



**Universidad
Europea**

UNIVERSIDAD EUROPEA DE MADRID

ESCUELA DE ARQUITECTURA, INGENIERÍA Y DISEÑO

MÁSTER UNIVERSITARIO EN ANÁLISIS DE DATOS MASIVOS

TRABAJO FIN DE MÁSTER

**Predicción del volumen neto de los desvíos
generación/demanda (MW) de la red
eléctrica de España.**

Edgar Romero Depablos

Dirigido por

Pablo del Saz-Orozco Huang

CURSO 2024 - 2025

TÍTULO: Predicción del volumen neto de los desvíos generación/demanda (MW) de la red eléctrica de España.

AUTOR: Edgar Romero Depablos

TITULACIÓN: MÁSTER UNIVERSITARIO EN ANÁLISIS DE DATOS MASIVOS

DIRECTOR DEL PROYECTO: Pablo del Saz-Orozco Huang

FECHA: Septiembre del 2025

RESUMEN

Este proyecto de investigación presenta el desarrollo de un modelo predictivo enfocado en el volumen neto de los desvíos generación/demanda del sistema eléctrico de España, con el objetivo de minimizar penalizaciones mediante la participación correcta en dos mercados: el continuo y el intradía. Cada mercado cuenta con un horizonte de predicción específico adaptado a sus características operativas.

El propósito es proveer a las comercializadoras de energía una herramienta que ayude a minimizar penalizaciones económicas relacionadas con los desequilibrios de sus compras.

Para ello se implementaron dos enfoques:

1. **Modelo de predicción horaria:** Se incorporaron variables exógenas (precio spot, producción eólica, temperatura, velocidad del viento, insolación, días festivos ponderados) y desvíos con un rezago de una hora. Este diseño fue evaluado mediante los algoritmos CatBoost y Prophet.
2. **Modelo multi-salida:** Con las mismas variables exógenas y una transformación de la variable objetivo para crear un dataframe con horizonte de 24 horas. La precisión se valoró cuantitativamente (RMSE) y cualitativamente (acierto de tendencia mediante matriz de confusión). Cabe destacar que fue evaluado con el modelo Catboost.

Los resultados muestran que, mientras el modelo de mercado continuo ofrece un rendimiento comparable al uso exclusivo de desvíos rezagados, el enfoque multi-salida consigue un ahorro medio del 5 % en costes para una comercializadora simulada, frente a estrategias sin modelado. El proyecto incluyó ingesta de datos públicos (ESIOS, AEMET, web scraping), análisis exploratorio, selección de variables y modelado estadístico y de machine learning en Python.

Palabras clave: modelo predictivo, mercado de desvíos, red eléctrica, machine learning, CatBoost, Prophet, web scraping.

ABSTRACT

This research project presents the development of a predictive model focused on the net volume of generation/demand deviations in the Spanish electrical system, aiming to minimize losses through correct participation in two markets: the continuous and the intraday. Each market has a specific forecasting horizon adapted to its operational characteristics.

The objective is to provide energy marketers with a tool that helps minimize economic penalties related to the imbalances in their purchases and, consequently, contribute to the overall stability of the electrical system.

Two approaches were implemented:

1. **Hourly prediction model:** Exogenous variables were incorporated (spot price, wind power production, temperature, wind speed, solar radiation, weighted holidays) as well as deviations with a one-hour lag. This design was evaluated using the CatBoost and Prophet algorithms.
2. **Multi-output model:** Using the same exogenous variables and a transformation of the target variable to create a dataframe with a 24-hour forecasting horizon. Accuracy was assessed quantitatively (RMSE) and qualitatively (trend accuracy using a confusion matrix). It is worth noting that this model was evaluated using CatBoost.

The results show that, while the continuous market model offers performance comparable to the exclusive use of lagged deviations, the multi-output approach achieves an average cost savings of 5% for energy marketers compared to non-modeling strategies. The project included the ingestion of public data (ESIOS, AEMET, web scraping), exploratory analysis, variable selection, and statistical and machine learning modeling in Python.

Keywords: predictive model, deviation market, power grid, machine learning, CatBoost, Prophet, web scraping.

Índice general

1. RESUMEN DEL PROYECTO	9
1.1. Contexto y justificación	9
1.2. Planteamiento del problema	9
1.3. Objetivos del proyecto	9
1.4. Resultados obtenidos	10
1.5. Estructura de la memoria	10
2. ANTECEDENTES / ESTADO DEL ARTE	11
2.1. Estado del arte	11
2.1.1. Modelos avanzados de machine learning y enfoques híbridos	11
2.1.2. Modelos estadísticos clásicos	11
2.1.3. Modelos estadísticos avanzados	12
2.1.4. Modelos de machine learning multi-salida	12
2.2. Contexto y justificación	13
2.3. Planteamiento del problema	14
2.4. Marco teórico	14
2.4.1. Sistema eléctrico	14
2.4.2. Agentes del sistema eléctrico	15
2.4.3. Mercados del sistema eléctrico	16
2.4.4. Equilibrio técnico y estabilidad del sistema eléctrico	16
2.4.5. Servicios de ajuste o servicios complementarios	16
2.4.6. Responsabilidad del equilibrio	17
2.4.7. Desvíos del sistema eléctrico	17
2.4.8. Liquidación de desvíos y precio de desvío	17
2.4.9. Fundamentos de series temporales	18
2.4.10. Modelo Prophet	18
2.4.11. Modelo CatBoost	20
2.4.12. Modelos híbridos y combinación de pronósticos	22
2.4.13. Enfoque multi-salida con CatBoost	22
2.4.14. Métricas de error en forecasting	23
2.4.15. Métricas de clasificación de tendencia	23
3. OBJETIVOS	24
3.1. Objetivos generales	24
3.2. Objetivos específicos	24
3.3. Beneficios del proyecto	24
4. DESARROLLO DEL PROYECTO	25
4.1. Planificación del proyecto	25
4.2. Descripción de la solución, metodologías y herramientas empleadas	25
4.2.1. Recopilación de datos	25
4.2.2. Análisis exploratorio de datos	27

4.3. Recursos requeridos	35
4.4. Resultados del proyecto	36
4.4.1. Análisis univariante	36
4.4.2. Análisis Multivariante	44
5. Discusión	52
5.1. Interpretación de los hallazgos	52
5.2. Decisiones metodológicas y adaptaciones	52
5.3. Limitaciones	52
5.4. Impacto práctico	52
6. Conclusiones	53
6.1. Conclusiones del trabajo	53
6.2. Conclusiones personales	53
7. Futuras líneas de trabajo	54
8. ANEXOS	58

Índice de Figuras

4.1. Serie temporal del volumen neto de desvíos. <i>Fuente: Cálculos propios.</i>	37
4.2. Histograma con densidad del volumen neto de los desvíos. <i>Fuente: Cálculos propios.</i>	37
4.3. Gráfico Q-Q del volumen neto de los desvíos. <i>Fuente: Cálculos propios.</i>	38
4.4. Función de autocorrelación (ACF). <i>Fuente: Cálculos propios.</i>	39
4.5. Función de autocorrelación parcial (PACF). <i>Fuente: Cálculos propios.</i>	39
4.6. Serie diaria del volumen neto de desvíos. <i>Fuente: Cálculos propios.</i>	41
4.7. Serie semanal del volumen neto de desvíos. <i>Fuente: Cálculos propios.</i>	42
4.8. Serie mensual del volumen neto de desvíos (2015). <i>Fuente: Cálculos propios.</i>	43
4.9. Serie mensual del volumen neto de desvíos (2016). <i>Fuente: Cálculos propios.</i>	43
4.10. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día. <i>Fuente: Cálculos propios.</i>	44
4.11. Matriz de correlación entre el volumen neto de desvíos y variables exógenas. <i>Fuente: Cálculos propios.</i>	45
8.1. Serie mensual del volumen neto de desvíos. <i>Fuente: Cálculos propios.</i>	58
8.2. Serie trimestral del volumen neto de desvíos. <i>Fuente: Cálculos propios.</i>	58
8.3. Serie semestral del volumen neto de desvíos. <i>Fuente: Cálculos propios.</i>	59
8.4. Serie anual del volumen neto de desvíos. <i>Fuente: Cálculos propios.</i>	59
8.5. Serie mensual del volumen neto de desvíos (2017). <i>Fuente: Cálculos propios.</i>	60
8.6. Serie mensual del volumen neto de desvíos (2018). <i>Fuente: Cálculos propios.</i>	60
8.7. Serie mensual del volumen neto de desvíos (2019). <i>Fuente: Cálculos propios.</i>	61
8.8. Serie mensual del volumen neto de desvíos (2020). <i>Fuente: Cálculos propios.</i>	61
8.9. Serie mensual del volumen neto de desvíos (2021). <i>Fuente: Cálculos propios.</i>	62
8.10. Serie mensual del volumen neto de desvíos (2022). <i>Fuente: Cálculos propios.</i>	62
8.11. Serie mensual del volumen neto de desvíos (2023). <i>Fuente: Cálculos propios.</i>	63
8.12. Serie mensual del volumen neto de desvíos (2024). <i>Fuente: Cálculos propios.</i>	63
8.13. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Enero). <i>Fuente: Cálculos propios.</i>	64
8.14. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Febrero). <i>Fuente: Cálculos propios.</i>	64
8.15. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Marzo). <i>Fuente: Cálculos propios.</i>	65
8.16. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Abril). <i>Fuente: Cálculos propios.</i>	65
8.17. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Mayo). <i>Fuente: Cálculos propios.</i>	66
8.18. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Junio). <i>Fuente: Cálculos propios.</i>	66

