

Predicción del consumo energético en espacios interiores de una edificación universitaria

Presentado por

Liliana Patricia Ortega Diaz

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones (E³T)

Maestría en Ingeniería Eléctrica

Bucaramanga

2025

Predicción del consumo energético en espacios interiores de una edificación universitaria

Presentado por

Liliana Patricia Ortega Diaz

Trabajo de investigación para optar al título de Magíster en Ingeniería Eléctrica

Director

German Alfonso Osma Pinto

Doctor en Ingeniería, Área Ingeniería Eléctrica

Co-Director

Jorge Luis Cardenas Rangel

PhD (c) en Ingeniería, Área Ingeniería Eléctrica

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones (E³T)

Maestría en Ingeniería Eléctrica

Bucaramanga

2025

Dedicatoria

A Dios y a la Virgen, por ser mi refugio en los momentos difíciles, por darme fortaleza cuando más la necesité y por iluminar mi camino. A mis padres, por su apoyo incondicional y por enseñarme el valor del esfuerzo y la constancia. A mis profesores y mentores, cuya guía ha sido fundamental en mi formación. En especial a mi director German Alfonso Osma Pinto por los aprendizajes y apoyo en las dificultades durante la investigación, a Alejandra Martínez Peñaloza por su acompañamiento y disposición para ayudarme cuando lo necesité y a mi codirector Jorge Luis Cárdenas Rangel por su compromiso y apoyo en este recorrido de formación. A mis compañeros del grupo GISEL por los momentos compartidos. A todos aquellos que, de alguna manera, han sido parte de este proceso, muchas gracias de corazón.

Agradecimientos

A Dios y a la Virgen por no desamparme, por guiarme e iluminarme siempre. A mi familia por estar siempre para mí sobretodo en los momentos difíciles, sin ellos no hubiese sido posible este gran logro. A mi alma mater, la UIS, por brindarme los espacios más bonitos y acogedores donde pude compartir momentos bonitos y aprender de las mejores personas.

A mi director German Osma, por su valiosa guía, compromiso y serenidad a lo largo de este camino. Sus consejos, dedicación y conocimiento fueron primordiales para el progreso de este trabajo. Aprecio grandemente el esfuerzo y tiempo que ha dedicado para orientarme, corregir mis errores y motivarme a mejorar en cada fase del proyecto. Su ayuda no solo ha enriquecido este proyecto, sino que también ha sido una fuente de crecimiento personal y aprendizaje. Gracias por su confianza en mi trabajo y colaboración, por sus sugerencias acertadas y por alentarme a dar siempre lo mejor de mí.

A la MSc. Alejandra Martínez, compañera y amiga, por su apoyo incondicional. Sus recomendaciones, tanto académicas como personales, fueron una guía invaluable que me ayudaron a superar obstáculos durante el proceso. Gracias por compartir conmigo su conocimiento, por su paciencia y por motivarme en los momentos difíciles. Su compañía y amistad hicieron que este proceso fuera más llevadero y enriquecedor. A el MSc. Jorge Cárdenas, por su orientación, continuo respaldo y dedicación a lo largo de esta investigación.

A mis compañeros del GISEL por los valiosos momentos compartidos, su entusiasmo y vocación han sido una fuente de inspiración en esta etapa de mi formación.

Contenidos

Introducción	14
1. Construcción de experimento y monitorización de variables	24
1.1. Estación meteorológica	24
1.1.1. Software y descarga de datos	26
1.2. Sensores Monnit	27
1.3. Medidores de energía	29
1.4. Cámaras de vídeo	33
2. Creación de base de datos	35
2.1. Procesamiento de los datos	36
2.2. Códigos empleados para la creación de la base de datos	36
2.2.1. Función consumo de energía	36
2.2.2. Función estación meteorológica	37
2.2.3. Función sensores Monnit 1	37
2.2.4. Función sensores Monnit 2	37
2.2.5. Función de las cámaras	38
2.2.6. Función de la tabla final	38
2.3. Normalización de datos	40
3. Entrenamiento de modelos	41
3.1. Regresión Lineal Múltiple (<i>Multiple Lineal Regression - MLR</i>)	41
3.2. Media Móvil Autorregresiva Integrada (<i>Autoregressive Integrated Moving Average - ARIMA</i>)	41
3.3. Red Neuronal Artificial Profunda (<i>Deep Neural Network - DNN</i>)	43

PREDICCIÓN DEL CONSUMO DE ENERGÍA	6
3.4. DesignBuilder (DB)	45
3.5. Modelo híbrido: DesignBuilder + DNN	46
4. Evaluación de modelos	48
5. Resultados de la investigación	49
5.0.1. Descripción de datos	49
5.0.2. Discusión	51
6. Conclusiones	58
Bibliografía	60
Anexo A	66
Anexo B	69
B.0.1. Cálculo consumo de energía	69
B.0.2. Función estación meteorológica	71
B.0.3. Función sensores Monnit 1	73
B.0.4. Función sensores Monnit 2	76
B.0.5. Función cámaras	77
B.0.6. Función final (concatena todas las variables)	77

Tablas

Tabla 1. Estudios relacionados: variables monitoreadas, equipos y frecuencia de muestreo.	17
Tabla 2. Ficha descriptiva del caso de estudio.	21
Tabla 3. Sensores integrados en la estación meteorológica.	25
Tabla 4. Variables estación meteorológica	25
Tabla 5. Sensores Monnit utilizados	28
Tabla 6. Cantidad de sensores ubicados por aula	28
Tabla 7. Circuitos medidos por el ACUREV 2020	32
Tabla 8. Circuitos medidos por el ACUREV 2100	33
Tabla 9. Valores de las variables categóricas y su significado.	39
Tabla 10. Índices de entrenamiento y prueba de cada división.	43
Tabla 11. Resumen del entrenamiento de redes neuronales	44
Tabla 12. Base de datos Design Builder	47
Tabla 13. Media y desviación estándar de los errores	53
Tabla 14. Análisis comparativo de los procesos y recursos utilizados en la formación de los modelos	57

Figuras

Figura 1. Edificio de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones (E3T).	20
Figura 2. Metodología general para la previsión del consumo energético en las aulas.	22
Figura 3. Estación meteorológica.	24
Figura 4. Vista del software WeatherLink.	26
Figura 5. Boletín e informe del clima (10/12/2024).	26
Figura 6. Esquemático de sensores IoT.	27
Figura 7. Vista de aplicativo iMonnit.	27
Figura 8. Ubicación de sensores IE404.	29
Figura 9. Ubicación de sensores IE405.	29
Figura 10. Ubicación y número de los sensores en las dos aulas.	30
Figura 11. Medidor de energía de aires acondicionados (Acurev2020).	30
Figura 12. Medidor de energía de circuitos de aulas de clase (Acurev2100).	31
Figura 13. Diagrama unifilar y conexión del medidor.	31
Figura 14. Conexión y configuración medidor de energía ACUREV2100.	32
Figura 15. Sistema de aire acondicionado en aulas.	33
Figura 16. Vista IE404 desde la aplicación EZVIZ.	34
Figura 17. Esquemático creación de la base de datos.	35
Figura 18. Base de datos final.	40
Figura 19. Estructura de la DNN.	44
Figura 20. Modelo de DesignBuilder.	45
Figura 21. Promedio del consumo de energía del IE404.	49
Figura 22. Promedio del consumo de energía del IE405.	50
Figura 23. Consumo de energía por día de la semana.	50

Figura 24. Mapa de correlaciones de variables. Superior a la diagonal IE405, inferior a la diagonal IE404.	51
Figura 25. Resultados del software DesignBuilder.	54
Figura 26. Errores del modelo híbrido.	55
Figura 27. Tablero de aires acondicionados.	66
Figura 28. Diagrama unifilar tablero AA.	67
Figura 29. Subtableros del cuarto piso del Edificio de Ingeniería Eléctrica.	67
Figura 4. Diagrama unifilar TGBT piso 4. (A) Circuitos de iluminación y tomacorrientes regulados. (B) Circuitos de tomacorrientes normales.	68

Anexos

Anexo A. Diagramas unifilares y fotos de los tableros monitorizados	66
Anexo B. Funciones de Python para la creación de la base de datos	69

Abreviaciones

EX	Extractores de aire
AA	Aire Acondicionado
IE404	Ingeniería Eléctrica 404
IE405	Ingeniería Eléctrica 405
T	Temperatura
H	Humedad
N	Número
X_{tiempo}	VARIABLES temporales
X_{ext}	VARIABLES exteriores
X_{int}	VARIABLES interiores
RLM	Regresión Lineal Múltiple (<i>Multiple Linear Regression</i>)
ARIMA	Modelo autorregresivo integrado de media móvil (<i>Autoregressive Integrated Moving Average</i>)
DNN	Red neuronal artificial profunda (<i>Deep Neural Network</i>)
LSTM	Red neuronal de memoria de corto y largo plazo (<i>Long Short-Term Memory</i>)
R^2	Coefficiente de determinación
MAE	Error absoluto medio (<i>Mean Absolute Error</i>)
RMSE	Raíz cuadrada del error cuadrático medio (<i>Root Mean Squared Error</i>)

Abstract

Title: Prediction of energy consumption in indoor spaces of a university building *

Author: Liliana Patricia Ortega Diaz **

Keywords: Energy efficiency, Energy consumption, Buildings, Comfort, Prediction.

Description: Buildings are one of the sectors with the greatest impact on energy consumption and, consequently, a significant source of greenhouse gas emissions. This highlights the importance of accurately predicting their energy use to improve efficiency and reduce their environmental impact. Energy consumption prediction enables efficient management through systems such as BEMS (Building Energy Management System), BAS (Building Automation System), and digital twins. Four main approaches are employed for this purpose: statistical, simulation-based, data-driven, and hybrid. This study predicts energy consumption in classrooms of an educational building located in a warm tropical climate. A representative model of each approach is trained to predict energy consumption in two classrooms. Continuous monitoring of climatic, operational, and energy consumption variables is carried out in these spaces, integrating the collected data into a database. The data is then processed and prepared for model training. Three evaluation scenarios are proposed. In the first, models based on historical data are compared, including RLM, ARIMA, DNN, and LSTM. The second scenario compares real data with those obtained through simulation in Design Builder. Finally, the third scenario compares simulated data with a hybrid model that incorporates a DNN to predict the weekly energy consumption of each classroom. The performance evaluation of the models is conducted using the metrics R^2 , RMSE, and MAE.

* MSc Thesis

** Faculty of Physic-Mechanical Engineering. Department of Electrical, Electronic and Telecommunications Engineering (E³T). Director: German Alfonso Osma Pinto, Ph.D. in Engineering.

Resumen

Título: Predicción del consumo energético en espacios interiores de una edificación universitaria *

Autor: Liliana Patricia Ortega Diaz **

Palabras clave: Eficiencia energética, Consumo de energía, Edificios, Confort, Predicción.

Descripción: Los edificios son uno de los sectores con mayor impacto en el consumo energético y, por ende, una fuente significativa de emisiones de gases de efecto invernadero. Esto subraya la importancia de predecir con precisión su uso de energía para mejorar la eficiencia y reducir su impacto ambiental. La predicción del consumo energético facilita una gestión eficiente a través de sistemas como BEMS (Building Energy Management System), BAS (Building Automation System) y gemelos digitales. Para ello, se emplean cuatro enfoques principales: estadístico, basado en simulación, basado en datos e híbrido. Este estudio realiza la predicción del consumo energético en aulas de un edificio educativo ubicado en un clima tropical cálido. Este estudio entrena un modelo representativo de cada enfoque para predecir el consumo de energía en dos aulas de clase. Se lleva a cabo un monitoreo continuo de variables climáticas, operativas y de consumo energético en estos espacios, integrando los datos recopilados en una base de datos. Posteriormente, los datos son procesados y preparados para el entrenamiento de los modelos. Se plantean tres escenarios de evaluación. En el primero, se comparan modelos basados en datos históricos, incluyendo RLM, ARIMA, DNN y LSTM. El segundo escenario realiza una comparación entre los datos reales y los obtenidos mediante simulación en Design Builder. Finalmente, el tercer escenario compara los datos simulados con un modelo híbrido que incorpora una DNN para predecir el consumo energético semanal de cada aula. La evaluación del desempeño de los modelos se lleva a cabo mediante las métricas R^2 , RMSE y MAE.

* Trabajo de investigación

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director: German Alfonso Osma Pinto, Doctor en Ingeniería, Área Ingeniería Eléctrica.

Introducción

Los edificios, la industria y el transporte son los sectores que más energía consumen en todo el mundo [1]. La construcción y el funcionamiento de los edificios representan el 38 % de las emisiones mundiales de gases de efecto invernadero. De manera particular, los sistemas de calefacción, ventilación y aire acondicionado (HVAC) son los que más contribuyen al consumo energético de los edificios [2]. La energía utilizada en el funcionamiento, calefacción y refrigeración de los edificios es de aproximadamente el 50%, la iluminación es de aproximadamente el 20% , y el funcionamiento de los electrodomésticos y equipos instalados en ellos es de aproximadamente el 20% [3].

Estrategias para mitigar las emisiones de carbono y reducir el consumo de energía en los edificios han sido propuestas, tales como: la integración de recursos energéticos distribuidos, mecanismos de comercio de carbono y la construcción de ciudades inteligentes. El éxito de estas estrategias depende de la capacidad de predecir con exactitud el consumo de energía en tiempo real [4]. Además, los planificadores urbanos y los responsables de las políticas energéticas utilizan modelos y análisis del rendimiento energético de los edificios para desarrollar planes estratégicos de energía sostenible [5].

La gestión eficiente de la energía en los edificios puede disminuir significativamente la demanda, especialmente cuando el consumo se predice con precisión. Para optimizar el uso energético, algunos edificios cuentan con sistemas como BEMS (*Building Energy Management System*), BAS (*Building Automation System*) y gemelos digitales, que permiten monitorear, gestionar y controlar el consumo de manera más efectiva. Los BEMS actúan como el cerebro del edificio, controlando el suministro de energía basándose en datos de sensores y algoritmos [6]. Los BAS son sistemas ciberfísicos diseñados para automatizar procesos, como la gestión de los controles de climatización, el acceso mediante cerraduras electrónicas y el control de la iluminación [7].

Los gemelos digitales mejoran la eficiencia energética y el confort térmico de los edificios

al proporcionar una representación digital en tiempo real del edificio y sus sistemas. Integran tres componentes: espacio físico, espacio virtual y datos conectados. Estos sistemas permiten la supervisión y el control predictivo; por ejemplo, Hosamo *et al.* [8] desarrollaron un marco que combina datos de sensores en tiempo real y aprendizaje automático para optimizar el aire acondicionado, reduciendo el consumo de energía y garantizando al mismo tiempo el confort térmico.

La predicción estima el consumo de energía en un instante específico del futuro. La literatura identifica cuatro enfoques principales para la predicción del consumo energético: métodos estadísticos, de caja blanca, de caja negra y de caja gris. Cada uno presenta puntos fuertes y débiles que motivan el uso de criterios para seleccionar el enfoque adecuado según el caso de estudio, el tiempo de predicción y el tipo de datos disponibles, entre otros [9].

El enfoque estadístico es un subcampo de las matemáticas que intenta encontrar relaciones entre variables para predecir un resultado. En el caso de la predicción del consumo de energía, este enfoque relaciona el consumo de energía con las distintas variables que lo afectan. Se necesitan series temporales con una resolución temporal específica para predecir valores futuros [10].

El enfoque de caja blanca también se conoce como enfoque de ingeniería; estos modelos utilizan los principios físicos de los edificios [11] para simular mediante software especializado el rendimiento energético de los edificios requieren datos geométricos, patrones de ocupación de los usuarios, cargas de equipos, sistemas de climatización y patrones de uso [5]. La construcción de este tipo de modelos requiere un nivel de detalle de datos tan elevado que a menudo resulta más compleja. Esto se debe a que se necesitan mayores niveles de conocimientos e información [12]. Sin embargo, son una buena alternativa para el diseño de edificios y el sector de la construcción.

El enfoque de caja negra o basado en datos está dentro del campo de la informática y el aprendizaje automático; utiliza datos históricos para entrenar modelos o algoritmos de aprendizaje automático y aprende y mejora su precisión. La ventaja de los modelos de aprendizaje automático de caja negra es que requieren un esfuerzo de modelado mínimo y, al mismo tiempo, capturan con precisión las relaciones entre los variables del modelo [12]. Estos modelos se han aplicado en diversos estudios relacionados con los edificios, como la optimización del funcionamiento eficiente,

la estimación de la ocupación, la detección de fallos y la predicción del consumo de energía [5], [6], [13]-[16].

El enfoque híbrido combina los modelos mencionados para mejorar la predicción y superar las desventajas. Las predicciones pueden realizarse combinando simulación y aprendizaje automático o modelos estadísticos y aprendizaje automático. En comparación con el enfoque de ingeniería, los modelos híbridos proporcionan una mayor precisión de predicción y facilitan el cálculo de los valores predichos.

