



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

FACULTAD DE INGENIERÍA

CENTRO DE INVESTIGACIÓN Y ESTUDIOS DE POSGRADO

**Predicción del confort térmico en viviendas ejemplares en
México que integren materiales de cambio de fase utilizando
algoritmos de Machine Learning**

TESIS

Para obtener el grado de:
Maestro en Ingeniería Mecánica

Presenta:
Ing. Brayan Gerardo Gamboa Loya

Asesor:
Dra. Geydy Luz Gutiérrez Urueta
Co-Asesor:
Dr. Robert Jäckel



San Luis Potosí, S.L.P.

Febrero 2025



20 de junio de 2024

**ING. BRAYAN GERARDO GAMBOA LOYA
P R E S E N T E.**

En atención a su solicitud de Temario, presentada por los **Dres. Geydy Luz Gutiérrez Urueta y Robert Jäckel**, Asesora y Coasesor de la Tesis que desarrollará Usted, con el objeto de obtener el Grado de **Maestro en Ingeniería Mecánica**. Me es grato comunicarle que en la sesión del H. Consejo Técnico Consultivo celebrada el día 20 de junio del presente, fue aprobado el Temario propuesto:

TEMARIO:

“Predicción del confort térmico en viviendas ejemplares en México que integren materiales de cambio de fase utilizando algoritmos de Machine Learning”

Introducción.

1. Estado del arte: confort térmico en edificios, materiales de cambio de fase y Machine Learning.
2. Modelado y simulación de condiciones térmicas en viviendas mexicanas ejemplares.
3. Evaluación del confort térmico.
4. Desarrollo de un modelo predictivo del confort térmico usando Machine Learning.
5. Análisis y discusión de resultados.

Conclusiones.

Referencias.

“MODOS ET CUNCTARUM RERUM MENSURAS AUDEBO”

A T E N T A M E N T E

DR. EMILIO JORGE GONZÁLEZ GALVÁN
DIRECTOR



UNIVERSIDAD AUTÓNOMA
DE SAN LUIS POTOSÍ
FACULTAD DE INGENIERÍA
DIRECCION

www.uaslp.mx

Av. Manuel Nava 8
Zona Universitaria • CP 78290
San Luis Potosí, S.L.P.
tel. (444) 826 2330 al 39
fax (444) 826 2336

Copia. Archivo
*etn.

“UASLP, con profesionistas abiertos al mundo”



FACULTAD DE
INGENIERÍA



CENTRO DE
INVESTIGACIÓN
Y ESTUDIOS
DE POSGRADO

UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ
FACULTAD DE INGENIERÍA
Área de Investigación y Estudios de Posgrado

DECLARACIÓN

El presente trabajo que lleva por título:

“Predicción del confort térmico en viviendas ejemplares en México que integren materiales de cambio de fase utilizando algoritmos de Machine Learning”

se realizó en el periodo mayo de 2023 a diciembre de 2024 bajo la dirección de la Dra. Geydy Luz Gutiérrez Urueta y del Dr. Robert Jäckel.

Originalidad

Por este medio aseguro que he realizado el trabajo reportado, y la escritura de este documento de tesis para fines académicos sin ayuda indebida de terceros y sin utilizar otros medios más que los indicados.

Las referencias e información tomadas directa o indirectamente de otras fuentes se han definido en el texto como tales y se ha dado el debido crédito a las mismas.

El autor exime a la UASLP de las opiniones vertidas en este trabajo escrito y asume la responsabilidad total del mismo.

Este trabajo no ha sido sometido como tesis o trabajo terminal a ninguna otra institución nacional o internacional en forma parcial o total, exceptuando el caso cuando existe un convenio específico de doble titulación celebrado entre ambas instituciones.

Se autoriza a la UASLP para que divulgue este documento para fines académicos.

El autor del trabajo escrito, Brayan Gerardo Gamboa Loya

Universidad Autónoma de San Luis Potosí
Facultad de Ingeniería
Centro de Investigación y Estudios de Posgrado



Presenta:

Ing. Brayan Gerardo Gamboa Loya

Sinodales:

Dra. Geydy Luz Gutiérrez Urueta
Asesora

Dr. Robert Jäckel
Co-asesor

Dr. Dirk Frederik De Lange
Sinodal

Dr. Francisco Gerardo Pérez Gutiérrez
Sinodal

Dr. Ricardo Romero Méndez
Sinodal

San Luis Potosí, S.L.P., Febrero 2025

Agradecimientos

Quiero agradecer primeramente a mis padres Gerardo y Adriana. Sin todo el apoyo y amor que he recibido de su parte durante toda mi vida, esto no hubiera sido posible. Siempre estaré agradecido por todo lo que han hecho y siguen haciendo para que yo pueda lograr mis metas.

A mi hermana Taty por su apoyo y por siempre compartir momentos de alegría conmigo.

A mis abuelos, Petra y Ramón, por ser mis segundos padres y estar conmigo en todo momento.

A mis asesores, Dra. Geydy Gutiérrez y Dr. Robert Jäckel por contribuir en gran medida a mi formación. Su apoyo, dedicación y disposición me han convertido en un mejor profesional e hicieron posible el desarrollo de este trabajo.

A la Dra. Cintia Monreal por su gran ayuda y contribución que fueron fundamentales en este trabajo.

Al Dr. Guillermo Barrios Del Valle por su apoyo.

A los maestros que tuve durante el curso de este programa académico que contribuyeron enormemente a mi formación.

A mis compañeros Aliena, Jonathan y Sandra por su apoyo durante el transcurso del posgrado.

A mis amigos Martha, Yan y Arturo por estar acompañándome durante todo este proceso.

A mi pareja Ricardo por estar conmigo en todo este proceso. Gracias por el apoyo y todos los momentos que hicieron que este proceso fuera muy feliz para mí.

Al Consejo Nacional de Humanidades Ciencia y Tecnología (CONAHCyT) por el apoyo económico brindado para mis estudios de posgrado.

Gracias a todos por hacer esto posible.

Resumen

En este trabajo se presenta un análisis del confort térmico en viviendas mexicanas mediante la integración de materiales de cambio de fase (PCM) y técnicas de Machine Learning para su predicción. A través de simulaciones térmicas realizadas en EnergyPlus, se generó una base de datos robusta destinada al entrenamiento de un modelo predictivo capaz de evaluar el impacto de estos materiales en las condiciones de confort térmico.

Se presentan los resultados del análisis del confort térmico en una vivienda representativa mexicana, tanto en escenarios con, como sin, la incorporación de PCM en las diferentes regiones climáticas encontradas en México. Se evaluaron diversos índices de malestar térmico, como el Índice de Malestar (DI) y el Cambio Total de Malestar (TDC). Los resultados demuestran que, en la mayoría de los casos, el uso de PCM ayuda a atenuar la temperatura interior. Este efecto fue particularmente beneficioso en ciudades como Monterrey durante el verano y Xalapa en invierno. Sin embargo, en Tuxtla Gutiérrez en verano y Mérida en invierno, el uso de PCM tuvo un impacto negativo en el confort térmico.

Asimismo, se describen los modelos predictivos desarrollados mediante técnicas de Machine Learning, empleando diversas métricas para evaluar su desempeño. El algoritmo que con un mejor resultado en la validación de desempeño es el árbol de regresión empaquetado. Los resultados indican que los algoritmos utilizados son adecuados, incluyendo árboles de regresión y redes neuronales, en la predicción del índice de malestar térmico. Se discuten además las estrategias implementadas para la reducción de errores y el mejoramiento de la precisión de las predicciones.

Por último, se analizan las implicaciones prácticas de los resultados y la relevancia de la interfaz gráfica desarrollada, diseñada para facilitar la visualización de datos y resultados. La comparación entre los datos simulados y los predichos valida la efectividad de los modelos en la representación de las condiciones térmicas. Este trabajo proporciona un modelo predictivo útil para arquitectos e ingenieros en la optimización del diseño de viviendas.

Contenido

Agradecimientos	I
Resumen	III
Nomenclatura	XI
Introducción	1
Objetivos	2
Objetivo general	2
Objetivos específicos	2
Metodología	3
Estructura de la tesis	4
Capítulo 1. Estado del arte: confort térmico en edificios, materiales de cambio de fase y Machine Learning	5
1.1 Comfort térmico en edificios	5
1.1.1 Agentes de malestar térmico	5
1.1.2 Generalidades del confort térmico	7
1.1.3 Condiciones favorables para el confort térmico	10
1.1.4 Indicadores de malestar térmico	12
1.1.5 Grados día	12
1.2 Materiales de cambio de fase	14
1.3 Machine Learning	17
1.3.1 Modelos predictivos de Machine Learning	17
1.3.2 Modelos White, Black y Grey Box	18
1.3.3 Algoritmos de Machine Learning seleccionados para el procesamiento de datos	19
1.4 Antecedentes	21
Capítulo 2. Modelado y simulación de condiciones térmicas en viviendas mexicanas ejemplares	23
2.1 Funcionamiento del software de simulación térmica	23
2.2 Modelos numéricos de EnergyPlus	27
2.2.1 Modelo matemático de conducción con resistencias térmicas en EnergyPlus	27
2.2.2 Algoritmos para calcular el coeficiente de convección en EnergyPlus	28
2.2.3 Modelo de diferencias finitas en conducción (Módulo PCM)	28
2.3 Creación del modelo ejemplar de vivienda mexicana en EnergyPlus	29
2.3.1 Planteamiento del problema	29
2.3.2 Construcción del modelo de vivienda	30
2.3.3 Selección de los PCMs	31
2.3.4 Selección de las ciudades	32

2.4 Validación de las simulaciones numéricas	34
2.5 Resultados y discusión del capítulo	36
Capítulo 3. Evaluación del confort térmico	39
3.1 Cálculo del confort térmico	39
3.2.1 Índice de malestar (DI)	39
3.1.2 Cambio total de malestar (TDC)	40
3.1.3 Metodología para el cálculo del TDC	41
3.2 Resultados y discusión del capítulo	44
Capítulo 4. Desarrollo de un modelo predictivo del confort térmico usando Machine Learning	53
4.1 Introducción y planteamiento del problema	53
4.2 Metodología para el desarrollo de los modelos de predicción	54
4.3 Desarrollo y validación de los modelos predictivos	57
4.3.1 Métricas de desempeño	57
4.3.2 Evaluación de desempeño y reducción de error en los primeros modelos . . .	59
4.3.3 Evaluación de desempeño de los modelos con reducción de error	65
4.3.4 Contribución por variable al desempeño del modelo predictivo	69
Capítulo 5. Análisis y discusión de resultados	75
5.1 Desarrollo de la interfaz gráfica	75
5.1.1 Pre-tratamiento de los datos de entrada	77
5.1.2 Estructura de la interfaz gráfica	78
5.2 Presentación de resultados específicos	78
5.3 Comparación entre los datos simulados y predichos	81
Conclusiones	85
Anexo A: Datos requeridos para replicar las simulaciones con EnergyPlus	87
Anexo B: Datos climáticos de entrada	90

Lista de figuras

1.	Metodología seguida durante el trabajo de tesis.	4
1.1.	Representación gráfica de las entradas y salidas de calor en una habitación.	8
1.2.	Ejemplo de carta psicrométrica; recuperado de [10].	9
1.3.	Zona de confort térmico con velocidad de aire elevada, según la norma ASHRAE 55 [9].	10
1.4.	Diagrama de cambio de fase para un PCM.	14
2.1.	Apartado de Sitio en OpenStudio.	24
2.2.	Apartado de Horarios en OpenStudio.	24
2.3.	Apartado de Construcción en OpenStudio.	25
2.4.	Apartado de Cargas en OpenStudio.	25
2.5.	Apartado de Espacios en OpenStudio.	26
2.6.	Apartado de Zonas Térmicas en OpenStudio.	26
2.7.	Representación de una resistencia térmica de una capa de pared en un modelo de EnergyPlus [35].	28
2.8.	Plano de la casa modelo de INFONAVIT.	30
2.9.	Ejemplo de una casa estándar de bajo costo en México y los materiales utilizados para el modelo de vivienda.	31
2.10.	Curvas de entalpía-temperatura adaptadas de [34].	33
2.11.	Ciudades seleccionadas en México, basadas en las zonas climáticas de Köppen, para realizar las simulaciones.	34
2.12.	Comparación entre el modelo de OpenStudio con la edificación en [46].	35
2.13.	Comparación entre las temperaturas interiores experimentales y simuladas	35
2.14.	Comparación entre las temperaturas interiores en Monterrey (BS) durante el invierno con y sin PCM, cuya $T_f = 21^{\circ}\text{C}$	37
3.1.	Representación esquemática para la determinación del TDC con respecto a los intervalos de confort.	42
3.2.	Diagrama de flujo utilizado para calcular el TDC.	43
3.3.	Variación de las temperaturas interiores.	46
3.4.	Variación de las temperaturas interiores.	46
3.5.	Variación del DI por hora en Monterrey (BS) durante el invierno, comparando con y sin PCM, junto con los valores de DC.	48
3.6.	Horas totales de malestar ponderadas en verano e invierno.	49
3.7.	TDC del PCM en cada ciudad y estación del período de simulación.	50
3.8.	Cantidad de horas del periodo de simulación dentro de un mismo rango de DI para las ciudades con mejor desempeño.	50
3.9.	Cantidad de horas del periodo de simulación dentro de un mismo rango de DI para las ciudades con peor desempeño.	51
3.10.	Comparación de los valores de TDC por PCM y ciudad en el periodo de simulación.	52
4.1.	Diagrama de la metodología para el desarrollo de los modelos de predicción.	55
4.2.	Diagrama de flujo de la metodología del entrenamiento y desarrollo de los modelos con algoritmos de ML.	56
4.3.	Introducción de la base de datos en la aplicación de MATLAB.	57

4.4.	Procesamiento de la base de datos con la definición de variables y porcentaje para validación.	58
4.5.	Comparación de los valores de respuesta real y predicho del SVM con la primer base de datos.	60
4.6.	Gráfica de la respuesta del SVM con la primer base de datos.	61
4.7.	Correlación entre el DI y el punto de rocío.	61
4.8.	Correlación entre el DI y la radiación difusa.	62
4.9.	Correlación entre el DI y la humedad relativa.	62
4.10.	Correlación entre el DI y la presión atmosférica.	63
4.11.	Correlación entre el DI y los grados día de calefacción.	64
4.12.	Correlación entre el DI y los grados día de refrigeración.	64
4.13.	Comparación del desempeño entre los algoritmos utilizados.	65
4.14.	Gráficas de regresión lineal y de ensamble.	66
4.15.	Gráficas de regresión para el SVM y la red neuronal.	67
4.16.	Comparación de los valores de respuesta real y predicho del SVM con la segunda base de datos.	67
4.17.	Gráfica de residuos para el ensamble de árboles empaquetados.	68
4.18.	Dependencia parcial entre los CDD y la respuesta de predicción.	70
4.19.	Dependencia parcial entre la humedad relativa y la respuesta de predicción.	71
4.20.	Dependencia parcial entre los HDD y la respuesta de predicción.	71
4.21.	Dependencia parcial entre la hora del día y la respuesta de predicción.	72
4.22.	Dependencia parcial entre el mes del año y la respuesta de predicción.	72
4.23.	Dependencia parcial entre la presión atmosférica y la respuesta de predicción.	72
4.24.	Dependencia parcial entre el punto de rocío y la respuesta de predicción.	73
4.25.	Dependencia parcial entre la temperatura de PCM y la respuesta de predicción.	73
4.26.	Dependencia parcial entre la radiación difusa y la respuesta de predicción.	74
4.27.	Dependencia parcial entre las zonas climáticas Köppen y la respuesta de predicción.	74
5.1.	Diagrama de flujo utilizado para calcular el TDC.	76
5.2.	Comparación entre las curvas de CDD en un día para Tuxtla Gutiérrez (Af) en verano	77
5.3.	Comparación entre las curvas de presión atmosférica en un día para Tuxtla Gutiérrez (Af) en verano	78
5.4.	Interfaz gráfica.	79
5.5.	Comparación del DI de predicción y base de datos para las ciudades con menor porcentaje de error relativo por estación.	80
5.6.	Comparación del DI de predicción y base de datos para las ciudades con mayor porcentaje de error relativo por estación.	81
5.7.	TDC del PCM 21 en cada ciudad y estación, comparando los valores de predicción y simulación.	83
6.1.	Observación de los datos de entrada en un archivo climático con Dview.	88
6.2.	Datos tabulados de entalpía-temperatura en EnergyPlus (parte 1).	89
6.3.	Datos tabulados de entalpía-temperatura en EnergyPlus (parte 2).	89

Lista de tablas

1.	Índices directos basados en temperaturas de bulbos seco y húmedo; recuperada de [4].	13
2.	Materiales utilizados y sus propiedades termo-físicas.	31
3.	Horarios de ocupación y parámetros considerados para el modelo de vivienda. . . .	32
4.	Ciudades seleccionadas para el estudio.	33
5.	Materiales utilizados y sus propiedades termofísicas para las simulaciones de validación.	36
6.	Criterio de confort con aumento de resolución según el índice de malestar.	40
7.	Comparación de los valores de TDC por ciudad y por PCM.	44
8.	Comparación de los valores de horas ponderadas de malestar por ciudad y por PCM.	45
9.	Tiempo requerido de cada algoritmo para completar el entrenamiento con la base de datos.	57
10.	Métricas de desempeño obtenidas con los modelos entrenados con la primer base de datos.	59
11.	Resultados obtenidos con los modelos entrenados con la base de datos para reducir el error.	65
12.	Valor de MRMR para cada variable de entrada.	69
13.	Comparación de los valores de TDC por ciudad utilizando los datos de predicción. .	79
14.	Comparación de los valores de TDC predichos y simulados por ciudad y estación. .	82
15.	Datos climáticos de entrada utilizados en el proceso de validación para la configuración A.	90
16.	Datos climáticos de entrada utilizados en el proceso de validación para la configuración B.	91
17.	Comparación de los resultados simulados y experimentales para la configuración A.	92
18.	Comparación de los resultados simulados y experimentales para la configuración B.	93

Nomenclatura

HVAC	Heating, Ventilation and Air Conditioning (calefacción, ventilación y aire acondicionado).
PCM	Phase change material (material de cambio de fase).
SENER	Secretaría de energía.
ASHRAE	Sociedad Americana de Ingenieros de Calefacción, Refrigeración y Aire Acondicionado.
DI	Índice de malestar.
CDD	Grados día de refrigeración.
HDD	Grados día de calefacción.
ML	Machine Learning.
SVM	Máquina de soporte vectorial.
CondFD	Diferencias finitas en conducción.
RT	Rubitherm.
INFONAVIT	Instituto del Fondo Nacional de la Vivienda para los Trabajadores.
TDC	Cambio total de malestar.
MSE	Error cuadrático medio.
RMSE	Raíz del error cuadrático medio.
MAE	Error absoluto medio.

Introducción

Con el auge de las formas de vida modernas, el confort térmico ha cobrado mayor importancia en el desarrollo de tecnología de la construcción, orientada a establecer un estándar de comodidad en los ambientes interiores. El concepto de confort térmico está relacionado con las entradas y salidas de energía térmica en un edificio, así como con las fuentes de calor internas. Existen varios factores que lo afectan, entre ellos: factores personales (metabolismo y vestimenta), factores ambientales (temperatura del aire, humedad y radiación solar) y factores contribuyentes (como la forma corporal de una persona). En una edificación, hay interacciones de energía térmica debido a la radiación solar que penetra por las ventanas, debido a la ventilación y particiones (principalmente paredes y ventanas); y también acumulación de energía térmica en el interior [1].

Los principales fenómenos naturales que afectan los cambios de temperatura en el interior de un inmueble son: transferencia de calor entre el aire interior y exterior, fugas o entradas de aire, radiación solar y generación de calor interna debido a las personas, aparatos electrónicos e iluminación [1]. En la actualidad, los sistemas HVAC (Heating, Ventilation and Air Conditioning) cubren el problema de los climas extremos que existen en varias regiones del planeta. En México, del consumo final total de energía en 2021, el 16.77 % es de consumo residencial y comercial [2]. En el último censo de población y vivienda del año 2020 se determinó que más del 17 % de los hogares mexicanos, que equivale a más de 6 millones de viviendas, cuentan con al menos un dispositivo de acondicionamiento de aire [3]. Al ser extremas las temperaturas en diferentes regiones del país, la calidad de vida de los habitantes se ve afectada notoriamente y suele tener consecuencias en el consumo de energía debido al uso de sistemas HVAC u otros aparatos o sistemas de acondicionamiento. Es importante mencionar la influencia que los materiales de cambio de fase (PCM) tienen en las condiciones térmicas de una vivienda. La utilización de PCM en la construcción puede sustituir en cierta medida la utilización de sistemas HVAC, al poder almacenar y liberar energía latente cuando las temperaturas ambientales se encuentran encima y debajo de su temperatura de fusión, respectivamente.

Por otra parte, en los últimos años la ciencia de datos y la inteligencia artificial han tenido un impacto y desarrollo sin precedentes, con lo que se han podido desarrollar modelos predictivos en diferentes industrias y áreas del conocimiento. La inteligencia artificial, en auge en diversas áreas de la investigación, se ha comenzado a usar para crear modelos que predigan variables relacionadas con la carga térmica dentro de edificios. Se espera que estos modelos puedan proporcionar predicciones en la utilización de energía, el gasto de energía con sistemas HVAC o la acumulación

de energía. La integración de modelos centrados en los datos y modelos centrados en la física ha comenzado a consolidarse. Esta metodología consiste en utilizar la modelación matemática y las simulaciones térmicas para proporcionar datos que alimenten algoritmos de Machine Learning capaces de predecir diversos aspectos relacionados con la carga térmica, como es el confort térmico. En el contexto mexicano, la simulación térmica para la creación de modelos predictivos es prácticamente inexistente. El desarrollo de este tipo de investigación sería de gran utilidad, ya que las temperaturas extremas en diversas regiones del país afectan la calidad de vida de los habitantes y repercuten en el consumo de energía debido al uso de sistemas HVAC.

A pesar de la existencia de modelos de predicción de carga térmica en edificios, este es un campo poco explorado, especialmente en el contexto mexicano. Esta tesis propone desarrollar un modelo para predecir las condiciones de confort térmico en viviendas mexicanas. El modelo será entrenado con diversos parámetros que afectan la carga térmica interior por región, con el fin de estimar el confort térmico del edificio y evaluar cómo cambiaría si se incorpora PCM. Esto incluye la simulación térmica de viviendas mexicanas ejemplares basada en modelos físicos para crear una base de entrenamiento, la selección y entrenamiento de un algoritmo de orden reducido, la validación del modelo de predicción y el desarrollo de una interfaz gráfica. Esta metodología permite una predicción más rápida y con menor costo computacional en comparación con la simulación física y numérica, ya que un modelo de Machine Learning reduce el orden y, por lo tanto, la complejidad del problema. Esta herramienta puede contribuir al desarrollo y mejora de tecnologías de construcción.

Objetivos

Objetivo general

Predecir las condiciones de confort térmico en viviendas ejemplares mexicanas que integren materiales de cambio de fase utilizando algoritmos de Machine Learning.

Objetivos específicos

- Revisar el estado del arte sobre el uso de materiales de cambio de fase (PCM) para mejorar el confort térmico en viviendas y sobre algoritmos de predicción de Machine Learning.
- Desarrollar un modelo que permita la simulación de las condiciones térmicas en viviendas ejemplares en México, considerando tanto escenarios con el uso de materiales de cambio de fase como sin ellos.
- Validar el modelo a través de experimentación o datos disponibles en la literatura científica.
- Emplear herramientas de simulación de confort térmico y, una vez validados sus resultados,

aplicarlos a efectos de la construcción de una base de datos de entrenamiento.

- Elegir un algoritmo de predicción en Machine Learning de orden reducido para predecir el confort térmico con base en los parámetros principales que gobiernan el problema.
- Entrenar y validar el algoritmo seleccionado de predicción en Machine Learning proporcionando casos simulados con y sin materiales de cambio de fase (PCM).
- Desarrollar una interfaz gráfica usando los parámetros principales como entrada para obtener la predicción del confort térmico.

Metodología

La metodología empleada en cada una de las etapas de este trabajo es la siguiente:

- Estudio de la literatura: herramientas de simulación de condiciones térmicas, análisis de parámetros climáticos y de construcción en diferentes entidades mexicanas, materiales de cambio de fase para el control de condiciones térmicas, medición del confort térmico, y métodos de reducción de orden basados en Machine Learning.
- Adquisición de habilidades en el uso de la herramienta de modelación EnergyPlus.
- Realización de simulaciones de las condiciones térmicas en viviendas mexicanas ejemplares en EnergyPlus, tanto con como sin el uso de materiales de cambio de fase.
- Validación de la capacidad de predicción de la herramienta de simulación térmica EnergyPlus a través de datos proporcionados por la literatura científica.
- Realización de cálculos de confort térmico utilizando los resultados de temperatura y humedad relativa, mediante un índice directo y una clasificación de malestar térmico.
- Creación de una amplia base de datos que relacione las características principales del problema con el confort térmico. Esta base de datos se utilizará como conjunto de entrenamiento para desarrollar un modelo de orden reducido basado en algoritmos de Machine Learning.
- Selección de modelos basados en Machine Learning, adecuados para ser entrenados con la base de datos creada, utilizando el Toolbox de Machine Learning y estadística en MATLAB.
- Entrenamiento y validación de un modelo supervisado para predecir el confort térmico en función de los principales parámetros que influyen en el comportamiento térmico de la vivienda.

Un resumen de la metodología se muestra en la Fig. 1.

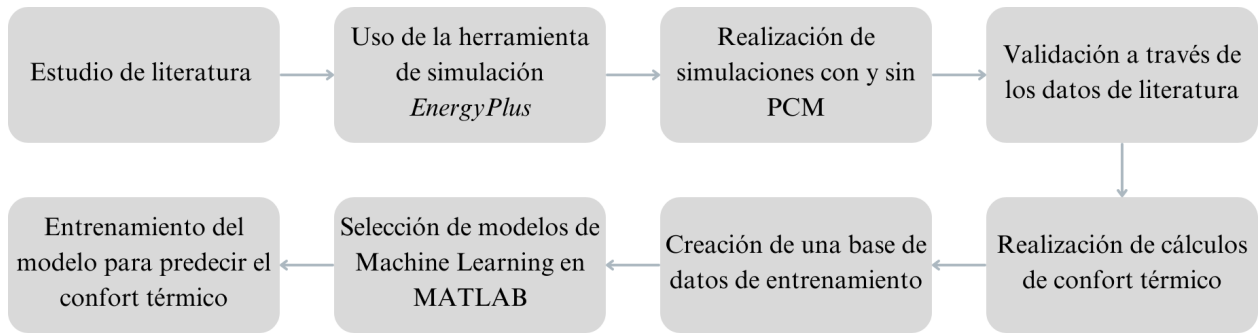


Figura 1: Metodología seguida durante el trabajo de tesis.

Estructura de la tesis

Este trabajo se encuentra organizado en 5 capítulos:

En el primer capítulo se presenta la revisión bibliográfica de los temas más relevantes para la realización de este trabajo.

Dado que este trabajo se basa en la simulación de las condiciones térmicas en México, en el Capítulo 2 se describen las condiciones climáticas y el tipo de construcción que influyen en la carga térmica de una vivienda mexicana estándar. Asimismo, se presenta el modelado de estas condiciones en EnergyPlus, el funcionamiento de esta herramienta de simulación y la realización de las simulaciones térmicas.

Las ecuaciones y las categorías de malestar térmico utilizadas para realizar los cálculos de confort térmico se muestran en el Capítulo 3, con el objetivo de comparar este malestar en el modelo de vivienda ejemplar mexicana, tanto con como sin PCM en su estructura.

En el Capítulo 4 se presenta la metodología detallada para la creación de modelos predictivos utilizando la herramienta de Machine Learning en MATLAB, incluyendo los procesos de entrenamiento, desarrollo y validación de estos modelos con diferentes algoritmos de aprendizaje supervisado.

Los resultados obtenidos en este trabajo, que incluyen el desarrollo de una interfaz gráfica con pretratamiento de datos, los valores de predicción del índice de malestar y la comparación entre los datos obtenidos mediante simulación y los modelos de predicción, se describen y analizan en el Capítulo 5.

Capítulo 1

Estado del arte: confort térmico en edificios, materiales de cambio de fase y Machine Learning

En diversos contextos de la investigación científica y tecnológica, se está comenzando a descubrir el potencial de las herramientas computacionales para fines de predicción. Es relevante desarrollar un modelo de predicción del confort térmico en una vivienda mexicana estándar, tanto con como sin PCM en su configuración, alimentado por datos obtenidos de simulaciones térmicas y modelos matemáticos. Para llevar a cabo esta tarea, es necesario comenzar revisando la literatura científica más relevante sobre confort térmico, materiales de cambio de fase y modelos de predicción.

1.1 Confort térmico en edificios

1.1.1 Agentes de malestar térmico

La temperatura ambiente por sí sola no siempre es el único factor que causa el malestar térmico [4]. Existen factores que definen el confort térmico en el ambiente térmico humano.

Los primeros son los *factores de comportamiento*, que se dividen en la tasa de metabolismo y la vestimenta, que afectan la respuesta humana al ambiente térmico:

- *Tasa de metabolismo.* Se refiere a la actividad química en las células que mantiene la temperatura corporal en 37°C , debido a la quema de alimentos como proteína, grasa y carbohidratos. La tasa de metabolismo basal es la requerida para que el cuerpo mantenga las funciones necesarias como la respiración y la circulación de sangre cuando el cuerpo se encuentra en reposo. La tasa de metabolismo aumenta con el nivel de actividad. Esta conversión que realiza el cuerpo de la energía química de la comida luego es disipada al ambiente; una persona que está realizando actividad física intensa puede aumentar más de 10 veces su tasa de metabolismo, proveyendo a una habitación de una cantidad de energía similar a la de un calentador de 1 kW [6]. La tasa de metabolismo se puede expresar por medio de la unidad **met**, que equivale a $58\text{ W}/\text{m}^2$.

- *Vestimenta.* La ropa es un factor muy importante para la disipación de calor. Existe una unidad llamada **clo** que corresponde a la cobertura aislante sobre todo el cuerpo, y equivale a una resistencia de $0.155 \text{ m}^2\text{K/W}$. 1 clo es el valor aislante de un traje de negocios normal y ropa interior de algodón; pantalones cortos con camisa de manga corta corresponde a 0.25 clo, y un traje de invierno con chaqueta equivale a 2 clo [1].

El resto de los factores que influyen en el confort térmico son los *factores ambientales*, que se dividen en la temperatura (ambiente y radiante), humedad y movimiento del aire:

- *Temperatura ambiental.* Es el indicador más importante de confort térmico. Como se mencionó en la sección 1.1, se han realizado múltiples trabajos de investigación para determinar una zona de confort térmico [6]. Este factor es medido por la temperatura de bulbo seco, determinando la disipación de calor convectivo junto con el movimiento del aire [1].
- *Temperatura radiante.* Los intercambios de radiación van a depender de la temperatura media de las superficies circundantes [1].
- *Humedad.* Se puede expresar como humedad relativa, humedad absoluta o contenido de humedad. Como se mencionó en la sección 1.1, la humedad relativa es la medida de la habilidad del aire de absorber humedad y afecta la cantidad de calor que un cuerpo puede disipar por evaporación. Un nivel alto de humedad relativa hace lento el proceso de rechazo de calor por evaporación, mientras que un nivel bajo de humedad relativa hace más rápido el proceso. El rango deseable de humedad relativa es entre 30 y 70 % [6].
- *Movimiento del aire.* Es medido por la velocidad y también afecta la disipación de calor por evaporación [1]. El malestar debido a este fenómeno es común para las personas que utilizan vestimenta ligera y que tienen una actividad sedentaria, y menos común en las personas que tienen un nivel alto de actividad. La velocidad del aire debe ser menor de 9 m/min en invierno y de 15 m/min en verano para minimizar el malestar por movimiento del aire. Se espera un nivel bajo de movimiento para que remueva el aire caliente y húmedo, y sea reemplazado por aire fresco [6].

El confort térmico es definido como la condición de la mente que expresa satisfacción con el ambiente térmico, ya que la comodidad es una sensación subjetiva [4]. Debido a que hay bastantes variaciones de persona a persona, tanto fisiológicas como psicológicas, es complicado que el ambiente térmico satisfaga a todas las personas dentro de un espacio [5]. Sin embargo, en la norma ASHRAE 55 [5] se menciona que se ha recolectado la suficiente información en laboratorios y campo que provee de la información estadística necesaria que define las condiciones en las que un porcentaje específico de personas dentro de un recinto se sentirán cómodos térmicamente.

1.1.2 Generalidades del confort térmico

A diferencia de otras especies, el ser humano nace con poca protección corporal contra las condiciones ambientales extremas del planeta. Por esta razón, la búsqueda del confort térmico ha sido una constante desde los inicios de la humanidad. El término **aire acondicionado** se utiliza comúnmente para referirse al enfriamiento, pero en realidad implica acondicionar el aire para que tenga características que garanticen el confort, como calentar, enfriar, humidificar, deshumidificar, limpiar y eliminar olores [6].

El confort humano depende de varios factores ambientales [6], como la temperatura, la humedad relativa y el movimiento del aire. Numerosas investigaciones se han llevado a cabo para determinar una zona de temperatura de confort. Aunque la temperatura es el indicador más importante del confort, la humedad relativa también influye, ya que determina la capacidad del aire para absorber humedad y afecta la forma en que el cuerpo humano disipa el calor por evaporación. El movimiento del aire, por su parte, impacta el confort térmico al provocar un enfriamiento localizado no deseado en el cuerpo. Se busca un nivel bajo de movimiento del aire para eliminar el aire caliente y húmedo sin resultar incómodo.

En una habitación tratada como un sistema cerrado (Fig. 1.1), cuyos límites con el exterior se denominan particiones, se producen interacciones energéticas con el entorno debido a la transferencia de calor [7]. Cualquier término (excepto la generación) puede actuar como entrada o salida de calor en función de la diferencia de temperatura entre el interior y exterior:

- Calor que entra por conducción desde el exterior ($\dot{Q}_g$)
- Radiación solar a través de las particiones ($\dot{Q}_s$)
- Generación interna debido a las personas, a dispositivos e iluminación ($\dot{Q}_w$)
- Ventilación (Q_v)
- Calor que sale hacia el exterior ($\dot{Q}_o$)
- Calor que sale a las habitaciones adyacentes ($\dot{Q}_a$)

Para aclimatar un edificio, es necesario mantener la temperatura interior dentro de un rango cómodo para el ser humano.

El balance de energía para el aire que se encuentra al interior de una habitación se muestra en la Ec. 1.1 [8].

$$\{\text{Razón de acumulación de energía}\}_{sistema} = \{\text{Suma de flujos de energía}\} \quad (1.1)$$

En forma diferencial, la ecuación se muestra en la Ec. 1.2.

$$\frac{dE}{dt}_{sistema} = \sum_{i=1}^n \dot{E}_i \quad (1.2)$$

Donde la suma representa las interacciones de energía, tanto salidas como entradas, dentro de la habitación.

Dado que la temperatura interior depende de las interacciones energéticas, para mantenerla dentro de un rango de confort, la energía en el sistema debe permanecer constante, lo que implica que su variación temporal es cero, y suponiendo que no hay trabajo neto que entre o salga del sistema. Así, la Ec. 1.2 es transformada por la Ec. 1.3, que representa el equilibrio térmico dentro de la habitación.

$$\sum_{i=1}^n \dot{E}_i = 0 \quad (1.3)$$

Babiarz y Szymanski [7] concluyen que, para mantener el confort térmico dentro de una habitación, las entradas y salidas de calor deben estar en equilibrio, como se indica en la Ec. 1.4.

$$\dot{Q}_g + \dot{Q}_s + \dot{Q}_w = \dot{Q}_o + \dot{Q}_a + \dot{Q}_v \quad (1.4)$$

En la Fig. 1.1 se muestra una representación gráfica de una habitación, donde se esquematizan las entradas y salidas de calor mencionadas en esta sección.