8.19. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Julio). <i>Fuente: Cálculos propios.</i>	67
8.20. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Agosto). <i>Fuente: Cálculos propios.</i>	67
8.21. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Septiembre). <i>Fuente: Cálculos propios.</i>	68
8.22. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Octubre). <i>Fuente: Cálculos propios.</i>	68
8.23. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Noviembre). <i>Fuente: Cálculos propios.</i>	69
8.24. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Diciembre). <i>Fuente: Cálculos propios.</i>	69

Índice de Tablas

4.1. Cronograma de actividades y esfuerzo del proyecto	25
4.2. Estadísticos descriptivos del volumen neto de desvíos	36
4.3. Resultados del test ADF (Dickey–Fuller aumentada) para $\alpha = 0,05$	40
4.4. Resultados del test de Ljung–Box (lag=10) para $\alpha = 0,05$	40
4.5. Resultados del test de Shapiro–Wilk para $\alpha = 0,05$	41
4.6. Hiperparámetros del modelo CatBoost	46
4.7. Importancia de características obtenida de CatBoost (%)	46
4.8. Mejores hiperparámetros encontrados por GridSearchCV	47
4.9. Desempeño en validación cruzada (5 folds)	47
4.10. Accuracy de signo en entrenamiento+validación	47
4.11. Mejores hiperparámetros de Prophet según $RMSE_{CV}$	48
4.12. Desempeño de Prophet en validación cruzada (5 folds)	48
4.13. Accuracy de signo de Prophet en entrenamiento+validación	48
4.14. Desempeño del modelo híbrido	48
4.15. Comparación de RMSE y accuracy entre modelos	49
4.16. Hiperparámetros finales del modelo multi-salida	49
4.17. Desempeño continuo por horizonte (RMSE)	49
4.18. Precisión de signo por horizonte	50
4.19. Comparativo RMSE y accuracy de signo: multi-salida vs. modelos dummy	50
4.20. Costes de penalización de la comercializadora	50

Capítulo 1. RESUMEN DEL PROYECTO

1.1 Contexto y justificación

En 2024, el coste de los servicios de ajuste en España alcanzó los 2 668 millones de euros, un 7,5 % superior al del año anterior, lo que supone una carga significativa para los agentes del mercado (Red Eléctrica de España, 2024).

Por otra parte, según Red Eléctrica de España (2025, 2024), el incremento en la producción de energías renovables ha aumentado la variabilidad en la programación energética y ha elevado el volumen (y coste) de los desvíos. En 2024, las renovables representaron el 56,8 % del mix eléctrico (Red Eléctrica de España, 2025) y el coste de los servicios de ajuste (indicador de penalizaciones por desvíos) alcanzó los 2 668 M€, un 7,5 % más que el año anterior (Red Eléctrica de España, 2024).

Por lo tanto, generar un modelo que permita predecir los desvíos futuros ayudaría a las comercializadoras a optimizar sus operaciones.

1.2 Planteamiento del problema

En el mercado eléctrico de España existen múltiples participantes que hacen posible su funcionamiento. Uno de esos agentes son las comercializadoras, las cuales se encargan de comprar la electricidad en el mercado y venderla a los consumidores (hogares y negocios). No obstante, cuando la cantidad de electricidad comprada resulta muy diferente de la consumida, se genera un desequilibrio en el sistema. Dicho desequilibrio se penaliza; es decir, todo agente que contribuya a la discrepancia entre lo consumido y lo programado será penalizado por el mercado para estabilizar el sistema (Conde Buezas, 2016).

Por lo tanto, predecir las desviaciones es fundamental para evitar pérdidas considerables de las partes involucradas. No obstante, las predicciones en este ámbito resultan un desafío técnico, dado que dependen de múltiples variables exógenas. Es importante resaltar que algunas son más difíciles de obtener que otras (Filgueira Fernández, 2024).

Por otra parte, capturar las relaciones entre las variables exógenas y la variable dependiente (volumen neto de los desvíos) puede ser complejo, ya que dichas relaciones pueden ser no lineales. Por lo tanto, se requiere el uso de algoritmos de machine learning y modelos estadísticos avanzados (López García, 2016).

1.3 Objetivos del proyecto

En esta investigación se busca comprender la importancia de prever el desvío del sistema de forma que se adecúe la compra de energía a las distribuidoras, utilizando modelos estadísticos y de machine learning. Para ello, se deben extraer, analizar y modelar distintas variables independientes, con el propósito de explicar su relevancia en el fenómeno de estudio y mejorar la precisión de los pronósticos.

1.4 Resultados obtenidos

Esta investigación destacó la importancia de realizar múltiples análisis exploratorios para identificar patrones en las variables. Al tratarse de un fenómeno temporal, la descomposición de la fecha y hora permitió configurar componentes como año, mes y hora, que aportaron significativamente a la precisión del modelo. Además, la realización de un análisis comparativo (benchmark) entre distintos modelos fue crucial para obtener predicciones robustas sin caer en sobreajuste. En términos cuantitativos, el modelo CatBoost obtuvo el mejor ajuste con un RMSE ≈ 398.5 MWh y una precisión en la dirección del desvío (accuracy de signo) del 84.4 %, frente al baseline (dummy 1 h) que presentó un RMSE de 424.3 MWh. Prophet alcanzó un RMSE ≈ 408.7 MWh (accuracy 84.2 %) y el ensamblado híbrido consiguió un RMSE intermedio ≈ 402.2 MWh (accuracy 84.37 %). El modelo multisalida, útil para pronosticar 24 h de forma directa, mostró un RMSE promedio ≈ 952.0 MWh y una accuracy global del 66 %, indicando pérdida de precisión en horizontes largos. Finalmente, la simulación económica basada en las predicciones reveló un ahorro absoluto de 104 889.48 EUR, equivalente a una reducción del 4.86 % en costes de penalización frente a una estrategia de compra fija.

1.5 Estructura de la memoria

El presente documento incluye un breve resumen de la investigación, explicando sus componentes y el motivo que generó este trabajo. Además, se presentan cada uno de los objetivos y una descripción detallada de cómo se alcanzaron. Por último, se exponen los resultados obtenidos, las conclusiones y las posibles mejoras.

Capítulo 2. ANTECEDENTES / ESTADO DEL ARTE

2.1 Estado del arte

En esta sección se exponen una serie de trabajos que anteceden este proyecto. Dichas investigaciones proporcionan una visión general de la problemática planteada en este trabajo y explican los avances logrados hasta la fecha. Además, se puede evidenciar las oportunidades de mejora y los beneficios que aporta esta investigación.

Cabe destacar que, para lograr este análisis del estado del arte, se revisaron una serie de fuentes oficiales como ScienceDirect y MDPI, aplicando palabras clave como forecasting, machine learning y electricity.

2.1.1. Modelos avanzados de machine learning y enfoques híbridos

Al Mamun et al. (2017) describen en su investigación la importancia de elegir cuidadosamente los diferentes factores que pueden afectar la precisión de un modelo de predicción de carga eléctrica, como lo son el tiempo, el clima y la economía. Por otra parte, mencionan que el uso de modelos híbridos como el SVM-BFGSA, entre otros, puede mejorar la precisión de las predicciones al integrar la fuerza de diferentes técnicas.

En adición, se comenta que las diferentes métricas de evaluación como RMSE y el MAPE son fundamentales para determinar el modelo más adecuado para el fenómeno de estudio. Además, el uso de diferentes modelos según el horizonte temporal requerido permite adaptar las técnicas a las necesidades específicas de cada sistema eléctrico.

En conclusión, el análisis de las variables exógenas y el uso de algoritmos robustos que permitan capturar relaciones no lineales puede ser de gran utilidad en esta investigación. De acuerdo con Al Mamun et al. (2017), los modelos de machine learning como el SVM-BFGSA mostraron una reducción del 9.63 % en el MAPE en comparación con el modelo ARIMA. Por lo tanto, es pertinente considerar aplicar técnicas avanzadas como el algoritmo CatBoost para la predicción del volumen neto de los desvíos, además, una combinación de modelos de machine learning y estadísticos podría generar un modelo aún más eficiente y robusto, como sugieren los autores.

2.1.2. Modelos estadísticos clásicos

En las zonas insulares de España, como las Islas Canarias y Baleares, Caro y Juan (2020) desarrollaron un modelo de serie temporal ARIMAX para realizar pronósticos horarios de la demanda de energía eléctrica (MW) con un horizonte de un día. Dicho modelo considera factores como la temperatura y los días festivos, con el propósito de adaptarse a cada sistema insular.

En su investigación, los autores compararon los resultados con un software de pronósticos utilizado por la Red Eléctrica de España (REE). Por otra parte, emplearon métricas como el error porcentual absoluto medio (MAPE) para evaluar la precisión de los modelos y la prueba de Diebold-Mariano para comparar la precisión de los modelos desarrollados con el software de referencia utilizado por REE.

Como señalan Caro y Juan (2020) en sus conclusiones, los modelos desarrollados son más precisos que el software de referencia, con mejoras considerables en las Islas Baleares. La hora y la temperatura tienen un efecto notable en dichas mejoras; es decir, considerar este tipo de variables en la investigación puede ser beneficioso para la precisión de los modelos.

En conclusión, dicho proyecto resulta beneficioso para esta investigación, dado que indica que los modelos estadísticos pueden generar buenos pronósticos en el sector energético y, además, si utilizamos las variables exógenas correctas, se puede aportar considerablemente al sector y a todas las partes involucradas.

2.1.3. Modelos estadísticos avanzados

Sanz Muñoz (2023) estudia la predicción de precios en el mercado diario de la electricidad de España durante el periodo 2018-2019 mediante la configuración y comparación de un conjunto de modelos, entre ellos el modelo Prophet en versiones univariantes y multivariantes.