Se ha investigado sobre la predicción sobre áreas clave de la gestión energética en los edificios. Tales estudios se han centrado en la predicción del consumo energético, la optimización del uso de la energía, la predicción de la ocupación, el confort de los ocupantes y el rendimiento energético. La Tabla 1 presenta una breve descripción de los trabajos relacionados con este estudio. Se hace un resumen detallado del conjunto de datos, el caso de estudio y los parámetros de entrada de los modelos. Esta revisión bibliográfica profundiza en el estudio de las investigaciones más relevantes sobre predicción del consumo energético en edificios. Destaca el uso de datos ambientales internos y externos, horarios de ocupación, número de ocupantes, medición del consumo energético y datos físicos y termofísicos.

Ni *et al.* [16] realizaron predicciones hora a hora del consumo energético y la carga de calefacción de dos edificios públicos. Utilizaron regresión lineal múltiple y siete técnicas de aprendizaje profundo. Los modelos TFT (transformada de fusión temporal) y MLR (regresión lineal múltiple) mostraron resultados prometedores. Henzel J *et al.* [25] utilizaron métodos ingenuos, regresión lineal, LSTM (*Long Sort Term Memory*), y el método Prophet para pronosticar el consumo de energía para el día siguiente utilizando datos únicos obtenidos de un modelo gemelo digital de un edificio, su objetivo es proponer un nuevo enfoque de acuerdo con los mejores resultados obtenidos. Mishra *et al.* [19] utilizaron redes neuronales LSTM para predecir el consumo de energía utilizando datos históricos y los compararon con modelos establecidos de regresión lineal, árbol de decisión y bosque aleatorio. En un edificio educativo, estos modelos de redes neuronales artificiales mostraron resultados de alto rendimiento con un R^2 de 0,97 y un MAE de 0,007 Wh.

Tabla 1. Estudios relacionados: variables monitoreadas, equipos y frecuencia de muestreo.

Ref	Año	País	Información del conjunto de datos	Estudio de caso	Características de entrada	Frecuencia de muestreo
[16]	2024	Suecia	4 años de datos (2016 - 2020)	2 edificios públicos: 1 museo y 1 teatro	Temperatura del bulbo seco, humedad relativa, temperatura del punto de rocío, precipitación, presión atmosférica, velocidad del viento, radiación global y consumo de energía	1 hora
[17]	2024	Egipto	Datos físicos, características de la envolvente del edificio, materiales de construcción, conductividad térmica, sistema de iluminación y consumo de energía	Mezquita	-	1 hora
[4]	2024	China	Consumo de electricidad, datos ambientales interiores y datos meteorológicos exteriores	7 edificios: 2 oficinas, 1 laboratorio, 2 bibliotecas y 2 educativos	Consumo eléctrico histórico, duración del uso de enchufes, temperatura interior, concentración de CO ₂ , humedad del aire, concentración de PM2.5, intensidad de radiación solar, temperatura exterior, humedad del aire y velocidad del viento	-
[18]	2024	India	Datos físicos, termo-físicos y de materiales de la envolvente	Habitación residencial	Temperatura, humedad y radiación solar	1 hora
[19]	2024	India	Datos energéticos, ocupación, factores ambientales y calendario académico con días laborables y no laborables	Edificios residenciales y comerciales	Días laborables y no laborables, temperatura y humedad	1 minuto
[14]	2023	Corea del Sur	Datos de sensores del ambiente interior	Edificios universitarios	Información interior y exterior, área del sitio, horarios de ocupación, temperatura exterior, temperatura interior, concentración de PM2.5, concentración de CO ₂ y humedad	-
[15]	2023	-	4 semanas	Aulas en un edificio educativo	Ocupación	-
[20]	2023	Investigación: China Edificio: Reino Unido	2 años	Aula	Mes del año, día de la semana, hora del día, temperatura exterior, temperatura del punto de rocío, velocidad del viento y consumo de energía del edificio	1 hora
[21]	2022	Singapur	Datos estáticos o independientes del tiempo (10,000 muestras)	Edificios residenciales	Datos climáticos, geometría del edificio, dispositivos de sombreado, construcción de muros y parámetros de acristalamiento	-
[13]	2022	EE. UU.	Datos públicos del Datathon 2022 de Women in Data Science (WiDS)	Edificios comerciales y residenciales	75,757 características continuas, discretas y categóricas	-
[22]	2020	Reino Unido	Consumo horario de electricidad y datos ambientales	Edificio educativo	Índice de día laborable (WDI), temperatura del bulbo seco, irradiancia solar, humedad relativa y velocidad del viento	1 hora
[23]	2020	EE. UU.	Monitoreo de energía y clima, y simulación del consumo energético de edificios	Edificios de oficinas y educativos	Datos de consumo energético sub-horario, consumo horario obtenido mediante simulaciones, área del edificio, tipo de sistemas de aire acondicionado y programas de ocupación	1 hora
[24]	2019	Singapur	Monitoreo de parámetros energéticos y ambientales (1,155 muestras)	Edificio educativo	Tipo de día, hora del día, temperatura exterior, radiación solar, estado de operación de bombas, parámetros del sistema de refrigeración y demanda de energía eléctrica	-

M. El-Maraghy *et al.* [17] utilizaron un software de simulación (IESVE) para generar el conjunto de datos energéticos. Se entrenó una CNN (red neuronal convolucional) para predecir el consumo energético anual de los edificios de mezquitas en función de varios escenarios operativos.

Por otra parte, Zhang *et al.* [4] predijeron el consumo energético de la carga enchufable de varios tipos de edificios. Utilizaron un sistema de monitorización del consumo energético y sensores ambientales en las habitaciones de los edificios. Los resultados mostraron que el sistema de predicción propuesto, Bi-LSTM mejorado mediante optimización de enjambre inteligente, funciona mejor que el tradicional. El sistema de predicción tradicional utiliza entradas convencionales y el modelo LSTM.

Zhu G *et al.* [11] realizaron un procedimiento mejorado para predecir la carga de los edificios con 24 horas de antelación, junto con un método de generación de perfiles climáticos de reserva. Utilizaron una RNA con retropropagación y el método de aprendizaje de Levenberg-Marquardt. Llegaron a la conclusión de que un modelo de predicción meteorológica es adecuado para predecir la carga de los edificios. Olu-Ajayi *et al.* [26] investigaron métodos de selección de características centrados en la predicción del consumo energético de los edificios mediante aprendizaje automático. Se entrenaron una DNN y 13 algoritmos más; se descubrió que la DNN funciona bien en muestras de gran tamaño para predecir el consumo energético de los edificios. Elbeltagi *et al.* utilizó un conjunto de datos generados a partir de simulaciones de consumo de energía en EnergyPlus, considerando múltiples opciones de diseño y variables de entrada aleatorias, para entrenar una RNA. El entrenamiento de la RNA mejora la precisión y las capacidades computacionales de los modelos de predicción del uso de la energía en edificios residenciales durante las primeras etapas de diseño.

La revisión de la literatura permitió identificar algunas lagunas en la investigación sobre predicción del consumo de edificaciones. Una de ellas es la poca investigación sobre el uso de la energía en espacios pequeños, como las aulas. Es necesario estudiar el comportamiento de estos lugares y la influencia de las condiciones climáticas, la ocupación y el uso del espacio. Es importante centrarse o profundizar en estudios de casos como estos para saber específicamente cómo

es el uso de la energía en cada zona del edificio. Además, los estudios se centran en condiciones climáticas templadas, por lo que los estudios en climas tropicales son escasos. El clima influye significativamente en la demanda energética, especialmente debido a los cambios en las temperaturas interiores y exteriores y el uso del sistema HVAC [16], [20]. Segundo, la investigación sobre los espacios interiores de los edificios se centra en la sensación térmica de los ocupantes, el número de ocupantes y el confort térmico, no en el consumo de energía. Hoz-Torres *et al.* [27] utilizan las aulas como caso de estudio, aunque estudian el ambiente térmico interior. Predominan los estudios predictivos del consumo energético de todo el edificio [22], [28]. Es difícil determinar qué técnica utilizar, ya que se desconoce la calidad de la exposición.

Dadas las lagunas existentes en la investigación, este estudio pretende identificar el modelo más preciso para predecir el consumo energético en las aulas de un edificio universitario situado en una ciudad de clima tropical cálido.

A continuación, se presenta el objetivo general y objetivos específicos de esta investigación se presentan:

Objetivo general

Desarrollar modelos de predicción del consumo energético en espacios interiores de una edificación educativa en clima tropical.

Objetivos específicos

- Realizar el montaje y puesta en marcha de un sistema de medición de variables climáticas, operativas y de consumo energético para espacios interiores.
- Crear una base de datos a partir de las mediciones realizadas en los espacios interiores seleccionados.
- Entrenar modelos de predicción de consumo de energía desde la aplicación de cinco técnicas

seleccionadas.

- Evaluar el desempeño de los modelos de predicción de consumo energético a partir de cálculo de métricas de rendimiento.

Alcances y limitaciones



El enfoque de la investigación se centra en el caso de estudio del edificio de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones (E3T). Este edificio se distingue por la implementación de medidas de ahorro energético, como una cubierta verde, sistemas de ventilación híbridos y fuentes de generación renovable. Las aulas de clase de este estudio fueron el IE404 e IE405, ubicadas en el cuarto piso de la edificación. Estas aulas tienen un alto flujo de personas de lunes a sábado, ya que en ellas se imparten clases tanto de pregrado como de posgrado. Se realizó la monitorización de variables meteorológicas, consumo de energía y variables internas u operativas de estos dos espacios. La Figura 1 muestra la fachada sur del edificio en mención y la Tabla 2 presenta la descripción general de los casos de estudio.



Figura 1. Edificio de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones (E3T).

Inicialmente, se plantea monitorear el consumo energético de todo el edificio; sin embargo, la infraestructura no cuenta con la cantidad suficiente de sensores y medidores para un seguimiento

Tabla 2. Ficha descriptiva del caso de estudio.

	Aula IE404	Aula IE405
Vista Interior		
Área	61,48 m ²	77,18 m ²
Capacidad	34 personas	
Equipos	2 aires acondicionados Fan Coil Cassette de 36000 BTU/h cada uno	
Iluminación	10 luminarias de 80 W (4 tubos + balasto electrónico regulable de 6 W) 60 x 60 cm 2 lámparas bañadoras de 28 W cada una	12 luminarias de 80 W (4 tubos + balasto electrónico regulable de 6 W) 60 x 60 cm 2 lámparas bañadoras de 28 W cada una
Otros equipos	Proyector 200 W Televisor 215,9 W Computador 240 W 2 extractores de 66 W cada uno	

completo. Por esta razón, se decide analizar el consumo de energía en dos aulas de clase. Las limitaciones del estudio están marcadas por fallos en el sistema de automatización, específicamente por el daño de una controladora central, responsable de gestionar las controladoras de iluminación y climatización, lo que afecta la captura de datos en las aulas. Para mitigar esta limitación, se instalan sensores Monnit dentro de las aulas. Asimismo, se incorporan dos medidores inteligentes para monitorear el consumo de los aires acondicionados y los circuitos eléctricos de las aulas.

Metodología y estructura del documento

Este estudio de enfoque experimental y de modelización se basa en la monitorización de diversas variables en dos aulas de un edificio universitario. La metodología de cuatro etapas incluye la recogida de datos, el preprocesamiento, el entrenamiento de modelos y la evaluación. La Figura 2 presenta un resumen de la metodología empleada en este estudio.

El Capítulo 1 describe el proceso de monitoreo de variables en las aulas de clase, incluyendo la instalación de diversos equipos como medidores de energía, sensores inalámbricos, cámaras

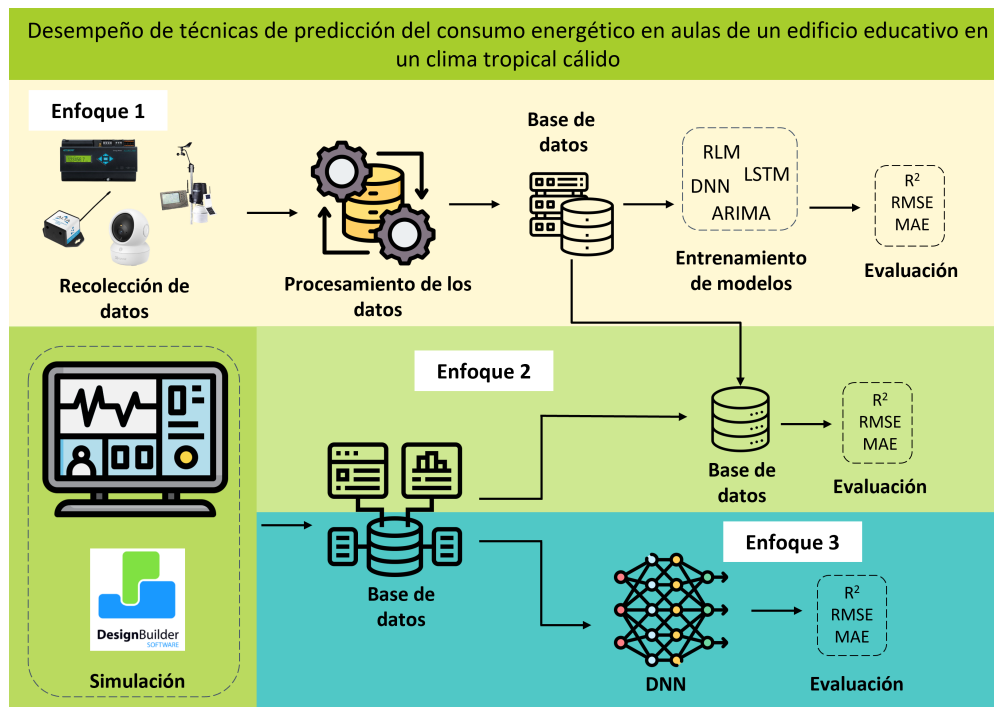


Figura 2. Metodología general para la previsión del consumo energético en las aulas.

de video y una estación meteorológica. Gracias a esta infraestructura, se logró monitorear el consumo de energía, las variables climáticas internas y externas, las variables operativas, el número de ocupantes y computadoras, así como las variables temporales. El Capítulo 2 se explica cómo se construyó la base de datos, detallando el agrupamiento de las variables para conformar el conjunto de datos final. El Capítulo 3 aborda el entrenamiento de los modelos seleccionados para la investigación, incluyendo sus características, parámetros y procesos de modelado. El Capítulo 4 se presentan los resultados de la evaluación de los modelos a través de tres escenarios o casos de estudio, utilizando métricas como el R^2 , el RMSE y el MAE. El Capítulo 5 expone los resultados de la investigación y las características de los datos. Finalmente, el Capítulo 6 recoge las conclusiones de la investigación, junto con recomendaciones y posibles líneas de trabajo futuro.

Divulgación científica y pasantías de investigación

- Publicación de un artículo científico denominado: “Estrategias de predicción de consumo

energético en edificaciones: una revisión” en la revista nacional TecnoLógicas. El doi del artículo es el siguiente: <https://doi.org/10.22430/22565337.2650>

- Publicación de un artículo denominado: “Performance of energy consumption prediction techniques in classrooms of an educational building in a warm tropical climate” en la revista Energy and Buildings (Categoría Q1 de Scimago). El doi del artículo es el siguiente: <https://doi.org/10.1016/j.enbuild.2025.115967>
- Publicación de un artículo denominado: “Estimation of the air conditioning energy consumption of a classroom using machine learning” en la revista Frontiers in Big Data (Categoría Q2 de Scimago). El doi del artículo es el siguiente: <https://doi.org/10.3389/fdata.2025.1520574>
- Pasantía de investigación de cinco meses en la Université de Sherbrooke, Canadá.

1. Construcción de experimento y monitorización de variables

Esta sección describe detalladamente el proceso de monitoreo de las variables en cada aula, incluyendo la instalación de los equipos y la consideración de los parámetros relevantes. Para este propósito, se emplearon dispositivos como medidores de energía, sensores IoT, una estación meteorológica y cámaras de video.

1.1. Estación meteorológica

Se utiliza la estación meteorológica DAVIS VANTAGE PRO2 PLUS, es una estación inalámbrica que envía los datos a una consola situada en el quinto piso del edificio. Consta de un conjunto de sensores encargados de monitorizar variables climáticas, está ubicada en la terraza del Edificio de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones (E3T) (Figura 3). En la Tabla 3 se explica cada uno de los sensores de la estación y la función que desempeñan.



Sistema Integrado de Sensores



Consola

Figura 3. Estación meteorológica.

De las 36 variables monitoreadas por la estación meteorológica, se seleccionaron 11 por diversas razones. En primer lugar, debido a la instalación de sensores de temperatura y humedad

Tabla 3. *Sensores integrados en la estación meteorológica.*

Sensor	Descripción
Veleta	Mide la dirección del viento.
Anemómetro	Mide la velocidad del viento.
Piranómetro	Mide la radiación solar, la evapotranspiración y los índices UV.
Pluviómetro	Mide las precipitaciones de la lluvia.
Panel solar	Energiza la estación durante el día. El super condensador incorporado proporciona energía durante la noche. La batería de litio proporciona respaldo cuando es necesario hasta por 8 meses sin luz solar.

dentro de las aulas, no se consideraron las variables relacionadas con las condiciones internas. Además, se descartaron las variables con valores constantes (como cero), aquellas irrelevantes para el análisis y las que representaban información redundante. Las variables seleccionadas, junto con sus unidades, se presentan en la Tabla 4.