El balance de energía presentado anteriormente se basa en la suposición de que la habitación se mantiene dentro del rango de confort térmico y que no existe trabajo neto. Sin embargo, en la realidad, el metabolismo de las personas y el uso de electricidad generan un trabajo neto en el sistema. Por ello, el estudio del confort térmico requiere análisis adicionales.

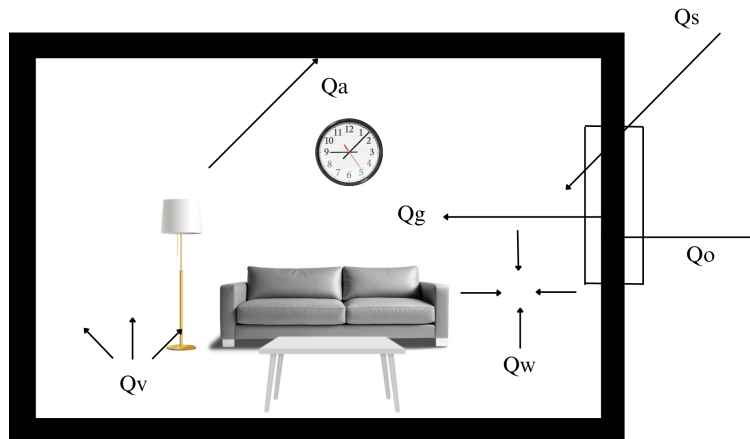
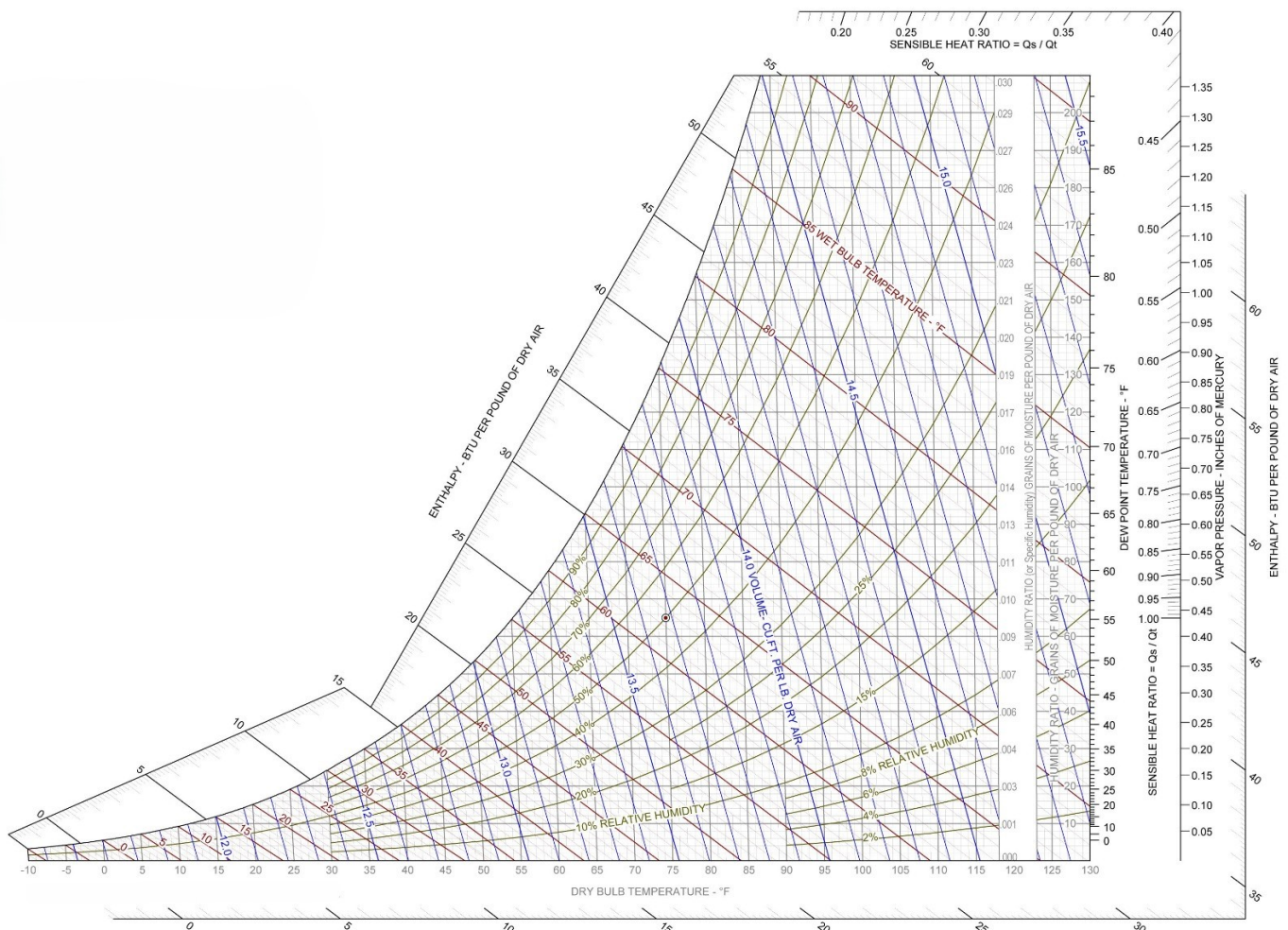


Figura 1.1: Representación gráfica de las entradas y salidas de calor en una habitación.

La carta psicrométrica es una representación gráfica de varias propiedades de la atmósfera, compuesta por una mezcla de aire y vapor de agua. Como se muestra en la Fig. 1.2, la carta

psicrométrica ilustra los siguientes componentes [1]:

- El eje horizontal es la **temperatura de bulbo seco** medida en $^{\circ}\text{C}$ y el eje vertical es el contenido de humedad o **humedad absoluta** medida en g/kg .
- Las curvas de **humedad relativa**, medidas en %, decrecen en dirección inferior derecha con respecto a la curva de saturación y en dirección de la temperatura y humedad absoluta.
- Las líneas rectas que van desde la curva de saturación, y cortan las rectas de temperatura de bulbo seco, son la **temperatura de bulbo húmedo**. Cuando el aire está saturado, no hay evaporación y ambas temperaturas son iguales, que son los puntos donde las líneas se cruzan.
- Las líneas de **entalpía específica** son casi paralelas a las líneas de temperatura de bulbo húmedo, pero la escala se muestra fuera del diagrama.
- Las líneas de **volumen específico**, que es el recíproco de la densidad, se indican mediante líneas inclinadas desde la curva de saturación.



definir las condiciones térmicas ambientales adecuadas para la ocupación humana [9]. Este documento incluye métodos para determinar el rango en la carta psicrométrica en el cual se considera que existe confort térmico bajo ciertas condiciones específicas. Por ejemplo, el método para determinar una zona de confort con velocidad de aire elevada establece el rango de confort para velocidades de viento superiores a 0.2 m/s , donde los ocupantes tienen una actividad metabólica entre 1 y 2 *met*, y un valor de aislamiento de ropa entre 0.65 y 1 *clo*. En este caso, la zona de confort se define en la carta psicrométrica mostrada en la Fig. 1.3.

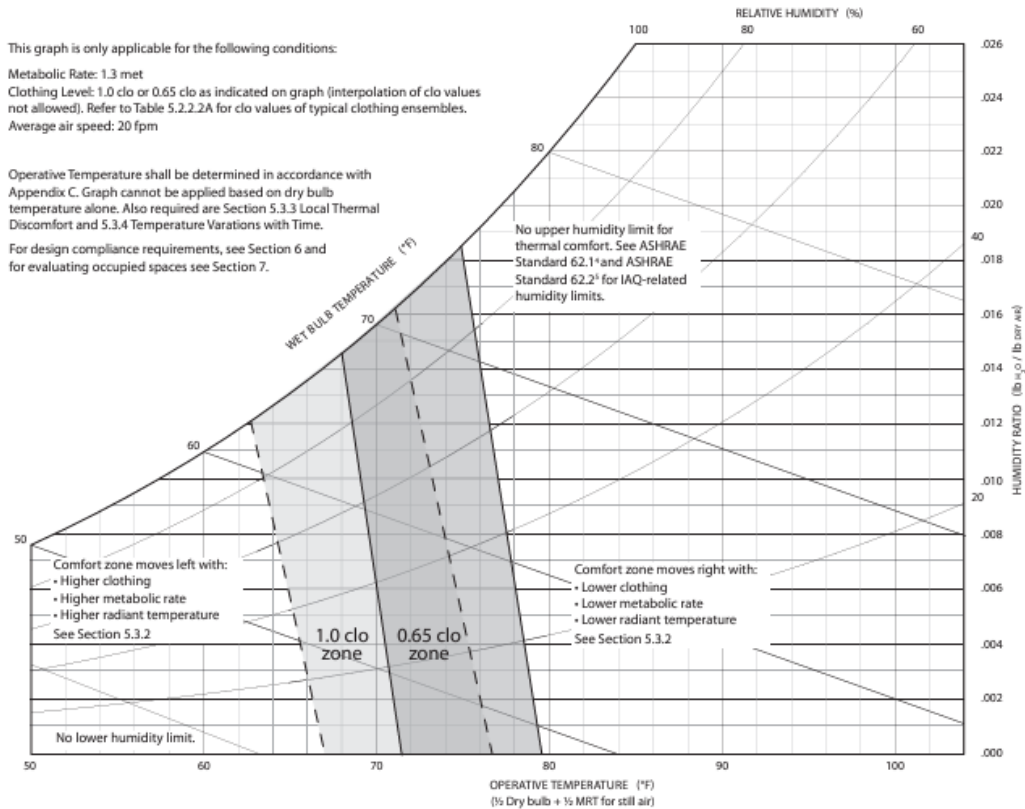


Figura 1.3: Zona de confort térmico con velocidad de aire elevada, según la norma ASHRAE 55 [9].

1.1.3 Condiciones favorables para el confort térmico

La norma *ASHRAE Standard 55-2004* [5] tiene como propósito el especificar las combinaciones de los factores ambientales térmicos interiores y factores personales que van a producir las condiciones de ambiente térmico que son aceptables para la mayoría de las personas que ocupan un espacio.

La gran mayoría de los datos disponibles que hay sobre confort térmico considera un nivel de actividad física casi sedentario, típico de trabajo de oficina. Sin embargo, se puede utilizar también para determinar las condiciones ambientales apropiadas para cuando existe una actividad física moderadamente elevada. Estos datos no sirven para personas que están durmiendo, ni hay información significativa sobre los requerimientos de confort para niños, personas con discapacidad o

enfermos, pero se puede aplicar si se toma a un grupo de personas en un espacio de forma homogénea [5].

Para ciertos valores de humedad, velocidad de viento, tasa de metabolismo y aislante de vestimenta, se pueden determinar zonas de confort. Esta zona es definida en términos de un rango de *temperaturas de operación* que provean de las condiciones aceptables de ambiente térmico que las personas encuentren agradables. En espacios donde los ocupantes tienen niveles de actividad que resultan en tasas de metabolismo para personas sedentarias o con actividad ligera, y un aislamiento térmico por vestimenta entre 1 y 1.5 clo, se pueden aplicar modelos computacionales para determinar zonas de confort [5].

La mayor parte de la información disponible sobre confort térmico se centra en niveles de actividad sedentarios o semi-sedentarios, debido a la complejidad de establecer criterios para actividades de mayor intensidad. Las zonas de confort definidas por la ASHRAE aplican únicamente a escenarios de baja actividad; sin embargo, es posible adaptar estos datos considerando diferentes niveles de aislamiento por vestimenta y condiciones ambientales [5].

La escala de sensación térmica de la ASHRAE [5] fue desarrollada para cuantificar la sensación térmica de las personas, y se define con una escala de puntuación:

- +3 Caliente
- +2 Tibio
- +1 Poco tibio
- 0 Neutral
- 1 Poco fresco
- 2 Fresco
- 3 Frío

El modelo de *valor medio predictivo (PMV)* utiliza principios del balance de calor para relacionar los agentes de malestar térmico con la respuesta media de las personas al ambiente térmico, basada en la escala mencionada anteriormente. El *porcentaje de personas insatisfechas (PPD)* está relacionado con el PMV, y está basado en el asumir que las personas que votan +2, +3, -2 y -3 en la escala de sensación térmica no se encuentran satisfechas. Se dice que existe el confort térmico cuando $PPD < 10$ y $-0.5 < PMV < +0.5$ [5].

1.1.4 Indicadores de malestar térmico

Un indicador de malestar térmico es un valor singular que integra los efectos de los parámetros básicos que afectan el ambiente térmico. Los indicadores se dividen en tres grupos [4]:

- **Indicadores racionales.** Están basados en cálculos que involucran la ecuación de transferencia de calor.
- **Indicadores empíricos.** Se sustentan en malestares objetivos y subjetivos.
- **Indicadores directos.** Son creados basándose en mediciones directas de variables ambientales.

Los primeros dos indicadores son los más complicados de implementar para realizar investigaciones de confort térmico en espacios ocupados determinados, debido a que involucran bastantes variables.

Se puede pensar que los indicadores más comprensibles son los racionales, ya que integran todas las variables ambientales y de comportamiento. Sin embargo, debido a que no hay una manera práctica de añadir todas las variables dentro de la ecuación de calor, algunos de los parámetros se deben asumir o determinar como constantes.

La tarea de construir un indicador y un rango de confort térmico es algo complicado, considerando todas las variables y fenómenos que se relacionan a que una persona se sienta cómoda en su ambiente térmico. Si se deben considerar los parámetros ambientales, la tasa de actividad y el nivel de vestimenta, entonces es importante considerar formas simplificadas de obtener una estimación del balance térmico que causa cambios en el contenido de calor del cuerpo humano. Epstein & Moran (2006) [4] mencionan que esto se puede lograr utilizando un indicador o *índice* que se basa en mediciones directas de variables ambientales. Por lo tanto, el uso de un índice directo junto con lineamientos para el nivel de actividad y la vestimenta es lo más práctico.

Existen diversos índices directos, como se puede observar en la Tabla 1, donde T_{db} es la temperatura de bulbo seco y T_{wb} es la temperatura de bulbo húmedo. De estos índices presentados, el que es de interés especial es el índice de malestar de Thom (DI) [11], debido a que es de uso extendido por más de 5 décadas.

1.1.5 Grados día

Los grados día son un concepto muy utilizado para determinar el confort térmico dentro de un lugar habitado, para así predecir la cantidad de energía que se puede consumir al utilizar dispositivos HVAC u otro tipo de instrumentos para calefacción o refrigeración.

Tabla 1: Índices directos basados en temperaturas de bulbos seco y húmedo; recuperada de [4].

Índice	Ecuación
Índice de Oxford (WD)	$0.15T_{db} + 0.85T_{wb} [^{\circ}C]$
Índice de malestar de Thom (DI)	$0.4T_{db} + 0.4T_{wb} + 15 [^{\circ}C]$
Índice de estres térmico (FITS)	$0.35T_{db} + 0.83T_{wb} + 5.08 [^{\circ}C]$
Índice modificado (MDI)	$0.3T_{db} + 0.75T_{wb} [^{\circ}C]$
Temperatura seca de bulbo húmedo (WBDT)	$0.6T_{db} + 0.4T_{wb} [^{\circ}C]$
Índice de Thom modificado (DI)	$T_{db} - 0.55(1 - 0.01RH)(T_{db} - 14.5) [^{\circ}C]$

Son una medida de la cantidad y el tiempo que la temperatura exterior se encuentra por debajo de un nivel específico [12], por lo que es una medida de qué tan frío es un lugar en un tiempo determinado, en el caso de los grados día de calefacción (HDD), y qué tan caliente es un lugar, en el caso de los grados día de refrigeración (CDD). Este concepto es muy utilizado en los estudios de consumo de energía debido a que mientras más alta o baja es la temperatura con respecto a un estándar, mayor será la cantidad de grados día y, por lo tanto, más probable será un aumento en el consumo de energía debido al uso de dispositivos de aire acondicionado [13, 14].

Los grados día, tanto de refrigeración como de calefacción, son un indicador que mide las desviaciones acumuladas de la temperatura con respecto a una temperatura base a lo largo del tiempo. Se calculan sumando las diferencias entre la temperatura exterior y una temperatura de referencia durante un periodo determinado [15].

El cálculo de grados día se realiza para cada uno de los intervalos de observación en el periodo de tiempo requerido. Por ejemplo, en este trabajo se desea obtener los grados día de un día determinado, y se tienen los datos de temperatura por hora del día, se utilizan las Ecs. 1.5 y 1.6, tomando el dato de temperatura como T_0 y $t = \frac{1}{24}$ debido a que se está realizando el cálculo para solamente una hora del día.

$$HDD_h = (T_{base} - T_0)^- t \quad (1.5)$$

$$CDD_h = (T_0 - T_{base})^+ t \quad (1.6)$$

donde h indica que son calculados por hora, T_{base} es la temperatura base, T_0 es la temperatura exterior en un tiempo determinado, los signos $+$ y $-$ indican que solo aplica para temperaturas por debajo o encima de la base, t es el tiempo para el cual aplica la temperatura exterior.

1.2 Materiales de cambio de fase

Hale et al. [17], en el manual técnico de la NASA sobre materiales de cambio de fase, definen los PCM como materiales que tienen la capacidad de absorber y liberar grandes cantidades de calor sin experimentar un cambio significativo en su temperatura. Esta absorción y liberación de calor ocurre durante el cambio de fase del material, que es el proceso de transición a un estado físico distinto, siendo comúnmente procesos de líquido a sólido, o de sólido a líquido. Los PCM pueden ser utilizados como un método pasivo de aclimatización, debido a que mantienen su temperatura dentro de un rango específico mediante propiedades termofísicas o consideraciones geométricas, sin necesidad de movimiento de partes, fluidos o potencia.

En cuanto a la diferencia entre los PCM y otros materiales de almacenamiento de calor, Hasenöhl [18] señala que, a diferencia de las rocas o el agua, el punto de fusión de los PCM se encuentra dentro de un rango de operación funcional, dependiendo de la aplicación requerida para el material.

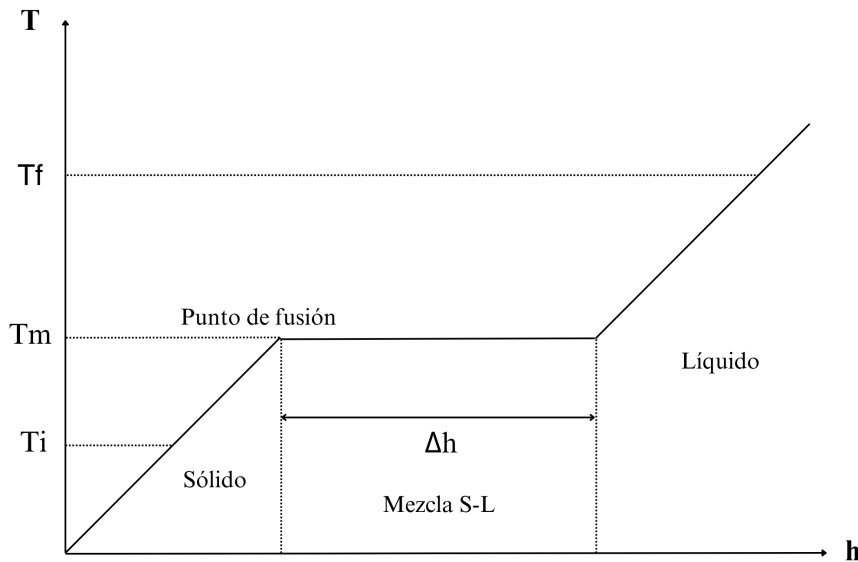


Figura 1.4: Diagrama de cambio de fase para un PCM.

Una cierta cantidad de calor, conocida como calor latente, se absorbe o se libera cuando el PCM se funde o solidifica, respectivamente. Este cambio de fase ocurre a una temperatura aproximadamente constante [18], como se ilustra en la Fig. 1.4.

Hasenöhl [18] detalla el cálculo del calor almacenado durante el cambio de fase, que es descrito en la Ec. 1.7.

$$Q = \int_{T_i}^{T_m} mC_p dT + m\Delta h + \int_{T_m}^{T_f} mC_p dT \quad (1.7)$$

donde m es la masa del material que cambia de fase, C_p es el calor específico a presión constante del material, Δh es el calor latente, T_i y T_f son temperaturas arbitrarias por debajo y por encima de la temperatura de fusión del PCM, y T_m es la temperatura de fusión.

Para que un PCM sea adecuado para aplicaciones en la construcción, debe poseer ciertas propiedades termofísicas [19]:

- Una temperatura de fusión adecuada para aplicaciones de construcción.
- Una alta conductividad térmica, que permita al PCM cargarse y descargarse en un tiempo adecuado.
- Un alto calor específico, que permita la disponibilidad de energía adicional en forma de calor sensible para el sistema de almacenamiento de energía térmica.
- El PCM debe derretirse completamente durante el cambio de fase, de modo que las fases líquida y sólida sean homogéneas, para evitar la reducción de su capacidad de almacenamiento cuando el cambio de fase no se realiza por completo [20].
- Un pequeño cambio en el volumen durante el cambio de fase y una baja presión de vapor a la temperatura de operación, para evitar problemas de contención.

Un PCM puede ser incorporado dentro de los materiales de construcción por medio de técnicas diversas, que son incorporación directa, inmersión, encapsulación y estabilización de forma [19]. Estas técnicas de incorporación se describen a continuación:

Incorporación directa.

Es la opción mas sencilla, económica y práctica, en donde el PCM se mezcla durante el proceso con los materiales de construcción, que pueden ser yeso o concreto. Sin embargo, con esta técnica, la filtración del PCM en estado líquido puede afectar las propiedades mecánicas y durabilidad del sistema.

Inmersión.

En esta técnica, los materiales de construcción, como pueden ser el concreto, la tablaroca o los ladrillos, son sumergidos en el PCM en estado líquido y absorben el PCM por la acción capilar. Sin embargo, las fugas del PCM son comunes con esta técnica, y también la durabilidad y las propiedades mecánicas de los materiales de construcción se ven afectadas.

Encapsulación.

El PCM debe de ser encapsulado antes de ser utilizado en los elementos de construcción. La encapsulación del PCM debe de cumplir los siguientes criterios: fuerza, durabilidad y estabilidad

térmica; actuar como barrera del PCM, para protegerlo de interacciones destructivas; tener un área suficiente para una transferencia de calor adecuada, y ser sencillo de manejar [21]. Existen dos tipos de encapsulación:

- *Microencapsulación.* Se concentran partículas pequeñas de PCM en cáscaras delgadas hechas de polímeros, desde $1\ \mu m$ hasta $1000\ \mu m$. Esta técnica previene el derramamiento del PCM y el cambio de volumen durante el cambio de fase, y provee de una alta tasa de transferencia de calor por área superficial.
- *Macroencapsulación.* Se almacena una cantidad significativa de PCM en contenedores como paneles, tubos y esferas. Esta técnica facilita el transporte del PCM, mejora su compatibilidad con los alrededores y minimiza los cambios externos de volumen. Sin embargo, reduce la conductividad térmica, requiere protección contra daños y demanda un mayor esfuerzo para integrarlo en las estructuras de construcción.

Estabilización de forma.

En esta técnica se utiliza un material de apoyo, como polietileno de alta densidad o butadieno. El PCM y material de apoyo son derretidos y mezclados a altas temperaturas y luego la mezcla se congela hasta volverlo sólido. Los materiales de apoyo deben ser porosos para que el PCM pueda ser absorbido por medio de capilaridad. Los PCM bajo esta técnica tienen un gran calor específico, una conductividad térmica adecuada y no necesita un contenedor.

Los PCMs se clasifican en orgánicos, inorgánicos y eutécticos [19], y aquellos que pueden aplicarse a la construcción son aquellos cuyos puntos de fusión están dentro del rango de $18\ ^\circ C$ y $40\ ^\circ C$. Esta clasificación se describe de la siguiente manera:

- *Orgánicos.* Se dividen en parafinas y no parafinas. Se encuentran en un rango alto de temperaturas, altos calores latentes de fusión, tienen cambios pequeños en el volumen durante el cambio de fase y son compatibles con los materiales de construcción; no obstante, su conductividad térmica suele ser baja (alrededor de $0.2\ W/m.K$), a diferencia de los inorgánicos que pueden tener valores entre 0.7 y $1.2\ W/m.K$. Los materiales más comunes de este tipo son parafinas de carbono, ácido láctico, poliglicol y estearato de vinilo.
- *Inorgánicos.* Se dividen en metales y sales hidratadas. Se destacan por tener un alto almacenamiento volumétrico de calor latente y conductividad térmica; sin embargo, suelen ser corrosivos, se degradan a largo plazo, tienen un alto cambio en el volumen y experimentan supercongelación cuando se solidifica. Los materiales más comunes de este tipo son sales hidratadas de calcio, cloro, sodio, litio y azufre.
- *Eutécticos.* Se dividen en orgánico-orgánico, inorgánico-inorgánico y orgánico-inorgánico. Son combinaciones de compuestos químicos que suelen tener temperaturas de fusión en el

rango operacional y una densidad volumétrica alta de almacenamiento térmico, pero hay información limitada en sus propiedades termofísicas.

1.3 Machine Learning

1.3.1 Modelos predictivos de Machine Learning

Machine Learning es una rama de la inteligencia artificial en la que se enseña a las máquinas a realizar tareas sin ser programadas explícitamente [22]. Machine Learning se enfoca en extraer conocimiento a partir de datos, lo cual tiene una gran influencia en la investigación centrada en datos (data-driven). Tiene aplicaciones en múltiples áreas y en diversos aspectos de la vida cotidiana; detrás de plataformas como Netflix y de redes sociales como Twitter, hay numerosos modelos de Machine Learning [22].

Las principales características de Machine Learning son las siguientes: la máquina se alimenta de datos y aprende las relaciones entre ellos; basándose en estas relaciones, puede predecir eventos para otros datos, y el desempeño de este proceso de aprendizaje y predicción mejora con la experiencia [22].

Müller et al. [23] hablan sobre los dos tipos de aprendizaje de los algoritmos de Machine Learning:

- *Aprendizaje supervisado*. El proceso automático de toma de decisiones se generaliza mediante ejemplos conocidos. El usuario proporciona al algoritmo ejemplos de entrada, etiquetados con las salidas deseadas; el objetivo es que el algoritmo aprenda por comparación. Se le llama *supervisado* porque se asemeja a la supervisión de un profesor que espera que su alumno aprenda de una manera específica; por lo tanto, supervisa tanto la información como el proceso de aprendizaje.
- *Aprendizaje no supervisado*. En este caso, solo se conocen los datos de entrada y estos no están etiquetados; por lo tanto, el algoritmo identifica aspectos en común entre los datos y los agrupa según sus similitudes.

Machine Learning utiliza métodos computacionales para aprender información directamente de una base de datos. Para el modelo predictivo que se plantea desarrollar, se utilizarán algoritmos de aprendizaje supervisado, los cuales entrenan el modelo con entradas y salidas de datos ya conocidas para predecir salidas futuras [24].

El aprendizaje supervisado utiliza técnicas de clasificación y regresión para desarrollar el modelo predictivo. Sin embargo, para los fines de este trabajo, se utilizarán exclusivamente técnicas de regresión, que se pueden clasificar en los siguientes algoritmos [23][24]:

- *Modelos lineales*: realizan una predicción utilizando una función lineal construida con las características de los datos de entrada.
- *Árboles de decisión*: aprenden una jerarquía o preguntas de estructura if-else, para tomar una decisión. Cada nodo en el árbol representa una pregunta que se responde con un verdadero o falso. Las respuestas suelen ser otras preguntas que concluyen con una decisión.
- *Máquinas de soporte vectorial*: se modelan los datos de entrada en un espacio multi-dimensional utilizando una transformación no lineal, calcular un valor inicial y observar las correlaciones entre las entradas y la salida aplicando regresión.
- *Redes neuronales*: son generalizaciones de modelos lineales que realizan múltiples pasos de procesamiento para llegar a una decisión. Cada nodo representa una variable de entrada y se le asigna un coeficiente aprendido que la conecta con la salida.

1.3.2 Modelos White, Black y Grey Box

La diversidad de variables de entrada, las condiciones climáticas y los parámetros de construcción aumentan la complejidad de los modelos matemáticos que buscan predecir de forma precisa las cargas térmicas de enfriamiento y calentamiento de un edificio. [25]. En los modelos de carga térmica, no siempre se incorporan suficientes datos de entrada, y diversas investigaciones llegan a conclusiones diferentes sobre la cantidad adecuada de datos para obtener un modelo preciso [25]. Por lo tanto, no existe un consenso sobre cuántos datos de entrada son suficientes ni sobre si ciertos tipos de datos son más importantes que otros.

Existen diferentes enfoques en el desarrollo de modelos predictivos: White Box, Black Box y Grey Box. El White Box modelling (también llamado physics-driven o centrado en la física) considera la lógica interna de la estructura de un modelo, lo que implica que el desarrollador debe comprender el fenómeno físico que se está modelando. El Black Box modelling (también llamado data-driven o centrado en los datos) se refiere al desarrollo del modelo sin conocimiento del funcionamiento interno de su estructura, es decir, no requiere entendimiento de la base física del modelo, solo de la parte computacional. El Grey Box modelling, también conocido como modelado de orden reducido, combina ambos enfoques: utiliza datos obtenidos mediante el modelado White Box para luego probarse con un modelo Black Box [26].

Cada uno de estos enfoques tiene sus ventajas y desventajas. Algunas ventajas del modelo Grey Box son que combina los beneficios de ambos enfoques y ofrece escenarios de prueba más amplios y precisos [26]. En la literatura científica, se han desarrollado modelos centrados en datos relacionados con la carga térmica dentro de edificios, pero estos suelen basarse en condiciones y conjuntos de datos muy específicos, como tendencias de temperatura o velocidad del viento, lo que resulta en interpretaciones poco generalizables, a diferencia de los modelos que consideran la física del problema [27]. Existen pocos modelos centrados en datos que se deriven de datos obtenidos

mediante modelación física (modelos Grey Box), ya que requieren un conocimiento especializado tanto en la física del problema como en el desarrollo de modelos estadísticos predictivos mediante Machine Learning [26].

Zhang et al. [27] realizaron una comparación entre modelos Black Box y Grey Box utilizando información de la operación normal de un edificio y datos de prueba de temperaturas para predecir el consumo de energía. Se concluyó que la precisión del modelo aumenta un 67 % en un escenario de prueba al utilizar datos térmicos, en comparación con un modelo que solo emplea datos de demanda y uso de energía.

1.3.3 Algoritmos de Machine Learning seleccionados para el procesamiento de datos

En este trabajo se utilizará el Toolbox de Estadística y Machine Learning en MATLAB, que provee de funciones y aplicaciones para describir, analizar y modelar datos; utilizando los algoritmos supervisados de regresión para obtener las inferencias entre datos y construir modelos predictivos [24].

Como se mencionó anteriormente, el aprendizaje supervisado es el tipo de Machine Learning más utilizado y consiste en la predicción de una salida determinada con una base de datos de entrada para entrenamiento. Para tareas de regresión, el objetivo es el de predecir un número continuo.

En aprendizaje supervisado, normalmente se desea construir modelos con datos de entrenamiento y que, con esto, sea capaz de realizar predicciones con nuevos datos que posean las mismas características que los datos de entrenamiento. Si un modelo logra esto, se dice que es capaz de generalizar.

A continuación se da una breve explicación de los algoritmos de Machine Learning utilizados para realizar el proceso de entrenamiento. Las matemáticas detrás de estos algoritmos están fuera de los alcances de este trabajo.

Máquina de soporte vectorial de núcleo Gaussiano.

Una máquina de soporte vectorial (SVM, por sus siglas en inglés) clasifican datos por medio de la búsqueda del mejor hiperplano que separe todos los puntos de datos en clases, siendo este el que tenga un margen mayor entre las clases de datos. Esto significa que se busca el espacio máximo entre el hiperplano y los puntos de datos interiores [24]. Durante el entrenamiento, la SVM aprende la importancia de cada uno de los puntos de datos de entrenamiento para representar el límite entre las clases. Típicamente, solo una parte de estos puntos de datos son importantes para definir este límite, que son aquellos que se encuentran en la frontera entre clases, y son llamados *vectores de soporte* [23].

Para realizar una predicción de un nuevo punto, se mide la distancia de cada vector de soporte. La decisión que toma para realizar la clasificación es basada en estas distancias [23].

El núcleo de una SVM es la forma en la que se mide esta distancia que hay entre cada vector de soporte. Cuando la SVM incluye un núcleo Gaussiano, la distancia entre vectores de soporte se calcula con la ecuación 1.8, donde x_1 y x_2 son puntos de datos, para los cuales se mide la distancia euclidiana [23].

$$k_g(x_1, x_2) = \exp(-||x_1 - x_2||^2) \quad (1.8)$$

Regresión lineal de interacción.

Un modelo lineal realiza una predicción utilizando una función lineal de las entradas. Existen diferentes modelos lineales de regresión, cuya diferencia entre ellos reside en cómo son aprendidos los parámetros de la función lineal en el modelo con los datos de entrenamiento y cómo se controla la complejidad del modelo [23].

La regresión lineal de interacción utiliza un modelo lineal que incluye términos que buscan las relaciones entre todas las variables de predicción. En la Ec. 1.9 se muestra la ecuación lineal de este tipo de regresión, incluyendo los términos de interacción, que multiplican cada una de las variables de predicción, que son x_n y x_m , y β_n y ε son los parámetros de regresión [28].

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \beta_{nm} x_n x_m + \varepsilon \quad (1.9)$$

Árbol de regresión empaquetado.

Un árbol de decisión es una jerarquía de preguntas *if/else* que deriva en una decisión. Estas preguntas son llamadas *pruebas*, que son utilizadas en datos continuos de forma que se van comparando, por ejemplo, si una característica i es más grande que un valor a . Para construir un árbol, el algoritmo busca todas las pruebas posibles y encuentra la que sea más informativa sobre la variable a predecir [23].

El *empaquetamiento* es un procedimiento que tiene el propósito de reducir la varianza de un árbol de decisión. Dado un conjunto de observaciones independientes $Z_1, Z_2, \dots, Z_n$, cada una con varianza de σ^2 , la varianza del promedio de estas observaciones se calcula con la Ec. 1.10, donde n son el número de observaciones [29].

$$\sigma_Z^2 = \frac{\sigma^2}{n} \quad (1.10)$$

En otras palabras, obtener el promedio de un conjunto de observaciones reduce la varianza, por lo que el empaquetamiento consiste en tomar conjuntos de la base de datos de entrenamiento y construir modelos de predicción por separado, para así calcular el promedio de estos modelos y reducir la varianza en el árbol de decisión, construyendo y ensamblando cientos de árboles en un solo procedimiento [29].

Red neuronal amplia.

Las redes neuronales son un conjunto de algoritmos que son conocidos como *deep learning*, y sus modelos se conocen como *perceptrones multicapa*, que se pueden ver como generalizaciones de modelos lineales que realizan múltiples pasos de procesamiento para llegar a una decisión [23].

En regresión lineal, la salida es la suma de las variables de entrada por los coeficientes de regresión que son aprendidos por el modelo. En un perceptrón multicapa, este proceso se repite múltiples veces, primero con la computación de capas ocultas que representan un paso intermedio de procesamiento. Este modelo requiere del cálculo de muchos coeficientes más que en un modelo de regresión lineal, por lo que, después de realizar el cómputo de la suma para cada capa oculta, se aplica una función no lineal de tangente hiperbólica al resultado [23].

El componente amplio de una red neuronal se refiere a que captura patrones conocidos e interacciones entre los datos de entrenamiento, que permite al modelo aprender las relaciones entre las variables de entrada. Esto resulta en un conjunto de relaciones binarias que indican la presencia o ausencia de interacción entre las variables, y posteriormente se incluyen en los modelos lineales [30].