El modelo Prophet en su versión multivariante incorporó variables exógenas como la demanda P48, la generación eólica P48 y el día de la semana, logrando el mejor MAE de todos los modelos (2,652 €/MWh). Esto demuestra la importancia de integrar variables exógenas en un modelo de series temporales para mejorar la precisión del mismo, al capturar de una mejor manera la estacionalidad y otros patrones.

Se puede concluir que el modelo Prophet resulta de gran utilidad para predecir valores relacionados con el sector energético, donde se pueden encontrar patrones cambiantes en la estacionalidad y es frecuente la presencia de valores atípicos.

2.1.4. Modelos de machine learning multi-salida

Miele, Ludwig y Corsini (2023) comentan en su investigación que la predicción multi-horizonte es de vital importancia porque permite anticipar varias horas futuras de una misma señal, apoyando de esa manera decisiones de operación y compra. Por consiguiente, los autores estudian este problema en el contexto de la potencia eólica a nivel de turbina, señalando que la variabilidad del viento y la necesidad de precisión exigen pronósticos de muchas horas adelante para planificar mantenimiento, despacho y coordinación con otros recursos.

La variable objetivo es la potencia activa generada por cada turbina, con una periodicidad horaria, y se busca poder predecir las siguientes 90 horas desde cada momento de partida. Para ello, se elabora un conjunto de datos donde cada registro contiene el historial reciente de la turbina, junto con las variables de salida correspondientes a cada intervalo de tiempo con horizonte de 90 horas (Miele, Ludwig & Corsini, 2023).

Cabe destacar que el conjunto de variables explicativas proviene de SCADA (Supervisory Control And Data Acquisition) y NWP (Numerical Weather Prediction). SCADA proporciona datos internos del equipo, como potencia instantánea, temperaturas internas y externas, vibraciones, entre otros indicadores de estado y rendimiento. Por otra parte, NWP proporciona valores como velocidad y dirección del viento, temperatura, presión y humedad en distintos niveles de la atmósfera y ubicaciones geográficas (Miele, Ludwig & Corsini, 2023).

El modelo utilizado para llevar a cabo esta investigación fue una red neuronal multimodal espaciotemporal basada en LSTM, utilizando métricas para evaluar los pronósticos, como el RMSE, el MAE y un skill score que mide la mejora porcentual frente a modelos base, como la regresión lineal. Los resultados obtenidos por Miele, Ludwig y Corsini (2023) indican que el error crece con el horizonte y reportan un skill score medio cercano al 25 % sobre el modelo base.

En conclusión, esta investigación deja un precedente en la importancia de usar modelos multi-salida en combinación con variables exógenas, señalando no solo su eficacia, sino la metodología para llevar a cabo este tipo de proyectos, desde la transformación de los datos hasta la forma de evaluar la precisión del modelo.

2.2 Contexto y justificación

En el sistema eléctrico español, las comercializadoras son las encargadas de comprar energía a las empresas generadoras en los mercados diarios e intradiarios, con el propósito de suministrarles a los consumidores, como los hogares y empresas, energía de manera constante según sus necesidades. Cuando la energía consumida difiere de lo programado, se genera una discrepancia entre la generación y la demanda. Dicha diferencia genera un desequilibrio en el sistema eléctrico (CNMC, 2019).

Por lo tanto, la entidad responsable de mantener el sistema eléctrico en óptimas condiciones (REE) debe corregir mediante penalizaciones en el mercado de desvíos. Dichas penalizaciones generan pérdidas económicas considerables a las comercializadoras y a todo agente responsable que contribuya con el desequilibrio.

Es importante resaltar que, en 2024, el coste de los servicios de ajuste en España alcanzó los 2 668 millones de euros, un 7,5 % superior al del año anterior, lo que supone una carga significativa para los agentes del mercado (Red Eléctrica de España, 2024).

Además, según Red Eléctrica de España (2024, 2025), el incremento en la producción de energías renovables ha aumentado la variabilidad en la programación energética y ha elevado el volumen (y coste) de los desvíos. En 2024, las renovables representaron el 56,8 % del mix eléctrico (Red Eléctrica de España, 2025).

En conclusión, el desarrollo de un modelo predictivo capaz de pronosticar el volumen neto de los desvíos es de vital importancia para todos los agentes involucrados, debido a que, con esa información, serán capaces de evitar penalizaciones en el mercado de desvíos y, como consecuencia, se producirá un equilibrio duradero en el sistema eléctrico, lo cual podría generar mejores precios para los consumidores y un servicio de mejor calidad.

2.3 Planteamiento del problema

Se ha evidenciado que el sistema eléctrico español está compuesto por múltiples agentes, desde los que generan la energía hasta los consumidores. Uno de ellos son las comercializadoras, agentes encargados de adquirir energía en el mercado energético (diario, intradiario y de desvíos) para poder suministrarla de forma eficiente a los diferentes tipos de consumidores. En este sistema, cuando las comercializadoras adquieren un volumen de energía diferente al realmente consumido, se genera un desvío que pone en desequilibrio todo el sistema. Por lo tanto, el agente encargado debe corregir dicho desequilibrio, y el método para hacerlo es a través del mercado de desvíos, donde los distintos agentes que presentaron desviaciones pueden comprar o vender según el caso, a un precio desfavorable para ellos debido a una estimación incorrecta de sus previsiones (CNMC, 2019).

Es importante tener en cuenta que un sistema eléctrico en óptimas condiciones genera beneficios para todos los agentes involucrados, lo cual no solo favorece a las comercializadoras, sino a todo el sector, incluidos los consumidores.

No obstante, a pesar de existir una cantidad considerable de investigaciones relacionadas con la demanda y la generación de energía por separado, son escasos los proyectos que específicamente evalúan el volumen neto de los desvíos (generación/demanda). Esto puede evidenciarse en la investigación de Caro y Juan (2020), donde, al desarrollar un modelo ARIMAX, lograron pronosticar la demanda de los sistemas insulares españoles con un RMSE satisfactorio. Además, Al Mamun et al. (2017) evaluaron un conjunto de técnicas predictivas individuales e híbridas, obteniendo resultados competitivos, pero ninguno de estos proyectos se enfocó en el pronóstico del volumen neto de los desvíos.

En trabajos de investigación recientes, Sanz Muñoz (2023) ejecutó el modelo Prophet, incorporando variables exógenas como la demanda P48, la generación eólica P48 y el día de la semana, con el propósito de predecir el precio del mercado diario español, obteniendo un MAE considerablemente bajo. Esto demuestra la utilidad de dicho modelo y su capacidad predictiva al incorporar regresores externos. Sin embargo, este proyecto tampoco evaluó el volumen de los desvíos, sino el precio de la energía.

Por lo tanto, en esta investigación se plantea desarrollar un modelo predictivo capaz de aportar una solución a esta problemática, permitiendo que las comercializadoras puedan mitigar sus pérdidas y que el sistema en general alcance intervalos de equilibrio más prolongados. Para ello, se evaluarán dos modelos con algoritmos completamente distintos, pero igual de eficaces. Dichos modelos se entrenarán y evaluarán con un histórico de datos del volumen neto de los desvíos desde 2015 hasta 2024.

2.4 Marco teórico

2.4.1. Sistema eléctrico

El sistema eléctrico se puede comprender como el conjunto de infraestructuras y procesos que garantizan la generación, transporte, distribución y suministro de electricidad, con el propósito de satisfacer la demanda de energía en todo momento (Red Eléctrica de España, s.f.).

Cabe destacar que en España coexisten varios subsistemas, tales como: Peninsular, Baleares, Canarias, Ceuta y Melilla. Sin embargo, el más grande es el subsistema peninsular, el cual está interconectado con países vecinos y registra una generación de energía que superó los 248 478 GWh en 2024 (Red Eléctrica de España, 2024). Por consiguiente, el sistema requiere la interacción de múltiples agentes y mercados para mantener el equilibrio entre generación y demanda de forma segura y eficiente (Red Eléctrica de España, s.f.).

2.4.2. Agentes del sistema eléctrico

(Fundación Endesa, s.f.) identifica los siguientes agentes principales del sistema:

- . **Generadores:** instalan, operan y mantienen centrales de generación eléctrica (térmicas, hidráulicas, renovables, etc.), y participan informando sus ofertas de cantidad y precio al mercado.
- . **Productores en régimen especial:** tipología de generadores sujetos a tratamientos económicos o regulatorios específicos por eficiencia o impacto ambiental.
- . **Transportistas:** gestionan la red de transporte de alta tensión, trasladando la energía desde los centros de generación hasta las redes de distribución, garantizando la capacidad y seguridad de la transmisión.
- . **Distribuidores:** operan y mantienen la red de media y baja tensión, llevando la electricidad hasta el punto de consumo y gestionando infraestructuras de distribución.
- . **Comercializadores:** empresas que adquieren electricidad en los mercados mayoristas y la venden a consumidores finales, asumiendo la contratación y gestión de precios.
- . **Consumidores calificados:** usuarios con un volumen de consumo que les otorga participación directa en el mercado mayorista.
- . **Operador del sistema:** en España, Red Eléctrica de España (REE) actúa como operador del sistema de transporte (TSO), responsable de la gestión técnica del sistema, coordinación continua de la red y mantenimiento del equilibrio entre generación y demanda (Fundación Endesa, s.f.; Red Eléctrica de España, s.f.).
- . **Operador del mercado:** gestiona la casación de ofertas y demandas para determinar los precios horarios; en el mercado diario OMIE (Operador del Mercado Ibérico de Energía) cumple esta función.
- . **Reguladores:** la Administración del Estado y la Comisión Nacional de los Mercados y la Competencia (CNMC) establecen y supervisan marcos regulatorios, metodologías de funcionamiento de mercados y liquidaciones (CNMC, 2019).