Tabla 4. *Variables estación meteorológica*

Nombre	Unidad
Temperatura exterior	°C
Humedad exterior	%
Punto de rocío	°C
Velocidad del viento	m/s
Dirección del viento	-
Índice de calor	°C
Presión atmosférica	mmHg
Cantidad de lluvia	mm/hr
Radiación solar	W/m ²
Índice UV	-
Grados-día de enfriamiento	°C

1.1.1. Software y descarga de datos

La configuración de la estación meteorológica puede efectuarse directamente desde la consola o mediante el software WeatherLink de Davis Instruments. En la Figura 4 se puede apreciar como se observan los datos desde el programa. Allí se pueden ver gráficos y boletines de las variables monitorizadas, ver la base de datos y descargar los datos en formato de texto (Figura 5).

Date	Time	Temp	Hum	Dew	Wind	Wind	Wind	Hum	Heat	THW	THSW	Rain	Solar	Solar	Solar	UV	UV	Heat	Cool	In	Out		
9/12/24	10:55a	22.2	22.2	22.2	84	19.4	0.0	---	0.00	0.0	---	22.2	23.3	23.3	22.0	753.2	0.00	0.0	0.0	0.000	0.014	25.6	73
9/12/24	11:00a	22.3	22.3	22.3	85	19.6	0.0	---	0.00	0.0	---	22.3	23.3	23.3	22.0	753.1	0.00	0.0	0.0	0.000	0.014	25.7	74
9/12/24	11:05a	22.3	22.3	22.2	85	19.6	0.0	---	0.00	0.0	---	22.3	23.3	23.3	22.0	753.0	0.00	0.0	0.0	0.000	0.014	25.7	66
9/12/24	11:10a	22.2	22.2	22.2	86	19.8	0.0	---	0.00	0.0	---	22.2	23.3	23.3	22.0	753.0	0.00	0.0	0.0	0.000	0.014	25.6	64
9/12/24	11:15a	22.1	22.1	22.1	86	19.7	0.0	---	0.00	0.0	---	22.1	23.2	23.2	21.9	753.0	0.00	0.0	0.0	0.000	0.013	25.4	69
9/12/24	11:20a	21.9	21.9	21.9	87	19.7	0.4	ENE	0.13	1.3	ENE	21.9	23.0	23.0	21.7	753.0	0.00	0.0	0.0	0.000	0.013	25.4	71
9/12/24	11:25a	21.8	21.8	21.8	88	19.8	0.4	ENE	0.13	1.3	ENE	21.8	23.0	23.0	21.7	753.0	0.00	0.0	0.0	0.000	0.012	25.5	73
9/12/24	11:30a	21.8	21.8	21.7	89	19.8	0.4	ENE	0.13	1.3	ENE	21.8	22.9	22.9	21.6	753.1	0.00	0.0	0.0	0.000	0.012	25.7	74
9/12/24	11:35a	21.7	21.7	21.6	89	19.9	0.4	ENE	0.13	1.3	ENE	21.7	22.9	22.9	21.5	753.1	0.00	0.0	0.0	0.000	0.012	25.8	71
9/12/24	11:40a	21.6	21.6	21.6	89	19.7	0.0	---	0.00	0.0	---	21.6	22.7	22.7	21.4	752.9	0.00	0.0	0.0	0.000	0.011	25.7	63
9/12/24	11:45a	21.5	21.5	21.5	89	19.6	0.0	---	0.00	0.0	---	21.5	22.6	22.6	21.3	752.9	0.00	0.0	0.0	0.000	0.011	25.5	66
9/12/24	11:50a	21.4	21.4	21.4	89	19.6	0.0	---	0.00	0.0	---	21.4	22.7	22.7	21.4	752.9	0.00	0.0	0.0	0.000	0.011	25.5	71
9/12/24	11:55a	21.3	21.3	21.3	88	19.6	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	72
9/12/24	12:00a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:05a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:10a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:15a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:20a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:25a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:30a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:35a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:40a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:45a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:50a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	12:55a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:00a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:05a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:10a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:15a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:20a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:25a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:30a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:35a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:40a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:45a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:50a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	1:55a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:00a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:05a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:10a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:15a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:20a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:25a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:30a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:35a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:40a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:45a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:50a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	2:55a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.000	0.012	25.6	73
9/12/24	3:00a	21.3	21.3	21.3	89	19.2	0.0	---	0.00	0.0	---	21.3	22.8	22.8	21.6	752.9	0.00	0.0	0.0	0.			

1.2. Sensores Monnit

El funcionamiento del sistema de sensores de Monnit se describe en la Figura 6. Los sensores envían datos a una puerta de enlace (Gateway) y este a su vez entrega los datos al Portal de Sensores Inalámbricos iMonnit. Desde este portal se pueden configurar alarmas, mensajes de texto, correos que alerten a los usuarios sobre una situación anormal (Figura 7).

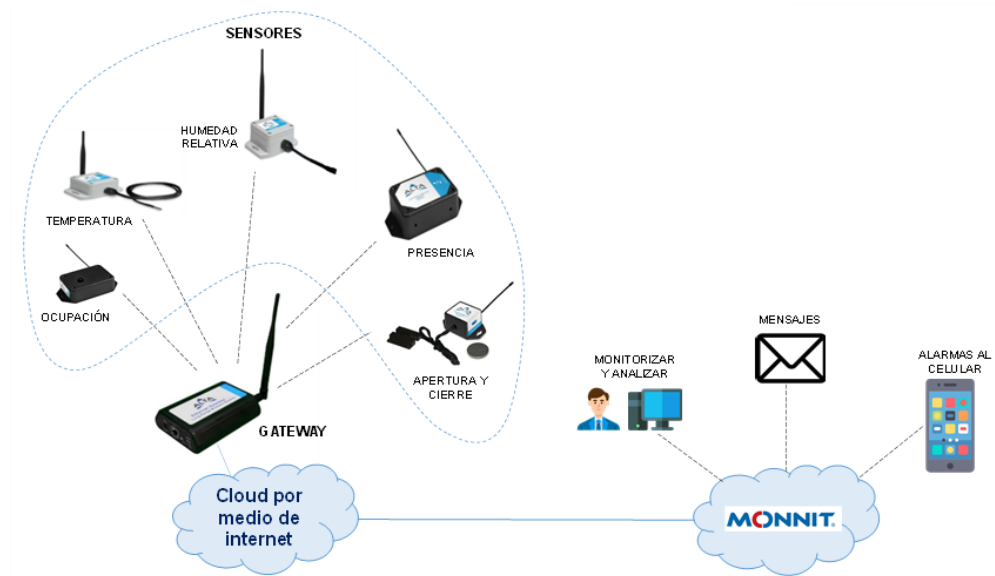


Figura 6. Esquemático de sensores IoT.

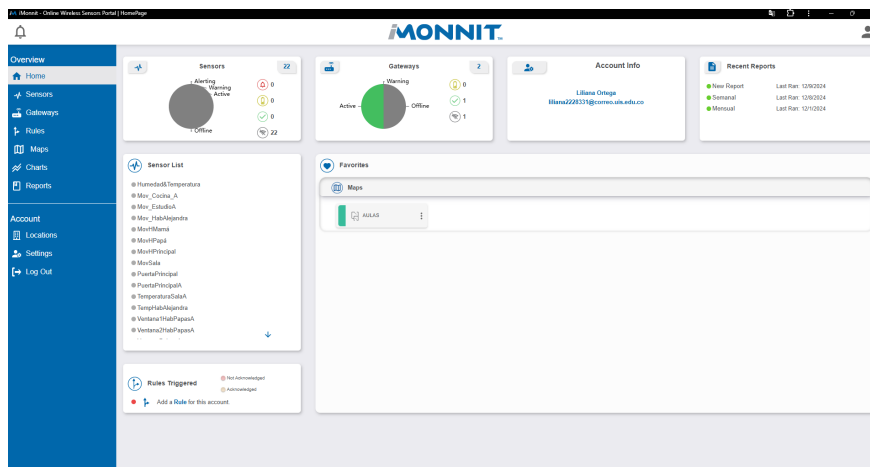


Figura 7. Vista de aplicativo iMonnit.

La Tabla 5 presenta la fotografía del sensor, su tipo y las variables que mide. Los sensores Monnit se clasifican en dos categorías: sensores de intervalos y sensores de eventos. La principal diferencia entre ambos radica en la manera en que realizan el sensado. Los sensores de eventos están programados para recopilar datos únicamente cuando ocurre una situación específica, mientras que los sensores de intervalos registran datos periódicamente, siguiendo un intervalo de tiempo configurable, por ejemplo, cada 10 minutos. La Tabla 6 muestra la cantidad de sensores de cada tipo en las aulas.

Tabla 5. *Sensores Monnit utilizados*




Sensor	Tipo	Mide
	Intervalo	Humedad y temperatura interior
	Eventos	Apertura y cierre de puertas y ventanas
	Eventos	Presencia

Tabla 6. *Cantidad de sensores ubicados por aula*

Aula	Sensor	Cantidad
IE404	Humedad	4
	Apertura y cierre	7
	Movimiento	4
IE405	Humedad	4
	Apertura y cierre	7
	Movimiento	3

Las figuras 8 y 9 muestran la ubicación de los sensores en las aulas IE404 e IE405. En la Figura 10 se observa la ubicación de los sensores en las aulas y la cantidad de sensores en cada una.

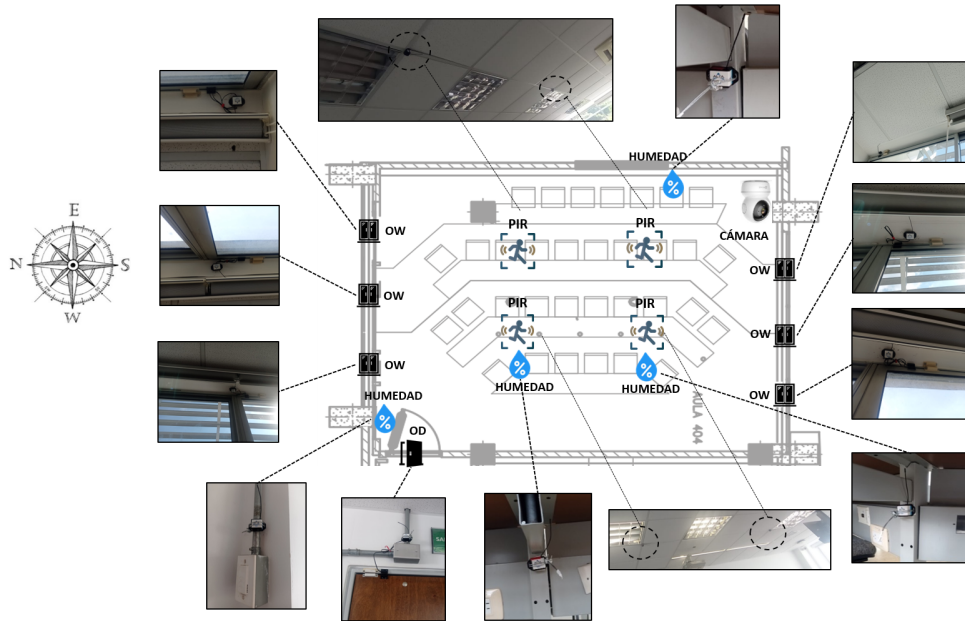


Figura 8. Ubicación de sensores IE404.

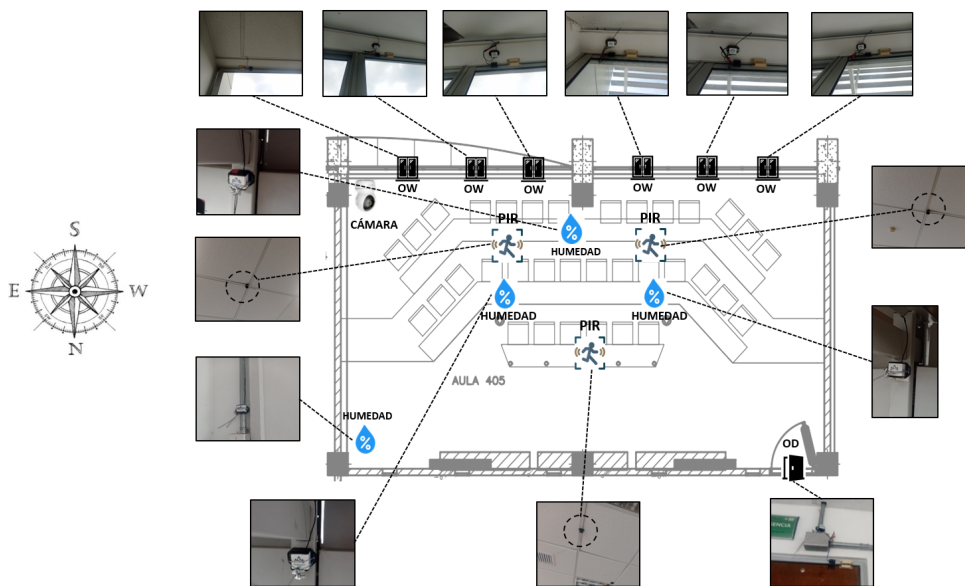


Figura 9. Ubicación de sensores IE405.

1.3. Medidores de energía

El consumo de energía de las aulas fue monitorizado a través de dos tableros eléctricos del edificio. El primero corresponde al tablero de aires acondicionados, ubicado en el quinto piso,

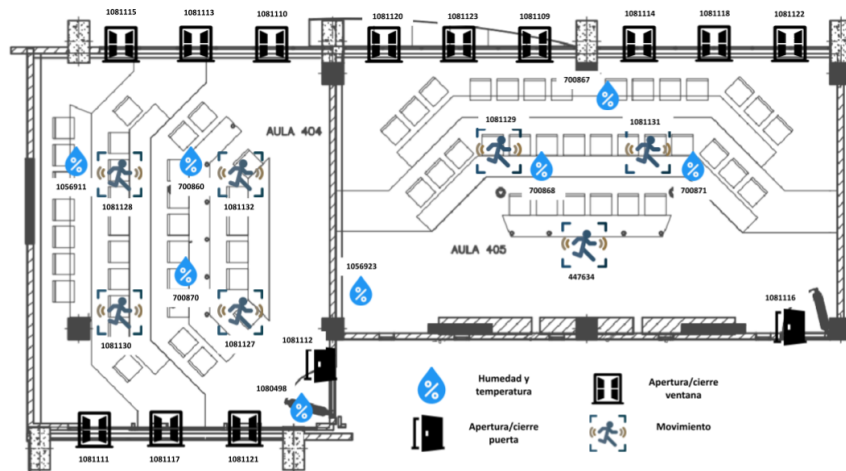


Figura 10. Ubicación y número de los sensores en las dos aulas.

mientras que el segundo es el tablero general de baja tensión del cuarto piso. En este último se identificaron los circuitos correspondientes a la iluminación y los tomacorrientes de las dos aulas.

La Figura 11 se muestra el medidor correspondiente a los aires acondicionados, y en la Figura 12, el medidor de los circuitos de iluminación y tomacorrientes. Los esquemas de conexión de estos medidores se presentan en las figuras 13 y 14. Además, las tablas 7 y 8 describen los circuitos monitorizados, incluyendo sus nombres, los canales del medidor asociados a cada circuito y la fase de alimentación.



Figura 11. Medidor de energía de aires acondicionados (Acurev2020).

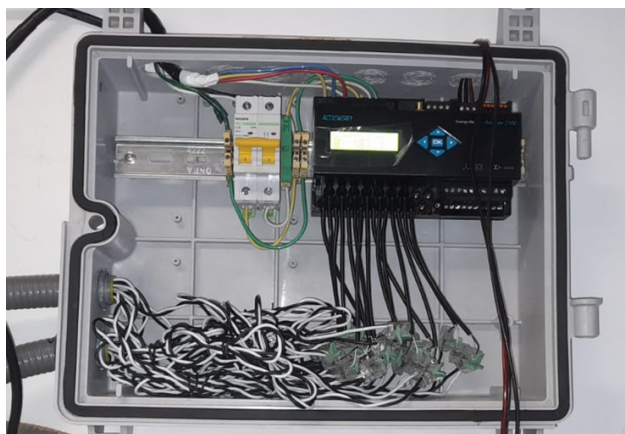


Figura 12. Medidor de energía de circuitos de aulas de clase (Acurev2100).

