupera la regresión lineal ordinaria, mientras que valores mayores de λ inducen una contracción progresiva de los coeficientes.

En este trabajo, la regresión Ridge se utilizó en combinación con características polinomiales, con el objetivo de aprovechar la flexibilidad de la expansión polinómica y, al mismo tiempo, controlar el crecimiento de coeficientes asociado a bases de alta dimensión.

Regresión por máquina de vectores

La Regresión por Vectores de Soporte (*Support Vector Regression*, SVR) es una extensión de la teoría de máquinas de soporte al contexto de regresión [27, 28]. Su idea central consiste en encontrar una función que aproxime la relación entre entradas y salidas manteniendo una complejidad controlada y tolerando un error de magnitud limitada alrededor de los valores observados. Para ello, SVR introduce una región de tolerancia ε , dentro de la cual los errores no son penalizados [28, 29].

En términos matemáticos, SVR busca encontrar una función de la forma

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$$

donde $\mathbf{x}$ representa el vector de entrada, $\phi(\mathbf{x})$ corresponde a una transformación hacia un espacio de características, $\mathbf{w}$ es el vector de pesos asociado a dicho espacio y b es el

término de sesgo. En este espacio transformado, el objetivo del modelo es aproximar la relación entre las variables de entrada y la variable de salida manteniendo una complejidad controlada.

Una de las principales fortalezas de SVR es su capacidad para modelar relaciones no lineales mediante el uso de funciones kernel, que permiten operar implícitamente en espacios de mayor dimensión sin necesidad de construir explícitamente dichas transformaciones [28, 29]. En consecuencia, SVR resulta especialmente útil en problemas donde la dependencia entre variables es altamente no lineal y no puede representarse adecuadamente mediante modelos lineales simples.

Perceptrón multicapa

El perceptrón multicapa (*Multilayer Perceptron*, MLP) es una red neuronal artificial *feedforward*, es decir, que la información de entrada fluye hacia adelante hasta la predicción. Dicha arquitectura está compuesta por una capa de entrada, una o más capas ocultas y una capa de salida [13]. Cada neurona realiza una combinación afín de las entradas recibidas, seguida por una función de activación no lineal, lo que le permite a la red capturar relaciones no lineales. El resultado es una composición sucesiva de transformaciones que permite construir modelos altamente flexibles capaces de aproximar relaciones funcionales complejas [13].

Cuando el MLP posee múltiples capas ocultas, se considera una arquitectura de aprendizaje profundo, en la medida en que puede aprender representaciones jerárquicas de la información [13, 14]. El ajuste de sus parámetros se realiza habitualmente mediante algoritmos basados en descenso de gradiente y retropropagación del error *Backpropagation* [13, 30].

En problemas de regresión, la capa de salida suele estar compuesta por una sola neurona con activación lineal, mientras que las capas ocultas utilizan funciones de activación no lineales, tales como ReLU, GELU, sigmoide o tangente hiperbólica que se definen en capítulos posteriores [13]. En este trabajo, el MLP se emplea como un modelo sustituto flexible para capturar dependencias no lineales entre parámetros geométricos y valores de factor de forma.

3.2.3. Consideraciones finales

Los modelos descritos anteriormente representan distintos niveles de complejidad, flexibilidad y costo computacional. Mientras que la regresión polinómica y Ridge ofrecen interpretabilidad y rapidez de entrenamiento, SVR y MLP permiten capturar relaciones no lineales más complejas [22, 13]. La comparación entre estos enfoques resulta relevante para determinar cuál de ellos ofrece el mejor compromiso entre precisión predictiva, generalización y eficiencia computacional en la estimación del factor de forma.

3.3. Técnicas de optimización en machine learning

Las técnicas de optimización en machine learning corresponden al conjunto de métodos matemáticos y computacionales utilizados para ajustar los parámetros de un modelo con el objetivo de minimizar una función de pérdida. Esta función cuantifica la diferencia entre las predicciones del modelo y los valores observados en los datos. En problemas de regresión, por ejemplo, una función de pérdida común es el error cuadrático medio, mientras que en problemas de clasificación suelen emplearse funciones como la entropía cruzada. En este trabajo se aborda la predicción de un factor de forma, por lo cual, el problema que se intenta resolver es una regresión.

En términos generales, el proceso de aprendizaje puede interpretarse como un problema de optimización. Dado un conjunto de datos de entrenamiento, el modelo busca encontrar los parámetros que permitan representar de la mejor manera posible la relación entre las variables de entrada y la variable de salida. En modelos simples, como la regresión lineal, este ajuste puede realizarse mediante soluciones analíticas, por ejemplo, utilizando mínimos cuadrados. Sin embargo, en modelos más complejos, como redes neuronales profundas, máquinas de soporte vectorial, suele ser necesario recurrir a métodos iterativos de optimización, métodos que se abordarán con mayor profundidad teórica y matemática más adelante.

Entre las técnicas de optimización más utilizadas se encuentran el descenso del gradiente y sus variantes. El descenso del gradiente clásico actualiza los parámetros del modelo en la dirección opuesta al gradiente de la función de pérdida, buscando reducir progresivamente el error. Una variante ampliamente utilizada es el descenso del gradiente estocástico, conocido como *Stochastic Gradient Descent* (SGD), el cual actualiza los parámetros utilizando subconjuntos pequeños de datos o incluso una sola observación por iteración. Esta estrategia permite reducir el costo computacional y facilita el entrenamiento de modelos con grandes volúmenes de datos.

También existen métodos adaptativos, como AdaGrad, RMSProp y Adam, que modifican dinámicamente la tasa de aprendizaje durante el entrenamiento. Estas técnicas ajustan el tamaño de los pasos de actualización de acuerdo con el comportamiento de los gradientes, lo que puede mejorar la estabilidad y velocidad de convergencia, dado que el acercamiento al óptimo se calibra con cada iteración.

Además de los optimizadores basados en gradientes, existen técnicas de optimización orientadas a la selección de hiperparámetros. A diferencia de los parámetros internos del modelo, los hiperparámetros no se aprenden directamente durante el entrenamiento, sino que deben definirse previamente. Entre ellos están la cantidad de capas y neuronas de una red, la tasa de aprendizaje de un optimizador, la regularización que se le aplica a un modelo para evitar un sobreajuste, entre otros.

Para ajustar estos valores se utilizan estrategias como *Grid Search*, *Random Search* y optimización bayesiana.

El *Grid Search* evalúa sistemáticamente todas las combinaciones posibles dentro de un conjunto discreto de valores definidos previamente. Aunque es fácil de implementar e interpretar, puede volverse costoso cuando el número de hiperparámetros aumenta. Por otro lado, el *Random Search* selecciona combinaciones de manera aleatoria dentro del espacio de búsqueda, lo que puede resultar más eficiente en espacios

de alta dimensión. La optimización bayesiana, en cambio, construye un modelo probabilístico del desempeño del algoritmo y utiliza dicha información para seleccionar de manera inteligente las siguientes combinaciones de hiperparámetros a evaluar.

Estas técnicas se utilizan para mejorar el desempeño predictivo, reducir el error de generalización, evitar problemas de sobreajuste y encontrar configuraciones adecuadas para cada modelo. En el contexto de modelos sustitutos aplicados a problemas de transferencia de calor por radiación, la optimización resulta especialmente relevante, ya que permite ajustar modelos capaces de aproximar con alta precisión relaciones geométricas no lineales asociadas al cálculo del factor de forma. De este modo, una adecuada estrategia de optimización contribuye tanto a la precisión del modelo como a su eficiencia computacional durante la etapa de inferencia.

En este capítulo se presentan las distintas técnicas utilizadas para el cálculo de factores de forma, así como los modelos de *Machine Learning* y *Deep Learning* empleados para aproximar dichos factores. El objetivo de estas metodologías es reducir el costo computacional asociado al cálculo directo de los factores de forma.

3.4. Fundamentos Matemáticos

El presente trabajo combina herramientas provenientes del cálculo vectorial, la geometría diferencial y la optimización matemática con el objetivo de modelar y aproximar factores de forma mediante técnicas de aprendizaje automático.

La formulación matemática del problema involucra conceptos de cálculo multivariable, integrales de contorno y métodos de optimización utilizados en el entrenamiento de modelos de Machine Learning. En esta sección se presentan los fundamentos matemáticos necesarios para comprender el modelo físico empleado para generar los datos y los algoritmos utilizados para entrenar los modelos predictivos.

3.4.1. Campos escalares y derivadas

Un campo escalar se define como una función

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

que asigna un valor real a cada punto del espacio. En muchas aplicaciones físicas, los campos escalares representan magnitudes tales como temperatura, presión o energía potencial [31].

Para funciones de varias variables, el cambio de la función respecto a cada una de sus variables se describe mediante derivadas parciales. Si

$$f(x_1, x_2, \dots, x_n)$$

es una función diferenciable, entonces sus derivadas parciales se definen como

$$\frac{\partial f}{\partial x_i}$$

las cuales miden la tasa de cambio de la función respecto a cada variable independiente.

Estas derivadas parciales permiten construir el gradiente de la función, concepto fundamental en problemas de optimización y aprendizaje automático.

3.4.2. Gradiente

El gradiente de una función escalar diferenciable

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

se define como el vector formado por todas sus derivadas parciales

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

El gradiente apunta en la dirección de mayor incremento de la función, mientras que el vector negativo del gradiente indica la dirección de máximo descenso [31].

Este concepto constituye la base de numerosos algoritmos de optimización utilizados en Machine Learning, particularmente el algoritmo de descenso de gradiente.

3.4.3. Matriz Jacobiana

Cuando se trabaja con funciones vectoriales de la forma

$$\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

la generalización del gradiente corresponde a la matriz Jacobiana. Sea

$$\mathbf{F}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

la matriz Jacobiana se define como

$$J_F(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

La matriz Jacobiana describe la sensibilidad de un sistema respecto a cambios en sus variables y es ampliamente utilizada en análisis numérico, optimización y métodos iterativos para resolver sistemas no lineales [32].

3.4.4. Matriz Hessiana

Para funciones escalares

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

la matriz Hessiana corresponde a la matriz formada por todas las derivadas parciales de segundo orden

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

Si la función es suficientemente regular, la matriz Hessiana es simétrica.

La Hessiana juega un rol fundamental en problemas de optimización. En particular, si la matriz Hessiana es definida positiva en un punto crítico, dicho punto

corresponde a un mínimo local, mientras que si es definida negativa corresponde a un máximo local [32, 33].

3.4.5. Descomposición en valores singulares

Sea $A \in \mathbb{R}^{m \times n}$. La descomposición en valores singulares (*Singular Value Decomposition*, SVD) establece que toda matriz puede factorizarse como

$$A = U \Sigma V^T$$

donde $U \in \mathbb{R}^{m \times m}$ y $V \in \mathbb{R}^{n \times n}$ son matrices ortogonales, mientras que Σ es una matriz diagonal que contiene los valores singulares σ_i de la matriz A , los cuales satisfacen

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0.$$

La descomposición SVD constituye una herramienta fundamental en análisis numérico y aprendizaje automático, ya que permite analizar propiedades geométricas y numéricas de las matrices asociadas a problemas de optimización y regresión [34].

El número de condición de una matriz mide la sensibilidad de la solución de un sistema lineal frente a perturbaciones en los datos. En términos de la norma inducida 2, el número de condición puede expresarse mediante los valores singulares de la matriz:

$$\kappa(A) = \frac{\sigma_{\max}}{\sigma_{\min}}$$

donde $\sigma_{\max}$ y $\sigma_{\min}$ corresponden al mayor y menor valor singular de la matriz A , respectivamente.

Cuando el número de condición es grande, el sistema se considera mal condicionado, lo que implica que pequeñas perturbaciones en los datos pueden generar grandes variaciones en la solución. En contraste, matrices con números de condición cercanos a uno corresponden a sistemas bien condicionados [32, 34].

3.4.6. Optimización en Machine Learning

El entrenamiento de modelos de Machine Learning puede formularse como un problema de optimización cuyo objetivo es minimizar una función de pérdida que mide la discrepancia entre los valores predichos por el modelo y los valores reales observados en los datos [21, 22].

En problemas de regresión, una de las funciones de pérdida más utilizadas es el error cuadrático medio (Mean Squared Error, MSE), definido como

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

donde y_i representa el valor real y $\hat{y}_i$ corresponde a la predicción del modelo.

3.4.7. Descenso de gradiente

El algoritmo de descenso de gradiente es uno de los métodos más utilizados para minimizar funciones de pérdida en Machine Learning. Este método actualiza iterativamente los parámetros del modelo en la dirección opuesta al gradiente de la función de pérdida [13, 32].

La actualización de los parámetros se define como

$$\theta_{k+1} = \theta_k - \eta \nabla L(\theta_k)$$

donde η corresponde a la tasa de aprendizaje.

Este algoritmo permite encontrar aproximaciones al mínimo de la función de pérdida mediante iteraciones sucesivas.

A continuación, se presenta un ejemplo visual del descenso de gradiente aplicado a una función no convexa. A partir de un punto inicial arbitrario, el algoritmo actualiza iterativamente las variables en la dirección opuesta al gradiente, reduciendo progresivamente el valor de la función objetivo hasta aproximarse a un mínimo local.

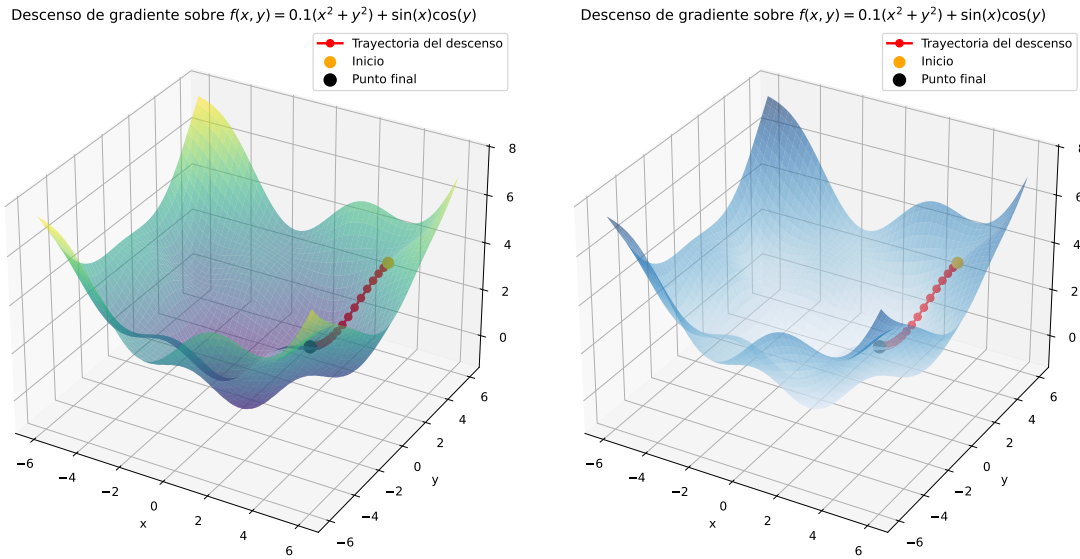


Figura 3.1: Visualización del descenso de gradiente aplicado a una función no convexa. La superficie representa la función objetivo, mientras que los puntos rojos muestran la trayectoria iterativa seguida por el algoritmo desde el punto inicial hasta el punto final. La figura de la izquierda muestra la superficie con un mapa de colores, mientras que la figura de la derecha utiliza una visualización más transparente para destacar la trayectoria de minimización.

3.4.8. Descenso de gradiente estocástico

Una variante ampliamente utilizada del descenso de gradiente es el descenso de gradiente estocástico (*Stochastic Gradient Descent*, SGD). A diferencia del descenso de gradiente clásico, que calcula el gradiente utilizando la totalidad del conjunto de

datos, el descenso de gradiente estocástico actualiza los parámetros utilizando una muestra individual o un pequeño subconjunto del conjunto de entrenamiento [13].

La actualización de los parámetros puede expresarse como

$$\theta_{k+1} = \theta_k - \eta \nabla L_i(\theta_k)$$

donde $L_i(\theta_k)$ corresponde a la función de pérdida evaluada en una muestra i del conjunto de datos y η representa la tasa de aprendizaje.

Este enfoque permite reducir significativamente el costo computacional por iteración y facilita el entrenamiento de modelos en conjuntos de datos de gran tamaño.

3.4.9. Método de Newton

Otra técnica ampliamente utilizada en optimización es el método de Newton, el cual utiliza información de segundo orden a través de la matriz Hessiana [32].

La actualización de los parámetros se define como

$$\theta_{k+1} = \theta_k - H^{-1} \nabla L(\theta_k)$$

donde H corresponde a la matriz Hessiana de la función de pérdida.

