



Trabajo de Fin de Máster

Máster Universitario en Ciberseguridad e Inteligencia de Datos

Evaluación de métodos de asignación de coeficientes de reparto en autoconsumos colectivos

*Evaluation of allocation coefficient methods in collective
self-consumption*

Paula Pérez Romero

La Laguna, 4 de julio de 2024

D. **Ricardo Luis Guerrero Lemus**, con N.I.F. 43358363M, Director del Máster de Energías Renovables, Prof. Dr. del Departamento de Física de la Universidad de La Laguna y Director de EnergyRIS, como cotutor.

D. **Carlos Javier Pérez González**, con N.I.F. 45452719G, miembro del consejo asesor de la Cátedra de Big Data, Open Data y Blockchain y Prof. Dr. del Departamento de Matemáticas, Estadística e Investigación Operativa de la Universidad de La Laguna, como cotutor.

C E R T I F I C A N

Que la presente memoria titulada:

"Evaluación de métodos de asignación de coeficientes de reparto en autoconsumos colectivos"

ha sido realizada bajo su dirección por Dña. **Paula Pérez Romero**, con N.I.F. 24451540X.

Y para que así conste, en cumplimiento de la legislación vigente y a los efectos oportunos firman la presente en La Laguna a 4 de julio de 2024

Agradecimientos

En primer lugar, quiero agradecer a mi familia su apoyo incondicional y su confianza.

A mis cotutores, Ricardo Guerrero y Carlos Pérez, les debo una inmensa gratitud por su tiempo.

A David Cañadillas, gracias por tus ideas innovadoras y tu ayuda en los momentos más cruciales.

Finalmente, quiero agradecer a EnergyRIS por compartir generosamente sus datos y brindarme la oportunidad de que este trabajo tenga una aplicación real.

Licencia

© Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.

Resumen

Este Trabajo de Fin de Máster tiene como objetivo principal identificar y comparar la actuación de diferentes métodos de asignación de coeficientes de reparto de energía entre los participantes de un autoconsumo colectivo. El propósito es determinar cuáles de estos métodos son más efectivos para garantizar un reparto 'justo' y beneficioso para todos los participantes, con un enfoque en minimizar los excedentes anuales de energía y optimizar el uso de los recursos generados.

Para este estudio, se utilizarán tanto datos reales como simulados de demanda energética, complementados con información proveniente de fuentes como el Photovoltaic Geographical Information System (PVGIS). Este enfoque dual permitirá una evaluación exhaustiva y precisa de los distintos métodos de reparto.

El análisis inicial considera la regulación vigente establecida por el Real Decreto 244/2019, proporcionando una base comparativa sólida. Adicionalmente, se explorarán propuestas innovadoras diseñadas para mejorar los coeficientes de reparto actuales de energía autoconsumida, subrayando su relevancia y el potencial de optimización que ofrecen.

Este estudio busca no solo evaluar la efectividad de los métodos actuales, sino también proponer mejoras significativas que puedan ser implementadas para beneficiar a los sistemas de autoconsumo colectivo, garantizando una distribución de energía más equitativa y eficiente.

Palabras clave: *Autoconsumo Colectivo, Coeficientes de Reparto, Minimización de Excedentes Energéticos, Sistemas Fotovoltaicos, Distribución de Energía, Real Decreto 244/2019.*

Abstract

This Master's Thesis aims to identify and compare the performance of different methods for allocating energy distribution coefficients among participants in a collective self-consumption system. The goal is to determine which of these methods are most effective in ensuring a 'fair' and beneficial distribution for all participants, focusing on minimizing annual energy surpluses and optimizing the use of generated resources.

This study will utilize both real and simulated energy demand data, supplemented with information from sources such as the Photovoltaic Geographical Information System (PVGIS). This dual approach will allow for a comprehensive and accurate evaluation of the different distribution methods.

The initial analysis considers the current regulations established by Royal Decree 244/2019, providing a solid comparative foundation. Additionally, innovative proposals designed to improve the current coefficients of self-consumed energy distribution will be explored, highlighting their relevance and optimization potential.

This study aims not only to evaluate the effectiveness of current methods but also to propose significant improvements that can be implemented to benefit collective self-consumption systems, ensuring a more equitable and efficient distribution of energy.

Keywords: *Collective Self-Consumption, Allocation Coefficients, Energy Surplus Minimization, Photovoltaic Systems, Energy Distribution, Royal Decree 244/2019.*

Índice general

Índice de Figuras	ii
Índice de Tablas	iv
Glosario	v
1. Introducción y objetivos	1
1.1. Introducción	1
1.2. Objetivos	3
2. Análisis de Datos: Preparación y Exploración	6
2.1. Datos de generación fotovoltaica	6
2.2. Datos de demanda	9
3. Metodos y técnicas	18
3.1. Cálculo de los coeficientes de reparto	19
3.2. Preparación de los datos para las predicciones	22
3.3. Modelo de predicción	23
4. Obtención de resultados del reparto energético	30
4.1. Cálculo de indicadores de reparto y excedentes	30
4.2. Resultados del reparto y excedentes	32
4.3. Resultados globales	32
4.4. Resultados por CUPS	34
5. Discusión, conclusiones y líneas futuras	37
5.1. Discusión	37
5.2. Conclusiones	38
5.3. Líneas Futuras	39
6. Summary and conclusions	40
6.1. Summary	40
6.2. Conclusions	40
A. Archivos, tablas y gráficas de los resultados	42
A.1. Enlaces a los archivos con datos	42
A.2. Análisis de los resultados de las predicciones	42

Índice de figuras

1.1. Cambios estructurales del sistema eléctrico consecuencia del prosumismo .	2
1.2. Gráfico de tendencias de búsqueda del término ' <i>self consumption</i> ' en Google	3
1.3. Esquema de las diez modalidades de reparto que se analizarán en este estudio	4
1.4. Esquema del autoconsumo colectivo simulado	5
1.5. Logo de EnergyRIS.	5
2.1. Fotografía de la planta fotovoltaica del SEGAI	6
2.2. Interfaz de PVGIS cuando accedes por primera vez a la aplicación web, antes de hacer nada.	7
2.3. Generación fotovoltaica horaria total (2021-2024)	8
2.4. Gráficas de la generación fotovoltaica mensual total (2021-2024)	8
2.5. Distribución de la demanda anual total por CUPS (2021-2024)	10
2.6. Fotografía del edificio del SEGAI	11
2.7. Demanda horaria media por año del CUPS SEGAI	11
2.8. Demanda horaria media por día de la semana del CUPS SEGAI	11
2.9. Demanda mensual por año del CUPS SEGAI	12
2.10 Fotografía de la Facultad de BBAA.	12
2.11 Demanda horaria media por año del CUPS BBAA	13
2.12 Demanda horaria media por día de la semana del CUPS BBAA	13
2.13 Demanda mensual por año del CUPS BBAA	13
2.14 Fotografía del C.M San Fernando	14
2.15 Demanda horaria media por año del CUPS C.M. San Fernando	14
2.16 Demanda horaria media por día de la semana del CUPS C.M. San Fernando	15
2.17 Demanda mensual por año del CUPS C.M. San Fernando	15
2.18 Fotografía del edificio de la Biblioteca General.	16
2.19 Demanda horaria media por año del CUPS Biblioteca General	16
2.20 Demanda horaria media por día de la semana del CUPS Biblioteca General .	16
2.21 Demanda mensual por año del CUPS Biblioteca General	17
3.1. Esquema de la metodología de trabajo del capítulo 3	18
3.2. Esquema del funcionamiento algoritmo de ' <i>XGBoost</i> '	24
3.3. Visualización del comportamiento de <i>TimeSeriesSplit</i>	26
3.4. Esquema del proceso de entrenamiento y predicción anual	27
3.5. Comparación de la distribución horaria de la demanda real de BBAA y las predicciones.	28
3.6. Comparación de la distribución mensual de la demanda real de BBAA y las predicciones.	28
3.7. Comparación de la distribución horaria por día de la semana de la demanda real de BBAA y las predicciones.	29

4.1. Excedentes evitables en el año 2022	33
4.2. Excedentes evitables en el año 2023	33
4.3. Excedentes evitables en el año 2024	34
4.4. Excedentes totales por modalidad de reparto y CUPS para el año 2022 . . .	35
4.5. Excedentes totales por modalidad de reparto y CUPS para el año 2023 . . .	35
4.6. Excedentes totales por modalidad de reparto y CUPS para el año 2024 . . .	36
A.1. Comparación de la distribución horaria de la demanda real de BBAA y las predicciones.	44
A.2. Comparación de la distribución mensual de la demanda real de BBAA y las predicciones.	44
A.3. Comparación de la distribución horaria por día de la semana de la demanda real de BBAA y las predicciones.	45
A.4. Comparación de la distribución horaria de la demanda real de SEGAI y las predicciones.	45
A.5. Comparación de la distribución mensual de la demanda real de SEGAI y las predicciones.	46
A.6. Comparación de la distribución horaria por día de la semana de la demanda real de SEGAI y las predicciones.	46
A.7. Comparación de la distribución horaria de la demanda real de C.M. San Fernando y las predicciones.	47
A.8. Comparación de la distribución mensual de la demanda real de C.M. San Fernando y las predicciones.	47
A.9. Comparación de la distribución horaria por día de la semana de la demanda real de C.M. San Fernando y las predicciones.	48
A.10 Comparación de la distribución horaria de la demanda real de Biblioteca y las predicciones.	48
A.11 Comparación de la distribución mensual de la demanda real de la biblioteca y las predicciones.	49
A.12 Comparación de la distribución horaria por día de la semana de la demanda real de la Biblioteca y las predicciones.	49

Índice de tablas

2.1. Resumen anual de la demanda global	9
3.1. Resultados de las métricas de evaluación del modelo de predicción	27
4.1. Nomenclatura para las modalidades de reparto	32
4.2. Excedentes totales y evitables por año [MWh]	33
4.3. Mejor modalidad de reparto y sus excedentes por CUPS y año	35
A.1. Resultados de la validación cruzada (MAE)	43
A.2. Mejores parámetros para cada modelo	43

Glosario

Autoconsumo De acuerdo con lo previsto en el artículo 9.1 de la Ley 24/2013, de 26 de diciembre (1), se entenderá por autoconsumo, el consumo por parte de uno o varios consumidores de energía eléctrica proveniente de instalaciones de producción próximas a las de consumo y asociadas a los mismos.

Autoconsumo colectivo Según lo estipulado en el RD 244/2019 (2), un sujeto consumidor participa en un autoconsumo colectivo cuando pertenece a un grupo de varios consumidores que se alimentan, de forma acordada, de energía eléctrica que proviene de instalaciones de producción próximas a las de consumo y asociadas a los mismos.

Prosumidor Consumidor de energía eléctrica que al mismo tiempo participa en la producción del misma.

Perfil de consumo Distribución horaria del consumo de energía de un consumidor a lo largo del día, mostrando cuándo y cuánto se consume en cada hora.

CUPS (Código Universal de Punto de Suministro) Es un identificador único asignado a cada punto de suministro eléctrico en España. Este código es esencial para la gestión y seguimiento del suministro eléctrico y es utilizado tanto por las compañías eléctricas como por los usuarios para diversas gestiones.

Coefficiente de reparto de la energía generada Proporción de la energía producida por una instalación de generación compartida (como paneles solares) que se asigna a cada uno de los miembros del colectivo, basado en criterios preestablecidos como la inversión, el consumo o acuerdos contractuales.

Energía consumida Cantidad total de energía eléctrica utilizada por un consumidor o un grupo de consumidores en un periodo determinado. Esta energía puede provenir tanto de la red eléctrica como de fuentes de generación propia.

Excedentes Energía eléctrica producida por una instalación de generación que supera el consumo de los usuarios asociados, y que puede ser inyectada a la red o almacenada para uso futuro.

Excedentes evitables Parte de la energía excedentaria que puede ser gestionada o minimizada mediante ajustes en la producción o consumo.

Excedentes inevitables Parte de la energía excedentaria que no puede ser gestionada ni minimizada.

Balance neto Diferencia entre la energía producida y la energía consumida en un periodo determinado para un consumidor o un grupo de consumidores, considerado para determinar la cantidad de energía que se puede considerar como excedente o déficit.

Capítulo 1

Introducción y objetivos

1.1. Introducción

En el contexto actual de transición energética, la reconversión de las fuentes de energía hacia alternativas más sostenibles se ha consolidado como una prioridad global. La producción de energía de origen renovable se presenta como una de las soluciones más viables frente al agotamiento de las reservas de combustibles fósiles y la necesidad urgente de reducir las emisiones de carbono. Dos factores clave han impulsado esta tendencia: la disminución de la dependencia energética externa¹ y la creciente preocupación medioambiental.

En primer lugar, la expansión de las energías renovables ha permitido a numerosos países **reducir su dependencia de la importación de combustibles fósiles**, aprovechando recursos locales para el autoabastecimiento energético. En segundo lugar, la creciente sensibilización medioambiental y las políticas internacionales han impulsado la adopción de objetivos de sostenibilidad que buscan un **desarrollo más ecológico y autosuficiente**. Entre las medidas destacadas se incluyen la reducción de emisiones de gases de efecto invernadero, la disminución de residuos tóxicos y radioactivos, y la mejora de la eficiencia energética.

La generación distribuida ha emergido como una estrategia clave para maximizar el impacto de las energías renovables a todos los niveles. **En este contexto, la creación de comunidades energéticas, que promueven el autoconsumo y el prosumismo, ha ganado relevancia.** El prosumidor, un agente que no solo consume sino también produce energía, se ha convertido en una figura crucial en el mercado eléctrico, facilitando la descentralización y democratización de la energía. En la última década, los costos de las tecnologías fotovoltaicas han disminuido significativamente, y junto con las políticas favorables al autoconsumo, han mejorado la rentabilidad del autoconsumo doméstico. Sin embargo, la distribución de la energía generada en modelos de autoconsumo colectivo aún presenta desafíos, afectando negativamente el rendimiento económico de estas instalaciones .

¹Según (3) en 2021, la UE importó de Rusia más del 40 % de su consumo total de gas, así como un 27 % y un 46 % de sus importaciones de petróleo y carbón, respectivamente.

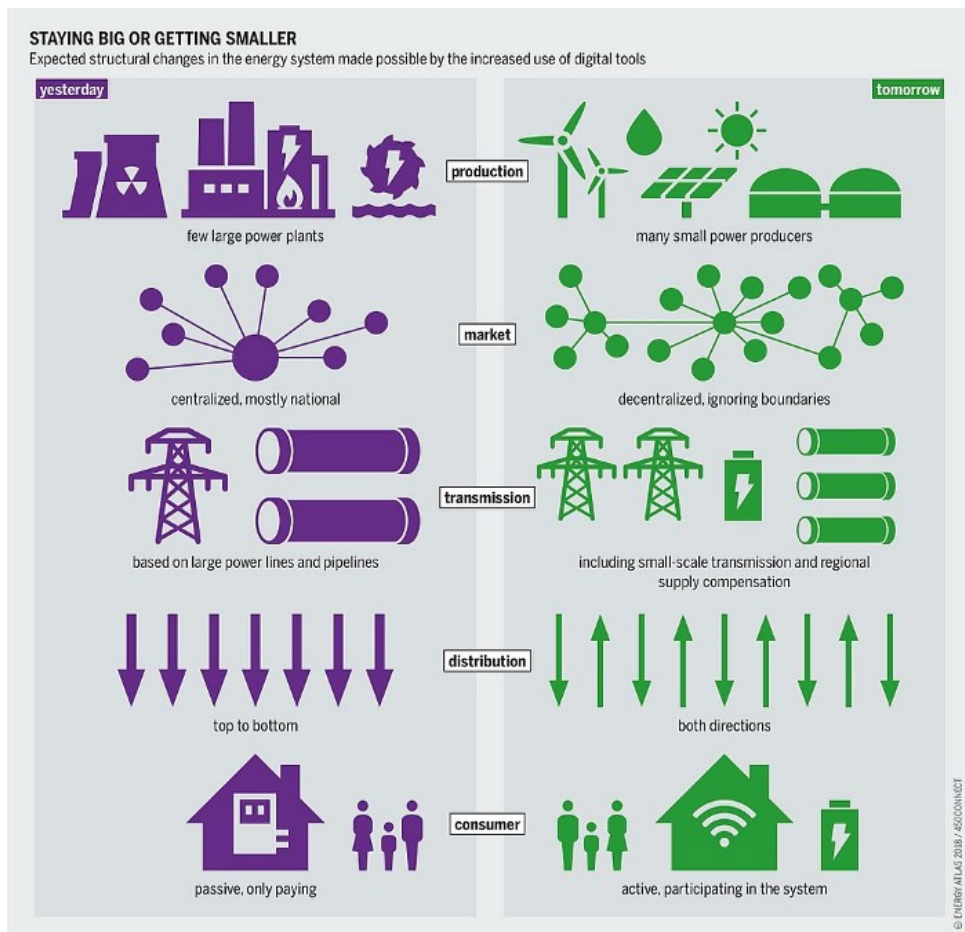


Figura 1.1: Representación esquemática de los cambios estructurales del sistema eléctrico consecuencia del prosumismo. Imagen de Green New Deal. **Fuente:** The Smart City Journal (4).

En Europa, el paquete legislativo *Clean Energy for All Europeans* (5), adoptado en 2019 y diseñado para cumplir con los compromisos del Acuerdo de París² ha sido un motor clave en la promoción de energías renovables y la reducción de emisiones. Esta transformación del sistema energético incluye la descarbonización, el desarrollo de almacenamiento a gran escala, la electrificación del transporte y la implementación de redes inteligentes y generación distribuida. En esta línea, **La Directiva de Energías Renovables II** (*Renewable Energy Directive II*, REDII) **ha proporcionado un marco normativo para impulsar el autoconsumo colectivo y las comunidades energéticas**, aunque la implementación varía significativamente entre los estados miembros. Estas variaciones abarcan aspectos técnicos y económicos, desde los requisitos de proximidad de las instalaciones hasta los regímenes de retribución de excedentes energéticos. Este hecho refleja una falta de armonización que podría optimizarse.

²El Acuerdo de París es un acuerdo dentro del marco de la Convención Marco de las Naciones Unidas sobre el Cambio Climático que establece medidas para la reducción de las emisiones de gases de efecto invernadero (GEI). El acuerdo busca mantener el aumento de la temperatura global promedio por debajo de los 2 °C por encima de los niveles preindustriales, y perseguir esfuerzos para limitar el aumento a 1.5 °C, reconociendo que esto reduciría significativamente los riesgos y efectos del cambio climático. La aplicabilidad del Acuerdo comenzó en 2020, tras la finalización de la vigencia del Protocolo de Kioto. El acuerdo fue negociado durante la XXI Conferencia sobre Cambio Climático (COP 21) por los 195 países miembros (6).

En España, el Real Decreto 244/2019 (2) regula el autoconsumo colectivo, estableciendo un marco para la creación de comunidades energéticas. En este proyecto se explorarán diversas metodologías para maximizar el rendimiento energético, considerando las restricciones establecidas por la normativa vigente mencionada y proponiendo posibles alternativas más allá de estas regulaciones.

El incremento de búsquedas en Google del término 'autoconsumo' refleja el creciente interés y relevancia de estos temas en la sociedad actual. La siguiente gráfica ilustra el aumento de la relevancia de estos términos en la última década, tanto a nivel global, evidenciando una tendencia al alza en la atención y adopción de soluciones energéticas sostenibles.



Figura 1.2: **Fuente:** Google Trends (7).

Este auge en la relevancia de estos conceptos no solo refleja una mayor conciencia y preocupación por el medio ambiente, sino también un cambio hacia modelos energéticos más sostenibles y autosuficientes, que buscan reducir la dependencia de los combustibles fósiles y minimizar el impacto ambiental. **La integración de energías renovables y la participación activa de los ciudadanos en la generación de energía son pilares fundamentales para alcanzar los objetivos de sostenibilidad y enfrentar los desafíos del cambio climático.**

1.2. Objetivos

Este estudio tiene como objetivo principal **identificar y comparar la actuación de diferentes métodos de asignación de coeficientes de reparto de energía entre los participantes de un autoconsumo colectivo.** Además de los métodos contemplados por la legislación vigente (RD244/2019), se han considerado para el análisis un método de creación propia, y otro propuesto en (8). Nos referiremos a ellos como 'tradicionales', 'reparto justo' y 'correlaciones de Pearson', respectivamente. Los dos primeros métodos

mencionados se calcularán de cuatro formas distintas mientras que tercero solo de dos ³, resultando en un total de **diez modalidades de reparto que serán analizadas** 1.3.