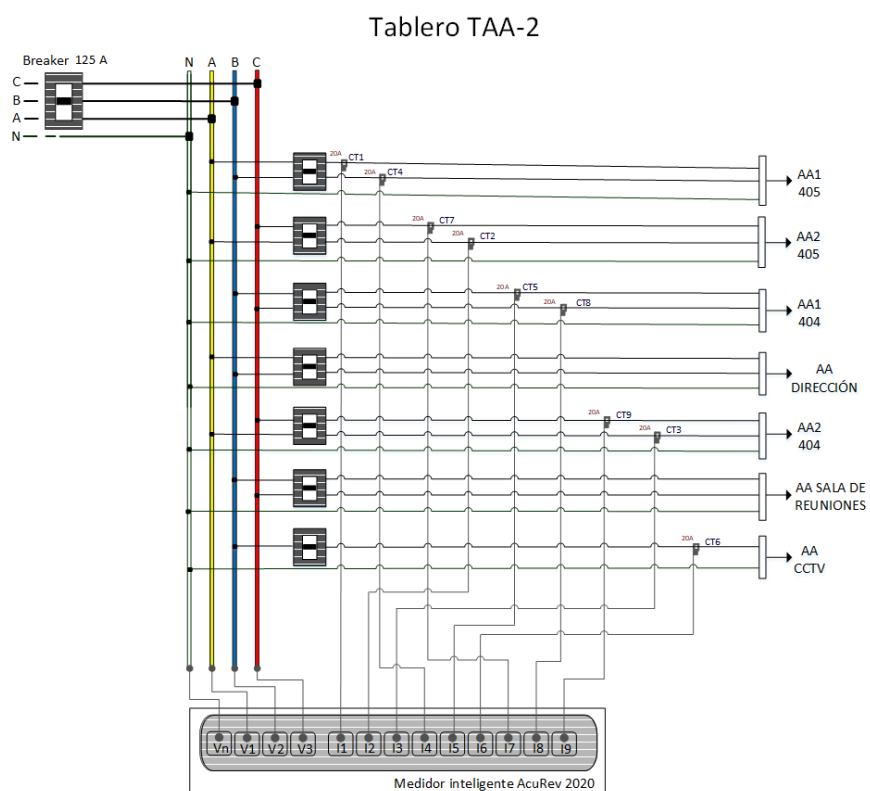


Figura 13. Diagrama unifilar y conexión del medidor.

Cada aula cuenta con dos aires acondicionados, que incluyen una manejadora (unidad interior) y una condensadora (unidad exterior). En la Figura 15 se ilustra el tipo de aire acondicionado

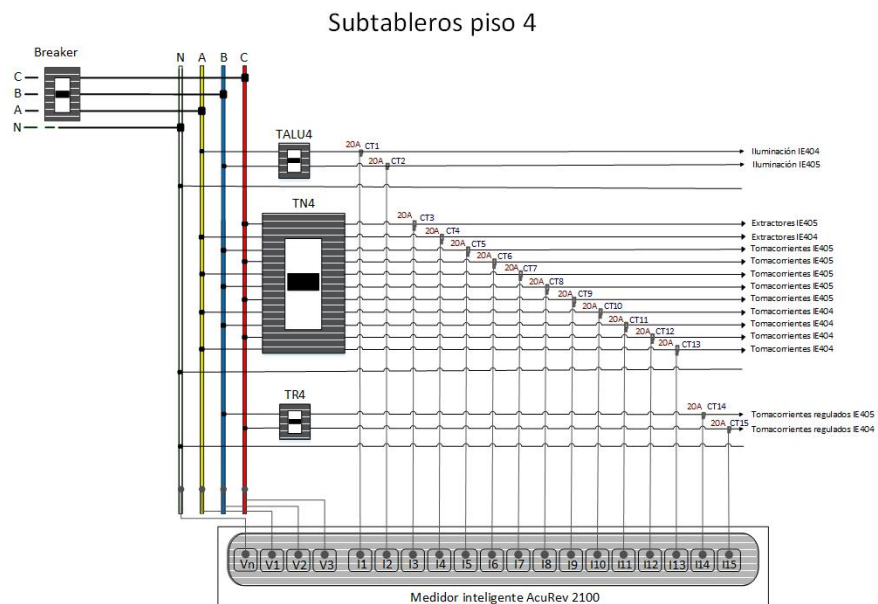


Figura 14. Conexión y configuración medidor de energía ACUREV2100.

Tabla 7. Circuitos medidos por el ACUREV 2020

Canal	Aula	Fase	Alimenta
1	IE405	A	AA1
2	IE405	B	AA1
3	IE404	C	AA1
4	IE405	A	AA2
5	IE404	B	AA1
6	CCTV	C	AA
7	IE405	A	AA2
8	IE404	B	AA2
9	IE404	C	AA2

instalado en las aulas. Finalmente, en el Anexo 1 se presentan las fotografías de los tableros de aires acondicionados y de los subtableros del cuarto piso, junto con el diagrama unifilar correspondiente.

Tabla 8. Circuitos medidos por el ACUREV 2100

Canal	Circuito	Tablero	Aula	Fase
1	1	TALU4	IE404	A
2	2	TALU4	IE405	B
3	27	EX IE405	IE405	C
4	28	EX IE404	IE404	A
5	11	TN4	IE405	B
6	12	TN4	IE405	C
7	13	TN4	IE405	A
8	14	TN4	IE405	B
9	15	TN4	IE405	C
10	16	TN4	IE404	A
11	17	TN4	IE404	B
12	18	TN4	IE404	C
13	19	TN4	IE404	A
14	5	TR4	IE405	B
15	6	TR4	IE404	C



Unidad Interior
(Evaporador)



Unidad exterior
(Condensador)

Figura 15. Sistema de aire acondicionado en aulas.

1.4. Cámaras de vídeo

Se instalaron cámaras de video en los espacios monitorizados, colocando una en cada aula y ubicándolas estratégicamente para facilitar el conteo eficiente de personas y computadoras. El tipo de cámara utilizada fue una Smart Home Camera con detección de movimiento, conversación bi-direccional y resolución de 1080p. Las grabaciones fueron revisadas cada cinco minutos mediante

la aplicación EZVIZ para realizar el conteo correspondiente. En la Figura 16 se observa una vista del aula IE404 desde la aplicación.

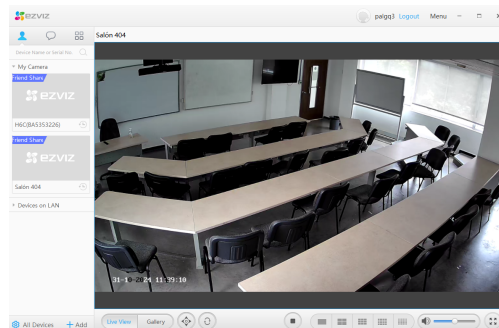


Figura 16. Vista IE404 desde la aplicación EZVIZ.

2. Creación de base de datos

Una vez completada la etapa de construcción del experimento y de la monitorización, se procedió a la creación de una base de datos. El periodo de monitoreo abarcó aproximadamente dos meses, desde el 19 de febrero hasta el 18 de abril de 2024. La Figura 17 presenta un esquema que ilustra el proceso de integración y fusión de datos provenientes de los distintos equipos de monitorización, con el objetivo de generar una base de datos consolidada. Para ello, se descargó un archivo CSV de cada una de las cinco fuentes de datos. Estos archivos fueron leídos y procesados mediante Python, lo que permitió generar un archivo CSV final que contiene las variables resultantes.

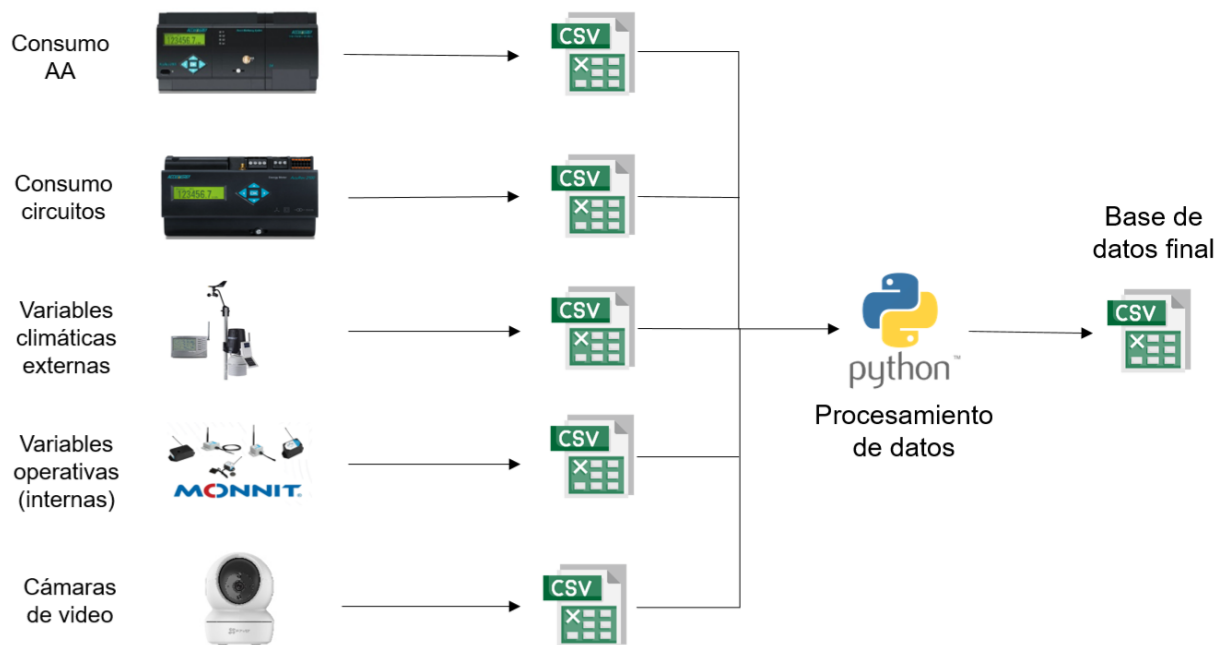


Figura 17. Esquemático creación de la base de datos.

2.1. Procesamiento de los datos

El consumo total del aula se calculó sumando los valores de los dos medidores de energía. Los datos de consumo energético, meteorológicos, de ocupación, temporales y operativos se combinaron para crear un conjunto de datos único. Sin embargo, los sensores Monnit tienen una frecuencia de muestreo de 10 minutos, mientras que las otras variables se miden con mayor frecuencia. Por esta razón, fue necesario ajustar el muestreo de los datos de Monnit a 5 minutos mediante el marco de fusión de datos en Python, utilizando la biblioteca Pandas [29], que permite combinar filas en un DataFrame basado en una o más claves. Los valores faltantes se interpolaron utilizando la función de interpolación temporal de Pandas.

2.2. Códigos empleados para la creación de la base de datos

Para la creación de la base de datos, se desarrollaron varias funciones en VSCode utilizando Python. En el Anexo 2 se encuentran los códigos de las funciones en Python. A continuación, se explican los nombres, las entradas y las salidas de cada una de las funciones:

2.2.1. Función consumo de energía

La función *calculo_consumo_energía(Dataframe,Dataframe1,start_time,end_time,time_interval)* crea una tabla con el consumo de energía de cada aula con un estampado de tiempo de 5 minutos.

Recibe cinco entradas:

- **Dataframe**: Archivo medidor Acurev 2020.
- **Dataframe1**: Archivo medidor Acurev 2100.
- **start_time**: Tiempo de inicio (para el filtrado de los datos).
- **end_time**: Tiempo de fin.
- **time_interval**: Estampado de tiempo.

Devuelve la tabla con los valores de consumo de energía.

2.2.2. Función estación meteorológica

La función *estacion(Raw_data,start_time,end_time,time_interval)* crea una tabla con las variables meteorológicas con un estampado de tiempo de 5 minutos. Recibe cuatro entradas:

- **Raw_data**: Archivo de la estación meteorológica.
- **start_time**: Tiempo de inicio (para el filtrado de los datos).
- **end_time**: Tiempo de fin.
- **time_interval**: Estampado de tiempo.

Devuelve la tabla con los valores de cada variable exterior seleccionada en el Capítulo 1.

2.2.3. Función sensores Monnit 1

La función *sensores(archivo, lista_sensores, lista_sensores_string,start_time,end_time,time_interval)* crea una tabla con los datos de cada variable interna utilizando un estampado de tiempo de 5 minutos. Recibe seis entradas:

- **archivo**: Archivo de sensores Monnit.
- **lista_sensores**: ID de los sensores.
- **lista_sensores_string**: Nombre de los sensores en la nueva tabla.
- **start_time**: Tiempo de inicio (para el filtrado de los datos).
- **end_time**: Tiempo de fin.
- **time_interval**: Estampado de tiempo.

Devuelve la tabla con los datos de todos los sensores por aula (el código se debe correr individualmente para cada aula).

2.2.4. Función sensores Monnit 2

La función *Sensor_Modified(Table)* crea una tabla con los datos de cada variable interna utilizando un estampado de tiempo de 5 minutos. Recibe una entrada:

- **Table**: Tabla generada en la función Sensores Monnit 1.

Devuelve la tabla con los valores de las variables, promedia la humedad y la temperatura, unifica

el movimiento y el estado de las ventanas.

2.2.5. Función de las cámaras

La función *camera_data(archivo,start_time,end_time)* crea una tabla con los datos de cada variable interna utilizando un estampado de tiempo de 5 minutos. Recibe tres entradas:

- **archivo:** Archivo de sensores Monnit.
- **start_time:** Tiempo de inicio (para el filtrado de los datos).
- **end_time:** Tiempo de fin.

Devuelve la tabla con los valores de las variables internas por aula individualmente.

2.2.6. Función de la tabla final

La función *Final_Table(Table_consumo,weather_station,Sensors,Cameras)* crea una tabla con los datos de cada variable interna utilizando un estampado de tiempo de 5 minutos. Recibe cuatro entradas:

- **Table_consumo:** Tabla del consumo de energía.
- **weather_station:** Tabla de la estación meteorológica.
- **Sensors:** Tabla de los sensores Monnit.
- **Cameras:** Tabla de las cámaras.

Devuelve la tabla final de la base de datos.

El conjunto final de datos consta de 22 variables, y las variables categóricas fueron transformadas en variables numéricas. La Tabla 9 detalla los valores y significados de las variables categóricas. La granularidad seleccionada para los datos permite una apreciación más precisa de los cambios en las variables durante intervalos cortos de tiempo. Después de tener la tabla o base de datos se agregan tres variables temporales utilizando librerías de Python. Se toma la fecha para crear las variables categóricas: Día de la semana, Hora del día y Día Laboral. La base de datos final para cada una de las aulas se muestra en la Figura 18, las variables se clasifican en tres grupos: X_{tiempo} son las variables temporales, X_{ext} son las variables climáticas en el exterior y X_{int} son las variables internas del aula. Y es la variable objetivo.

Tabla 9. *Valores de las variables categóricas y su significado.*

Variable	Valor	Significado
Día de la semana	0-6	Lunes a Domingo
Hora del día	0-23	0 horas a 23 horas
Día laboral	0;1	Día no laborable
		Día laborable
Dirección del viento	0	N
	22,5	NNE
	45	NE
	67,5	ENE
	90	E
	112,5	ESE
	135	SE
	157,5	SSE
	180	S
	202,5	SSW
	225	SW
	247,5	WSW
	270	W
	292,5	WNW
315	NW	
337,5	NNW	
360	---	
Estado de la puerta	1;0	Cerrada; Abierta
Estado de la ventana	0-6	Número de ventanas cerradas
Movimiento	1;0	Movimiento;No movimiento
Ocupación	1;0	Ocupado;Desocupado

Xtiempo			Xext													Xint				Y	
Día de la semana	Hora del día	Día laboral	T. Exterior [°C]	H. Exterior [%]	Punto de rocío [°C]	Velocidad del viento [m/s]	Dirección del viento	Índice de calor [°C]	Presión atmosférica [mmHg]	Cantidad de lluvia [mm/hr]	Radiación solar [W/m²]	Índice UV	Grados día de enfriamiento [°C]	Estado de la puerta	Estado de ventanas	T. Interior [°C]	H. Interior [%]	Movimiento	N. de ocupantes	N. de computadores	Consumo de energía [Wh]
2	15	1	27,8	74	22,7	1,8	157,5	30,9	751,3	0	279	1,9	0,03	1	6	24,28	66,09	0	4	4	96,50
2	15	1	27,8	75	22,9	2,2	157,5	31,1	751,4	0	520	1,9	0,03	1	6	24,28	66,09	1	4	4	263,58
2	15	1	27,9	74	22,9	1,3	157,5	31,3	751,4	0	455	1,6	0,03	1	6	24,28	66,31	1	2	0	265,83
2	15	1	28,4	75	23,5	2,2	157,5	32,4	751,4	0	336	1,3	0,04	1	6	24,28	66,31	0	0	0	74,00
2	16	1	27,9	74	22,9	1,3	157,5	31,3	751,4	0	203	0,9	0,03	1	6	24,05	65,43	0	0	0	230,08

Figura 18. Base de datos final.

2.3. Normalización de datos

Antes de entrenar los modelos, es esencial normalizar los datos, ya que las variables tienen rangos significativamente diferentes. La normalización mejora la integridad de los datos, mitiga los problemas causados por las características dominantes y ayuda a lograr la convergencia en el entrenamiento de los modelos. La normalización Min-Max Scaling transforma cada característica para que esté en el rango de 0 a 1 [30].

3. Entrenamiento de modelos

Este capítulo presenta los modelos que se emplean para entrenar y predecir el consumo de energía, describiendo brevemente las metodologías seleccionadas y su aplicación en el contexto del estudio.