1.4 Antecedentes

Chen et al. [25] desarrollaron un modelo Grey Box para predecir el consumo eléctrico en oficinas estadounidenses, utilizando datos obtenidos a partir de modelos físicos, que luego fueron procesados con un algoritmo de Machine Learning seleccionado. En investigaciones recientes, como las mencionadas anteriormente, se ha otorgado una gran importancia al desarrollo de modelos predictivos precisos para la carga térmica de edificios, combinando modelos centrados en datos y modelos centrados en la física, con el fin de mejorar el confort térmico en las nuevas edificaciones. Realizaron un modelo físico en simulación para obtener un conjunto de datos a partir de prototipos de edificios. Utilizaron diferentes tipos de edificios, con variaciones en el número de pisos y área superficial, basados en un prototipo de oficina del Departamento de Energía de Estados Unidos. Los datos obtenidos se procesaron y se emplearon como entradas para un modelo en Python, utilizando parámetros como la hora del día, el día de la semana, el número de ocupantes y las condiciones climáticas (temperatura, humedad y radiación solar). Concluyeron que un modelo centrado en datos (Black Box) puede representar con precisión uno centrado en la física.

Zhussupbekov et al. [31] construyeron un modelo Grey Box para predecir el consumo energético de edificios residenciales integrados con PCM en la región climática del Mediterráneo. Para ello, realizaron simulaciones térmicas a fin de obtener datos de entrada como el punto de fusión de los PCM, las condiciones ambientales y las características de construcción del edificio. La base de datos se creó utilizando nueve tipos diferentes de edificios bajo las condiciones climáticas de siete ciudades del Mediterráneo. El análisis de resultados consistió en identificar los parámetros de diseño con mayor influencia en el consumo de energía, el tipo de PCM que generó el mayor ahorro energético y el algoritmo de Machine Learning con mayor confiabilidad estadística, utilizando el 70 % de los datos para el entrenamiento del modelo y el 30 % restante para validación y pruebas.

Existen diversas investigaciones que abordan el desarrollo de modelos centrados en datos para la predicción térmica [32][33]. Seyedzadeh et al. [32] destacan los vacíos en la creación de modelos de Machine Learning para la predicción del uso energético, señalando que estos modelos suelen enfocarse en la elección de algoritmos adecuados y en el desempeño estadístico, sin considerar la física del problema. Para mejorar estos modelos, se utilizaron algoritmos de optimización. Kumar et al. [33] mejoran otros modelos al incluir datos de atributos de construcción, como la geometría, áreas de piso, techo y paredes, y la altura del edificio. Además, emplean un tipo de algoritmo conocido como Extreme Machine Learning, que utiliza grandes redes neuronales con asignación aleatoria de entradas, centrándose también exclusivamente en los datos.

Adilkhanova et al. [34] investigaron el potencial de los PCM para mejorar el confort térmico en viviendas. Utilizaron *EnergyPlus* para desarrollar simulaciones basadas en un modelo sencillo de vivienda y evaluaron las condiciones térmicas en varias ciudades de Kazajistán. Para ello, emplearon un índice de confort térmico y analizaron el impacto de combinar PCM con ventilación natural. Los resultados mostraron que esta combinación mejora significativamente el confort térmico, ya que el uso de PCM por sí solo puede aumentar la temperatura interior si no se activa adecuadamente.

Capítulo 2

Modelado y simulación de condiciones térmicas en viviendas mexicanas ejemplares

La simulación térmica permite modelar las interacciones energéticas dentro de un espacio determinado mediante la descomposición de la ecuación de calor en una zona térmica. En este trabajo se utilizó el software de simulación térmica *EnergyPlus*, acompañado de su editor y la interfaz gráfica *OpenStudio*. Por ello, resulta fundamental explorar el funcionamiento interno del software y las funciones necesarias para simular las condiciones térmicas en una vivienda mexicana representativa que incorpora PCM en su estructura. Esto requiere definir las características de una vivienda ejemplar, analizar las condiciones ambientales específicas del país y determinar el tipo de PCM más adecuado para este caso.

2.1 Funcionamiento del software de simulación térmica

A continuación, se describirán algunas de las funciones y conceptos básicos de OpenStudio, con el objetivo de explicar el funcionamiento general de este software de simulación térmica.

- Sitio

EnergyPlus requiere la adición de un archivo climático, que contenga datos ambientales de una ciudad específica, como lo son datos históricos de temperatura, velocidad del viento, radiación, humedad relativa, punto de rocío, entre otras, o ser creado por el usuario a partir de mediciones climáticas. Este archivo se integra en el apartado de sitio (ver Fig. 2.1) y se obtiene de bases de datos climáticas o se genera mediante softwares especializados en la creación de este tipo de datos.

- Horarios

Los horarios se refieren a los momentos en los que hay actividad en los diferentes espacios del edificio: actividad humana, uso de equipos electrónicos, de gas, iluminación y sistemas de climatización (HVAC) (ver Fig. 2.2). Para esto, se construyeron distintos tipos de horarios, que se asociarán a la actividad en el edificio:

Always On. Horario que tiene un valor de 1 por default las 24 horas, que será asociado a definiciones que siempre están encendidas o tienen actividad durante todo el día.

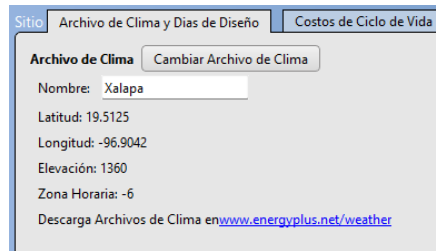


Figura 2.1: Apartado de Sitio en OpenStudio.

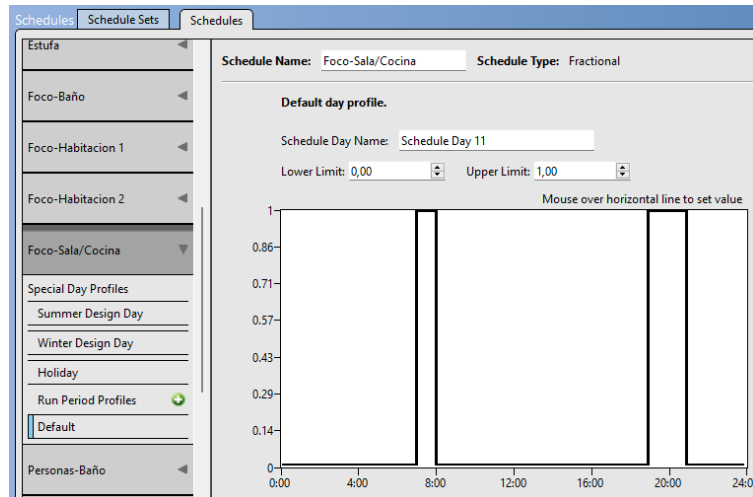


Figura 2.2: Apartado de Horarios en OpenStudio.

Always Off. Contrario a Always On.

Definitions Schedule. Horarios definidos para aparatos electrónicos que se agregan como lavadoras, microondas, computadoras, etc.

HTG/CLG HVAC. Sistema HVAC activado 24/7 que mantiene la zona a una temperatura determinada de calentamiento/enfriamiento respectivamente.

Clothing Schedule. Horario en el que todos los habitantes del edificio utilizan ropa de invierno. En fechas de primavera y otoño se tiene un valor de 1 en las horas de la mañana y noche, y durante el día, un valor menor de 1; en verano se tiene un valor menor a 1 durante todo el día, y en invierno, se tiene un valor de 1 durante el día.

■ Construcción

En el apartado de construcción se emplean diversos tipos de materiales que pueden asignarse a las diferentes partes del edificio: paredes, techo exterior, techo interior, suelo, ventanas, puertas y tragaluces (ver Fig. 2.3). Esta sección cuenta con tres pestañas: sets de construcción, construcción y materiales.

Sets de construcción. Son un grupo de ensambles de construcción que serán aplicados al edificio: material interior y exterior de paredes, pisos y techo, material subsuperficial, contacto con el piso, puertas de vidrio, tragaluz, etc.

Construcción. Se muestran a mayor detalle los ensambles de construcción. Cada ensamble tiene capas de material.

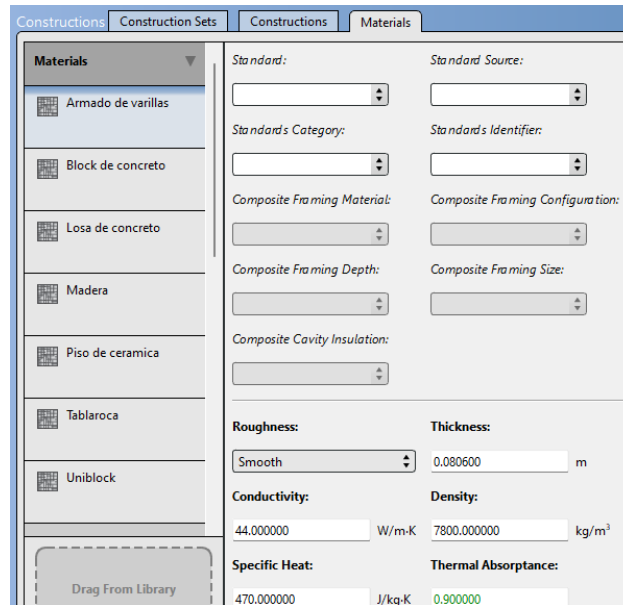


Figura 2.3: Apartado de Construcción en OpenStudio.

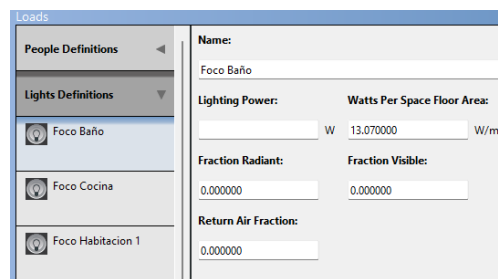


Figura 2.4: Apartado de Cargas en OpenStudio.

Materiales. Se tienen las propiedades mecánicas y térmicas de los materiales utilizados en cada ensamble, que pueden modificarse. Los materiales que no se encuentren en la librería se pueden cargar o agregar por sus propiedades físicas.

■ Cargas

En el apartado de cargas, es necesario definir las actividades energéticas dentro del edificio (ver Fig. 2.4). Cada carga agregada debe tener asignado un horario.

Definiciones de masa interna. Cálculo de la masa térmica basada en la densidad de los materiales localizados dentro del edificio.

Definiciones de personas. Densidades de ocupación dentro de varios espacios del edificio con el calor y generación de CO_2 que cada persona provee.

Definiciones de iluminación. Potencia generada por iluminación dentro del edificio.

Definiciones de equipo eléctrico. Potencia generada por aparatos en el edificio.

■ Espacios

El modelo se compone de superficies y subsuperficies. Las superficies corresponden a las principales áreas del edificio, como paredes, techo y suelo; las subsuperficies incluyen puertas, ventanas, puertas de garaje y tragaluces. También existen las particiones interiores, que

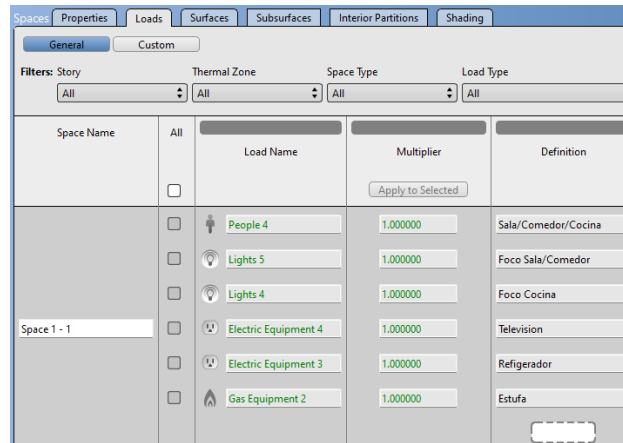


Figura 2.5: Apartado de Espacios en OpenStudio.

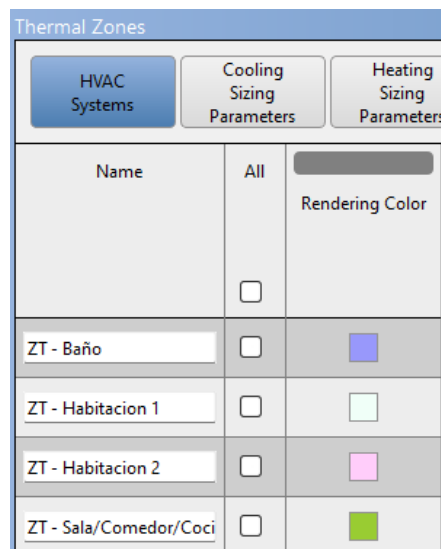


Figura 2.6: Apartado de Zonas Térmicas en OpenStudio.

suelen ser paredes de altura parcial, como en los cubículos de oficina. Los tipos de espacios se asignarán a la geometría y representan distintas habitaciones o áreas del edificio. Para cada espacio, se deben agregar sets de construcción y horarios predeterminados, así como especificar el aire exterior y la forma de infiltración de aire (ver Fig. 2.5).

■ Zonas térmicas

Cada habitación que compone el modelo corresponde a una zona térmica, y estas interactúan térmicamente entre sí. Cada una tendrá resultados de temperatura y humedad relativa interior. Se deben asignar los horarios de los sistemas HVAC u otros dispositivos de climatización en cada espacio del edificio, si corresponde (ver Fig. 2.6).

2.2 Modelos numéricos de EnergyPlus

2.2.1 Modelo matemático de conducción con resistencias térmicas en Energy-Plus

El método utilizado por EnergyPlus [35] para calcular la transferencia de calor por conducción es llamado *método de espacio de estados* definido por las siguientes ecuaciones matriciales lineales:

$$\frac{d[\mathbf{x}]}{dt} = [\mathbf{A}][\mathbf{x}] + [\mathbf{B}][\mathbf{u}] \quad (2.1)$$

$$[\mathbf{y}] = [\mathbf{C}][\mathbf{x}] + [\mathbf{D}][\mathbf{u}] \quad (2.2)$$

donde

$\mathbf{x}$ es el vector de variables de estado $\mathbf{x} = [T_1 \ T_2 \ \dots \ T_n]$.

$\mathbf{u}$ es el vector de entradas $\mathbf{u} = [T_i \ T_o]$.

$\mathbf{y}$ es el vector de salida $\mathbf{y} = [q_i'' \ q_o'']$.

$\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ y $\mathbf{D}$ son matrices de coeficientes.

La formulación se puede utilizar para resolver la ecuación transiente de conducción de calor aplicando un mallado de diferencias finitas sobre las capas de los elementos del edificio siendo analizado. Las temperaturas nodales son las variables de estado ($\mathbf{x}$), las temperaturas ambientales son las entradas ($\mathbf{u}$), y los flujos de calor exterior e interior de superficie son las salidas ($\mathbf{y}$).

Para una capa simple del edificio con dos nodos interiores y convección en ambos lados, como se observa en la Fig. 2.7, las ecuaciones resultantes son:

$$\text{Flujo de calor interno } q_i'' = h(T_i - T_2) \quad (2.3)$$

$$\text{Flujo de calor externo } q_o'' = h(T_1 - T_o) \quad (2.4)$$

$$\text{Primer nodo interno } C \frac{dT_1}{dt} = hA(T_o - T_1) + \frac{T_2 - T_1}{R} \quad (2.5)$$

$$\text{Segundo nodo interno } C \frac{dT_2}{dt} = hA(T_i - T_2) + \frac{T_1 - T_2}{R} \quad (2.6)$$

donde;

$R = \frac{l}{kA}$, es la resistencia de conducción, que depende de la longitud que separa los nodos, la conductividad térmica y el área de superficie.

$C = \frac{\rho c_p l A}{2}$ es el coeficiente de almacenamiento para la ecuación de difusión transitoria, que depende de la densidad, el calor específico, la longitud y el área de superficie.

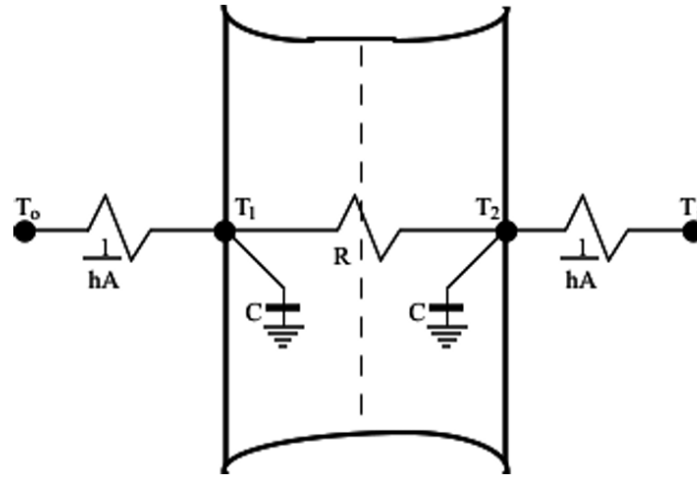


Figura 2.7: Representación de una resistencia térmica de una capa de pared en un modelo de EnergyPlus [35].

2.2.2 Algoritmos para calcular el coeficiente de convección en EnergyPlus

Se utiliza un algoritmo para el cálculo del coeficiente de convección (h) para el aire interior y exterior del edificio que es el objeto de estudio [35]. Este algoritmo es llamado **TARP** (Thermal Analysis Research Program), que se basa en experimentos de placa plana y correlaciones. En este modelo, la convección es dividida en componentes forzado (h_f) y natural (h_n):

$$h_c = h_f + h_n \quad (2.7)$$

$$h_f = 2.537 W_f R_f \left(\frac{P V_z}{A} \right)^{1/2} \quad (2.8)$$

$$h_n = \frac{1.810 |\Delta T|^{1/3}}{1.382 + |\cos \theta|} \quad (2.9)$$

donde

W_f es un modificador que depende de la dirección del viento.

R_f es un multiplicador que depende de la rugosidad de la superficie.

P es el perímetro de la superficie.

V_z es la velocidad del viento.

ΔT es la diferencia de temperaturas entre el aire y la superficie.

A es el área de la superficie.

θ es el ángulo de inclinación con respecto a la superficie horizontal.

2.2.3 Modelo de diferencias finitas en conducción (Módulo PCM)

EnergyPlus utiliza un algoritmo específico para simular la contribución térmica de PCMs en recintos, llamado **CondFD** (Conduction Finite Differences); este realiza una discretización de las paredes, suelo y techo en nodos y utiliza un esquema implícito [36], que se muestra en la ecuación

2.10 para una construcción con un material homogéneo y espacio nodal uniforme.

$$C_p \rho \Delta x \frac{T_i^{j+1} - T_i^j}{\Delta t} = k_W \frac{T_{i+1}^{j+1} - T_i^{j+1}}{\Delta x} + k_E \frac{T_{i-1}^{j+1} - T_i^{j+1}}{\Delta x} \quad (2.10)$$

donde

k_W es la conductividad térmica al interior de la construcción.

k_E es la conductividad térmica al exterior de la construcción.

C_p , ρ son el calor específico y la densidad del material, respectivamente.

Δx , Δt son el grosor de la capa de material y el paso temporal, respectivamente.

i es el nodo siendo modelado.

j es el paso de tiempo anterior.

Para el PCM, se utiliza una función de entalpía dependiente de la temperatura $h = h(T)$ para que el algoritmo calcule un calor específico equivalente para cada paso de tiempo [36], que será el que se utilice en la ecuación 2.10, y se calcula por medio de la Ec. 2.11.

$$C_p^* = \frac{h_i^j - h_i^{j-1}}{T_i^j - T_i^{j-1}} \quad (2.11)$$

2.3 Creación del modelo ejemplar de vivienda mexicana en Energy-Plus

2.3.1 Planteamiento del problema

Este trabajo evalúa el uso de materiales de cambio de fase para mejorar el confort térmico dentro de una vivienda en México. Para ello, se creó un modelo de una vivienda mexicana estándar y se realizó una simulación numérica utilizando el software EnergyPlus, empleando datos climáticos de ocho ciudades mexicanas que representan los principales tipos de clima según la clasificación de Köppen.

Las simulaciones se realizaron con dos tipos de modelos: uno sin materiales de cambio de fase y otro con PCMs (Rubitherm PCMs RT 21HC, RT 26 y RT 31 [37]) integrados en las paredes y el techo. El objetivo fue comparar el confort térmico en ambos casos y evaluar el desempeño de los PCMs en la mejora del confort térmico en cada ciudad mexicana. Se utilizó un índice directo, acompañado de una categorización de malestar, para medir el nivel de incomodidad térmica dentro del modelo de la vivienda.

2.3.2 Construcción del modelo de vivienda

La construcción de la geometría se basó en los planos proporcionados por INFONAVIT para que el modelo de vivienda sea representativo de una casa estándar de bajo costo en México, como se presenta en las Figuras 2.8 y 2.9. Para apoyar a los trabajadores mexicanos con menores recursos económicos en la construcción de sus viviendas, INFONAVIT ofrece documentos que contienen planos y especificaciones técnicas para diferentes tipos de viviendas, incluyendo la expansión y construcción de espacios para sus propios negocios. Los planos para construir la geometría de una vivienda de bajo costo en México se obtienen de un manual técnico de INFONAVIT [38], que consiste en un edificio rectangular de 48 m^2 , cuyas dimensiones son 4.8 m por 10 m y 2.5 m de altura. El modelo de vivienda contiene 4 habitaciones. El frente de la casa está en la pared oeste, y contiene una puerta de 2 m por 1 m y una ventana de dimensiones 1 m por 1.2 m (Fig. 2.8). Asimismo, en la pared norte se incluyeron una puerta y una ventana de las mismas dimensiones, mientras que en las paredes este y sur se incluyó una ventana de las mismas dimensiones, añadiendo una ventana de 0.6 m por 0.5 m en la pared sur (ver Fig. 2.9). En México, alrededor del 42 % de las viviendas tienen entre 30 y 75 m^2 ; por lo tanto, el modelo construido es representativo de una vivienda mexicana estándar [39].

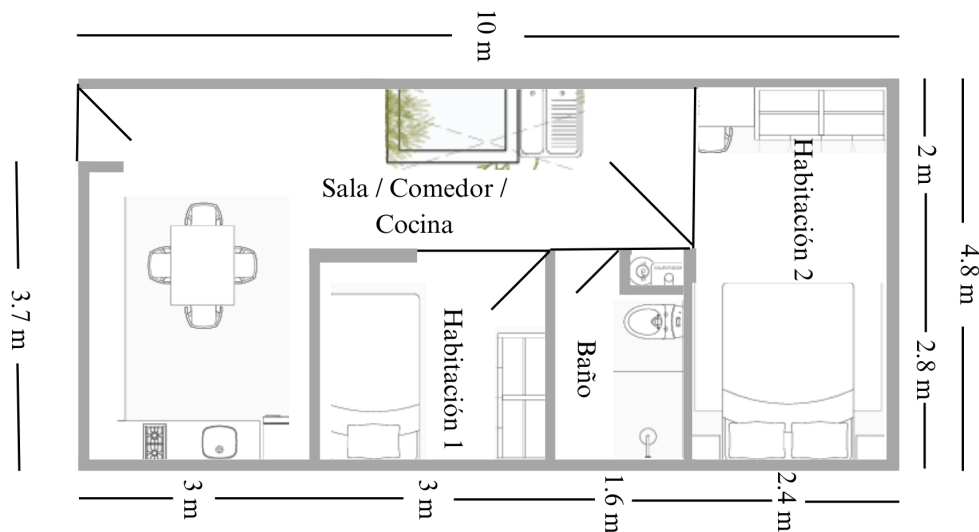


Figura 2.8: Plano de la casa modelo de INFONAVIT.

Los materiales introducidos en el modelo fueron seleccionados con base en los materiales comunes utilizados en la construcción de viviendas en México [40]. Para las ventanas, se utilizó vidrio incluido en las bibliotecas de EnergyPlus. Los materiales utilizados y sus propiedades termo-físicas se encuentran en la Tabla 2.

El periodo de simulación utilizado para la realización de cálculos térmicos fue un mes de invierno, del 10 de enero al 10 de febrero, y un mes de verano, del 15 de julio al 15 de agosto, ambos del año 2020. La selección de este periodo es debido a que se requiere de un rango ilustrativo a ambas estaciones del año sin tener una base de datos demasiado grande; este periodo se ampliará en el Capítulo 4, para construir la base de datos para entrenamiento. El modelo de vivienda,

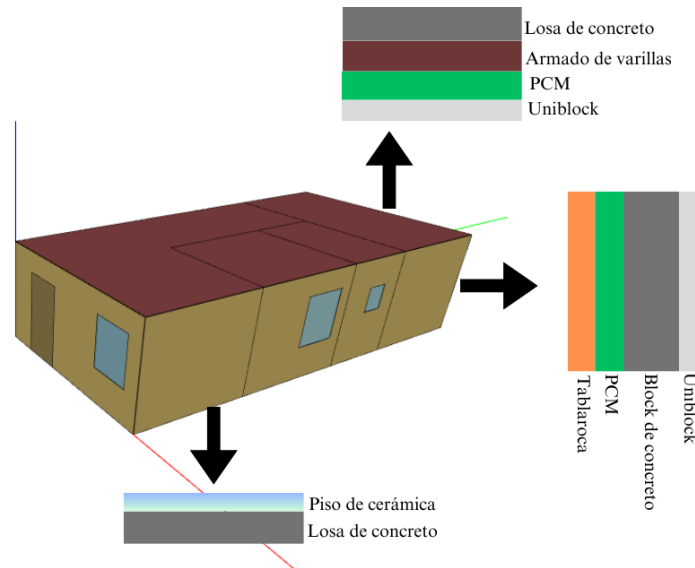


Figura 2.9: Ejemplo de una casa estándar de bajo costo en México y los materiales utilizados para el modelo de vivienda.

Tabla 2: Materiales utilizados y sus propiedades termo-físicas.

Componente	Material	Espesor (m)	Cond. Térmica ($W.m^{-1}K^{-1}$)	Densidad ($kg.m^{-3}$)	Calor Esp. ($J.kg^{-1}K^{-3}$)
Pared	Block de concreto	0.12	0.92	2100	880
Pared, Techo	Tablaroca	0.0127	0.1445	615.75	1090
	PCM	0.02	0.2	880	2000
	Uniblock (revestimiento)	0.007	0.2101	1811.27	840
Techo	Armado de varillas	0.008	44	7800	470
Techo, Suelo	Losa de concreto	0.08	1.5	2400	1050
Suelo	Piso de cerámica	0.01	1.7	2500	323
Puerta	Madera	0.05	0.22	400	2300
Ventana	Vidrio teórico	0.003	2.1073	0	0

al ser un edificio con múltiples habitaciones, que se refieren como *zonas térmicas* en el software EnergyPlus, incluye horarios de ocupación, uso de iluminación, equipos eléctricos y uso de electrodomésticos y gas, como se muestra en la Tabla 3.

Estas cargas se han calculado para una vivienda ocupada por cuatro personas, siguiendo los horarios establecidos en la Tabla 3. Según estos horarios, la ocupación varía a lo largo del día: una persona en el baño, entre una y dos personas en cada habitación, y de una a cuatro personas en la sala/comedor/cocina.

2.3.3 Selección de los PCMs

El criterio principal para la selección de PCM es el punto de fusión del material. Los PCM almacenan y liberan energía durante el proceso de cambio de fase, que ocurre cuando el material está expuesto a una temperatura cercana a su punto de fusión. Esta temperatura debe ser una que aparezca frecuentemente en el entorno. Entonces, se seleccionaron tres PCMs con temperaturas de fusión de $21^{\circ}C$, $26^{\circ}C$ y $31^{\circ}C$, que son tres temperaturas que se encuentran frecuentemente en el ambiente mexicano para las estaciones de verano e invierno.

Tabla 3: Horarios de ocupación y parámetros considerados para el modelo de vivienda.

Zona térmica	Tipo de carga	Horario	Potencia (W)
Baño	Actividad	6:00 - 8:00; 12:00 - 12:05; 14:00 - 14:05; 16:00 - 16:05; 19:30 - 19:40; 20:40 - 21:00	50/persona
	Iluminación	6:00 - 8:00; 19:30 - 19:40; 20:40 - 21:00	13.07/m ²
Habitación 1	Actividad	0:00 - 6:00; 21:00 - 24:00	20/persona
		6:00 - 8:00; 14:00 - 18:00	80/persona
	Iluminación	6:00 - 8:00; 20:50 - 21:00	7.66/m ²
	Equipo (Computadora)	15:00 - 18:00	15.33/m ²
Habitación 2	Actividad	0:00 - 6:00; 21:00 - 24:00	20/persona
		6:00 - 8:00; 14:00 - 18:00	80/persona
	Iluminación	6:00 - 8:00; 20:50 - 21:00	4.81/m ²
Sala/Comedor/Cocina	Actividad	7:00 - 8:00; 14:00 - 21:00	120/persona
		12:00 - 14:00	60/persona
	Iluminación 1	7:00 - 8:00; 19:00 - 21:00	6.21/m ²
	Iluminación 2	7:00 - 8:00; 19:00 - 21:00	4.46/m ²
	Equipo (Refrigerador)	Todo el día	20.83/m ²
	Equipo (Televisión)	7:00 - 8:00; 13:00 - 21:00	9.53/m ²
	Equipo (Estufa)	7:00 - 8:00; 13:00 - 14:30	66.6/m ²
		19:00 - 20:00	

Los PCMs seleccionados para su incorporación en las simulaciones son comerciales, elegidos de la sección de PCMs orgánicos versátiles de Rubitherm [37] con un grosor de 2 cm. En el caso las simulaciones sin PCM, solamente se omite la parte de PCM en la configuración de pared y techo. Los PCMs se denominan RT 21 HC, RT 26 y RT 31.

Dado que EnergyPlus requiere la entalpía del PCM como una función dependiente de la temperatura para las simulaciones del modelo de vivienda, las curvas de entalpía-temperatura del PCM RT 21 HC, PCM RT 26 HC y PCM RT 31 se adaptaron de [34], como se muestra en la Figura 2.10. Para estudiar la capacidad del PCM de afectar las condiciones térmicas dentro del modelo de vivienda, no se utiliza ningún sistema de HVAC ni ventilación en las simulaciones.

2.3.4 Selección de las ciudades

Para este trabajo, se consideró la categorización generalizada de los diferentes tipos de clima y zonas climáticas en el mapa de Köppen-Geiger [41]. Hay nueve regiones climáticas en México, que son tropical con lluvias durante todo el año (Af), tropical con lluvias de monzón (Am), tropical con lluvias de verano (Aw), seco estepario (BS), seco desértivo (BW), templado con lluvias durante todo el año (Cf), templado con lluvias de verano (Cw), templado con lluvias de invierno (Cs), y polar de alta montaña (EB) [42].

Para el desarrollo de las simulaciones, se deben seleccionar ciudades que representen la diversidad de climas en México. Para la selección se consideró el número de habitantes. El clima polar fue omitido porque este tipo de clima no es representativo de ninguna ciudad mexicana, ya que solo se encuentra en las partes altas de algunas montañas del país. La Tabla 4 muestra las ciudades

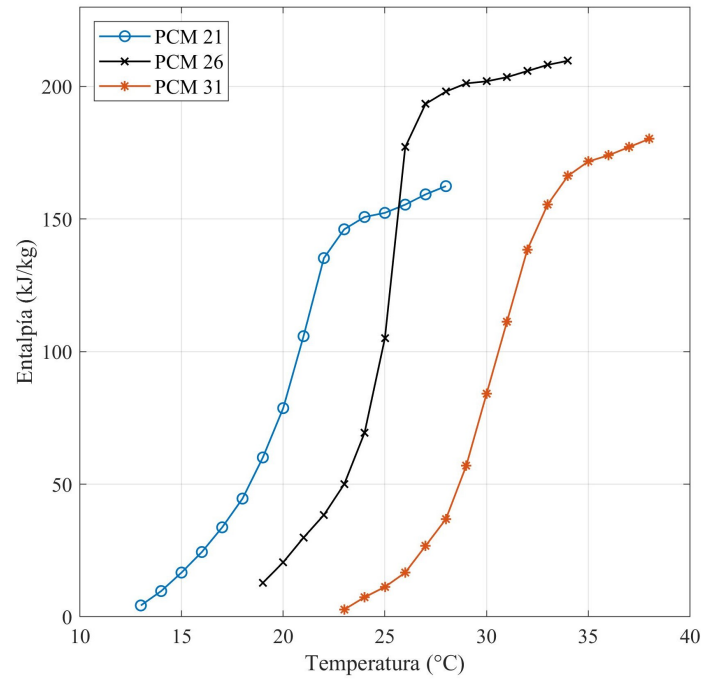


Figura 2.10: Curvas de entalpía-temperatura adaptadas de [34].

seleccionadas.

En la Fig. 2.11, se muestra un mapa de México donde se destacan las ciudades seleccionadas para las simulaciones, que fueron escogidas debido a su mayor densidad poblacional e importancia, por lo que se consideran representativas dentro del grupo climático. Los archivos de datos climáticos se obtuvieron de la base de datos *Climate One Building* [43]. En esta base de datos, la información meteorológica se deriva de varias fuentes climáticas públicas, principalmente de la Base de Datos Integrada de Superficie (ISD) [44].

Tabla 4: Ciudades seleccionadas para el estudio.

Grupo climático	Región climática	Ciudad
Tropical	Af	Tuxtla Gutiérrez, Chiapas
	Am	Villahermosa, Tabasco
	Aw	Mérida, Yucatán
Seco	BS	Monterrey, Nuevo León
	BW	Hermosillo, Sonora
Templado	Cf	Xalapa, Veracruz
	Cw	Guadalajara, Jalisco
	Cs	Toluca, Estado de México
Polar	EB	Omitido



Figura 2.11: Ciudades seleccionadas en México, basadas en las zonas climáticas de Köppen, para realizar las simulaciones.

2.4 Validación de las simulaciones numéricas

Para comprobar de primera mano la confiabilidad y el uso correcto de EnergyPlus en combinación con su interfaz gráfica OpenStudio, se realizó una validación adaptando las condiciones experimentales de los datos reportados en [45] y [46] a un modelo en *EnergyPlus*. Las mediciones de temperatura se realizaron en una edificación sencilla construida con poliestireno expandido, madera barnizada y dos PCMs.

Los experimentos fueron realizados con dos configuraciones de pared, techo y suelo. La primera, llamada A, contenía un PCM con punto de fusión de 40°C , mientras que la segunda, llamada B, contenía dos PCM con puntos de fusión de 40°C y 33°C . Se replicó el experimento en un modelo de *EnergyPlus* utilizando las especificaciones del problema.

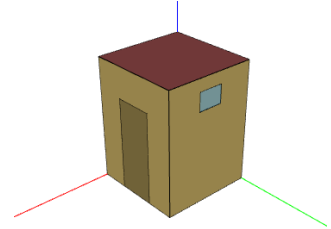
En la figura 2.12 se puede observar una comparación entre el modelo de OpenStudio basado en la edificación creada por Zhuang, et al. [46] contra una fotografía de la edificación real. Los experimentos se realizaron en la ciudad de Chongqing, China en 2007, por lo que un archivo climático correspondiente a esa ciudad y año fue utilizado.

Con el fin de que las condiciones de simulación fueran lo más similares posibles a las reportadas en el artículo original, se asumieron los siguientes parámetros: el orden de las capas, de exterior a interior, es madera barnizada, PCM 40, EPS y madera barnizada; los coeficientes convectivos

interno y externo son de 8.7 y $23 \text{ W/m}^2 - K$, respectivamente; el piso se consideró adiabático; se utilizó un color gris cemento con una absorptancia solar de 0.5 ; y el valor de infiltración de aire fue de 0.5 . Asimismo, las condiciones climáticas correspondientes al archivo climático, así como los resultados de temperatura de las simulaciones, se presentan en el Anexo B.



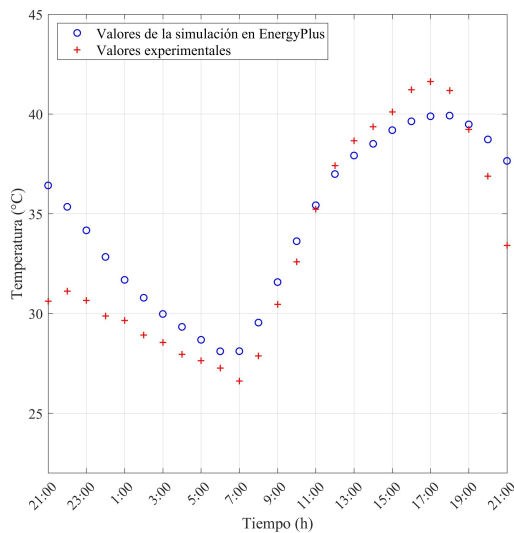
(a) Edificación utilizada en [46].