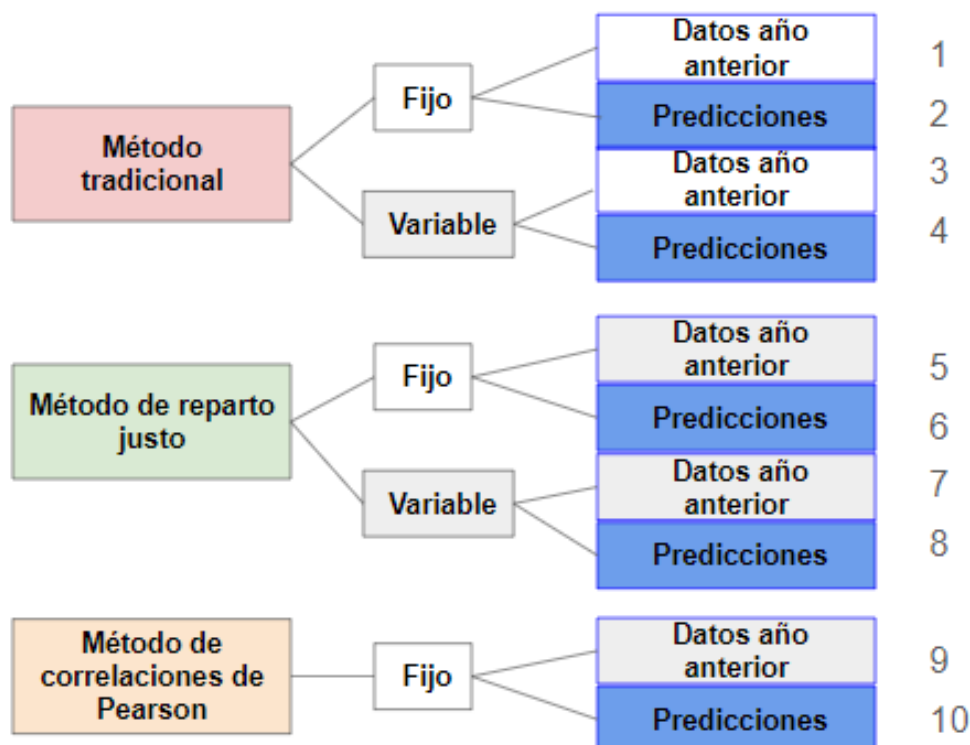


Figura 1.3: Esquema de las diez modalidades de reparto que se analizarán en este estudio. **Fuente:** Elaboración propia.

Partimos de dos hipótesis iniciales. **La primera hipótesis es que los coeficientes variables, ajustados a la demanda horaria, podrían ofrecer mejores resultados en comparación con los coeficientes fijos.** Esto se debe a la capacidad de los coeficientes variables para adaptarse de manera más eficiente a los patrones de la demanda energética. **La segunda hipótesis es que los métodos que utilizan predicciones se presuponen superiores a los métodos calculados con datos del año anterior.** Esto se debe a que las predicciones podrían adaptarse mejor a las variaciones actuales y futuras en la producción y el consumo de energía.

Para analizar estas hipótesis, se evaluará el desempeño de los diez métodos mencionados en un autoconsumo colectivo concreto. **Minimizar los excedentes anuales de energía para optimizar el uso de los recursos generados será el criterio principal de evaluación.**

El estudio se llevará a cabo en una comunidad de autoconsumo colectivo formada por cuatro participantes y siete plantas de generación fotovoltaica agregadas que actuarán como una única fuente de generación. Es importante mencionar que aunque los datos utilizados son reales, el autoconsumo colectivo no existe en la realidad y ha sido simulado para este trabajo.

Los cuatro participantes serán cuatro CUPS que pertenecen a la Universidad de La

³El procedimiento de cálculo será detallado en el capítulo 3.

Laguna: la Facultad de Bellas Artes, el Edificio del SEGAI (Servicios Generales de Apoyo a la Investigación), la Biblioteca General y el Colegio Mayor San Fernando. Para la generación fotovoltaica agregada, se han seleccionado siete de las plantas fotovoltaicas también de la Universidad de La Laguna: Aparejadores, Biología, Bellas Artes, Derecho, Estabulario, Periodismo y SEGAI.

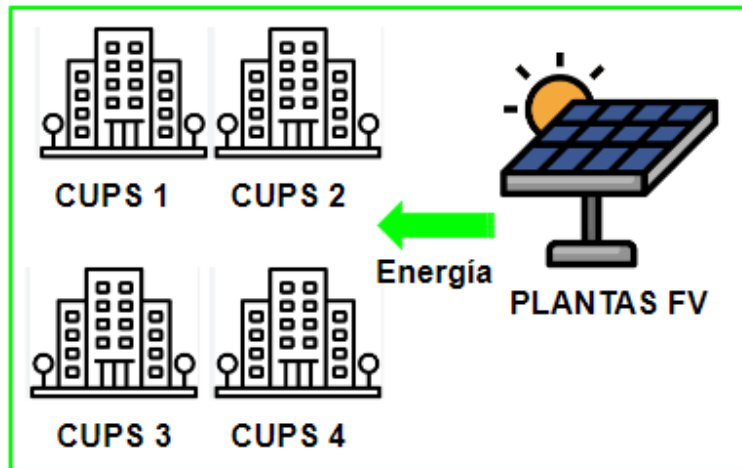


Figura 1.4: Esquema del autoconsumo colectivo simulado. **Fuente:** Elaboración propia.

El acceso a los datos de demanda y los datos de las plantas fotovoltaicas ha sido facilitado por EnergyRIS, la spin-off de la Universidad de La Laguna dirigida por el Prof. Dr. Ricardo Guerrero, cotutor de este TFM.



Figura 1.5: Logo de EnergyRIS.

Capítulo 2

Análisis de Datos: Preparación y Exploración

En este capítulo se detalla el proceso de obtención, preparación y exploración de los datos utilizados en este estudio. La primera sección está dedicada a los datos de generación fotovoltaica y la segunda a los datos de demanda de los cuatro CUPS del autoconsumo colectivo: la Facultad de Bellas Artes, el Edificio del SEGAI, la Biblioteca General y el Colegio Mayor San Fernando.

2.1. Datos de generación fotovoltaica

Los datos de generación fotovoltaica utilizados en este estudio se han obtenido mediante la API de PVGIS (Photovoltaic Geographical Information System). PVGIS ¹ es una herramienta desarrollada por la Comisión Europea que proporciona datos y mapas sobre el rendimiento de los sistemas fotovoltaicos en todo el mundo.



Figura 2.1: Fotografía de la planta fotovoltaica del SEGAI. Fuente: Portal de la Universidad de La Laguna

Para la obtención de los datos, se ha empleado un código de Python con técnicas de **web**

¹Para más información sobre la API de PVGIS, consulte: https://joint-research-centre.ec.europa.eu/photovoltaic-geographical-information-system-pvgis/getting-started-pvgis/api-non-interactive-service_en

scraping² para extraer los datos horarios de generación fotovoltaica desde PVGIS. La función que permite realizar solicitudes a la API requiere algunos parámetros específicos. Estos parámetros incluyen:

- Latitud y longitud: coordenadas geográficas de cada planta fotovoltaica.
- Año de inicio y fin: periodo para el cual se desean obtener los datos horarios.
- Inclinación y azimut de la Superficie: ángulos que definen la orientación de los paneles solares.
- Potencia pico: potencia máxima instalada de cada planta fotovoltaica ³.
- Pérdidas del Sistema: factor que considera las pérdidas de eficiencia del sistema.

La API devuelve un conjunto de datos horarios que incluyen la curva de generación en vatios (W). Estos datos se convierten a kilovatios hora (kWh) para facilitar su análisis y comparación.

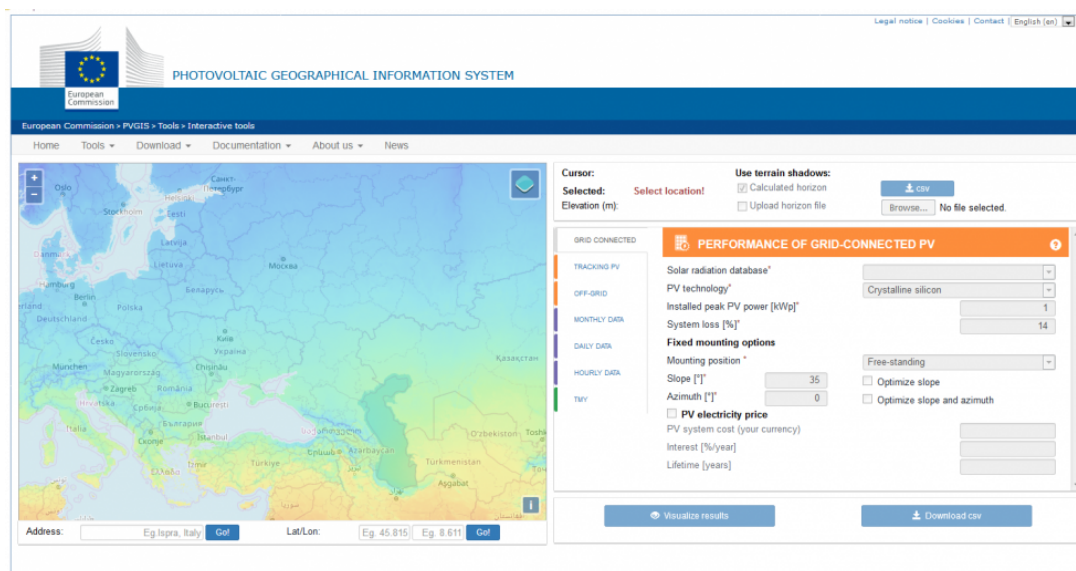


Figura 2.2: Interfaz de PVGIS cuando accedes por primera vez a la aplicación web, antes de hacer nada. **Fuente:**PVGIS.

De esta forma, con los parámetros enumerados cuyos valores fueron proporcionados por EnergyRIS, se recopilaron los datos horarios PV de generación fotovoltaica **desde el 1 de mayo de 2021 hasta el 28 de mayo de 2024** ⁴. En total,

$$7 \text{ plantas fotovoltaicas} \times 26.952 \text{ horas} = 188.664 \text{ datos}$$

Con el fin de analizar la generación de energía de manera global, se calculó la **generación total horaria agregada de las 7 plantas fotovoltaicas**. Al calcular la generación

²El web scraping es una técnica utilizada para extraer datos de sitios web de manera automatizada.

³La potencia pico (también conocida como potencia máxima) es la cantidad máxima de energía que un panel solar puede producir. Esta cantidad se especifica en los datos técnicos del panel solar y se mide en vatios (W). Esta potencia se mide a una temperatura de 25°C y una radiación solar de 1.000 W/m².

⁴Se seleccionó este periodo para que coincidiera con el de los datos de demanda de los CUPS.

total horaria, se puede evaluar cómo cada método de asignación de coeficientes afecta la distribución de energía entre los participantes del autoconsumo colectivo. De esta forma, el número de datos se redujo a **26.952 datos**.

Es importante tener en cuenta que los años 2021 y 2024 no están completos, lo cual debe ser considerado en el análisis.

Análisis de la generación horaria

La siguiente gráfica muestra la distribución de la generación fotovoltaica horaria total para los años 2021, 2022, 2023 y 2024. Se utiliza un diagrama de cajas y bigotes para cada hora del día, proporcionando una visión detallada de la variabilidad y distribución de la generación de energía a lo largo del día.

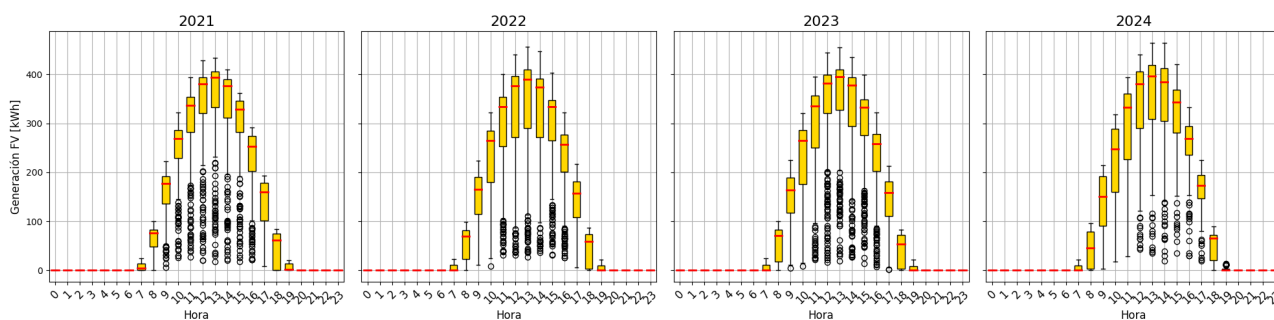


Figura 2.3: Generación fotovoltaica horaria total (2021-2024).

Analizando las gráficas, se comprueba la coherencia de los datos. Se observa que la generación sigue un patrón claro y consistente a lo largo de los años, aumentando a partir de las 7:00 am, alcanzando su pico máximo entre las 11:00 am y las 2:00 pm, y disminuyendo gradualmente hasta las 7:00 pm.

Análisis de la generación mensual

Observando la siguiente gráfica se observa que la generación sigue un patrón estacional claro y consistente a lo largo de los años, con picos durante los meses de verano (junio, julio y agosto) y valles en los meses de invierno (diciembre, enero y febrero).

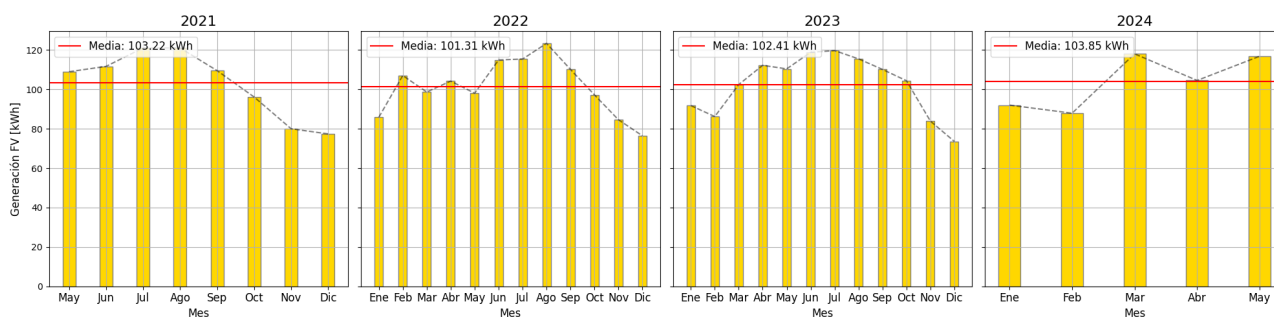


Figura 2.4: Gráfica de la generación fotovoltaica mensual total. Las barras representan la generación mensual en kilovatios hora (kWh). La línea horizontal roja indica la media anual de generación.

2.2. Datos de demanda

Para la obtención de los datos de demanda de los cuatro CUPS de la Universidad de La Laguna se utilizó la API de Datadis⁵. De nuevo, se automatizó la descarga de los datos con un código de Python.

Tras el preprocesamiento de los datos, se obtuvo un dataset con un total de

4 CUPS x 26.952 horas = 107.808 datos

de demanda energética horaria. El periodo de los datos abarca **desde el 1 de mayo de 2021 hasta el 28 de mayo de 2024**. Como ya se ha mencionado, es importante destacar que los años 2021 y 2024 no están completos. A continuación, se presenta un resumen de la demanda total para los años 2021 a 2024.

Año	Total [MWh]	Total sol [MWh]	Mes min	Mes max	Día min	Día max
2021	1220,41	734,82	08	10	11/08	04/10
2022	1699,15	970,70	08	05	31/12	09/06
2023	1571,67	945,50	08	10	21/08	09/10
2024	611,57	345,41	05	05	28/05	15/01

Tabla 2.1: Demanda anual (Total) y demanda en horas de sol (Total sol), junto con los meses y días de mínima y máxima demanda total para los años 2021 a 2024. Los valores están expresados en megavatios hora (MWh), y las fechas en formato dd/mm.

La diferencia entre la demanda total y la demanda en horas de sol es un indicador clave de cómo se distribuye el consumo energético a lo largo del día y la noche. En 2021 (de mayo a diciembre), el 60,23 % de la demanda total se produjo durante las horas de sol. En 2022, este porcentaje disminuyó a 57,13 %. En 2023, la demanda en horas de sol volvió a aumentar al 60,16 %. Finalmente, en 2024 (de enero a mayo), el 56,48 % de la demanda se produjo durante las horas de sol. En general, se observa que **la demanda en horas de sol implica aproximadamente el 60 % de la demanda anual**. Este hecho recalca el potencial del ahorro energético del autoconsumo.

Contemplando los años completos de 2022 y 2023, se observa que el mes con menor demanda es agosto, lo cual es coherente con el hecho de que es un mes de vacaciones en el que la Universidad permanece cerrada.

⁵La API de Datadis es de acceso privado, por lo que los datos usados para este estudio no se pueden compartir. Más información en Datadis.

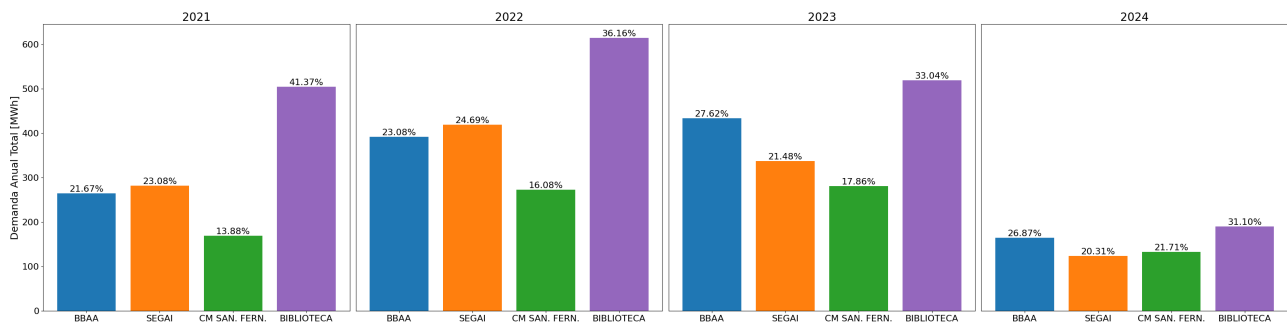


Figura 2.5: La gráfica muestra la demanda anual total en megavatios hora (MWh) para cada CUPS.

En 2021, la demanda del CUPS de la Biblioteca representaba el 41,37 % de la demanda total, disminuyendo a 31,10 % en 2024. Por otro lado, el CUPS de la Facultad de BBAA incrementó su participación del 21,67 % en 2021 a 26,87 % en 2024. El CUPS del SEGAI tuvo una participación estable, del 23,08 % en 2021 al 20,31 % en 2024. Por último, el CUPS del C.M. San Fernando, tuvo la menor participación en 2021 (13,88 %). Esta aumentó a 21,71 % en 2024.

A continuación, se hará un análisis para cada uno de los CUPS participantes del autoconsumo colectivo de manera individual. Para todos ellos, se analizarán las siguientes gráficas de boxplots:

- Gráfica de la demanda horaria media por año: muestra los boxplots de la demanda horaria media anual. Cada subgráfica representa un año específico, y los datos se agrupan por horas del día, de 0 a 23 horas. Esta visualización permite observar las fluctuaciones en la demanda horaria a lo largo del día y comparar cómo estas varían entre diferentes años.
- Gráfica de la demanda horaria media por día de la semana: muestra los boxplots de la demanda horaria media para cada día de la semana. Cada subgráfica representa un día específico, desde lunes hasta domingo, y los datos se agrupan por horas del día, de 0 a 23 horas. Esta visualización ayuda a identificar patrones y variaciones en la demanda energética a lo largo de la semana, mostrando cómo la demanda cambia entre días laborables y fines de semana.
- Gráfica de la demanda mensual por año: muestra los boxplots de la demanda mensual anual. Cada subgráfica representa un año específico, y los datos se agrupan por meses, de enero a diciembre. Esta visualización permite observar la estacionalidad en la demanda energética, identificando meses con mayor o menor consumo y facilitando la comparación de tendencias entre diferentes años.

Edificio del SEGAI

El Edificio SEGAI de la Universidad de La Laguna es una instalación destinada a proporcionar soporte y servicios a la comunidad investigadora de la universidad. Aquí se encuentran diversos laboratorios y unidades especializadas que facilitan la realización de investigaciones en distintas disciplinas científicas.



Figura 2.6: Fotografía del edificio del SEGAI.

Observando la Fig.2.7 se aprecia un cambio en la distribución horaria de la demanda de 2022 a 2023. Mientras que para los dos primeros años esta es casi constante a lo largo del día, con ligeros picos en las horas de la tarde, en 2023 la demanda se desplazó a las horas de sol, causando una disminución de la demanda base. Esta tendencia se mantiene para los meses de 2024 contemplados. Se desconoce la causa de este fenómeno. Se estima que puede estar causado por la implementación de nuevos experimentos y proyectos que requieren un uso intensivo de equipos durante las horas de sol. Además, es posible que haya cambiado el modo de trabajo en el SEGAI, con una mayor concentración de actividades durante el día para maximizar el uso de la luz solar y reducir el consumo en horas nocturnas.