2.4.3. Mercados del sistema eléctrico

Los mercados eléctricos se organizan en diferentes tipos para permitir la compra y venta de energía en distintos horizontes temporales y de esa forma gestionar el balance en tiempo real (CNMC, s.f.; Fundación Endesa, s.f.).

- **Mercados a plazo:** contratos bilaterales o subastas a futuro para cubrir volúmenes con antelación.
- **Mercado diario:** casación de ofertas de compra y venta para cada una de las 24 horas del día siguiente. Todas las unidades de generación disponibles participan obligatoriamente, y el precio resultante es marginalista, beneficiando la eficiencia del sistema (Fundación Endesa, s.f.; Endesa, 2025).
- **Mercado intradiario:** permite ajustes tras el cierre del mercado diario, organizándose en varias sesiones donde generadores y comercializadoras corrigen posiciones según nueva información de demanda o disponibilidad.
- **Mercado continuo:** es un segmento dentro del mercado intradiario en el cual se permite la negociación en tiempo casi real, mediante un sistema de casación continua, a diferencia del mercado diario y de las sesiones intradiarias que funcionan por subasta. En este mercado, los agentes pueden modificar sus previsiones de generación o demanda y ajustar sus posiciones comerciales conforme se acercan a la hora real de entrega, permitiendo una mayor flexibilidad operativa (OMIE, s.f.).
- **Mercados de servicios complementarios o de balance:** incluyen mecanismos para resolver desequilibrios en tiempo real o casi real. Dentro de estos está el llamado “mercado de desvíos” o “servicios de ajuste”, donde se compensan las diferencias entre la energía programada y la realmente inyectada o consumida para mantener el equilibrio técnico del sistema (CNMC, 2019; Fundación Endesa, s.f.).

2.4.4. Equilibrio técnico y estabilidad del sistema eléctrico

“El equilibrio instantáneo entre la electricidad generada y la demandada es imprescindible, ya que cualquier diferencia podría comprometer la estabilidad de la red. Para garantizar este balance, el operador del sistema (REE) monitoriza en tiempo real los flujos eléctricos, activa reservas y utiliza mecanismos de ajuste para absorber desviaciones” (Red Eléctrica de España, s.f.).

2.4.5. Servicios de ajuste o servicios complementarios

“Los servicios de ajuste o complementarios permiten al operador del sistema mantener la calidad del suministro eléctrico y compensar rápidamente desviaciones imprevistas. Entre

ellos se encuentran la regulación secundaria (respuesta en segundos), la terciaria (respuesta en minutos) y los servicios de gestión de restricciones técnicas"(REE, s.f.).

2.4.6. Responsabilidad del equilibrio

"La normativa española y europea establece que todos los agentes participantes en el sistema eléctrico son responsables de mantener su propio equilibrio entre inyecciones y extracciones de energía, debiendo asumir las consecuencias económicas de los desvíos"(CNMC, 2019).

2.4.7. Desvíos del sistema eléctrico

El desvío en el sistema eléctrico se puede definir como la diferencia entre la energía medida y la energía programada en el mercado. Desde la perspectiva del agente generador, el cálculo de los desvíos es igual a la energía medida en barras de central menos la energía programada en el horario de liquidación (MagnusCMD, s.f.).

$$\text{Desvío} = \text{Energía medida} - \text{Energía programada}$$

Por otra parte, se puede mencionar que los desvíos generan obligaciones de pago o derechos de cobro con el operador del sistema (REE), independientemente del operador de mercado (OMIE), y se liquidan mediante mecanismos específicos. Cabe destacar que la gestión de desvíos es crítica porque las penalizaciones o recompensas afectan directamente la rentabilidad de generadores y comercializadoras. En adición, contribuyen a la estabilidad del sistema eléctrico al incentivar comportamientos que reduzcan desequilibrios (MagnusCMD, s.f.).

2.4.8. Liquidación de desvíos y precio de desvío

La liquidación de los desvíos está regulada por el Procedimiento Operativo PO-14.4 de Red Eléctrica de España, en el cual los precios de desvío se calculan de forma marginalista para cada período de liquidación (Red Eléctrica de España, s.f.). En detalle:

- **Desvíos a subir:** si un agente inyecta menos energía de la programada, paga la diferencia al precio marginal de "desvíos a subir" publicado por el operador del sistema (Red Eléctrica de España, s.f.).
- **Desvíos a bajar:** si inyecta más energía de la programada, recibe la diferencia al precio marginal de "desvíos a bajar" (Red Eléctrica de España, s.f.).

Este mecanismo tiene en consideración la situación del sistema en cada momento y el volumen concreto desviado. Por lo tanto, la penalización o compensación varía según el contexto operativo y no es uniforme para todos los agentes (Red Eléctrica de España, s.f.).

2.4.9. Fundamentos de series temporales

La predicción de los desvíos de la red eléctrica se encuentra dentro del análisis de series temporales, donde es pertinente comprender la estructura de dichas series. Hyndman, R. J., & Athanasopoulos, G. (2018) comentan que la estructura se conforma por la tendencia, estacionalidad y ruido. Además, entender las propiedades, como si la serie es estacionaria, presenta autocorrelación y si se distribuye normal.

Para poder determinar las propiedades antes mencionadas, se pueden aplicar las siguientes pruebas:

- **Pruebas de estacionariedad:** se emplea, entre otros, el test de Dickey-Fuller aumentado (ADF) para evaluar si la serie tiene raíz unitaria (Dickey & Fuller, 1979). Una serie no estacionaria puede requerir diferenciación o transformaciones antes de modelar.
- **Autocorrelación y correlación parcial:** el análisis de ACF/PACF y pruebas como Ljung-Box ayudan a identificar dependencia temporal y retardos relevantes (Box, Jenkins, Reinsel, & Ljung, 2015).
- **Normalidad de residuos o de variables:** tests como Shapiro-Wilk o Kolmogórov-Smirnov informan si la distribución de errores o de la variable se ajusta a normalidad, aunque muchos métodos de machine learning no exigen normalidad estricta, conocer la distribución ayuda a entender riesgos extremos (Shapiro & Wilk, 1965).

2.4.10. Modelo Prophet

Prophet es un modelo estadístico, el cual se basa en un modelo aditivo para series temporales, descomponiendo la observación en componentes de tendencia, estacionalidad, efectos de días festivos y ruido (Taylor & Letham, 2018):

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

donde:

- $g(t)$ es la función de tendencia,
- $s(t)$ modela la estacionalidad,
- $h(t)$ representa los efectos de días festivos o eventos,
- ε_t es el término de error (ruido), asumido con media aproximadamente cero.

1. Tendencia: $g(t)$

Prophet admite principalmente dos tipos de tendencia: lineal con puntos de cambio (piecewise linear) o logística. En el caso lineal con cambios de pendiente en tiempos $\{t_c\}$, se define:

$$g(t) = \left(k + \sum_{j=1}^S \delta_j a_j(t) \right) (t - t_0) + \left(m + \sum_{j=1}^S \gamma_j a_j(t) \right)$$

donde:

- k es la pendiente inicial,
- m es la intersección inicial,
- $\{t_c\}_{j=1}^S$ son los tiempos de cambio (*changepoints*),
- $a_j(t) = \begin{cases} 1, & \text{si } t > t_{cj} \\ 0, & \text{si } t \leq t_{cj} \end{cases}$,
- δ_j es el cambio de pendiente en t_{cj} ,
- Para mantener la continuidad de $g(t)$, se ajusta el intercepto con $\gamma_j = -t_{cj}\delta_j$,
- t_0 es un punto de referencia temporal (por ejemplo, el tiempo inicial).

La estimación de δ_j se realiza con regularización bayesiana (*Bayesian prior*) para evitar un exceso de cambios de pendiente y sobreajuste (Taylor & Letham, 2018).

En la versión logística, la tendencia se define como:

$$g(t) = \frac{C}{1 + \exp(-k(t - m))}$$

o en una forma análoga con *changepoints*.

2. Estacionalidad: $s(t)$

La estacionalidad se modela mediante series de Fourier para capturar patrones periódicos (diarios, semanales, anuales). Por ejemplo, para una estacionalidad de período P , se utiliza la siguiente representación:

$$s(t) = \sum_{n=1}^N \left[a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right) \right]$$

donde:

- P es el período de la estacionalidad (por ejemplo, $P = 24$ para estacionalidad diaria u $P = 7$ para estacionalidad semanal),
- N es el número de términos (orden) de la serie de Fourier, que controla el grado de complejidad capturado,
- a_n y b_n son los coeficientes de la serie, estimados junto con el resto de parámetros del modelo,
- La estimación puede realizarse minimizando la función de pérdida o mediante inferencia bayesiana aproximada (Taylor & Letham, 2018).

3. Efectos de días festivos: $h(t)$

Los efectos de días festivos se incorporan como regresores binarios o mediante ventanas de influencia. Si D es el conjunto de fechas festivas, se define para cada festivo $d \in D$ un indicador:

$$D_d(t) = \begin{cases} 1, & \text{si } t \text{ coincide o está en una ventana alrededor de } d \\ 0, & \text{en otro caso} \end{cases}$$

Entonces, el componente de festivos se modela como:

$$h(t) = \sum_{d \in D} \beta_d D_d(t)$$

donde:

- $D_d(t)$ es el indicador binario para el festivo d ,
- β_d es el coeficiente que cuantifica el impacto del festivo d sobre la serie,
- Estos regresores también pueden representar eventos especiales o campañas que alteren el comportamiento normal de la serie.

2.4.11. Modelo CatBoost

CatBoost es un algoritmo de gradient boosting sobre árboles de decisión que introduce innovaciones para manejar variables categóricas y reducir el sesgo en el cálculo de gradientes (Prokhorenkova et al., 2018). A continuación se exponen los fundamentos matemáticos del boosting y las particularidades de CatBoost.