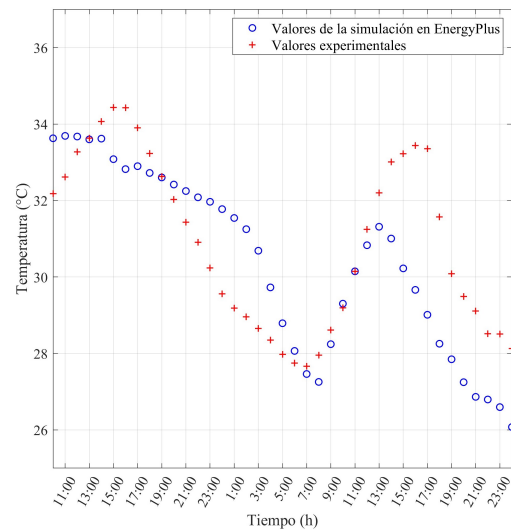
(b) Modelo de edificación creado en OpenStudio.

Figura 2.12: Comparación entre el modelo de OpenStudio con la edificación en [46].

En la figura 2.13, se puede observar una comparación entre las temperaturas interiores obtenidas experimentalmente por Zhuang, et al. [46] y aquellas obtenidas en las simulaciones utilizando *EnergyPlus* por los días en los que se realizaron los experimentos, que fueron 14 y 15 de agosto de 2007 para A, y 23 y 24 de agosto de 2007 para B. Los materiales que fueron introducidos solamente para el modelo de validación y sus propiedades termo-físicas se muestran en la Tabla 5. Los PCMs utilizados para el modelo de validación fueron introducidos con curvas genéricas de entalpía-temperatura para sus puntos de fusión.



(a) Temperatura interior con la configuración A.



(b) Temperatura interior con la configuración B.

Figura 2.13: Comparación entre las temperaturas interiores experimentales y simuladas

Tabla 5: Materiales utilizados y sus propiedades termofísicas para las simulaciones de validación.

Material	Espesor (m)	Conductividad térmica (W/m-K)	Densidad (kg/m ³)	Calor específico (J/kg-K)
Madera barnizada	0.005	0.17	796	1731
Poliestireno EPS	0.05	0.042	56	1189
PCM 33	0.01	0.3	860	1970
PCM 40	0.01	0.3	860	1970

Se puede observar en las gráficas que existen discrepancias entre las temperaturas experimentales y simuladas, principalmente debido a la falta de información en [45] y [46] para que sea posible replicar la simulación. Esta información incluye las curvas de entalpía-temperatura de los PCMs y los datos termo-físicos completos de los materiales. Adicionalmente, surgen errores debido a las discrepancias de los datos climáticos y posibles errores en la medición experimental. A pesar de lo anterior, las desviaciones entre los valores experimentales y simulados, calculadas por el **error absoluto medio (MAE)**, expresados según la Ec. 2.12. Para cada configuración, el MAE resultante es de 1.78 °C para A, y 1.4 °C para B.

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_{i,sim} - X_{i,exp}| \quad (2.12)$$

2.5 Resultados y discusión del capítulo

Las simulaciones en EnergyPlus proporcionaron la temperatura y la humedad relativa dentro del modelo de vivienda durante el período de simulación, tanto sin PCM como con PCM. El tiempo para realizar cada simulación está entre 1 minuto en los casos sin PCM y de 3 a 5 minutos cuando incluye PCM.

Para evaluar el efecto del PCM en la temperatura interior y el funcionamiento del módulo PCM en *EnergyPlus*, la Fig. 2.14 presenta una comparación de las temperaturas interiores con y sin el uso del PCM 21, junto con la temperatura exterior, que representa los datos climáticos. Esta comparación abarca 5 días de invierno en la ciudad de Monterrey (BS), muestra que durante los picos altos de temperatura exterior, que afecta directamente las condiciones ambientales en el interior y oscila entre los 13 y 27 °C, la temperatura interior con PCM disminuye en comparación con la temperatura interior sin PCM.

Este comportamiento corresponde al efecto amortiguador del PCM, debido a la absorción o liberación de calor latente durante temperaturas exteriores cercanas a la temperatura de fusión del PCM, lo que permite que la temperatura interior con PCM se mantenga más constante y cercana a una temperatura de confort. En el caso del verano en Monterrey (BS), este efecto se intensifica porque la temperatura exterior tiende a fluctuar alrededor del punto de fusión del PCM, que es de 21 °C. La temperatura interior sin PCM a veces se encuentra por debajo de la temperatura exterior se puede deber a la liberación de calor por medio de radiación desde las paredes y techo debido a que en el modelo se tiene una vivienda aislada.

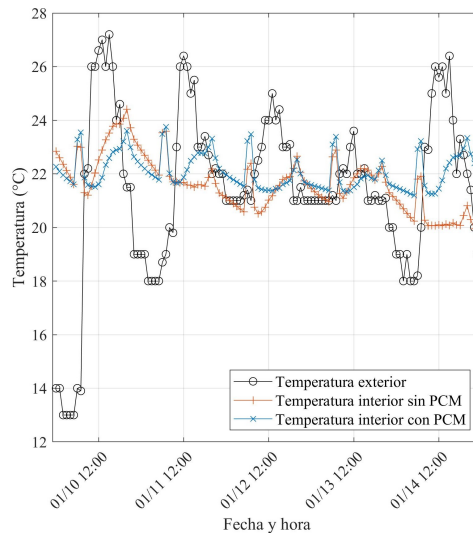


Figura 2.14: Comparación entre las temperaturas interiores en Monterrey (BS) durante el invierno con y sin PCM, cuya $T_f = 21^{\circ}\text{C}$.

Se concluye que el módulo PCM en *EnergyPlus* es eficaz para simular la inclusión de un PCM con una curva específica de entalpía-temperatura. Asimismo, permite analizar el efecto de amortiguamiento térmico que estos materiales generan en la temperatura interior de una edificación.

Capítulo 3

Evaluación del confort térmico

Tras la validación y prueba de las simulaciones térmicas con PCM, el paso siguiente es realizar un análisis de confort térmico aplicado al periodo de simulación. Para ello, se emplean dos conceptos clave en este análisis. El primero es el *índice de malestar (DI)*, ya mencionado en la sección 1.1, que clasifica el tipo de malestar térmico que experimenta un grupo de personas bajo condiciones específicas de temperatura y humedad relativa. El segundo es el *cambio total de malestar (TDC)*, que cuantifica el efecto de los PCMs en el confort térmico.

3.1 Cálculo del confort térmico

3.1.1 Índice de malestar (DI)

El objetivo de realizar simulaciones térmicas es predecir el confort térmico dentro del modelo de vivienda. Por lo tanto, se requiere un indicador matemático para medir dicho confort. Las variables obtenidas por las simulaciones de EnergyPlus son la temperatura interior de bulbo seco y la humedad relativa interior. Por lo tanto, se puede utilizar un índice directo para esta tarea.

El Índice de Malestar (DI, por sus siglas en inglés) [4] es una forma de medir el confort térmico, mediante la ecuación 3.1 [47]. Este corresponde al índice de malestar modificado de Thom, presentado en la sección 1.1.4.

$$DI = T_{db} - 0.55(1 - 0.01RH)(T_{db} - 14.5) \quad (3.1)$$

donde T_{db} es la temperatura de bulbo seco y RH es la humedad relativa en porcentaje, ambas medidas dentro del edificio y determinadas a través de simulaciones en EnergyPlus.

Este índice fue seleccionado debido a su uso estandarizado en investigaciones sobre confort térmico, ya que presenta una alta correlación con otros indicadores y una fuerte relación con los aspectos fisiológicos del confort térmico, validados experimentalmente [4].

Para los días de verano e invierno, existe un criterio de confort, asumiendo vestimenta normal, dependiendo del valor del DI, determinado a través de la clasificación presentada en la Tabla 1.

Las condiciones de incomodidad térmica debido al frío se determinaron basándose en la clasificación existente para altas temperaturas. La clasificación mostrada en la Tabla 6 es un rango propuesto del Índice de malestar basado en el utilizado en la literatura [47]. Para los propósitos de este trabajo, se aumentó la resolución de los rangos de DI con el objetivo de visualizar de manera más detallada los cambios en la clasificación del malestar térmico al emplear PCM, debido a que estos cambios no siempre suelen ser muy grandes. El aumento de la resolución tiene como objetivo reducir el margen de separación en las clasificaciones de malestar térmico. Esto se debe a que, en la literatura, estos rangos suelen abarcar un porcentaje demasiado amplio de malestar térmico, lo que dificulta el análisis de los cambios provocados por el uso de PCM y su impacto en el confort térmico.

Tabla 6: Criterio de confort con aumento de resolución según el índice de malestar.

Rango de DI	Etiqueta	Condiciones de malestar	Ponderación
DI < 7	A	Malestar severo por frío con posible riesgo para la salud	1
7 < DI < 8.5	B	Malestar grande por frío	8/9
8.5 < DI < 9.5	C	Todas las persona sienten malestar por frío	7/9
9.5 < DI < 11	D	La mayoría de las personas experimentan malestar por frío	2/3
11 < DI < 12.5	E	Más del 80 % de las personas sienten malestar por frío	5/9
12.5 < DI < 14	F	Entre 60 % a 80 % de las personas sienten malestar por frío	4/9
14 < DI < 15.5	G	Entre 40 % a 60 % de las personas sienten malestar por frío	1/3
15.5 < DI < 17	H	Entre 20 % a 40 % de las personas sienten malestar por frío	2/9
17 < DI < 18.5	I	Menos del 20 % de las personas sienten malestar por frío	1/9
18.5 < DI < 20.5	J	No hay malestar	0
20.5 < DI < 22	K	Menos del 20 % de las personas sienten malestar por calor	1/9
22 < DI < 23.5	L	Entre 20 % a 40 % de las personas sienten malestar por calor	2/9
23.5 < DI < 25	M	Entre 40 % a 60 % de las personas sienten malestar por calor	1/3
25 < DI < 26.5	N	Entre 60 % a 80 % de las personas sienten malestar por calor	4/9
26.5 < DI < 28	O	Más del 80 % de las personas sienten malestar por calor	5/9
28 < DI < 29.5	P	La mayoría de las personas experimentan malestar por calor	2/3
29.5 < DI < 30.5	Q	Todas las persona sienten malestar por calor	7/9
30.5 < DI < 32	R	Malestar grande por calor	8/9
DI > 32	S	Malestar severo por calor con posible riesgo para la salud	1

3.1.2 Cambio total de malestar (TDC)

El Cambio Total de Malestar (TDC, por sus siglas en inglés) es un concepto utilizado para medir el impacto térmico de los PCM en las viviendas. Este se mide estableciendo las horas durante el período de simulación en las que se mantiene un rango específico del Índice de Malestar [47]. El TDC se mide con la ecuación 3.2:

$$TDC = \sum_1^k n(\text{Reducción de malestar}) - \sum_1^k n(\text{Aumento de malestar}) \quad (3.2)$$

Donde k representa el número total de horas en el período de simulación y n es un número entero natural que representa cada una de las observaciones u horas simuladas.

La reducción de malestar representa la suma de horas en las que el malestar disminuye de un nivel superior a un nivel inferior, según los rangos mostrados en la Tabla 6, mientras que el aumento de malestar representa la suma de horas en las que el malestar aumenta de un nivel inferior a un nivel superior. Los valores de TDC se obtienen comparando las categorías de malestar para cada hora del período de simulación usando PCM y comparándolos con las horas en las que no se usa PCM. Para calcular el total de horas de malestar en el período de simulación, se asignó un peso a cada categoría de malestar. Tanto para invierno como para verano, se asignó un peso de $1/9$, añadiendo $1/9$ más para cada categoría a medida que se aleja del rango de confort térmico.

La Figura 3.1 ilustra gráficamente cómo se obtiene el TDC para una sola categoría. En el caso sin PCM, el valor de DI obtenido para una hora específica se encuentra en un rango específico de malestar térmico. Para la misma hora, en el caso con PCM, si el valor de DI transita a un rango de malestar inferior, el TDC suma un valor positivo. Por el contrario, si el uso de PCM provoca una transición a un rango de malestar superior, el TDC suma un valor negativo. El valor numérico del TDC para una hora específica representa el número de rangos que salta. El PCM con el valor de TDC más alto será el que tenga el mejor desempeño, ya que esto indica que reduce la mayor cantidad de horas de malestar térmico.

Para evitar confusiones, a partir de este punto se utilizará el término *TDC* para referirse a la suma total de los cambios en el malestar térmico debido al uso de PCM. Por otro lado, el término *DC* (Cambio de Malestar) se empleará para describir el cálculo según la Ec. 3.2 en una hora específica, sin aplicar la suma.

3.1.3 Metodología para los cálculos del TDC

Para calcular el cambio total de malestar para cada PCM en todas las ciudades durante el período de simulación, es necesario determinar primero el cambio de malestar para cada hora simulada. Para ello, se desarrolló un programa de MATLAB que sigue la lógica descrita en el diagrama de flujo mostrado en la Fig. 3.2.

La metodología empleada para calcular el cambio total de malestar se detalla a continuación:

1. Para cada hora del periodo de simulación, se obtienen de *EnergyPlus* la temperatura interior y la humedad relativa interior simuladas. Con esta información, se utiliza la Ec. 3.1 para calcular el índice de malestar por hora. Este cálculo se realiza tanto para el caso sin PCM como para los casos con PCM.
2. El valor de índice de malestar calculado en el paso anterior se clasifica dentro del rango especificado en la Tabla 6.

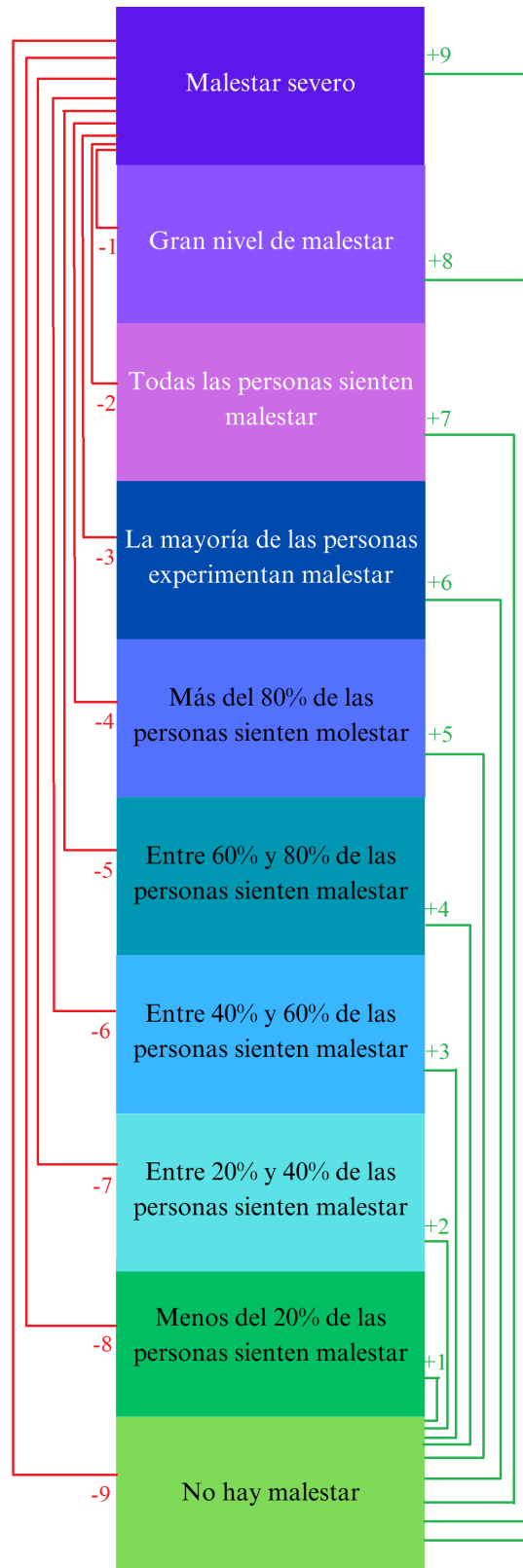


Figura 3.1: Representación esquemática para la determinación del TDC con respecto a los intervalos de confort.

- Según el rango del índice de malestar correspondiente a una hora específica del periodo de simulación, se asigna la ponderación mostrada en la Tabla 6.
3. Para comparar los rangos del índice de malestar correspondientes a una hora del periodo de

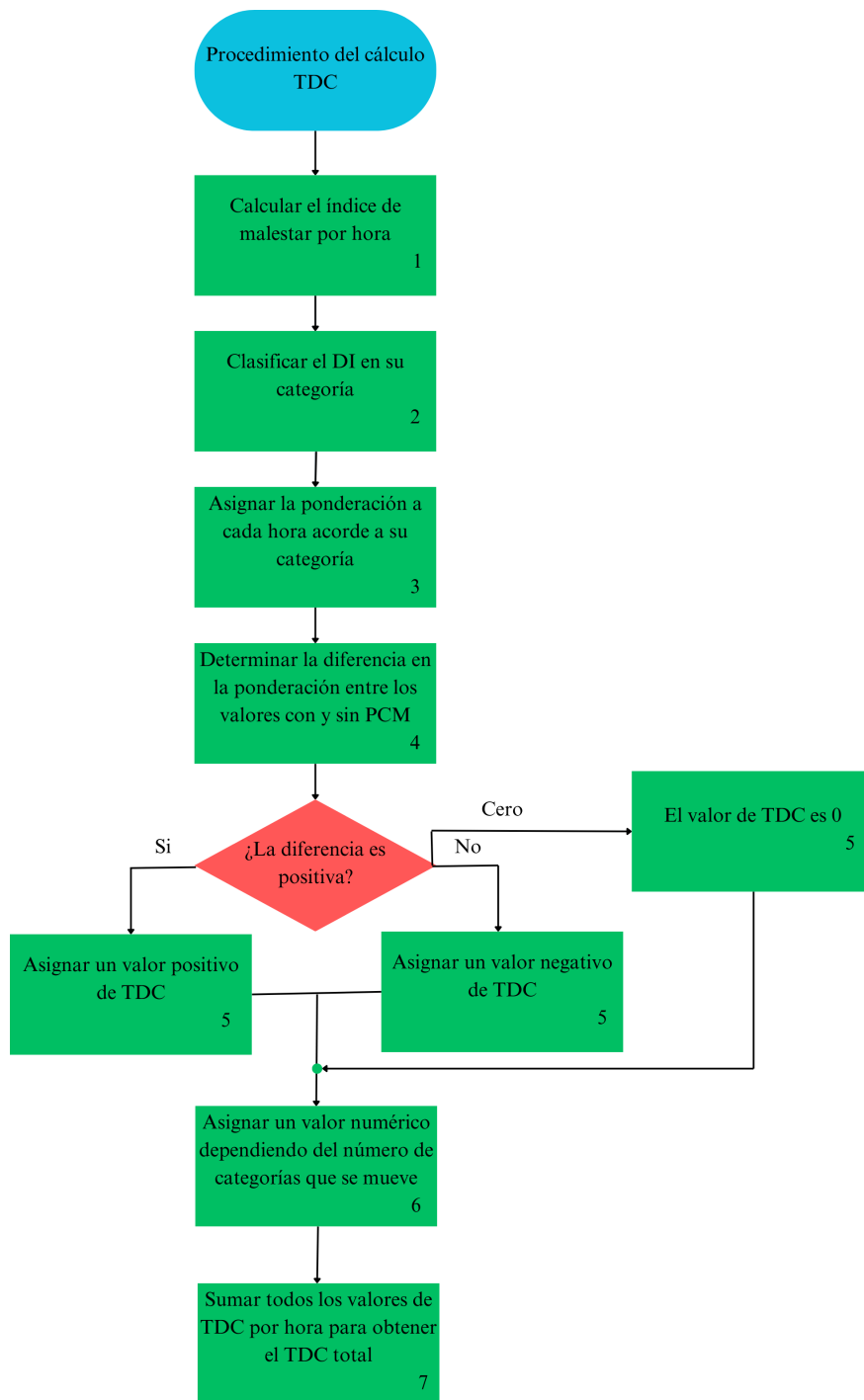


Figura 3.2: Diagrama de flujo utilizado para calcular el TDC.

simulación sin PCM con aquellos del caso en el que se incluye PCM, se restan los valores de ponderación asignados a ambos casos para esa misma hora.

4. El valor de esta diferencia determina el signo y el valor numérico del cambio de malestar.
5. Si la diferencia es positiva, significa que el uso de PCM provocó un desplazamiento hacia un rango de menor malestar térmico, resultando en un cambio de malestar positivo. Si la diferencia es negativa, indica que el uso de PCM causó un desplazamiento hacia un rango de mayor malestar térmico, lo que genera un cambio de malestar negativo. Finalmente, si la

diferencia es igual a 0, implica que el uso de PCM no produjo un cambio de rango.

6. El valor numérico del cambio dependerá del número de rangos que haya sido afectado por el uso de PCM, como se ilustra en la Fig. 3.1.
7. Una vez calculados todos los valores de cambio de malestar, el TDC se obtiene mediante una suma aritmética de estos valores para el mismo PCM y ciudad durante el periodo de simulación.

3.2 Resultados y discusión del capítulo

Las tablas 7 y 8 contienen comparaciones entre los valores de TDC y las horas totales de malestar con ponderación por ciudad y por PCM para el periodo de simulación.

Tabla 7: Comparación de los valores de TDC por ciudad y por PCM.

Ciudad	Estación	PCM 21	PCM 26	PCM 31
Tuxtla Gutiérrez (Af)	Verano	-2235	-2395	-2403
Villahermosa (Am)		1410	1329	1127
Mérida (Aw)		-1938	-1919	-2272
Monterrey (BS)		1660	1628	1268
Hermosillo (BW)		-308	-327	-627
Xalapa (Cf)		-43	-161	-214
Guadalajara (Cw)		115	-27	-10
Toluca (Cs)		1533	1188	993
Tuxtla Gutiérrez (Af)	Invierno	478	498	430
Villahermosa (Am)		832	956	832
Mérida (Aw)		-463	-459	-494
Monterrey (BS)		1065	809	641
Hermosillo (BW)		478	400	300
Xalapa (Cf)		1106	875	738
Guadalajara (Cw)		236	167	188
Toluca (Cs)		1064	752	605

En ambas tablas se puede observar que, en general, el PCM 21 tiene un mejor desempeño en cuanto al cambio total de malestar y la disminución de horas ponderadas de malestar.

En la Figura 3.3a, se muestra la variación de la temperatura interna del modelo de vivienda simulado en Monterrey (BS) durante el invierno, hora por hora, durante 168 horas (7 días) de simulación. Esta gráfica presenta la temperatura en la pared correspondiente a la zona de sala/cocina/comedor, tanto para el caso con PCM como para el caso sin PCM; se observa que la pared sin PCM alcanza picos de temperatura más elevados que la pared con PCM, aunque ambas mantienen una fluctuación en torno a una temperatura media semejante. Esto sugiere que, cuando las condiciones climáticas están próximas al rango de fusión del PCM (21 °C), el material puede absorber

Tabla 8: Comparación de los valores de horas ponderadas de malestar por ciudad y por PCM.

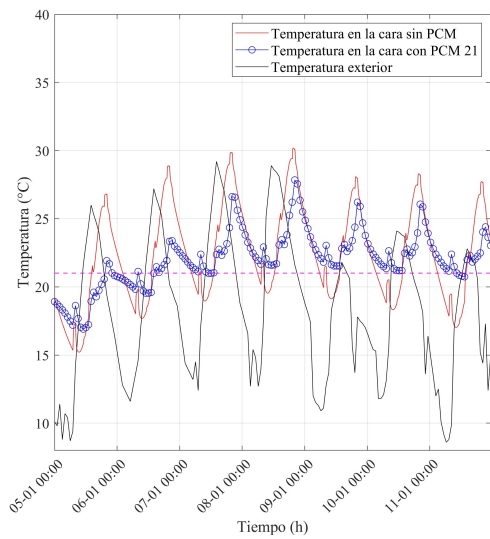
Ciudad	Estación	No PCM	PCM 21	PCM 26	PCM 31
Tuxtla Gutiérrez (Af)	Verano	887.73	1128.76	1146.3	1147.18
Villahermosa (Am)		2055.31	1898.38	1907.25	1929.2
Mérida (Aw)		1577.58	1791.02	1788.92	1827.35
Monterrey (BS)		1994.09	1810.18	1813.69	1852.79
Hermosillo (BW)		2280.56	2312.95	2315.04	2347.35
Xalapa (Cf)		999.28	1003.98	1016.85	1022.68
Guadalajara (Cw)		994.65	981.94	997.51	995.65
Toluca (Cs)		475.44	306.79	344.74	366.2
Tuxtla Gutiérrez (Af)	Invierno	986.83	933.99	931.76	939.2
Villahermosa (Am)		1138.74	1046.92	1033.19	1046.77
Mérida (Aw)		897.88	948.66	948.13	951.96
Monterrey (BS)		566.23	448.96	477.09	495.68
Hermosillo (BW)		491.36	438.71	447.19	458.36
Xalapa (Cf)		537.85	416.14	441.5	456.67
Guadalajara (Cw)		476.68	450.67	458.1	456
Toluca (Cs)		603.66	486.53	520.85	537.02

o ceder calor de manera efectiva, contribuyendo a limitar los aumentos bruscos de temperatura en la pared.

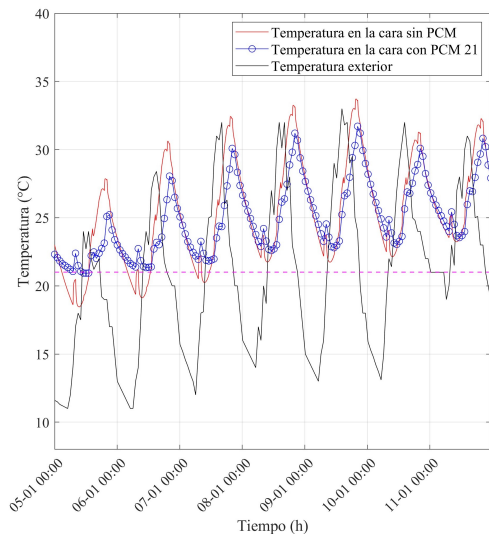
Al comparar las dos curvas (con y sin PCM), se aprecia que la incorporación del PCM suaviza las fluctuaciones. Dicho de otro modo, el material actúa como un amortiguador térmico, en la medida en que almacena calor cuando la temperatura supera ligeramente los 21 °C y lo libera cuando la temperatura desciende de ese umbral. De esta forma, el interior de la vivienda experimenta variaciones de temperatura menos pronunciadas durante las horas más cálidas del verano de Monterrey (BS).

La Figura 3.4a muestra la evolución de la temperatura en 10 nodos distribuidos a lo largo de la pared. Note que el nodo 1 se ubica más cerca del ambiente exterior, mientras que el nodo 10 se encuentra próximo al ambiente interior. Aunque EnergyPlus no especifica la posición exacta de cada nodo, se infiere que el nodo 10 es el más cercano al PCM al presentar un comportamiento más estable y amortiguado. Este patrón sugiere que la capacidad de almacenamiento y liberación de calor del PCM influye de manera notable en la porción de pared más próxima al interior, mejorando la moderación térmica dentro de la edificación.

Por su parte, la Fig. 3.3b usa un esquema análogo para ilustrar las condiciones de temperatura en la ciudad de Tuxtla Gutiérrez (Af) durante el invierno (168 horas de simulación). Esta gráfica muestra nuevamente la temperatura de la pared en la zona sala/cocina/comedor con y sin PCM, presentando un comportamiento distinto al de Monterrey en lo que respecta a los valores máximos alcanzados. Se observa que a lo largo de este periodo de observación, la temperatura exterior se va desviando hacia una región encima de la temperatura de fusión del PCM, lo cual reduce progresivamente su efectividad para estabilizar el perfil térmico de la pared y, por extensión, del espacio

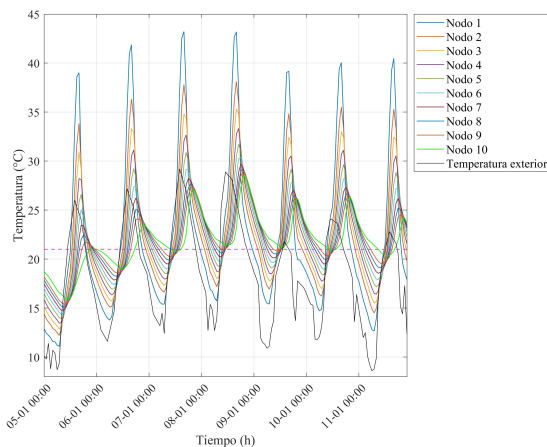


(a) Variación de la temperatura interior en una de las paredes en la ciudad de Monterrey (BS) durante el invierno, por hora, con y sin PCM.

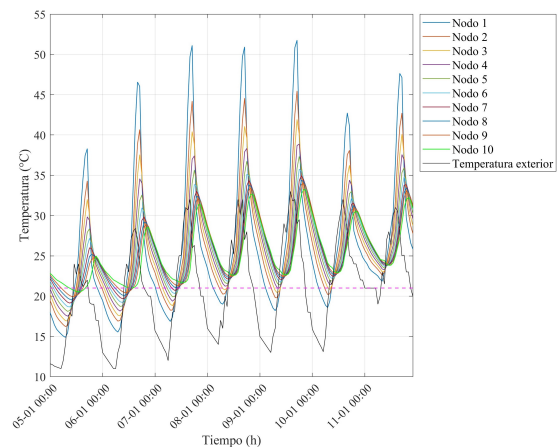


(b) Variación de la temperatura interior en una de las paredes en la ciudad de Tuxtla Gutiérrez (Af) durante el invierno, por hora, con y sin PCM.

Figura 3.3: Variación de las temperaturas interiores.



(a) Variación de la temperatura en cada nodo dentro de la configuración de la pared incluyendo PCM en Monterrey (BS).



(b) Variación de la temperatura en cada nodo dentro de la configuración de la pared incluyendo PCM en Tuxtla Gutiérrez (Af).

Figura 3.4: Variación de las temperaturas interiores.

interior. Esto se aprecia claramente en los últimos días de la comparación de las temperaturas, donde la diferencia entre el caso con y sin PCM se reduce progresivamente.

La Figura 3.4b vuelve a mostrar la temperatura en los 10 nodos de la pared, siguiendo la misma lógica de ubicación (nodo 1 como el más externo y nodo 10 como el más cercano al interior/PCM). Al mantenerse la pared expuesta a temperaturas elevadas y alejadas de los 21 °C, el PCM no encuentra tantas oportunidades para realizar la transición de fase y, en consecuencia, no puede atenuar de manera significativa las fluctuaciones térmicas de la pared.

Al comparar las Figs. 3.3a y 3.3b, se observa que el PCM resulta más eficiente en Monterrey (BS), puesto que las temperaturas externas en esta localidad tienden a oscilar alrededor de los 21 °C, favoreciendo la acción amortiguadora del material. En contraste, en Tuxtla Gutiérrez (Af), la temperatura ambiente supera ampliamente dicho rango de fusión, lo que disminuye la capacidad del PCM para absorber y liberar calor de manera útil. De este modo, puede concluirse que la eficacia del PCM depende de la relación entre la temperatura de fusión del material y las condiciones climáticas de la región, lo que explica su mayor o menor rendimiento en cada caso.

En la Figura 3.5, se muestra la variación del DI a lo largo del tiempo durante un período de simulación de cuatro días en Monterrey (BS) durante el invierno. Se comparan las curvas resultantes con y sin el uso de PCM, junto con los valores de DC por hora. El valor de DC indica el cambio en el rango de incomodidad térmica cuando se utiliza PCM en el modelo de vivienda. El valor numérico del DC representa el número de rangos de DI que cambian con el uso de PCM, mientras que el signo indica si el cambio disminuye (signo positivo) o aumenta (signo negativo) el malestar. Se observan cambios en las categorías del DI, con líneas horizontales que representan las categorías de malestar térmico. Por ejemplo, a la hora 10, el DI sin PCM indica una categoría de malestar H (20-40 % de malestar por frío), mientras que con PCM, cambia a la categoría I (menos del 20 % de malestar por frío), indicando una disminución del malestar y por lo tanto un valor de DC de +1 para esa hora. Por el contrario, a la hora 80, el DI sin PCM indica la categoría de malestar nulo J, mientras que con PCM indica entre 22 – 23.5 °C (20-40 % de malestar por calor), lo que resulta en un aumento del malestar y por lo tanto un valor de DC de -2 debido a un cambio de dos categorías.

En las horas donde los valores del DI permanecen en la misma categoría, el valor de DC es 0, lo que indica que el rango de malestar térmico permanece constante.

Para visualizar el rendimiento general del uso de PCM, la Fig. 3.6 presenta un gráfico de barras que muestra las horas totales de malestar ponderadas en verano. Estas horas representan los periodos durante los cuales los residentes de la vivienda experimentan algún grado de malestar térmico, calculado por la suma de todos los pesos asignados a cada hora de simulación, según la categoría correspondiente de malestar térmico. En verano se observa que en Villahermosa (Am), Monterrey (BS) y Toluca (Cs), el uso de PCM reduce las horas totales de malestar térmico. En Xalapa (Cf) y Guadalajara (Cw), las horas de malestar son similares con y sin PCM, mientras que

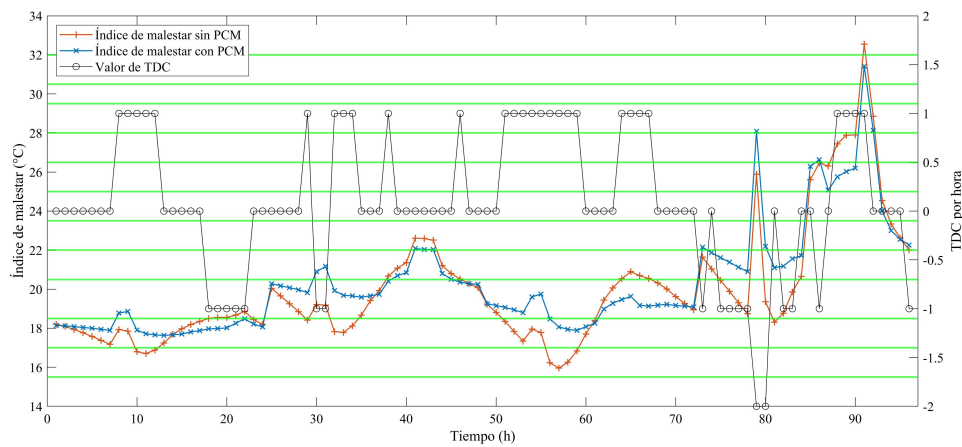


Figura 3.5: Variación del DI por hora en Monterrey (BS) durante el invierno, comparando con y sin PCM, junto con los valores de DC.

en Tuxtla Gutiérrez (Af) (como se muestra en la Fig. 3.6), Mérida (Aw) y Hermosillo (BW), el uso de PCM aumenta las horas de malestar térmico. En [48] mencionan que resultados similares se reportan tanto en la literatura como en sus propios estudios para ciudades en Kazajistán durante el verano, donde las temperaturas también son altas. En ciudades con altas temperaturas en verano, el uso de PCM puede aumentar las horas de malestar térmico, ya que el PCM tiende a liberar calor durante las horas de baja temperatura externa, generalmente por la noche, lo que puede infiltrar calor hacia el interior de la vivienda y afectar negativamente el confort térmico. Sin embargo, en algunas ciudades seleccionadas durante el verano, como Tuxtla Gutiérrez (Af) y Hermosillo (BW), la temperatura exterior durante el día permanece principalmente por encima del punto de fusión del PCM de 21 °C, manteniendo el PCM en estado líquido. Posteriormente, por la noche, la temperatura exterior se acerca al punto de fusión del PCM, lo que provoca que el material libere calor y aumente la temperatura interior. En Mérida (Aw), las temperaturas exteriores durante el verano son en su mayoría más altas que el punto de fusión del PCM, por lo tanto, el PCM puede actuar como otra capa de material en la vivienda y causar un aumento de la temperatura interior.