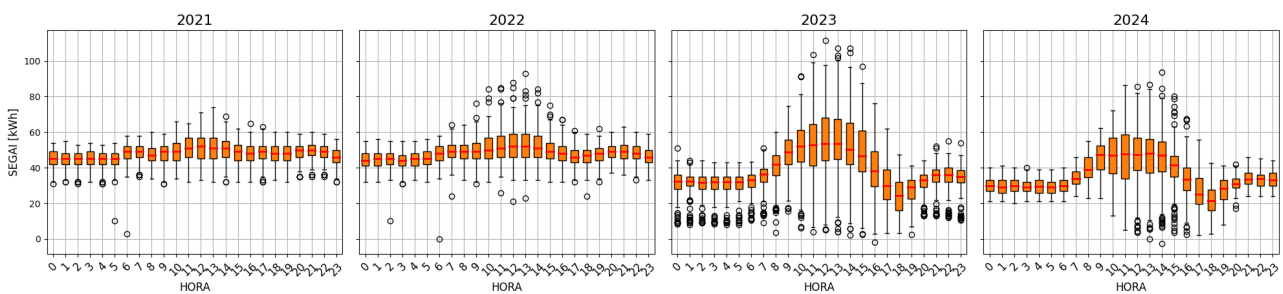


Figura 2.7: Demanda horaria media por año del CUPS SEGAI.

Como se observa en 2.9 la distribución de la demanda sigue un patrón bastante constante durante los días laborales, con ligeros picos en las horas centrales del día. Los sábados y domingos muestran una demanda más reducida y menos variable, reflejando una menor actividad durante el fin de semana. Sin embargo, hay picos ocasionales que sugieren la realización de algunas actividades o experimentos especiales incluso en estos días.

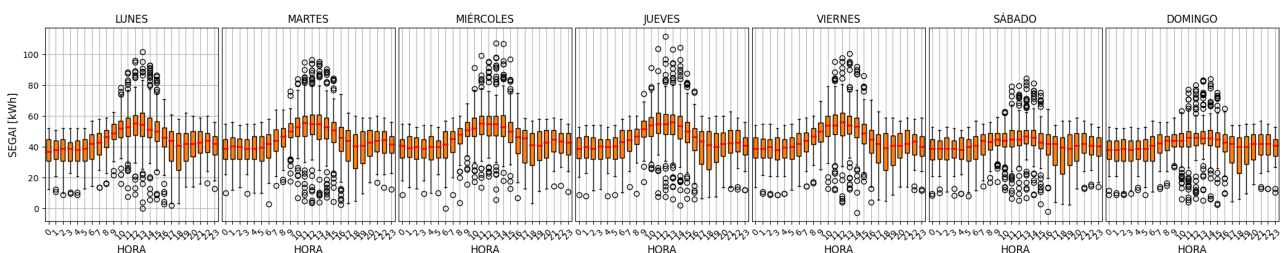


Figura 2.8: Demanda horaria media por día de la semana del CUPS SEGAI.

Analizando la distribución de su demanda mensual 2.9, se observa un incremento significativo en la variabilidad y en los niveles de demanda a lo largo de los años, especialmente en 2023. Como ya se ha mencionado, esta tendencia podría ser un reflejo de cambios en las dinámicas de trabajo y en la implementación de actividades que demandan más energía.

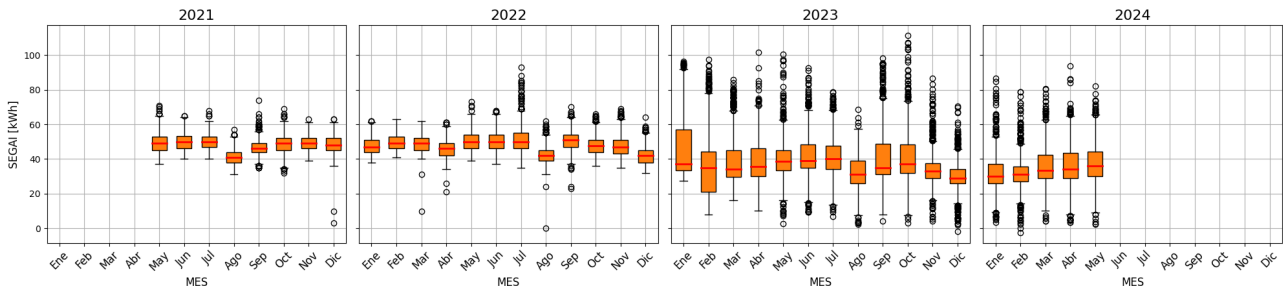


Figura 2.9: Demanda mensual por año del CUPS SEGAI.

Facultad de Bellas Artes

La Facultad de Bellas Artes de la Universidad de La Laguna es un centro académico dedicado a la formación en diversas disciplinas artísticas: grado en bellas artes, másteres y programas de posgrado, exposiciones, etc.



Figura 2.10: Fotografía de la Facultad de BBAA.

La siguiente gráfica 2.11 muestra algunas tendencias interesantes en la distribución del consumo energético a lo largo de estos años. En los años 2021 y 2022 se observan patrones similares: una demanda que es baja en las primeras horas del día y que alcanza cierta estabilidad durante las horas de mayor actividad (de 9 a 20 horas), con un promedio de consumo que se mantiene alrededor de los 50 kWh. Sin embargo, en 2022 hay una mayor variabilidad en la demanda especialmente en las horas de mayor consumo, lo que indica posibles fluctuaciones en la actividad o el uso de instalaciones.

La tendencia descrita cambia en 2023, donde se observa un cambio notable en el patrón de consumo energético. La demanda aumenta rápidamente desde las 7 horas, alcanzando un pico muy pronunciado entre las 10 y las 14 horas, y luego disminuye de manera

considerable en la tarde. Esta variabilidad sugiere que ha habido una concentración de actividades principalmente en la mañana. Una posible explicación para este cambio es una reestructuración del horario académico, donde las clases y actividades que antes se distribuían uniformemente entre la mañana y la tarde se han concentrado más en las horas de la mañana. Esta tendencia se mantiene en 2024.

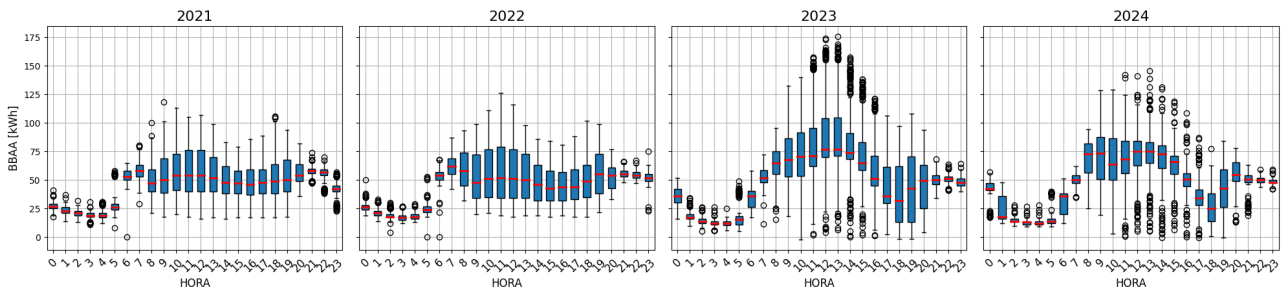


Figura 2.11: Demanda horaria media por año del CUPS BBA

La distribución de la demanda semanal es clara, de lunes a viernes, la demanda energética sigue un patrón consistente de aumento en la mañana y un pico entre las 10 y las 14 horas, seguido de una disminución en la tarde. Por otro lado, los sábados y domingos tienen una demanda significativamente menor, con picos menos pronunciados y una variabilidad más alta durante las horas de actividad.

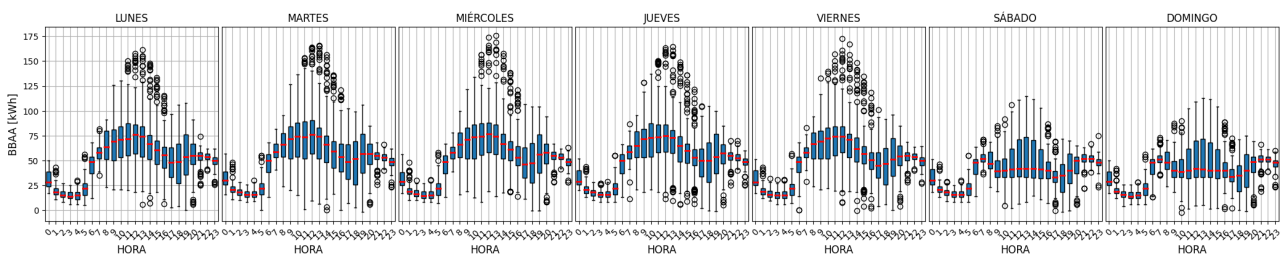


Figura 2.12: Demanda horaria media por día de la semana del CUPS BBA.

El análisis de los patrones de demanda energética mensual de Fig.2.13, revela una clara disminución en la demanda energética durante los meses de verano (julio y agosto) cada año, consistente con el período vacacional. Destaca la mayor variabilidad y los picos de demanda de 2023, posiblemente debido a eventos específicos o cambios en la programación de actividades.

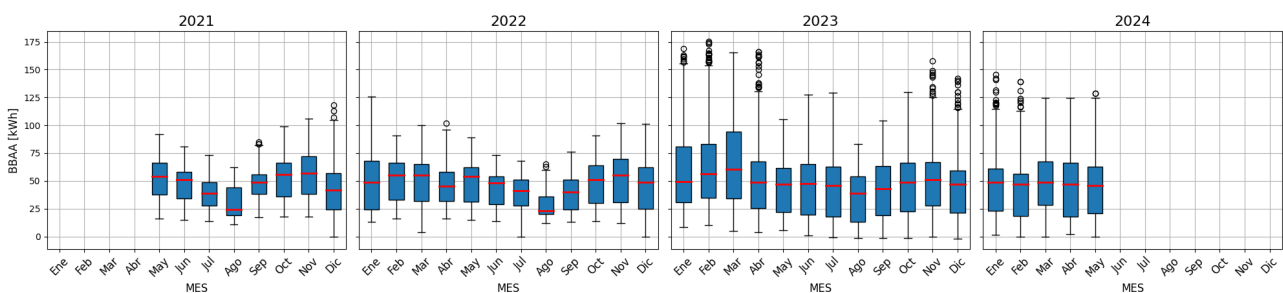


Figura 2.13: Demanda mensual por año del CUPS BBA.

Colegio Mayor San Fernando

Este colegio mayor proporciona alojamiento a más de 100 estudiantes universitarios. Además de las habitaciones, las instalaciones cuentan con comedores, salas de estudio, biblioteca, lavandería y áreas recreativas.



Figura 2.14: Fotografía del C.M San Fernando.

Comparando las gráficas anuales de la Fig.2.15, se observa un incremento general en la demanda energética y en la variabilidad del consumo a lo largo del tiempo. Mientras que 2021 muestra un patrón más estable y predecible, los años siguientes, especialmente 2023 y 2024, presentan mayores picos de demanda y una variabilidad más alta, lo que sugiere un aumento en la intensidad y duración del uso de las instalaciones del colegio mayor.

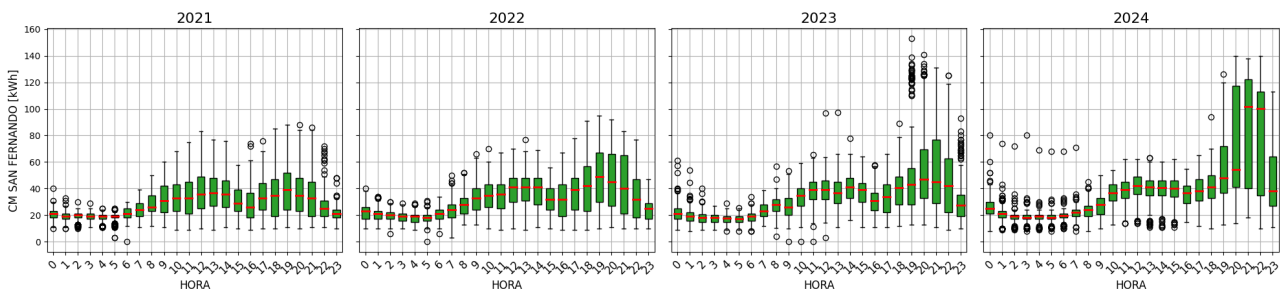


Figura 2.15: Demanda horaria media por año del CUPS C.M. San Fernando

El análisis de la demanda energética horaria por día de la semana en el revela patrones similares a los estudiados en el resto de CUPS: de lunes a viernes, la demanda sigue un patrón consistente con picos significativos alrededor del mediodía y últimas horas de la tarde. Durante el fin de semana se reduce notablemente la demanda en las horas pico pero se mantiene la misma tendencia. Destaca el número de outliers observados el viernes por la noche. El notable número de outliers en la demanda energética durante la noche del viernes indica picos elevados de consumo asociados probablemente a actividades sociales y recreativas, así como preparaciones para el fin de semana.

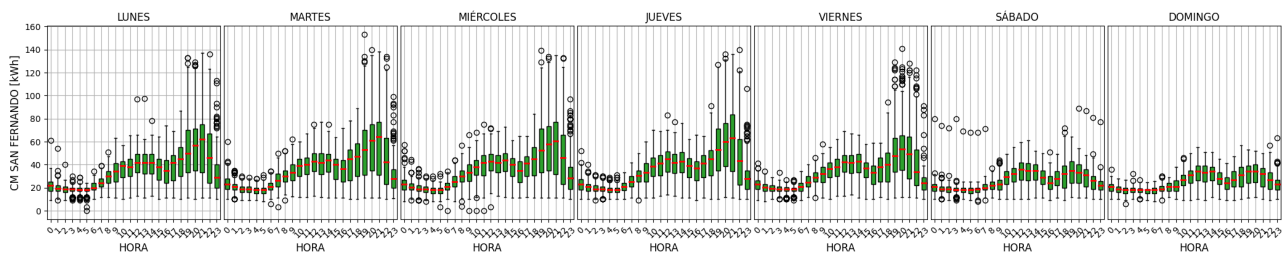


Figura 2.16: Demanda horaria media por día de la semana del CUPS C.M. San Fernando.

En cuanto a la evolución de la distribución de la demanda mensual a lo largo de los años mostrada en la Fig.2.9, se aprecia un aumento general y una mayor variabilidad en los últimos años.

Mientras que 2021 y 2022 presentan patrones más estables y predecibles, 2023 y 2024 se caracterizan por picos más frecuentes y valores atípicos, lo que sugiere un uso más intensivo y fluctuante de las instalaciones.

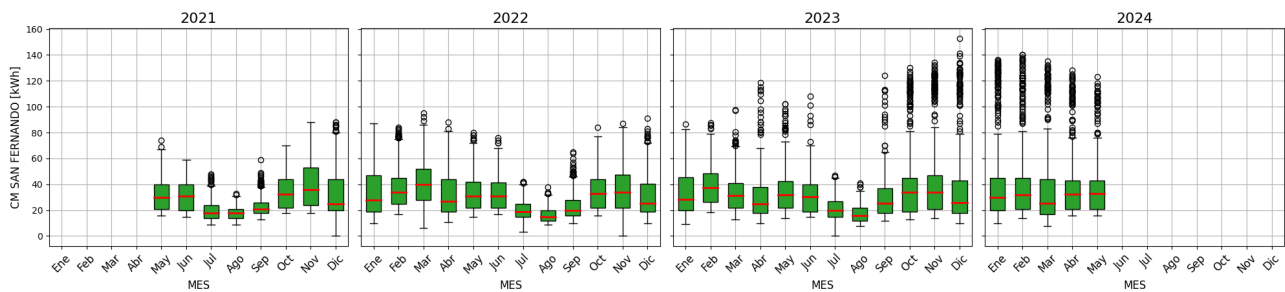


Figura 2.17: Demanda mensual por año del CUPS C.M. San Fernando.

Biblioteca General

Este edificio no solo sirve como un espacio para el estudio y la investigación, sino que también alberga una gran cantidad de recursos y servicios destinados a apoyar a la comunidad universitaria. Tiene horarios de apertura extendidos para acomodar las necesidades de los usuarios, especialmente durante los períodos de exámenes. Sin embargo, hay variaciones en los horarios según la época del año y según el año. Por ejemplo, durante el verano, la biblioteca puede tener horarios reducidos o incluso cerrar durante algunos días específicos.



Figura 2.18: Fotografía del edificio de la Biblioteca General.

Analizando la distribución de la demanda horaria 2.19 se observan cambios notables en 2023, donde la variabilidad y los picos de demanda aumentan significativamente. Además, destacar la disminución de la demanda desde las 14:00 hasta las 16:00 de la tarde, que aunque se intuye en 2022, es clara a partir de 2023.

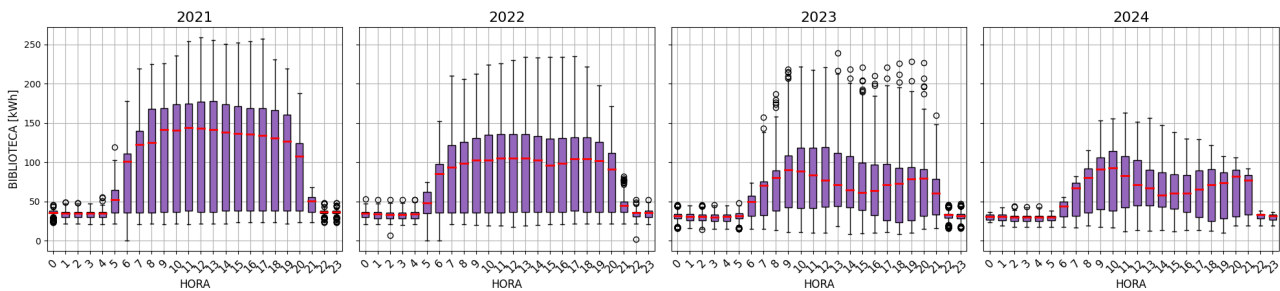


Figura 2.19: Demanda horaria media por año del CUPS Biblioteca General.

La Fig.2.20 muestra una demanda alta y constante durante los días laborales, con picos en las horas de mayor uso y una demanda mínima (se mantiene cierta demanda básica por actividades de mantenimiento o sistemas esenciales) durante los fines de semana debida al cierre de la biblioteca.

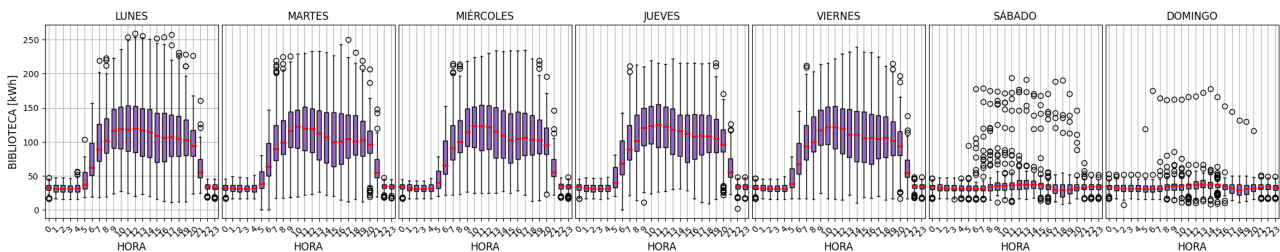


Figura 2.20: Demanda horaria media por día de la semana del CUPS Biblioteca General.

En la gráfica de Fig.2.21 se observa una alta variabilidad a lo largo de los años, con picos en los meses de julio y una reducción en septiembre a partir de 2022. Este cambio podría ser atribuible a la modificación en el calendario académico, donde las recuperaciones se trasladaron de septiembre a julio, y a los horarios de operación de la

biblioteca, que influyen significativamente en el uso de las instalaciones. En todos los años estudiados la demanda en agosto disminuye debido al cierre parcial. En resumen, los patrones de demanda de la Biblioteca están fuertemente influenciados por los cambios en su horario de operación y el calendario académico.