1. Fundamentos matemáticos del boosting

Gradient Boosting construye un modelo de forma aditiva, ajustando cada nuevo árbol a los residuos del modelo anterior (Friedman, 2001). Formalmente, dada una función de pérdida L , el procedimiento es el siguiente:

1. Se inicia con una predicción constante:

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^N L(y_i, \gamma)$$

2. Para cada iteración $m = 1, 2, \dots, M$:

- Se calculan los pseudo-residuos como el gradiente negativo de la función de pérdida:

$$r_i^{(m)} = - \left. \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right|_{F=F_{m-1}}$$

- Se ajusta un nuevo árbol de regresión $h_m(x)$ para predecir los pseudo-residuos $r_i^{(m)}$.
- Se actualiza el modelo:

$$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x)$$

donde $\eta \in (0, 1]$ es la tasa de aprendizaje, que controla cuánto contribuye cada nuevo árbol.

En el caso particular de pérdida cuadrática:

$$L(y, F) = \frac{1}{2}(y - F)^2$$

los pseudo-residuos son simplemente:

$$r_i^{(m)} = y_i - F_{m-1}(x_i)$$

Para funciones de pérdida genéricas (como log-loss para clasificación), se utilizan los gradientes y en algunos algoritmos como XGBoost también los hessianos (segunda derivada) para construir los splits de los árboles (Chen & Guestrin, 2016).

2. Particularidades de CatBoost

- **Ordered Boosting:** evita el *target leakage* al tratar variables categóricas. En lugar de usar toda la muestra para calcular estadísticas como medias condicionales, CatBoost utiliza permutaciones internas y estima las estadísticas en orden aleatorio. Esto reduce el sesgo en los gradientes durante el entrenamiento y mejora la generalización del modelo (Prokhorenkova et al., 2018).
- **Árboles simétricos:** cada árbol es construido de forma que todos los splits a una misma profundidad usan la misma característica. Esta estructura acelera la predicción, facilita la paralelización y estabiliza el modelo al reducir la varianza.
- **Regularización interna:** el modelo incluye hiperparámetros como `depth` (profundidad del árbol), `learning_rate` (tasa de aprendizaje) y `l2_leaf_reg` (equivalente al parámetro λ en la regularización de las hojas), que ayudan a controlar el sobreajuste y mejorar el rendimiento fuera de muestra.
- **Manejo nativo de variables categóricas:** CatBoost procesa internamente las variables categóricas mediante codificaciones basadas en estadísticas relativas a la variable objetivo, por lo que no es necesario aplicar una codificación manual tipo one hot encoding que aumente mucho la dimensionalidad del conjunto de datos. Esto simplifica el preprocesado, reduce el riesgo de errores por transformación y permite aprovechar la información de las categorías de forma más directa y eficiente (Prokhorenkova et al., 2018).

2.4.12. Modelos híbridos y combinación de pronósticos

La combinación de pronósticos busca mejorar precisión al aprovechar fortalezas complementarias de modelos distintos (Clemen, 1989). Se puede mencionar que al combinar Prophet (fuerte en capturar tendencia/estacionalidad) y CatBoost (fuerte en no linealidades y exógenas) mediante ponderación basada en errores históricos, se puede reducir sesgos de cada modelo y mejorar la precisión de los pronósticos.

Métodos de combinación: ponderación lineal de predicciones:

$$\hat{y} = w_1 \hat{y}_{\text{Prophet}} + w_2 \hat{y}_{\text{CatBoost}}$$

con pesos w_i determinados por inversa de RMSE en validación o mediante optimización (e.g., mínimos cuadrados).

Ventajas:

- puede mejorar robustez y generalización;
- mitiga fallos cíclicos de uno u otro modelo en ciertas condiciones.

2.4.13. Enfoque multi-salida con CatBoost

Para predecir 24 horas futuras de desvíos, se pueden usar:

- **Modelos independientes:** un CatBoost por cada horizonte h , minimizando

$$\sum_i L(y_{i,t+h}, F_h(x_i)).$$

Esto ignora correlaciones entre horas.

- **Modelo “horizonte como feature”:** reformular el problema como un dataset “aplanoado”: cada fila es un par (x_i, h) y objetivo $y_{i,t+h}$. Entonces se entrena un único modelo $F(x, h)$ que capta información compartida entre horizontes (Hyndman & Athanasopoulos, 2018).
- **Multi-output directo:** si la librería o un wrapper lo permite, entrenar un solo CatBoost que devuelva un vector de 24 predicciones simultáneas, aprovechando correlaciones entre outputs.

2.4.14. Métricas de error en forecasting

Root Mean Squared Error (RMSE) mide la raíz cuadrada del promedio de los errores al cuadrado entre valores reales y_i y predicciones \hat{y}_i :

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

Es sensible a errores grandes debido al cuadrado de la desviación y se expresa en las mismas unidades de la variable de interés (Hyndman & Athanasopoulos, 2018).

Symmetric Mean Absolute Percentage Error (SMAPE) valora el error porcentual simétrico y mitiga el sesgo cerca de cero:

$$\text{SMAPE} = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{\frac{1}{2}(|y_i| + |\hat{y}_i|)}.$$

Penaliza por igual sobre- y subestimaciones y resulta útil cuando los valores reales pueden ser cercanos a cero (Hyndman & Athanasopoulos, 2018).

2.4.15. Métricas de clasificación de tendencia

Para evaluar la capacidad del modelo de predecir la dirección del desvío (positivo/negativo), se utiliza la **matriz de confusión**:

	Predicción Positiva	Predicción Negativa
Real Positiva	Verdaderos Positivos (VP)	Falsos Negativos (FN)
Real Negativa	Falsos Positivos (FP)	Verdaderos Negativos (VN)

De ella se derivan numerosas métricas de rendimiento (precisión, exhaustividad, F1, etc.) (Hastie et al., 2009).

Precisión se define como la proporción de predicciones correctas (tanto positivas como negativas) sobre el total de observaciones:

$$\text{Precisión} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{FP} + \text{FN} + \text{VN}}.$$

Aunque intuitiva, puede resultar engañosa en casos de clases muy desbalanceadas (Hastie et al., 2009).

Capítulo 3. OBJETIVOS

3.1 Objetivos generales

El objetivo general del presente Trabajo Fin de Máster consiste en desarrollar e implementar un modelo predictivo del volumen neto de los desvíos generación/demanda en la red eléctrica peninsular española, mediante la comparación de un algoritmo de machine learning (CatBoost) y un modelo de series temporales (Prophet) con variables exógenas, con el fin de optimizar la estrategia de compra y venta en el mercado intradiario y minimizar las penalizaciones económicas asociadas a los desequilibrios.

3.2 Objetivos específicos

- Realizar análisis exploratorio de datos, a nivel univariante, para comprender las principales características de la serie temporal objetivo, y, por otra parte, realizar análisis multivariante con las potenciales variables exógenas para detectar el grado de correlación entre las variables, patrones y tendencias.
- Implementar un modelo de boosting (CatBoost) y un modelo de series temporales (Prophet), incorporando las variables más significativas, y, además, descifrar el mejor conjunto de hiperparámetros de cada modelo, mediante grid-search y validación cruzada de series temporales.
- Comparar el RMSE obtenido de cada modelo con el conjunto de validación, después de obtener el mejor conjunto de hiperparámetros para cada modelo mediante la optimización con el conjunto de train. Finalmente, poner en producción el mejor modelo para este estudio y evaluar con el conjunto de test su desempeño.

3.3 Beneficios del proyecto

Este proyecto aporta a las comercializadoras una herramienta predictiva capaz de anticipar con alta precisión el volumen neto de los desvíos generación/demanda en la red eléctrica peninsular. Con ella, los agentes podrán ajustar proactivamente sus posiciones en el mercado intradiario, minimizando las penalizaciones asociadas a desequilibrios.

Además, al aplicar los ajustes necesarios en el mercado intradiario de forma proactiva, los agentes participantes contribuyen a mantener la estabilidad del sistema eléctrico, lo que refuerza su seguridad.

Capítulo 4. DESARROLLO DEL PROYECTO

4.1 Planificación del proyecto

Nº	Actividad	Inicio	Fin	Dur. (horas)
1	Revisión del estado del arte	01-04-2025	05-04-2025	10
2	Recopilación y unificación de datos	06-04-2025	28-04-2025	40
3	Preprocesamiento y limpieza	29-04-2024	04-05-2025	10
4	Ingeniería de variables	05-05-2024	19-05-2025	20
5	Implementación de modelos	20-05-2024	25-05-2025	10
6	Validación y optimización	26-05-2024	30-05-2025	10
7	Evaluación y análisis económico	31-05-2024	04-06-2025	10
8	Redacción final y defensa	05-06-2024	02-09-2025	40

Tabla 4.1. Cronograma de actividades y esfuerzo del proyecto

- **Revisión del estado del arte:** Búsqueda y análisis crítico de proyectos e investigaciones relacionados con modelos de machine learning y modelos estadísticos, cuyo objetivo es el pronóstico de series temporales asociadas a la red eléctrica.
- **Recopilación y unificación de datos:** Descarga de series temporales de la red eléctrica, datos meteorológicos y datos de los días festivos del país. Cálculo de medias ponderadas de dichos datos por población para obtener una medida general a nivel nacional.
- **Preprocesamiento y limpieza:** Imputación de valores faltantes, unificación de datos y detección de valores atípicos.
- **Ingeniería de variables:** Generación de regresores temporales (hora, estación), agregación de variables meteorológicas y análisis univariantes y multivariantes para detectar patrones en las series temporales.
- **Implementación de modelos:** Desarrollo de pipelines CatBoost y Prophet con inclusión de variables exógenas.
- **Validación y optimización:** Validación “walk-forward” y búsqueda de hiperparámetros por grid-search respetando el orden temporal.
- **Evaluación de los modelos:** Cálculo del RMSE de los modelos y estimación del volumen neto de los desvíos, con análisis de costes por penalizaciones.
- **Redacción final y defensa:** Escritura de capítulos, revisión ortográfica y preparación de la presentación de defensa.