3.1. Regresión Lineal Múltiple (*Multiple Linear Regression - MLR*)

El método de regresión lineal relaciona variables y ofrece algunas ventajas, como la facilidad de uso y la interpretabilidad. El modelo de regresión lineal es definido en 1, donde \hat{y} es el valor predicho, β_0 , β_1 , β_2 y β_n son los coeficientes de regresión y X son los datos observados. La relación entre las variables medidas y la variable de respuesta no sugiere necesariamente causalidad, sino que existe una correlación significativa entre las variables [31]. El consumo de energía depende de actividades específicas de los ocupantes (como la diferencia entre días festivos y laborables), de las condiciones meteorológicas y de acontecimientos inesperados [32].

$$\hat{y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (1)$$

Para la selección de características se dividieron los datos en 80 % para el entrenamiento y 20 % para la prueba. Luego se utilizó el método Selector Secuencial de Características con una dirección hacia atrás, tomando R^2 como métrica de seguimiento. Una vez seleccionadas las variables de entrada se aplicó validación cruzada para evaluar el modelo calculando la media y la desviación estándar.

3.2. Media Móvil Autorregresiva Integrada (*Autoregressive Integrated Moving Average - ARIMA*)

ARIMA es un método que utiliza datos históricos para analizar series temporales y predecir valores futuros [33]. Los componentes del modelo ARIMA son la parte autorregresiva (AR), la

diferenciación (I) y la media móvil (MA) [34]. Inicialmente en la Ecuación 2 se observa la forma general del modelo ARMA.

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

Donde:

- Y_t : El valor de la serie temporal en el tiempo t .
- μ : La constante o intercepto del modelo, que representa el valor promedio de la serie temporal si se incluye.
- $\phi_1, \phi_2, \dots, \phi_p$: Los coeficientes autoregresivos (AR) del modelo, que indican la influencia de los valores pasados de la serie sobre el valor actual.
- p : El orden del componente autoregresivo (AR). Representa el número de rezagos de la serie temporal considerados en el modelo.
- ε_t : El error o perturbación en el tiempo t , también llamado “residuo”, que refleja la parte de la serie no explicada por el modelo.
- $\theta_1, \theta_2, \dots, \theta_q$: Los coeficientes del modelo de media móvil (MA), que indican la influencia de los errores pasados sobre el valor actual.
- q : El orden del componente de media móvil (MA). Representa el número de rezagos de los errores considerados en el modelo.
- t : El índice temporal o periodo en el que se observa el valor de la serie.

El modelo se convierte en un modelo ARIMA cuando la serie de tiempo Y_t debe ser diferenciada para que sea estacionaria. Se debe diferenciar d veces hasta que se cumpla con las condiciones de estacionariedad (media, varianza y covarianza constantes). Se aplica entonces ARMA a la serie de tiempo diferenciada Y_t .

Para entrenar este modelo se utiliza TimeSeriesSplit ya que, al tratarse de una serie temporal, no es adecuado utilizar la validación cruzada tradicional porque se perdería la naturaleza temporal de los datos. Se utilizó la función auto_arima de la librería estadística **pmdarima**. Se eligieron cuatro divisiones y el modelo se entrenó con estos datos. También se calcularon la media y la desviación estándar. La Tabla 10 explica cómo se distribuyeron los datos de cada separación.

Tabla 10. *Índices de entrenamiento y prueba de cada división.*

División	Índice de entrenamiento	Índice de prueba
1	[0; 1; ...; 2878]	[2879; 2880; ...; 5755]
2	[0; 1; ...; 5755]	[5756; 5757; ...; 8632]
3	[0; 1; ...; 8632]	[8633; 8634; ...; 11509]
4	[0; 1; ...; 11509]	[11510; 11511; ...; 14386]

3.3. Red Neuronal Artificial Profunda (*Deep Neural Network* - DNN)

Las redes neuronales artificiales (ANN) se basan en la organización estructural del cerebro humano. Las ANN suelen considerarse cajas negras porque su funcionamiento interno no es directamente observable. Se necesita una gran cantidad de datos para entrenar la red neuronal artificial; su estructura distribuida masivamente paralela y sus capacidades de generalización y aprendizaje permiten resolver problemas complejos [35]. Las redes neuronales profundas (DNN) son más complejas porque tienen más capas ocultas y requieren más recursos computacionales. Las LSTM son un tipo de DNN. Tienen una célula de memoria controlada por tres puertas clave: la puerta de entrada, que regula qué información se añade a la célula de memoria; la puerta de olvido, que determina qué información se elimina de la célula de memoria; y la puerta de salida, que gestiona la información que se elimina de la célula de memoria. La Tabla 11 resume el entrenamiento de los modelos DNN y LSTM. Para la DNN, se aplicó validación cruzada (4 pliegues) y se calcularon la media y la desviación estándar de los errores. La Figura 19 ilustra la arquitectura, mostrando el número de capas de entrada, ocultas y de salida y los nodos en cada capa.

Tabla 11. Resumen del entrenamiento de redes neuronales

Parámetro	DNN	LSTM
Selección de características	Regresión Lasso: LassoCV	Serie temporal: Consumo de energía
Partición de los datos	Validación cruzada: 4 pliegues	División de las series de tiempo: 3 divisiones
Devoluciones de llamada	Parada temprana Punto de control del modelo ReduceLRonPlateau	Parada temprana Punto de control del modelo ReduceLRonPlateau
Compilación	Optimizador: RMSprop Pérdida: MSE	Optimizador: Adam Pérdida: MSE
Entrenamiento	Épocas: 1000 Tamaño del lote: 64	Épocas: 53 Tamaño del lote: 64

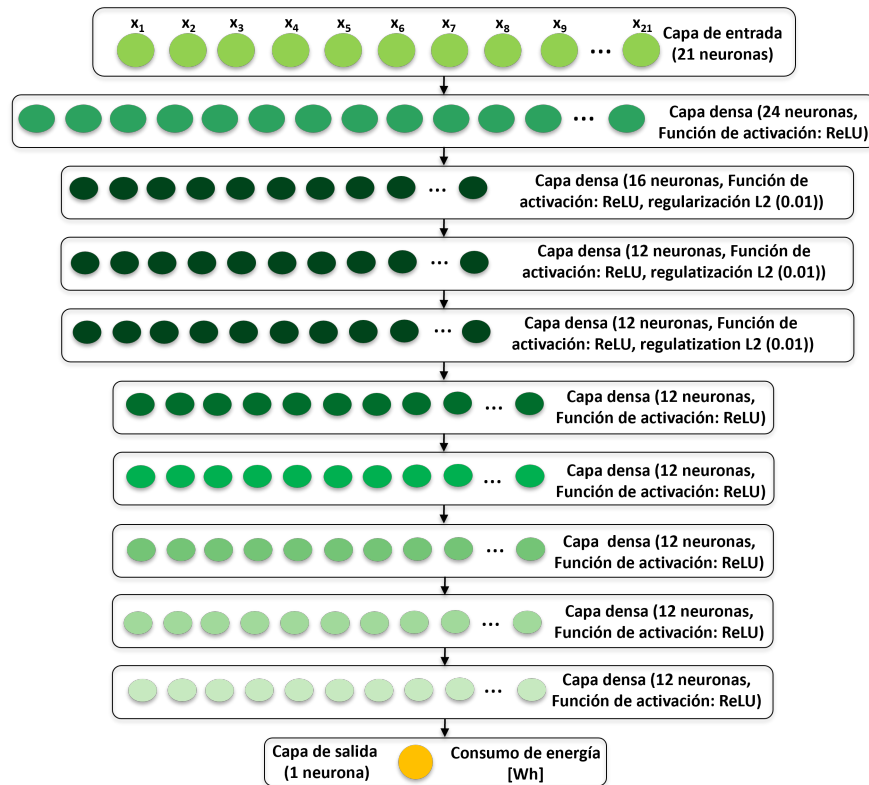


Figura 19. Estructura de la DNN.

La red neuronal LSTM entrenada en este estudio se basó inicialmente en la red descrita

en [19]. Tras varios entrenamientos para optimizar la red neuronal, el modelo se estructuró de la siguiente manera: dos capas LSTM de 32 unidades cada una con función de activación ReLU, dos capas densas de 5 unidades cada una, y una capa densa de una neurona para la salida. Los datos se ajustaron como secuencias con un paso de tiempo de 10. Se guardaron los mejores modelos de cada división y se utilizaron para evaluar los datos de prueba, calculando la media y la desviación estándar de las métricas.

3.4. DesignBuilder (DB)

DesignBuilder es un programa informático basado en EnergyPlus. El programa permite evaluar el comportamiento medioambiental de edificios nuevos y existentes. Los modelos integrados permiten realizar análisis de rendimiento en materia de energía, confort, sistemas de climatización, iluminación natural, costes, optimización del diseño, CFD y créditos BREEAM/LEED. Además, genera informes que cumplen varias normativas de construcción y estándares de certificación nacionales [36].

El modelo completo del edificio está disponible; fue necesario identificar y separar las dos aulas (IE404 e IE405) de las demás zonas del edificio. Después de esto, se actualizaron los horarios y las densidades de ocupación, el uso de equipos, el uso de HVAC y el uso de iluminación para las dos aulas. Se simuló una semana del 4 al 10 de marzo de 2024. La Figura 20 muestra el modelo del edificio en el software y el edificio real.



Figura 20. *Modelo de DesignBuilder.*

3.5. Modelo híbrido: DesignBuilder + DNN

El modelo híbrido integra los datos generados en DB como entradas a una DNN. Este enfoque optimiza el tiempo de predicción y mejora el rendimiento del modelo, ya que los métodos basados en simulaciones de ingeniería suelen ser costosos desde el punto de vista computacional. El software proporciona series temporales de datos de comportamiento energético, que se procesan y utilizan como entradas de la DNN previamente entrenada. La variable objetivo del modelo es el consumo de energía, que se define como la suma del consumo de iluminación, refrigeración y otros usos energéticos. Las variables de entrada de la DNN, descritas en la Tabla 12, incluyen el nombre y la definición de cada variable. Los datos generados por DB tienen un intervalo de tiempo de 5 minutos, acumulando un total de 2016 registros por variable (correspondientes a una semana de datos). Estas variables reflejan, en general, las pérdidas o ganancias de calor debidas a las condiciones específicas del entorno simulado. Para el entrenamiento del modelo, los datos se dividieron en un (80%) para el entrenamiento y un (20%) para las pruebas, garantizando una adecuada generalización del modelo.

Tabla 12. *Base de datos Design Builder*

Característica	Definición
Acrilamiento	Instalación de cristales en ventanas, puertas o fachadas.
Paredes	Construcción que permite delimitar o dividir un espacio.
Techo (int)	La superficie superior interior que cubre los límites superiores de un aula.
Suelos (int)	La superficie del fondo de un espacio sobre la que se camina.
Particiones (int)	Panel o pared no estructural utilizado para dividir espacios interiores.
Refrigeración sensible	Proceso de reducción de la temperatura del aire en el espacio mediante la eliminación del calor sensible.
Infiltración externa	Entrada involuntaria o incontrolada de aire exterior en un espacio cerrado.
Computador + Equipos	Equipamiento electrónico de las aulas.
Ocupación	Uso o presencia de personas en un edificio o espacio.
Ganancias solares de las ventanas exteriores	Aumento del calor en el interior de un edificio como consecuencia del paso de los rayos solares a través de sus ventanas.
Temperatura del aire	Medida de la calidez o frialdad del aire.
Temperatura radiante	Calor emitido en forma de radiación por los elementos del entorno.
Temperatura exterior de bulbo seco	Temperatura del aire ambiente medida en el exterior de un edificio.
Temperatura Operativa	Valor medio calculado entre la temperatura del aire seco y la temperatura radiante media de los recintos del aula (paredes, suelo y techo).

4. Evaluación de modelos

Los parámetros para evaluar el rendimiento de los modelos son el coeficiente de determinación (R^2), el error medio absoluto (MAE) y el error cuadrático medio (RMSE). El R^2 representa la proporción prevista de la varianza de la variable dependiente que pueden explicar las variables independientes. MAE es la diferencia media entre las variables esperadas y las reales. RMSE es la magnitud promedio del error entre las predicciones y los valores reales. Las métricas anteriores se definen del siguiente modo:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

Donde y_i son los datos reales, \hat{y}_i son los valores predichos, \bar{y} es el valor medio de la variable dependiente, y n es el número de observaciones.

Al combinar estos indicadores, se obtiene una visión más completa y detallada que permite una evaluación exhaustiva del modelo. Por ejemplo, dado que el RMSE es más sensible a los valores atípicos que el MAE, una diferencia significativa entre ambos puede indicar la presencia de estos valores extremos. En contraste, el R^2 refleja la capacidad del modelo para predecir con precisión muestras futuras [37]. Cada uno de estos índices de evaluación proporciona una perspectiva única sobre la efectividad de los modelos predictivos.

5. Resultados de la investigación

Este capítulo proporciona una descripción detallada de los datos, presenta los resultados obtenidos en esta investigación y ofrece un análisis y discusión de los mismos.

5.0.1. Descripción de datos

Es esencial ser consciente de algunos patrones y problemas en los datos antes de revisar los resultados de las predicciones del modelo. El consumo promedio de energía para las dos aulas según el día de la semana y la hora del día se muestran en las figuras 21 y 22; las dos aulas tienen un alto consumo de energía por las tardes, especialmente IE405 los martes de 2:00 p.m. a 7 p.m. Figura 23 muestra el boxplot del consumo de energía para cada día de la semana, el consumo medio de energía en IE405 está entre 200 Wh y 300 Wh en días laborables, en IE404 el consumo es más regular los sábados que en IE405. Los datos se filtraron para incluir sólo los registros con ocupación en las aulas.

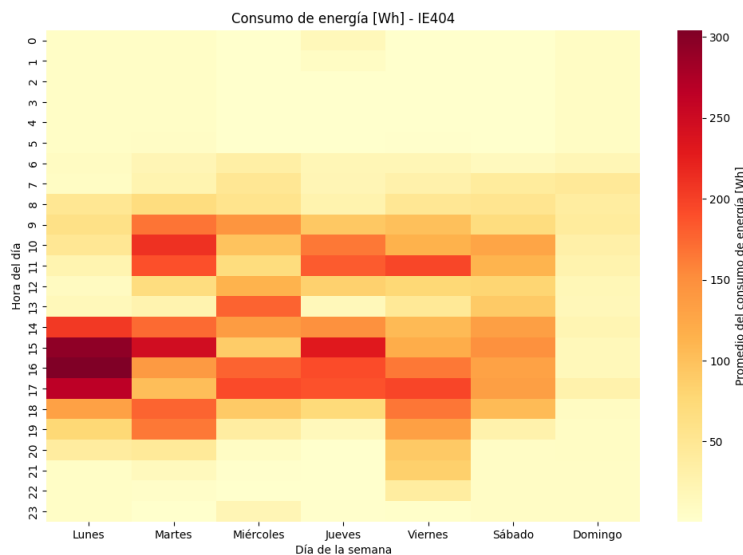


Figura 21. Promedio del consumo de energía del IE404.

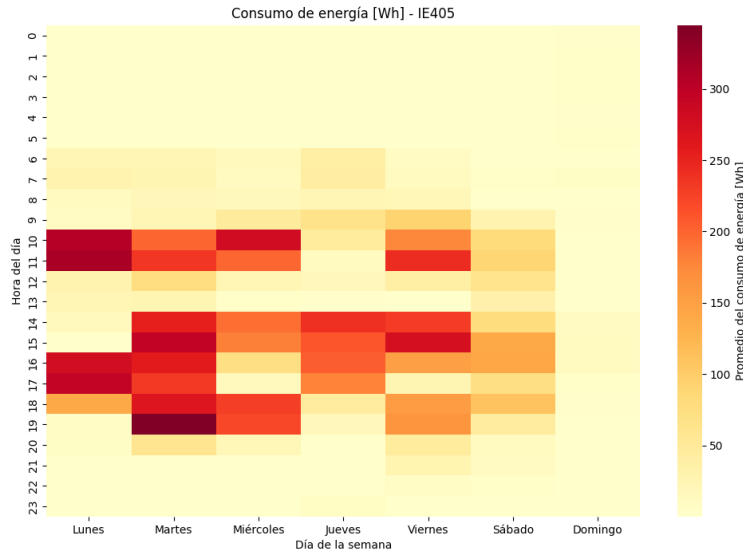


Figura 22. Promedio del consumo de energía del IE405.

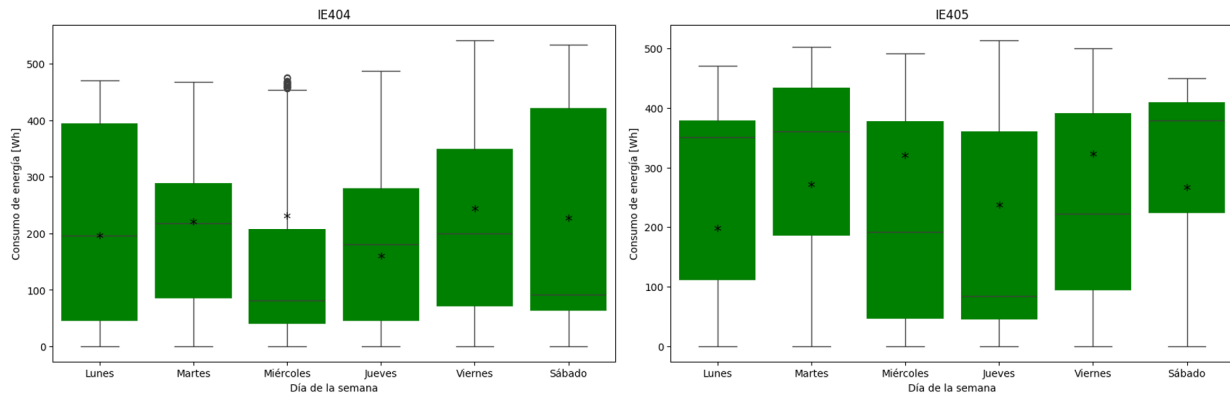


Figura 23. Consumo de energía por día de la semana.