Por otro lado, en la Fig. 3.6 también se muestra un gráfico de barras con las horas totales de malestar ponderadas en invierno. En este caso, se observa que, excepto en Mérida (Aw), todas las ciudades experimentaron una reducción en las horas de malestar al utilizar PCM. Esto indica que el uso de PCM es más efectivo para reducir el malestar térmico en invierno que en verano.

La Figura 3.7 muestra el TDC total del PCM en cada ciudad y en cada temporada del periodo de simulación. Se puede observar que el valor del TDC se corresponde con las horas de malestar mostradas en la Figura 3.6, donde las ciudades con mayores horas de malestar al usar PCM tienen un TDC negativo. Un TDC negativo indica un aumento en el malestar térmico con el uso de PCM, mientras que un TDC positivo indica lo contrario. En la Figura 3.7, se observa que el mayor TDC para el verano se registra en la ciudad de Monterrey (BS), mientras que para el invierno es en la ciudad de Xalapa (Cf). En Monterrey (BS) se redujeron 183.91 horas de malestar en verano (9.22 %), mientras que en Xalapa (Cf) durante el invierno se redujeron 121.71 horas de malestar

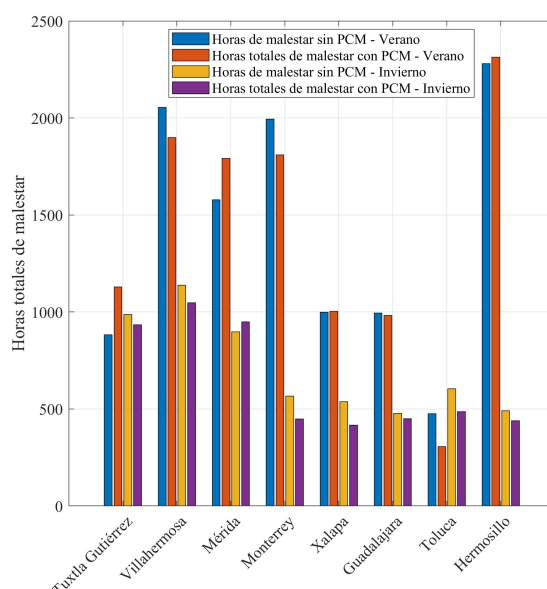


Figura 3.6: Horas totales de malestar ponderadas en verano e invierno.

(22.63 %).

Por otro lado, el TDC más bajo para el verano se encuentra en la ciudad de Tuxtla Gutiérrez (Af), mientras que para el invierno es en la ciudad de Mérida (Aw). También es notable que las ciudades de Villahermosa (Am), Monterrey (BS), Guadalajara (Cw) y Toluca (Cs) mostraron un TDC positivo en ambas estaciones durante el periodo de simulación, mientras que la ciudad de Mérida (Aw) tuvo un TDC negativo en ambos casos. Se puede concluir que Monterrey (BS) mostró un mejor rendimiento mejorando el confort térmico con el uso de PCM, dado que tuvo el valor más alto de TDC combinando los valores de invierno y verano, mientras que Mérida (Aw) exhibió el peor desempeño, con un TDC negativo en ambas estaciones. En Tuxtla Gutiérrez (Af) aumentaron 246.03 horas de malestar en verano (27.8 %), mientras que en Mérida (Aw) durante el invierno se redujeron 121.71 horas de malestar (5.66 %).

Para analizar en detalle cómo el PCM afecta el rango de malestar térmico durante el periodo de simulación, se elaboraron gráficas de barras que muestran las horas correspondientes a cada rango de malestar térmico en distintas ciudades. Las Figuras 3.8a y 3.8b presentan estas gráficas para Monterrey (BS) en verano y Xalapa (Cf) en invierno, respectivamente, que representan las ciudades con mejor desempeño del PCM en la reducción del malestar térmico. Por otro lado, las Figuras 3.9a y 3.9b muestran las gráficas para Tuxtla Gutiérrez (Af) en verano y Mérida (Aw) en invierno, las cuales representan las ciudades con el peor desempeño.

En la Fig. 3.8a se puede observar que, a pesar de que no hay horas en el rango de no malestar, debido a las altas temperaturas que se experimentan en Monterrey durante el verano, la mayoría de las horas de malestar con PCM disminuyen significativamente en los rangos de malestar alto por calor. Estas horas se concentran principalmente en el rango de 60 % – 80 % de las personas en el

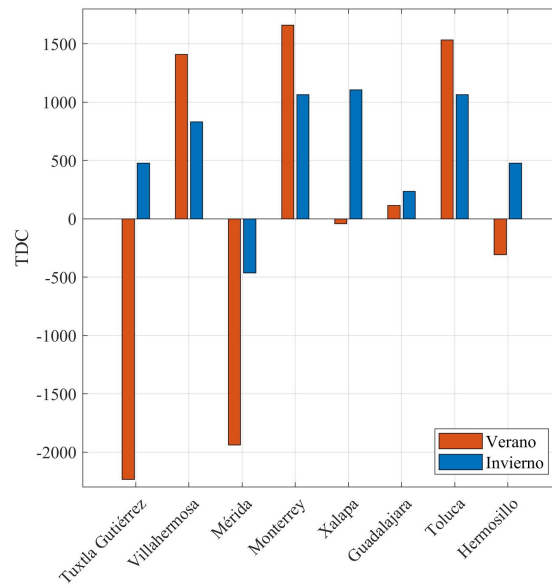
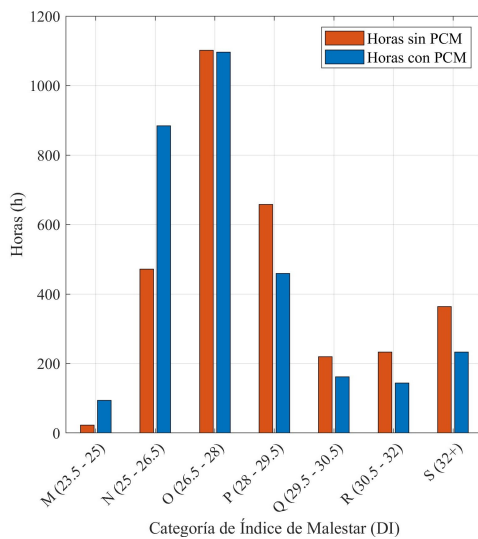
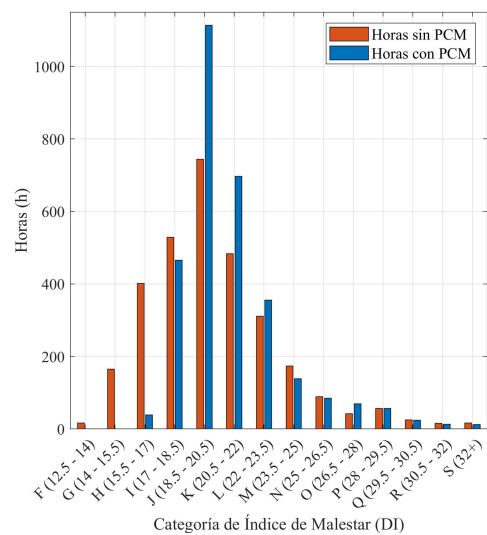


Figura 3.7: TDC del PCM en cada ciudad y estación del período de simulación.

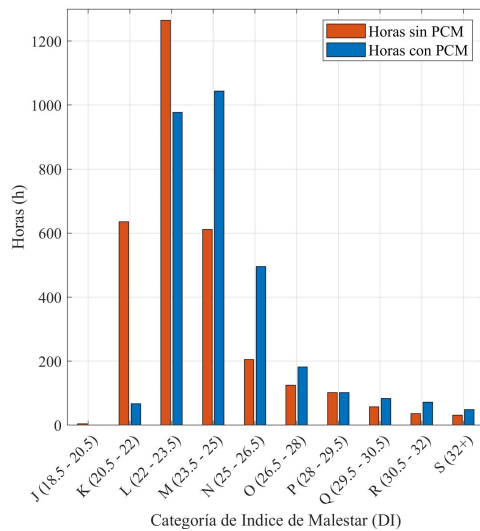


(a) Cantidad de horas dentro de un mismo rango para Monterrey (BS) en verano.

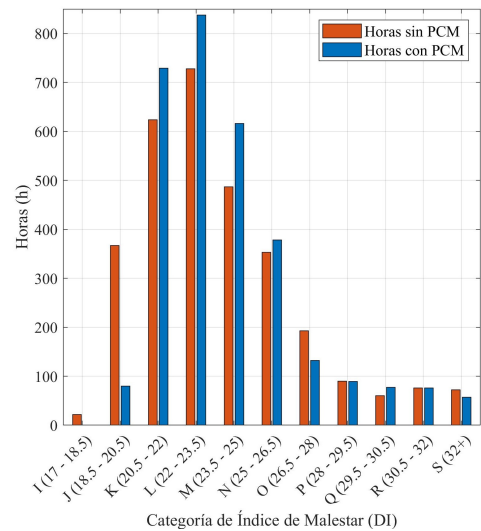


(b) Cantidad de horas dentro de un mismo rango para Xalapa (Cf) en invierno.

Figura 3.8: Cantidad de horas del periodo de simulación dentro de un mismo rango de DI para las ciudades con mejor desempeño.



(a) Cantidad de horas dentro de un mismo rango para Tuxtla Gutiérrez (Af) en verano.



(b) Cantidad de horas dentro de un mismo rango para Mérida (Aw) en invierno.

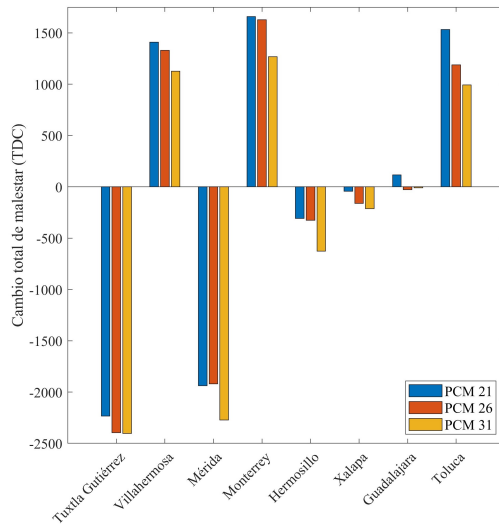
Figura 3.9: Cantidad de horas del periodo de simulación dentro de un mismo rango de DI para las ciudades con peor desempeño.

recinto que sienten malestar por calor. Esta mejora es significativa ya que reduce el malestar severo causado por las altas temperaturas. En la Fig. 3.8b se observa que al utilizar PCM, aumentan las horas en los rangos de no malestar y de menos del 20 % de personas que sienten malestar. Sin embargo, las horas de malestar significativo también aumentan ligeramente. Además, se reducen las horas de malestar por frío.

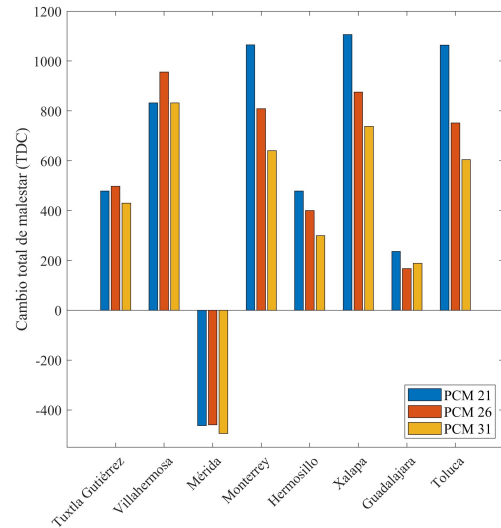
Por el contrario, en la Figura 3.9a se observa que, al utilizar PCM, las horas en los rangos de menor malestar por calor disminuyen. Esto se debe a que el amortiguamiento del PCM hace que estas horas tiendan a concentrarse en el rango de 40 % – 60 % de las personas que experimentan malestar por calor. En la Figura 23 se observa que, de manera similar a la Figura 22, las horas en los rangos de no malestar disminuyen, concentrándose en los rangos de 20 % – 40 % y 40 % – 60 % de personas que sienten malestar por calor. Esto se debe al efecto de amortiguamiento del PCM, que provoca una mayor centralización de las horas en estas categorías, aumentando con ello el malestar térmico.

En la Fig. 3.10 se presenta una comparación de los valores del cambio total de malestar para cada uno de los tres PCM y para cada ciudad durante el periodo de simulación.

La Figura 3.10a muestra la comparación de los valores de TDC para el periodo de verano, donde se observa que el PCM 21 presenta el mejor desempeño en comparación con los demás, ya que sus valores de TDC son los más positivos en todas las ciudades. Por otro lado, en el periodo de invierno, como se ilustra en la Fig. 3.10b, el PCM 26 registra los valores de TDC más positivos en las ciudades de Tuxtla Gutiérrez (Af) y Villahermosa (Am). En contraste, para el resto de las ciudades, el PCM 21 sigue mostrando los valores más positivos de TDC.



(a) Comparación del TDC por PCM y ciudad en verano.



(b) Comparación del TDC por PCM y ciudad en invierno.

Figura 3.10: Comparación de los valores de TDC por PCM y ciudad en el periodo de simulación.

Debido a que, como se observa en la Fig. 3.10, los resultados obtenidos para los PCM 26 y PCM 31 en las simulaciones son bastante similares entre sí, se omite un análisis más detallado de los cálculos de confort térmico para estos materiales.

La base de datos analizada en este capítulo fue ampliada para abarcar todo el periodo de verano e invierno del año 2020 (seis meses) con el fin de tener una cantidad mayor de información para tener lista la base de datos de entrenamiento.

Capítulo 4

Desarrollo de un modelo predictivo del confort térmico usando Machine Learning

4.1 Introducción y planteamiento del problema

Entre los objetivos de este trabajo se incluye el uso de algoritmos de predicción basados en Machine Learning para estimar el confort térmico que brinda la utilización de PCM. Estos algoritmos son entrenados y validados utilizando una base de datos generada a partir de casos simulados. Este capítulo se compone de los siguientes rubros: metodología del entrenamiento y desarrollo de los modelos, evaluación del desempeño de los modelos, cambio de parámetros para reducción de error y discusión de resultados.

Se desarrollaron cuatro modelos basados en algoritmos de Machine Learning, los cuales fueron entrenados utilizando la base de datos generada en el capítulo 2. Esta base de datos fue ampliada para abarcar todo el periodo de verano e invierno del año 2020 (julio, agosto y septiembre para verano; enero, febrero y marzo para invierno), ya que un aumento en la cantidad de datos puede contribuir a mejorar la precisión de los modelos. Estos fueron desarrollados con la herramienta de *Estadística y Machine Learning* de MATLAB, empleando cuatro tipos de algoritmos: árbol de regresión empaquetado, regresión lineal con interacción, máquina de soporte vectorial con núcleo Gaussiano y red neuronal amplia. Las variables de entrada consistieron en diversos parámetros que gobiernan el problema de difusión de calor dentro de una edificación con PCM. La salida del modelo corresponde al índice de malestar (DI). En términos técnicos, esto implica la creación de un modelo *black box* derivado de un modelo *white box*, que son las simulaciones térmicas en *Energy-Plus*.

La Fig. 4.1 muestra un diagrama que generaliza la metodología desarrollada para la creación de los modelos de predicción, que consiste en lo siguiente:

- 1 Incorporación de los siguientes parámetros que influyen en el problema:

Climáticos: temperatura exterior, velocidad del viento, grados día, punto de rocío, radiación, humedad relativa y presión atmosférica. Estos se encuentran dentro de un archivo

climático.

No ambientales (espaciales/temporales): hora del día, zona climática Köppen y mes (estación del año).

Los PCM que se utilizarán en el problema, a través de sus temperaturas de fusión: 21 °C, 26 °C y 31 °C.

- 2 Una vez definidos los parámetros, se desarrollaron las simulaciones de la carga térmica en una vivienda mexicana estándar, con y sin materiales de cambio de fase (PCM). Estas simulaciones, presentadas en el capítulo 2, se extendieron a todo el periodo de verano e invierno de 2020.
- 3 A partir de las simulaciones, se obtienen los cálculos de temperatura interior y humedad relativa, ambos en resolución horaria.
- 4 Con base en estos cálculos, se desarrollan los índices de malestar (DI) y el cambio total de malestar (TDC).
- 5 Una vez definidos los parámetros que rigen el problema y realizados los cálculos, se construye la base de datos de entrenamiento. Los parámetros actúan como datos de entrada, mientras que los cálculos constituyen los datos de salida para los modelos. Con esto se determinan los parámetros ambientales que rigen el problema y se simplifican para reducir el modelo.
- 6 Se entrenan los modelos de predicción utilizando algoritmos de aprendizaje supervisado (Machine Learning). Para ello, se emplea el 80% de la base de datos en el entrenamiento y el 20% restante para la validación y prueba de los modelos.
- 7 Finalizada la fase de entrenamiento supervisado, se obtienen los modelos predictivos, evaluados mediante el RMSE y el coeficiente de determinación (R^2) para verificar su confiabilidad estadística.
- 8 De ser necesario, se iteran los pasos del 5 al 7 para mejorar el desempeño de los modelos cambiando los parámetros de entrada.

4.2 Metodología para el desarrollo de los modelos de predicción

En la Fig. 4.2 se muestra un diagrama de flujo con el procedimiento realizado para el entrenamiento y creación de los modelos predictivos utilizando algoritmos de Machine Learning.

El desarrollo de los modelos de predicción utilizando algoritmos de Machine Learning se realizó con ayuda de la herramienta de Estadística y Machine Learning en MATLAB. La metodología seguida se detalla a continuación.

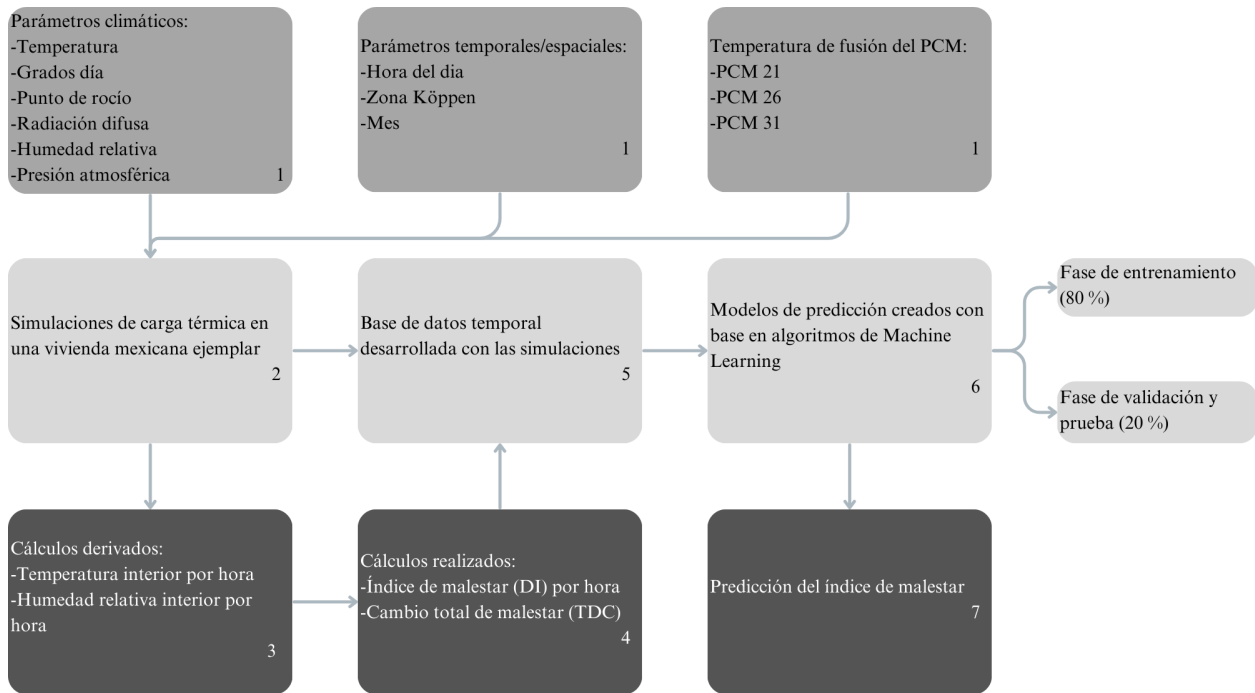


Figura 4.1: Diagrama de la metodología para el desarrollo de los modelos de predicción.

- La base de datos se completó a partir del trabajo presentado en el Capítulo 2, ampliando el desarrollo de simulaciones con y sin PCM para cubrir todo el periodo de invierno y verano del año 2020. Además, se calculó el índice de malestar por hora con los resultados obtenidos de las simulaciones.
- Una vez finalizadas las simulaciones, se organizó la base de datos en Excel, estructurándola de manera que cada columna representara una variable de entrada o la variable de salida. Las variables de entrada incluyen: hora del día, zona climática Köppen, estación del año, temperatura de PCM, grados día, velocidad del viento, temperatura exterior, punto de rocío, radiación difusa, humedad relativa y presión atmosférica. La variable de salida corresponde al índice de malestar calculado para cada hora del día con los valores específicos de las variables de entrada. Cada fila de la base de datos representa una hora específica de un día. La cantidad de datos por variable corresponde a 173 días del periodo de simulación por las 24 horas del día por las 8 zonas climáticas Köppen por los 3 PCM, resultando en 99648 ($173 \times 24 \times 8 \times 3$) datos en total por parámetro.
- En la aplicación de Estadística y Machine Learning de MATLAB se introdujo y procesó la base de datos en Excel, como se puede observar en las Figs. 4.3 y 4.4, donde se definen las variables de entrada, la variable de salida y el porcentaje de datos que se utilizarán para la validación.
- Una vez que se introducen y procesan los datos, se seleccionan los algoritmos que se desean utilizar para poder comenzar el entrenamiento. En este caso, se seleccionaron todos los algoritmos de tipo regresión lineal, árboles de decisión, máquinas de soporte vectorial (SVM) y redes neuronales.

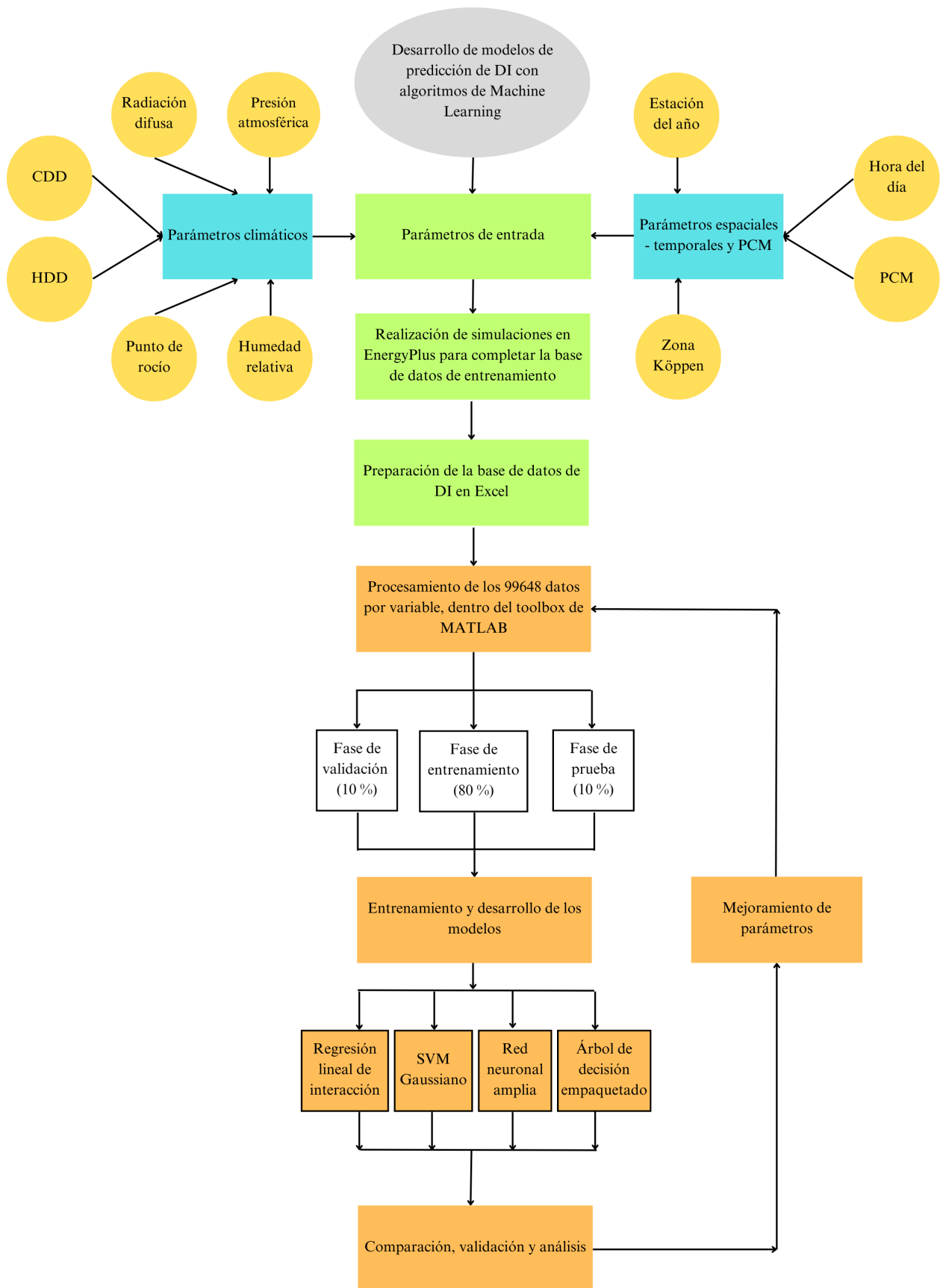


Figura 4.2: Diagrama de flujo de la metodología del entrenamiento y desarrollo de los modelos con algoritmos de ML.

- Después de seleccionar un algoritmo, se comienza con el proceso de entrenamiento con todos estos algoritmos. El tiempo de entrenamiento para cada algoritmo, que depende de la capacidad computacional, se muestra en la Tabla 9.
- Al terminar el proceso de entrenamiento, se obtienen los modelos con cada algoritmo y los resultados de la validación, que incluye el error de root mean squared (RMSE), la R^2 , el error absoluto medio (MAE) y el error cuadrático medio (MSE).
- Se obtiene una métrica de rendimiento preliminar, para después redefinir los parámetros de entrada para reducir el error y aumentar la capacidad predictiva de los modelos.

Tabla 9: Tiempo requerido de cada algoritmo para completar el entrenamiento con la base de datos.

Algoritmo	Tiempo (min)
Regresión lineal de interacción	1
SVM Gaussiano fino	120
Árbol de regresión empaquetado	80
Red neuronal amplia	28

IMPORT

Selection/History

A2:K99649...

Variable Names Row

1

☐ No Variable Names

Name

Base_de_datos5

Type

Table

Table Options

Replace unimportable cells with NaN

UNIMPORTABLE CELLS

Import Selection

IMPORT

SELECTION

IMPORTED VARIABLE

UNIMPORTABLE CELLS

IMPORT

Base_de_datos5.xlsx

	A	B	C	D	E	F	G	H	I	J	K
	Hora_d_a	Zona_K_p...	PCM	Mes	CDD	HDD	P_roc_o	Rad_dif	H_rel_ext	P_atm	DI_Z1
	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number
1	Hora_dia	Zona_Köpp...	PCM	Mes	CDD	HDD	P_rocio	Rad_dif	H_rel_ext	P_atm	DI_Z1
2	1	1	21	1	0.18333	0	19.8	0	73	896.48	23.3745
3	2	1	21	1	0.16667	0	19.4	0	73	896.29	23.2887
4	3	1	21	1	0.14583	0	19	0	73	896.11	23.1963
5	4	1	21	1	0.12917	0	18.6	0	74	895.92	23.1029
6	5	1	21	1	0.10417	0	20	0	83	898.68	23.0143
7	6	1	21	1	0.0625	0	20	0	88	895.88	22.9151
8	7	1	21	1	0.095833	0	18.9	22	79	896.26	24.0027
9	8	1	21	1	0.10417	0	20	70	83	898.68	24.1571
10	9	1	21	1	0.1875	0	20	124	74	898.11	23.0478
11	10	1	21	1	0.31667	0	20.2	148	62	899.29	22.9123
12	11	1	21	1	0.35417	0	19	152	55	900.8	23.0228
13	12	1	21	1	0.4375	0	23.9	194	66	898.42	23.1881

Figura 4.3: Introducción de la base de datos en la aplicación de MATLAB.

4.3 Desarrollo y validación de los modelos predictivos

4.3.1 Métricas de desempeño

Para poder evaluar la precisión de cada uno de los modelos se utilizan métricas establecidas para determinar el error que contienen los datos de predicción contra los datos de entrenamiento, que son el MSE, RMSE, MAE y la R^2 [49].

Data set

Data Set Variable
Base_de_datos5 99648x11 table

Response
DI_21 double 13.9352...

Predictors

	Name	Type	Range
☒	Hora_d_a	double	1 .. 24
☒	Zona_Kppen	double	1 .. 8
☒	PCM	double	21 .. 31
☒	Mes	double	1 .. 6
☒	CDD	double	0 .. 1.0625
☒	HDD	double	0 .. 2

Add All Remove All

[How to prepare data](#)

Validation

Validation Scheme
Holdout Validation

Recommended for large data sets. The app selects a percentage of the data to use as a validation set.

Percent held out 20

[Read about validation](#)

Test

☒ Set aside a test data set

Percent set aside 10

Use a test set to evaluate model performance after tuning and training models. To import a separate test set instead of partitioning the current data set, use the Test Data button after starting an app session.

[Read about test data](#)

Start Session Cancel

Figura 4.4: Procesamiento de la base de datos con la definición de variables y porcentaje para validación.

Error absoluto medio (MAE)

Este error, ya utilizado en la sección 2.4, representa el promedio de las diferencias absolutas entre los valores de predicción y los reales dentro de la base de datos, y mide el valor de los residuos dentro de esta. La Ec. 4.1 es la forma de calcular el MAE, donde x_i son los valores reales y $\hat{x}$ son los valores de predicción.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}| \quad (4.1)$$

Error cuadrático medio (MSE)

Este tipo de error es definido como el promedio del cuadrado de la diferencia entre los valores originales y de predicción en la base de datos, que mide la varianza de los residuos. El MSE se puede ver como la diferencia cuadrática +/- entre ambos tipos de valores. La Ec. 4.2 es la forma de calcular el MSE, donde x_i son los valores reales y $\hat{x}$ son los valores de predicción.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x})^2 \quad (4.2)$$

Error de raíz cuadrática media (RMSE)

Es la raíz cuadrada del error cuadrático medio, y mide la desviación estándar de los residuos. El RMSE se puede interpretar como la diferencia promedio +/- entre los valores real y de predicción. La Ec. 4.3 es la forma de calcular el RMSE, donde x_i son los valores reales y $\hat{x}$ son los valores de predicción.

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x})^2} \quad (4.3)$$

Coefficiente de determinación (R^2)

Representa la proporción de la varianza en la variable dependiente que es explicada por el modelo de predicción, y se mide en porcentaje. Es una medida de la correlación de los datos de predicción con el objetivo de esta, que es un modelo perfecto. En general, este coeficiente se puede ver como el porcentaje más cercano a un modelo perfecto para los datos de predicción en comparación con el uso del valor promedio de los datos. La Ec. 4.4 es la forma de calcular la R^2 , donde x_i son los valores reales, $\hat{x}$ son los valores de predicción, y $\bar{x}$ es el valor promedio de los datos en la base de datos.

$$R^2 = 1 - \frac{\sum (x_i - \hat{x})^2}{\sum (x_i - \bar{x})^2} \quad (4.4)$$

4.3.2 Evaluación de desempeño y reducción de error en los primeros modelos

En el proceso del desarrollo de los modelos de predicción, se comenzó utilizando una base de datos que incluía como variables de entrada la temperatura interior, la radiación directa y horizontal, la velocidad y dirección del viento y la capacidad calorífica, además del punto de rocío, radiación difusa, presión atmosférica y humedad relativa, que se mencionan en la sección anterior. Los resultados obtenidos con el entrenamiento y validación de estos modelos entrenados con la primer base de datos se muestran en la Tabla 10.

Tabla 10: Métricas de desempeño obtenidas con los modelos entrenados con la primer base de datos.

Algoritmo	RMSE	MSE	MAE	R^2
Regresión lineal de interacción	1.1782	1.3882	0.9286	0.8664
SVM Gaussiano fino	0.994	0.9881	0.7395	0.9049
Árbol de regresión empaquetado	0.6409	0.4107	0.4931	0.9605
Red neuronal amplia	0.9481	0.8988	0.7418	0.9135

Para el SVM que se muestra en la Tabla 10, se observa en la Fig. 4.5 la gráfica de respuesta del modelo, donde se agrupan todos los datos de respuesta desde la base de datos, que es el DI y se comparan con los datos de predicción, donde el eje x representa el número de dato en la base de datos, ya que se tienen cerca de 100000, y el eje y es el valor de respuesta de DI en $^{\circ}\text{C}$. En la Fig. 4.6 se aprecia la gráfica de validación sobre qué tanto se acercan los datos a una predicción perfecta.

En la Tabla 10 se encuentran los cuatro tipos de errores mencionados en la sección 4.2 para este modelo, donde se aprecia que se tienen resultados bastante aceptables para un modelo entrenado con una base de datos amplia, como se puede observar en las Figs. 4.5 y 4.6 que los datos de predicción se acercan a la base de datos original, y tienden a acercarse a la predicción perfecta con poca dispersión en los datos, con $R^2 = 90.4\%$. Sin embargo, con el afán de desarrollar modelos predictivos con mejores resultados, reducir los errores y mejorar el valor de R^2 , se eliminaron y cambiaron parámetros de entrada.

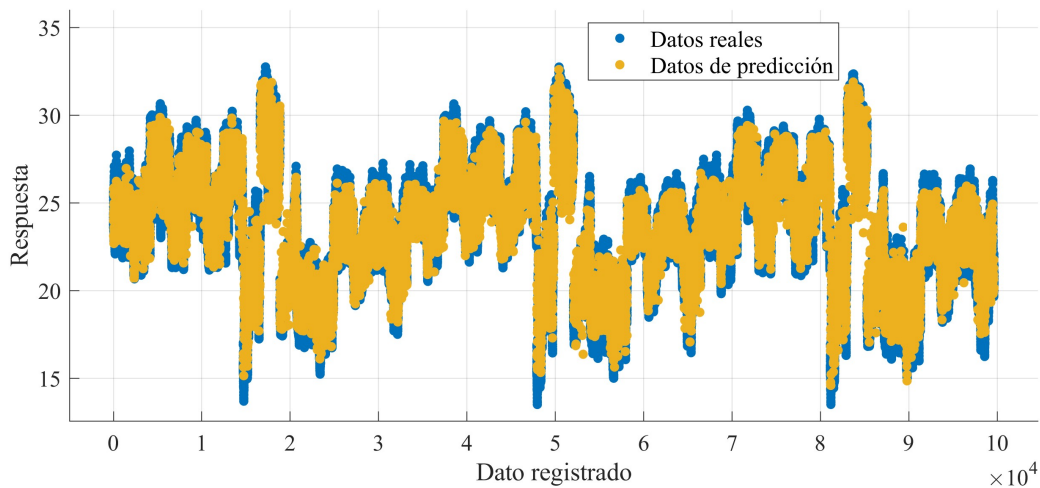


Figura 4.5: Comparación de los valores de respuesta real y predicho del SVM con la primer base de datos.

Para poder mejorar el desempeño de los modelos, se comenzó con la eliminación de variables, debido a que tener demasiados parámetros de entrada puede propiciar a que haya ruido en el modelo y aumente el valor del error, causando sobreajustes. Para esto se crearon gráficas de correlación, en donde se puede observar el comportamiento que tienen las variables climáticas con respecto a la variable de salida, que es el DI. Se conservaron el punto de rocío, la radiación difusa, la humedad relativa exterior y la presión atmosférica. En las Figs. 4.7, 4.8, 4.9 y 4.10 se muestran las correlaciones de estos parámetros de entrada y la salida.