Este método suele presentar una convergencia más rápida que el descenso de gradiente cuando la Hessiana es bien condicionada.

3.4.10. Optimizador Adam

En el entrenamiento de redes neuronales profundas, uno de los algoritmos de optimización más utilizados es Adam (*Adaptive Moment Estimation*). Este método combina las ventajas del descenso de gradiente con momento y la adaptación automática de la tasa de aprendizaje para cada parámetro del modelo.

El algoritmo mantiene estimaciones exponencialmente ponderadas del primer y segundo momento del gradiente. El primer momento corresponde al promedio del gradiente, mientras que el segundo momento estima la varianza de dicho gradiente. Estas estimaciones permiten ajustar dinámicamente la magnitud de las actualizaciones durante el proceso de entrenamiento.

La actualización de los parámetros se expresa como

$$\theta_{t+1} = \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

donde α corresponde a la tasa de aprendizaje, $\hat{m}_t$ representa la estimación corregida del primer momento del gradiente y $\hat{v}_t$ corresponde a la estimación corregida del segundo momento.

Este método presenta una rápida convergencia y una gran estabilidad en problemas de optimización de alta dimensionalidad, razón por la cual es ampliamente utilizado en el entrenamiento de redes neuronales profundas [35].

3.4.11. Curvas parametrizadas

Una curva en el espacio tridimensional puede representarse mediante una función vectorial

$$\mathbf{r}(t) = (x(t), y(t), z(t)), \quad t \in I$$

donde t corresponde al parámetro que describe la curva.

El vector tangente a la curva está dado por

$$\mathbf{r}'(t)$$

el cual describe la dirección de la curva en cada punto [31].

Las curvas parametrizadas se utilizan en este trabajo para describir los contornos de las superficies involucradas en el cálculo de los factores de forma.

3.4.12. Integrales de línea

Sea C una curva parametrizada por

$$\mathbf{r}(t)$$

la integral de línea de un campo vectorial $\mathbf{F}$ a lo largo de la curva se define como

$$\int_C \mathbf{F} \cdot d\mathbf{r}.$$

Este tipo de integrales permite evaluar magnitudes físicas acumuladas a lo largo de una trayectoria y constituye la base matemática para la formulación de los factores de forma utilizada en este trabajo [36].

3.4.13. Teorema de Stokes

El teorema de Stokes establece una relación fundamental entre integrales de superficie e integrales de línea. En términos generales, el teorema establece que

$$\iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint_{\partial S} \mathbf{F} \cdot d\mathbf{r}.$$

Este resultado permite transformar integrales de superficie en integrales de contorno, reduciendo la dimensionalidad del problema [36].

En el contexto del cálculo de factores de forma, esta propiedad permite reformular integrales de superficie dobles en términos de integrales de contorno, lo que facilita significativamente su evaluación numérica.

3.5. Factores de Forma

En esta sección se presentan las formulaciones matemáticas empleadas para el cálculo de los factores de forma. Dichas formulaciones se abordan desde dos perspectivas complementarias: mediante integrales de superficie y mediante integrales de contorno, derivadas a partir del teorema de Stokes. Ambas aproximaciones permiten caracterizar rigurosamente el intercambio radiativo entre superficies [8].

3.6. Integrales de Línea y Superficie

Los factores de forma implementados en esta tesis serán abordados desde dos perspectivas importantes:

Por integrales de superficie:

$$F_{1 \rightarrow 2} = \frac{1}{\pi A_1} \int \int_{A_1 A_2} \frac{\cos(\theta_1) \cos(\theta_2)}{R^2} ds_1 ds_2 \quad (3.1)$$

Por Stokes:

$$F_{1 \rightarrow 2} = \frac{1}{2\pi A_1} \int \int_{C_1 C_2} \ln(R) dr_1 dr_2 \quad (3.2)$$

Donde R es la distancia entre superficies.

3.7. Factor de forma por integrales de superficies.

Dada la parametrización de las superficies 1 y 2:

$$r_1(u, v), \quad r_2(\theta, \phi) \quad (3.3)$$

Se define:

$$m_1 = r_2(\theta, \phi) - r_1(u, v), \quad m_2 = r_1(u, v) - r_2(\theta, \phi) \quad (3.4)$$

La norma euclidiana nos da:

$$\|m_1\| = \|m_2\| \quad (3.5)$$

Del producto punto:

$$n_1 \cdot m_1 = \|n_1\| \|m_1\| \cos(\theta_1), \quad n_2 \cdot m_2 = \|n_2\| \|m_2\| \cos(\theta_2) \quad (3.6)$$

3.7.1. Vectores Normales a las Superficies

Los vectores normales se calculan como:

$$\vec{n}_1 = \frac{\partial r_1}{\partial u} \times \frac{\partial r_1}{\partial v}, \quad \vec{n}_2 = \frac{\partial r_2}{\partial \theta} \times \frac{\partial r_2}{\partial \phi} \quad (3.7)$$

Se obtiene la relación entre los cosenos de los ángulos:

$$\cos(\theta_1) = \frac{\vec{n}_1 \cdot \vec{m}_1}{\|\vec{n}_1\| \|\vec{m}_1\|}, \quad \cos(\theta_2) = \frac{\vec{n}_2 \cdot \vec{m}_2}{\|\vec{n}_2\| \|\vec{m}_2\|} \quad (3.8)$$

3.7.2. Expresión Final del Factor de Forma

Reordenando:

$$\cos(\theta_1) \cos(\theta_2) = \frac{(\vec{n}_1 \cdot \vec{m}_1)(\vec{n}_2 \cdot \vec{m}_2)}{\|\vec{n}_1\| \|\vec{n}_2\| \|\vec{m}_1\| \|\vec{m}_2\|} \quad (3.9)$$

La forma final del factor de forma se expresa como:

$$F_{1 \rightarrow 2} = \frac{1}{\pi A_1} \int \int \frac{(\vec{n}_1 \cdot \vec{m}_1)(\vec{n}_2 \cdot \vec{m}_2)}{R^4} dA_1 dA_2 \quad (3.10)$$

la cual corresponde a una formulación clásica empleada en transferencia radiativa entre superficies arbitrarias [8].

3.8. Factor de forma por Integrales de Línea (Stokes)

El factor de forma entre dos superficies finitas A_1 y A_2 , con contornos cerrados $C_1 = \partial A_1$ y $C_2 = \partial A_2$, puede expresarse mediante integrales dobles de contorno al aplicar el teorema de Stokes de forma sucesiva (reduciendo la integral de superficie original). La formulación general puede escribirse como:

$$F_{1 \rightarrow 2} = \frac{1}{2\pi A_1} \oint_{C_2} \oint_{C_1} \ln(R) d\vec{r}_1 \cdot d\vec{r}_2, \quad (3.11)$$

donde R es la distancia euclidiana entre puntos de los contornos:

$$R = \|\vec{r}_2(v) - \vec{r}_1(u)\|. \quad (3.12)$$

3.8.1. Parametrización de los contornos

Sea C_1 parametrizado por $u \in [a, b]$ y C_2 por $v \in [c, d]$:

$$\vec{r}_1(u), \quad u \in [a, b], \quad \vec{r}_2(v), \quad v \in [c, d]. \quad (3.13)$$

Los diferenciales vectoriales de línea quedan:

$$d\vec{r}_1 = \vec{r}'_1(u) du, \quad d\vec{r}_2 = \vec{r}'_2(v) dv. \quad (3.14)$$

3.8.2. Norma euclidiana de ambos vectores.

Definimos la diferencia entre puntos de contorno como la diferencia de vectores definida por:

$$\vec{m}(u, v) = \vec{r}_2(v) - \vec{r}_1(u), \quad (3.15)$$

con norma:

$$R(u, v) = \|\vec{m}(u, v)\|. \quad (3.16)$$

3.8.3. Producto entre vectores tangentes

La integral doble de contorno depende del producto punto entre los vectores tangentes a los contornos:

$$d\vec{r}_1 \cdot d\vec{r}_2 = \vec{r}'_1(u) \cdot \vec{r}'_2(v) du dv. \quad (3.17)$$

Reemplazando en (3.11), se obtiene la forma computable:

$$F_{1 \rightarrow 2} = \frac{1}{2\pi A_1} \int_c^d \int_a^b \ln(R(u, v)) \left(\vec{r}'_1(u) \cdot \vec{r}'_2(v) \right) du dv. \quad (3.18)$$

3.8.4. Descomposición por segmentos.

En aplicaciones prácticas, los contornos suelen representarse como la unión de segmentos parametrizables (rectas, arcos, circunferencias, hélices, etc.). Si

$$C_1 = \bigcup_{i=1}^{N_1} C_{1,i}, \quad C_2 = \bigcup_{j=1}^{N_2} C_{2,j},$$

entonces el factor de forma se expresa como suma de integrales dobles:

$$F_{1 \rightarrow 2} = \frac{1}{2\pi A_1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \int_{C_{2,j}} \int_{C_{1,i}} \ln(\|\vec{r}_{2,j}(v) - \vec{r}_{1,i}(u)\|) d\vec{r}_{1,i} \cdot d\vec{r}_{2,j}. \quad (3.19)$$

Esta descomposición es la base para implementar el cálculo numérico, ya que permite integrar cada par de segmentos de forma independiente y luego sumar los aportes.

3.8.5. Expresión final del factor de forma

De forma compacta, el enfoque por integrales de línea queda determinado por:

$$F_{1 \rightarrow 2} = \frac{1}{2\pi A_1} \int_c^d \int_a^b \ln(\|\vec{r}_2(v) - \vec{r}_1(u)\|) \left(\vec{r}'_1(u) \cdot \vec{r}'_2(v) \right) du dv \quad (3.20)$$

donde $\vec{r}_1(u)$ y $\vec{r}_2(v)$ parametrizan los contornos C_1 y C_2 , respectivamente. Esta formulación reemplaza la integral doble de superficie por una integral doble de contorno, reduciendo la dimensionalidad del problema y facilitando su evaluación numérica en geometrías donde los contornos pueden parametrizarse de forma simple.

3.9. Modelos de Machine Learning y Deep Learning

Esta sección presenta un resumen de los modelos de aprendizaje automático utilizados en el estudio. La selección de estos modelos se basó en su capacidad para resolver problemas de regresión lineal y no lineal y en su amplia aplicación documentada en la literatura para tareas predictivas. A continuación, se proporciona una breve descripción de sus principales características.

3.9.1. Regresión lineal múltiple con interacciones

El modelo polinomial utilizado en este trabajo puede justificarse formalmente a través del *teorema multinomial*, el cual generaliza la expansión binomial a más de dos variables. Por ejemplo, la expansión de la potencia de una suma de tres variables reales está dada por:

$$(x_1 + x_2 + x_3)^d = \sum_{n_1+n_2+n_3=d} \binom{d}{n_1, n_2, n_3} x_1^{n_1} x_2^{n_2} x_3^{n_3} \quad (3.21)$$

donde el coeficiente multinomial se define como:

$$\binom{d}{n_1, n_2, n_3} = \frac{d!}{n_1! n_2! n_3!} \quad (3.22)$$

Esta expresión genera todos los monomios posibles de grado total d formados por la interacción de las tres variables, incluyendo términos cruzados como $x_1 x_2^2 x_3^3$. Por lo tanto, un modelo polinomial completo de grado d en tres variables incluirá todos los monomios de la forma:

$$x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \quad \text{tal que} \quad \alpha_1 + \alpha_2 + \alpha_3 \leq d$$

Esto proporciona la base matemática para incorporar tanto no linealidades como interacciones entre variables en el modelo de regresión.

La regresión polinomial extiende el modelo lineal clásico al incorporar términos de orden superior e interacciones entre variables, permitiendo aproximar relaciones no lineales complejas. Sea $\mathbf{x} = (x_0, x_1, x_2) \in \mathbb{R}^3$ el vector de variables de entrada, y sea $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2) \in \mathbb{N}_0^3$ un multi-índice de enteros no negativos. El modelo completo de regresión polinomial de grado hasta n puede expresarse de manera compacta como:

$$\hat{y} = \underbrace{T_0}_{\text{Grado 0}} + \underbrace{T_1}_{\text{Grado 1}} + \underbrace{T_2}_{\text{Grado 2}} + \underbrace{T_3}_{\text{Grado 3}} + \cdots + \underbrace{T_n}_{\text{Grado } n} \quad (3.23)$$

donde cada término T_g agrupa todos los monomios de grado total g , y se define como:

$$T_g = \sum_{\alpha_0+\alpha_1+\alpha_2=g} \beta_{\alpha_0, \alpha_1, \alpha_2}^{(g)} \cdot x_0^{\alpha_0} x_1^{\alpha_1} x_2^{\alpha_2} \quad (3.24)$$

Para completar, el modelo total que incluye todos los monomios de grado hasta 9 puede escribirse de manera equivalente como:

$$\hat{y} = \sum_{\substack{\alpha_0+\alpha_1+\alpha_2 \leq 9 \\ \alpha_0, \alpha_1, \alpha_2 \in \mathbb{N}_0}} \beta_{\alpha_0, \alpha_1, \alpha_2} \cdot x_0^{\alpha_0} x_1^{\alpha_1} x_2^{\alpha_2} \quad (3.25)$$

En los casos analizados en este trabajo, se implementaron dos modelos de regresión polinomial con interacciones para estimar factores de forma a partir de configuraciones geométricas: (1) un modelo clásico basado en mínimos cuadrados ordinarios (OLS), y (2) un modelo regularizado utilizando regresión Ridge. Ambos modelos incluyen no linealidades e interacciones entre variables hasta un grado máximo n , el cual se selecciona según la configuración evaluada.

3.9.2. Formulación del problema

Dada la expansión polinomial completa hasta grado n , el modelo puede escribirse en forma matricial como [37]:

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} \quad (3.26)$$

donde $\mathbf{X}$ corresponde a la matriz de diseño que contiene las variables predictoras y sus términos polinomiales, $\boldsymbol{\beta}$ representa el vector de parámetros del modelo, y $\mathbf{y}$ corresponde al vector de respuestas observadas asociadas a cada instancia del conjunto de datos.

3.10. Estimación de parámetros vía mínimos cuadrados

Sea $\mathbf{X} \in \mathbb{K}^{m \times n}$ con $m \gg n$. El sistema lineal

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{y} \quad (3.27)$$

es un sistema **sobredeterminado**, es decir, posee más ecuaciones que incógnitas. En general, este sistema no admite solución exacta.

El sistema tendrá solución exacta **si y solo si** el vector $\mathbf{y}$ pertenece a la imagen de $\mathbf{X}$, es decir,

$$\mathbf{y} \in \text{Im}(\mathbf{X}), \quad (3.28)$$

lo que equivale a que $\mathbf{y}$ pueda expresarse como combinación lineal de las columnas de $\mathbf{X}$.

Si $\mathbf{y} \notin \text{Im}(\mathbf{X})$, entonces no existe un vector $\boldsymbol{\beta}$ tal que

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{y}. \quad (3.29)$$

En este contexto, el vector de coeficientes $\boldsymbol{\beta}$ se estima mediante el método de los **mínimos cuadrados ordinarios** (OLS), el cual busca minimizar la norma euclidiana del residuo [22]:

$$\min_{\boldsymbol{\beta}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2. \quad (3.30)$$

Definimos la función objetivo como:

$$J(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2. \quad (3.31)$$

Recordando que la norma euclidiana está inducida por el producto interno,

$$\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}, \quad (3.32)$$

se obtiene:

$$J(\beta) = (\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta). \quad (3.33)$$

Desarrollando el producto:

$$J(\beta) = \mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\beta - (\mathbf{X}\beta)^T \mathbf{y} + (\mathbf{X}\beta)^T (\mathbf{X}\beta). \quad (3.34)$$

Utilizando que $(\mathbf{X}\beta)^T = \beta^T \mathbf{X}^T$, se obtiene:

$$J(\beta) = \mathbf{y}^T \mathbf{y} - 2\beta^T \mathbf{X}^T \mathbf{y} + \beta^T \mathbf{X}^T \mathbf{X} \beta. \quad (3.35)$$

Para minimizar $J(\beta)$, derivamos respecto a β e igualamos a cero:

$$\nabla_{\beta} J(\beta) = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \beta = 0. \quad (3.36)$$

De aquí se obtienen las **ecuaciones normales**:

$$\mathbf{X}^T \mathbf{X} \beta = \mathbf{X}^T \mathbf{y}. \quad (3.37)$$

Si $\mathbf{X}^T \mathbf{X}$ es invertible (equivalentemente, si $\mathbf{X}$ tiene rango columna completo), la solución analítica viene dada por:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (3.38)$$

En este caso, la solución es única y corresponde al **proyector ortogonal** de $\mathbf{y}$ sobre el subespacio columna de $\mathbf{X}$, es decir, sobre $\text{Im}(\mathbf{X})$ [38].