La Figura 24 muestra la correlación lineal entre las variables monitorizadas. Conocer y comprender esta correlación entre cada variable es fundamental, especialmente en lo que respecta al consumo de energía. Muestra la influencia de cada variable en el consumo de energía y entre sí. Proporciona una visión preliminar de las posibles variables de entrada para los modelos de aprendizaje automático. Durante el entrenamiento del modelo se utilizaron técnicas específicas para seleccionar características más detalladas. En la matriz de correlaciones, los valores por encima

de la diagonal corresponden al aula IE405, mientras que los valores por debajo de la diagonal representan las correlaciones del aula IE404.

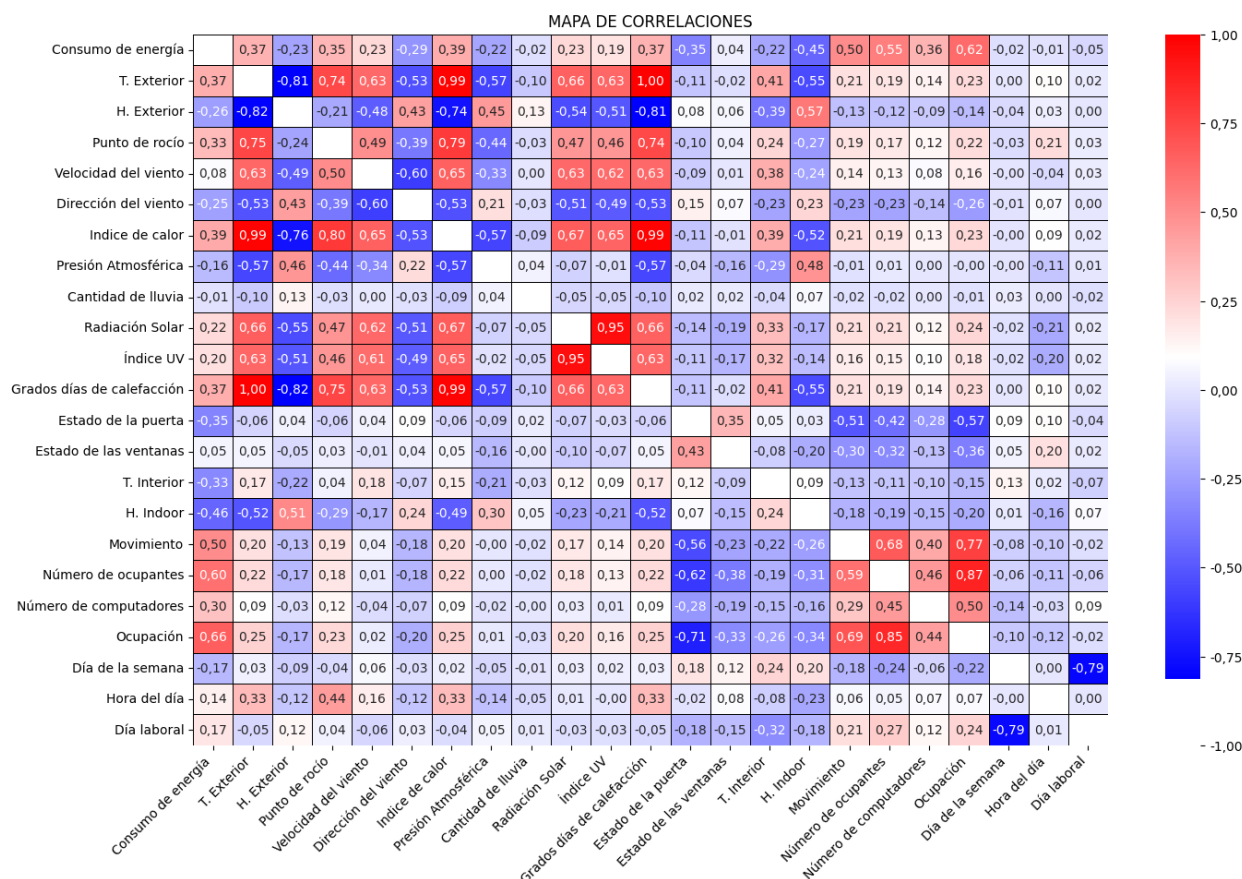


Figura 24. Mapa de correlaciones de variables. Superior a la diagonal IE405, inferior a la diagonal IE404.

5.0.2. Discusión

A diferencia de otros estudios que tuvieron en cuenta un número relativamente limitado de variables, en este se consideró un número considerablemente mayor de variables [4], [19]. Además, en otro estudio, no se tuvo en cuenta el número de ocupantes a pesar de considerar variables climáticas internas y externas y el consumo histórico de energía. En su lugar, se emplearon funciones de probabilidad de uso de la carga. Tampoco se incluyeron variables temporales ni el movimiento de las personas [4]. En este estudio se llevó a cabo una monitorización exhaustiva, lo que permitió

un análisis minucioso y detallado del comportamiento de las aulas.

Se realizaron tres análisis en el entrenamiento de los modelos debido a las diferencias en el contexto temporal y las variables de entrada. El primero son los resultados obtenidos por los modelos de aprendizaje automático y estadístico, estos son: MLR, DNN, ARIMA y LSTM. En este caso, los resultados de la media y la desviación estándar de los errores para cada modelo permiten generar una estimación robusta del rendimiento del modelo.

La Tabla 13 muestra los valores medios y la desviación estándar de los errores; este proceso se realiza para optimizar un buen modelo que tenga un buen rendimiento para los nuevos datos. En el aula IE405, el modelo DNN mostró valores medios de RMSE de 0,1 y R^2 de 0,86. En IE404, ARIMA mostró un rendimiento de error medio de 0,11, 0,01 y 0,74 de RMSE, MAE y R^2 , respectivamente. El modelo ARIMA mostró un mejor rendimiento en IE404, lo que podría deberse a que los datos presentaban un comportamiento temporal estacionario. En cambio, los datos muestran una dinámica no lineal más compleja en IE405, lo que permite un mejor rendimiento de la DNN. Esta diferencia podría explicarse por factores como el tipo de espacio utilizado, los patrones de encendido y apagado del aire acondicionado o la variabilidad del comportamiento de los ocupantes.

El modelo RLM no puede superar a DNN y ARIMA, ya que su capacidad para captar patrones no lineales dentro del conjunto de datos es limitada. Del mismo modo, LSTM no supera a DNN, ya que puede requerir un conjunto de datos más extenso para el entrenamiento debido a la complejidad inherente del modelo y a que los datos no tienen fuertes dependencias temporales.

Tabla 13. Media y desviación estándar de los errores

		Entrenamiento				
Aula	Error		RLM	ARIMA	DNN	LSTM
IE405	RMSE	promedio	0,14	0,12	0,1	0,18
		std	0,01	0,01	0,02	0,08
	MAE	promedio	0,07	0,01	0,04	0,02
		std	0,01	0	0	0,03
	R^2	promedio	0,66	0,75	0,86	0,32
		std	0,8	0,1	0,7	0,25
IE404	RMSE	promedio	0,12	0,11	0,12	0,13
		std	0,01	0,01	0,01	0,07
	MAE	promedio	0,06	0,01	0,05	0,06
		std	0,01	0	0,02	0,01
	R^2	promedio	0,62	0,74	0,7	0,52
		std	0,15	0,02	0,08	0,03

Un segundo análisis de este estudio consiste en comparar los resultados de la simulación en DesignBuilder con los datos reales. En este caso, los resultados para la semana considerada se muestran gráficamente en la Figura 25. Se puede observar que los datos obtenidos por el programa siguen unos patrones de consumo concretos, pero no recogen picos significativos o consumos elevados, y se muestran consumos en momentos en los que no debería haberlos. En el aula IE405, el RMSE fue de 0,322 y el MAE de 0,174. También tiene un R^2 de -0,243 indicando que el modelo no se ajusta bien a los datos y en IE404 el RMSE fue de 0,337, el MAE de 0,193 y, un R^2 de -0,64. La discrepancia entre los datos simulados y los reales puede atribuirse a la dependencia del modelo respecto a las entradas proporcionadas. A pesar de la posibilidad de modelizar el comportamiento de los usuarios en cuanto al uso del espacio y los equipos, se puede incurrir en imprecisiones que ocasionan incertidumbre en el modelo. Además, presentan otros inconvenientes importantes:

requieren múltiples comprobaciones y ajustes complejos para garantizar una fiabilidad adecuada de los resultados previstos. Dado que se necesita una gran cantidad de datos, es difícil recopilar los parámetros del edificio con el nivel de detalle necesario [38]. Se necesitan muchos datos, incluida información detallada sobre la estructura del edificio, los sistemas del edificio, el comportamiento de los ocupantes y las condiciones meteorológicas. La naturaleza impredecible del comportamiento de los ocupantes puede dar lugar a discrepancias entre las predicciones teóricas y el uso real de la energía. La incorporación precisa de datos meteorológicos en los modelos es difícil debido a la variabilidad de las condiciones microclimáticas, que pueden afectar a las predicciones de uso de la energía [39], [40].

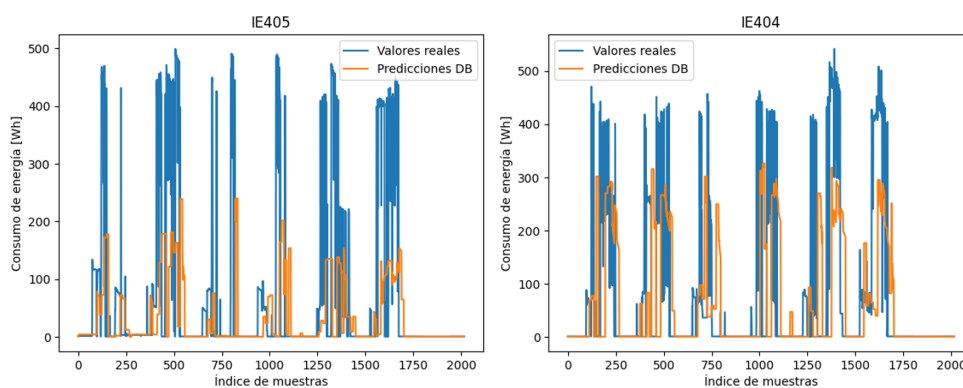


Figura 25. Resultados del software DesignBuilder.

El modelo híbrido combina el modelado físico del software de simulación, que genera datos sobre el consumo de energía y las variables físicas que lo afectan, con una red neuronal profunda (DNN). Estos datos se utilizan como entrada a la DNN, lo que permite identificar patrones complejos que el modelo físico no logra captar. Los resultados del modelo híbrido se comparan con los datos de consumo energético proporcionados por DesignBuilder. La evaluación de este modelo con los datos de prueba se puede analizar en la Figura 26; se puede observar que los valores de RMSE fueron inferiores a 0,065, y el MAE fue inferior a 0,035. Por otra parte, el R^2 fue superior a 0,93. En general, estos resultados indican que este modelo funciona de forma óptima, capturando la mayor parte de la variabilidad en el consumo energético de las aulas en función de las variables

de entrada. El apoyo de las redes neuronales mejora los procesos de entrenamiento de los modelos de simulación. El modelo híbrido puede utilizarse cuando se dispone de un modelo de simulación fiable. En tal caso, aunque los datos proporcionados por el software para una sola semana no sean óptimos, el entrenamiento de una DNN se aproximará mucho al consumo calculado por el software. Este tipo de modelo puede utilizarse para calibrar simulaciones, realizar predicciones en tiempo real y aumentar la precisión de los resultados. En un estudio anterior, se utilizó un modelo híbrido para predecir el consumo energético de una mezquita, combinando el software IESVE para el modelado energético con un modelo de aprendizaje profundo para la predicción. Este enfoque alcanzó un R^2 de 0,98, un valor muy cercano a los obtenidos en el presente estudio, que fueron de 0,93 y 0,94 [17].

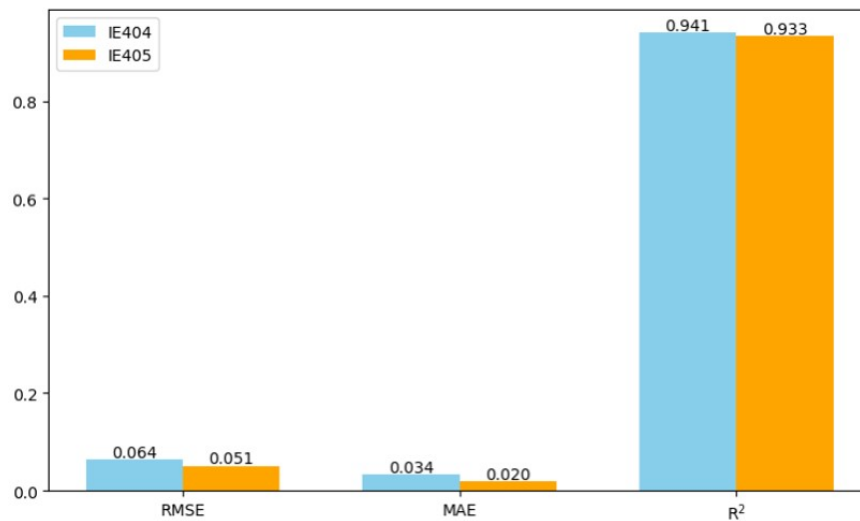


Figura 26. Errores del modelo híbrido.

Este trabajo pretende apoyar a los gestores de edificios en el desarrollo de planes estratégicos para una gestión energética sostenible. Los modelos propuestos en esta investigación son integrables en sistemas avanzados como BAS, BEMS y gemelos digitales, facilitando así la toma de decisiones informadas y la optimización del consumo energético. La metodología propuesta en este estudio puede extenderse a otros espacios de un edificio mediante la instalación de los sensores y equipos mostrados en este estudio, considerando cuidadosamente las especificaciones

técnicas de cada dispositivo y la disposición espacial de los sensores. Para el desarrollo del modelo de simulación, es esencial disponer de información detallada sobre las características físicas, las condiciones climáticas y los equipos instalados en el edificio. A diferencia de estudios similares realizados en climas no tropicales, la principal diferencia radica en los sistemas de calefacción, factor clave a tener en cuenta en países con este tipo de clima [16]. Por otra parte, Yunho Kim *et al.* [14] realizaron un estudio sobre la predicción del consumo de energía en edificios situados en Corea del Sur. En ese contexto, el diseño arquitectónico de los edificios y los patrones de consumo y ocupación varían significativamente en función de la estación del año.

El rendimiento de los modelos podría mejorar considerablemente si se dispusiera de un mayor volumen de datos. Según la literatura especializada en inteligencia artificial, un aumento de la cantidad de datos contribuye generalmente a un aprendizaje más eficaz por parte de los modelos. Por lo tanto, para futuras investigaciones, se recomienda ampliar el periodo de monitorización y capturar datos con una granularidad de media hora o una hora, lo que permitiría una mejor representación de los patrones subyacentes en el consumo de energía.

La metodología empleada en este estudio difiere de las utilizadas para predecir el consumo energético de todo un edificio debido a la complejidad que entraña la instalación de los equipos de vigilancia. Esto dificulta la instalación de más sensores, contadores y cámaras. Aunque los modelos resultantes tendrían más entradas o características, su estructura interna no se vería afectada significativamente. El modelo de simulación requiere más detalles sobre el edificio y sus sistemas. Esto incluye considerar una mayor variedad de escenarios o espacios del edificio y establecer cómo integrarlos eficazmente para la supervisión y el análisis.

Aunque los datos utilizados en este estudio han resultado útiles, es esencial señalar que las limitaciones pueden afectar a la generalizabilidad de los resultados. El tamaño de la muestra era relativamente pequeño, lo que podría afectar a la representación estadística y hacer que los resultados fueran más sensibles a los valores atípicos. Además, la recogida de datos se limitó a un entorno específico con un clima tropical cálido, lo que restringe la aplicabilidad de las conclusiones a otros climas o contextos. También es esencial tener en cuenta posibles sesgos en el proceso de recogida

de datos o la ausencia de determinadas variables vitales, ya que podrían afectar a la precisión de los modelos utilizados. A pesar de estas limitaciones, los resultados constituyen una base sólida para futuras investigaciones. Se recomienda que los estudios futuros se centren en ampliar la base de datos, recopilar datos en una variedad de contextos y realizar pruebas adicionales para validar los modelos en entornos más amplios. De este modo, se podrá reforzar la generalizabilidad de los resultados y superar los obstáculos identificados.

La Tabla 14 presenta un análisis comparativo de los modelos entrenados en este estudio. Los aspectos evaluados incluyen el tiempo de modelado, correspondiente al periodo necesario para ajustar los modelos y garantizar su fiabilidad, y el tiempo de entrenamiento, definido como el tiempo necesario para que el modelo realice predicciones. Se observa que los modelos basados en la simulación son los que más tiempo consumen, tanto en términos de modelado como de entrenamiento. Además, también se tienen en cuenta factores como el coste de los programas utilizados para la modelización, las herramientas informáticas empleadas durante el proceso, el software seleccionado y el nivel de experiencia del usuario necesario para llevar a cabo todo el desarrollo del modelo.