En estas gráficas de variables climáticas vs DI, se agrupa cada una de estas variables de la base de datos por ciudad seleccionada (o zona climática Köppen) y estación del año. Se hace la clasificación por color de todas las combinaciones mencionadas, para observar el comportamiento de cada zona climática y estación. Para cada zona Köppen se utiliza un color distintivo. Por ejemplo, en el caso de Tuxtla Gutiérrez (Zona Köppen 1), se utiliza siempre el color rojo. Ade-

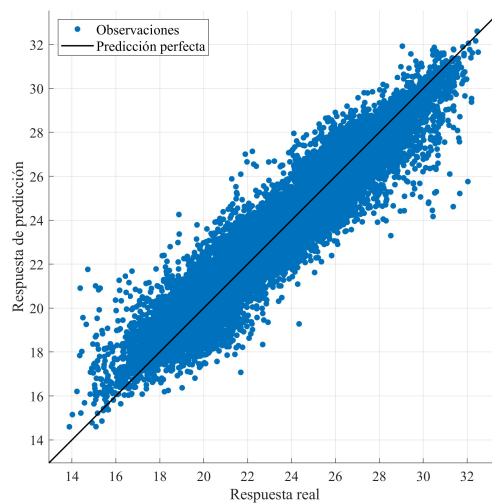


Figura 4.6: Gráfica de la respuesta del SVM con la primer base de datos.

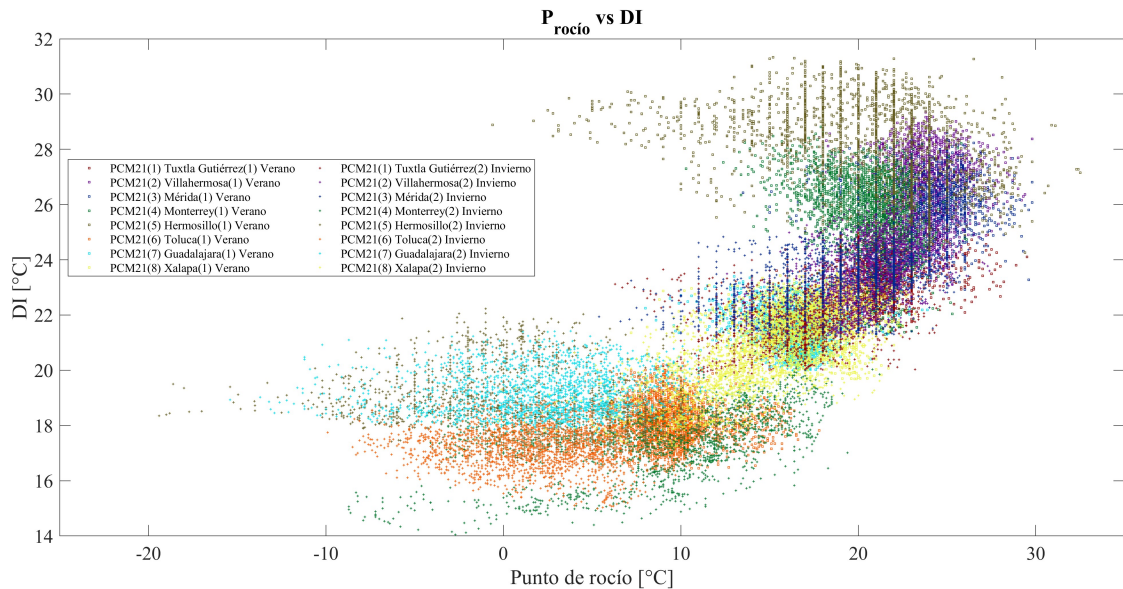


Figura 4.7: Correlación entre el DI y el punto de rocío.

más, para diferenciar las estaciones del año, se representan con cuadros en verano y con cruces en invierno. De este modo, los cuadros rojos indican la Zona 1 durante el verano, mientras que las cruces rojas representan la Zona 1 en invierno. Esta codificación visual facilita la identificación de la posible relación entre la zona Köppen, la estación del año y el parámetro climático seleccionado.

El propósito de esta actividad es el de identificar el tipo de relación que cada variable climática tiene con el DI que es la variable de salida, para así poder descartar aquellas que tienen un comportamiento casi constante y eliminar el posible ruido dentro del modelo que pudiera afectar negativamente su precisión.

La Fig. 4.7 muestra una gráfica de punto de rocío vs DI, donde se puede apreciar el comportamiento no constante de esta variable con respecto a la de respuesta, por lo que se conservó como

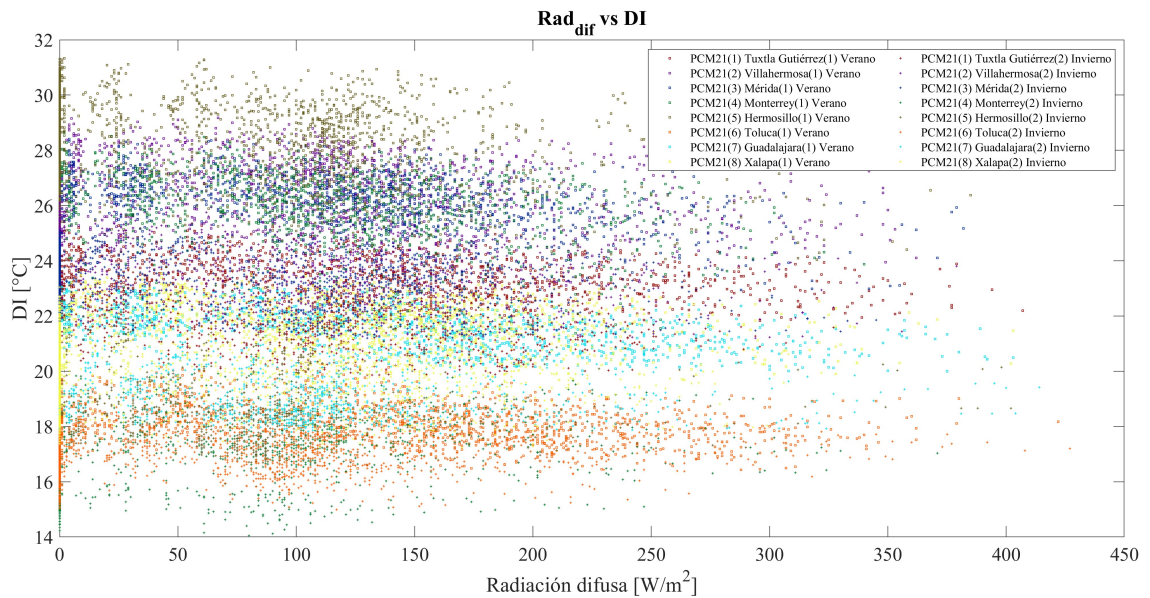


Figura 4.8: Correlación entre el DI y la radiación difusa.

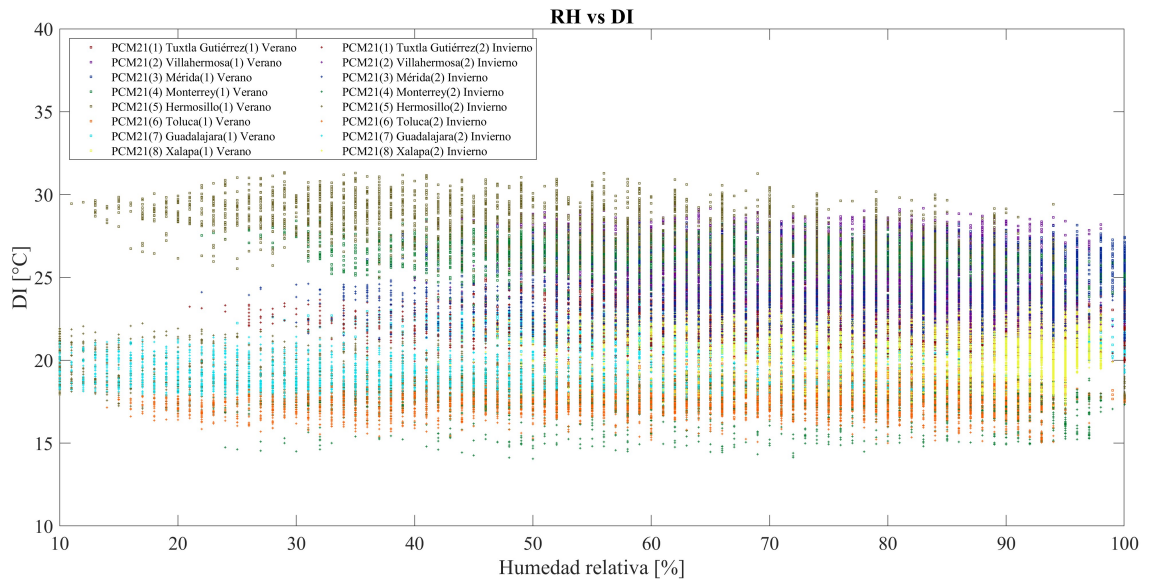


Figura 4.9: Correlación entre el DI y la humedad relativa.

parte de la base de datos para entrenamiento. La siguiente gráfica, mostrada en la Fig. 4.8 es la radiación difusa vs DI, donde se aprecia la relación lineal inversa que esta variable climática posee con respecto a la salida, conservando este tipo de radiación a diferencia de la horizontal o directa debido a que estas tienen un comportamiento más constante. En la Fig. 4.9 se presenta una gráfica de humedad relativa vs DI, donde se observa la relación lineal inversa entre ambas variables, mientras que la Fig. 4.10 se muestra una gráfica de presión atmosférica vs DI, donde se aprecia una relación lineal más sutil pero variable, por lo que se conservó dentro de la base de datos.

Al conservar estas variables, se cambió la utilización de la temperatura interior por un concepto utilizado dentro de la literatura sobre modelos *grey box* relacionados con cargas térmicas dentro de

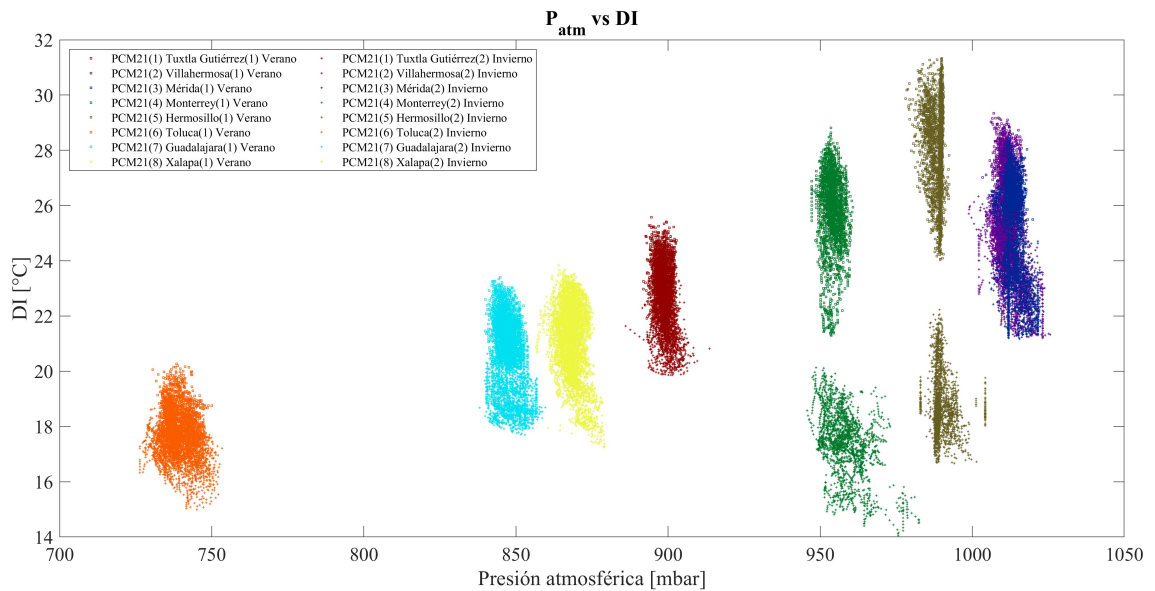


Figura 4.10: Correlación entre el DI y la presión atmosférica.

edificios, conocido como *grados día*. Se decidió incluir las variables de grados día de calefacción (HDD) y grados día de refrigeración (CDD), que se mencionan en la sección 1.1.5, y que son muy utilizados para determinar la demanda de energía [31].

La temperatura base que se utilizó para la obtención de los grados día de calefacción es de 18.5°C , que es la cota inferior del rango de DI que representa el confort térmico en la Tabla 6. Mientras que la temperatura base que se utilizó para la obtención de los grados día de refrigeración es de 20.5°C , que es la cota superior del rango de DI que representa el confort térmico en la Tabla 6.

Para el entrenamiento de los modelos, se considera únicamente la diferencia de temperatura en intervalos de una hora ($1/24$ de un día), sin realizar una suma acumulativa. Si en un momento determinado la temperatura supera los 20.5°C , el HDD será 0. Por otro lado, si la temperatura es inferior a 18.5°C , el CDD será 0. Si la temperatura se encuentra dentro de este rango, tanto el CDD como el HDD tendrán un valor de 0.

En este trabajo, el concepto de grados día se emplea para comparar las temperaturas exteriores con las temperaturas base en un solo día, sin considerar la suma acumulativa ni el efecto integrativo utilizado en la predicción del consumo energético. El objetivo es identificar si el malestar térmico se debe al frío o al calor, considerando las temperaturas dentro del rango de confort como 0.

La Fig. 4.11 es una gráfica de HDD vs DI, donde se muestra la relación inversa que hay entre ambas variables, por lo que se justifica su presencia dentro de la base de datos para entrenamiento.

En la Fig. 4.12, que es una gráfica de CDD vs DI, se puede observar que existe una relación entre los grados día de refrigeración y el índice de malestar, para así añadir los CDD como variable de entrada.

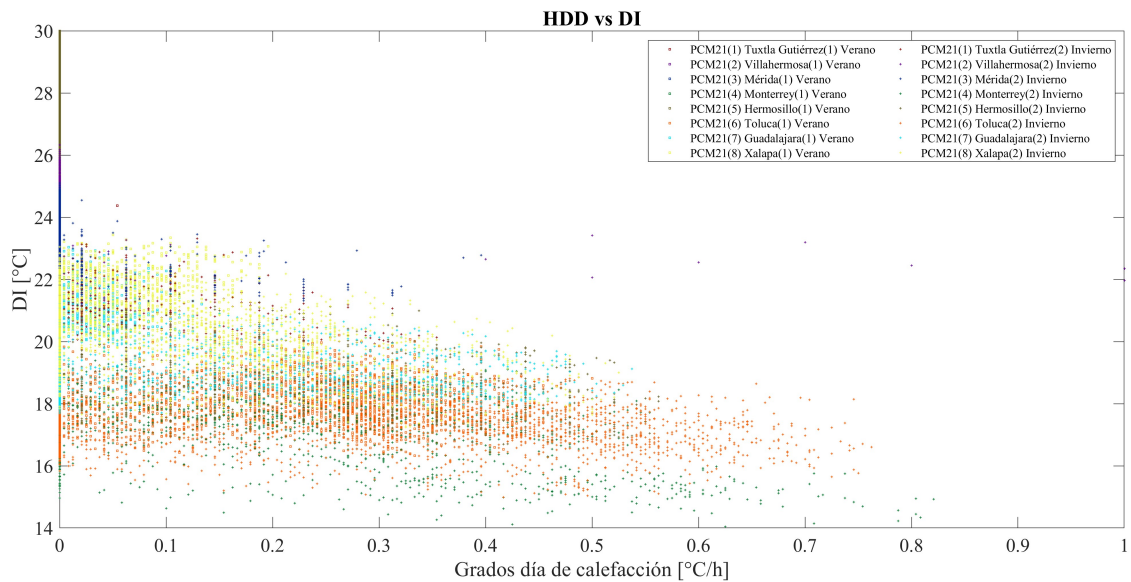


Figura 4.11: Correlación entre el DI y los grados día de calefacción.

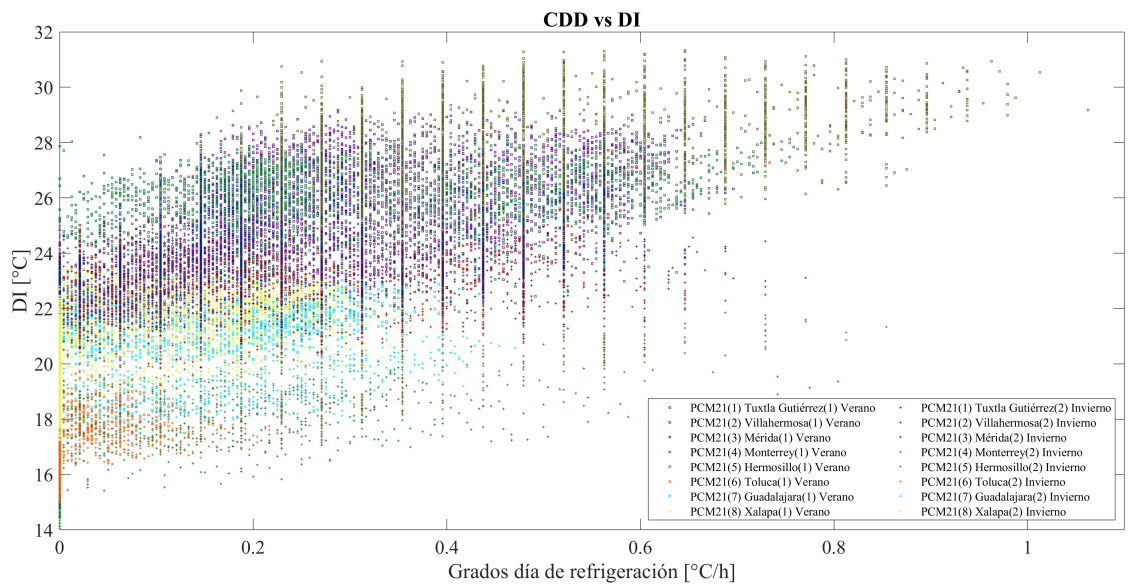


Figura 4.12: Correlación entre el DI y los grados día de refrigeración.

En resumen, las cantidad de variables de entrada con la base de datos ajustada se dividen en tres tipos:

- **Parámetros climáticos:** grados día de refrigeración (CDD), grados día de calefacción (HDD), punto de rocío, radiación difusa, humedad relativa y presión atmosférica.
- **Parámetros temporales/espaciales:** hora del día, zona climática Köppen y estación del año.
- **Temperatura de fusión del PCM:** PCM 21, PCM 26 y PCM 31.

4.3.3 Evaluación de desempeño de los modelos con reducción de error

En la Tabla 11 se muestran los resultados de validación obtenidos con los modelos entrenados con la base de datos más reciente mencionada en la subsección anterior.

Tabla 11: Resultados obtenidos con los modelos entrenados con la base de datos para reducir el error.

Algoritmo	RMSE	MSE	MAE	R^2
Regresión lineal de interacción	1.2284	1.5089	0.9399	0.8779
SVM Gaussiano fino	0.6371	0.4058	0.4695	0.9672
Árbol de regresión empaquetado	0.4331	0.1878	0.3224	0.9848
Red neuronal amplia	0.647	0.4186	0.4939	0.9661

En la tabla anterior, se observa que, en comparación con la Tabla 10, los algoritmos SVM Gaussiano, el árbol de regresión empaquetado y la red neuronal fina obtuvieron mejores resultados, logrando reducir exitosamente los valores de error y aumentar el valor de R^2 . Por otro lado, aunque el algoritmo de regresión lineal presentó un ligero incremento en el valor de R^2 , también mostró un aumento en los valores de error.

El algoritmo con mejores resultados fue el árbol de regresión empaquetado, que alcanzó los valores de error más bajos y un coeficiente de determinación R^2 superior al 98 %. Esto se aprecia en la Fig. 4.13, donde se presentan, en una gráfica de barras, los valores de error y R^2 para cada algoritmo. En contraste, el algoritmo de regresión lineal de interacción mostró los valores de error más altos y un R^2 cercano al 88 %.

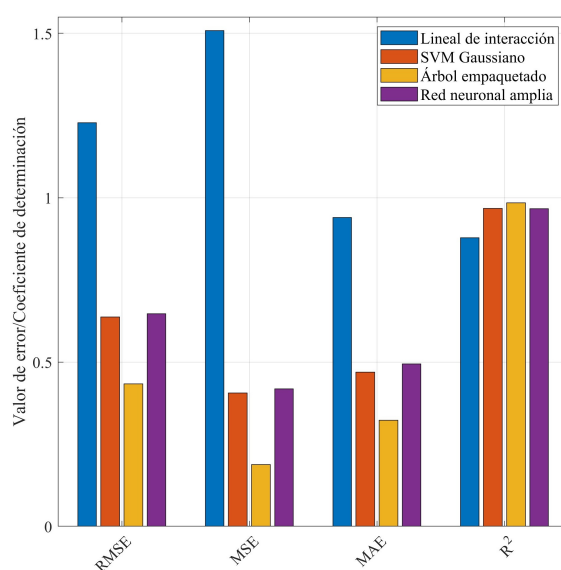
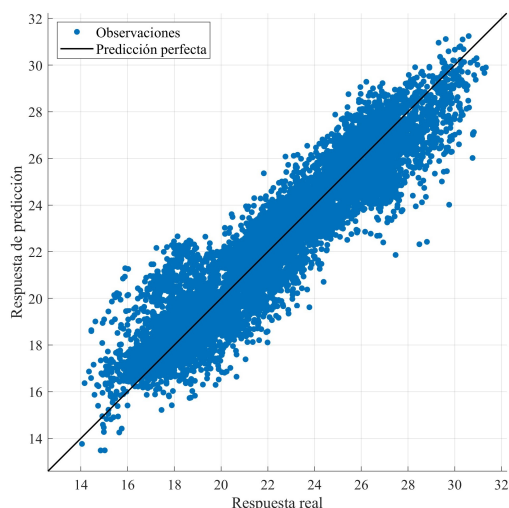
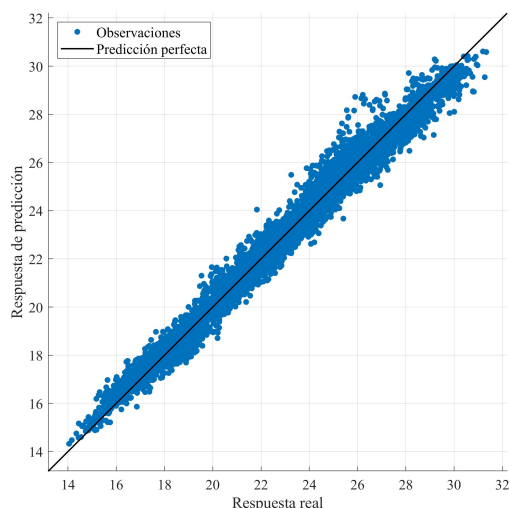


Figura 4.13: Comparación del desempeño entre los algoritmos utilizados.



(a) Gráfica de regresión para el algoritmo de regresión lineal de interacción.



(b) Gráfica de regresión para el algoritmo de ensamble de árboles empaquetados.

Figura 4.14: Gráficas de regresión lineal y de ensamble.

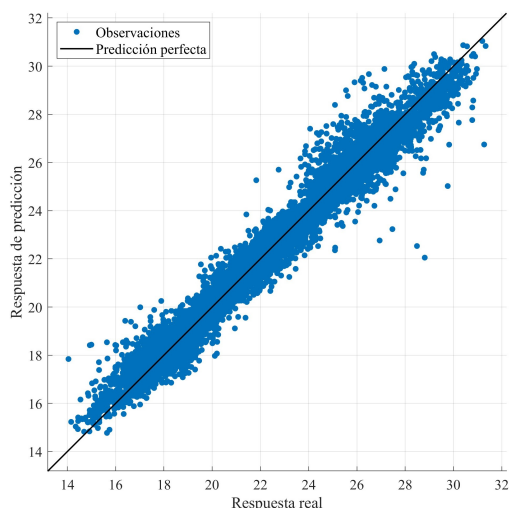
En las Figs. 4.14a, 4.14b, 4.15a y 4.15b se presentan las gráficas de regresión para cada uno de los algoritmos evaluados. En la Fig. 4.14a, muestra la gráfica de regresión para el algoritmo de regresión lineal de interacción, que obtuvo los resultados de validación menos satisfactorios. Esto se refleja en una mayor dispersión de los datos, el menor valor de R^2 , y los mayores valores de error.

Por otro lado, las Figs. 4.15a y 4.15b presentan las gráficas de dispersión para el SVM Gaussiano fino y la red neuronal amplia, respectivamente. Ambas muestran una reducción en la dispersión de los datos en comparación con la Fig. 4.14a, lo que refleja resultados más consistentes entre estos algoritmos.

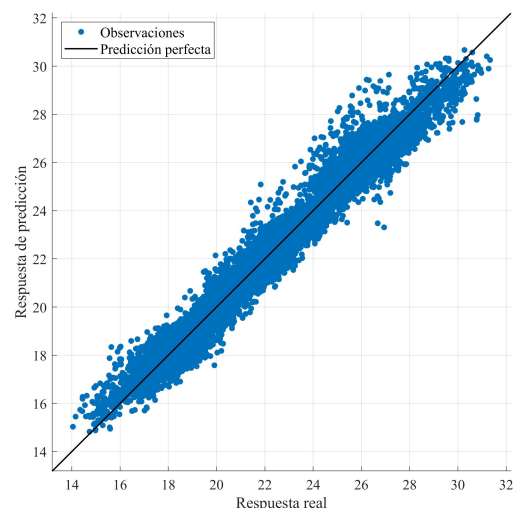
Finalmente, la gráfica de dispersión para el ensamble de árbol empaquetado presentada en la Fig. 4.14b muestra la menor dispersión entre todas las gráficas. Esto se debe a que este algoritmo obtuvo los mejores resultados de validación, logrando los errores más bajos y el valor más alto de R^2 .

La Fig. 4.16 muestra una fuerte correlación entre los datos observados y las predicciones del modelo de ensamble de árboles empaquetados, representado por el índice de dispersión (DI). Este resultado, que confirma los hallazgos de la Fig. 4.5, evidencia el excelente desempeño del modelo en la tarea de predicción. La baja dispersión observada indica un ajuste preciso del modelo a los datos.

La gráfica de residuos para el ensamble de árboles empaquetados (Fig. 4.17) muestra una distribución aleatoria alrededor de cero, sin patrones claros. Esto indica que el modelo se ajusta bien a los datos y que los errores de predicción son aleatorios.



(a) Gráfica de regresión para el algoritmo de SVM Gaussiano fino



(b) Gráfica de regresión para el algoritmo de red neuronal amplia.

Figura 4.15: Gráficas de regresión para el SVM y la red neuronal.

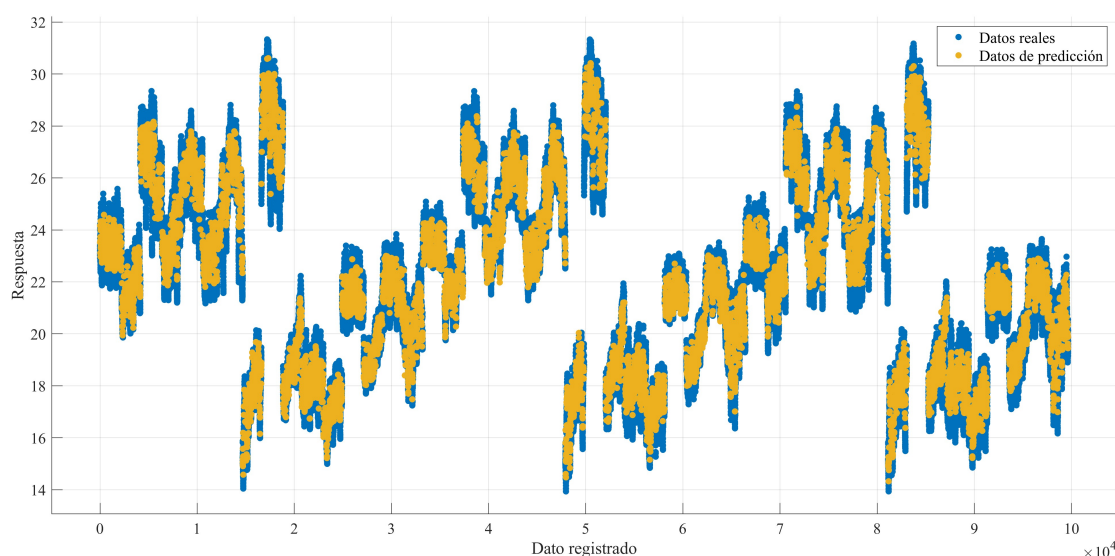


Figura 4.16: Comparación de los valores de respuesta real y predicho del SVM con la segunda base de datos.

Los tres algoritmos con mejor desempeño utilizados en esta tesis para el aprendizaje de regresión fueron: la máquina de soporte vectorial con núcleo Gaussiano (SVM Gaussiana), los árboles de regresión empaquetados y la red neuronal amplia.

El modelo seleccionado para los siguientes resultados fue la SVM, ya que presentó el menor error en las predicciones y permitió evitar la tendencia al sobreajuste observada en los árboles de regresión. A diferencia de una red neuronal, la SVM no requiere capas ni pesos ponderados; en su lugar, resuelve un problema de optimización para clasificar los datos. En el contexto de esta tesis, durante el entrenamiento, la SVM busca generar una *curva* que separe y clasifique de manera

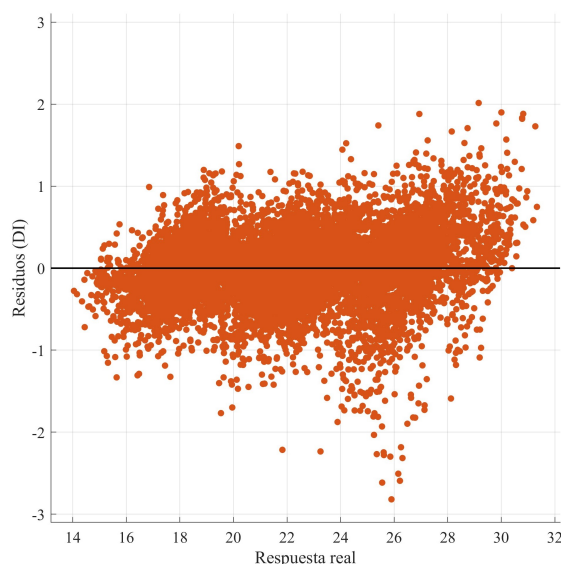


Figura 4.17: Gráfica de residuos para el ensamble de árboles empaquetados.

óptima los valores de DI.

Por ejemplo, si el DI tuviera dos clasificaciones posibles (verde y rojo) y dependiera únicamente de la temperatura, el algoritmo buscaría una línea ($n = 1$) que separe los datos donde DI = verde. Para lograrlo, la SVM transforma los datos a una dimensión superior, por ejemplo, elevándolos al cuadrado, y luego aplica procesos de optimización para determinar los parámetros que permitan encontrar la mejor línea de separación. Este procedimiento se lleva a cabo durante la fase de entrenamiento. Una vez obtenidos estos parámetros, la capacidad del modelo para realizar predicciones se evalúa comparando los datos predichos con los datos de entrenamiento.

Una máquina de soporte vectorial (SVM) traza un límite de separación entre dos tipos de datos. Este límite debe ser lo suficientemente sensible a los datos de entrenamiento para evitar sesgos en la clasificación y, al mismo tiempo, garantizar una baja varianza en los datos. Cuando existen datos mal clasificados, el concepto de *margen suave* determina la distancia de estos puntos con respecto al límite de separación. A través de la validación cruzada, se encuentra el margen óptimo que permite cierto grado de error en la clasificación, priorizando que la mayoría de los datos sean correctamente categorizados. Los puntos de datos que se encuentran dentro del margen suave o en sus bordes se denominan vectores de soporte. El clasificador o límite de decisión es un hiperplano, que en dimensiones menores a tres se convierte en un subespacio de este.

En una máquina de soporte vectorial (SVM), es necesario encontrar el clasificador óptimo para los vectores de soporte según el tipo de datos disponibles. Para lograrlo, los datos se transforman de manera que puedan clasificarse en solo dos categorías. La función Kernel es fundamental en este proceso, ya que permite mapear los datos a un espacio de mayor dimensión donde sean separables. En este caso, se emplea una función Kernel Gaussiana para mejorar la capacidad de clasificación

del modelo.

Los datos de entrenamiento consisten en N pares $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, con $x_i \in \mathbb{R}^n$ y $y_i \in \{-1, 1\}$. Se define un hiperplano, que es el límite entre los datos, por medio de la Ec. 4.5 [51].

$$\{x : f(x) = x^T \beta + \beta_0 = 0\} \quad (4.5)$$

Donde β es un vector unitario: $\|\beta\| = 1$. Debido a que ambas clases son separables, se puede encontrar una función $f(x) = x^T \beta + \beta_0$ con $y_i f(x_i) > 0$. Se desea encontrar el hiperplano que cree el *margen* (M) más grande entre los puntos de entrenamiento para las clases 1 y -1. Se define un problema de optimización que se muestra en las Ecs. 4.6 y 4.7.

$$\max_{\beta, \beta_0, \|\beta\|=1} M \quad (4.6)$$

$$\min_{\beta, \beta_0} \|\beta\| \quad (4.7)$$

Sujeto a $y_i(x_i^T \beta + \beta_0) \geq M$ y $y_i(x_i^T \beta + \beta_0) \geq 1, i = 1, \dots, N$.

4.3.4 Contribución por variable al desempeño del modelo predictivo

Para poder analizar la contribución de cada una de las variables utilizadas en el entrenamiento del modelo, se utiliza el algoritmo de *Redundancia Mínima Relevancia Máxima* (MRMR) en MATLAB, que es un algoritmo iterativo que, como lo indica su nombre, mide la relevancia y la redundancia de un predictor dentro del modelo, asignándole un valor de importancia a cada uno, basado en la razón entre relevancia y redundancia [50].

En la Tabla 12 se muestra el valor de importancia MRMR por cada una de las variables de entrada en el modelo de predicción.

Tabla 12: Valor de MRMR para cada variable de entrada.

Variable predictora	MRMR
Presión atmosférica	0.9501
Mes del año	0.7119
Grados día de calefacción	0.1128
Punto de rocío	0.1021
Zona climática Köppen	0.081
Grados día de refrigeración	0.0705
Humedad relativa externa	0.0429
Radiación difusa	0.0258
Hora del día	0.0201
Temperatura de fusión del PCM	0

Se puede observar en la Tabla 12 que la presión atmosférica y el mes del año son los predictores

con mayor relevancia dentro del modelo, mientras que la radiación difusa y hora del día son los de menor relevancia. La temperatura de fusión del PCM tuvo un valor MRMR de 0, lo que quiere decir que es una variable sin relevancia dentro del modelo, y esto se puede deber a que las variables climáticas son las mismas para los tres PCM, variando únicamente su valor de respuesta, que es el DI.

En las Figs. 4.18, 4.19, 4.20, 4.21, 4.22, 4.23, 4.24, 4.25, 4.26 y 4.27 se muestran gráficas de dependencia parcial entre cada una de las variables predictoras de entrada y la variable de salida (DI). Una gráfica de dependencia parcial muestra la relación promedio entre el predictor y el valor de predicción en el modelo entrenado. Las líneas en verticales en el eje x indican el valor del predictor en la base de datos.

La Fig. 4.18 ilustra la relación entre el DI de predicción y los grados día de refrigeración (CDD), mostrando una proporción directa. Este comportamiento es coherente con la definición de CDD, donde un aumento en su valor implica un incremento en la temperatura dentro de un recinto, lo que a su vez eleva el DI.