En base a lo anterior, se definen las funciones de pérdida correspondientes a los modelos de regresión múltiple con términos de interacción y de regresión múltiple con interacciones regularizadas, las cuales se presentan a continuación.

Modelo 1: Regresión Polinomial Clásica. Este modelo estima $\beta \in \mathbb{R}^n$ minimizando el criterio de mínimos cuadrados ordinarios:

$$J(\beta) = \|\mathbf{y} - \mathbf{X}\beta\|^2 \quad (3.39)$$

Modelo 2: Regresión Polinomial Ridge. Con el fin de capturar interacciones geométricas más complejas sin incurrir en sobreajuste, se introduce un término de regularización, dando lugar a la función de pérdida Ridge:

$$J_{\text{ridge}}(\beta) = \|\mathbf{y} - \mathbf{X}\beta\|^2 + \lambda \|\beta\|^2 \quad (3.40)$$

El parámetro de regularización $\lambda \geq 0$ será optimizado específicamente para cada configuración evaluada, con el fin de ajustar adecuadamente el modelo y prevenir el sobreajuste.

3.10.1. Crecimiento combinatorial en modelos polinomiales

Los modelos de regresión polinómica con términos de interacción presentan una limitación importante asociada al crecimiento combinatorial del número de términos del modelo a medida que aumenta el grado del polinomio o el número de variables predictoras.

Considerando la expansión multinomial

$$(x_1 + x_2 + x_3)^d = \sum_{n_1+n_2+n_3=d} \binom{d}{n_1, n_2, n_3} x_1^{n_1} x_2^{n_2} x_3^{n_3} \quad (3.41)$$

donde los coeficientes multinomiales están definidos como

$$\binom{d}{n_1, n_2, n_3} = \frac{d!}{n_1! n_2! n_3!}. \quad (3.42)$$

Esta expresión ilustra cómo el número de términos crece rápidamente con el grado d . En general, para un modelo polinómico de grado d con p variables predictoras, el número total de términos posibles del modelo viene dado por

$$N(p, d) = \binom{p+d}{d}. \quad (3.43)$$

Este crecimiento combinatorial implica que, incluso para valores moderados de p y d , el número de parámetros del modelo puede incrementarse significativamente. Como consecuencia, el modelo puede volverse altamente complejo, aumentando el riesgo de sobreajuste y generando posibles problemas de estabilidad numérica durante la estimación de los coeficientes.

En particular, cuando se incluyen términos de interacción entre variables, la matriz de diseño del modelo puede volverse mal condicionada, lo que provoca coeficientes de gran magnitud o altamente sensibles a pequeñas perturbaciones en los datos.

Debido a estas limitaciones, en este trabajo se emplearán modelos polinomiales de bajo grado con el fin de evitar un crecimiento excesivo del número de variables predictoras. En aquellos casos donde se requiera incrementar el grado del polinomio, se recurrirá a modelos regularizados con el objetivo de mitigar problemas de sobreajuste (*overfitting*) y mantener la estabilidad y confiabilidad de las predicciones.

3.11. Support Vector Regression (SVR)

Support Vector Regression (SVR) es una extensión del marco de las Support Vector Machines (SVM) al problema de regresión, propuesta originalmente por Vapnik [27, 29]. El objetivo del SVR es encontrar una función que aproxime la relación entre las variables predictoras y la variable objetivo, manteniendo simultáneamente una buena capacidad de generalización.

3.11.1. Formulación del problema

Sea un conjunto de entrenamiento

$$\{(x_i, y_i)\}_{i=1}^n, \quad x_i \in \mathbb{R}^d, \quad y_i \in \mathbb{R}. \quad (3.44)$$

En su forma lineal, SVR busca una función:

$$f(x) = w^T x + b \quad (3.45)$$

donde

$$w \in \mathbb{R}^d, \quad b \in \mathbb{R}. \quad (3.46)$$

El modelo se construye bajo el principio de margen máximo, minimizando la complejidad del modelo y permitiendo errores menores a un umbral ε .

3.11.2. Función de pérdida ε -insensible

SVR utiliza la función de pérdida ε -insensible definida como:

$$L_\varepsilon(y, f(x)) = \max(0, |y - f(x)| - \varepsilon) \quad (3.47)$$

Equivalentemente,

$$L_\varepsilon(y, f(x)) = \begin{cases} 0 & \text{si } |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & \text{en otro caso} \end{cases} \quad (3.48)$$

Esta función penaliza únicamente los errores que exceden una banda de tolerancia ε , lo que induce robustez frente a pequeñas perturbaciones [28].

3.11.3. Problema primal

El problema de optimización en su forma primal es:

$$\min_{w, b, \xi_i, \xi_i^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (3.49)$$

sueto a:

$$y_i - w^T x_i - b \leq \varepsilon + \xi_i \quad (3.50)$$

$$w^T x_i + b - y_i \leq \varepsilon + \xi_i^* \quad (3.51)$$

$$\xi_i \geq 0, \quad \xi_i^* \geq 0 \quad (3.52)$$

donde

$$C > 0 \quad (3.53)$$

es el parámetro de regularización.

El término

$$\frac{1}{2} \|w\|^2 \quad (3.54)$$

actúa como regularización tipo Ridge, favoreciendo soluciones con menor norma y mejor capacidad de generalización.

3.11.4. Formulación dual

Mediante multiplicadores de Lagrange y aplicando las condiciones KKT, el problema puede reformularse en su versión dual:

$$\max_{\alpha_i, \alpha_i^*} -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) x_i^T x_j - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \quad (3.55)$$

sueto a:

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \quad (3.56)$$

$$0 \leq \alpha_i, \alpha_i^* \leq C \quad (3.57)$$

La solución final adopta la forma:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) x_i^T x + b \quad (3.58)$$

Lo anterior muestra que la solución depende únicamente de un subconjunto de muestras denominadas *support vectors*.

3.11.5. Extensión no lineal y kernel trick

Para modelar relaciones no lineales, se emplea una transformación implícita

$$\phi : \mathbb{R}^d \rightarrow \mathcal{H} \quad (3.59)$$

hacia un espacio de mayor dimensión, utilizando una función kernel:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (3.60)$$

La función de regresión queda entonces:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (3.61)$$

3.11.6. Ventajas y propiedades

SVR presenta las siguientes propiedades relevantes:

- Control explícito del margen y la complejidad del modelo.
- Robustez frente a outliers mediante la pérdida ε -insensible.
- Solución convexa global.
- Dependencia únicamente de vectores soporte.

Estas características lo convierten en un método competitivo frente a modelos polinomiales de alto grado o redes neuronales en escenarios con tamaño de muestra moderado [22].

3.12. Perceptrón Multicapa (MLP)

El último modelo implementado en este estudio corresponde a una red neuronal multicapa, conocida como Perceptrón Multicapa (MLP). Este tipo de arquitectura pertenece a la familia de redes neuronales feedforward y fue formalizado en el marco del aprendizaje conexionista [30], constituyendo actualmente uno de los modelos fundamentales del aprendizaje profundo [13].

Un MLP está compuesto por una capa de entrada, una o más capas ocultas y una capa de salida. Todas las capas están completamente conectadas (fully conectad), es decir, cada neurona de una capa se conecta con todas las neuronas de la capa siguiente.

Sea el vector de entrada definido como:

$$\mathbf{x} \in \mathbb{R}^m \quad (3.62)$$

3.12.1. Propagación hacia adelante (Forward Propagation)

Para una red con una sola capa oculta, la propagación hacia adelante se define mediante transformaciones afines seguidas de funciones de activación no lineales.

En notación matricial:

$$\mathbf{Z}^{(h)} = \mathbf{X}\mathbf{W}^{(h)} + \mathbf{b}^{(h)} \quad (3.63)$$

$$\mathbf{A}^{(h)} = \sigma(\mathbf{Z}^{(h)}) \quad (3.64)$$

$$\mathbf{Z}^{(out)} = \mathbf{A}^{(h)}\mathbf{W}^{(out)} + \mathbf{b}^{(out)} \quad (3.65)$$

$$\hat{\mathbf{y}} = \sigma^{(out)}(\mathbf{Z}^{(out)}) \quad (3.66)$$

donde:

- $\mathbf{W}^{(h)}$ y $\mathbf{W}^{(out)}$ corresponden a las matrices de pesos,
- $\mathbf{b}^{(h)}$ y $\mathbf{b}^{(out)}$ son los vectores de sesgo,
- $\sigma(\cdot)$ es la función de activación en la capa oculta,
- $\sigma^{(out)}(\cdot)$ es la función de activación en la capa de salida.

Las funciones de activación introducen no linealidad en el modelo, permitiendo aproximar funciones complejas [15]. Una observación importante es que, en ausencia de funciones de activación, la salida de la red neuronal se reduciría a una simple combinación lineal de las entradas

3.12.2. Función de pérdida

El entrenamiento del modelo consiste en minimizar una función de pérdida que cuantifica la diferencia entre las predicciones y los valores reales. En este trabajo se utiliza el Error Cuadrático Medio (MSE):

$$\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.67)$$

Esta función es diferenciable y adecuada para problemas de regresión [21].

3.12.3. Funciones de activación en redes neuronales

Las redes neuronales utilizan funciones de activación con el objetivo de introducir no linealidad en el modelo, permitiendo representar relaciones complejas entre las variables de entrada y salida. Sin la presencia de funciones de activación, una red neuronal multicapa se reduciría a una transformación lineal equivalente a una regresión lineal [13].

Las funciones de activación se aplican a la salida de cada neurona y transforman la combinación lineal de las entradas en una salida no lineal que es propagada hacia las siguientes capas de la red.

En este trabajo se utilizaron las funciones de activación *Leaky ReLU* y *GELU*, ampliamente empleadas en redes neuronales modernas debido a su estabilidad durante el entrenamiento.

Leaky ReLU

La función Leaky ReLU es una variante de la función ReLU que permite mantener un pequeño gradiente para valores negativos de la entrada. Su expresión matemática se define como

$$f(x) = \begin{cases} x & x > 0 \\ \alpha x & x \leq 0 \end{cases}$$

donde α corresponde a una constante pequeña que controla la pendiente en la región negativa.

Esta función permite evitar el problema conocido como *dying ReLU*, en el cual algunas neuronas dejan de activarse durante el entrenamiento al recibir valores negativos constantes [20].

GELU

La función GELU (*Gaussian Error Linear Unit*) es una función de activación basada en la distribución normal acumulada. Su definición puede escribirse como

$$\text{GELU}(x) = x\Phi(x)$$

donde $\Phi(x)$ corresponde a la función de distribución acumulada de una variable normal estándar.

En la práctica, suele utilizarse la siguiente aproximación computacional:

$$\text{GELU}(x) = 0,5x \left(1 + \tanh \left(\sqrt{\frac{2}{\pi}}(x + 0,044715x^3) \right) \right)$$

Esta función introduce una transición suave entre regiones activas e inactivas, lo que mejora la estabilidad del entrenamiento en redes neuronales profundas [13].

3.12.4. Retropropagación (Backpropagation)

El algoritmo de retropropagación permite calcular los gradientes de la función de pérdida respecto a cada parámetro del modelo mediante la aplicación sistemática de la regla de la cadena [30].

Para la capa de salida, el gradiente respecto a la preactivación es:

$$\delta^{(out)} = \frac{\partial \mathcal{L}}{\partial \mathbf{Z}^{(out)}} = (\hat{\mathbf{y}} - \mathbf{y}) \odot \sigma^{(out)'}(\mathbf{Z}^{(out)}) \quad (3.68)$$

donde $\odot$ denota el producto elemento a elemento.

Para la capa oculta:

$$\delta^{(h)} = (\delta^{(out)} \mathbf{W}^{(out)T}) \odot \sigma'(\mathbf{Z}^{(h)}) \quad (3.69)$$

Los gradientes respecto a los parámetros quedan entonces:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{W}^{(out)}} = \mathbf{A}^{(h)T} \delta^{(out)} \quad (3.70)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{b}^{(out)}} = \sum_{i=1}^n \delta_i^{(out)} \quad (3.71)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{W}^{(h)}} = \mathbf{X}^T \delta^{(h)} \quad (3.72)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{b}^{(h)}} = \sum_{i=1}^n \delta_i^{(h)} \quad (3.73)$$

3.12.5. Actualización de parámetros

Los parámetros se actualizan mediante un algoritmo de optimización basado en gradiente, como descenso por gradiente estocástico (SGD) o variantes adaptativas como Adam:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_{\theta} \mathcal{L} \quad (3.74)$$

donde η es la tasa de aprendizaje y θ representa el conjunto de parámetros del modelo.

3.12.6. Capacidad de aproximación

Las redes neuronales multicapa poseen capacidad de aproximación universal, es decir, pueden aproximar cualquier función continua bajo condiciones adecuadas de número de neuronas [15]. Esta propiedad justifica su uso en problemas de regresión no lineal compleja, como el modelamiento del factor de forma considerado en este trabajo.



4 | Metodología

El objetivo de este trabajo es desarrollar modelos sustitutos basado en técnicas de *Machine Learning* y *Deep Learning* capaz de aproximar factores de forma en problemas de transferencia de calor radiativa.

La metodología propuesta combina modelamiento físico, generación de datos sintéticos mediante integración numérica y entrenamiento de modelos supervisados para construir aproximaciones eficientes de los factores de forma.

El procedimiento general se divide en las etapas que se describen a continuación.

4.1. Modelamiento físico de factores de forma

El punto de partida del presente estudio corresponde a la formulación matemática del factor de forma entre superficies. En este trabajo se utiliza una formulación basada en integrales de contorno obtenida a partir de la aplicación del teorema de Stokes a la expresión clásica de transferencia radiativa.

Mediante esta reformulación, el factor de forma entre dos superficies A_1 y A_2 , delimitadas por los contornos C_1 y C_2 , puede expresarse como

$$F_{1 \rightarrow 2} = \frac{1}{2\pi A_1} \int_{C_2} \int_{C_1} \ln(R_{12}) d\mathbf{r}_1 d\mathbf{r}_2 \quad (4.1)$$

donde $R_{12} = \|\mathbf{r}_1 - \mathbf{r}_2\|$ corresponde a la distancia euclidiana entre puntos de los contornos que delimitan ambas superficies, y $d\mathbf{r}_1$, $d\mathbf{r}_2$ representan los vectores tangentes asociados a las curvas que definen dichos contornos.

Esta expresión surge de la aplicación sucesiva del teorema de Stokes a la formulación clásica del factor de forma basada en integrales de superficie, permitiendo reducir la dimensionalidad del problema desde una integral doble sobre superficies a una integral doble sobre contornos.

La reducción de dimensionalidad que proporciona esta formulación resulta particularmente ventajosa desde el punto de vista computacional, ya que permite evaluar los factores de forma mediante integración numérica con un costo computacional significativamente menor.

En este trabajo, esta formulación constituye el modelo físico de referencia utilizado para generar los conjuntos de datos sintéticos empleados en el entrenamiento de los modelos de aprendizaje automático. Para ello, las integrales de contorno fueron evaluadas numéricamente para múltiples configuraciones geométricas, obteniendo así

valores de factor de forma que posteriormente fueron utilizados como datos de entrenamiento para los modelos de Machine Learning y Deep Learning.

A continuación se presentan los principales métodos utilizados en el cálculo del factor de forma, incluyendo métodos tales como integración sobre superficies, integración sobre contornos, monte carlo y modelos de machine learning.

Tabla 4.1: Comparación de métodos empleados para el cálculo de factores de forma.

Método	Descripción	Complejidad computacional	Ventajas	Limitaciones
Soluciones analíticas cerradas	Uso de expresiones exactas disponibles para configuraciones geométricas simples.	Baja, cuando existe una solución cerrada.	Alta precisión y bajo costo de evaluación.	Aplicable solo a geometrías específicas y casos altamente idealizados.
Integración doble sobre superficies	Evaluación directa de la definición del factor de forma mediante integración sobre las superficies emisora y receptora.	Alta, debido a la integración en dominios bidimensionales acoplados.	Formulación general y físicamente directa.	Elevado costo computacional, especialmente en análisis paramétricos o geometrías complejas.
Métodos de Monte Carlo	Estimación estadística del intercambio radiativo mediante el trazado aleatorio de rayos o muestras.	Media a alta, dependiendo del número de muestras requerido.	Flexible para geometrías complejas y problemas tridimensionales.	Presenta error estadístico y puede requerir muchas muestras para lograr alta precisión.
Métodos numéricos basados en discretización	División de las superficies en elementos finitos o paneles para aproximar el intercambio radiativo entre pares de elementos.	Alta, especialmente cuando aumenta el número de elementos.	Permite tratar geometrías complejas y condiciones de borde variadas.	La precisión depende de la malla; puede requerir alto costo de memoria y tiempo de cómputo.
Integrales de contorno mediante el teorema de Stokes	Transformación de la integral superficial en una integral sobre los contornos que delimitan las superficies.	Media a baja, al reducir la dimensionalidad del problema.	Disminuye el costo computacional manteniendo alta precisión numérica; adecuada para generar datos de referencia.	Requiere disponer de una parametrización adecuada de los contornos y puede ser menos directa para geometrías arbitrarias.
Modelos sustitutos basados en aprendizaje supervisado	Entrenamiento de modelos de Machine Learning o Deep Learning para aproximar el factor de forma a partir de variables geométricas.	Alta durante el entrenamiento, pero muy baja durante la inferencia.	Permiten predicciones rápidas una vez entrenados; útiles para análisis paramétricos, optimización e ingeniería inversa.	Requieren datos de entrenamiento confiables y su capacidad de generalización depende del dominio considerado.