Tabla 14. *Análisis comparativo de los procesos y recursos utilizados en la formación de los modelos*

Modelo	Tiempo de modelado	Tiempo de entrenamiento	Costo	Herramientas computacionales	Software	Experiencia requerida
RLM	1 día	≈ 20 segundos	Gratuito	Google Colab, computador*, RAM de 12 GB y conexión a internet	Python	Programación, ciencia de datos y aprendizaje automático
DNN	1 semana	≈ 4 minutos	Gratuito	Google Colab, computador*,		
LSTM	1 semana	≈ 3.5 minutos	Gratuito	GPU, RAM de 12 GB y conexión a internet		
ARIMA	1 semana	≈ 20 minutos	Gratuito			
Simulación	1 a 2 meses	≈ 1 hora	Licencia de pago*	Ordenador de alto rendimiento*	DesignBuilder	Simulación y modelado energético
Híbrido	1 a 2 meses	≈ 1.06 horas		Ordenador de alto rendimiento*, Google Colab, RAM de 12 GB, y conexión a internet	DesignBuilder Python	Simulación, modelado energético, programación, ciencia de datos y aprendizaje automático

Nota: * Fueron adquiridos con anterioridad a este estudio y fueron financiados por la UIS.

6. Conclusiones

Este trabajo predice el consumo energético de las aulas de un edificio educativo en un clima tropical cálido. Se utilizaron los cuatro enfoques presentados en la literatura. Cada enfoque tiene sus características, estructura y forma de estimar el consumo. Antes de entrenar los modelos, se realizó un análisis de correlación para averiguar las correlaciones más relevantes. Los resultados indicaron que la ocupación y el número de ocupantes eran los factores que más influían en la demanda energética de las dos aulas.

El modelo que mostró el mejor resultado entre los modelos de aprendizaje automático y estadístico fue DNN en el aula IE405. Este modelo específico aprende de las relaciones no lineales de los datos y puede describir el comportamiento del consumo energético, considerando variables internas, operativas y climatológicas. Considera lo que está sucediendo e intenta replicar la realidad.

Los resultados son un coeficiente de determinación R^2 de 0,86 para el aula IE405 y de 0,74 para el aula IE404. Las redes neuronales LSTM no muestran buenos resultados en comparación con otros modelos ya que, al ser complejas, requieren una cantidad de datos más significativa para su entrenamiento. Por otro lado, al entrenar el modelo de simulación se observó que, a pesar de conseguir errores RMSE y MAE bajos, los valores de R^2 eran desfavorables. El valor negativo de R^2 indica que el modelo no está simulando con precisión el consumo de energía basado en datos reales, posiblemente porque es imposible incorporar completamente al modelo los patrones de interacción de los ocupantes con el espacio.

Los valores negativos de R^2 indican que este modelo no es adecuado para este caso de estudio. Sin embargo, este modelo demuestra una excelente capacidad para tener en cuenta la influencia del clima en los datos y ha sido validado internacionalmente para evaluar la energía de los edificios. En este caso, no pueden descartarse posibles deficiencias en la introducción de datos o la necesidad de semanas adicionales de simulación para lograr resultados más óptimos. Por

último, el modelo híbrido proporciona unas métricas de rendimiento excelentes. La integración de datos simulados con redes neuronales artificiales (RNA) reduce significativamente el tiempo de simulación y los requisitos computacionales.

En particular, incluso cuando la precisión de los datos simulados es limitada, su incorporación a una red neuronal puede producir resultados sólidos y satisfactorios. Este tipo de modelo se utiliza principalmente en la fase de diseño de proyectos de construcción más que para simular estructuras existentes. Las conclusiones extraídas se limitan exclusivamente a dos aulas concretas sin extenderse a una generalización de la predicción energética en edificios educativos. Se sugiere la necesidad de seguir investigando en aulas de diferentes tipos climáticos y espacios arquitectónicos para extender la generalización a un contexto más amplio.

Bibliografía

- [1] IEA. «Energy system.» (2024), dirección: <https://www.iea.org/energy-system> (visitado 2024).
- [2] C. Zhang, Y. Zhao, C. Fan, T. Li, X. Zhang y J. Li, «A generic prediction interval estimation method for quantifying the uncertainties in ultra-short-term building cooling load prediction,» *Applied Thermal Engineering*, vol. 173, pág. 115 261, 2020, ISSN: 1359-4311. DOI: <https://doi.org/10.1016/j.applthermaleng.2020.115261>.
- [3] W. E. Forum. «Transforming Energy Demand.» (2024), dirección: <https://www.weforum.org/publications/transforming-energy-demand/>.
- [4] C. Zhang, L. Ma, Z. Luo, X. Han y T. Zhao, «Forecasting building plug load electricity consumption employing occupant-building interaction input features and bidirectional LSTM with improved swarm intelligent algorithms,» *Energy*, vol. 288, pág. 129 651, 2024, ISSN: 0360-5442. DOI: <https://doi.org/10.1016/j.energy.2023.129651>.
- [5] U. Ali, S. Bano, M. H. Shamsi et al., «Urban building energy performance prediction and retrofit analysis using data-driven machine learning approach,» *Energy and Buildings*, vol. 303, pág. 113 768, 2024, ISSN: 0378-7788. DOI: <https://doi.org/10.1016/j.enbuild.2023.113768>.
- [6] R. Wang, R. Rayhana, M. Gholami, O. E. Herrera, Z. Liu y W. Mérida, «Multi-task deep learning for large-scale buildings energy management,» *Energy and Buildings*, vol. 307, pág. 113 964, 2024, ISSN: 0378-7788. DOI: <https://doi.org/10.1016/j.enbuild.2024.113964>.
- [7] C. Morales-Gonzalez, M. Harper, M. Cash et al., «On building automation system security,» *High-Confidence Computing*, vol. 4, n.º 3, pág. 100 236, 2024, ISSN: 2667-2952. DOI: <https://doi.org/10.1016/j.hcc.2024.100236>.

- [8] H. Hosamo, M. H. Hosamo, H. K. Nielsen, P. R. Svennevig y K. Svidt, «Digital Twin of HVAC system (HVACDT) for multiobjective optimization of energy consumption and thermal comfort based on BIM framework with ANN-MOGA,» *Advances in Building Energy Research*, vol. 17, n.º 2, págs. 125-171, 2023. DOI: 10.1080/17512549.2022.2136240.
- [9] y. G. O.-P. L. Ortega-Diaz J. Cárdenas-Rangel, «Estrategias de predicción de consumo energético en edificaciones: una revisión,» *TecnoLógicas*, vol. 26, n.º 56, e2650, 2023. DOI: <https://doi.org/10.22430/22565337.2650>.
- [10] Q. Qiao, A. Yunusa-Kaltungo y R. Edwards, «Hybrid method for building energy consumption prediction based on limited data,» en *2020 IEEE PES/IAS PowerAfrica*, 2020, págs. 1-5. DOI: 10.1109/PowerAfrica49420.2020.9219915.
- [11] T. N. Zhu G Chow T-T, «Short-term load forecasting coupled with weather profile generation methodology,» *Building Services Engineering Research and Technology*, vol. 39, págs. 1310-327, 2018. DOI: 10.1177/0143624417740858.
- [12] A. Alyakoob, S. Hartono, T. Johnson y A. Middel, «Estimating cooling loads of Arizona State University buildings using microclimate data and machine learning,» *Journal of Building Engineering*, vol. 64, pág. 105 705, 2023. DOI: <https://doi.org/10.1016/j.jobe.2022.105705>.
- [13] N. Njimbouom, Soualihou, Lee et al., «Predicting Site Energy Usage Intensity Using Machine Learning Models,» *Sensors*, vol. 23, 2023. DOI: <https://doi.org/10.3390/s23010082>.
- [14] K. Yunho, Y. Park, H. Seo y J. Hwang, «Load Prediction Algorithm Applied with Indoor Environment Sensing in University Buildings,» *Energies*, vol. 16, 2023. DOI: <https://doi.org/10.3390/en16020999>.
- [15] N. Haidar, N. Tamani, Y. Ghamri-Doudane y A. Boujou, «Selective reinforcement graph mining approach for smart building energy and occupant comfort optimization,» *Building and Environment*, vol. 28, pág. 109 806, 2023. DOI: <https://doi.org/10.3390/en16020999>.

- [16] Z. Ni, C. Zhang, M. Karlsson y S. Gong, «A study of deep learning-based multi-horizon building energy forecasting,» *Building and Environment*, vol. 303, pág. 113 810, 2024. DOI: <https://doi.org/10.1016/j.enbuild.2023.113810>.
- [17] M. El-Maraghy, M. Metawie, M. Safaan et al., «Predicting energy consumption of mosque buildings during the operation stage using deep learning approach,» *Energy and Buildings*, vol. 303, pág. 113 829, 2024. DOI: <https://doi.org/10.1016/j.enbuild.2023.113829>.
- [18] A. D. C. A. Babu, R. S. Srivastava y A. C. Rai, «Impact of climate change on the heating and cooling load components of an archetypical residential room in major Indian cities,» *Building and Environment*, vol. 250, pág. 111 181, 2024. DOI: <https://doi.org/10.1016/j.buildenv.2024.111181>.
- [19] A. Mishra, H. R. Lone y A. Mishra, «DECODE: Data-driven energy consumption prediction leveraging historical data and environmental factors in buildings,» *Energy and Buildings*, vol. 307, pág. 113 950, 2024. DOI: <https://doi.org/10.1016/j.enbuild.2024.113950>.
- [20] X. Liang, S. Chen, X. Zhu, X. Jin y Z. Du, «Domain knowledge decomposition of building energy consumption and a hybrid data-driven model for 24-h ahead predictions,» *Applied Energy*, vol. 344, pág. 121 244, 2023. DOI: <https://doi.org/10.1016/j.apenergy.2023.121244>.
- [21] H. Yan, G. Ji y K. Yan, «Data-driven prediction and optimization of residential building performance in Singapore considering the impact of climate change,» *Building and Environment*, vol. 226, pág. 109 735, 2022. DOI: <https://doi.org/10.1016/j.buildenv.2022.109735>.
- [22] B. Akbar, K. P. Amber, A. Kousar, M. W. Aslam, M. A. Bashir y M. S. Khan, «Data-driven predictive models for daily electricity consumption of academic buildings,» *AIMS Energy*, vol. 8, págs. 783-801, 2024. DOI: <https://doi.org/10.3934/energy.2020.5.783>.
- [23] H. Yan, G. Ji y K. Yan, «Data-driven prediction and optimization of residential building performance in Singapore considering the impact of climate change,» *Building and Environment*, vol. 226, pág. 109 735, 2022. DOI: <https://doi.org/10.1016/j.buildenv.2022.109735>.

- [24] T. Chaudhuri, Y. C. Soh, H. Li y L. Xie, «A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings,» *Applied Energy*, vol. 248, págs. 44-53, 2019. DOI: <https://doi.org/10.1016/j.apenergy.2019.04.065>.
- [25] H. Joanna, W. Łukasz, F. Marcin y S. Marek, «Energy Consumption Forecasting for the Digital-Twin Model of the Building,» *Energies*, vol. 15, 12:4318, 2022. DOI: <https://doi.org/10.3390/en15124318>.
- [26] R. Olu-Ajayi, H. Alaka, I. Sulaimon et al., «Building energy performance prediction: A reliability analysis and evaluation of feature selection methods,» *Expert Systems with Applications*, vol. 225, pág. 120 109, 2023. DOI: <https://doi.org/10.1016/j.eswa.2023.120109>.
- [27] M. L. de la Hoz-Torres, A. J. Aguilar, D. P. Ruiz y M. D. Martínez-Aires, «An investigation of indoor thermal environments and thermal comfort in naturally ventilated educational buildings,» *Journal of Building Engineering*, vol. 84, pág. 108 677, 2024. DOI: <https://doi.org/10.1016/j.jobe.2024.108677>.
- [28] A. Almalaq, J. Hao, J. J. Zhang y F.-Y. Wang, «Parallel building: a complex system approach for smart building energy management,» *IEEE/CAA Journal of Automatica Sinica*, vol. 6, págs. 1452-1461, 2024. DOI: <https://doi.org/10.1109/JAS.2019.1911768>.
- [29] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. "O'Reilly Media, Inc.", 2012.
- [30] A. Navlani, A. Fandango e I. Idris, *Python Data Analysis: Perform data collection, data processing, wrangling, visualization, and model building using Python*. Packt Publishing, 2021, ISBN: 9781789953459. dirección: <https://books.google.com.co/books?id=DN4SEAAAQBAJ>.
- [31] R. Olu-Ajayi, H. Alaka, I. Sulaimon, F. Sunmola y S. Ajayi, «Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques,» *Journal of Building Engineering*, vol. 45, pág. 103 406, 2022, ISSN: 2352-7102. DOI: <https://doi.org/10.1016/j.jobe.2021.103406>.

- [32] Y. Ma, J.-q. Yu, C.-y. Yang y L. Wang, «Study on Power Energy Consumption Model for Large-Scale Public Building,» en *2010 2nd International Workshop on Intelligent Systems and Applications*, 2010, págs. 1-4. DOI: 10.1109/IWISA.2010.5473608.
- [33] Statsmodels. «ARIMA.» (2024), dirección: <https://www.statsmodels.org/stable/generated/statsmodels.tsa.arima.model.ARIMA.html> (visitado 2024).
- [34] K. K, M. R, B. W, G.-D. K y K. E, «Prediction of Energy Consumption on Example of Heterogenic Commercial Buildings,» *Energies*, vol. 17, 13:3220, 2024. DOI: <https://doi.org/10.3390/en17133220>.
- [35] M. Zhang, M.-A. Millar, S. Chen, Y. Ren, Z. Yu y J. Yu, «Enhancing hourly heat demand prediction through artificial neural networks: A national level case study,» *Energy and AI*, vol. 15, pág. 100 315, 2024. DOI: <https://doi.org/10.1016/j.egyai.2023.100315>.
- [36] M. M. Attia, B. Elboshy, A. S. Mohamed, M. Eita y K. E. Shakra, «Improving the energy efficiency performance of residential buildings in hot arid regions using lightweight concrete bricks,» *Case Studies in Construction Materials*, vol. 21, e03474, 2024. DOI: <https://doi.org/10.1016/j.cscm.2024.e03474>.
- [37] S. Havaeji, P. Ghanizadeh Anganeh y T. Esfahani, «A comparative analysis of machine learning techniques for building cooling load prediction,» *Journal of Building Pathology and Rehabilitation*, vol. 9, pág. 119, 2024. DOI: <https://doi.org/10.1007/s41024-024-00466-8>.
- [38] J. Cárdenas-Rangel, J. Jaramillo-Ibarra y G. Osmá-Pinto, «Restrictions and alternatives for developing data-based energy prediction models in buildings located in tropical climate: Literature review,» *Building and Environment*, vol. 262, pág. 111 786, 2024. DOI: <https://doi.org/10.1016/j.buildenv.2024.111786>.
- [39] H. Liu, J. Liang, Y. Liu y H. Wu, «A Review of Data-Driven Building Energy Prediction,» *Buildings*, vol. 13, n.º 2, 2023, ISSN: 2075-5309. DOI: 10.3390/buildings13020532. dirección: <https://www.mdpi.com/2075-5309/13/2/532>.

- [40] Y. Jiaqi, C. Wen-Shao y D. Yu, «Building Energy Prediction Models and Related Uncertainties: A Review,» *Buildings*, vol. 12, pág. 1284, 2022. DOI: <https://www.mdpi.com/2075-5309/12/8/1284>.

Anexo A

Anexo A. Diagramas unifilares y fotos de los tableros monitorizados

Este anexo muestra los diagramas unifilares y las fotos de los dos tableros monitorizados por los medidores de energía. La Figura 27 muestra la foto del tablero de aires acondicionados ubicado en el CCTV (quinto piso) y la Figura 28 el diagrama unifilar del tablero de aires acondicionados junto con la ubicación de los CTs del medidor de energía.