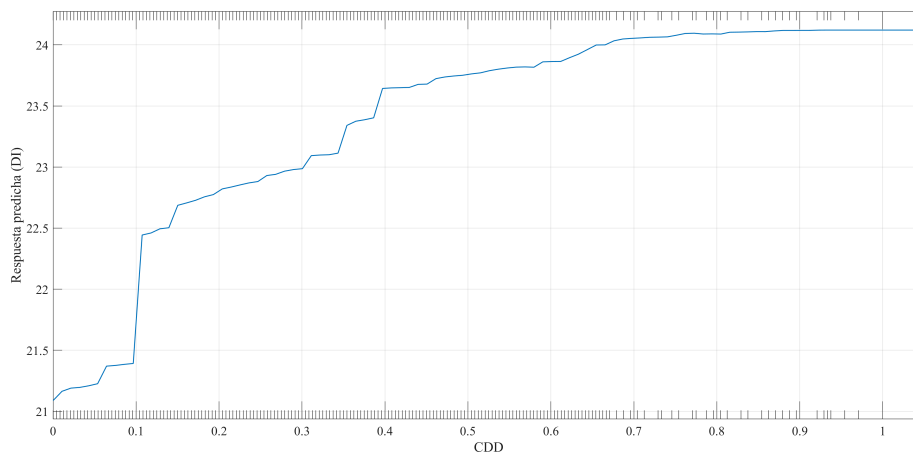


Figura 4.18: Dependencia parcial entre los CDD y la respuesta de predicción.

En la Fig. 4.19, se analiza la relación con la humedad relativa, revelando una proporción inversa. Esto es consistente con la conexión directa entre el DI y la humedad relativa; a medida que la humedad aumenta, el DI tiende a disminuir.

La gráfica de la Fig. 4.20 presenta el comportamiento en relación con los grados día de calefacción (HDD), evidenciando un comportamiento exponencial y una proporción inversa. Un valor elevado de HDD indica temperaturas interiores bajas, lo que se traduce en un DI también bajo. Además, se destaca que en el rango de HDD entre 0 y 0.4, el valor del DI en la predicción es más alto.

En la Fig. 4.21 se aprecia la dependencia parcial entre el DI de predicción y la hora del día. En esta gráfica se muestra que, en promedio para la base de datos, existen aumentos de temperatura

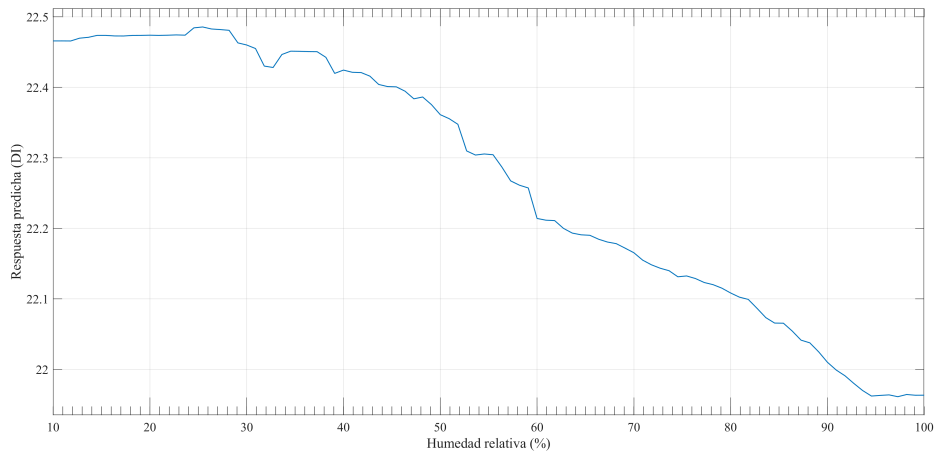


Figura 4.19: Dependencia parcial entre la humedad relativa y la respuesta de predicción.

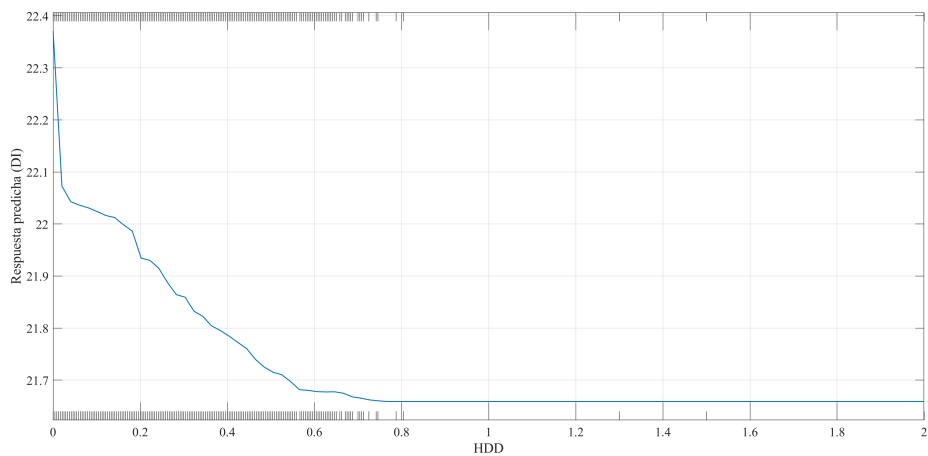


Figura 4.20: Dependencia parcial entre los HDD y la respuesta de predicción.

durante las horas de la mañana, entre 6 am y 9 am, y en las tardes entre 3 pm y 8 pm. Las bajas de temperatura en el día se encuentran en las horas de madrugada, entre 1 am y 5 am, y las horas del día entre 10 am y 3 pm que la temperatura se mantiene casi constante antes de aumentar.

La Fig. 4.22 muestra cómo varía el DI a lo largo de los meses del año. Durante 1 a 3, que corresponden al verano, el DI alcanza su valor máximo, mientras que de 4 a 6, que son meses de invierno, se presenta un valor mínimo. Este comportamiento es coherente con las condiciones climáticas de cada estación.

En la Fig. 4.23 se observa la relación con la presión atmosférica, donde se evidencia una proporción directa. Un aumento en la presión se asocia con un incremento en la temperatura, lo que también eleva el DI.

La Fig. 4.24 presenta el comportamiento en relación con el punto de rocío, mostrando un aumento a partir de una temperatura de punto de rocío de 10 °C. Este comportamiento es consistente

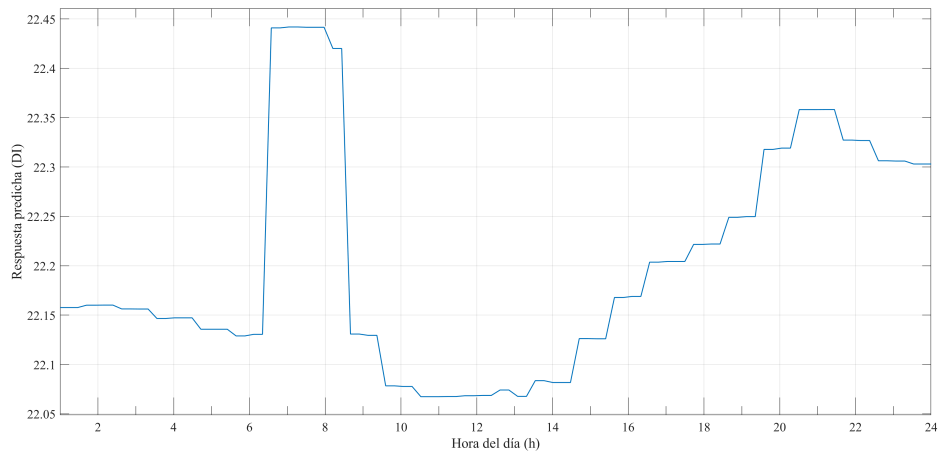


Figura 4.21: Dependencia parcial entre la hora del día y la respuesta de predicción.

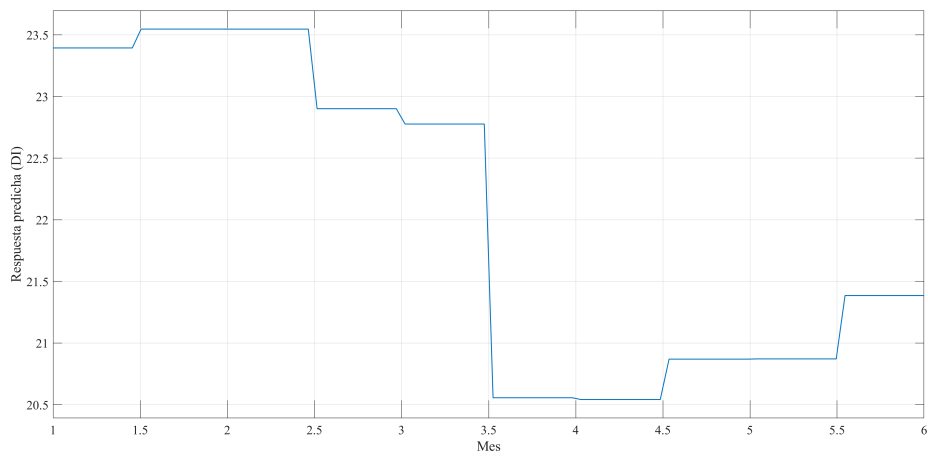


Figura 4.22: Dependencia parcial entre el mes del año y la respuesta de predicción.

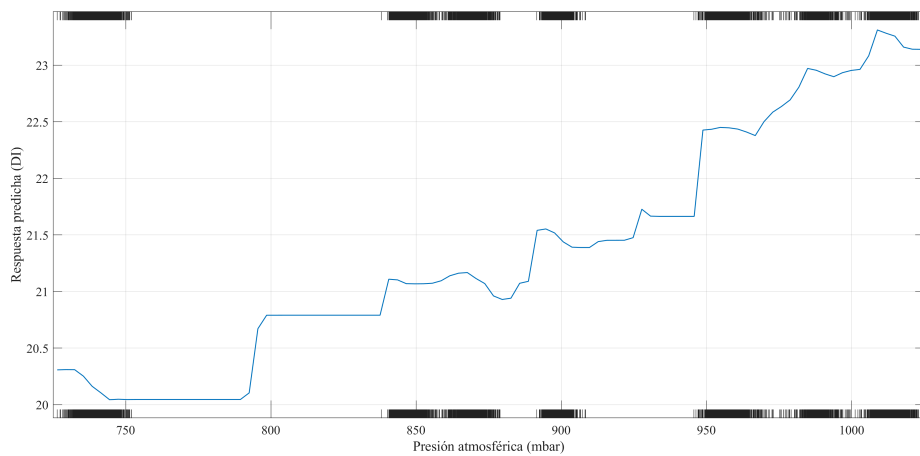


Figura 4.23: Dependencia parcial entre la presión atmosférica y la respuesta de predicción.

con la relación directa entre la humedad relativa y el punto de rocío, indicando que un alto valor en este predictor incrementará el DI.

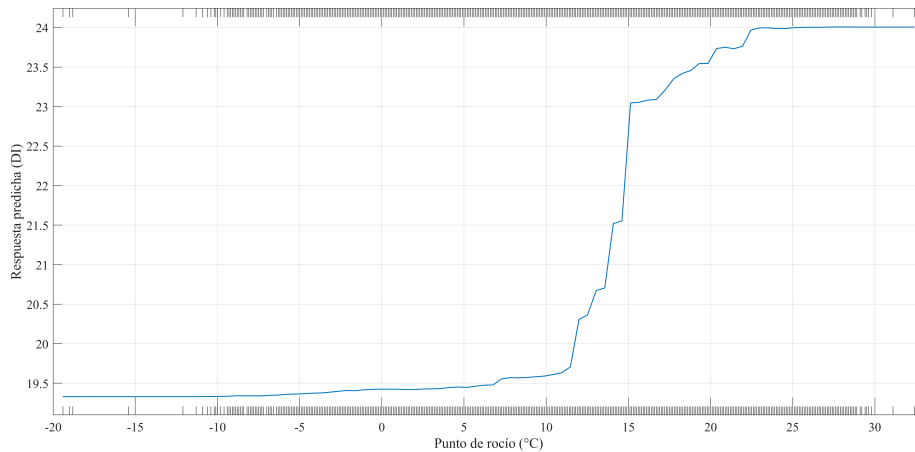


Figura 4.24: Dependencia parcial entre el punto de rocío y la respuesta de predicción.

La Fig. 4.25 muestra la relación con la temperatura de fusión del PCM. En esta gráfica, se observa que la respuesta se mantiene casi constante en función de la temperatura de los tres PCM utilizados, aunque el DI experimenta un ligero aumento con la temperatura de fusión. Este cambio mínimo en el valor de la respuesta de predicción en relación con el predictor explica el valor de MRMR que se presenta en la Tabla 12.

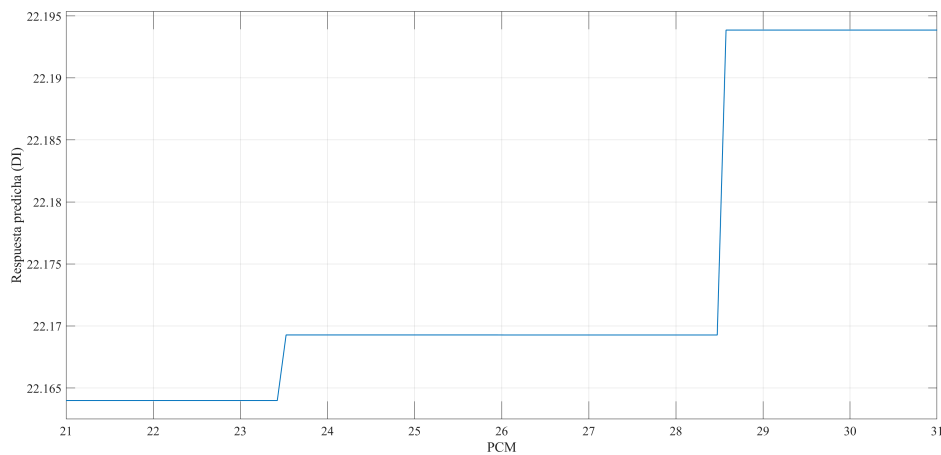


Figura 4.25: Dependencia parcial entre la temperatura de PCM y la respuesta de predicción.

La dependencia parcial entre la radiación difusa y el DI de predicción se muestra en la Fig. 4.26. En esta gráfica se puede observar que existe una proporción inversa entre el predictor y la respuesta de predicción. Lo anterior no es consistente debido a que al aumento de radiación dentro de la base de datos está relacionada con el aumento de la temperatura exterior. Se puede decir que el comportamiento de este predictor dentro del modelo está asociado al error que existe en este.

La Fig. 4.27 ilustra la relación entre las zonas climáticas Köppen y el DI de predicción. Se

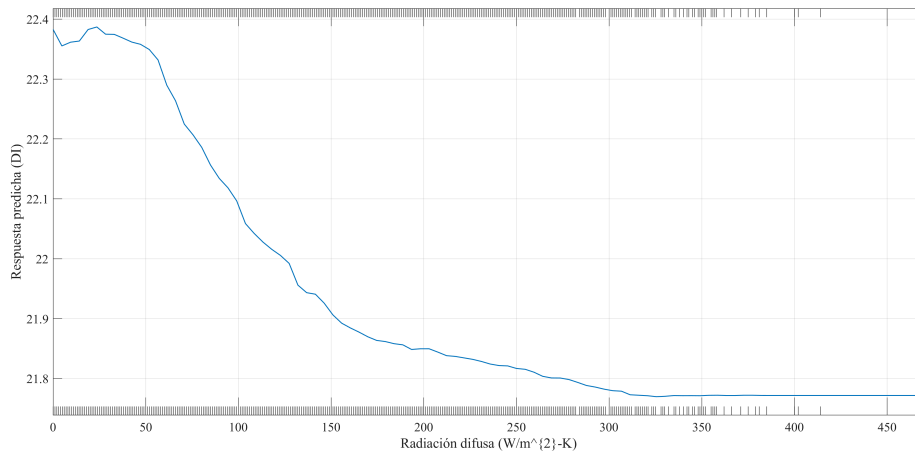


Figura 4.26: Dependencia parcial entre la radiación difusa y la respuesta de predicción.

observa que las primeras tres zonas, que corresponden a regiones climáticas tropicales, están asociadas con un valor alto de DI. En contraste, las siguientes dos zonas, que representan regiones secas, muestran un valor de DI ligeramente más bajo. Por último, las tres zonas restantes, que pertenecen a regiones templadas, presentan los valores más bajos de DI. Este comportamiento es coherente con la información contenida en la base de datos.

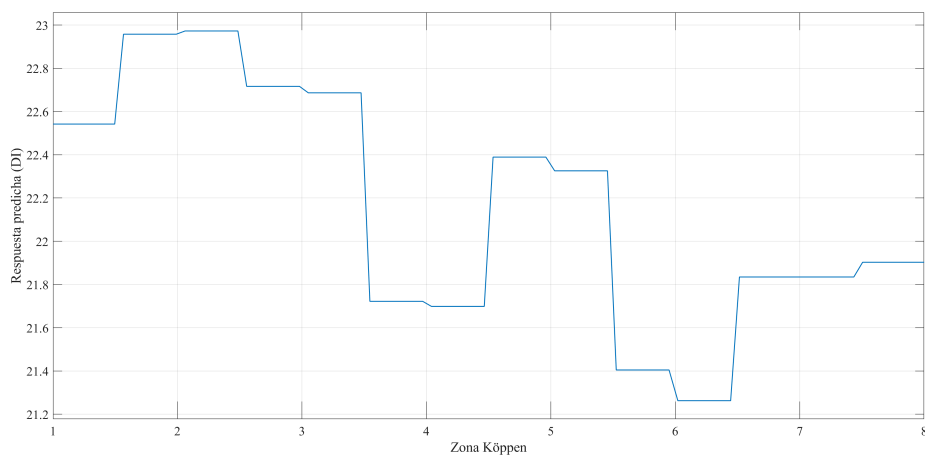


Figura 4.27: Dependencia parcial entre las zonas climáticas Köppen y la respuesta de predicción.

Capítulo 5

Análisis y discusión de resultados

El índice de malestar (DI) es la variable de salida de los modelos de predicción. Esto permite desarrollar una interfaz gráfica capaz de calcular el TDC de un PCM con pocas entradas del usuario, eliminando la necesidad de realizar simulaciones físicas. Con esto se puede desarrollar un análisis de los resultados obtenidos con la predicción del DI.

5.1 Desarrollo de la interfaz gráfica

La interfaz gráfica es un derivado directo del desarrollo de modelos para la predicción del DI. Una vez entrenados los modelos, se procedió a realizar las predicciones de los valores del DI utilizando la metodología ilustrada en la Fig. 5.1. El objetivo del programa es permitir que el usuario acceda a una predicción del DI para el interior de una vivienda mexicana promedio en un día específico. Para ello, el usuario sólo necesita introducir la zona climática, la temperatura del PCM y la estación del año, eliminando la necesidad de realizar simulaciones térmicas en EnergyPlus.

Para este propósito, se generaron números aleatorios de las variables climáticas relevantes, como el punto de rocío, los grados día de calefacción y refrigeración, la presión atmosférica, la humedad relativa y la radiación difusa.

El objetivo del programa explicado en la Fig. 5.1 es que un usuario pueda acceder a una predicción del DI para el interior en un día de una vivienda mexicana promedio, introduciendo solamente la zona climática, la temperatura del PCM y la estación del año, sin tener que realizar las simulaciones térmicas en EnergyPlus. Para esto se generaron números aleatorios, dentro de un rango determinado por el promedio y la variación estándar en cada punto, de las variables climáticas, que son el punto de rocío, los grados día de calefacción y refrigeración, la presión atmosférica, humedad relativa y radiación difusa.

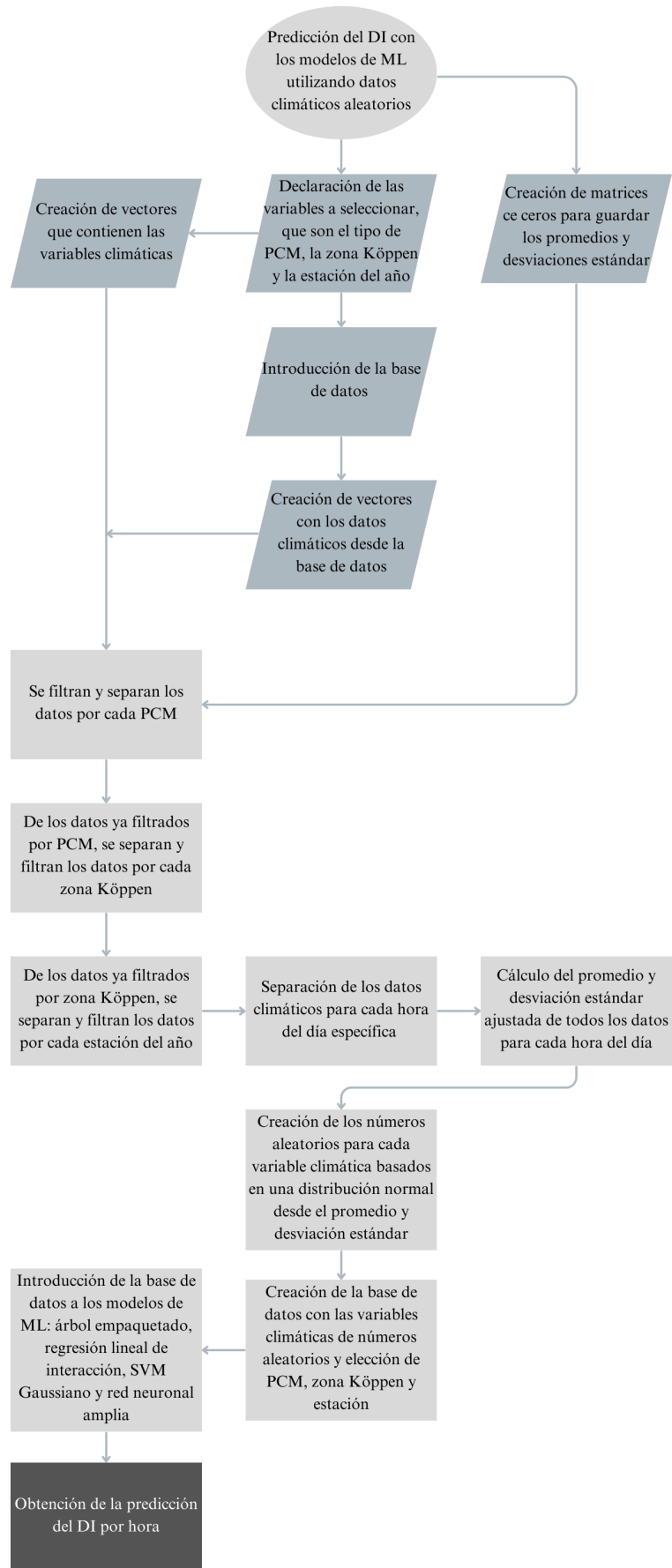


Figura 5.1: Diagrama de flujo utilizado para calcular el TDC.

5.1.1 Pre-tratamiento de los datos de entrada

Dada la complejidad de buscar y utilizar archivos de bases de datos climáticas, así como de manejar una gran cantidad de datos ambientales, la interfaz gráfica incorpora un pre-tratamiento de datos. Este proceso genera datos aleatorios que representan el promedio diario dentro del periodo de simulaciones, asegurando que sigan una distribución normal y se mantengan dentro de un rango definido. El objetivo de generar estos datos aleatorios es permitir al usuario seleccionar una zona climática y una estación del año específica para predecir el comportamiento del DI con el uso de un PCM, sin necesidad de acceder directamente a una base de datos climática.

Para generar estos números aleatorios se introdujeron los datos desde la base de datos para calcular el promedio y desviación estándar para un solo día y generar los datos aleatorios con una distribución normal. Luego de esto, se genera una tabla con los datos de entrada para introducirla a los modelos ya entrenados, y así se obtiene la predicción del DI para un día promedio con las condiciones introducidas.

Para garantizar la consistencia de los datos aleatorios, se aseguró que la tendencia de cada parámetro fuese coherente con el transcurso de las horas del día. Por ejemplo, que la temperatura aumente conforme avanza la madrugada hacia el mediodía y descienda después del atardecer.

En las Figs. 5.2 y 5.3 se muestra una comparación entre los datos promedio, los datos aleatorios y los datos de un día tomado directamente de la base de datos en la ciudad de Tuxtla Gutiérrez (Af) en verano para los parámetros climáticos de CDD y presión atmosférica, respectivamente. La curva verde representa los datos de un día seleccionado directamente de la base de datos; la curva roja representa el promedio calculado por cada hora del día desde la base de datos, y la curva azul representa los datos aleatorios generados.

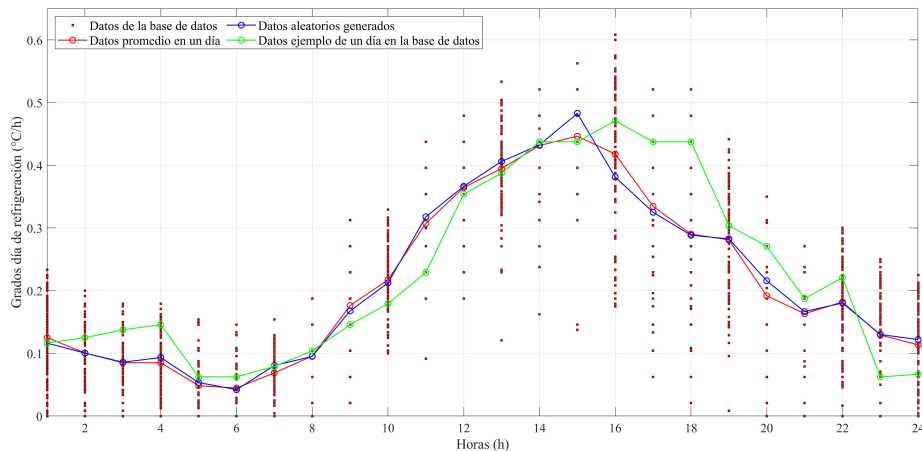


Figura 5.2: Comparación entre las curvas de CDD en un día para Tuxtla Gutiérrez (Af) en verano

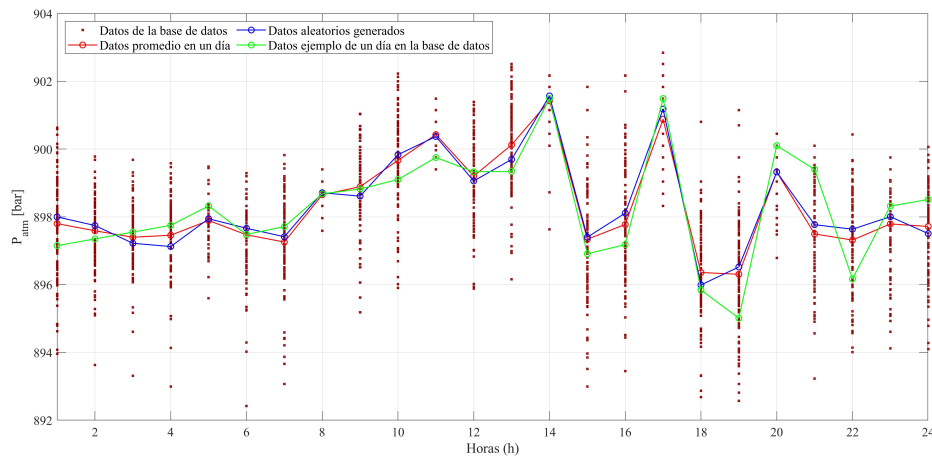


Figura 5.3: Comparación entre las curvas de presión atmosférica en un día para Tuxtla Gutiérrez (Af) en verano

Se puede observar en ambas figuras que la variación en los datos es pequeña y que los datos aleatorios se apegan a los datos promedio, por lo que se puede determinar la confiabilidad en el método utilizado para que así el usuario pueda generar los datos de DI sin necesidad de obtener los datos climáticos de una base de datos y solamente obtener los datos de un día aleatorio y no para la estación del año completa.

5.1.2 Estructura de la interfaz gráfica

Para desarrollar la interfaz gráfica se adaptó el programa a la aplicación de *AppDesigner* en MATLAB, añadiendo las funciones propias de esta como generar botones y texto. La interfaz se puede observar en la Fig. 5.4, donde se tienen las instrucciones para utilizar la aplicación, los menús desplegables donde se pueden seleccionar el tipo de PCM, la estación del año y la zona climática, el botón para generar los datos del DI, la zona donde se genera la gráfica de estos datos y el botón donde se tiene la opción de exportar los datos de predicción calculados.

Es importante mencionar que los datos predichos generados por la interfaz gráfica son ligeramente distintos siempre que se selecciona la opción de Calcular DI, debido a que cada vez genera diferentes parámetros climáticos aleatorios y la predicción del DI para un día promedio con estos parámetros en específico.

5.2 Presentación de resultados específicos

La Tabla 13 muestra los errores relativos de los valores de DI promedio simulados con respecto a los de predicción obtenidos de la interfaz grafica, tomando como referencia los simulados. En esta Tabla se puede observar que las únicas ciudades con un error relativo mayor a 5 % son Monterrey (BS) en ambas estaciones y Hermosillo (BW) en invierno.

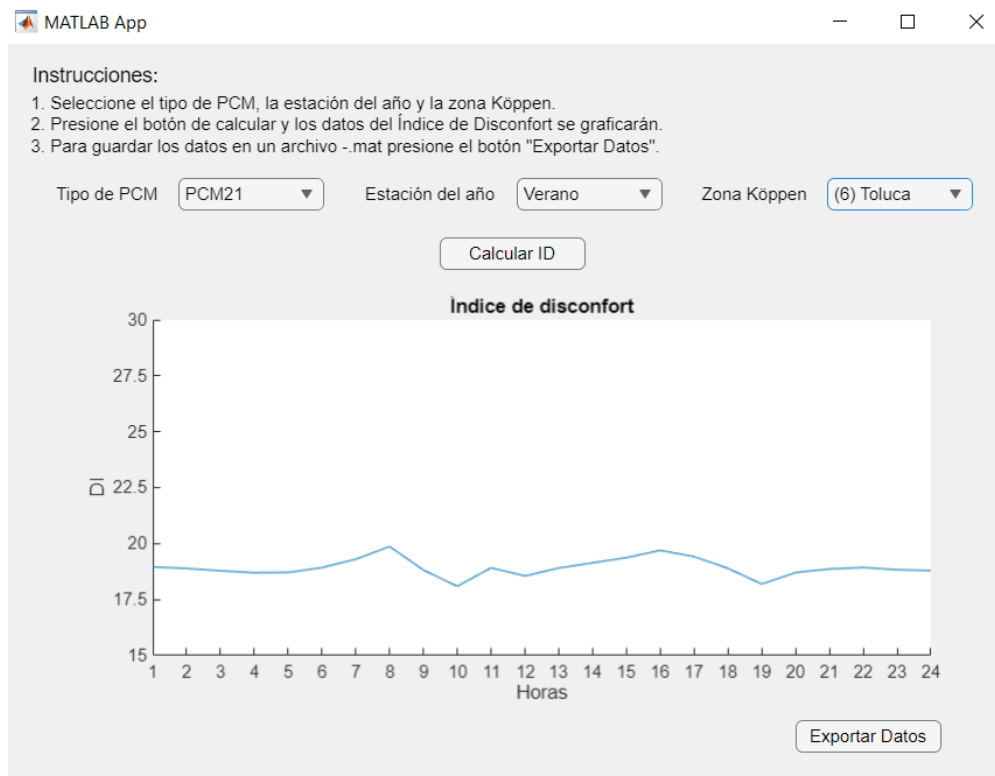
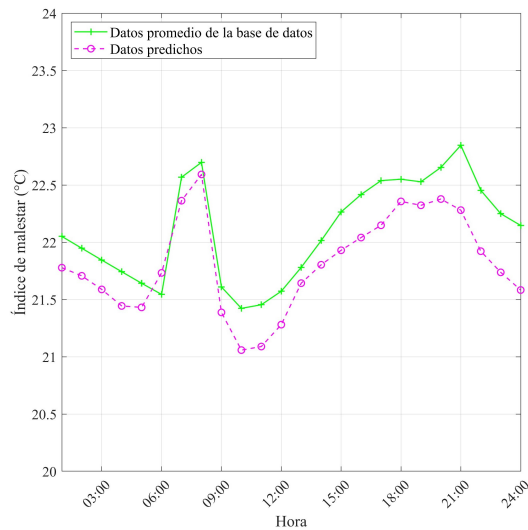


Figura 5.4: Interfaz gráfica.

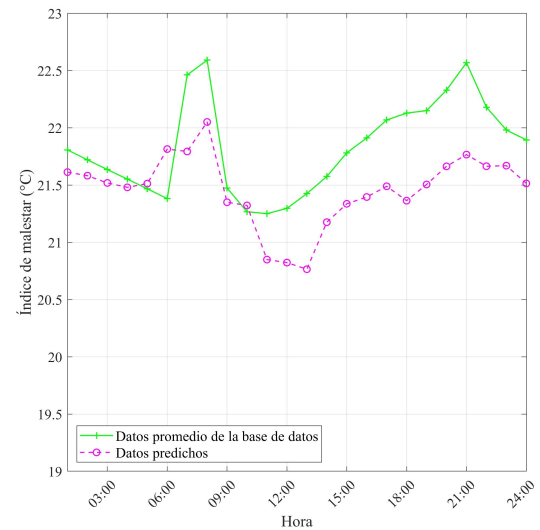
Tabla 13: Comparación de los valores de TDC por ciudad utilizando los datos de predicción.

Ciudad	Estación	Promedio de error relativo (%)
Tuxtla Gutiérrez (Af)	Verano	2.004
Villahermosa (Am)		1.4439
Mérida (Aw)		3.4017
Monterrey (BS)		5.9349
Hermosillo (BW)		2.6436
Xalapa (Cf)		1.378
Guadalajara (Cw)		3.5652
Toluca (Cs)		1.8121
Tuxtla Gutiérrez (Af)	Invierno	1.8868
Villahermosa (Am)		2.9364
Mérida (Aw)		2.4745
Monterrey (BS)		5.4695
Hermosillo (BW)		7.146
Xalapa (Cf)		2.107
Guadalajara (Cw)		4.8003
Toluca (Cs)		3.1119

Con la utilización de la interfaz gráfica, como se menciona en la sección anterior, se obtienen las gráficas de DI para un día promedio dentro de la zona climática y estación del año indicada para los tres tipos de PCM, derivando en una predicción de cómo se comporta el confort térmico para un día promedio bajo las condiciones climáticas ya determinadas.



(a) Comparación del DI de predicción y base de datos para Xalapa (Cf) en verano.



(b) Comparación del DI de predicción y base de datos para Tuxtla Gutiérrez (Af) en invierno.

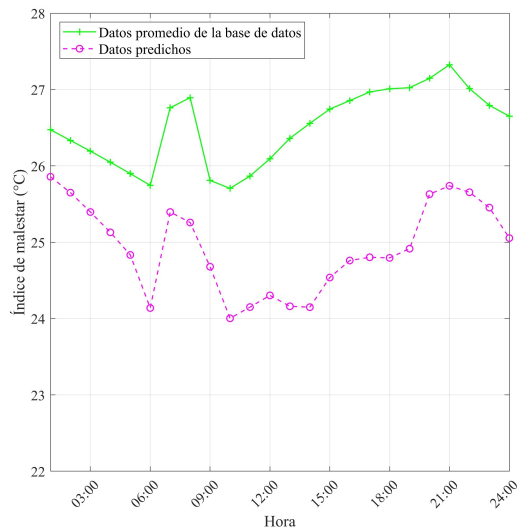
Figura 5.5: Comparación del DI de predicción y base de datos para las ciudades con menor porcentaje de error relativo por estación.

En la Fig. 5.5 se presentan algunas de las gráficas generadas a partir de la interfaz gráfica para las ciudades de Xalapa (Cf) en verano y Tuxtla Gutiérrez (Af) en invierno. Estas ciudades, en sus respectivas estaciones del año, registraron el menor porcentaje de error relativo. Las gráficas muestran una comparación entre el promedio de los valores del DI, obtenidos a partir de los datos en la base de datos, y los valores predichos por el modelo.

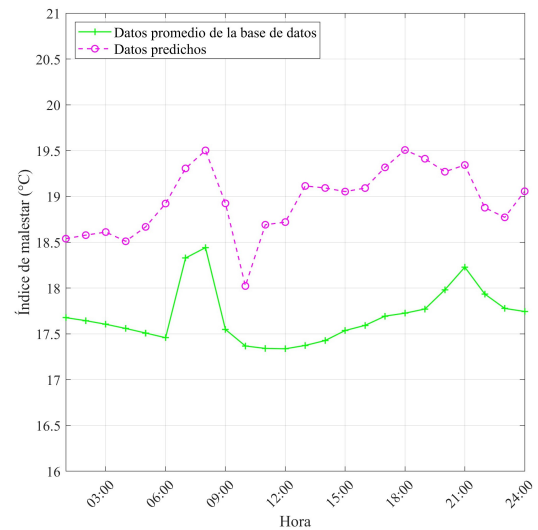
Por otro lado, en la Fig. 5.6 se incluyen gráficas análogas para las ciudades de Monterrey (BS) en verano y Hermosillo (BW) en invierno, las cuales presentaron el mayor porcentaje de error relativo.