4.2. Casos de estudio

Con el fin de evaluar la generalidad del enfoque propuesto, se analizan tres configuraciones geométricas representativas:

- **Caso 1:** Factor de forma entre una placa rectangular y un disco.
- **Caso 2:** Factor de forma entre un heater cilíndrico y un elemento diferencial.
- **Caso 3:** Factor de forma entre un cono calorímetro y un elemento diferencial.

Estos casos fueron seleccionados debido a su relevancia en aplicaciones de ingeniería relacionadas con transferencia de calor radiativa y experimentos de ciencia del fuego.

4.3. Generación del conjunto de datos

Para cada caso de estudio se definió una familia paramétrica de configuraciones geométricas, en la cual las variables geométricas relevantes fueron modificadas dentro de rangos preestablecidos. De este modo, cada combinación de parámetros define una geometría particular y, por consiguiente, un valor específico del factor de forma.

Las integraciones fueron implementadas en Python utilizando métodos de integración numérica adaptativa. Para cada combinación de parámetros geométricos se evaluaron las integrales que definen el factor de forma, generando así un conjunto de datos sintético que representa la relación funcional entre la geometría del sistema y el intercambio radiativo.

Por lo tanto, el conjunto de datos no corresponde a repeticiones de una única geometría fija, sino a un muestreo del espacio de parámetros geométricos asociado a cada configuración analizada.

El conjunto de datos obtenido constituye la referencia física utilizada para entrenar los modelos de aprendizaje automático.

4.4. Preprocesamiento de datos

Antes del entrenamiento de los modelos, las variables de entrada *Generadas mediante integrales de contorno* fueron sometidas a un proceso de escalado. El escalado de datos corresponde a un conjunto de transformaciones aplicadas a las variables de entrada con el fin de llevarlas a magnitudes comparables, evitando que diferencias arbitrarias de escala entre atributos influyan indebidamente en el proceso de aprendizaje. Esta etapa de preprocesamiento es especialmente importante cuando las variables de entrada presentan órdenes de magnitud distintos, ya que, en tal caso, aquellas con valores numéricamente más grandes pueden dominar ciertas operaciones algebraicas y sesgar el ajuste del modelo.

En términos generales, el escalado permite mejorar la interpretabilidad relativa de las variables dentro del espacio de entrada y contribuye a un entrenamiento más estable desde el punto de vista numérico. En particular, ayuda a reducir problemas asociados al mal condicionamiento de la matriz que contiene los datos de entrada, favoreciendo operaciones matriciales más robustas y disminuyendo la posibilidad de que pequeñas perturbaciones numéricas se amplifiquen durante el ajuste. Asimismo, en modelos basados en optimización iterativa, como las redes neuronales o la regresión regularizada, el escalado favorece trayectorias de descenso en la función de pérdida más equilibradas y una convergencia más eficiente.

Existen diversas técnicas de escalado, entre ellas la normalización min-max, el escalado robusto y la estandarización. En este trabajo se utilizó la estandarización,

debido a que preserva la forma general de la distribución de cada variable y transforma los datos para que tengan media cero y desviación estándar unitaria. Matemáticamente, cada variable x fue transformada según

$$z = \frac{x - \mu}{\sigma},$$

donde μ representa la media muestral de la variable y σ su desviación estándar. Como resultado, las variables transformadas quedan centradas en torno a cero y expresadas en una escala comparable.

La aplicación de esta transformación resulta necesaria en el presente estudio porque los parámetros geométricos considerados pueden encontrarse en rangos distintos. Sin este preprocesamiento, tales diferencias podrían introducir sesgos de escala en el entrenamiento y afectar tanto el desempeño predictivo como la estabilidad numérica de los modelos. Por ello, la estandarización se adoptó como una etapa previa al ajuste, con el propósito de asegurar una representación más balanceada de las variables de entrada y una construcción más robusta de los modelos sustitutos.

4.5. Modelos de Machine Learning

Con el objetivo de aproximar la relación entre los parámetros geométricos y el factor de forma, se evaluaron distintos modelos de regresión supervisada:

- Regresión polinómica con interacciones cruzadas
- Regresión polinómica con iteracciones cruzadas y regularizadas.
- Regresores por maquinas de vectores (SVR)
- Redes neuronales tipo perceptrón multicapa (MLP)

Estos modelos fueron seleccionados debido a su capacidad para representar relaciones no lineales entre variables y su amplio uso en problemas de aproximación funcional.

4.5.1. Optimización de hiperparámetros

El rendimiento de los modelos depende en gran medida de la selección adecuada de hiperparámetros. En este trabajo se emplearon diversas estrategias de optimización, incluyendo:

- Grid Search con validación cruzada
- Optimización bayesiana
- Criterios de información AIC y BIC

Estas técnicas permiten encontrar configuraciones de hiperparámetros que maximizan la capacidad predictiva de los modelos y minimizan el riesgo de sobreajuste.

4.5.2. Evaluación del desempeño

El desempeño de los modelos fue evaluado mediante diversas métricas de error comúnmente utilizadas en problemas de regresión, incluyendo:

- Mean Squared Error (MSE)
- Coeficiente de determinación (R^2)
- Absolute Percentage Error (APE)
- Mean Absolute Percentage Error (MAPE)

Estas métricas permiten cuantificar la precisión del modelo y evaluar la capacidad de generalización de cada uno de los modelos implementados.

4.5.3. Comparación con el modelo físico

Finalmente, los resultados obtenidos mediante los modelos de Machine Learning fueron comparados con los valores calculados a partir de las integrales de contorno con la finalidad de evaluar que tan bien es la capacidad de generalización de cada uno de los modelos implementados.

Adicionalmente, se evaluó la eficiencia computacional de cada enfoque, comparando los tiempos de cálculo requeridos para estimar los factores de forma mediante integración numérica y mediante los modelos sustitutos entrenados.

Este análisis permite determinar el potencial de los modelos de aprendizaje automático como herramientas eficientes para la estimación rápida de factores de forma en aplicaciones de ingeniería.



5 | Casos de estudio

En este capítulo se analizan las formulaciones matemáticas correspondientes a los tres casos previamente descritos. Posteriormente, se presentan los resultados obtenidos, donde se evalúa el desempeño de los modelos de *Machine Learning* y *Deep Learning* utilizados en el estudio.

Como se menciona en la sección de metodología, el primer paso consiste en calcular los factores de forma mediante formulaciones con integrales de contorno. Estos valores serán utilizados posteriormente para generar el conjunto de datos sintético (dataset), el cual servirá como base para el entrenamiento y validación de los modelos de Machine Learning desarrollados en este trabajo.

Como se mencionó en capítulos anteriores, los factores de forma serán calculados mediante integrales de contorno basadas en el teorema de Stokes. Dado el elevado costo computacional asociado al cálculo de integrales de superficie, la aplicación del teorema de Stokes permite transformar dichas integrales en integrales de línea a lo largo de los contornos que delimitan las superficies, reduciendo significativamente el esfuerzo computacional requerido.

En consecuencia, el factor de forma entre dos superficies finitas delimitadas por los contornos Γ_1 y Γ_2 , respectivamente, puede obtenerse de manera eficiente aplicando el teorema de Stokes en dos ocasiones, tal como se expresa en [8, 10]:

$$F_{A_i \rightarrow A_j} = \frac{1}{2\pi A_i} \oint_{\Gamma_j} \oint_{\Gamma_i} \ln(R_{ij}) d\mathbf{r}_i d\mathbf{r}_j \quad (5.1)$$

donde $R_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$ representa la distancia euclidiana entre los puntos de los contornos, y $d\mathbf{r}_i d\mathbf{r}_j = \mathbf{r}'_i(u) \cdot \mathbf{r}'_j(v) du dv$ corresponde al producto punto entre los vectores tangentes asociados a dichos contornos.

5.1. Configuraciones geométricas consideradas

A Continuación se presenta la formulación matemática de cada una de las configuraciones utilizadas en este trabajo.

5.1.1. Placa rectangular y disco

El factor de forma entre una placa rectangular y un disco será calculado mediante la evaluación de integrales de contorno y posteriormente comparado con los resultados obtenidos mediante técnicas de aprendizaje automático. Como se mencionó

previamente, estas superficies finitas serán representadas mediante sus respectivos contornos de borde.

La Figura 5.1(a) muestra el sistema de coordenadas y la orientación de las superficies finitas utilizadas en esta formulación. La Figura 5.1(b) identifica los vértices a partir de los cuales se generan las siguientes líneas paramétricas:

$$\begin{aligned}
 \Gamma_1 : \mathbf{r}_1(v) &= \left(\frac{L}{2}, \frac{P}{2} - Pv, 0 \right), 0 \leq v \leq 1 \\
 \Gamma_2 : \mathbf{r}_2(v) &= \left(\frac{L}{2} - Lv, -\frac{P}{2}, 0 \right), 0 \leq v \leq 1 \\
 \Gamma_3 : \mathbf{r}_3(v) &= \left(-\frac{L}{2}, -\frac{P}{2} + Pv, 0 \right), 0 \leq v \leq 1 \\
 \Gamma_4 : \mathbf{r}_4(v) &= \left(-\frac{L}{2} + Lv, \frac{P}{2}, 0 \right), 0 \leq v \leq 1
 \end{aligned} \tag{5.2}$$

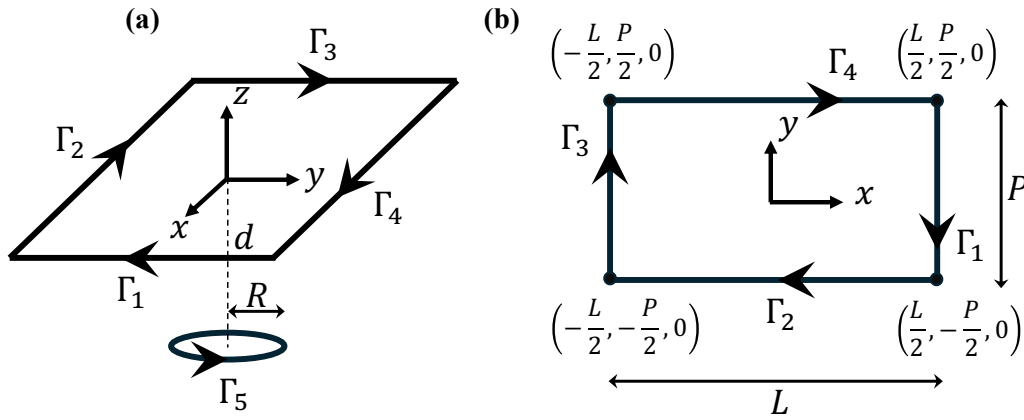


Figura 5.1: Representación esquemática de los contornos de la placa y el disco.

Adicionalmente, el disco es representado mediante un contorno circular (mostrado en la Figura 5.1(a)), cuyo centro se encuentra en las coordenadas (h, k) , lo que permite posicionar esta superficie enfrentada a la placa rectangular.

$$\Gamma_5 : \mathbf{r}_5(u) = (R \cos(u) + h, R \sin(u) + k, -d), 0 \leq u \leq 2\pi \tag{5.3}$$

El producto punto entre los vectores tangentes, $d\mathbf{r}_i d\mathbf{r}_5$, está dado por:

$$\begin{aligned}
 d\mathbf{r}_1 d\mathbf{r}_5 &= -PR \cos(u) \, dudv \\
 d\mathbf{r}_2 d\mathbf{r}_5 &= LR \sin(u) \, dudv \\
 d\mathbf{r}_3 d\mathbf{r}_5 &= PR \cos(u) \, dudv \\
 d\mathbf{r}_4 d\mathbf{r}_5 &= -LR \sin(u) \, dudv
 \end{aligned} \tag{5.4}$$

Finalmente, el factor de forma entre la placa rectangular y el disco está dado por la suma de las siguientes integrales dobles:

$$F_{pl \rightarrow disk} = \frac{1}{2\pi PL} \left(\int_0^1 \int_0^{2\pi} \ln(R_{15}) d\mathbf{r}_1 d\mathbf{r}_5 + \int_0^1 \int_0^{2\pi} \ln(R_{25}) d\mathbf{r}_2 d\mathbf{r}_5 \right. \\ \left. + \int_0^1 \int_0^{2\pi} \ln(R_{35}) d\mathbf{r}_3 d\mathbf{r}_5 + \int_0^1 \int_0^{2\pi} \ln(R_{45}) d\mathbf{r}_4 d\mathbf{r}_5 \right) \quad (5.5)$$

5.1.2. Calefactor cilíndrico y elemento diferencial

Un calefactor cilíndrico eléctrico, comúnmente utilizado en estudios previos como un firebrand idealizado para investigar el tiempo de retardo de ignición de lechos de combustible forestal mediante transferencia de calor radiativa [17, 18, 19], se emplea aquí como caso de estudio para evaluar los enfoques propuestos de aprendizaje automático y aprendizaje profundo. El objetivo principal es determinar la distribución espacial del flujo de calor radiativo incidente a lo largo del portamuestras.

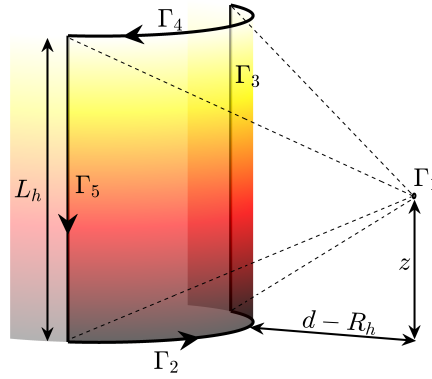


Figura 5.2: Esquema del calefactor cilíndrico (I-FIT)

La distribución espacial del flujo de calor incidente a lo largo del portamuestras (en la dirección z , como se puede observar en la Fig. 5.2) será representada mediante el factor de forma entre el heater cilíndrico eléctrico y un elemento diferencial del lecho de combustible forestal. Este elemento diferencial será representado como una pequeña circunferencia, de la siguiente manera:

$$\Gamma_1 : \mathbf{r}_1(u) = \left(\frac{p}{2} \cos(u), d, \frac{p}{2} \sin(u) + z + \frac{p}{2} \right), 0 \leq u \leq 2\pi \quad (5.6)$$

El heater eléctrico será descrito mediante los siguientes cuatro contornos:

$$\begin{aligned} \Gamma_2 : \mathbf{r}_2(v) &= (R_h \cos(v), R_h \sin(v), 0), \theta \leq v \leq \pi - \theta \\ \Gamma_3 : \mathbf{r}_3(v) &= (-R_h \cos(\theta), R_h \sin(\theta), L_h v), 0 \leq v \leq 1 \\ \Gamma_4 : \mathbf{r}_4(v) &= (R_h \cos(v), R_h \sin(v), L_h), \pi - \theta \leq v \leq \theta \\ \Gamma_5 : \mathbf{r}_5(v) &= (R_h \cos(\theta), R_h \sin(\theta), L_h - L_h v), 0 \leq v \leq 1 \end{aligned} \quad (5.7)$$

donde $\theta = \arctan \left(R_h / \sqrt{d^2 - R_h^2} \right)$ corresponde al campo de visión entre el elemento diferencial y el heater cilíndrico.