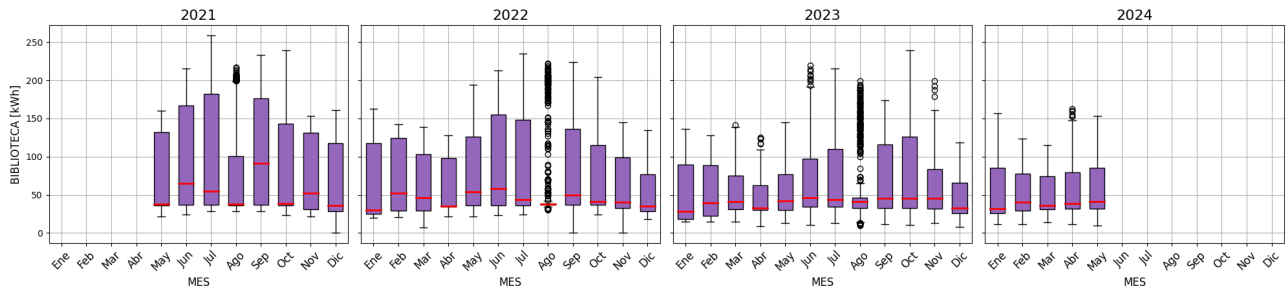


Figura 2.21: Demanda mensual por año del CUPS Biblioteca General.

Capítulo 3

Metodos y técnicas

En este capítulo se detallan los métodos y las técnicas utilizadas para extraer nuevos datos a partir de los datos de demanda y generación fotovoltaica descritos en el capítulo 2. Estos nuevos datos serán necesarios para la obtención de los resultados del estudio.

En la **primera sección** se explican las metodologías para el cálculo de los coeficientes de reparto energético que fueron mencionados en el capítulo 1. En la **segunda sección** se exponen las técnicas de preprocesamiento de datos que fueron aplicadas para preparar el dataset para realizar predicciones con el modelo de aprendizaje automático supervisado¹ XGBRegressor. La **tercera y última sección** se centra en este modelo predictivo. Tras una breve introducción teórica, se describirá su implementación. Para facilitar el seguimiento de este capítulo se facilita el siguiente esquema:

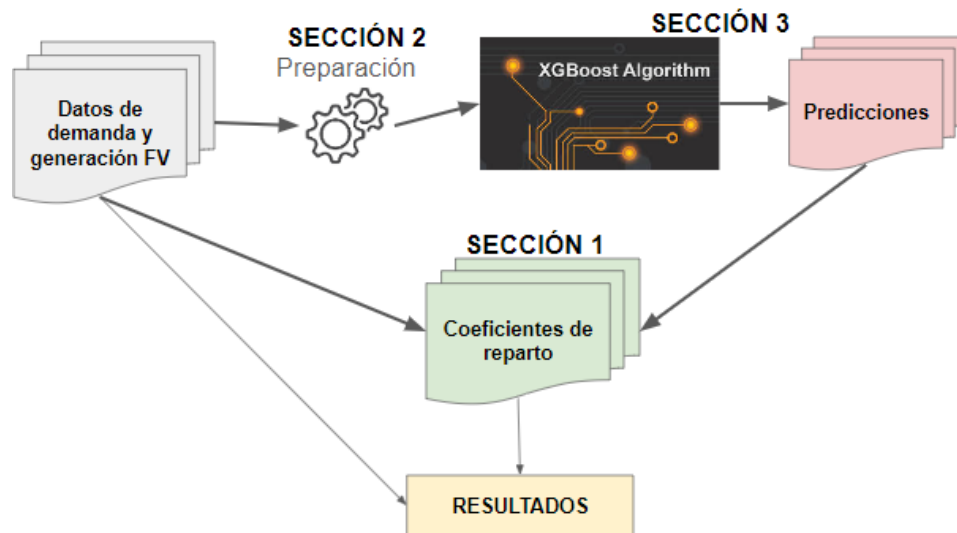


Figura 3.1: Esquema de la metodología de trabajo del capítulo 3. **Fuente:** Elaboración propia.

¹En el aprendizaje supervisado, el modelo se entrena utilizando un conjunto de datos etiquetados, es decir, cada entrada de datos tiene una etiqueta o resultado conocido. El objetivo del modelo es aprender la relación entre las características de entrada (variables independientes) y las etiquetas (variables dependientes) para poder predecir las etiquetas de nuevas entradas de datos.

3.1. Cálculo de los coeficientes de reparto

En esta sección **se explican los diferentes enfoques para calcular los coeficientes de reparto**. Además de los métodos de cálculo contemplados por la legislación española vigente, RD244/2019 ('coeficientes tradicionales'), se considerarán dos métodos adicionales: el primero, que es de creación propia, se denominará 'coeficientes de reparto justo' y el segundo, propuesto en (8), al que llamaremos 'coeficientes con correlaciones de Pearson'.

En todos los casos, la suma horaria de los coeficientes debe ser igual a uno. Pues esto asegura que la cantidad de energía a repartir se reparta completamente entre los participantes. Para facilitar el seguimiento del estudio realizado se proponen las siguientes clasificaciones:

1. Tipo de coeficiente:

- **Coeficientes fijos:** son constantes para todas las horas del periodo de análisis y para su cálculo se emplean datos agregados sobre un periodo de tiempo específico.
- **Coeficientes variables:** estos coeficientes se calculan para cada hora. Por lo tanto varían en función de las condiciones temporales horarias.

2. Datos usados para el cálculo de los coeficientes:

- **Datos del año anterior** al de aplicación del reparto: los datos reales de generación fotovoltaica y demanda de los participantes de cierto año se usarán para calcular los coeficientes que se encargarán del reparto energético en el año siguiente.
- **Predicciones** de los datos para el año de aplicación del reparto: con los datos de cierto año, se estiman la demanda y la generación del año siguiente. Los coeficientes se calculan con estas estimaciones.

3.1.1. Coeficientes de reparto bajo el marco regulatorio del RD 244/2019

Se detallan las fórmulas matemáticas para el cálculo de los coeficientes de reparto energético propuestos por el RD 244/2019 del 5 de abril. En ambos casos, dichos coeficientes solo dependen de la demanda de los participantes.

Coeficientes fijos

Para el cálculo del coeficiente para cada participante, se emplea la siguiente fórmula:

$$C_i = \frac{D_i}{\sum_{j=1}^n D_j} \quad (3.1)$$

donde C_i es el coeficiente de reparto para el participante i , D_i es la demanda total del participante i durante el periodo de análisis y n es el número total de participantes.

Coefficientes variables

Para el cálculo de los coeficientes horarios para cada participante se emplea la siguiente fórmula:

$$C_{i,h} = \frac{D_{i,h}}{\sum_{j=1}^n D_{j,h}} \quad (3.2)$$

donde $C_{i,h}$ es el coeficiente de reparto para el participante i en la hora h , $D_{i,h}$ es la demanda del participante i en la hora h y n es el número total de participantes.

3.1.2. Coeficientes de Reparto Justo

Se plantea un método alternativo para el cálculo de los coeficientes de reparto que permite asegurar que la energía generada se distribuye de manera justa y proporcional entre los consumidores, adaptándose a las condiciones de generación y demanda en cada hora. Por lo tanto, los coeficientes calculados dependerán tanto de las demandas como de la generación fotovoltaica. Antes del cálculo de los coeficientes de reparto, es necesario considerar los siguientes casos:

- **Caso 1 - Generación suficiente para cubrir la demanda total de la hora h.** En las horas en las que la generación fotovoltaica es mayor o igual a la suma de las demandas se asigna a cada consumidor la energía que demanda.
- **Caso 2 - Generación insuficiente para cubrir la demanda total en la hora h.** En las horas en las que la generación es menor que la suma de las demandas, la asignación de la energía generada se hace según el algoritmo descrito en 1.

$$D_h^{\text{cubierta}} = \begin{cases} D_h = (D_{1,h}, \dots, D_{n,h}) & \text{si } G_h \geq \sum_{i=1}^n D_{i,h} \\ D_h^{\text{cubierta}(K)} = (D_{1,h}^{\text{cubierta}(K)}, \dots, D_{n,h}^{\text{cubierta}(K)}) & \text{si } G_h < \sum_{i=1}^n D_{i,h} \end{cases} \quad (3.3)$$

donde D_h^{cubierta} es el vector de longitud n (participantes autoconsumo colectivo) de la demanda cubierta con la asignación. D_h es el vector con las demandas horarias de los n participantes para el caso 1 y $D_h^{\text{cubierta}(K)}$ es el vector la demanda cubierta para el caso 2. G_h es la generación de la hora h .

A continuación se describen los pasos del algoritmo para el cálculo del vector $D_h^{\text{cubierta}(K)}$ en el caso 2.

1. En **6** se calcula la cantidad de energía a distribuir inicialmente como el mínimo entre la generación dividida por el número de consumidores y la menor demanda.
2. En **7** se asigna esta distribución inicial a cada consumidor y se ajustan las demandas y la generación restante. Si aún queda energía por distribuir después de la distribución inicial, se redistribuye el excedente de la misma forma.
3. En **8** y **9** se recalculan los inputs.
4. En **10** se suma una iteración.

Considerando el reparto descrito, se calcularán de nuevo los coeficientes fijos y variables.

Algoritmo 1: Cálculo de la demanda curbierta para el caso 2

1 **Input:** $G_h, D_{i,h}$ para $i = 1, \dots, n$
2 **Output:** $D_{i,h}^{\text{cubierta}(K)}$ para $i = 1, \dots, n$
3 **Inicialización:** $K = 1$;
4 $D_h^{\text{cubierta}(1)} = (0, \dots, 0)$;
5 **while** $G_h > 0$ **do**
6 $R = (\text{mín}(\frac{G_h}{n}, \text{mín}_i\{D_{i,h}\}), \dots, \text{mín}(\frac{G_h}{n}, \text{mín}_i\{D_{i,h}\}))$ con $D_{i,h} \neq 0$;
7 $D_{i,h}^{\text{cubierta}} = D_{i,h}^{\text{cubierta}} + R \quad \forall i$;
8 $G_h = G_h - R \times n$;
9 $D_{i,h} = D_{i,h} - D_{i,h}^{\text{cubierta}} \quad \forall i$;
10 $K = K + 1$;

Coefficientes fijos

En este caso el cálculo del coeficiente para el participante i se hará con las demandas cubiertas en base al reparto detallado:

$$C_i = \frac{D_i^{\text{cubierta}}}{\sum_{j=1}^n D_j^{\text{cubierta}}} \quad (3.4)$$

donde C_i es el coeficiente de reparto para el participante i , D_i^{cubierta} es su demanda cubierta total durante el periodo de análisis y n es el número total de participantes.

Coefficientes variables

En este caso, se repite el cálculo anterior usando la demanda cubierta en cada hora:

$$C_{i,h} = \frac{D_{i,h}^{\text{cubierta}}}{\sum_{j=1}^n D_{j,h}^{\text{cubierta}}} \quad (3.5)$$

siendo $C_{i,h}$ y $D_{i,h}^{\text{cubierta}}$ el coeficiente de reparto horario y la demanda cubierta horaria del participante i , respectivamente.

3.1.3. Coeficientes con Correlaciones de Pearson

Esta otra alternativa se dispone a realizar el reparto energético usando los coeficientes de correlación de Pearson entre las demandas de los participantes, $D_{i,h}$, y la generación FV, G_h .

$$\rho_i = \frac{\text{cov}(D_{i,h}, G_h)}{\sigma_{D_{i,h}} \sigma_{G_h}} \quad (3.6)$$

- ρ_i es el coeficiente de correlación de Pearson entre la curva de demanda del participante i , $D_{i,h}$, y la curva de generación fotovoltaica, G_h .
- $\text{cov}(D_{i,h}, G_h)$ es la covarianza entre $D_{i,h}$ y G_h .
- $\sigma_{D_{i,h}}$ y σ_{G_h} son las desviaciones estándar de $D_{i,h}$ y G_h , respectivamente.

Coefficientes fijos

Para este cálculo se usarán los coeficientes de correlación de Pearson. Una vez escalados al rango $[0, 1]$ se podrán interpretar como proporciones².

$$C_i = \frac{\rho_i^{\text{escalado}}}{\sum_{j=1}^n \rho_j^{\text{escalado}}} \quad (3.7)$$

donde ρ_i^{escalado} es el coeficiente de Pearson escalado correspondiente al participante i .

3.2. Preparación de los datos para las predicciones

En esta sección se describirán las principales técnicas realizadas para preparar el dataset original para su uso en la obtención de predicciones: extracción de características y escalado de los datos.

3.2.1. Extracción de características

La extracción de características consiste en generar características relevantes a partir del conjunto de datos original. Se trata de una etapa crítica en el preprocesamiento de datos para el entrenamiento de modelos de aprendizaje automático. Estas características ayudan a capturar patrones temporales y eventos específicos que pueden influir en la demanda y generación de energía.

Distinción de días festivos y fines de semana

Se implementó una clase personalizada, "*CanaryHolidaysCalendar*"³, para definir los festivos no laborables de Tenerife, incluyendo la lógica para ajustar las fechas al día laboral más cercano si caen en un fin de semana. Además, se detectan los fines de semana de cada año, asegurando que el modelo se adapte correctamente a las variaciones anuales en los días no laborables. Esta característica se añadió como una columna binaria al dataset original.

Generación de características temporales

Se añadieron características como la hora del día, el día de la semana y el mes del año utilizando transformaciones trigonométricas (coseno) que capturan la naturaleza cíclica de estas variables. Estas transformaciones permiten representar de manera continua y periódica las variaciones temporales, facilitando así que el modelo de aprendizaje automático pueda captar patrones cíclicos inherentes a los datos temporales.

1. **Hora del día:** se transformó utilizando la fórmula (3.10). Esto convierte las horas, que originalmente oscilan entre 0 y 23, en una función cíclica que refleja la repetición diaria de las actividades humanas y los patrones de consumo energético. Esta

²Los coeficientes de correlación de Pearson, ρ_i , pueden variar entre -1 y 1.

³Se utilizaron las siguientes bibliotecas para implementar las funcionalidades descritas: *AbstractHolidayCalendar*, *Holiday* y *nearest_workday* de *pandas.tseries.holiday*, *easter* de *dateutil.easter* y *relativedelta* de *dateutil.relativedelta*.

transformación ayuda a que el modelo entienda que la hora 23 y la hora 0 están próximas en el ciclo diario.

$$hora_{dia} = \cos\left(\frac{2\pi \cdot hora}{24}\right) \quad (3.8)$$

2. **Día de la semana:** similarmente, se aplicó la transformación (3.11) para los días de la semana. Esta transformación convierte los días, que van de 0 (lunes) a 6 (domingo), en una función cíclica. Esto es útil para capturar patrones semanales en el comportamiento de la demanda y la generación de energía, como la diferencia entre días laborables y fines de semana.

$$dia_{semana} = \cos\left(\frac{2\pi \cdot dia}{7}\right) \quad (3.9)$$

3. **Mes del año:** para los meses del año, se utilizó la transformación (3.12). Esta transformación convierte los meses, que van de 1 a 12, en una función cíclica que refleja la estacionalidad anual. Esto permite al modelo capturar variaciones estacionales en la demanda y generación de energía, como las diferencias entre verano e invierno.

$$mes_{año} = \cos\left(\frac{2\pi \cdot mes}{12}\right) \quad (3.10)$$

3.2.2. Escalado de datos

Se utilizó "*StandardScaler*" de "*scikit-learn*" para escalar las características, normalizando los datos para que tengan media cero y desviación estándar uno. Este paso es esencial para que los algoritmos de aprendizaje automático funcionen de manera óptima, pues asegura que todas las características contribuyan de manera equilibrada al modelo, evitando que las de mayor rango dominen sobre aquellas con rango menor.

$$z = \frac{x - \mu}{\sigma} \quad (3.11)$$

Donde z es el valor escalado, x el original, μ su media y σ su desviación estándar.

3.3. Modelo de predicción

Como se ha descrito en la clasificación expuesta en 2, los coeficientes de reparto también se calcularán empleando las predicciones de la demanda y la generación fotovoltaica correspondientes al año de aplicación de los mismos. El interés de este método reside en la posibilidad de anticiparse a las variaciones en la demanda y la generación que puedan ocurrir en el año de aplicación.

Para esta tarea, se ha seleccionado el algoritmo **XGBRegressor** ('Extreme Gradient Boosting Regressor'). A continuación se introducirá dicho algoritmo y se detallará su implementación.

3.3.1. XGBRegressor

El algoritmo XGBRegressor es una implementación de **XGBoost** ('Extreme Gradient Boosting' o Refuerzo de Gradientes Extremo) ⁴, específicamente diseñada para tareas de regresión. Es parte de la librería de Python 'xgboost' (10) (usa la API de 'scikit-learn'), que es una biblioteca optimizada para el aprendizaje automático basada en árboles de decisión conocida por su eficiencia, flexibilidad y velocidad. Como su nombre indica, su funcionamiento está basado en el '*boosting*', técnica que combina predicciones de varios modelos débiles para crear un modelo fuerte. En 'XGBoost', esto se realiza secuencialmente, ajustando cada nuevo árbol para corregir los errores de los anteriores.

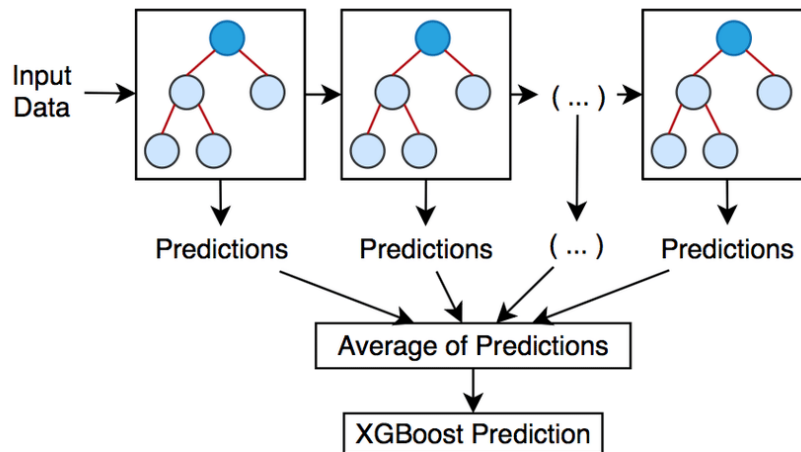


Figura 3.2: Esquema del funcionamiento algoritmo de 'XGBoost'. **Fuente:** Imagen obtenida de (11).

Motivos para la elección del algoritmo

1. Eficiencia: es conocido por su capacidad de manejar grandes conjuntos de datos de manera eficiente, utilizando técnicas avanzadas de optimización y paralelización para acelerar el proceso de entrenamiento del modelo (9).
2. Precisión: este algoritmo ha demostrado un alto rendimiento en diversas competiciones de aprendizaje automático y estudios académicos, destacándose por su capacidad para producir modelos altamente precisos (12).
3. Flexibilidad: permite la optimización de varios hiperparámetros, lo que ofrece la posibilidad de ajustar el modelo para obtener el mejor rendimiento posible para un conjunto de datos específico. Esta flexibilidad es crucial para mejorar la precisión de las predicciones en tareas complejas (10).
4. Prevención del overfitting: incluye parámetros que ayudan a prevenir el sobreajuste del modelo a los datos de entrenamiento, lo que resulta en mejores predicciones en datos no vistos.

⁴XGBoost' es una implementación de código abierto muy eficiente del algoritmo de *boosting* de gradiente. Fue desarrollado inicialmente por Tianqi Chen y descrito por Chen y Carlos Guestrin en su paper de 2016 titulado 'XGBoost: A Scalable Tree Boosting System' (9).

Funcionamiento

1. Ensemble de árboles de decisión: XGBoost crea un conjunto (ensemble) de árboles de decisión secuenciales. Cada árbol se entrena para corregir los errores cometidos por el árbol anterior.
2. Función de pérdida y gradientes: en cada iteración, el algoritmo minimiza una función de pérdida (por ejemplo, el error cuadrático medio) utilizando el gradiente descendente. Los gradientes indican la dirección y magnitud de los ajustes necesarios para minimizar los errores.
3. Regularización: XGBoost incluye términos de regularización L1 (α) y L2 (γ) en su función de pérdida para prevenir el sobreajuste. La regularización ayuda a mantener el modelo simple y generalizable.
4. Peso de los árboles: cada árbol en el ensemble tiene un peso que se ajusta durante el entrenamiento. Este peso refleja la contribución del árbol al resultado final. El aprendizaje se realiza con una tasa de aprendizaje (η), que controla la contribución de cada árbol nuevo.
5. Manejo de datos faltantes: XGBoost puede manejar automáticamente los valores faltantes en los datos de entrada, asignando direcciones óptimas en los nodos de decisión basados en estadísticas de entrenamiento.

3.3.2. Implementación del modelo de predicción

A continuación se detallarán los pasos seguidos para la implementación del modelo de predicción.

Selección de características y etiquetas

Para ambos conjuntos de datos (entrenamiento y predicción), se separaron las características y las etiquetas. Las características son las variables independientes que se utilizarán para entrenar el modelo, mientras que las etiquetas son las variables objetivo que el modelo intentará predecir.