En la Fig. 5.5, que es la comparación de los valores del DI para las ciudades con menor error relativo en su respectiva estación, se puede observar que las curvas son muy similares y tienen la misma tendencia en cuanto a los puntos máximos y mínimos de temperatura durante el día. Por el contrario, la Fig. 5.6 muestra la comparación del DI en las ciudades con mayor error relativo en su respectiva estación, donde se puede observar que las curvas tienen un desfase debido al error asociado en los valores de predicción, pero la forma y la tendencia de las curvas son muy similares entre sí.

A pesar del error que contienen los valores de predicción del DI, las gráficas anteriores demuestran que el modelo de Machine Learning es capaz de realizar una predicción lógica y de comprender el comportamiento del DI dependiendo de las condiciones climáticas, hora del día, estación del año, temperatura del PCM y zona climática, que son las variables de entrada del entrenamiento.



(a) Comparación del DI de predicción y base de datos para Monterrey (BS) en verano.



(b) Comparación del DI de predicción y base de datos para Hermosillo (BW) en invierno.

Figura 5.6: Comparación del DI de predicción y base de datos para las ciudades con mayor porcentaje de error relativo por estación.

Las Figs. 5.5 y 5.6 no deben interpretarse directamente como una representación del error entre el valor correcto y el valor predicho. Es importante recordar que las líneas correspondientes a los datos predichos se generan a partir de valores que varían cada vez que el usuario presiona el botón *Calcular DI*. Por lo tanto, en una ejecución podrían aproximarse considerablemente a los datos promedio, mientras que en otra podrían divergir de ellos. Los valores promedio no reflejan la variación presente entre los datos individuales. Por ejemplo, los datos de temperatura para un mismo caso podrían oscilar entre 15 y 20 °C. Si el promedio es 17.5 °C y el valor aleatorio generado es 20 °C, esto no indica que el cálculo sea incorrecto. Más bien, significa que nos encontramos en un caso extremo, aunque posible, dentro del rango considerado.

Con base en lo anterior, estas gráficas deben considerarse únicamente como una herramienta de verificación visual, destinada a comprobar que el comportamiento de los resultados obtenidos coincide con las simulaciones realizadas en *EnergyPlus*. Además, permiten confirmar que los resultados son lógicos y se encuentran dentro de los valores esperados.

5.3 Comparación entre los datos simulados y predichos

Para poder analizar y comparar el desempeño del PCM en la predicción con respecto a los cálculos realizados directamente con los datos obtenidos con las simulaciones, se calculó el valor de TDC para el PCM 21 en cada una de las ciudades representativas a las zonas Köppen y para ambas estaciones del año. Para esto se utilizaron los datos de DI promedio para 24 horas, obtenidos directamente de la interfaz gráfica, para calcular el TDC.

En la Tabla 14 se muestra una comparación entre el valor de TDC predicho y simulado para cada ciudad representativa en ambas estaciones.

Tabla 14: Comparación de los valores de TDC predichos y simulados por ciudad y estación.

Ciudad	Estación	TDC predicho	TDC simulado
Tuxtla Gutiérrez (Af)	Verano	-15	-15
Villahermosa (Am)		13	15
Mérida (Aw)		-3	-14
Monterrey (BS)		41	15
Hermosillo (BW)		7	-3
Xalapa (Cf)		7	1
Guadalajara (Cw)		18	3
Toluca (Cs)		13	11
Tuxtla Gutiérrez (Af)	Invierno	13	6
Villahermosa (Am)		-3	5
Mérida (Aw)		-9	-2
Monterrey (BS)		15	5
Hermosillo (BW)		27	4
Xalapa (Cf)		6	9
Guadalajara (Cw)		14	-5
Toluca (Cs)		6	0

En la Fig. 5.7 se presenta una gráfica de barras, similar a la presentada en el Capítulo 3, donde se muestran los valores de TDC por ciudad y estación, pero esta vez únicamente para el periodo de un día promedio de las simulaciones realizadas.

La Fig. 5.7 muestra que los valores de TDC predichos son similares a los simulados en la Fig. 3.6, aunque con algunas diferencias. En verano, las ciudades de Tuxtla Gutiérrez (Af) y Mérida registraron valores negativos de TDC, mientras que Hermosillo (BW) y Xalapa (Cf) tuvieron valores positivos, en contraste con la Fig. 3.6. En invierno, Villahermosa (Am) presentó un valor negativo de TDC, a diferencia de lo observado en la misma figura.

Estas diferencias se deben a que la base de datos utilizada en este análisis se redujo a 24 datos, representando un solo día promedio en lugar del periodo completo de simulación del Capítulo 3, como se menciona al final de la sección 5.2. A pesar de esto, los valores de TDC en Hermosillo y Xalapa durante el verano, así como en Villahermosa en invierno, son pequeños en comparación con los de otras ciudades, como se aprecia también en la Fig. 3.6. Por ello, las variaciones se atribuyen a la reducción en el número de datos.

A pesar de lo anterior, se tienen las siguientes similitudes entre las Figs. 3.6 y 5.7:

- La ciudad de Tuxtla Gutiérrez (Af) tiene el valor más negativo de TDC durante el verano, lo cual confirma el impacto negativo que tiene el PCM en esta zona y estación del año debido

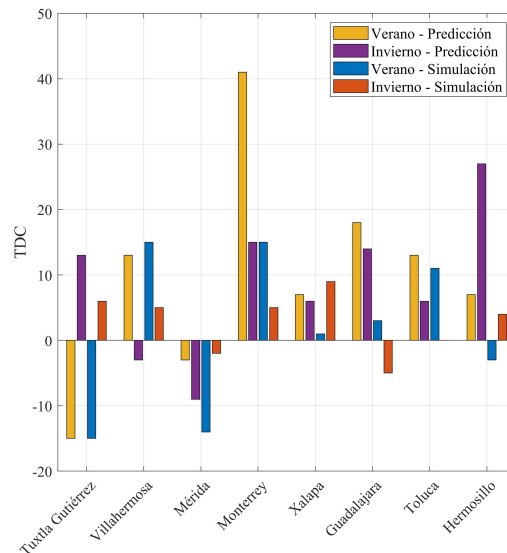


Figura 5.7: TDC del PCM 21 en cada ciudad y estación, comparando los valores de predicción y simulación.

al calor liberado del PCM durante el día.

- La ciudad de Monterrey (BS) tiene el valor más positivo de TDC durante el verano, lo cual confirma el impacto positivo que tiene el PCM en esta zona y estación del año.
- Las ciudades de Toluca (Cs), Guadalajara (Cw) y Monterrey (BS) tienen un TDC positivo para ambas estaciones, confirmando la utilidad del PCM en la mejora del confort térmico en estas zonas climáticas.
- La ciudad de Mérida (Aw) tiene un TDC negativo en ambas estaciones, lo cuál confirma el impacto negativo en el confort térmico que tiene el uso de PCM en esta zona climática.

A diferencia de la Fig. 3.6, en la Fig. 5.7 se observa que los valores positivos de TDC son más numerosos, y el impacto negativo del PCM en las ciudades con valores negativos de TDC es más pequeño. Por lo tanto, se puede decir que para el promedio de días dentro del periodo de simulación y con una confiabilidad aceptable al observar la Tabla 14 y el R^2 de 98 % en el modelo, se predice que el uso de PCM tiene un impacto positivo en la mejora del confort térmico para un día promedio de verano e invierno en las ciudades de Monterrey (BS), Xalapa (Cf), Guadalajara (Cw), Toluca (Cs) y Hermosillo (BW).

Conclusiones

En el presente trabajo de tesis se desarrollaron modelos capaces de predecir el confort térmico dentro de una vivienda mexicana representativa, utilizando algoritmos de Machine Learning. Estos modelos fueron entrenados con una base de datos generada a partir de simulaciones térmicas realizadas en *EnergyPlus*. Además, se estableció una metodología clara para evaluar la mejora del confort térmico en edificaciones que incorporan PCM, utilizando el concepto de cambio total de malestar. A partir de este trabajo, se pueden extraer las siguientes conclusiones:

- Se realizaron la validación y las pruebas de funcionamiento para el módulo de PCM en *EnergyPlus*, determinando su capacidad para realizar simulaciones térmicas con edificios que involucran PCM, debido a su función para simular el efecto de amortiguamiento que poseen los PCMs.
- El uso de PCM disminuye el malestar térmico con mayor medida en invierno que en verano, debido a las altas temperaturas en verano que mantienen el PCM en estado líquido. Cuando las temperaturas en el día son más bajas, la liberación de calor desde el PCM impacta negativamente al confort térmico en verano. Este es el caso para las ciudades de Tuxtla Gutiérrez (Af) y Hermosillo (BW).
- Es recomendable elegir un PCM con un punto de fusión que permita su activación durante el día, debido a que, como es el caso de Mérida (Aw) en verano, la temperatura exterior se mantiene constantemente por encima del punto de fusión del PCM, lo que hace que este se comporte solamente como otra capa de aislante.
- El refinamiento propuesto de los rangos de DI provocó que el sistema fuera más sensible a capturar los cambios en el malestar térmico y en visualizar mejor el desempeño del PCM en el cambio del confort térmico dentro del modelo de vivienda.
- Se desarrollaron cuatro modelos de Machine Learning: árbol de regresión empaquetado, regresión lineal con interacción, máquina de soporte vectorial con núcleo Gaussiano y red neuronal amplia. Los modelos fueron validados utilizando Error Cuadrático Medio (MSE) y el coeficiente de determinación (R^2), teniendo un mejor desempeño en predicción del DI cuando se eliminaron variables de entrada y se cambió la temperatura exterior por los grados día de calefacción y refrigeración. Estos cambios permitieron tener un modelo con $R^2= 98\%$ y error en el valor de predicción del DI menor a 0.5, siendo este un ensamble de árboles empaquetados.

- Se determinó la dependencia parcial de cada uno de los parámetros de entrada con la predicción del DI, lo cual brinda confiabilidad en el modelo de predicción.
- Se creó una interfaz gráfica que permite calcular el índice de malestar (DI) para un día promedio durante el periodo de simulación, considerando las condiciones específicas de la zona climática, el tipo de PCM y la estación del año. La utilización de una interfaz gráfica permite que un usuario externo pueda acceder a la predicción del confort térmico en una vivienda que incluye PCM sin necesidad de acceder a software de simulación térmica.

Trabajo a futuro

Para trabajos futuros, se propone comparar diferentes puntos de fusión de los materiales de cambio de fase (PCM) e investigar la evaluación del confort térmico cuando se combina el uso de PCM con ventilación natural. Se recomienda analizar la combinación de ventilación natural y el uso de PCM en regiones donde, bajo las condiciones climáticas de México, el uso exclusivo de PCM no resultó favorable para el confort térmico dentro de la vivienda.

Además, para el desarrollo de un modelo predictivo entrenado con datos de simulaciones térmicas, se evaluar más a fondo la comparación entre tipos de algoritmos, la optimización de las variables dentro de la base de datos, y el análisis de las capacidades predictivas de los modelos. Se pueden utilizar técnicas computacionales fuera del alcance de este trabajo.

Finalmente, la interfaz gráfica es un área de mejora importante, debido a que se puede hacer más amigable al usuario, realizar comparaciones para los casos con y sin PCM, calcular el TDC de forma directa, optimizar el pre-tratamiento de los parámetros climáticos. Una aplicación de este tipo puede ser de gran utilidad para predecir las capacidades de mejora en el confort térmico al incluir PCM en la construcción de viviendas.

Anexo A: Datos requeridos para replicar las simulaciones con EnergyPlus

Para poder replicar el modelo de forma exacta se requiere de lo siguiente:

Archivos climáticos utilizados.

Los archivos climáticos que fueron utilizados se obtuvieron de la base de datos Climate.OneBuilding, con los nombres siguientes:

- MEX_CHP_Tuxtla.Gutierrez.768430_TMYx.2007-2021
- MEX_JAL_Guadalajara-Chapalita.766120_TMYx.2007-2021
- MEX_MEX_Toluca.de.Lerdo.766750_TMYx.2007-2021
- MEX_SON_Hermosillo.761600_TMYx.2007-2021
- MEX_TAB_Villahermosa.767430_TMYx.2007-2021
- MEX_TAB_Villahermosa.767430_TMYx.2007-2021
- MEX_YUC_Merida-Rejon.Intl.AP.766440_TMYx.2007-2021
- MEX_NLE_Monterrey.763930_TMYx.2007-2021

Los datos de entrada se pueden observar con ayuda del software de libre acceso *Dview*, como se observa en la Fig. 6.1.

Especificaciones y medidas del diseño de la geometría.

Como se especifica en la sección 2.3.2, el modelo de vivienda consiste en un edificio rectangular de 48 m^2 , cuyas dimensiones son 4.8 m por 10 m y 2.5 m de altura. El modelo de vivienda contiene 4 habitaciones. El frente de la casa está en la pared oeste, y contiene una puerta de 2 m por 1 m y una ventana de dimensiones 1 m por 1.2 m (Fig. 2.8). Asimismo, en la pared norte se incluyeron una puerta y una ventana de las mismas dimensiones, mientras que en las paredes este y sur se incluyó una ventana de las mismas dimensiones, añadiendo una ventana de 0.6 m por 0.5 m en la pared sur (ver Fig. 2.9).

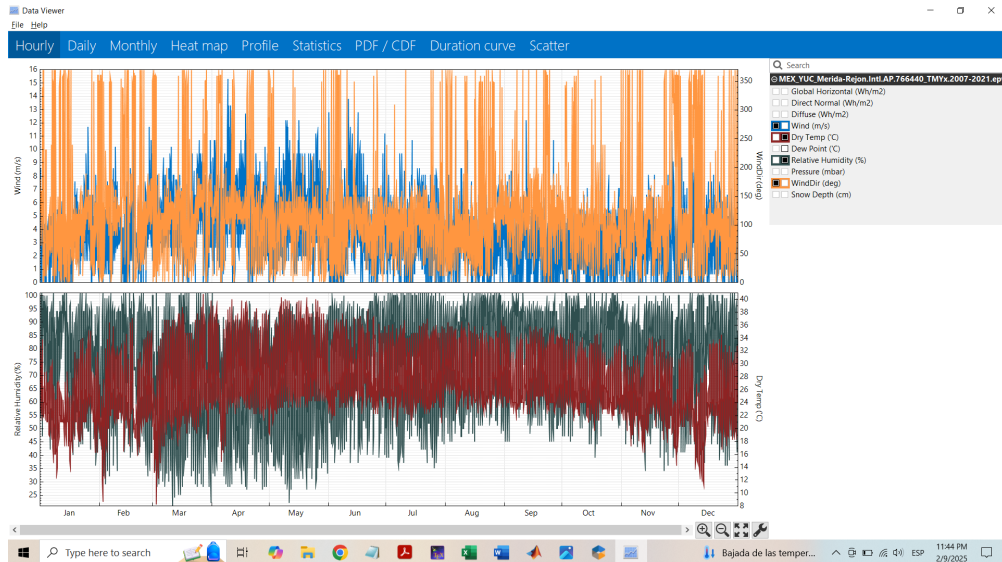


Figura 6.1: Observación de los datos de entrada en un archivo climático con Dview.

La pared frontal se encuentra sobre el eje y, siendo esta la orientación de la casa.

Datos termo-físicos de los materiales.

Los datos termofísicos de cada uno de los materiales utilizados se encuentra en la Tabla 2.

Capas de material en techo, suelo y paredes.

La configuración de las capas de material se pueden observar en la Fig. 2.9. Estas se muestran desde la capa interna hacia la externa. En los casos donde no existe PCM, únicamente se omite esta capa.

Horarios y valores de las cargas.

Los horarios y valores de cada carga añadida se muestran en la Tabla 2.2.

Curvas entalpía-temperatura de los PCMs.

Las curvas de entalpía-temperatura de los PCMs utilizados se aprecian en la Fig. 2.10. Estas curvas se introducen por medio de cada uno de sus puntos tabulados, como se muestran en las Figs. 6.2 y 6.3.

Configuración de los algoritmos para realizar los cálculos.

La configuración de los algoritmos dentro de EnergyPlus para realizar los cálculos es:

- Pasos de tiempo por hora: 20

Field	Units	Obj1	Obj2	Obj3
Name		PCM-21	PCM-26	PCM-31
Temperature Coefficient for Thermal Conductivity	W/m-K2			
Temperature 1	C	13	19	23
Enthalpy 1	J/kg	4263.57	12790.7	2713.18
Temperature 2	C	14	20	24
Enthalpy 2	J/kg	9689.92	20542.64	7364.34
Temperature 3	C	15	21	25
Enthalpy 3	J/kg	16666.67	29844.96	11240.31
Temperature 4	C	16	22	26
Enthalpy 4	J/kg	24418.6	38372.09	16666.67
Temperature 5	C	17	23	27
Enthalpy 5	J/kg	33720.93	50000	26744.18
Temperature 6	C	18	24	28
Enthalpy 6	J/kg	44573.64	69379.84	36821.7
Temperature 7	C	19	25	29
Enthalpy 7	J/kg	60077.52	105038.76	56976.74
Temperature 8	C	20	26	30
Enthalpy 8	J/kg	78682.17	177131.78	84108.53
Temperature 9	C	21	27	31
Enthalpy 9	J/kg	105813.95	193410.85	111240.31
N20		22	28	32
N21		135271.32	198062	138372.09
N22		23	29	33
N23		146124.03	201162.79	155426.36
N24		24	30	34
N25		150775.19	201937.98	166279.07

Figura 6.2: Datos tabulados de entalpía-temperatura en EnergyPlus (parte 1).

N26		25	31	35
N27		152325.58	203488.37	171705.43
N28		26	32	36
N29		155426.36	205813.95	174031
N30		27	33	37
N31		159302.33	208139.53	177131.78
N32		28	34	38
N33		162403.1	209689.92	180232.56

Figura 6.3: Datos tabulados de entalpía-temperatura en EnergyPlus (parte 2).

- Algoritmo de convección: TARP (Thermal Analysis Research Program)
- Algoritmo de balance de calor: ConductionFiniteDifference

Anexo B: Datos climáticos de entrada

En las Tablas 15 y 16 se observan los parámetros climáticos introducidos en la simulación para poder obtener y replicar los resultados obtenidos en la Fig. 2.13.

Tabla 15: Datos climáticos de entrada utilizados en el proceso de validación para la configuración A.

Fecha y hora	T_{ext} (°C)	H_{rel} (%)	v_{viento} (m/s)	R_{dif} (Wh/m ²)	R_{glob} (Wh/m ²)	R_{norm} (Wh/m ²)	P_{rocio} (°C)	P_{atm} (mbar)	dir_{viento} (°)
13/08 21:00	25	83	4	0	0	0	22	966.27	280
13/08 22:00	25.7	81	2	0	0	0	22.2	959.27	320
13/08 23:00	25	83	2	0	0	0	22	966.27	240
14/08 00:00	25	83	3	0	0	0	22	966.27	270
14/08 01:00	24.8	85	3	0	0	0	22.1	959.42	320
14/08 02:00	23	94	2	0	0	0	22	965.96	280
14/08 03:00	24	89	3	0	0	0	22	966.12	240
14/08 04:00	24.5	85	3	0	0	0	21.8	959.38	290
14/08 05:00	23	94	3	0	0	0	22	965.96	240
14/08 06:00	23	94	3	6	6	0	22	965.96	210
14/08 07:00	24.7	85	1	17	17	0	22	960.74	250
14/08 08:00	24	89	4	44	44	0	22	966.12	250
14/08 09:00	24	83	3	61	61	0	21	966.12	270
14/08 10:00	25.8	79	1	93	93	0	21.9	961.67	180
14/08 11:00	25	83	2	166	176	11	22	966.27	270
14/08 12:00	25	89	2	211	238	28	23	966.27	130
14/08 13:00	25.5	82	1	185	208	25	22.2	960.48	270
14/08 14:00	24	94	2	193	251	72	23	966.12	170
14/08 15:00	25	89	2	202	261	77	23	966.27	213
14/08 16:00	28.6	70	1	129	153	48	22.6	959.23	90
14/08 17:00	26	79	2	63	84	59	22	966.42	110
14/08 18:00	26	79	1	21	27	23	22	966.42	129
14/08 19:00	26.8	77	1	0	3	0	22.4	958.96	270
14/08 20:00	25	89	1	0	0	0	23	966.27	60
14/08 21:00	25	89	2	0	0	0	23	966.27	20

En las Tablas 17 y 18 se muestran los resultados de temperatura interior, tanto experimentales como los simulados en *EnergyPlus*.

Tabla 16: Datos climáticos de entrada utilizados en el proceso de validación para la configuración B.

Fecha y hora	T_{ext} (°C)	H_{rel} (%)	v_{viento} (m/s)	R_{dif} (Wh/m ²)	R_{glob} (Wh/m ²)	R_{norm} (Wh/m ²)	P_{rocio} (°C)	P_{atm} (mbar)	dir_{viento} (°)
23/08 10:00	25.2	75	1	271	439	198	20.5	965.02	250
23/08 11:00	24	78	3	303	515	230	20	966.12	250
23/08 12:00	25	55	3	311	512	213	15.3	966.27	250
23/08 13:00	26.3	54	2	239	339	107	16.2	963.85	290
23/08 14:00	25	56	4	165	178	15	15.6	966.27	310
23/08 15:00	24	74	5	130	138	11	19	966.12	310
23/08 16:00	24.1	62	4	114	132	39	16.4	963.03	320
23/08 17:00	23	78	6	86	105	64	19	965.96	300
23/08 18:00	22	83	4	27	34	34	19	965.81	290
23/08 19:00	22.9	83	3	2	2	0	19.9	963.71	290
23/08 20:00	22	83	5	0	0	0	19	965.81	280
23/08 21:00	21	83	4	0	0	0	18	965.65	270
23/08 22:00	22.1	82	3	0	0	0	18.9	965.39	290
23/08 23:00	21	83	4	0	0	0	18	965.65	270
24/08 00:00	21	88	3	0	0	0	19	965.65	280
24/08 01:00	21.5	85	3	0	0	0	18.9	964.73	290
24/08 02:00	21	83	5	0	0	0	18	965.65	300
24/08 03:00	21	83	2	0	0	0	18	965.65	290
24/08 04:00	21.5	81	2	0	0	0	18.1	964.44	290
24/08 05:00	20	83	3	0	0	0	17	965.49	270
24/08 06:00	20	83	4	20	22	7	17	965.49	270
24/08 07:00	22.4	77	2	73	98	61	18.2	965.63	180
24/08 08:00	21	78	3	154	221	125	17	965.65	260
24/08 09:00	22	78	3	236	360	173	18	965.81	270
24/08 10:00	23.1	74	2	284	422	167	18.2	966.22	290
24/08 11:00	23	69	3	279	366	96	17	965.96	280
24/08 12:00	24	65	4	291	358	70	17	966.12	270
24/08 13:00	23.9	71	2	240	292	56	18.3	964.91	320
24/08 14:00	23	69	4	156	184	33	17	965.96	290
24/08 15:00	23	69	4	141	157	21	17	965.96	300
24/08 16:00	24.3	68	1	103	106	6	18	963.83	340
24/08 17:00	23	69	4	58	60	6	17	965.96	320
24/08 18:00	23	69	2	21	22	4	17	965.96	290
24/08 19:00	23.1	73	1	1	1	0	18	964.02	290
24/08 20:00	22	73	2	0	0	0	17	965.81	290
24/08 21:00	22	73	2	0	0	0	17	965.81	290
24/08 22:00	22.2	79	2	0	0	0	18.4	965.03	290
24/08 23:00	21	83	2	0	0	0	18	965.65	290
25/08 00:00	21	83	2	0	0	0	18	965.65	290

Tabla 17: Comparación de los resultados simulados y experimentales para la configuración A.

Fecha y hora	T_{int} simulada ($^{\circ}C$)	T_{int} experimental ($^{\circ}C$)
13/08 21:00	36.42	30.61
13/08 22:00	35.35	31.12
13/08 23:00	34.17	30.66
14/08 00:00	32.84	29.88
14/08 01:00	31.69	29.65
14/08 02:00	30.79	28.91
14/08 03:00	29.98	28.55
14/08 04:00	29.33	27.95
14/08 05:00	28.68	27.63
14/08 06:00	28.1	27.26
14/08 07:00	28.11	26.62
14/08 08:00	29.55	27.87
14/08 09:00	31.58	30.46
14/08 10:00	33.63	32.59
14/08 11:00	35.43	35.23
14/08 12:00	36.99	37.41
14/08 13:00	37.92	38.66
14/08 14:00	38.51	39.36
14/08 15:00	39.19	40.1
14/08 16:00	39.63	41.21
14/08 17:00	39.89	41.63
14/08 18:00	39.92	41.17
14/08 19:00	39.48	39.23
14/08 20:00	38.73	36.88
14/08 21:00	37.65	33.41

Tabla 18: Comparación de los resultados simulados y experimentales para la configuración B.

Fecha y hora	T_{int} simulada ($^{\circ}C$)	T_{int} experimental ($^{\circ}C$)
23/08 10:00	33.63	32.18
23/08 11:00	33.68	32.61
23/08 12:00	33.67	33.27
23/08 13:00	33.6	33.63
23/08 14:00	33.62	34.07
23/08 15:00	33.08	34.43
23/08 16:00	32.82	34.42
23/08 17:00	32.9	33.9
23/08 18:00	32.72	33.23
23/08 19:00	32.6	32.63
23/08 20:00	32.42	32.03
23/08 21:00	32.25	31.43
23/08 22:00	32.08	30.91
23/08 23:00	31.96	30.23
24/08 00:00	31.77	29.56
24/08 01:00	31.54	29.18
24/08 02:00	31.25	28.96
24/08 03:00	30.69	28.65
24/08 04:00	29.73	28.35
24/08 05:00	28.79	27.97
24/08 06:00	28.07	27.74
24/08 07:00	27.46	27.66
24/08 08:00	27.25	27.95
24/08 09:00	28.24	28.61
24/08 10:00	29.3	29.19
24/08 11:00	30.15	30.15
24/08 12:00	30.83	31.25
24/08 13:00	31.31	32.2
24/08 14:00	31	33.01
24/08 15:00	30.22	33.22
24/08 16:00	29.66	33.43
24/08 17:00	29.01	33.35
24/08 18:00	28.25	31.57
24/08 19:00	27.84	30.09
24/08 20:00	27.25	29.49
24/08 21:00	26.86	29.11
24/08 22:00	26.79	28.51
24/08 23:00	26.6	28.51
25/08 00:00	26.08	28.13

Bibliografía

- [1] Aulicems, A., Szokolay, S. (2007). *Thermal Comfort*. PLEA: Passive and Low Energy Architecture International. Brisbane, 2nd edition.
- [2] Secretaría de Energía. (2022). *Balance Nacional de Energía 2021*. Gobierno de México. Subsecretaría de Planeación y Transición Energética.
- [3] Revista Mundo HVAC & R. (2023). *Estimaciones de crecimiento para el mercado HVAC en el 2023*.
- [4] Epstein, Y., Moran, D. (2006). *Thermal comfort and the heat stress indices*. Industrial Health, vol. 44.
- [5] ANSI/ASHRAE Standard 55. (2004). *Thermal Environmental Conditions for Human Occupancy*.
- [6] Cengel, Y. A. (2002). *Heat transfer. A practical approach*. 2nd edition. McGraw-Hill: New York.
- [7] Babiarz, B. and Szymanski, W. (2020). *Introduction to the dynamics of heat transfer in buildings*. Energies, vol. 13. doi:10.3390/en13236469
- [8] Jäckel, R. (2015). *Modeling and simulation of thermal comfort conditions in a building using a passive acclimation system based on incorporated phase change materials*. Integration project. Universidad Autónoma Metropolitana.
- [9] ASHRAE Addendum d to ASHRAE 55. (2017). *Thermal Environmental Conditions for Human Occupancy*.
- [10] *Procesos y carta psicrométrica*. (2023). <https://enfrifactor.com/psicometria-procesos/>
- [11] E. C. Thom. (1959). *The Discomfort Index*. Weatherwise vol. 12, 2 doi: 10.1080/00431672.1959.9926960
- [12] Bizee. *Degree Days - An Introduction*. Wheater data for energy saving. <https://www.degreedays.net/introduction>
- [13] U.S. Energy Information Administration. *Degree Days*. <https://www.eia.gov/energyexplained/units-and-calculators/degree-days.php>

- [14] Day, A. R., Maidment, G., Ratcliffe, M. (1999). *Cooling degree-days and their applicability to building energy estimation*. School of Engineering Systems and Design, South Bank University.
- [15] Mistry, M. (2018). *Cooling Degree days and Heating degree days*. <https://www.energy-a.eu/cooling-degree-days-and-heating-degree-days/>
- [16] Day, A. R. & Karayiannis, T. (1998). *Degree-days: Comparison of calculation methods*. Building Services Engineering Research and Technology, vol. 19. doi: 10.1177/014362449801900102
- [17] Hale, D., Hoover, M., O'Neill, M. (1971). *Phase change materials handbook*. NASA - George C. Marshall Space Flight Center. Alabama.
- [18] Hasenöhrl, T. (2009) *An introduction to phase change materials as heat storage mediums*. Project report. Dept. of Energy Science. Lund, Sweden.
- [19] Memon, S. A. (2014). *Phase change materials integrated in building walls: A state of the art review*. Renewable and Sustainable Energy Reviews, vol. 31. <http://dx.doi.org/10.1016/j.rser.2013.12.042>
- [20] Pasupathy, A., Velraj, R., Seeniraj, R. V. (2006). *Phase change material-based building architecture for thermal management in residential and commercial establishments*. Renewable and Sustainable Energy Reviews vol. 12.
- [21] Zhang, Y., Rao, Z., Wang, S., Zhang, Z., Li, X. (2012). *Experimental evaluation on natural convection heat transfer of microencapsulated phase change materials slurry in a rectangular heat storage tank*. Energy Conversion and Management, vol. 59.
- [22] *Python for Machine Learning: The Complete Beginner's Course*. <https://www.udemy.com/course/python-for-machine-learning-beginners>
- [23] Müller, A., Guido, S. (2017). *Introduction to Machine Learning with Python. A guide for data scientists*. O'Riley: Primera edición. California, U.S.
- [24] MATLAB. (2024). *Statistics and Machine Learning Toolbox. User's Guide*.
- [25] Chen, Y., Ye, Y., Liu, J., Zhang, L., Li, W., Mohtaram, S. (2023). *Machine learning approach to predict building thermal load considering feature variable dimensions: An office building case study*. Buildings 2023, 13, 312, <https://doi.org/10.3390/buildings13020312>
- [26] Khan, M., Khan, F. (2012). *A comparative study of White Box, Black Box and Grey Box Testing Techniques*. International Journal of Advanced Computer Science and Applications. Vol. 3, N°6, pp. 22-25
- [27] Zhang, L., Wen, J. (2021). *Active learning strategy for high fidelity short-term data-driven building energy forecasting*. Energy & Buildings 244. <https://doi.org/10.1016/j.enbuild.2021.111026>

- [28] Statistics Knowledge Portal. *Multiple Linear Regression with Interactions*. https://www.jmp.com/en_fi/statistics-knowledge-portal/what-is-multiple-regression/mlr-with-interactions.html
- [29] Steorts, R. *Tree Based Methods: Bagging, Boosting, and Regression Trees*. Duke University.
- [30] Grigoryan, A. (2023). *Exploring Wide and Deep Networks*. Medium. <https://thegrigorian.medium.com/exploring-wide-and-deep-networks-b6af7f0a5d3a>
- [31] Zhussupbekov, M., Memon, S. A., Khawaja, S., Nazir, K., Kim, J. (2023). *Forecasting energy demand of PCM integrated residential buildings: A machine learning approach*. Journal of Building Engineering, vol. 70. <https://doi.org/10.1016/j.jobbe.2023.106335>
- [32] Seyedzadeh, S., Rahimian, F., Rastogi, P., Glesk, I. (2019). *Turning machine learning models for prediction of building energy loads*. Sustainable Cities and Society, vol. 47. <https://doi.org/10.1016/j.scs.2019.101484>
- [33] Kumar, S., Pal, S., Singh, R. (2018). *A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes*. Energy & Buildings. <https://doi.org/10.1016/j.enbuild.2018.06.056>
- [34] Adilkhanova, I., Memon, S., Kim, J., Sheriyev, A. (2020). *A novel approach to investigate the thermal confort of the lightweight relocatable building integrated with PCM in different climates of Kazakhstan during summertime*. Energy, <https://doi.org/10.1016/j.energy.2020.119390>
- [35] U.S. Department of Energy. (2023). *Engineering Reference*. EnergyPlus Version 23.1.0 Documentation.
- [36] Tabares-Velasco, P., Christensen, C., Bianchi, M., Booten, C. (2012). *Verification and validation of EnergyPlus Conduction Finite Differences and Phase Change Material models for opaque wall assemblies*. National Renewable Energy Laboratory.
- [37] Rubitherm (2024). Recuperado de: <https://www.rubitherm.eu/en/productcategory/organische-pcm-rt>
- [38] INFONAVIT (2024). *La casa 3. Manual de vivienda progresiva*. Recuperado de: <https://decide-construye-autodiagnostico.ruv.org.mx/assets/docs/Casa3-Planos.pdf>
- [39] INEGI. (2020). *Encuesta nacional de vivienda (ENVI), 2020. Principales resultados*. Recuperado de: <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021>
- [40] Solano, N. (2022). *Materiales en la edificación*. Material oplementario para las asignaturas Sistemas Ambientales I y II. Universidad Nacional Autónoma de México.
- [41] Peel, M., Finlayson, B., McMahon, T. (2007). *Updated world map of the Köppen-Geiger climate classification*. Hydrology and Earth System Sciences 11. Copernicus Publications.

- [42] *Clasificación y distribución de los climas en México*. Recuperado de www.cursosinea.conevyt.org.mx/cursos/cnaturales_v2/interface/main/recursos/antologia/cnant_4_13.htm
- [43] Lawrie, L. K. & Drury B. C. (2022). *Repository of free climate data for building performance simulation*. <https://climate.onebuilding.org/>
- [44] *National Centers of Environmental Information*. (2024). Integrated Surface Dataset. Global Hourly; ISD.
- [45] Li, B., Zhuang, C., Deng, A., Li, S., Shen X. (2009). *Improvement of indoor thermal environment in light weight building combining phase change material wall and night ventilation*. Journal of Civil, Architectural and Environmental Engineering, 31(3), 109-113.
- [46] Zhuang, C., Deng, A., Chen Y., Li, S., Zhang, H., Fan, G. (2010). *Validation of Veracity on Simulating the Indoor Temperature in PCM Light Weight Building by EnergyPlus*. https://doi.org/10.1007/978-3-642-15621-2_53
- [47] Siami, L., Ramadhani, A. (2019). *Climatology of Discomfort Index for Decade in Bandar Lampung, Indonesia*. KnE Social Sciences.
- [48] Sheriyev, A., Memon, S. A., Adilkhanova, I., Kim, J. (2021). *Effect of Phase Change Materials on the Thermal Performance of Residential Building Located in Different Cities of a Tropical Rainforest Climate Zone*. Energies, 14. <https://doi.org/10.3390/en14092699>
- [49] Chugh, A. (2020). *MAE, MSE, RMSE, Coefficient of Determination, Adjusted R Squared — Which Metric is Better?* Medium.
- [50] *MRMR - Minimum Redundancy Maximum Relevance*. https://feature-engine.trainindata.com/en/1.8.x/user_guide/selection/MRMR.html
- [51] Hastie, T., Tibshirani, R., Friedman, J. (2008). *The Elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer: 2nd edition. California, U.S.
- [52] Cengel, Y., Boles, M., Kanoglu, M. (2019). *Thermodynamics: An Engineering Approach*. McGraw Hill, 9th edition: NYC, USA.