El producto punto entre los vectores tangentes $d\mathbf{r}_1 d\mathbf{r}_j$ está dado por la siguiente expresión:

$$\begin{aligned} d\mathbf{r}_1 d\mathbf{r}_2 &= d\mathbf{r}_1 d\mathbf{r}_4 = \frac{pR_h}{2} \sin(u) \sin(v) \, dudv \\ d\mathbf{r}_1 d\mathbf{r}_3 &= \frac{pL_h}{2} \cos(u) \, dudv \\ d\mathbf{r}_1 d\mathbf{r}_5 &= -\frac{pL_h}{2} \cos(u) \, dudv \end{aligned} \quad (5.8)$$

De esta manera, de acuerdo con la Ec. 5.1, el factor de forma entre el elemento diferencial y el heater cilíndrico está dado por las siguientes integrales dobles:

$$\begin{aligned} F_{dA_z \rightarrow h} &= \frac{2}{p^2 \pi^2} \left(\int_{\theta}^{\pi-\theta} \int_0^{2\pi} \ln(R_{12}) d\mathbf{r}_1 d\mathbf{r}_2 + \int_0^1 \int_0^{2\pi} \ln(R_{13}) d\mathbf{r}_1 d\mathbf{r}_3 \right. \\ &\quad \left. + \int_{\pi-\theta}^{\theta} \int_0^{2\pi} \ln(R_{14}) d\mathbf{r}_1 d\mathbf{r}_4 + \int_0^1 \int_0^{2\pi} \ln(R_{15}) d\mathbf{r}_1 d\mathbf{r}_5 \right) \end{aligned} \quad (5.9)$$

5.1.3. Cono calorímetro y elemento diferencial.

Desarrollado por el National Bureau of Standards (NBS) [39], el cono calorímetro (Cone Calorimeter) es ampliamente utilizado para cuantificar métricas fundamentales de inflamabilidad tales como la tasa de liberación de calor (HRR), el tiempo de ignición y las emisiones [40]. Estos datos son fundamentales para comprender el comportamiento del fuego y orientar el desarrollo de materiales resistentes al fuego.

Un estudio reciente introduce un modelo de contorno helicoidal (HC) para el factor de forma del calorímetro de cono [41], validado frente al modelo tradicional de cono truncado (TC) ([42, 43]) mediante integración analítica y experimentos con radiómetros. Este enfoque mejora la predicción espacial de la transferencia de calor radiativa y puede extenderse a materiales complejos y a diferentes configuraciones experimentales. La representación geométrica se muestra en la Fig. 5.3.

$$\begin{aligned} \Gamma_2 : \mathbf{r}_2(v) &= \left(f(v) \cos v, f(v) \sin v, \frac{Hv}{19\pi} \right), \quad 17\pi \leq v \leq 0, \\ \Gamma_3 : \mathbf{r}_3(v) &= \mathbf{r}_2(v), \quad 2\pi \leq v \leq 19\pi, \\ \Gamma_4 : \mathbf{r}_4(v) &= \left(\frac{D}{2} \cos v - R_1 - \frac{18R_2}{19}, 0, \frac{D}{2} \sin v + \frac{H}{19} \right), \quad \alpha - \pi \leq v \leq \alpha, \\ \Gamma_5 : \mathbf{r}_5(v) &= \left(\frac{D}{2} \cos v + R_1 + \frac{R_2}{19}, 0, \frac{D}{2} \sin v + \frac{18H}{19} \right), \quad \pi - \alpha \leq v \leq 2\pi - \alpha \end{aligned} \quad (5.10)$$

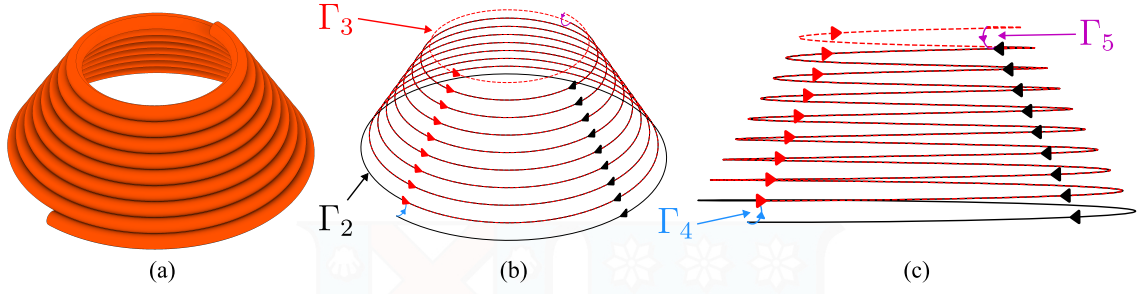


Figura 5.3: (a) Representación isométrica 3D de la geometría de la bobina helicoidal. (b) Vista isométrica de los contornos Γ_2 y Γ_4 correspondientes a la bobina helicoidal y la curva. Las orientaciones se representan mediante flechas. (c) Vista lateral de (b) mostrando los contornos Γ_4 y Γ_5 . Figura tomada de la Ref. [41].

En esta formulación, $\alpha = \arctan(H/R_2)$, la función lineal $f(v)$ y el diámetro del elemento del heater D se describen de la siguiente manera:

$$f_2(v) = -R_1 - R_2 + \frac{R_2 v}{19\pi} \text{ y } D = \frac{2\sqrt{R_2^2 + H^2}}{19} \approx 8 \text{ mm} \quad (5.11)$$

Finalmente, el factor de forma entre el sensor (de forma circular o disco) y el heater (cono calorímetro) está dado por la siguiente ecuación:

$$F_{disk \rightarrow c} = \frac{2}{\pi^2 p^2} \left(\int_{-17\pi}^0 \int_0^{2\pi} \ln(R_{12}) d\mathbf{r}_1 d\mathbf{r}_2 + \int_{2\pi}^{19\pi} \int_0^{2\pi} \ln(R_{13}) d\mathbf{r}_1 d\mathbf{r}_3 \right. \\ \left. + \int_{\alpha-\pi}^{\alpha} \int_0^{2\pi} \ln(R_{14}) d\mathbf{r}_1 d\mathbf{r}_4 + \int_{\pi-\alpha}^{2\pi-\alpha} \int_0^{2\pi} \ln(R_{15}) d\mathbf{r}_1 d\mathbf{r}_5 \right) \quad (5.12)$$

donde $d\mathbf{r}_1 d\mathbf{r}_j = \mathbf{r}'_1(u) \cdot \mathbf{r}'_j(v) du dv$ para $j = 2, \dots, 5$.

Es importante destacar que la suma de integrales presentada en Ecuación 5.12 se resuelve numéricamente con un error de integración despreciable, utilizando métodos adaptativos tanto en *Mathematica* como en *Python*, mediante técnicas de integración numérica adaptativa

5.2. Modelos sustitutos

Con el objetivo de evaluar la capacidad predictiva y la eficiencia computacional de los modelos propuestos, se analizaron tres configuraciones geométricas definidas previamente en los capítulos anteriores. Estas configuraciones permiten estudiar el desempeño de los modelos en distintos escenarios de transferencia de calor por radiación, considerando variaciones en la geometría del sistema y en las variables de entrada asociadas al cálculo del factor de forma.

- Caso 1: Placa rectangular y disco.
- Caso 2: Calefactor cilíndrico y elemento diferencial.
- Caso 3: Cono calorímetro y elemento diferencial.

En todos los casos, los valores de referencia del factor de forma fueron obtenidos mediante integración de contornos basada en el teorema de Stokes, lo que garantiza alta fidelidad numérica en la generación del dataset.

Para los resultados se analizarán diversos aspectos de cada modelo, que se enumeran a continuación.

- Coeficiente de determinación.
- Costo computacional.
- Capacidad de generalización.
- Estabilidad del modelo.

5.3. Caso 1: Placa rectangular y disco

El primer caso corresponde a la configuración más estructuralmente simple, aunque presenta dependencia no lineal entre el radio del disco y la distancia de separación.

Los modelos implementados en este caso son los siguientes:

- Regresión lineal múltiple con interacciones.
- regresión lineal múltiple regularizada.
- Support vector regressor.
- multilayer perceptron

Una vez analizados los cuatro modelos bajo la primera configuración, se seleccionarán los dos con mejor rendimiento para el entrenamiento bajo la segunda configuración.

5.3.1. Generación del dataset.

Se generó un conjunto de datos sintético fue generado mediante un proceso de integración numérica basado en integrales de contorno asociadas al cálculo de los factores de forma. Para cada configuración geométrica, el factor de forma fue calculado evaluando de manera iterativa una integral doble sobre las superficies radiativas involucradas. Este procedimiento fue implementado en Python y repetido para un gran número de combinaciones de parámetros (por ejemplo, radios y alturas), con el fin de garantizar una amplia cobertura del espacio de diseño.

El conjunto final de datos para este caso contiene 30.000 instancias y constituye la referencia (*ground truth*) para el entrenamiento y validación de los modelos de aprendizaje automático. Posteriormente, el dataset fue dividido en tres subconjuntos disjuntos: uno para entrenamiento, uno para validación y uno para prueba, siguiendo las prácticas estándar de *machine learning* con el objetivo de asegurar una evaluación robusta del modelo y una adecuada capacidad de generalización.

5.3.2. Construcción y entrenamiento de los modelos

En la siguiente tabla se observan los hiperparámetros encontrados mediante métodos de optimización y utilizados en el entrenamiento de cada uno de los modelos.

Tabla 5.1: Resumen de los hiperparámetros seleccionados para cada modelo evaluado, incluyendo el grado polinomial, la configuración del SVR y la arquitectura y configuración de entrenamiento del MLP para el caso de estudio 1.

Model	Hyperparameters
Polynomial	Degree = 3
Ridge	Degree = 8
SVR	kernel = rbf, $C = 598,95$, $\epsilon = 0,13$, $\gamma = 0,31$
MLP	Hidden layers: [64, 64, 32] Activation function: GELU Optimizer: AdamW, learning rate = 10^{-3} weight decay = 10^{-4} Number of training epochs: 300

5.4. Caso 2: Calentador cilíndrico y elemento diferencial

De acuerdo con el desempeño obtenido por los modelos en el Caso 1, se seleccionaron aquellos que presentaron los mejores resultados predictivos y computacionales. En particular, los modelos seleccionados fueron:

- Regresión polinomial múltiple con interacciones y regularización.
- Perceptrón multicapa (MLP).

5.4.1. Generación del dataset.

Para el desarrollo del Caso 2, se generó un conjunto de datos sintético compuesto por 40.000 observaciones, obtenidas a partir de la evaluación numérica de las integrales múltiples que definen el factor de forma para la configuración geométrica considerada. Este proceso fue implementado mediante métodos computacionales de integración numérica, permitiendo construir una base de datos representativa de la relación entre las variables geométricas del sistema y el valor correspondiente del factor de forma.

El conjunto de datos generado fue utilizado para entrenar y evaluar los modelos seleccionados a partir del análisis previo, correspondientes a la regresión múltiple con términos de interacción y regularización Ridge, y al perceptrón multicapa (MLP, *Multilayer Perceptron*). De esta manera, se busca comparar la capacidad de ambos modelos para aproximar el comportamiento del factor de forma, reduciendo el costo computacional asociado al cálculo directo mediante integrales múltiples.

5.4.2. Construcción y entrenamiento de los modelos

En esta etapa se llevó a cabo la construcción y entrenamiento de los distintos modelos considerados en este estudio. Para cada uno de ellos se determinó la configuración óptima de hiperparámetros mediante procedimientos de búsqueda sobre el conjunto de entrenamiento y validación.

Las configuraciones seleccionadas se presentan en las Tablas 5.2 y 5.3, donde se resumen los hiperparámetros óptimos obtenidos para cada modelo. En particular, para el modelo de regresión polinómica con regularización Ridge se determinó el grado del polinomio y el parámetro de regularización α , mientras que para el modelo basado en redes neuronales se definió la arquitectura de la red, incluyendo número de capas ocultas, cantidad de neuronas por capa y función de activación.

Estas configuraciones fueron posteriormente utilizadas para entrenar los modelos finales, los cuales se evaluaron mediante los conjuntos de validación y prueba con el fin de analizar su desempeño predictivo y capacidad de generalización.

Tabla 5.2: Hiperparámetros óptimos obtenidos para el modelo Polinomial con regularización Ridge utilizando BayesSearchCV para el Caso de Estudio 2.

Componente	Hiperparámetros	Mejor valor
Regresión polinomial	Grado	5
Ridge	α (L2)	1.25×10^{-2}

Tabla 5.3: Configuración óptima de hiperparámetros del modelo MLP con mejor desempeño en el Caso de Estudio 2.

Componente	Hiperparámetros	Mejor valor
Capa entrada	Neuronas	2
Capa oculta	Número de capas	3
Capa oculta	Neuronas por capa	84
Capa oculta	Función activación	LeakyReLU
Capa salida	Neuronas	1

5.5. Caso 3: Cono calorímetro y elemento diferencial

Finalmente, para el Caso 3 se seleccionó el modelo de perceptrón multicapa (MLP, *Multilayer Perceptron*), debido a que en los casos anteriores presentó el mejor desempeño tanto en la etapa de entrenamiento como en la de validación. Esta elección se fundamenta en su capacidad para aproximar relaciones no lineales complejas entre las variables geométricas del sistema y el factor de forma, manteniendo un adecuado nivel de precisión predictiva.

5.5.1. Generación del conjunto de datos

Para el Caso 3, correspondiente a la configuración del cono calorímetro y el elemento diferencial, se generó un conjunto de datos sintético mediante un marco numérico desarrollado en *Python*, específicamente diseñado para calcular el factor de forma entre un elemento diferencial y una superficie cónica. Esta geometría representa de manera realista el comportamiento radiativo característico de un cono calorímetro utilizado en ensayos experimentales de transferencia de calor por radiación.

Los valores de referencia del factor de forma fueron obtenidos a partir de la evaluación numérica de las integrales dobles asociadas a la formulación analítica presentada en la Sección 2. Para ello, se empleó la rutina `dblquad` de la biblioteca *SciPy*, la cual permite realizar integración adaptativa en dos dimensiones, asegurando una estimación precisa del valor de F_{ij} para cada combinación de parámetros geométricos considerada.

Las variables de entrada fueron muestreadas aleatoriamente dentro de rangos físicamente consistentes y representativos del dominio experimental del cono calorímetro. Este procedimiento permitió cubrir de manera adecuada el espacio paramétrico, garantizando diversidad geométrica en el conjunto de datos y facilitando que el modelo sustituto aprendiera con precisión la dependencia no lineal entre la con-

figuración geométrica y el comportamiento radiativo subyacente en todo el dominio de estudio.

5.5.2. Construcción y entrenamiento del modelo

La siguiente tabla presenta los hiperparámetros utilizados en el entrenamiento del modelo. Estos parámetros fueron determinados mediante un proceso de optimización automatizada empleando algoritmos de búsqueda eficientes implementados en la biblioteca Optuna, lo que permitió explorar el espacio de hiperparámetros y seleccionar una configuración que favorece un desempeño adecuado del modelo durante el proceso de entrenamiento.

Tabla 5.4: Hiperparámetros empleados en la arquitectura del modelo MLP

Hiperparámetros	Valores
Número capas ocultas	2
Neuronas por capa oculta	64
Función de activación	ReLU
Tasa de Dropout	– (none)
Normalización por lotes	No
Optimizador	Adam
Tasa de aprendizaje	$1,0 \times 10^{-3}$
Weight decay	0
Tamaño de lote	Full batch
Función de pérdida	MSE
Épocas	300

De acuerdo con la Tabla 5.4, entre los hiperparámetros más relevantes se encuentran el número de capas ocultas y la cantidad de neuronas por capa, los cuales determinan la capacidad representativa de la red. Asimismo, se especifican la función de activación utilizada, el algoritmo de optimización, la tasa de aprendizaje y la función de pérdida empleada durante el proceso de entrenamiento. En conjunto, estos hiperparámetros definen tanto la arquitectura de la red neuronal como la dinámica de su proceso de aprendizaje.

La configuración final fue obtenida mediante el algoritmo de optimización de hiperparámetros implementado en la biblioteca *Optuna*, el cual permite identificar la combinación de parámetros que minimiza el valor de la función de pérdida durante el entrenamiento del modelo.

6 | Resultados

En este capítulo se presentan y analizan los resultados obtenidos a partir de los distintos casos de estudio previamente definidos. Para cada caso se evalúa el desempeño de los diferentes modelos de *Machine Learning* y *Deep Learning* utilizados en este trabajo.

El análisis de los resultados permite comparar la capacidad predictiva de los modelos, así como su precisión y comportamiento frente a los datos generados mediante los métodos numéricos tradicionales. De esta manera, se busca evaluar el potencial de estas técnicas como herramientas para aproximar el cálculo de los factores de forma, reduciendo el costo computacional asociado a su evaluación directa.

6.1. Caso 1: Placa rectangular y disco

Los resultados que se presentan a continuación corresponden al desempeño obtenido por los modelos de aprendizaje supervisado considerados en este estudio. En particular, se evaluaron los siguientes modelos:

- Regresión lineal múltiple con términos de interacción.
- Regresión lineal múltiple con términos de interacción y regularización.
- Regresión de vectores de soporte (SVR, *Support Vector Regression*).
- Perceptrón multicapa (MLP).

6.1.1. Comparación de perfiles del factor de forma

Con el objetivo de analizar y comparar la capacidad de generalización de los distintos modelos utilizados, se generaron perfiles representativos para cada uno de ellos, los cuales se presentan a continuación.