4.2 Descripción de la solución, metodologías y herramientas empleadas

4.2.1. Recopilación de datos

Para construir el conjunto de datos necesario para entrenar y evaluar los modelos predictivos, se reunieron las siguientes fuentes, cubriendo el mismo horizonte temporal (01-01-2015 a 30-11-2024) para cada variable:

- **Datos de desvíos, generación eólica y precio diario:** descargados manualmente desde el portal ESIOS de Red Eléctrica de España, usando la funcionalidad de exportación a CSV que ofrece la interfaz web (REE, 2024).
- **Variables meteorológicas (temperatura, horas de insolación):** obtenidas vía API del portal de datos abiertos de AEMET, mediante consultas automáticas que devolvían series horarias por provincia (AEMET, 2024).
- **Días festivos:** extraídos mediante web-scraping de los calendarios laborales provinciales y nacionales (CalendariosLaborales, 2024), asignando un indicador binario y ponderado por población.
- **Población por provincia:** descargada en formato CSV desde la web del Instituto Nacional de Estadística (INE, 2024)), para calcular medias ponderadas de las variables climáticas.

A continuación se describen con más detalle los procedimientos de extracción y unificación:

1. ESIOS (desvíos, eólica, precio) Los datos horarios se obtuvieron manualmente seleccionando el rango de fechas deseado y pulsando “Exportar CSV” en el portal de ESIOS. De este modo se aseguró una descarga completa y homogénea de todas las series: desvíos netos de generación/demanda, previsión eólica y precios del mercado diario (REE, 2024). Cada serie se exportó y guardó en un fichero separado.

A continuación, se verificó la integridad y coherencia de los datos. Al importar los ficheros en Python, se realizó una inspección inicial de cada DataFrame, detectándose inconsistencias debidas a cambios de hora (horario de verano/invierno) que provocaban saltos temporales. Para corregirlo, la columna de fecha y hora se normalizó eliminando la información de zona horaria y convirtiendo todas las marcas temporales a UTC, garantizando así una referencia única y continua para el análisis.

2. API AEMET (clima) Utilizando la librería `requests` de Python, se desarrolló un script que, para cada provincia española, consultaba la API de AEMET y extraía la temperatura media diaria y las horas de insolación. Cada respuesta JSON se parseaba y volcaba a un DataFrame de pandas.

A continuación, estos archivos se cargaron en Python para inspeccionar la calidad de los datos. Como se debía combinar esta información con la población provincial, se acordó usar como clave de unión el nombre de la provincia y el año. Para ello, primero se estandarizó el campo `provincia`: todos los nombres se convirtieron a minúsculas, sin espacios ni caracteres especiales. Este mismo criterio de limpieza se aplicó más adelante a otros conjuntos de variables.

Durante la inspección se detectaron valores faltantes en las columnas de temperatura media e insolación. Para cuantificar el problema, se calculó el porcentaje de NaN en cada serie. En el caso de temperatura, los valores ausentes eran escasos y dispersos, por lo que se resolvió imputándolos con el valor inmediatamente anterior (forward-fill), es decir:

$$T(t) = T(t - 1).$$

Para la variable de insolación, varios intervalos de tiempo carecían por completo de datos en ciertas provincias, por lo que el método forward-fill no era adecuado. En este caso, cada NaN se reemplazó por el promedio de las demás provincias en esa misma fecha. Si en la fecha t había datos de insolación en 14 provincias y faltaban en 5, dichos cinco valores se imputaron con la media calculada de las 14 provincias disponibles.

Una vez unificados los nombres y corregidos los valores faltantes, se creó una nueva columna concatenando año y provincia limpia. Esta clave permitió fusionar el DataFrame meteorológico con el de población y, a partir de ahí, calcular las medias ponderadas de temperatura e insolación a nivel nacional.

3. Web-scraping de festividades Utilizando la librería BeautifulSoup de Python, se recorrieron las páginas de calendarios laborales provinciales. Se identificaron las fechas festivas y se generó un indicador numérico:

$$\text{festivo}_{\text{provincia}}(t) = \begin{cases} 1, & \text{si es festivo nacional o domingo,} \\ 0,5, & \text{si es sábado y no es festivo nacional,} \\ \frac{P_{\text{prov}}}{P_{\text{total}}}, & \text{si es festivo regional.} \end{cases}$$

Es importante resaltar que, una vez descargados los datos y unificados en un solo dataframe, se estandarizó la columna "provincia" para así poder crear la columna entre el año y la provincia.

4. Descarga de población (INE) Se descargó un fichero CSV con la población de cada provincia para el intervalo de tiempo de 2015–2024. Estos valores permitieron calcular los pesos demográficos $\frac{P_{\text{prov}}}{P_{\text{total}}}$ utilizados en la agregación de variables climáticas y días festivos, como se explicó en los apartados anteriores.

Con estas cuatro fuentes se dispuso de un dataset unificado, alineado en frecuencia horaria y geográfica, listo para su limpieza y posterior análisis.

4.2.2. Análisis exploratorio de datos

Análisis univariante temporal y estacional Para comprender el comportamiento de la serie de volumen neto de desvíos, en primer lugar se generó un conjunto de variables exógenas a partir de la marca temporal. De la columna de fecha y hora se obtuvieron las variables: hora, día del mes, semana del año, día de la semana, mes, trimestre y semestre. Estas características permitieron investigar posibles patrones y estacionalidades en la serie principal.

Acto seguido, se trazó gráficamente la variable objetivo utilizando matplotlib y seaborn. La serie temporal se representó con un gráfico de líneas, su distribución se examinó mediante un histograma y se evaluó el ajuste a la normalidad con un QQ-plot. Esta exploración inicial ofreció una visión clara de tendencias, picos y colas de la distribución.

Para determinar la estacionariedad, se aplicó el test de Dickey–Fuller aumentado usando el modulo de Python "statsmodels". El estadístico resultante y su p-valor permitieron evaluar si la media y la varianza se mantenían constantes a lo largo del tiempo. A continuación, se empleó la prueba de Ljung–Box para verificar la autocorrelación serial y detectar dependencias significativas entre valores separados por distintos lags.

Asimismo, la normalidad de la serie de desvíos se contrastó mediante las pruebas de Shapiro–Wilk y Kolmogórov–Smirnov, utilizando el módulo de Python scipy. Estas pruebas permitieron evaluar si la suposición de normalidad era razonable o si resultaba necesario aplicar transformaciones a los datos. Cabe destacar que, si bien los modelos empleados en esta investigación no requieren estrictamente que se cumpla el supuesto de normalidad, conocer si la distribución empírica se aproxima a una normal resulta igualmente importante, dado que esta información puede ser crucial para la interpretación de resultados y la aplicación de técnicas estadísticas complementarias.

Para profundizar en las correlaciones temporales, se generaron los correlogramas (ACF y PACF). Mediante ellos se identificaron los retardos (lags) en los que los valores pasados de la serie o sus errores explicaban parte de la varianza futura, lo cual sirve de guía para la selección de órdenes en modelos ARIMA o para el diseño de ventanas de retardos en algoritmos de machine learning.

Posteriormente, se calcularon estadísticas descriptivas básicas y avanzadas: media, mediana, percentiles 25 y 75, mínimo, máximo, desviación típica, coeficiente de asimetría y curtosis. Este resumen cuantitativo complementó la inspección gráfica y los contrastes de hipótesis, revelando posibles sesgos, dispersión y colas en la distribución de los desvíos, y orientó la elección de las familias de modelos predictivos más adecuadas.

A fin de suavizar las fluctuaciones diarias y poner de relieve patrones de mayor escala, la serie de volumen neto de desvíos se re-muestreó con frecuencia semanal y se representó año por año. Este procedimiento permitió atenuar la variabilidad a corto plazo y, al comparar cada temporada anual por separado, identificar posibles comportamientos recurrentes en determinados meses.

Seguidamente, se generaron mapas de calor (heatmaps) del volumen neto de desvíos usando como ejes el día de la semana y la hora del día. De este modo se visualizaron con claridad las franjas horarias y jornadas en que los desequilibrios tienden a concentrarse. Para profundizar en la estacionalidad semanal y mensual, se elaboró un heatmap global —abarcando todo el intervalo 2015–2024— y, además, un heatmap individual para cada mes del año. Cabe destacar, que estas representaciones gráficas se analizarán en la sección de resultados, donde se destacarán los picos de desvío asociados a horas punta y a eventos estacionales.

Análisis multivariante temporal y estacional Con el fin de evaluar las relaciones entre la variable objetivo (volumen neto de desvíos) y el conjunto de variables exógenas, así como de identificar posibles problemas de multicolinealidad, se siguió el siguiente protocolo metodológico:

Se construyó una matriz de correlación de Pearson que incluyó todas las variables numéri-

cas (regresores temporales, climáticos y de mercado). En los casos de alta correlación entre regresores, se seleccionó el predictor con mayor correlación con la serie de desvíos, garantizando así un conjunto de variables explicativas parsimonioso. Sin embargo, no se asegura que el conjunto de datos esté libre de multicolinealidad, pero sí se reduce considerablemente el riesgo de que esta se presente.