Validación cruzada con 'TimeSeriesSplit'

La técnica de 'TimeSeriesSplit' se utiliza para la validación cruzada respetando la secuencia temporal de los datos. Esta es esencial en series temporales porque asegura que el modelo se entrene en datos del pasado y se pruebe en datos del futuro, evitando la fuga de datos. Para esto, divide los datos en múltiples particiones, donde cada partición de prueba está cronológicamente después de la partición de entrenamiento.

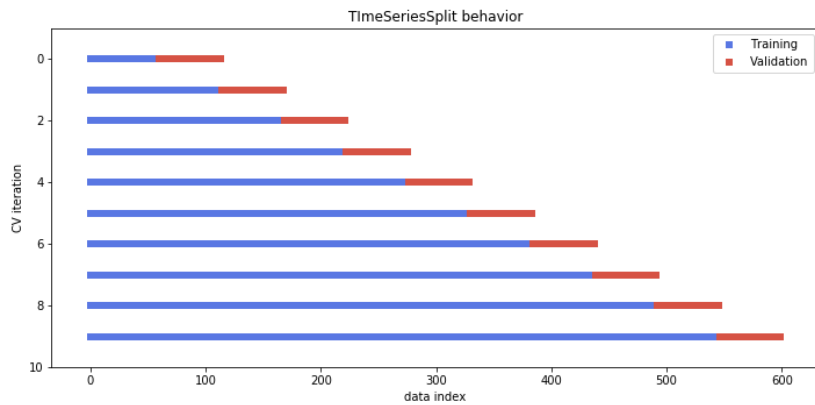


Figura 3.3: Visualización del comportamiento de TimeSeriesSplit. Las barras azules representan los datos de entrenamiento y las barras rojas representan los datos de validación para cada iteración de validación cruzada. Esta técnica de división asegura que los datos de validación siempre sean posteriores a los datos de entrenamiento, respetando así el orden temporal de los datos. **Fuente:** Datacamp.

Selección de hiperparámetros con 'GridSearchCV'

La técnica de 'GridSearchCV' (búsqueda de cuadrícula) permite evaluar una serie de combinaciones de hiperparámetros especificadas mediante un proceso de búsqueda exhaustiva. Cada combinación se evalúa utilizando validación cruzada, y se selecciona la combinación de hiperparámetros que minimiza el error absoluto medio (MAE).

Entrenamiento y predicción

Se entrenó un modelo para cada año y cada variable a predecir utilizando los mejores hiperparámetros encontrados en la búsqueda de cuadrícula. El modelo se entrenó exclusivamente con los datos del año anterior para ajustarse más a las situaciones reales de aplicación en el mercado, donde no es común disponer de datos de demanda de varios años. Se realizaron predicciones para el año siguiente y se almacenaron tanto los modelos optimizados como los escaladores utilizados. Los resultados de las predicciones se compararon con los datos reales del año correspondiente.

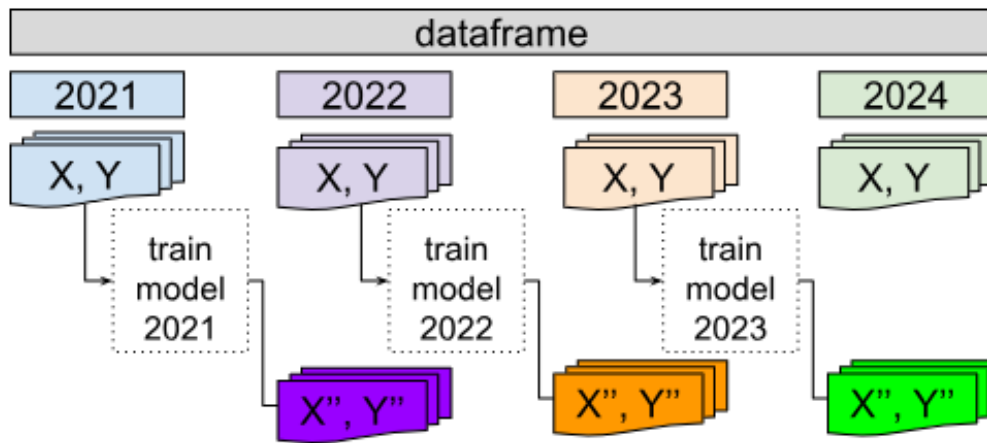


Figura 3.4: Esquema del proceso de entrenamiento y predicción anual. Cada modelo entrenado con los datos de un año (X, Y) y se utiliza para predecir las demandas y generación fotovoltaica del año siguiente (X'', Y''). Las predicciones de cada año se comparan con los datos reales del año predicho. **Fuente:** Esquema de elaboración propia.

Evaluación del modelo

Para evaluar la actuación del modelo y la calidad de sus predicciones se calcularon las siguientes métricas a partir de los datos reales de demanda y generación, y las predicciones: **MAE** (Mean Absolute Error), **MSE** (Mean Square Error) y **RMSE** (Root Mean Square Error). En la tabla 3.1 se muestran los resultados obtenidos para cada uno de los modelos.

Año	Variable	MAE (kWh)	MSE (kWh ²)	RMSE (kWh)
2022	BBA	8,14	133,45	11,55
	SEGA	4,42	32,47	5,70
	CM San Fernando	8,24	159,68	12,64
	Biblioteca	21,47	1081,28	32,88
	Generación (gen)	24,76	2778,31	52,71
2023	BBA	17,64	633,63	25,17
	SEGA	13,71	258,27	16,07
	CM San Fernando	8,52	195,70	13,99
	Biblioteca	21,12	907,02	30,12
	Generación (gen)	24,31	2435,28	49,35
2024	BBA	11,38	308,12	17,55
	SEGA	7,61	114,75	10,71
	CM San Fernando	11,29	342,69	18,51
	Biblioteca	12,48	353,40	18,80
	Generación (gen)	28,49	2996,42	54,74

Tabla 3.1: Resultados de las métricas de evaluación del modelo de predicción.

Además, se analizaron los resultados de cada modelo de forma gráfica. A continuación se muestra a modo de ejemplo las gráficas que fueron analizadas para el CUPS de la Facultad de BBA. Las gráficas para el resto de CUPS se pueden encontrar en el anexo A.

Análisis gráfico de las predicciones para BBAA

En todas las figuras presentadas en esta sección se sigue la siguiente estructura: distribución de la demanda de real (primera fila), de la demanda predicha (segunda fila) y del error relativo asociado a las predicciones (tercera fila).

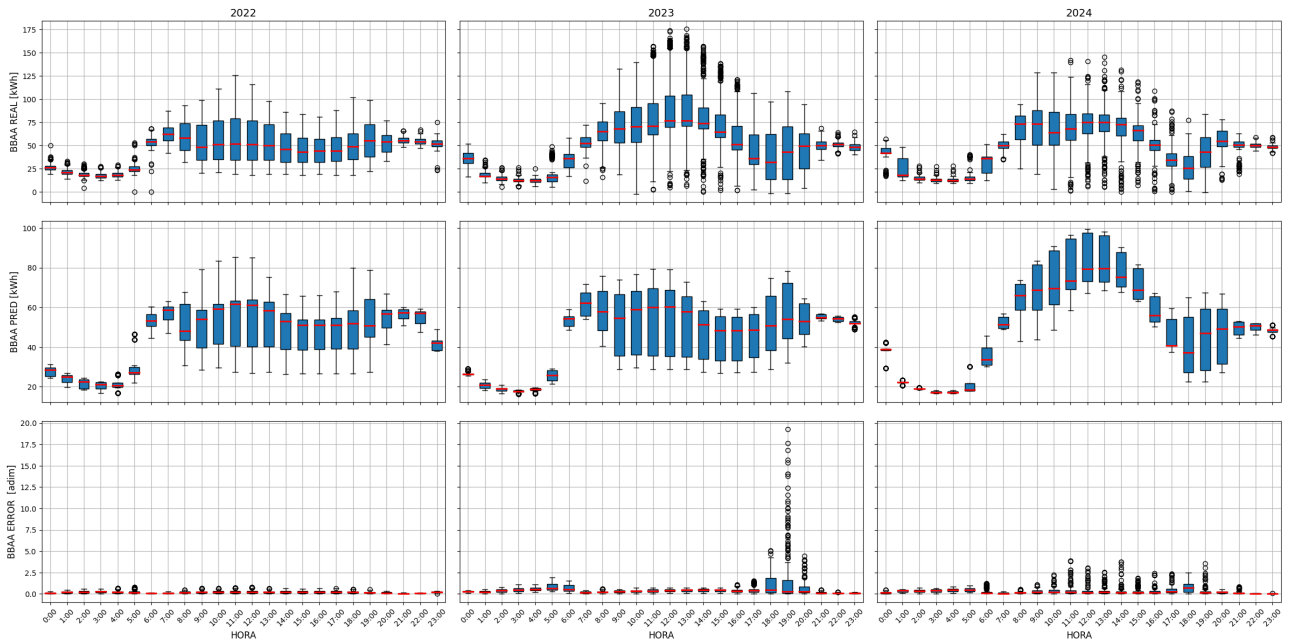


Figura 3.5: Comparación de la distribución horaria de la demanda real de BBAA y las predicciones.

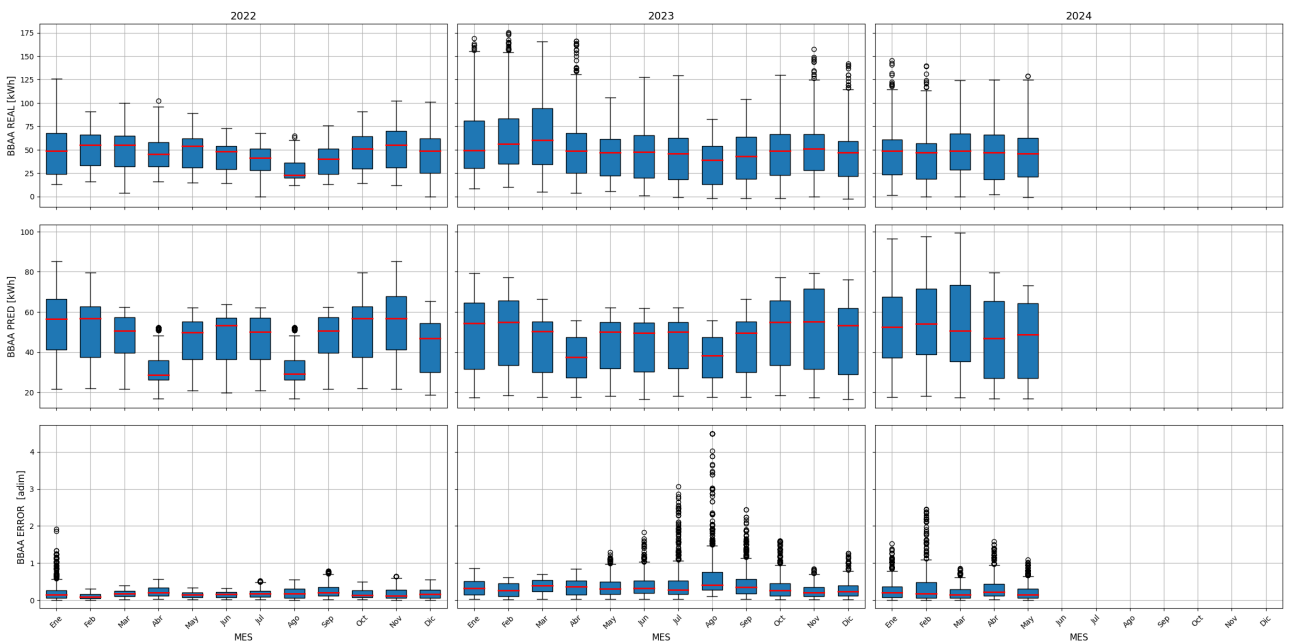


Figura 3.6: Comparación de la distribución mensual de la demanda real de BBAA y las predicciones.

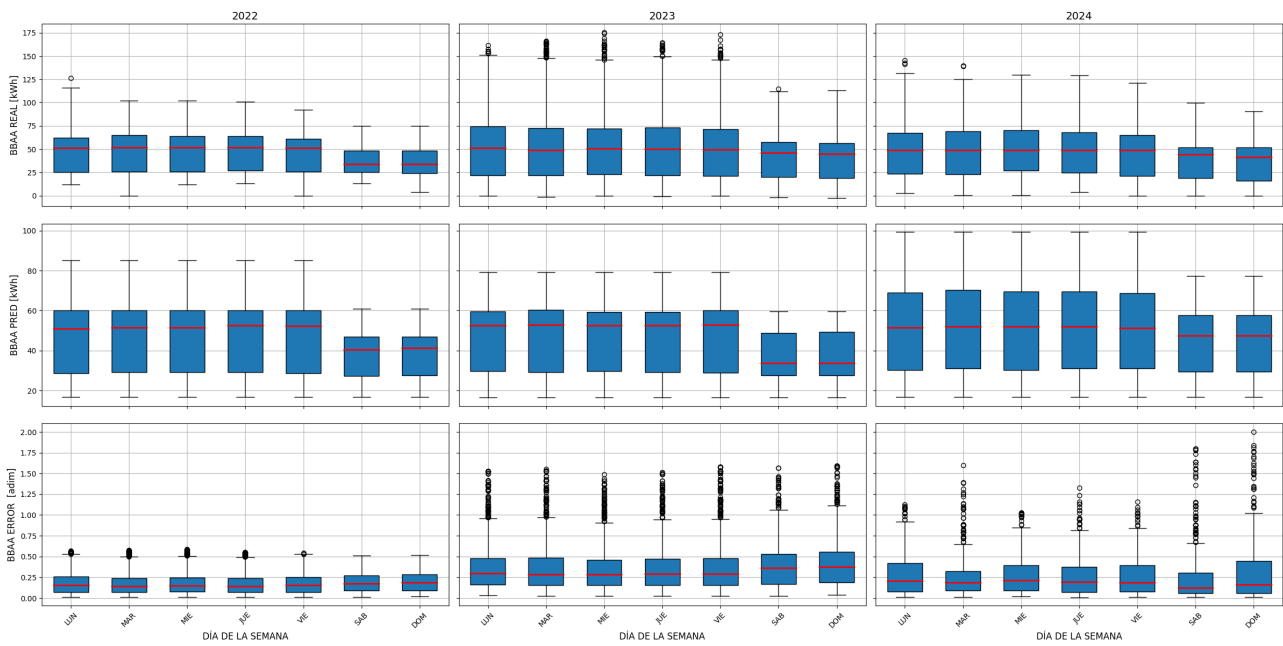


Figura 3.7: Comparación de la distribución horaria por día de la semana de la demanda real de BBAA y las predicciones.

En las gráficas presentadas (3.7, 3.6 y 3.5) se observan **errores relativos considerablemente bajos en comparación con los valores reales** (de uno o dos órdenes de diferencia). Esto avala la **calidad de las predicciones permitiendo su uso para la obtención de resultados.**

Capítulo 4

Obtención de resultados del reparto energético

Este capítulo se dividirá en dos secciones. En la primera, se describirán los cálculos realizados para obtener los resultados del reparto energético usando tanto los datos presentados en el capítulo 2 como los obtenidos siguiendo la metodología descrita en el capítulo 3. Dichos resultados serán posteriormente presentados en la segunda sección.

4.1. Cálculo de indicadores de reparto y excedentes

En esta sección se detalla el procedimiento para calcular el reparto energético y sus excedentes usando los coeficientes de reparto calculados según lo descrito en 3. Además, se incluirá una pequeña subsección en la que se definirá la nomenclatura que se usará en la presentación de los resultados en la sección 4.2.

4.1.1. Reparto energético

El reparto energético se realiza para asegurar que la energía generada se distribuye de manera eficiente y justa. Para este propósito, se utilizan los coeficientes de reparto previamente acordados entre los participantes del autoconsumo. A continuación, se detallan los pasos para realizar el reparto energético:

1. En cada hora h , se determina la cantidad de energía generada por el sistema fotovoltaico, denotada como G_h .
2. La energía generada se reparte entre los participantes utilizando los coeficientes de reparto¹ $C_{i,h}$ calculados según los métodos descritos en secciones anteriores.

$$E_{i,h}^A = C_{i,h} \times G_h \quad (4.1)$$

donde $E_{i,h}^A$ es la cantidad de energía asignada al participante i en la hora h .

4.1.2. Excedentes del reparto

Los excedentes de energía son la cantidad de energía generada que no se utiliza para satisfacer la demanda de los participantes del sistema de autoconsumo colectivo. Estos excedentes se pueden calcular tras el reparto.

¹Para los coeficientes fijos, se aplicará el mismo coeficiente para todas las horas del año: $C_{i,h} = C_i \quad \forall h$

1. Cálculo de los excedentes por CUPS y hora: se calculan los excedentes del reparto energético para cada CUPS en cada hora. Si la energía asignada es menor que la demandada $E_{i,h}^A < D_{i,h}$ se tiene que $EX_{i,h}^T = 0$. En el caso contrario, los excedentes para i en la hora h se calcularán con la siguiente fórmula:

$$EX_{i,h}^T = E_{i,h}^A - D_{i,h} \quad (4.2)$$

2. Cálculo de los excedentes totales por hora: se calcularán como la suma de los excedentes individuales:

$$EX_{total,h}^T = \sum_{i=1}^n EX_{i,h} \quad (4.3)$$

Estos excedentes se pueden clasificar en dos categorías: excedentes evitables y excedentes inevitables.

Excedentes inevitables

En nuestro sistema, se considerarán de esta categoría los excedentes de las horas en las que la generación fotovoltaica supere la suma de las demandas de todos los participantes. Estos excedentes ocurren cuando, a pesar de una gestión eficiente, la producción de energía es tan alta que excede la capacidad de consumo de los usuarios. Se calculan de la siguiente manera:

1. Determinación de la demanda total horaria: en cada hora h , se suma la demanda de todos los participantes para obtener la demanda total $D_{total,h}$:

$$D_{total,h} = \sum_{i=1}^n D_{i,h} \quad (4.4)$$

Donde n es el número total de participantes en el sistema de autoconsumo colectivo.

2. Comparación con la generación: se compara la demanda total $D_{total,h}$ con la generación fotovoltaica G_h . Si $G_h > D_{total,h}$ se calcula el excedente inevitable como:

$$EX_{total,h}^I = G_h - D_{total,h} \quad (4.5)$$

Donde $EX_{total,h}^I$ es el excedente de energía en la hora h .

Excedentes evitables

Los excedentes evitables son aquellos que se producen debido a un reparto energético imperfecto. Estos excedentes pueden ocurrir cuando la energía generada no se distribuye de manera óptima entre los participantes, por ejemplo, debido a errores en la predicción de la demanda o en la implementación de los coeficientes de reparto. Por ello, serán los que nos indiquen el método de reparto que mejor funciona. Para obtener estos excedentes se calcula la diferencia entre los excedentes totales y los inevitables:

$$EX_{total,h}^E = EX_{total,h}^T - EX_{total,h}^I \quad (4.6)$$

donde $EX_{total,h}^E$ es el excedente de energía evitable en la hora h .

4.1.3. Nomenclatura

Para presentar los resultados de manera clara, se ha establecido una nomenclatura específica para las modalidades de reparto analizadas. Esta nomenclatura se crea seleccionando una opción de cada una de las secciones de la Tabla 4.1. Primero, se selecciona un tipo de coeficiente (C, Cepsilon, o Cpearson), que representan respectivamente los coeficientes de cálculo tradicional, coeficientes de reparto justo, y coeficientes calculados con correlaciones de Pearson. Luego, se determina el tipo (v o f), indicando si los coeficientes son variables o fijos anuales. Finalmente, se especifica la modalidad de cálculo, donde se puede optar por coeficientes calculados con datos del año anterior o con predicciones, representado por 'p'. Así, el nombre completo se forma combinando estas tres elecciones, proporcionando una identificación clara y específica para cada modalidad de reparto.

NOMBRE COEFICIENTE
C
Cepsilon
Cpearson

(a) Nombre coeficiente

TIPO
v
f

(b) Tipo

MODALIDAD DE CÁLCULO
-
p

(c) Modalidad de cálculo

Tabla 4.1: Nomenclatura para las modalidades de reparto.

4.2. Resultados del reparto y excedentes

En este capítulo se presentan los resultados del reparto energético y los excedentes obtenidos con cada uno de los métodos de cálculo de coeficientes detallados en el capítulo 3. Los resultados intermedios de las predicciones y del cálculo de los coeficientes se pueden consultar en el apéndice A.

4.3. Resultados globales

A continuación se exponen los resultados del cálculo de los excedentes obtenidos con los distintos métodos de reparto para cada uno de los periodos contemplados. No se hace distinción por CUPS.