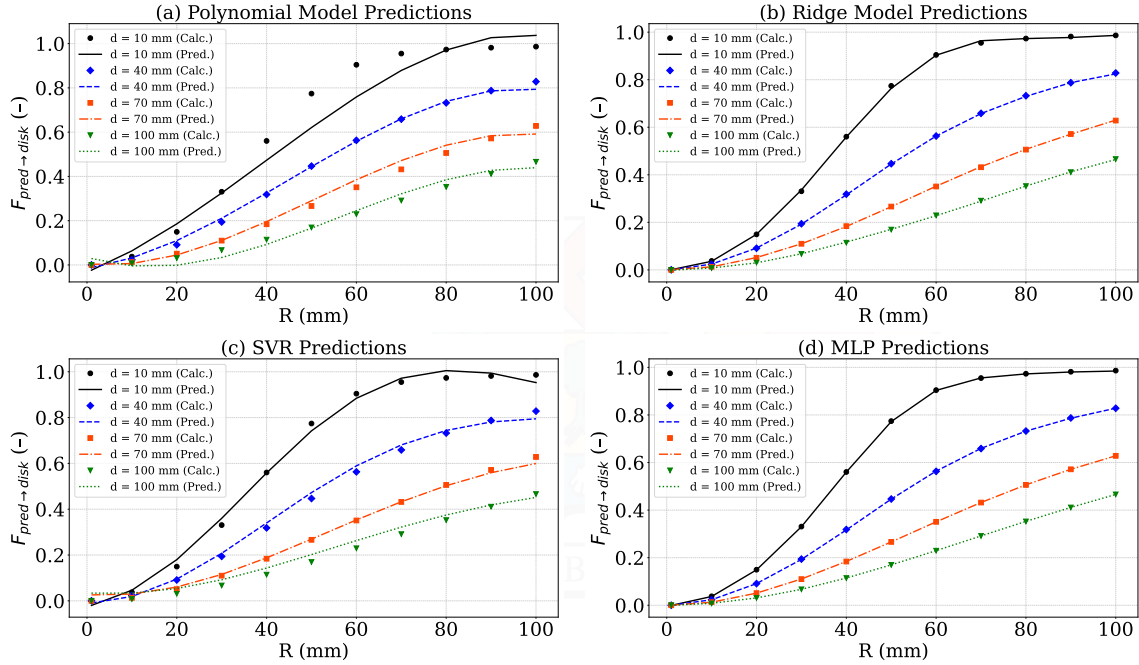


Figura 6.1: Factores de forma predichos desde la placa hacia el disco ($F_{pl \rightarrow disk}$) en función del radio (R) y la distancia (d). Los valores de VF calculados mediante el método analítico se muestran como puntos dispersos, mientras que los valores predichos se representan mediante líneas continuas para: (a) Regresión polinómica con términos de interacción cruzada, (b) Regresión polinómica con regularización Ridge, (c) Regresión por Vectores de Soporte (SVR), y (d) Perceptrón Multicapa (MLP).

6.1.2. Comparación del desempeño de los modelos predictivos

La Figura 6.1 presenta la comparación entre los valores del factor de forma $F_{pl \rightarrow disk}$ obtenidos mediante integración analítica por contornos y las predicciones generadas por los distintos modelos de aprendizaje automático evaluados: regresión polinomial con interacciones, regresión Ridge, Support Vector Regression (SVR) y red neuronal multicapa (MLP).

En cada subfigura se representa la variación del factor de forma en función del radio R del disco, para distintas separaciones d entre superficies (10, 40, 70 y 100 mm). Los puntos discretos corresponden a los valores calculados mediante el método integral, mientras que las líneas continuas representan las predicciones del modelo entrenado.

En la subfigura (a), correspondiente al modelo polinomial clásico, se observa que, si bien el modelo reproduce la tendencia general del fenómeno radiativo, presenta desviaciones apreciables en regiones de mayor radio y para distancias intermedias. Estas discrepancias evidencian la limitada capacidad del modelo para capturar la naturaleza fuertemente no lineal del mapeo geométrico subyacente, especialmente cuando no se incorpora regularización.

La subfigura (b), asociada al modelo Ridge, muestra una mejora sustancial en la aproximación. La inclusión del término de regularización L_2 permite estabilizar la estimación de coeficientes de alto orden, reduciendo el sobreajuste y mejorando la coherencia física de la superficie de regresión. Se aprecia una mayor suavidad y una mejor correspondencia con la solución analítica en todo el dominio evaluado.

En la subfigura (c), correspondiente al modelo SVR con kernel RBF, se observa una buena capacidad de generalización, capturando adecuadamente la forma global de las curvas. No obstante, en regiones de mayor curvatura o en valores extremos de R , el modelo tiende a suavizar excesivamente la respuesta, lo que genera pequeñas subestimaciones en comparación con la referencia analítica. Este comportamiento es característico de los métodos basados en kernels, donde existe un compromiso entre suavidad y precisión local.

Finalmente, la subfigura (d), correspondiente al modelo MLP, evidencia una concordancia prácticamente perfecta con los valores obtenidos por integración analítica. Las curvas predichas se superponen casi completamente con los puntos calculados, lo que indica que la red neuronal logró capturar de manera precisa la relación no lineal entre los parámetros geométricos y el factor de forma. Este comportamiento confirma la elevada capacidad de aproximación funcional de las redes neuronales profundas para modelar fenómenos físicos continuos y suaves en dominios compactos.

En conjunto, la figura demuestra que, si bien los modelos polinomiales y el SVR ofrecen aproximaciones razonables, la arquitectura MLP presenta el mejor desempeño predictivo, reproduciendo con alta fidelidad tanto la tendencia global como los detalles locales del comportamiento radiativo. Esto justifica su selección como modelo sustituto para los casos de estudio posteriores.

6.1.3. Comparación métricas de validación y prueba

A continuación se presentan dos tablas que contienen las métricas de desempeño obtenidas para los conjuntos de validación y prueba de cada uno de los modelos evaluados.

Tabla 6.1: Métricas de desempeño obtenidas en el conjunto de validación para los modelos evaluados en el Caso de Estudio 1, incluyendo MSE, RMSE, MAE, R^2 , MAPE y el porcentaje de predicciones con error inferior al 5 %.

Model	MSE (-)	RMSE (-)	MAE (-)	R^2 (-)	MAPE (%)	$\hat{y} < 5\%$ error (%)
Polynomial	$2,0856 \times 10^{-14}$	$1,4347 \times 10^{-7}$	$8,7645 \times 10^{-8}$	0.9664	7.04	62.31
Ridge	$5,3851 \times 10^{-16}$	$2,3083 \times 10^{-8}$	$1,3454 \times 10^{-8}$	0.9991	1.03	97.74
SVR	$2,8629 \times 10^{-15}$	$5,3536 \times 10^{-8}$	$4,5186 \times 10^{-8}$	0.9953	3.41	77.00
MLP	$2,9190 \times 10^{-17}$	$5,4030 \times 10^{-9}$	$3,8200 \times 10^{-9}$	0.9999	0.28	99.99

Tabla 6.2: Métricas de desempeño obtenidas en el conjunto de prueba para los modelos evaluados en el Caso de Estudio 1, incluyendo MSE, RMSE, MAE, R^2 , MAPE y el porcentaje de predicciones con error inferior al 5 %.

Model	MSE (-)	RMSE (-)	MAE (-)	R^2 (-)	MAPE (%)	$\hat{y} < 5\%$ error (%)
Polynomial	$2,0715 \times 10^{-14}$	$1,4393 \times 10^{-7}$	$8,8494 \times 10^{-8}$	0.9658	7.13	61.03
Ridge	$5,7275 \times 10^{-16}$	$2,3932 \times 10^{-8}$	$1,3725 \times 10^{-8}$	0.9995	1.06	97.59
SVR	$2,8517 \times 10^{-15}$	$5,3402 \times 10^{-8}$	$4,4871 \times 10^{-8}$	0.9953	3.46	77.31
MLP	$3,2992 \times 10^{-17}$	$5,7442 \times 10^{-9}$	$3,9550 \times 10^{-9}$	0.9999	0.31	99.99

6.1.4. Análisis de métricas de validación y prueba

Las Tablas 6.1 y 6.2 presentan las métricas de desempeño obtenidas durante las etapas de validación y prueba para los distintos modelos evaluados en el Caso 1. Las métricas consideradas incluyen el Error Cuadrático Medio (MSE), la Raíz del Error Cuadrático Medio (RMSE), el Error Absoluto Medio (MAE), el coeficiente de determinación (R^2), el Error Porcentual Absoluto Medio (MAPE) y el porcentaje de predicciones con error inferior al 5 %.

En la etapa de validación, el modelo de regresión polinomial clásica mostró el desempeño más bajo entre los modelos evaluados ($R^2 = 0,9664$, MAPE = 7.04 %), evidenciando limitaciones para capturar la fuerte no linealidad presente en la relación geométrica del factor de forma. Aunque el error absoluto es numéricamente pequeño debido a la escala del problema, el error relativo refleja una menor precisión comparativa.

La incorporación de regularización mediante Ridge produjo una mejora significativa en todas las métricas ($R^2 = 0,9991$, MAPE = 1.03 %), confirmando que la

penalización L_2 permite controlar la varianza del modelo y mejorar su capacidad de generalización sin comprometer la flexibilidad estructural. Asimismo, el porcentaje de predicciones con error inferior al 5 % aumentó a 97.74 %, lo que evidencia una mayor estabilidad predictiva.

El modelo SVR con kernel RBF alcanzó un desempeño intermedio ($R^2 = 0,9953$, $\text{MAPE} = 3.41\%$), logrando capturar adecuadamente la tendencia global de los datos. No obstante, la suavización inherente al kernel radial puede introducir pequeñas discrepancias en regiones de mayor curvatura, lo que se traduce en un leve incremento en el error porcentual respecto al modelo Ridge.

El modelo MLP presentó el mejor desempeño global durante la validación, con $R^2 = 0,9999$ y MAPE inferior a 0.3 %. Además, el 99.99 % de las predicciones se ubicaron dentro de un margen de error inferior al 5 %, lo que demuestra una capacidad sobresaliente para aproximar la función subyacente con alta fidelidad.

Los resultados obtenidos en el conjunto de prueba confirman la robustez de las conclusiones anteriores. El modelo polinomial mantiene un desempeño limitado ($R^2 = 0,9658$, $\text{MAPE} = 7.13\%$), mientras que el modelo Ridge conserva una elevada precisión ($R^2 = 0,9995$, $\text{MAPE} = 1.06\%$). El modelo SVR mantiene un comportamiento consistente con la validación ($R^2 = 0,9953$, $\text{MAPE} = 3.46\%$), evidenciando estabilidad pero con menor precisión relativa frente a los modelos regularizados y neuronales.

Finalmente, el modelo MLP vuelve a destacar en el conjunto de prueba ($R^2 = 0,9999$, $\text{MAPE} = 0.31\%$), con prácticamente la totalidad de las predicciones dentro del margen de error del 5 %. La consistencia entre validación y prueba indica que el modelo no presenta sobreajuste y posee una elevada capacidad de generalización.

En conjunto, el análisis comparativo de las métricas confirma que, aunque los modelos lineales con regularización ofrecen un desempeño competitivo, la arquitectura MLP proporciona la mejor combinación entre precisión, estabilidad y capacidad de generalización para la estimación del factor de forma en este caso de estudio.

6.1.5. Análisis de desempeño

Con el objetivo de analizar el comportamiento de entrenamiento y la capacidad de generalización del modelo seleccionado, se presentan las curvas de aprendizaje y el gráfico de ajuste ideal correspondientes al modelo MLP. Estas representaciones permiten evaluar la convergencia del proceso de optimización y la coherencia entre valores predichos y reales.

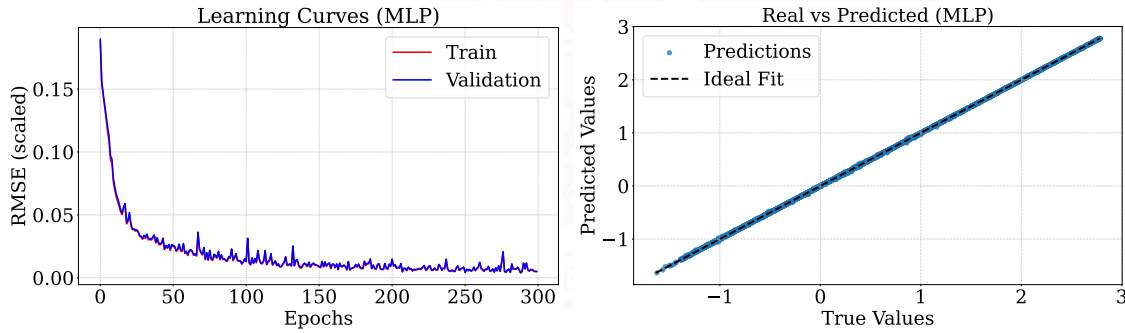


Figura 6.2: Curvas de aprendizaje de entrenamiento y validación (izquierda) y ajuste ideal (derecha) para el modelo MLP. Las curvas muestran la evolución de la pérdida RMSE a lo largo de 300 épocas, mientras que el histograma ilustra la frecuencia de los errores de predicción.

Como se observa en la Figura 6.2, la pérdida RMSE disminuye rápidamente durante las primeras épocas y posteriormente converge de manera estable, sin evidencias de sobreajuste, dado que la brecha entre entrenamiento y validación es mínima. Asimismo, el gráfico de ajuste ideal muestra una alineación prácticamente perfecta entre valores reales y predichos, lo que confirma la alta precisión y capacidad de generalización del modelo MLP.

6.1.6. Contraste entre valores reales y predicción.

Con el fin de evaluar la capacidad del modelo seleccionado para reproducir la distribución espacial del factor de forma, se presenta la comparación entre los valores obtenidos mediante el método integral y las predicciones generadas por el modelo MLP.

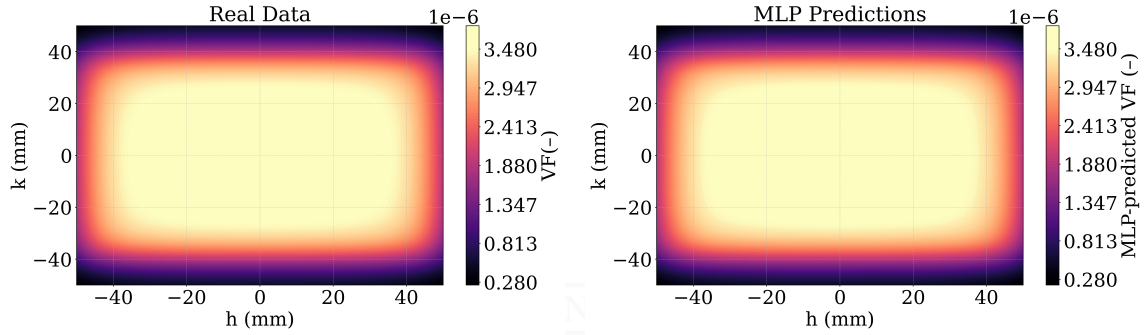


Figura 6.3: Mapas de calor de los valores del factor de forma (VF) del Caso 1: (izquierda) valores de referencia calculados mediante el método integral, y (derecha) predicciones obtenidas con el modelo MLP.

Como se aprecia en la figura, existe una concordancia prácticamente total entre la solución numérica (mediante integrales de línea) y la predicción del modelo neuronal. La distribución espacial, la simetría del fenómeno y la atenuación progresiva hacia los bordes son correctamente capturadas, lo que confirma la alta fidelidad del modelo MLP para aproximar el comportamiento físico subyacente.

6.1.7. Análisis de eficiencia computacional

La siguiente tabla muestra una comparación de los tiempos de cálculo obtenidos mediante la evaluación de las integrales de contorno y los modelos de aprendizaje automático empleados

Tabla 6.3: Comparación del costo computacional entre la evaluación mediante integrales de contorno y los modelos sustitutos basados en aprendizaje automático. El método de integrales de contorno se utiliza como referencia base para el cálculo del speed-up..

Método	Tiempo entrenamiento (s)	Tiempo de predicción por muestra (s)	Speed-up f/modelo
Integral de contorno	—	$6,91 \times 10^{-2}$	Baseline
Regresión polinomial	0.1431	$2,71 \times 10^{-6}$	$2,55 \times 10^4$
Ridge	0.1690	$2,02 \times 10^{-5}$	$3,42 \times 10^3$
SVR	58.2051	$1,46 \times 10^{-4}$	$4,73 \times 10^2$
MLP	176.779	$6,80 \times 10^{-7}$	$1,02 \times 10^5$

Los resultados muestran que los modelos de aprendizaje automático presentan tiempos de predicción significativamente menores en comparación con el método basado en integrales de contorno. En particular, el modelo MLP alcanza el menor tiempo de predicción por muestra, seguido por los modelos Ridge y SVR. En consecuencia, todos los modelos evaluados logran factores de aceleración (*speed-up*) de varios órdenes de magnitud respecto al método de referencia.