Por otra parte, a partir de las variables exógenas originales, se crearon un conjunto de nuevos atributos con el fin de facilitar la captura de patrones no lineales y profundizar los análisis exploratorios:

- **Tipo de clima:** basado en la distribución de la temperatura media. Se calcularon los percentiles 25, 50 y 75 de la serie; los valores por debajo del p25 se etiquetaron como “frío”, los entre p25 y p50 como “templado”, los entre p50 y p75 como “cálido” y los superiores a p75 como “muy cálido”.
- **Grupo_hora:** resultado del análisis de los mapas de calor. Las horas entre las 07:00 y las 20:00 se agruparon bajo la etiqueta “día” y el resto como “noche”, para diferenciar patrones de comportamiento diurno y nocturno.
- **Estación del año:** categorización de los meses en las cuatro estaciones: primavera (marzo–mayo), verano (junio–agosto), otoño (septiembre–noviembre) e invierno (diciembre–febrero).
- **Festivo (agrupado):** versión no ponderada de la variable festivo, donde sábados y festivos regionales se consideran un único grupo “medio festivo” y los festivos nacionales otro grupo “festivo”.
- **Rezago 1h del volumen de desvíos:** variable de retraso de una hora de la serie objetivo, incorporada tras observar en los correlogramas que el lag 1 aportaba información significativa para la predicción.

Con estas transformaciones se amplió el espacio de características, facilitando a los modelos la detección de efectos estacionales, horarios y de temperatura sobre el volumen neto de desvíos.

Sin embargo, una inspección visual resultaba pertinente para analizar cómo las nuevas características capturaban la variabilidad del volumen neto de desvíos y las relaciones con las exógenas. Para ello, se siguió este procedimiento:

- **Diagramas de caja (boxplots)** del volumen neto de desvíos agrupado por *tipo de clima*, *estación del año* y *grupo_hora*. Estos gráficos permiten apreciar de un vistazo la dispersión, los cuartiles y la existencia de valores atípicos en cada categoría.
- **Gráficas de líneas** comparativas entre el volumen neto de desvíos y variables exógenas clave (previsión eólica y precio del mercado diario). Cada par de series se representó con distintas agregaciones temporales: diaria, semanal, mensual, trimestral, semestral y anual, con el fin de revelar tendencias y sincronías a diferentes escalas.
- **Escalado de variables:** antes de trazar las series, todas las variables numéricas (volumenes, precios, temperatura, viento, horas de insolación) se normalizaron para compartir rango y unidad estadística. De este modo, las gráficas de líneas comparativas resultaron más informativas y facilitaron la detección de correlaciones.
- El mismo proceso de escalado y visualización se aplicó a pares de variables explica-

tivas (temperatura vs. velocidad del viento, temperatura vs. insolación), con objeto de identificar co-movimientos y posibles redundancias.

Estas representaciones gráficas, junto con el escalado previo, proporcionan una visión clara de los patrones conjuntos y sirven de guía para la selección y parametrización de los modelos predictivos.

Para concluir con los análisis exploratorios se considero cuantificar la dependencia temporal entre el volumen neto de desvíos y la temperatura media. Para ello, se empleó un análisis de correlación cruzada (cross-correlation). Mediante una función implementada en Python, se generó un vector de retardos $\ell \in [-24, 24]$ y, para cada ℓ , se calculó el coeficiente de correlación de Pearson entre la serie de desvíos y la serie de temperatura desplazada ℓ horas. Este enfoque permite determinar qué valores pasados de la temperatura (lags positivos) guardan mayor asociación con el valor actual de los desvíos, y viceversa (lags negativos).

Cabe destacar, que el mismo protocolo metodológico se replicó para otras variables exógenas relevantes: previsión eólica, precio del mercado diario e insolación horas. Para cada una, se construyó la curva de correlación cruzada en el rango de -24 a $+24$ horas, de modo que los picos en lags positivos indiquen los desfases temporales más informativos para la predicción del volumen neto de desvíos y sirvan de guía para la selección de rezagos en los modelos predictivos.

Modelado con CatBoost y análisis de importancia de variables Para cuantificar la contribución de cada variable explicativa a la predicción del volumen neto de desvíos, se entrenó un modelo de CatBoost Regressor siguiendo este procedimiento:

1. **Partición temporal de los datos:** el conjunto completo se dividió en dos subconjuntos entrenamiento (90 %) y prueba (10 %)—respetando el orden cronológico para evitar filtrado de información futura.
2. **Selección de características:** las variables predictoras se agruparon en dos listas:
 - *Númericas:* temperatura, insolación, velocidad del viento, precio de mercado, previsión eólica, y variables temporales (hora, día, mes, semana, trimestre, semestre, año) junto al rezago de desvíos a una hora.
 - *Catégoricas:* tipo de clima, grupo_hora, estación y festivo.
3. **Configuración del modelo:**
 - Iteraciones: 500
 - Tasa de aprendizaje (learning rate): 0.05
 - Profundidad máxima de árboles: 6
 - Función de pérdida y métrica de evaluación: RMSE
 - Early stopping con 50 rondas de paciencia, usando el conjunto de validación

Estos valores se eligieron para garantizar que el modelo entrenara lo suficiente como para estabilizar las importancias sin entrar en una búsqueda exhaustiva de hiperparámetros. Un número alto de iteraciones (500) asegura que cada variable tenga oportuni-

dad de aportar a la reducción del error, mientras que una tasa de aprendizaje moderada (0.05) y una profundidad limitada (6) evitan un sobreajuste excesivo. El uso de RMSE como función de pérdida y métrica de evaluación permite cuantificar directamente la desviación media en las mismas unidades de la serie, y el early stopping con paciencia de 50 rondas previene el entrenamiento innecesario una vez que el RMSE en validación deja de mejorar.

4. **Entrenamiento:** se construyó un `Pool` de CatBoost para entrenamiento y validación, indicando explícitamente cuáles variables eran categóricas, y se ajustó el modelo optimizando el RMSE sobre el conjunto de validación.

CatBoost maneja internamente las variables categóricas mediante codificaciones basadas en estadísticas de la propia variable objetivo, por lo que pasar un `Pool` con esa información permite explotar esa característica sin convertir manualmente cada categoría en dummies (Prokhorenkova et al., 2018). Optimizar sobre el RMSE en validación asegura que las predicciones futuras conserven la capacidad de capturar la tendencia y los patrones generales de la serie, ya que la evaluación se realiza con datos que no se usaron directamente para ajustar los pesos del modelo (Prokhorenkova et al., 2018).

5. **Extracción de importancia:** una vez finalizado el fit, se obtuvo el *feature importance* de CatBoost y se representó gráficamente mediante un diagrama de barras horizontales para visualizar qué variables aportan con mayor peso al poder predictivo.

Este enfoque permitió identificar de forma robusta las características más relevantes —tanto numéricas como categóricas— y sirvió de base para depurar el conjunto de regresores antes de la fase final de evaluación.

Es importante señalar que este procedimiento fue llevado a cabo con el lenguaje de programación Python, específicamente con los módulos `sklearn` y `catboost`.

Una vez realizado el análisis de importancia de variables, se procedió con la configuración del modelo CatBoost.

Entrenamiento y ajuste de hiperparámetros con CatBoost Para optimizar la capacidad predictiva de CatBoost sobre el volumen neto de desvíos, se aplicó el siguiente protocolo:

1. **Definición de variables.** Se estableció como variable objetivo `volumen_netos_desvios` y como predictores un conjunto mixto de características numéricas y categóricas:
 - Numéricas: volumen neto desvíos lag1, hora, año, temperatura en celsius, insolación horas, semana, precio mercado, mes, prevision eólica mw.
 - Categóricas: grupo hora.
2. **Partición temporal.** El DataFrame completo se dividió en dos subconjuntos cronológicos:
 - Entrenamiento y Validación (90 % de las observaciones).
 - Prueba (10 %).

Esta división asegura que no haya filtrado de información futura durante el ajuste de hiperparámetros.

3. **Validación cruzada para series temporales.** Se empleó `TimeSeriesSplit` con 5 particiones para respetar la dependencia temporal durante la búsqueda de parámetros.
4. **Grid search de hiperparámetros.** Se configuró un espacio de búsqueda que incluía:
 - `iterations`: [300, 500]
 - `learning_rate`: [0.05, 0.1]
 - `depth`: [4, 6, 8]
 - `l2_leaf_reg`: [1, 3, 5]
 - `bagging_temperature`: [0, 1, 3]

Usando `GridSearchCV` con `scoring` de RMSE (negativo para maximizar) y procesamiento en paralelo, se identificaron los valores de hiperparámetros que minimizan el error de validación.

Con el objetivo de poder comparar estos resultados con las diferentes metodologías, se extrajo el RMSE positivo y, además, se calculó el SMAPE, para de esa forma obtener el error porcentual.

5. **Entrenamiento del modelo final.** Con los mejores parámetros hallados, se reentrenó CatBoost sobre el conjunto completo de entrenamiento (90 % de los datos) y se calculó la precisión del modelo para capturar los signos de los desvíos, ya sean positivos o negativos. Posteriormente, se obtuvo la precisión del modelo (accuracy).

Modelado con Prophet El protocolo aplicado al modelo Prophet resultó muy similar al del modelo CatBoost, en cuanto a la definición de variables, la partición temporal y la validación cruzada para series temporales. Esto se debió a que se buscó una metodología similar con el propósito de comparar la efectividad de ambos modelos; por lo tanto, dichos procedimientos se llevaron a cabo de forma similar.

No obstante, los parámetros e hiperparámetros del modelo Prophet son distintos y, por lo tanto, requieren un proceso diferente.

Para afinar los hiperparámetros del modelo se utilizó un grid con tres valores para la flexibilidad de la tendencia y tres para la fuerza de la estacionalidad.