Tabla 4.2: Excedentes totales y evitables por año [MWh]

Modalidad de reparto	Excedentes totales			Excedentes evitables		
	2022	2023	2024	2022	2023	2024
C_v_p	160,53	223,44	103,44	8,12	37,21	5,32
C_f_p	167,57	226,45	104,74	15,16	40,23	6,67
Cepsilon_v_p	163,99	228,79	104,54	11,58	42,56	6,41
Cepsilon_f_p	166,23	228,45	110,65	22,32	43,86	6,14
Cpearson_f_p	194,03	249,60	113,35	41,62	63,58	15,23
C_v	170,47	237,54	113,22	18,06	51,31	15,10
C_f	167,34	226,44	104,74	15,16	40,23	6,67
Cepsilon_v	175,47	240,80	113,55	23,05	54,85	9,00
Cepsilon_f	166,18	227,35	111,42	12,74	42,34	6,41
Cpearson_f	193,88	244,69	112,00	41,47	54,86	13,88

Las siguientes gráficas muestran los valores totales de excedentes evitables en MWh para diferentes modalidades de reparto. Se incluyen los porcentajes relativos respecto al valor más bajo para proporcionar una mejor comparación visual entre las modalidades.

Figura 4.1: Gráfico de excedentes evitables en el año 2022.

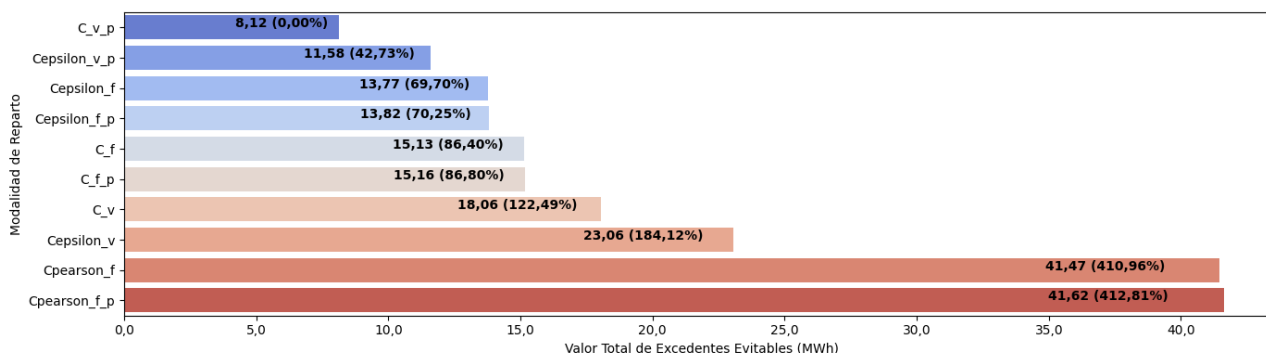


Figura 4.2: Gráfico de excedentes evitables en el año 2023.

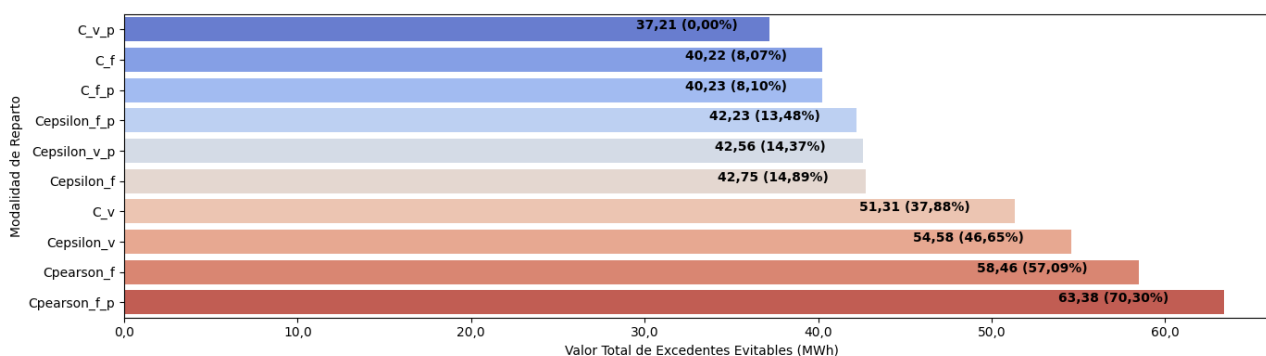
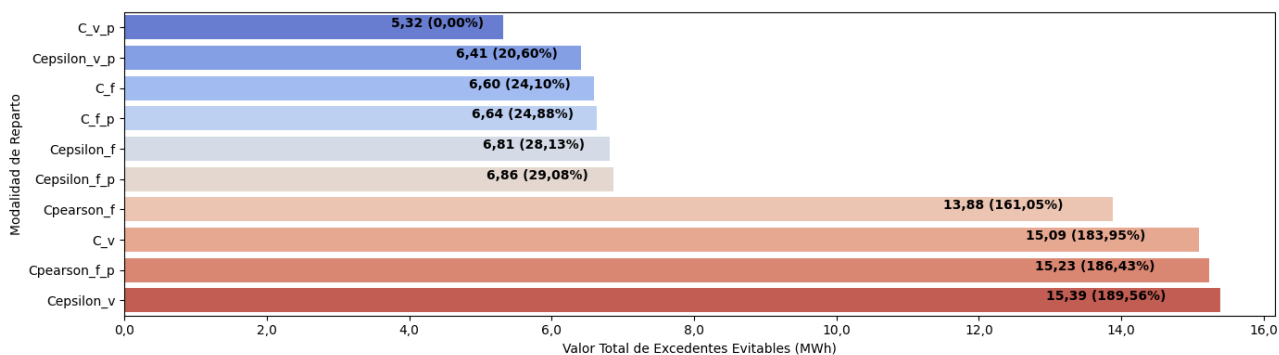


Figura 4.3: Gráfico de excedentes evitables en el año 2024.



Para el método C, la mejor opción en todos los años es **C_v_p**, mostrando consistentemente los valores más bajos de excedentes evitables. La comparación de este método con su modalidad **C_v** y su modalidad fija, **C_f_p**, muestra mejoras significativas en todos los años, siendo especialmente notable la reducción de excedentes evitables en el año 2024, donde la reducción respecto a **C_v** es del 64,75 %.

Para método Cepsilon, la mejor modalidad en todos los años es **Cepsilon_v_p**, mostrando consistentemente menores valores de excedentes evitables. La peor modalidad varía cada año pero siempre muestra un aumento significativo respecto a la mejor, especialmente en el año 2024 donde **Cepsilon_v** tiene un 140,09 % más de excedentes evitables que **Cepsilon_v_p**.

Por otro lado, en el caso de **Cpearson**, la mejor modalidad varía entre **Cpearson_f** y **Cpearson_f_p** dependiendo del año. Las diferencias entre la mejor y la peor modalidad son menores en comparación con Cepsilon, mostrando una mayor consistencia entre las variantes de **Cpearson**. La diferencia más notable se observa en el año 2023, donde la peor modalidad tiene un 7,76 % menos de excedentes evitables que la mejor.

En resumen, en los tres años analizados (2022, 2023 y 2024), el método **C_v_p** resulta ser el que produce menos excedentes evitables. Estas observaciones sugieren que el método **C_v_p** es el más eficiente de los estudiados en la minimización de los excedentes globales evitables.

4.4. Resultados por CUPS

En la siguiente tabla se recogen las mejores modalidades de reparto con sus excedentes asociados para cada CUPS en los años 2022, 2023 y 2024.

Tabla 4.3: Mejor modalidad de reparto y sus excedentes (en MWh) por CUPS y año.

CUPS	Mejor modalidad			Excedentes		
	2022	2023	2024	2022	2023	2024
BBAA	Cepsilon_f_p	C_v_p	C_f	32,34	11,34	22,48
SEGA I	Cepsilon_f_p	Cepsilon_f_p	C_v_p	23,70	48,23	21,88
C.M. San Fernando	C_v_p	Cv_f_p	C_v_p	20,56	40,90	13,89
Biblioteca	Cpearson_f_p	Cpearson_f_p	Cpearson_f_p	33,19	51,89	16,34

En las siguientes gráficas se muestran la participación de cada uno de los CUPS en los excedentes totales (en MWh) obtenidos con las distintas modalidades de reparto para los años 2022, 2023 y 2024. Los excedentes están representados por barras horizontales apiladas, donde cada color corresponde a CUPS.

Figura 4.4: Excedentes totales por modalidad de reparto y CUPS para el año 2022.

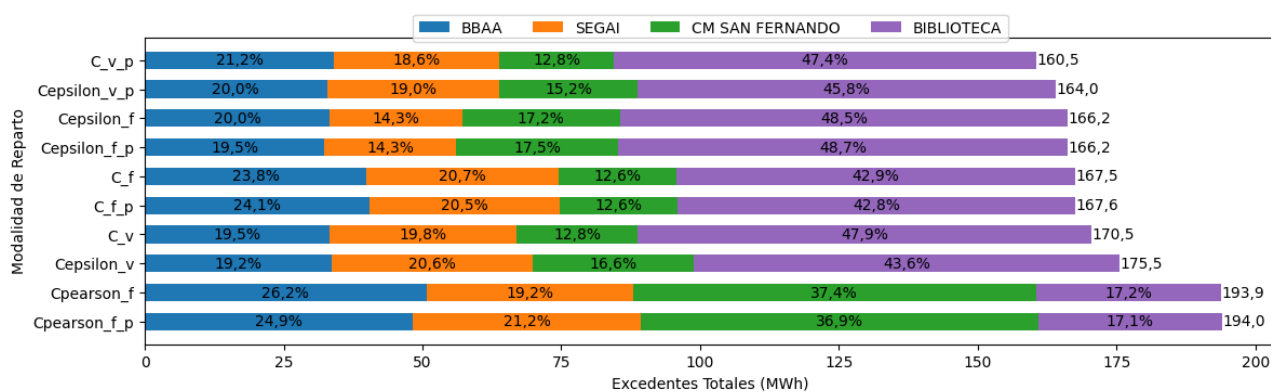


Figura 4.5: Excedentes totales por modalidad de reparto y CUPS para el año 2023.

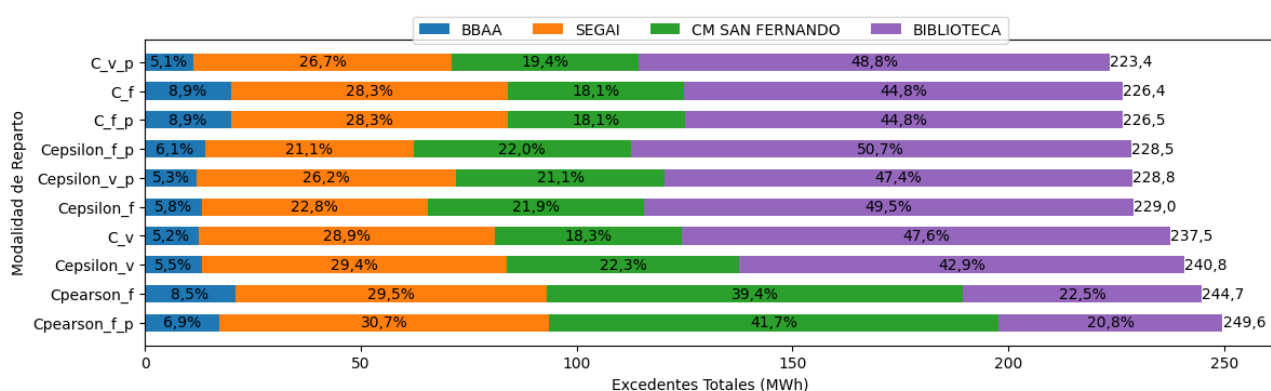
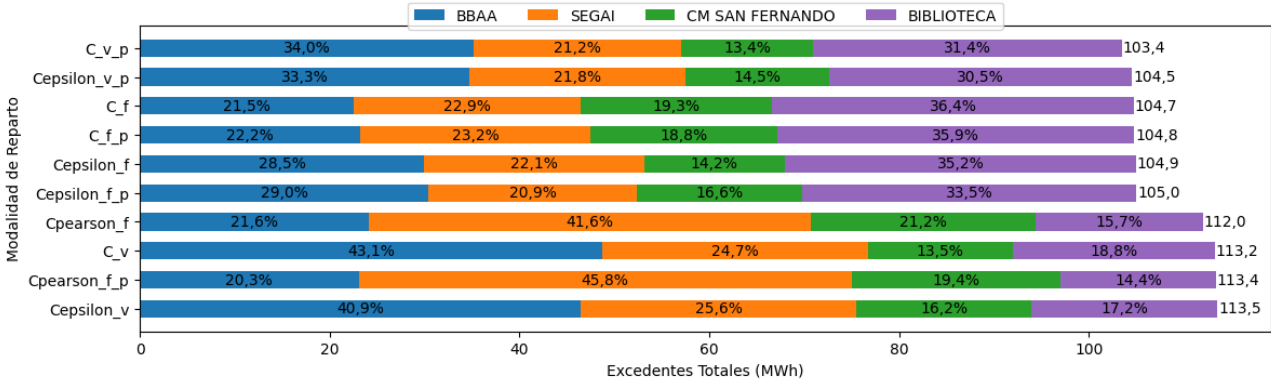


Figura 4.6: Excedentes totales por modalidad de reparto y CUPS para el año 2024.



El análisis de los excedentes por método y CUPS a lo largo de los años muestra que la Biblioteca fue consistentemente el mayor contribuyente en la mayoría de los métodos, especialmente en los años 2022 y 2023, mientras que la Facultad de BBA ganó protagonismo en 2024. Por otro lado, el C.M. San Fernando aunque presentó menores aportaciones en varios métodos, tuvo una notable presencia en los métodos Cpearson_f y Cpearson_f_p.

Capítulo 5

Discusión, conclusiones y líneas futuras

En este capítulo se discutirán los resultados obtenidos expuestos en 4. Además se presentarán las conclusiones finales de este trabajo y se propondrán interesantes líneas futuras para la continuación de las investigaciones realizadas.

5.1. Discusión

En esta sección se analiza y discute los resultados obtenidos en la investigación y presentados en capítulo 4 de este trabajo. La discusión se estructura en torno a los siguientes aspectos: eficiencia de los métodos de reparto energético, análisis por CUPS, comparación con estudios previos y limitaciones del estudio.

5.1.1. Eficiencia de los métodos de reparto energético

El método de coeficientes variables con predicciones (C_{v_p}) ha demostrado ser el mejor en la reducción de los excedentes evitables en todos los años del estudio. Sin embargo, es crucial evaluar si la mejora en los resultados justifica el uso de este método considerando el significativo aumento en el coste y complejidad computacional. En general, es necesario sopesar si los recursos adicionales necesarios para implementar predicciones y coeficientes variables realmente valen la pena.

5.1.2. Análisis por CUPS

El análisis de los datos por CUPS ha revelado una notable variabilidad en los resultados. Por ejemplo, la Facultad de Bellas Artes y el SEGAI mostraron diferentes patrones de demanda que influyeron significativamente en los excedentes generados. Estos resultados sugieren que los patrones de demanda específicos de cada CUPS son cruciales al diseñar estrategias de reparto energético. De esta forma se plantea la duda de si es viable generalizar los resultados para todos los sistemas de autoconsumo colectivo, o si sería más conveniente llevar a cabo un estudio previo de cada sistema para clasificarlos y, a partir de esta clasificación, elegir el método de reparto que mejor se ajuste a sus características específicas.

5.1.3. Comparación con estudios previos

Si bien se han encontrado artículos como (8), (13) y (14) que mencionan algunos de los métodos utilizados en este trabajo, no se ha encontrado ninguno que los aplique a datos reales como en este estudio. La ausencia de estudios comparativos similares en la literatura existente resalta la importancia de ampliar el alcance de este trabajo para obtener resultados más concluyentes y generalizables como una base sólida para futuras investigaciones. Esto ayudaría a establecer directrices más efectivas para la implementación de sistemas de reparto energético en diversos contextos.

5.1.4. Limitaciones del estudio

El dataset utilizado en este estudio estaba limitado a cuatro curvas de consumo de los CUPS de la Universidad de La Laguna. Aunque estos datos proporcionan una base inicial para la investigación, es posible que no sean las más adecuadas para generalizar los resultados a otros tipos de autoconsumo colectivo. Hubiera sido más interesante combinar demandas de viviendas, oficinas y otros tipos de edificaciones que presentan curvas de demanda muy diferentes tanto en distribución como en media. Esta diversidad podría haber ofrecido una visión más amplia y aplicable a diferentes contextos de autoconsumo colectivo.

5.2. Conclusiones

Este estudio ha evaluado la eficacia de diferentes métodos de asignación de coeficientes de reparto de energía en sistemas de autoconsumo colectivo, enfocándose en la minimización de los excedentes energéticos y la optimización del uso de los recursos generados por plantas fotovoltaicas.

Una de las conclusiones más destacadas es que el método de coeficientes variables con predicciones (C_v_p) ha demostrado ser el mejor en la reducción de los excedentes evitables en todos los años del estudio, con un aumento promedio de excedentes del 34,51 % para el segundo mejor método de cada año. Por otro lado, el análisis por CUPS ha revelado una variabilidad significativa en los resultados obtenidos para los distintos escenarios. En general, la Biblioteca y la Facultad de Bellas Artes fueron los principales contribuyentes a los excedentes en distintos años, lo que sugiere que los patrones de demanda específicos de cada participante son cruciales al diseñar estrategias de reparto energético. Por este motivo, los resultados no se puede generalizar para otros autoconsumos colectivos con patrones de demanda diferentes a los estudiados.

Sin embargo, también se ha identificado que los métodos dinámicos, aunque más eficientes, implican un mayor coste computacional y una complejidad adicional en su implementación. Este incremento en los recursos necesarios, tanto en términos de hardware como de gestión humana, debe ser considerado al decidir qué método adoptar. La evaluación de la eficiencia debe, por tanto, incluir no solo los beneficios en la reducción de excedentes sino también los recursos adicionales que estos métodos requieren.

Otra conclusión relevante es que, aunque se podía intuir a priori que los métodos variables funcionarían mejor que los fijos y que los métodos con predicciones serían superiores

a los calculados con los datos del año anterior, este estudio no confirma esta hipótesis de manera concluyente. Si bien el mejor método respalda esta idea, no se aprecia ningún patrón consistente para el resto de los casos. Es decir, aunque el método de coeficientes variables con predicciones mostró la mayor eficiencia, los resultados para otros métodos no siguen una tendencia clara que permita generalizar el resultado.

Aunque el estudio se ha llevado a cabo con datos reales, similares a los que se usarían en una aplicación práctica, proporcionando una primera aproximación valiosa para guiar la investigación, es necesario ampliar su alcance. Un aspecto positivo de las investigaciones realizadas es que se han desarrollado e implementado algoritmos efectivos y generalizables para el cálculo de los coeficientes y la distribución de la energía, lo que facilita su aplicación en otros sistemas de autoconsumo colectivo.

En definitiva, es fundamental realizar investigaciones adicionales en diversos sistemas de autoconsumo colectivo, donde los miembros presenten diferentes curvas de demanda. Aplicar estos métodos a un número representativo de sistemas permitiría obtener resultados más concluyentes y generalizables.

5.3. Líneas Futuras

A partir de las conclusiones obtenidas en este estudio, se plantean varias líneas futuras de investigación para seguir mejorando la eficiencia de los sistemas de autoconsumo colectivo y ampliar el alcance de los resultados obtenidos.

En primer lugar, es crucial realizar investigaciones adicionales en diversos sistemas de autoconsumo colectivo, donde los miembros presenten diferentes curvas de demanda. Aplicar estos métodos a un número representativo de sistemas permitirá obtener resultados más concluyentes y generalizables, y ayudará a comprender mejor cómo los patrones de demanda específicos de cada participante afectan la eficiencia de los distintos métodos de reparto energético.

Para continuar avanzando en la optimización de sistemas de autoconsumo colectivo, se propone desarrollar y probar sistemas que permitan la implementación de coeficientes de reparto dinámicos en tiempo real, utilizando técnicas avanzadas de monitoreo y predicción. La creación de algoritmos capaces de ajustarse automáticamente a las variaciones de demanda y generación será esencial para mejorar la eficiencia de estos sistemas.

Por último, se propone una línea de investigación prometedora y cada vez más relevante: la integración de sistemas de almacenamiento de energía, como baterías, para gestionar los excedentes energéticos. El uso de baterías es, a día de hoy, una de las mejores herramientas capaces de equilibrar las variaciones de demanda y generación al almacenar el exceso de energía producido para su uso posterior. Esto no solo reduce los excedentes energéticos, sino que también mejora la eficiencia general del sistema, ofreciendo así una solución más sostenible y efectiva.

Capítulo 6

Summary and conclusions

6.1. Summary

This Master's Thesis investigates the effectiveness of various methods for allocating energy distribution coefficients in collective self-consumption systems, focusing on minimizing energy surpluses and optimizing the use of resources generated by photovoltaic plants. The study employs both variable and fixed methods, utilizing real demand data as well as generation data downloaded from the PVGIS API.

To achieve this, generalizable algorithms were developed and implemented for calculating the coefficients for each distribution method. For predictive purposes, the XGBoost algorithm was used. The analysis of these methods involved calculating the avoidable surpluses in each case, which was the key metric targeted for minimization.