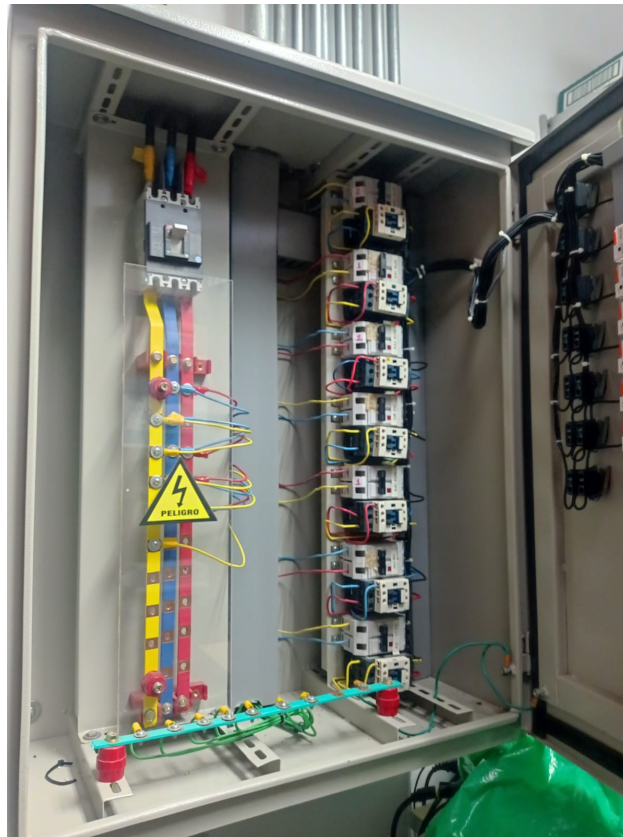


Figura 27. *Tablero de aires acondicionados.*

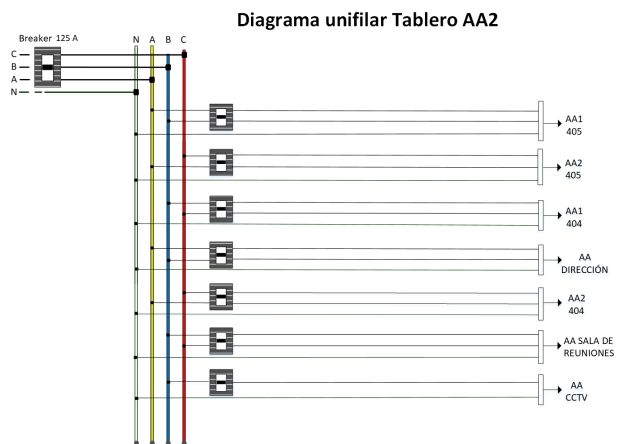


Figura 28. Diagrama unifilar tablero AA.

La Figura 4 muestra los tableros de circuitos del cuarto piso del edificio.

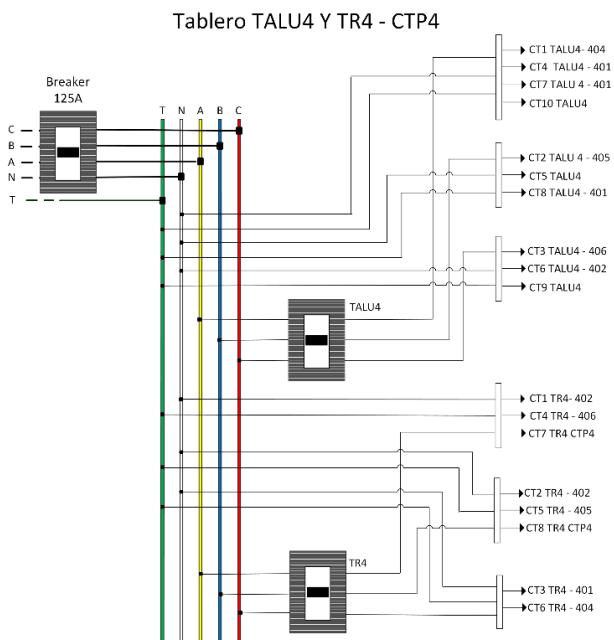


(a) Subtablero de tomacorrientes normales.

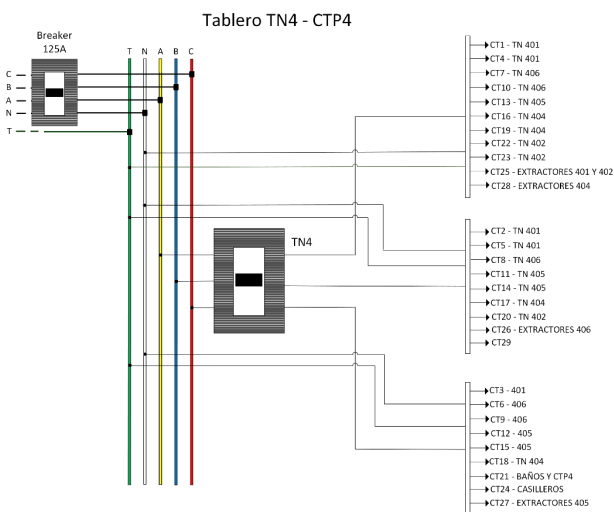
(b) Subtablero de iluminación.

(c) Subtablero de tomacorrientes regulados.

Figura 29. Subtableros del cuarto piso del Edificio de Ingeniería Eléctrica.



(a) Diagrama unifilar circuitos del tablero de iluminación y cargas reguladas piso 4.



(b) Diagrama unifilar circuitos del tablero de tomacorrientes normales piso 4.

Figura 4. Diagrama unifilar TGBT piso 4. (A) Circuitos de iluminación y tomacorrientes regulados. (B) Circuitos de tomacorrientes normales.

Anexo B

Anexo B. Funciones de Python para la creación de la base de datos

B.0.1. Cálculo consumo de energía

```
def calculo_consumo_energia(Dataframe, Dataframe1, start_time, end_time,
                             time_interval):

    #El archivo Dataframe es el archivo del medidor ACUREV2020 de AA
    #El archivo Dataframe1 es el archivo del medidor ACUREV2100 de
    circuitos del aula

    # Convertir la columna 'Fecha' al formato de fecha de Pandas
    Dataframe['Date'] = pd.to_datetime(Dataframe['Date'],
                                       format='%d/%m/%Y %H:%M')

    # Filtrar el DataFrame para incluir solo datos dentro del rango de
    #tiempo
    Dataframe= Dataframe[Dataframe['Date'].isin(pd.date_range(start=
        start_time, end=end_time, freq=time_interval))]

    #Las potencias están en kW
    #Se suman los canales de los aires acondicionados de cada aula
    ARCHIVO PLATAFORMA MEDIDORES
    Dataframe.loc[:, 'Potencia_AA_404'] = Dataframe.loc[:, 'P3 (channel 3
        real power)'] + Dataframe.loc[:, 'P5 (channel 5 real power)']
```

```

+Dataframe.loc[:, 'P8 (channel 8 real power) ']+Dataframe.loc[:, 'P9
    (channel 9 real power) ']
Dataframe.loc[:, 'Potencia_AA_405'] = Dataframe.loc[:, 'P1 (channel 1
    real power) ']+Dataframe.loc[:, 'P2 (channel 2 real power) ']+
+Dataframe.loc[:, 'P4 (channel 4 real power) ']+Dataframe.loc[:, 'P7
    (channel 7 real power) ']

#Las potencias est n en kW
# #Se suman los canales de los aires acondicionados de cada aula
    ARCHIVO medidor
# Dataframe.loc[:, 'Potencia_AA_404'] = Dataframe.loc[:, 'R2> Watt A
    (CH103) ']+Dataframe.loc[:, 'R2> Watt A (CH105) ']+
# +Dataframe.loc[:, 'R2> Watt A (CH108) ']+Dataframe.loc[:, 'R2> Watt
    A (CH109) ']+
# Dataframe.loc[:, 'Potencia_AA_405'] = Dataframe.loc[:, 'R2> Watt A
    (CH101) ']+Dataframe.loc[:, 'R2> Watt A (CH102) ']+
# +Dataframe.loc[:, 'R2> Watt A (CH104) ']+Dataframe.loc[:, 'R2> Watt
    A (CH107) ']+

# #Las potencias est n en kW
# #Se suman los canales de cada circuito por aula
Dataframe1.loc[:, 'Potencia_circuitos_404'] = Dataframe1.loc[:, '
    P001_kW ']+Dataframe1.loc[:, 'P004_kW ']+Dataframe1.loc[:, 'P010_kW '
    ]
+Dataframe1.loc[:, 'P011_kW ']+Dataframe1.loc[:, 'P012_kW ']+Dataframe1
    .loc[:, 'P013_kW ']+Dataframe1.loc[:, 'P015_kW ']+
Dataframe1.loc[:, 'Potencia_circuitos_405'] = Dataframe1.loc[:, '
    P002_kW ']+Dataframe1.loc[:, 'P003_kW ']+Dataframe1.loc[:, 'P005_kW '
    ]

```

```

+Dataframe1.loc[:, 'P006_kW'] + Dataframe1.loc[:, 'P007_kW'] + Dataframe1
    .loc[:, 'P008_kW'] + Dataframe1.loc[:, 'P009_kW'] + Dataframe1.loc[:, '
    P014_kW']

df=pd.DataFrame()
#Suma de potencias y calculo de la total. Calculo de consumo y la
    energ a queda en Wh
df["Potencia 404"] = Dataframe['Potencia_AA_404'] + Dataframe1['
    Potencia_circuitos_404']
df["Consumo 404"] = df["Potencia 404"] * (5/60) * 1000
df["Potencia 405"] = Dataframe['Potencia_AA_405'] + Dataframe1['
    Potencia_circuitos_405']
df["Consumo 405"] = df["Potencia 405"] * (5/60) * 1000

df["Date"] = Dataframe["Date"]
df = df[df['Date'].isin(pd.date_range(start=start_time, end=
    end_time, freq=time_interval))]
df.set_index("Date", inplace=True)
return df

```

B.0.2. Función estación meteorológica

```

def estacion(archivo, start_time, end_time, time_interval):

    archivo['Date_Time'] = archivo['Date'] + ' ' + archivo['Time']
    print(archivo['Date_Time'])
    archivo['Date_Time'] = archivo['Date_Time'].str.replace('a. m.', '
        AM').str.replace('p. m.', 'PM').str.replace('a. m .', 'AM').str.
        replace('p. m .', 'PM')

```

```
#archivo['Date_Time'] = archivo['Date_Time'].astype(str).apply(
    lambda x: re.sub(r'\b(a\. m\.|p\. m\.)\b', lambda m: m.group(1).
        upper(), x))
print(archivo['Date_Time'])

archivo["Date_Time"] = pd.to_datetime(archivo["Date_Time"], format=
    '%d/%m/%Y %I:%M %p')
#archivo["Date_Time"] = pd.to_datetime(archivo["Date_Time"],
    infer_datetime_format=True)
#print(archivo['Date_Time'])

archivo = archivo[(archivo['Date_Time'] >= start_time) & (archivo['
    Date_Time'] <= end_time)]

archivo.set_index("Date_Time", inplace=True)

# Reindexar el DataFrame para tener una serie de tiempo continua
archivo = archivo.resample('5T').asfreq()

# Llenar los valores faltantes utilizando el método de
    interpolación lineal
archivo = archivo.interpolate(method="linear")

# Restablecer el índice si es necesario
archivo.reset_index(inplace=True)

#Paso de variables de dirección de viento a datos numéricos (Se
    usaron los grados)
```

```

#La estaci n muestra (---) cuando en el intervalo de tiempo no hay
viento, le puse un valor de 360
archivo[ 'Dir' ] = archivo[ 'Dir' ].map({ 'N': 0, 'NNE': 22.5, 'NE': 45,
    'ENE':67.5, 'E':90, 'ESE':112.5, 'SE':135, 'SSE':157.5, 'S':180,
    'SSW':202.5, 'SW':225, 'WSW':247.5, 'W':270, 'WNW':292.5, 'NW'
    :315, 'NNW':337.5, '---': 360}).astype( float )

#Selecci n de variables de la estaci n
archivo = archivo[[ 'Date_Time', 'Temp Out', 'Out Hum', 'Dew Pt.', 'Wind
    Speed', 'Dir', 'HeatIndex', 'Bar', 'Rain', 'SolarRad.', 'UV Index', '
    CoolD-D' ]]

archivo.set_index("Date_Time", inplace=True)
return archivo

```

B.0.3. Función sensores Monnit 1

```

def sensores(archivo, lista_sensores, lista_sensores_string, start_time,
end_time, time_interval):

    date_range = pd.date_range(start=start_time, end=end_time, freq='5T
        ')
    Raw_data = archivo
    # Convertir la columna 'Fecha' al formato de fecha de Pandas
    Raw_data[ 'Date' ] = pd.to_datetime( Raw_data[ 'Date' ], format='%m/%d/%
        Y %I:%M %p' )

    # Crear una nueva columna 'Fecha_Formateada' con el formato deseado
    y omitiendo la hora
    Raw_data[ 'New_Date' ] = Raw_data[ 'Date' ].dt.strftime( '%d/%m/%Y' )

```

```
# Crear una nueva columna 'Hora' con la hora en formato HH:MM
Raw_data['Hour'] = Raw_data['Date'].dt.strftime('%I:%M %p')

#List_sensors = list(set(Raw_data.iloc[:,1]))
List_sensors_404 = lista_sensores
String_404 = lista_sensores_string
# List_sensors_404 = [1081112]

# ESTO DEBE AJUSTARSE EN EL CAMBIO DE SALON
Table_Data = pd.DataFrame(np.zeros((3536, len(String_404))))

count = 0
for list in List_sensors_404:
    print(list)
    Sensor_Data = Raw_data[Raw_data['SensorID'] == list]

# Establece la columna de tiempo como ndice
Sensor_Data.set_index('Date', inplace=True)

# Resampla los datos a intervalos de cinco minutos y utiliza
    la ltima observaci n
df_resampled = Sensor_Data.resample(time_interval).last()

# Reindexa el DataFrame para incluir todas las fechas en el
    rango
df_resampled = df_resampled.reindex(date_range)
```

```
df_resampled = df_resampled.loc[(df_resampled.index >=
    start_time) & (df_resampled.index <= end_time)]
print(df_resampled.iloc[:,0])

# Usa forward fill para propagar la última observación a los
# intervalos vacíos
df_resampled.ffill(inplace=True)

df_resampled.bfill(inplace=True)

if df_resampled["Sensor Name"][10].split("-")[0] != "Humidity "
:
    Table_Data.iloc[:,count] = df_resampled["Value"].values
    count += 1

else:
    patron = r'@ (\d+(\.\d+)?)'
    Table_Data.iloc[:,count] = df_resampled["Value"].values

    for y in range(len(Table_Data.iloc[:,1])):
        Table_Data.iloc[y,count+1] = float(re.findall(patron,
            df_resampled["Formatted Value"][y])[0][0])

    count += 2

# ESTO DEBE AJUSTARSE EN EL CAMBIO DE SALON
if list == 1081131:
    Table_Data.iloc[:,len(String_404)-1] = df_resampled.index
```

```

Table_Data.columns = String_404
Table_Data.set_index("Date", inplace=True)
return Table_Data

```

B.0.4. Función sensores Monnit 2

```

def Sensors_Modified(Table):
    ### 404
    # df = pd.DataFrame()
    # df['Door_Status'] = Table['1081112-OD']
    # df['Windows_Status'] = Table['1081121-OW']+Table["1081117-OW"]+
        Table["1081111-OW"]+Table["1081115-OW"]+Table["1081113-OW"]+
        Table["1081110-OW"]
    # df['Average_Temp'] = (Table["1080498-T"]+Table["700870-T"]+Table
        ["700860-T"]+Table["1056911-T"])/4
    # df['Average_Humi'] = (Table["1080498-H"]+Table["700870-H"]+Table
        ["700860-H"]+Table["1056911-H"])/4
    # df['suma'] = Table["1081130-PIR"] + Table["1081127-PIR"] + Table
        ["1081132-PIR"] + Table["1081128-PIR"]
    # df['PIR'] = 0
    # df.loc[df['suma'] != 0, 'PIR'] = 1
    # df.drop(columns=['suma'], inplace=True)

    ### 405
    df = pd.DataFrame()
    df['Door_Status'] = Table['1081116-OD']
    df['Windows_Status'] = Table['1081120-OW']+Table["1081123-OW"]+
        Table["1081109-OW"]+Table["1081114-OW"]+Table["1081118-OW"]+
        Table["1081122-OW"]

```

```

df[ 'Average_Temp' ] = ( Table[ "1056923-T" ]+ Table[ "700868-T" ]+ Table[ "
700871-T" ]+ Table[ "700867-T" ] )/4
df[ 'Average_Humi' ] = ( Table[ "1056923-H" ]+ Table[ "700868-H" ]+ Table[ "
700871-H" ]+ Table[ "700867-H" ] )/4
df[ 'suma' ] = Table[ "447634-PIR" ] + Table[ "1081129-PIR" ] + Table[ "
1081131-PIR" ]
df[ 'PIR' ] = 0
df.loc[ df[ 'suma' ] != 0, 'PIR' ] = 1
df.drop( columns=[ 'suma' ], inplace=True)

```

B.0.5. Función cámaras

```

def camera_data( archivo , start_time , end_time ):
    date_range = pd.date_range( start=start_time , end=end_time , freq='5T
    ')
    archivo[ 'Date' ] = pd.to_datetime( archivo[ 'Date' ] )
    archivo.set_index( 'Date' , inplace=True)
    archivo = archivo.loc[ ( archivo.index >= start_time ) & ( archivo.
    index <= end_time ) ]
    archivo.reset_index( drop=True , inplace=True)
    archivo.set_index( date_range , inplace=True)
    return archivo

```

B.0.6. Función final (concatena todas las variables)

```

def Final_Table_404( Table_consumo , weather_station , Sensors_404 , Cameras ):
    df_concat = pd.concat( [ Table_consumo[ 'Potencia 405' ] ,
    Table_consumo[ 'Consumo 405' ] , weather_station , Sensors_404 , Cameras ] ,
    axis=1)
    return df_concat

```