6.2. Caso 2: Calentador cilíndrico y elemento diferencial

Los resultados que se presentan a continuación corresponden al desempeño obtenido por los modelos de aprendizaje supervisado considerados en este estudio. En particular, los modelos que tuvieron mejor desempeño en el caso 1.

- Regresión lineal múltiple con términos de interacción y regularización.
- Perceptrón multicapa (MLP).

6.2.1. Comparación métricas de validación y prueba

A continuación se presentan dos tablas que contienen las métricas de desempeño obtenidas para los conjuntos de validación y prueba de cada uno de los modelos evaluados.

Tabla 6.4: Métricas de desempeño en el conjunto de validación para los modelos evaluados, incluyendo MSE, RMSE, MAE, R^2 , MAPE y el porcentaje de predicciones con error inferior al 5 % para el caso de estudio 2.

Model	MSE (-)	RMSE (-)	MAE (-)	R^2 (-)	MAPE (%)	$\hat{y} < 5\%$ error (%)
Ridge	$2,2417 \times 10^{-5}$	$4,7347 \times 10^{-3}$	$3,7080 \times 10^{-3}$	0.9988	2.53	89.75
MLP	$2,9981 \times 10^{-6}$	$1,7315 \times 10^{-3}$	$8,1633 \times 10^{-4}$	0.9998	0.58	99.08

Tabla 6.5: Resumen de las métricas de desempeño obtenidas en el conjunto de prueba para los modelos evaluados en el Caso de Estudio 2, incluyendo MSE, RMSE, MAE, R^2 , MAPE y el porcentaje de predicciones con error inferior al 5 %..

Model	MSE (-)	RMSE (-)	MAE (-)	R^2 (-)	MAPE (%)	$\hat{y} < 5\%$ error (%)
Ridge	$2,2417 \times 10^{-5}$	$4,7347 \times 10^{-3}$	$3,7080 \times 10^{-3}$	0.9988	2.49	88.12
MLP	$3,3350 \times 10^{-6}$	$1,8262 \times 10^{-3}$	$8,1546 \times 10^{-4}$	0.9998	0.58	99.12

6.2.2. Análisis métricas de validación y prueba

Las Tablas 6.4 y 6.5 presentan las métricas de desempeño obtenidas para el Caso 2, correspondiente a la configuración del calefactor cilíndrico y el elemento diferencial. En este escenario se comparan los modelos Ridge y MLP, seleccionados tras el análisis previo realizado en el Caso 1.

Durante la etapa de validación, ambos modelos exhiben un desempeño altamente satisfactorio, con coeficientes de determinación elevados ($R^2 = 0,9988$ para Ridge y $R^2 = 0,9998$ para MLP), lo que indica una excelente capacidad para capturar la relación no lineal entre los parámetros geométricos y el factor de forma. No obstante, el modelo MLP presenta errores significativamente menores en todas las métricas consideradas. En particular, el MAPE del MLP alcanza un valor de 0.58 %, en comparación con 2.53 % obtenido por Ridge, y el porcentaje de predicciones con error inferior al 5 % asciende a 99.08 %, superando el 89.75 % alcanzado por el modelo regularizado.

Los resultados obtenidos en el conjunto de prueba confirman la consistencia del comportamiento observado en validación. El modelo Ridge mantiene un desempeño robusto ($R^2 = 0,9988$, MAPE = 2.49 %), lo que demuestra una adecuada capacidad de generalización. Sin embargo, el modelo MLP vuelve a mostrar una superioridad clara, con $R^2 = 0,9998$ y un MAPE de 0.58 %, manteniendo además más del 99 % de las predicciones dentro del margen de error del 5 %.

La estabilidad de las métricas entre validación y prueba indica que ninguno de los modelos presenta sobreajuste significativo. Sin embargo, la reducción consistente del error absoluto y porcentual observada en el MLP confirma su mayor capacidad de aproximación funcional frente al modelo polinomial regularizado. En consecuencia, para el Caso 2, la arquitectura MLP se consolida como el modelo con mejor equilibrio entre precisión y generalización.

6.2.3. Análisis de desempeño

Para analizar el comportamiento de entrenamiento y la capacidad de generalización del modelo seleccionado, se presentan las curvas de aprendizaje y el gráfico de ajuste ideal del modelo MLP.

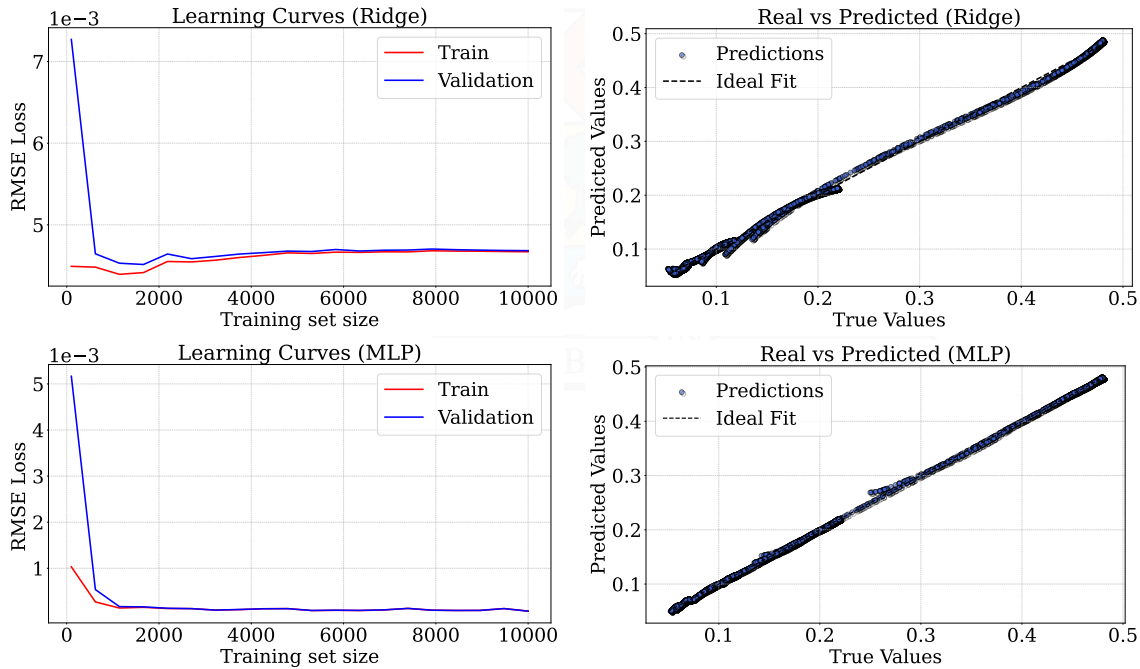


Figura 6.4: Curvas de aprendizaje y desempeño predictivo para el Caso 2. (a) Curvas de aprendizaje del modelo Ridge en función del tamaño del conjunto de entrenamiento, (b) Comparación entre valores predichos y valores reales para el modelo Ridge, (c) Curvas de aprendizaje del modelo MLP en función del tamaño del conjunto de entrenamiento, y (d) Comparación entre valores predichos y valores reales para el modelo MLP.

Con el objetivo de analizar el comportamiento de convergencia y la capacidad de generalización de los modelos evaluados en el Caso 2, se presentan las curvas de aprendizaje y los gráficos de ajuste ideal correspondientes a los modelos Ridge y MLP. Las curvas muestran la evolución de la pérdida RMSE en función del tamaño del conjunto de entrenamiento, mientras que los diagramas de dispersión permiten comparar directamente los valores reales del factor de forma con sus respectivas predicciones.

Como se observa en la figura, ambos modelos presentan una rápida reducción del error durante las primeras etapas de entrenamiento, seguida de una estabilización progresiva a medida que aumenta el tamaño de los datos. En el caso del modelo Ridge, se aprecia una convergencia estable con una brecha mínima entre entrenamiento y validación, lo que indica ausencia de sobreajuste significativo. Sin embargo, el modelo MLP alcanza niveles de error considerablemente menores y mantiene una mayor consistencia entre ambas curvas, evidenciando una mejor capacidad de aproximación.

En los gráficos de valores reales versus predichos, ambos modelos muestran una alineación cercana a la recta identidad. No obstante, el modelo MLP presenta una

dispersión prácticamente imperceptible respecto a la línea ideal, mientras que el modelo Ridge exhibe ligeras desviaciones en regiones de menor magnitud del factor de forma. Estos resultados confirman la superioridad del modelo MLP en términos de precisión predictiva y fidelidad en la representación del comportamiento físico subyacente.

6.2.4. Comparación de perfiles del factor de forma

Con el propósito de analizar el comportamiento espacial del factor de forma en el Caso 2, se presentan los perfiles de $F_{dA \rightarrow cyl}$ evaluados a distintas distancias radiales d y en función de la coordenada axial h . Esta comparación permite examinar la capacidad de los modelos para reproducir la distribución física del intercambio radiativo a lo largo del eje del sistema.

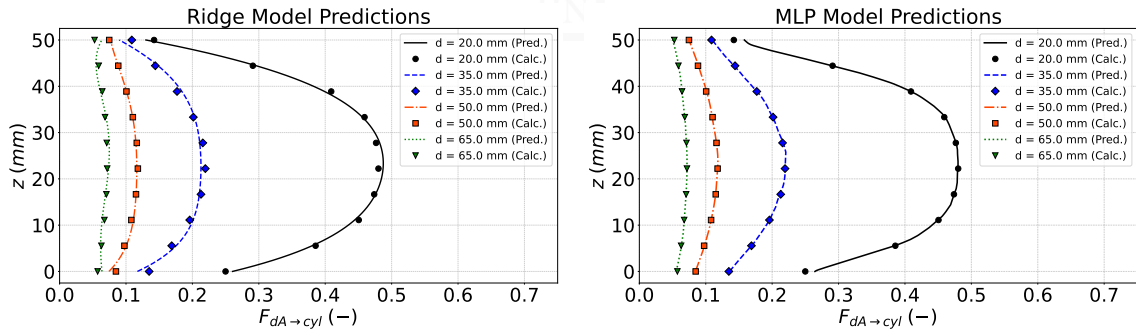


Figura 6.5: Perfiles del factor de forma ($F_{dA \rightarrow cyl}$) para el Caso de Estudio 2 a diferentes distancias radiales d , evaluados a lo largo de la coordenada axial h en el rango de 0–50 mm. El panel izquierdo muestra las predicciones obtenidas con el modelo de regresión Ridge, mientras que el panel derecho presenta los perfiles correspondientes estimados mediante el modelo MLP.

Como se observa en la figura, ambos modelos logran capturar la forma general de los perfiles del factor de forma, incluyendo la simetría axial y la variación progresiva con la distancia radial. No obstante, el modelo MLP reproduce con mayor precisión la amplitud y curvatura de las curvas en todo el rango de h , mientras que el modelo Ridge presenta ligeras desviaciones en regiones cercanas a los extremos axiales. Estos resultados refuerzan la superioridad del modelo neuronal para representar con mayor fidelidad la estructura no lineal del fenómeno radiativo en esta configuración geométrica.

Con el fin de cuantificar la eficiencia computacional de los modelos sustitutos propuestos, la Tabla 6.6 presenta una comparación entre el tiempo de evaluación del método analítico basado en integrales de contorno y los modelos Ridge y MLP. El método integral se adopta como referencia base para el cálculo del factor de aceleración (speed-up).

6.2.5. Análisis de eficiencia computacional

Para evaluar la eficiencia computacional de los modelos propuestos, se realizó una comparación entre el tiempo requerido para calcular los factores de forma mediante el método numérico tradicional y el tiempo necesario para obtener las predicciones utilizando los modelos de *Machine Learning*.

La Tabla 6.6 presenta el factor de aceleración (*speed-up*) obtenido para cada modelo, evidenciando la reducción significativa en el tiempo de cómputo al utilizar modelos de aprendizaje automático como aproximadores del cálculo numérico.

Tabla 6.6: Comparación del desempeño computacional entre la formulación analítica basada en integrales de contorno y los modelos sustitutos. El método de integrales de contorno se utiliza como referencia base para el cálculo del factor de aceleración (*speed-up*).

Método	Tiempo de entrenamiento (s)	Tiempo de predicción por muestra (s)	Speed-up f/modelo
Integral de contorno	—	$7,42 \times 10^{-2}$	Baseline
Ridge	0.0573	$9,0 \times 10^{-7}$	$8,24 \times 10^4$
MLP	47.6490	$4,68 \times 10^{-7}$	$1,58 \times 10^5$

Los resultados evidencian una reducción drástica del costo computacional al utilizar modelos de aprendizaje automático. Mientras que la evaluación mediante integración analítica requiere del orden de 10^{-2} segundos por muestra, los modelos Ridge y MLP alcanzan tiempos de inferencia del orden de 10^{-7} segundos, lo que se traduce en factores de aceleración de $8,24 \times 10^4$ y $1,58 \times 10^5$, respectivamente. Aunque el entrenamiento del MLP implica un mayor costo inicial, su tiempo de predicción es el más bajo, consolidándolo como la alternativa más eficiente para aplicaciones que requieren evaluaciones masivas o simulaciones en tiempo real.

6.3. Caso 3: Cono calorímetro y elemento diferencial

De acuerdo con los resultados obtenidos en los casos anteriores, el modelo que presentó el mejor desempeño global fue el perceptrón multicapa (MLP). Esta superioridad se observó tanto en términos de capacidad predictiva como en su comportamiento durante las etapas de entrenamiento y validación. Por esta razón, dicho modelo fue seleccionado para ser aplicado en la tercera configuración geométrica analizada.

La elección del MLP se justifica por su capacidad para aproximar relaciones no lineales complejas entre las variables geométricas del sistema y el factor de forma. En este sentido, su aplicación en el Caso 3 permite evaluar si el modelo mantiene un desempeño adecuado en una configuración geométrica distinta y potencialmente más compleja que las estudiadas previamente.

6.3.1. Comparación métricas de validación y prueba

En la siguiente tabla, se muestra una comparación entre las métricas de validación y testeo de la red neuronal dispuesta para el caso del cono calorímetro.

Tabla 6.7: Métricas de desempeño del modelo MLP en los conjuntos de validación y prueba, incluyendo MSE, RMSE, MAE, R^2 , MAPE y el porcentaje de predicciones con error inferior al 5 %.

Set	MSE (-)	RMSE (-)	MAE (-)	R^2 (-)	MAPE (%) (%)	$\hat{y} < 5\%$ (%)
Validation	$1,39 \times 10^{-5}$	$3,73 \times 10^{-3}$	$2,03 \times 10^{-3}$	0.9941	0.35	99.55
Test	$1,18 \times 10^{-5}$	$3,44 \times 10^{-3}$	$1,96 \times 10^{-3}$	0.9949	0.34	99.76

La Tabla 6.7 resume las métricas de desempeño obtenidas por el modelo MLP en los conjuntos de validación y prueba. Los resultados muestran valores muy bajos para las métricas de error (MSE, RMSE y MAE), lo que indica una alta precisión en la estimación del factor de forma. En particular, el coeficiente de determinación (R^2) alcanza valores cercanos a 1 en ambos conjuntos, evidenciando que el modelo explica prácticamente toda la variabilidad presente en los datos.

Asimismo, el error porcentual absoluto medio (MAPE) se mantiene por debajo del 1 %, lo que confirma la capacidad del modelo para realizar predicciones altamente precisas en todo el dominio de estudio. Finalmente, se observa que más del 99 % de las predicciones presentan un error inferior al 5 %, lo que demuestra la robustez y confiabilidad del modelo MLP para aproximar el comportamiento radiativo del sistema analizado.

6.3.2. Análisis de desempeño

Para evaluar el proceso de entrenamiento y la capacidad de generalización del modelo seleccionado, se presentan las curvas de aprendizaje y el gráfico de ajuste ideal correspondientes al modelo MLP.