The results indicate that the method of variable coefficients with predictions (C_v_p) achieved the best performance in reducing avoidable surpluses. However, the study concludes that this result is not generalizable. Significant variability was observed in the results across different CUPS, suggesting that specific demand patterns heavily influence the effectiveness of each energy distribution strategy. Despite the superior performance of the C v p method, its general applicability across diverse collective self-consumption systems remains uncertain, underscoring the need for further research in this area.

6.2. Conclusions

This study has evaluated the effectiveness of different methods for allocating energy distribution coefficients in collective self-consumption systems, focusing on minimizing energy surpluses and optimizing the use of resources generated by photovoltaic plants.

One of the most notable conclusions is that the method of variable coefficients with predictions (C_v_p) has proven to be the best in reducing avoidable surpluses in all the years of the study, with an average increase in surpluses of 34.51 % for the second-best method each year. On the other hand, the analysis by CUPS revealed significant variability in the results obtained for the different scenarios. In general, the Library and the Faculty of Fine Arts were the main contributors to surpluses in different years, suggesting that the specific demand patterns of each participant are crucial when designing energy distri-

bution strategies. For this reason, the results cannot be generalized to other collective self-consumption systems with different demand patterns than those studied.

However, it has also been identified that dynamic methods, although more efficient, entail higher computational costs and additional complexity in their implementation. This increase in the required resources, both in terms of hardware and human management, must be considered when deciding which method to adopt. Therefore, the evaluation of efficiency should include not only the benefits in reducing surpluses but also the additional resources these methods require.

Another relevant conclusion is that although it could be intuitively assumed that variable methods would perform better than fixed ones and that methods with predictions would be superior to those calculated with the previous year's data, this study does not conclusively confirm this hypothesis. While the best method supports this idea, no consistent pattern is observed for the other cases. That is, although the method of variable coefficients with predictions showed the highest efficiency, the results for other methods do not follow a clear trend that allows generalizing the outcome.

Although the study has been conducted with real data, similar to what would be used in a practical application, providing a valuable first approximation to guide the research, it is necessary to broaden its scope. A positive aspect of the conducted research is that effective algorithms have been developed and implemented for the calculation of coefficients and the distribution of energy, facilitating their application in other collective self-consumption systems.

In conclusion, it is essential to conduct additional research in various collective self-consumption systems, where members present different demand curves. Applying these methods to a representative number of systems would yield more conclusive and generalizable results.

Apéndice A

Archivos, tablas y gráficas de los resultados

En este apéndice se presentan por un lado, los enlaces a los archivos con extensión '.csv' con los datos usados en este trabajo. Además, se ofrecerán tablas y gráficas relacionadas con el modelo de predicción. Estas serán útiles para consultar de forma detallada la calidad de las predicciones.

A.1. Enlaces a los archivos con datos

Estos archivos contienen toda la información necesaria para replicar y verificar los resultados presentados en el capítulo 4.

- Datos de generación fotovoltaica descargados de PVGIS con procesamiento
- Datos de la demanda energética de los participantes del autoconsumo colectivo (con la generación fotovoltaica total)
- Resultados de las predicciones de demanda y generación para los años 2022, 2023 y 2024.
- Coeficientes de reparto para cada una de las modalidades estudiadas.

A.2. Análisis de los resultados de las predicciones

En esta sección se presentarán algunas tablas y gráficas útiles para analizar la calidad de las predicciones.

Año	Variable	MAE (kWh)
2022	BBA	12,00
	SEGA	5,06
	CM San Fernando	8,00
	Biblioteca	22,26
	Generación (gen)	21,34
2023	BBA	9,13
	SEGA	4,71
	CM San Fernando	9,41
	Biblioteca	18,95
	Generación (gen)	28,26
2024	BBA	15,16
	SEGA	9,18
	CM San Fernando	9,44
	Biblioteca	18,73
	Generación (gen)	28,34

Tabla A.1: Resultados de la validación cruzada (MAE) para cada modelo entrenado.

Año	Variable	P1	P2	P3	P4	P5
2022	BBA	0,8	0,01	9	300	0,8
	SEGA	0,8	0,01	9	100	1,0
	CM San Fernando	1,0	0,1	9	100	0,8
	Biblioteca	1,0	0,1	3	100	1,0
	Generación (gen)	0,8	0,1	9	200	0,8
2023	BBA	0,8	0,05	6	100	0,8
	SEGA	1,0	0,05	3	300	1,0
	CM San Fernando	0,8	0,01	3	300	0,8
	Biblioteca	0,8	0,05	3	100	0,8
	Generación (gen)	0,8	0,1	9	100	0,8
2024	BBA	0,8	0,01	6	200	0,8
	SEGA	1,0	0,01	3	100	1,0
	CM San Fernando	0,8	0,05	3	200	0,8
	Biblioteca	1,0	0,01	6	200	0,8
	Generación (gen)	1,0	0,05	3	100	1,0

Tabla A.2: Mejores Parámetros para cada modelo. Los parámetros optimizados son P1: Colsample_bytree, P2: Learning_Rate, P3: Max_Depth, P4: N_Estimators, P5: Subsample.

A.2.1. Análisis gráfico de las predicciones

En todas las figuras presentadas en esta sección se sigue la siguiente estructura: Distribución de la demanda de real (primera fila), de la demanda predicha (segunda fila) y del error relativo asociado a las predicciones (tercera fila). En todos los casos se observan errores relativos considerablemente bajos en comparación a los valores reales, llegando incluso a diferenciarse en varios orden de magnitud.

BBAA

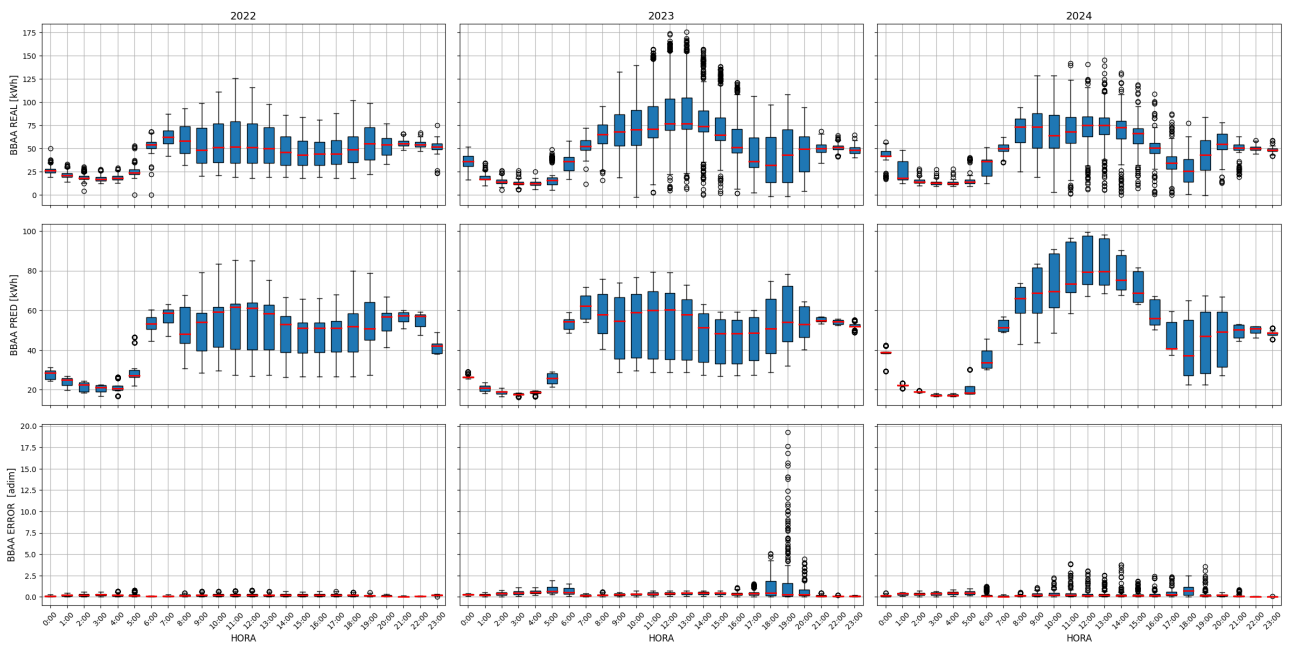


Figura A.1: Comparación de la distribución horaria de la demanda real de BBAA y las predicciones.

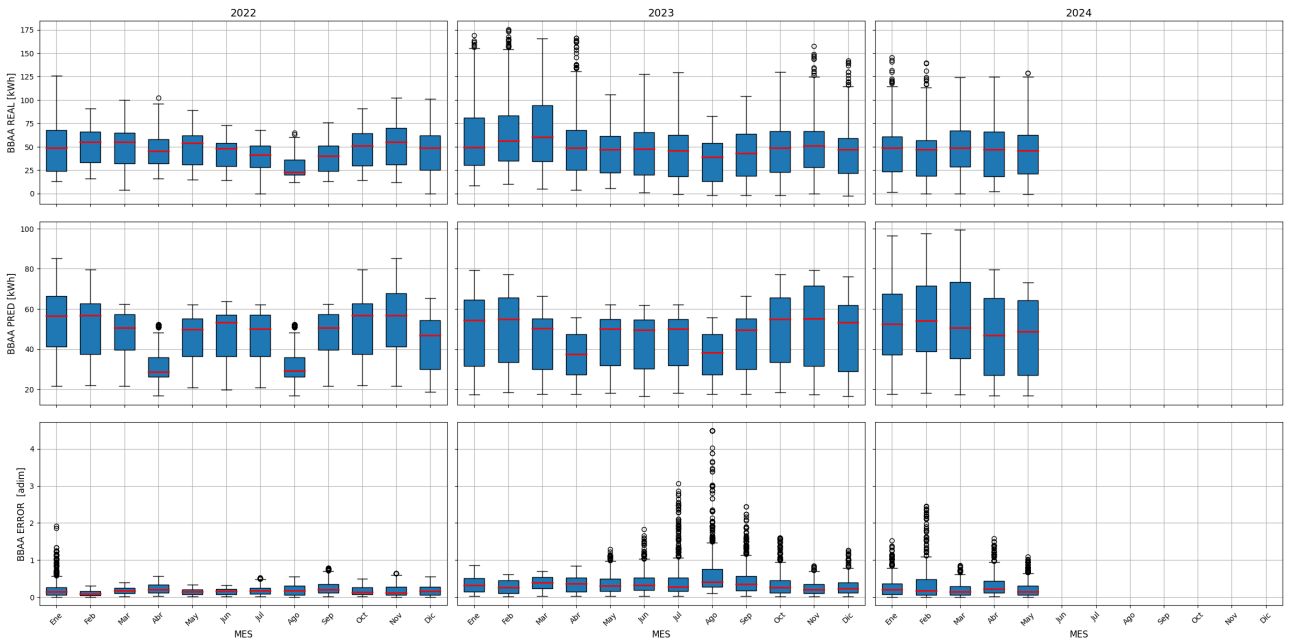


Figura A.2: Comparación de la distribución mensual de la demanda real de BBAA y las predicciones.

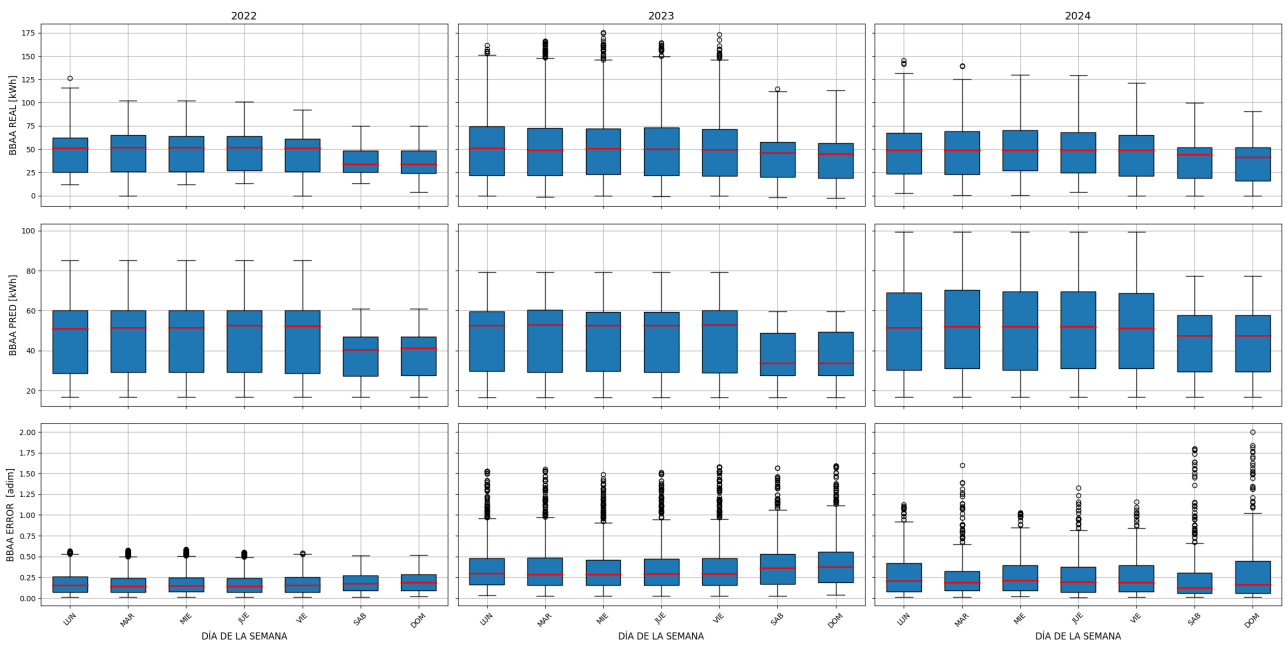


Figura A.3: Comparación de la distribución horaria por día de la semana de la demanda real de BBA y las predicciones.

SEGAI

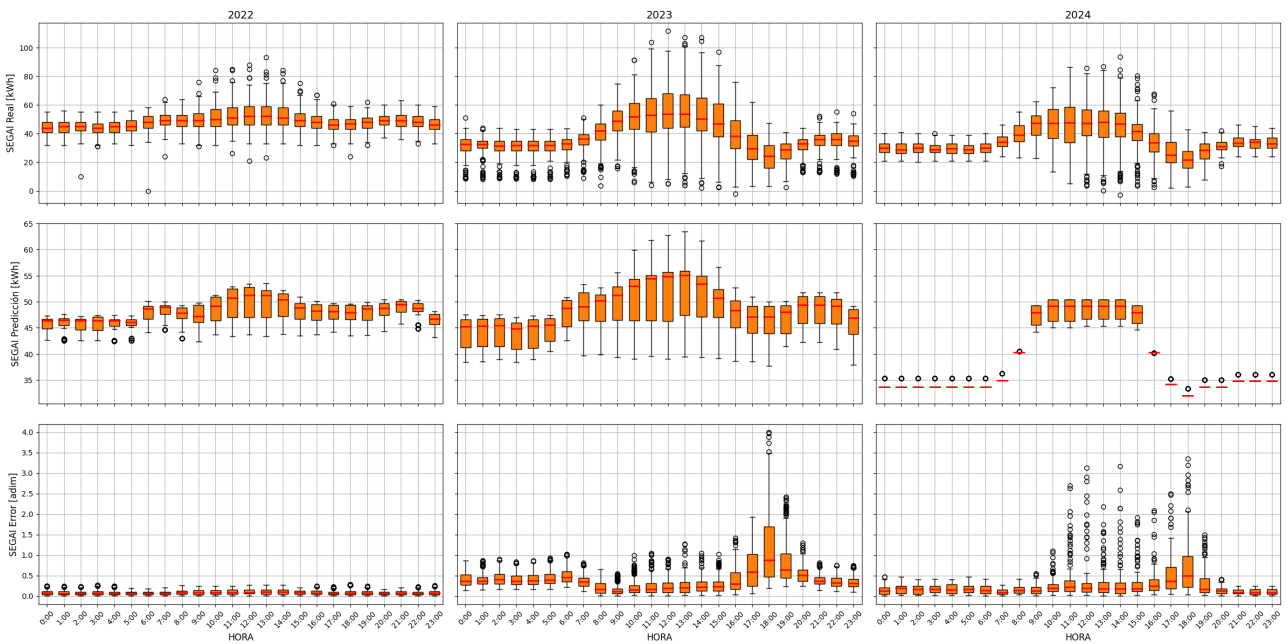


Figura A.4: Comparación de la distribución horaria de la demanda real de SEGAI y las predicciones.

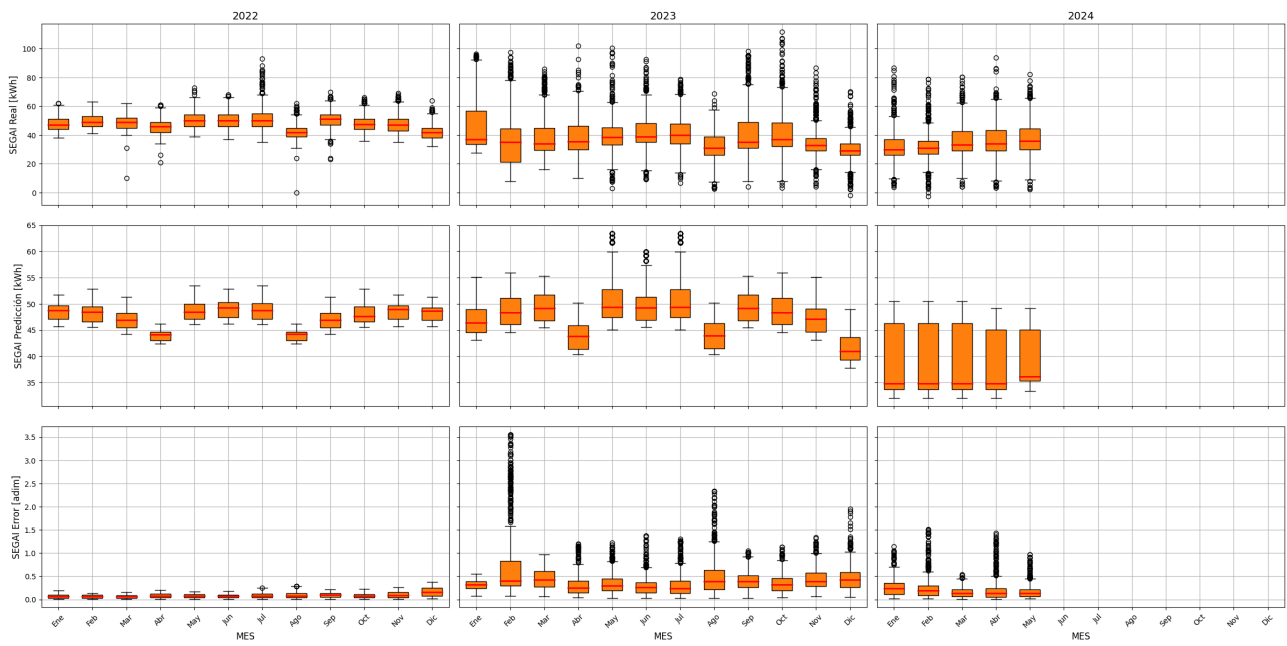


Figura A.5: Comparación de la distribución mensual de la demanda real de SEGAI y las predicciones.

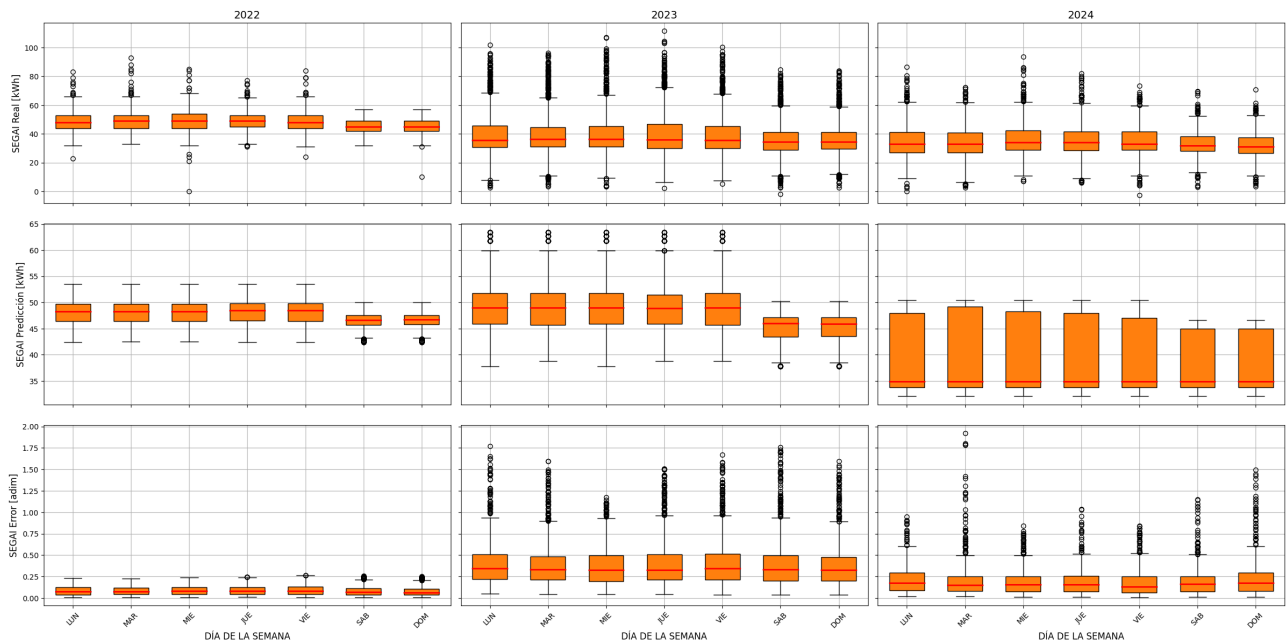


Figura A.6: Comparación de la distribución horaria por día de la semana de la demanda real de SEGAI y las predicciones.

C.M. San Fernando

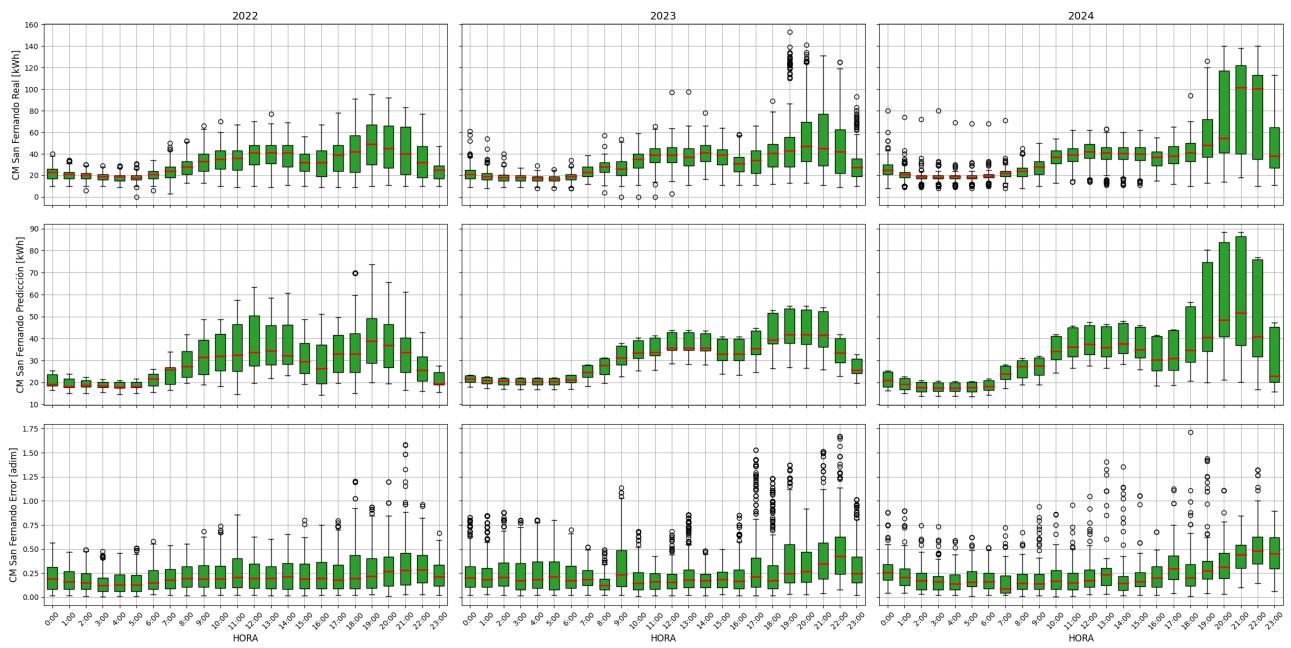


Figura A.7: Comparación de la distribución horaria de la demanda real de C.M. San Fernando y las predicciones.

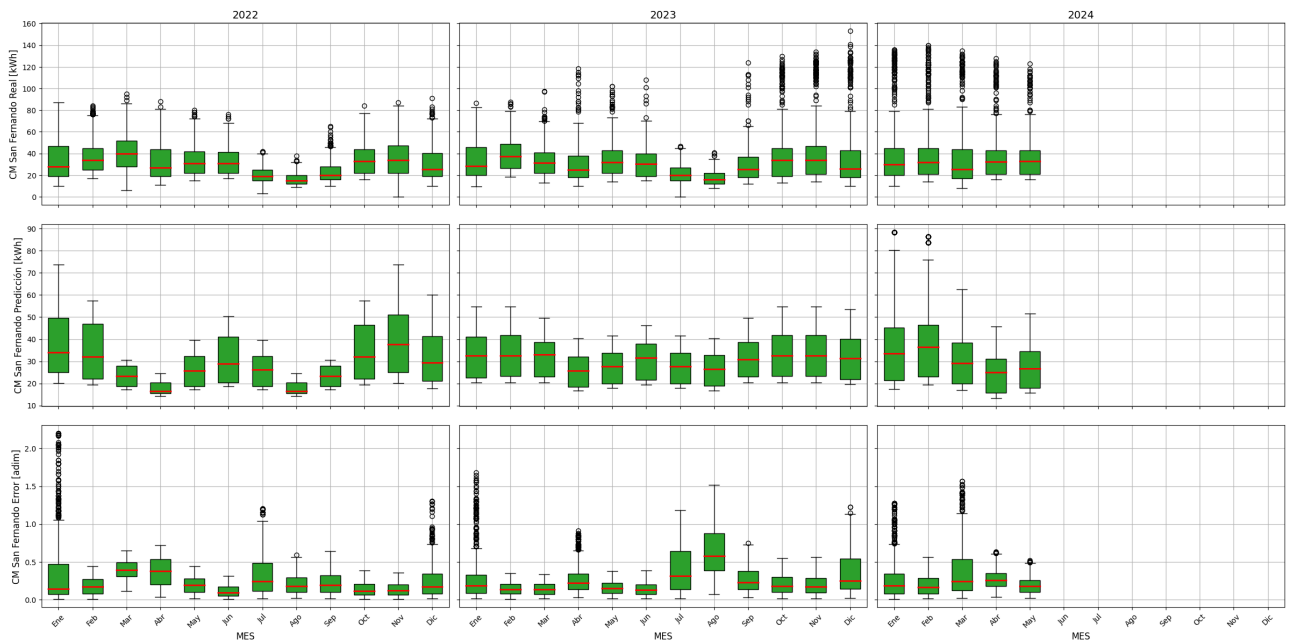


Figura A.8: Comparación de la distribución mensual de la demanda real de C.M. San Fernando y las predicciones.

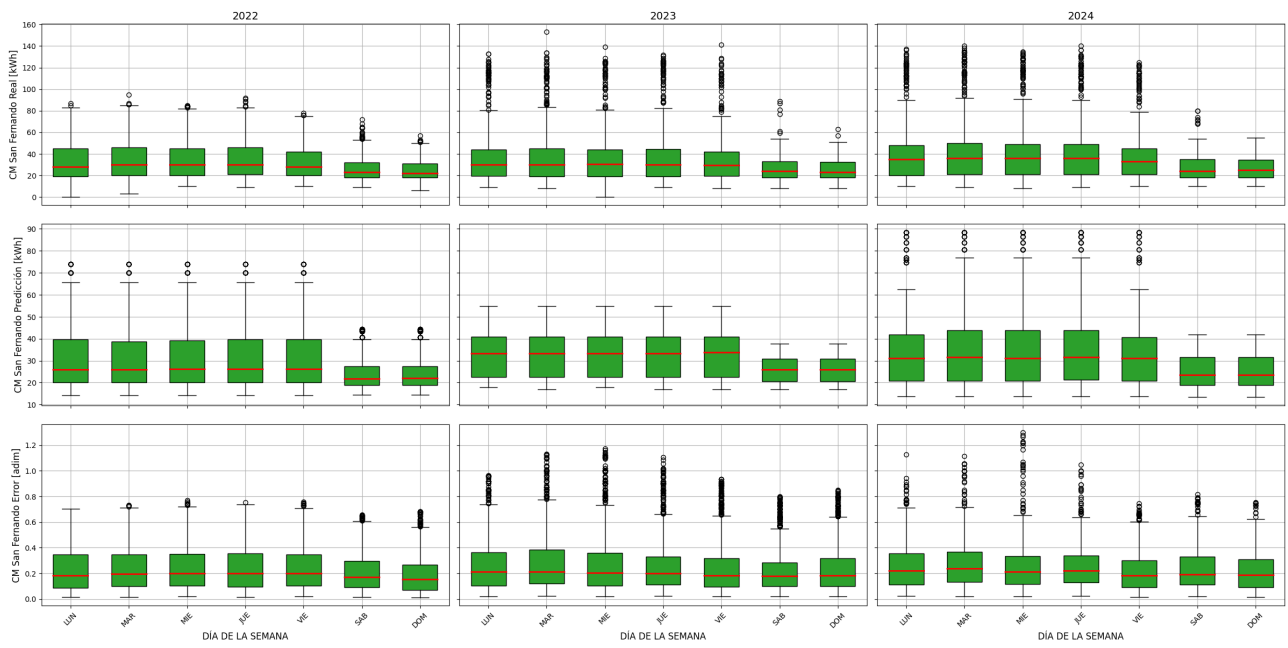


Figura A.9: Comparación de la distribución horaria por día de la semana de la demanda real de C.M. San Fernando y las predicciones.

Biblioteca

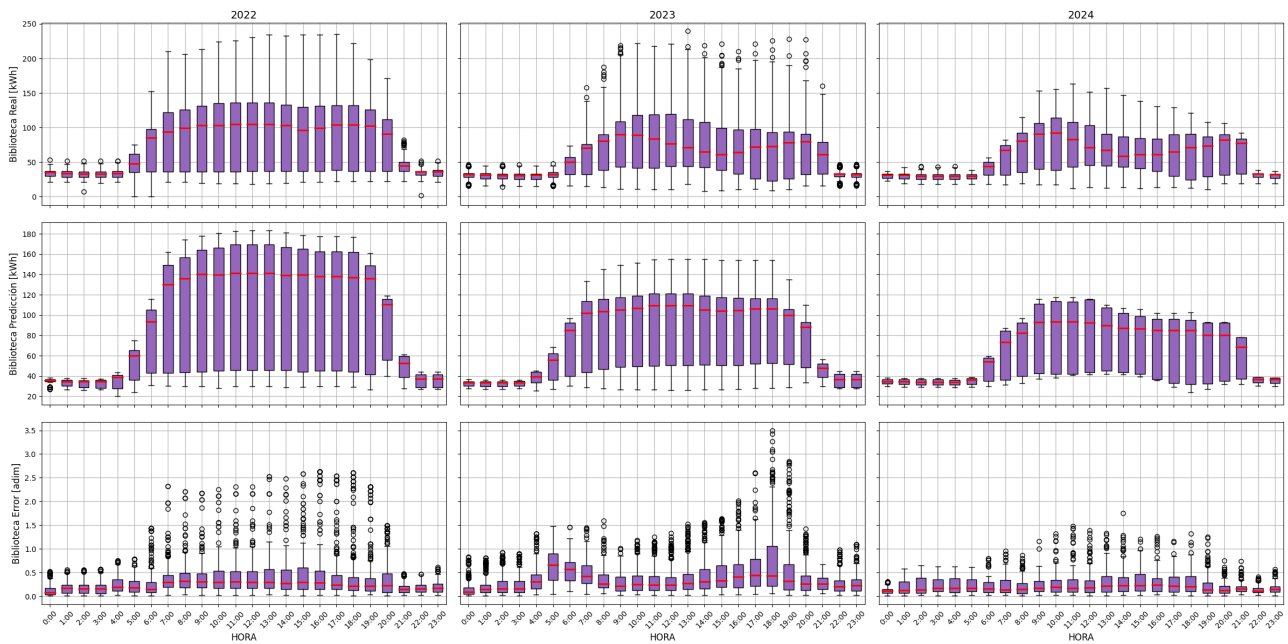


Figura A.10: Comparación de la distribución horaria de la demanda real de Biblioteca y las predicciones.

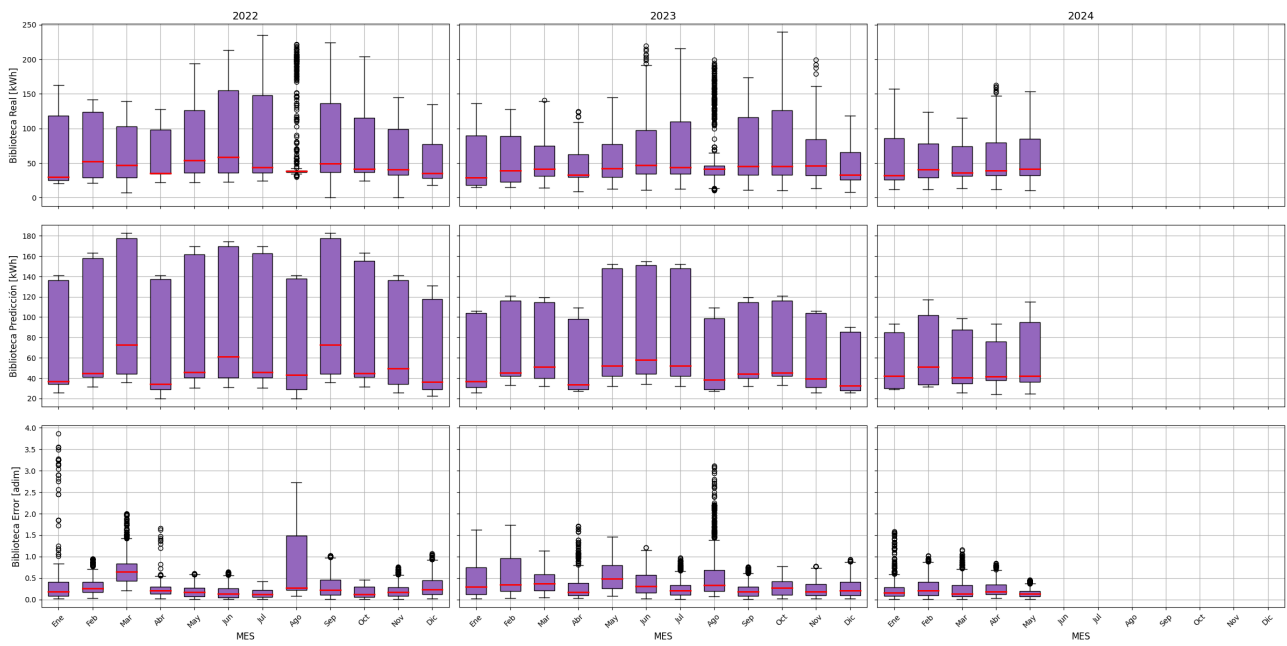


Figura A.11: Comparación de la distribución mensual de la demanda real de la biblioteca y las predicciones.

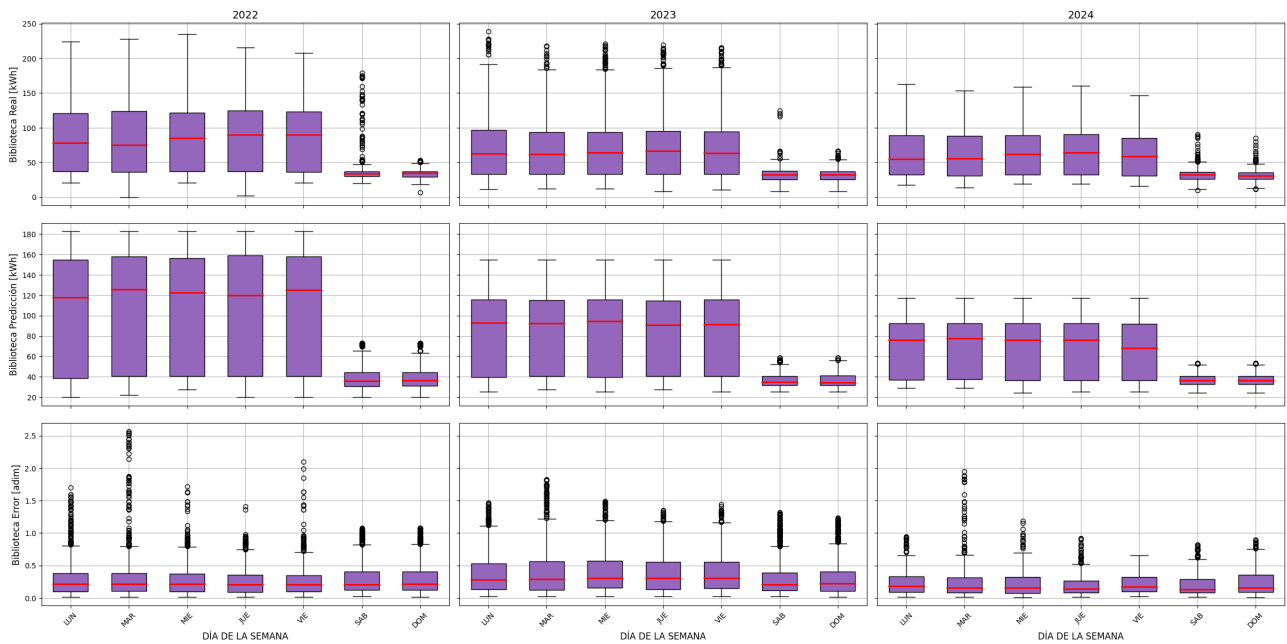


Figura A.12: Comparación de la distribución horaria por día de la semana de la demanda real de la Biblioteca y las predicciones.

Bibliografía

- [1] “BOE-A-2013-13645 ley 24/2013, de 26 de diciembre, del sector eléctrico.” <https://www.boe.es/buscar/act.php?id=BOE-A-2013-13645>. Accessed: 2024-5-24.
- [2] “BOE-A-2019-5089 real decreto 244/2019, de 5 de abril, por el que se regulan las condiciones administrativas, técnicas y económicas del autoconsumo de energía eléctrica.” <https://www.boe.es/buscar/doc.php?id=BOE-A-2019-5089>. Accessed: 2024-5-22.
- [3] “Tema destacado: reducir la dependencia de la UE de los combustibles fósiles importados.” https://commission.europa.eu/news/focus-reducing-eus-dependence-imported-fossil-fuels-2022-04-20_es. Accessed: 2024-6-3.
- [4] Smart City, “Energy communities: How private citizens are saving the planet.” <https://www.thesmartcityjournal.com/en/green-new-deal/energy-communities-how-private-citizens-are-saving-the-planet>, Feb. 2022. Accessed: 2024-6-3.
- [5] “Clean energy for all europeans package.” https://energy.ec.europa.eu/topics/energy-strategy/clean-energy-all-europeans-package_en. Accessed: 2024-6-3.
- [6] Wikipedia contributors, “Acuerdo de París.” https://es.wikipedia.org/w/index.php?title=Acuerdo_de_Par%C3%ADs&oldid=160082677. Accessed: 2024-06-3.
- [7] G. T. (2024), “Exploración del término ‘self consumption’.” <https://trends.google.es/trends/explore?date=2010-01-06%202024-01-06&q=self%20consumption>, 2024.
- [8] F. Gianaroli, M. Ricci, P. Sdringola, M. Alessandra Ancona, L. Branchini, and F. Melino, “Development of dynamic sharing keys: Algorithms supporting management of renewable energy community and collective self consumption,” *Energy Build.*, vol. 311, no. 114158, p. 114158, 2024.
- [9] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” 2016.
- [10] “Python API reference — xgboost 2.1.0 documentation.” https://xgboost.readthedocs.io/en/stable/python/python_api.html. Accessed: 2024-6-22.
- [11] L. M. Demajo, “Explainable ai for interpretable credit scoring.” Thesis for: M.Sc. in Artificial Intelligence, 2020.

- [12] "An introduction to XGBoost regression." <https://www.kaggle.com/code/carlmcbrideellis/an-introduction-to-xgboost-regression>, Mar. 2023. Accessed: 2024-6-22.
- [13] A. Ogando-Martínez, X. García-Santiago, S. Díaz García, F. Echevarría Camarero, G. Blázquez Gil, and P. Carrasco Ortega, "Optimization of energy allocation strategies in spanish collective self-consumption photovoltaic systems," *Sustainability*, vol. 15, no. 12, p. 9244, 2023.
- [14] J. T. Villalonga Palou, J. Serrano González, and J. M. Riquelme Santos, "Estimation of energy distribution coefficients in collective self-consumption using meta-heuristic optimization techniques," *Sustainability*, vol. 16, no. 7, p. 2741, 2024.