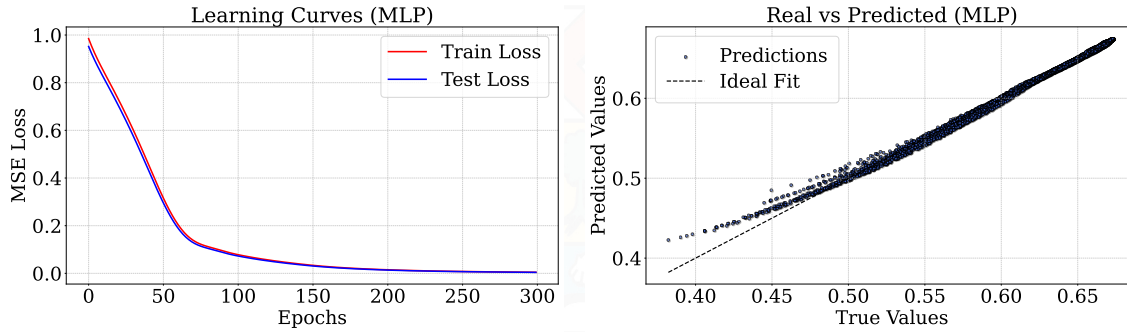


Figura 6.6: Curvas de aprendizaje (panel izquierdo) y comparación entre valores predichos y valores de referencia (panel derecho) del modelo MLP para el Caso de Estudio 3.

La Figura 6.6 presenta el desempeño del modelo de red neuronal multicapa (MLP) para el Caso 3. En el panel izquierdo se muestran las curvas de aprendizaje correspondientes a la pérdida durante el entrenamiento y la validación a lo largo de las épocas. Ambas curvas exhiben una disminución rápida en las primeras iteraciones, seguida de una convergencia progresiva hacia valores cercanos a cero, lo que indica que el modelo logra aprender adecuadamente la relación subyacente entre las variables de entrada y el factor de forma. Además, la cercanía entre las curvas de entrenamiento y prueba sugiere que el modelo presenta una buena capacidad de generalización y no evidencia problemas significativos de sobreajuste.

En el panel derecho se muestra la comparación entre los valores predichos por el modelo y los valores de referencia calculados mediante el método integral. La línea discontinua representa el ajuste ideal, correspondiente a una coincidencia perfecta entre predicción y valor real. Se observa que la mayoría de los puntos se distribuyen muy cerca de esta línea, lo que confirma que el modelo MLP logra reproducir con alta precisión los valores del factor de forma en el dominio analizado.

6.3.3. Contraste entre valores reales y predicción

Con el propósito de comparar los valores de referencia obtenidos mediante el método integral con las predicciones generadas por el modelo de aprendizaje automático, se construyeron dos mapas de calor que permiten visualizar la distribución espacial del factor de forma en el dominio de estudio.

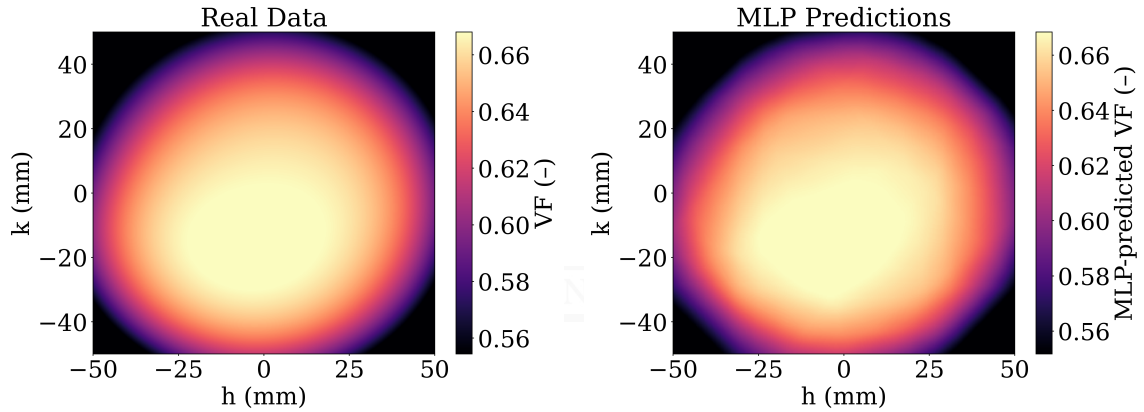


Figura 6.7: Mapas de calor de los valores del factor de forma (VF) para el Caso de Estudio 3: (izquierda) valores de referencia calculados mediante el método de integrales, y (derecha) predicciones obtenidas con el modelo MLP.

En la Figura 8.7 se presentan ambos mapas de calor. El panel izquierdo corresponde a los valores de referencia calculados numéricamente a partir de la formulación integral, mientras que el panel derecho muestra las predicciones obtenidas mediante el modelo MLP entrenado. La comparación visual entre ambos resultados evidencia una alta similitud en los patrones espaciales del factor de forma, observándose que las regiones de mayor y menor intensidad se reproducen de manera consistente en ambas representaciones. En particular, la distribución radial característica del fenómeno radiativo se mantiene en las predicciones del modelo, lo que indica que la red neuronal logra capturar adecuadamente el comportamiento físico subyacente del sistema. Estas similitudes confirman la capacidad del modelo para aproximar con precisión el campo del factor de forma dentro del dominio analizado.

6.3.4. Análisis de eficiencia computacional

En la siguiente tabla se presentan los tiempos de entrenamiento del modelo MLP y los tiempos de predicción por muestra, los cuales se comparan con el tiempo de cálculo requerido por el método basado en integrales de contorno.

Tabla 6.8: Comparación del costo computacional para el Caso de Estudio 3 entre la evaluación analítica basada en integrales de contorno y el modelo sustituto MLP. El método de integrales de contorno se adopta como referencia base para el cálculo del factor de aceleración (*speed-up*), y el valor reportado refleja la aceleración relativa alcanzada mediante el modelo sustituto MLP.

Método	Tiempo de entrenamiento (s)	Tiempo de predicción por muestra (s)	Speed-up $\int/\text{modelo}$
integral contorno	–	$3,61 \times 10^{-1}$	Baseline
MLP	15.48	$3,36 \times 10^{-7}$	$1,08 \times 10^6$

Los resultados muestran que, aunque el modelo MLP requiere un tiempo de entrenamiento inicial, su tiempo de predicción por muestra es varios órdenes de magnitud menor que el del método basado en integrales de contorno. En particular, el modelo sustituto logra un factor de aceleración del orden de 10^6 , lo que evidencia una reducción significativa del costo computacional al utilizar el enfoque basado en aprendizaje automático.

7 | Conclusiones

En este trabajo se presentó una metodología basada en aprendizaje automático para la estimación eficiente de factores de forma en transferencia de calor radiativa. A partir de valores de referencia obtenidos mediante formulaciones analíticas basadas en integrales de contorno derivadas del teorema de Stokes, se generaron conjuntos de datos sintéticos que permitieron entrenar diferentes modelos de aprendizaje supervisado para aproximar el comportamiento del factor de forma en distintas configuraciones geométricas.

Los resultados obtenidos demuestran que los modelos de aprendizaje automático son capaces de reproducir con alta precisión los valores de referencia calculados mediante integración numérica. En particular, el modelo de red neuronal multicapa (MLP) mostró el mejor desempeño entre todos los enfoques evaluados, alcanzando valores del coeficiente de determinación superiores a $R^2 > 0,99$ y errores porcentuales medios inferiores al 1 %, lo que evidencia su capacidad para capturar las relaciones no lineales entre los parámetros geométricos y el comportamiento radiativo del sistema. Estos resultados confirman que las redes neuronales constituyen aproximadores universales capaces de modelar funciones continuas complejas en dominios compactos [15].

Además de su alta precisión, los modelos desarrollados presentan ventajas computacionales significativas. Mientras que el cálculo directo de factores de forma mediante integración numérica puede resultar costoso desde el punto de vista computacional, especialmente cuando se requieren evaluaciones repetidas dentro de estudios paramétricos o procesos de optimización, los modelos sustitutos basados en aprendizaje automático permiten realizar predicciones prácticamente instantáneas una vez entrenados. En el presente estudio, se observaron mejoras de varios órdenes de magnitud en el tiempo de cómputo en comparación con el método integral de referencia, lo que coincide con resultados reportados en la literatura sobre el uso de modelos sustitutos en problemas de transferencia de calor y simulación física [44, 45].

Los mapas de calor generados a partir de las predicciones del modelo MLP mostraron una notable concordancia con las distribuciones espaciales obtenidas mediante el modelo analítico, reproduciendo adecuadamente las simetrías geométricas y las regiones de máxima intensidad radiativa. Esto demuestra que el modelo no solo logra aproximar valores puntuales con alta precisión, sino también capturar el comportamiento físico global del fenómeno radiativo.

Desde una perspectiva aplicada, estos resultados adquieren especial relevancia para el desarrollo de modelos computacionales simplificados orientados al análisis

rápido de fenómenos de propagación de incendios. En este tipo de problemas, la radiación térmica cumple un papel fundamental en el precalentamiento e ignición de combustibles cercanos, por lo que la estimación eficiente del intercambio radiativo entre llamas, superficies expuestas y materiales combustibles puede constituir un componente relevante dentro de modelos predictivos de propagación. En este sentido, los modelos sustitutos desarrollados en este trabajo no buscan reemplazar modelos físicos completos de dinámica del fuego, sino proporcionar una herramienta rápida y precisa para estimar uno de sus componentes térmicos fundamentales: el factor de forma asociado al intercambio radiativo.

En consecuencia, la metodología propuesta podría contribuir al desarrollo de herramientas computacionales de bajo costo capaces de evaluar rápidamente múltiples configuraciones geométricas y escenarios de exposición térmica. Esto resulta particularmente útil en aplicaciones donde se requiere explorar una gran cantidad de escenarios en tiempos reducidos, tales como análisis preliminares de riesgo, evaluación de zonas expuestas, estudios paramétricos, diseño de experimentos o integración con modelos simplificados de propagación de incendios. Bajo esta perspectiva, la estimación rápida de factores de forma mediante aprendizaje automático puede transformarse en un módulo auxiliar dentro de sistemas más amplios destinados a predecir, de manera aproximada y eficiente, la transferencia radiativa que favorece la propagación del fuego.

En conjunto, los resultados obtenidos evidencian que los modelos de aprendizaje automático, y en particular las redes neuronales profundas, constituyen una herramienta poderosa para construir modelos sustitutos capaces de reemplazar evaluaciones numéricas costosas en problemas de ingeniería térmica. Este enfoque abre nuevas oportunidades para la realización de estudios paramétricos, optimización de sistemas térmicos, análisis inversos y desarrollo de modelos computacionales sencillos orientados a la estimación rápida de procesos radiativos asociados a incendios.

Como trabajo futuro, sería interesante explorar arquitecturas de redes neuronales más avanzadas, como redes convolucionales o modelos basados en aprendizaje físico informado (Physics-Informed Neural Networks), así como ampliar el dominio geométrico de los casos de estudio para evaluar la capacidad de generalización del modelo en configuraciones más complejas. Asimismo, la integración de estos modelos dentro de herramientas de simulación térmica o modelos simplificados de propagación de incendios podría permitir acelerar significativamente procesos de diseño, análisis de riesgo y predicción rápida de escenarios en aplicaciones de ingeniería y ciencia del fuego.

Bibliografía

- [1] P. Higuera and J. Abatzoglou. Record-setting climate enabled the extraordinary 2020 fire season in the western united states. *Global Change Biology*, 27(1):1–2, 2020.
- [2] S. L. Manzello, R. Blanchi, M. J. Gollner, D. Gorham, S. McAllister, E. Pastor, E. Planas, P. Reszka, and S. Suzuki. Summary of workshop large outdoor fires and the built environment. *Fire Safety Journal*, 100:76–92, 2018.
- [3] CONAF. Estadísticas de incendios forestales en chile, 2023. Disponible en: <http://www.conaf.cl>.
- [4] F. Guerrero, C. Hernández, M. Toledo, L. Espinoza, Y. Carrasco, A. Arriagada, A. Muñoz, L. Taborga, J. Bergmann, and C. Carmona. Leaf thermal and chemical properties as natural drivers of plant flammability. *International Journal of Environmental Research and Public Health*, 18(13):7191, 2021.
- [5] O. Toy-Opazo, A. Fuentes-Ramirez, and V. Palma-Soto. Flammability features of native and non-native woody species. *Fire Ecology*, 20(21), 2024.
- [6] L. Zárate, J. Arnaldos, and J. Casal. Establishing safety distances for wildland fires. *Fire Safety Journal*, 43(8):565–575, 2008.
- [7] P. Pinto, M. Littin, J. I. Rivera, G. Severino, J. J. Cruz, and A. Fuentes. A simpler tractable contour technique to model thermal radiation from buoyant diffusion flames. *Experimental Thermal and Fluid Science*, 149:111027, 2023.
- [8] M. F. Modest. *Radiative Heat Transfer*. Academic Press, 3 edition, 2013.
- [9] J. R. Howell, M. P. Menguc, and R. Siegel. *Thermal Radiation Heat Transfer*. CRC Press, 2015.
- [10] E. M. Sparrow. Radiation view factors for arbitrary geometries. *International Journal of Heat and Mass Transfer*, 6(5):415–424, 1963.
- [11] A. Forrester, A. Sobester, and A. Keane. *Engineering Design via Surrogate Modelling: A Practical Guide*. Wiley, 2009.
- [12] S. Razavi, B. A. Tolson, and D. H. Burn. Review of surrogate modeling in water resources. *Water Resources Research*, 48(7):W07401, 2012.

- [13] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [14] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [15] K. Hornik. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.
- [16] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, 1989.
- [17] J. Rivera, N. Hernández, J.-L. Consalvi, P. Reszka, J. Contreras, and A. Fuentes. Ignition of wildland fuels by idealized firebrands. *Fire Safety Journal*, 120: 103036, 2021. doi: 10.1016/j.firesaf.2020.103036.
- [18] C. Álvarez, G. Moreno, F. Valenzuela, J. I. Rivera, F. Ebersperger, P. Reszka, and A. Fuentes. Use of an electric heater as an idealized firebrand to determine ignition delay time of Eucalyptus globulus leaves. *Fire Safety Journal*, 141: 103923, 2023. doi: 10.1016/j.firesaf.2023.103923.
- [19] M. Reveco, C. Álvarez, J. Gallardo, F. Valenzuela, G. Severino, A. Fuentes, P. Reszka, and R. Demarco. Effect of moisture content on the spotting ignition of live wildland fuels. *Proceedings of the Combustion Institute*, 40(1–4):105275, 2024. doi: 10.1016/j.proci.2024.105275.
- [20] S. Raschka, Y. Liu, and V. Mirjalili. *Machine Learning with PyTorch and Scikit-Learn*. Packt Publishing, 2022.
- [21] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [22] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer, 2009.
- [23] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2 edition, 2018.
- [24] D. C. Montgomery, E. A. Peck, and G. G. Vining. *Introduction to Linear Regression Analysis*. Wiley, 6 edition, 2021.
- [25] G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*. Wiley, 2 edition, 2003.
- [26] A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.
- [27] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer, 1995.
- [28] A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14:199–222, 2004.

- [29] V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.
- [30] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.
- [31] J. Marsden and A. Tromba. *Vector Calculus*. W. H. Freeman, 5 edition, 2003.
- [32] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 2006.
- [33] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [34] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 2013.
- [35] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *ICLR*, 2014.
- [36] G. B. Arfken, H. J. Weber, and F. E. Harris. *Mathematical Methods for Physicists*. Academic Press, 7 edition, 2013.
- [37] A. Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, 2 edition, 2019.
- [38] G. Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 5 edition, 2016.
- [39] V. Babrauskas. Development of the cone calorimeter—a bench-scale heat release rate apparatus based on oxygen consumption. *Fire and Materials*, 8(2):81–95, 1984. doi: 10.1002/fam.810080206.
- [40] V. Babrauskas. The cone calorimeter. In *SFPE Handbook of Fire Protection Engineering*, pages 952–980. Springer, 5 edition, 2016. doi: 10.1007/978-1-4939-2565-0_28.
- [41] P. E. Pinto, J. Valdivia, A. Singh, X. Xi, J. Cuevas, and J. L. Urban. Investigating radiation heat transfer from the cone calorimeter heater: A new view factor model and uncertainty quantification. *International Journal of Heat and Mass Transfer*, 245:126976, 2025. doi: 10.1016/j.ijheatmasstransfer.2025.126976.
- [42] M. T. Wilson, B. Z. Dlugogorski, and E. M. Kennedy. Uniformity of radiant heat fluxes in cone calorimeter. *Fire Safety Science*, 7:815–826, 2003. doi: 10.3801/IAFSS.FSS.7-815.
- [43] S. Kang, S. Choi, and J. Yoon. View factor in cone calorimeter testing. *International Journal of Heat and Mass Transfer*, 93:217–227, 2016. doi: 10.1016/j.ijheatmasstransfer.2015.09.067.
- [44] M. Raissi, P. Perdikaris, and G. Karniadakis. Physics-informed neural networks. *Journal of Computational Physics*, 378:686–707, 2019.

- [45] S. L. Brunton. Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics*, 52:477–508, 2020.