- `changepoint_prior_scale`: [0.01, 0.05, 0.1]
- `seasonality_prior_scale`: [1.0, 10.0, 20.0]

Posteriormente, se emplearon cinco particiones temporales sucesivas para aplicar validación cruzada, en las que cada partición entrenaba al modelo con los datos más antiguos y se validaba con el conjunto de datos inmediatamente posterior. Después de recorrer las particiones, se calculaban el RMSE y el SMAPE promedio provenientes de cada conjunto de evaluación; de esa forma, se obtenía una métrica robusta, capaz de prevenir el sobreajuste.

Además, luego de haberse ejecutado cada una de las posibles combinaciones de hiperparámetros y de haberse obtenido, para cada combinación, sus respectivas métricas de bondad

de ajuste, se escogió el RMSE más bajo y sus respectivos hiperparámetros.

Para evaluar qué tan bien el modelo Prophet, ya calibrado con los mejores parámetros, captura la dirección de los desvíos, se reentrenó con el conjunto de datos de entrenamiento. Una vez ajustado el modelo final, se generaron predicciones sobre el mismo conjunto de datos para obtener la precisión del pronóstico de la tendencia de la serie; es decir, identificar si el volumen neto de los desvíos es negativo o positivo y compararlo con modelos previamente configurados (como el modelo CatBoost).

Modelo híbrido Catboost-Prophet Con el objetivo de mejorar la precisión de los pronósticos, se combinaron las predicciones de CatBoost y Prophet mediante un ensamblado ponderado, de forma tal que el modelo más preciso recibiera mayor peso. Primero, se calculó para cada modelo un peso proporcional a la inversa de su RMSE obtenido mediante validación cruzada, es decir:

$$w_i = \frac{1}{\text{RMSE}_i}, \quad w_i^{(n)} = \frac{w_i}{\sum_j w_j},$$

Posteriormente, se normalizaron los pesos para que sumaran 1. Con los pesos normalizados, se generó la predicción híbrida en cada instante mediante la siguiente fórmula:

$$\hat{y}_{\text{híbrido}} = w_{\text{Cat}}^{(n)} \hat{y}_{\text{CatBoost}} + w_{\text{Prop}}^{(n)} \hat{y}_{\text{Prophet}}.$$

Después, se calculó el RMSE de esta nueva serie y la precisión que tuvo el modelo para capturar las tendencias, con el propósito de corroborar si la unión de ambos modelos derivaba en un modelo más robusto y eficiente.

Modelo Catboost Multi-salida Con el propósito de implementar el enfoque multi-salida con CatBoost y poder pronosticar las próximas 24 horas del volumen neto de desvíos, se llevó a cabo una transformación de datos en la que cada fila corresponde a la referencia de las 00:00 de un día. Por otra parte, se tomaron los valores rezagados de la variable objetivo como variables explicativas—en este caso, las últimas 48 horas. En paralelo, se incluyeron las variables exógenas utilizadas en los modelos anteriores, con la excepción del rezago de una hora. Como variable objetivo, se tienen los valores reales de los desvíos desde las 00:00 hasta las 23:00 horas. Cabe destacar que cada hora corresponde a una columna.

Además, se aplicó una transformación one-hot encoding a las variables categóricas y, posteriormente, todas las variables explicativas se transformaron al tipo float. A continuación, se dividió el conjunto de datos en entrenamiento y prueba con una proporción de 90 % y 10 %, respectivamente. Después, se separaron las variables explicativas del conjunto de variables objetivo.

Cabe destacar que el pipeline de preprocesamiento incluyó una estandarización y, dado que existe una alta dimensionalidad, se agregó un PCA manteniendo el 95 % de la varianza de los

datos. Además, se realizó una validación cruzada con 3 particiones respetando el orden cronológico. Por último, después de haber seleccionado el mejor conjunto de hiperparámetros, se entrenó el conjunto de entrenamiento para obtener el modelo final.

Con el modelo final entrenado, se evaluó su precisión con los datos de prueba, obteniendo las métricas RMSE y SMAPE para cada hora; es decir, se obtuvieron 24 valores por métrica. Adicionalmente, se evaluó la precisión para estimar si el desvío sería negativo o positivo.

Simulación de penalizaciones y gastos Para poner a prueba el valor de las predicciones obtenidas por el modelo multi-salida, se realizó una simulación de cómo habría operado una comercializadora siguiendo la estrategia sugerida por el modelo, comparándola contra una estrategia de compra con un valor constante.

Para ello, se construyó un conjunto de datos con el precio del mercado, el precio de las penalizaciones por desvíos al alza y a la baja, las predicciones del volumen neto de los desvíos y los valores reales, dentro del intervalo de tiempo correspondiente a los datos de prueba.

Una vez cargados los datos y configurado el conjunto (dataset), se calcularon las penalizaciones de la siguiente manera:

$$\pi_{\text{subir}}(t) = p_{\text{spot}}(t) - p_{\text{desvíos, subir}}(t), \quad \pi_{\text{bajar}}(t) = p_{\text{desvíos, bajar}}(t) - p_{\text{spot}}(t).$$

$$C(t) = \begin{cases} |\Delta v(t)| \pi_{\text{subir}}(t), & \Delta v(t) \leq 0, \\ \Delta v(t) \pi_{\text{bajar}}(t), & \Delta v(t) > 0, \end{cases}$$

Posteriormente, se simuló la compra, en donde se plantearon 2 estrategias. En la primera, la comercializadora compraría un volumen constante de 100 MW cada hora, sin importar la volatilidad de los desvíos. Por otra parte, si se consideran los desvíos pronosticados como parte de la estrategia de compra, la comercializadora comprará 100 MW cada hora más la centésima parte de la desviación pronosticada. Cabe destacar que se toma la centésima parte porque los desvíos pronosticados son del sistema eléctrico de España y no de un agente en específico. Por lo tanto, solo se toma una parte de los mismos.

El cálculo de la compra se expresa de la siguiente forma:

$$\text{compra_con_estrategia}(t) = L + \frac{\hat{v}(t)}{100} \quad (4.1)$$

- $L = 100$ MW
- $\hat{v}(t)$ = volumen neto de desvíos predicho en la hora t

A continuación, se generó un vector de desviaciones a partir de una distribución normal para emular variabilidad en el consumo real de la empresa. Los parámetros de la distribución son ficticios y se seleccionaron con el criterio de que la probabilidad de observar un desvío supe-

rior a 45 MW sea aproximadamente del 0,3 %. A partir de ese valor, se calculó el consumo real, el cual es el desvío real simulado más la compra base (100 MW). Teniendo el consumo real, se determinó el desvío con estrategia; el mismo es la diferencia entre el consumo real menos la compra con estrategia. Expresado en fórmulas, serían las siguientes:

$$v_{\text{real}}(t) \sim \mathcal{N}(0, 15^2), \quad (4.2)$$

$$c_{\text{real}}(t) = v_{\text{real}}(t) + L, \quad (4.3)$$

$$\Delta v_{\text{con_estrategia}}(t) = c_{\text{real}}(t) - \text{compra_con_estrategia}(t) \quad (4.4)$$

Por último, a partir del desvío con estrategia, se obtuvo el coste de las operaciones con estrategia y el coste sin estrategia. Dichos cálculos se pueden representar mediante las siguientes fórmulas:

$$C_{\text{original}}(t) = \begin{cases} |v_{\text{real}}(t)| \pi_{\text{subir}}(t), & v_{\text{real}}(t) \leq 0, \\ v_{\text{real}}(t) \pi_{\text{bajar}}(t), & v_{\text{real}}(t) > 0, \end{cases} \quad (4.5)$$

$$C_{\text{con_estrategia}}(t) = \begin{cases} |\Delta v_{\text{con_estrategia}}(t)| \pi_{\text{subir}}(t), & \Delta v_{\text{con_estrategia}}(t) \leq 0, \\ \Delta v_{\text{con_estrategia}}(t) \pi_{\text{bajar}}(t), & \Delta v_{\text{con_estrategia}}(t) > 0. \end{cases} \quad (4.6)$$

4.3 Recursos requeridos

A continuación se enumeran los recursos utilizados para la ejecución de este proyecto:

- **Hardware**
 - Ordenador personal con procesador Ryzen 5, 8 GB de RAM y almacenamiento SSD.
 - Conexión a Internet para descarga de datos y librerías.
- **Software**
 - Python 3.10.0 y entornos virtuales (venv) para instalación de librerías.
 - Librerías de análisis y modelado:
 - pandas, NumPy, scikit-learn
 - CatBoost, fbprophet
 - matplotlib, seaborn
 - VS Code como entorno de desarrollo.
 - Distribución \LaTeX (TeX Live / Overleaf) para la redacción del documento.
- **Datos**
 - Series de volumen neto de desvíos y variables exógenas obtenidas de ESIOS, AEMET y scraping web.
 - Precios del mercado spot e intradiario descargados de OMIE y REE.
- **Asistencia de expertos**
 - Tutor académico, con reuniones periódicas para revisión de avances metodológi-

cos y de redacción.

4.4 Resultados del proyecto

En esta sección se presentan los resultados obtenidos durante la fase de análisis de datos y modelado. El objetivo es mostrar de forma ordenada y precisa los siguientes apartados:

- **Análisis univariante:** Se analiza la serie de volumen neto de desvíos, incluyendo pruebas estadísticas de hipótesis y visualizaciones a diferentes escalas temporales.
- **Análisis multivariante:** Se examinan las relaciones entre la variable objetivo y las variables exógenas.
- **Evaluación de modelos:** Se emplean métricas de precisión y comparativos entre los modelos Prophet, CatBoost y un modelo híbrido. También se evalúa el modelo CatBoost con enfoque multi-salida.
- **Simulación económica:** Se calcula el impacto financiero de aplicar la estrategia de predicción frente a una política de compra constante.

4.4.1. Análisis univariante

Estadísticas descriptivas

Se presentan los estadísticos descriptivos correspondientes al volumen neto de los desvíos. Dichos indicadores permiten la comprensión de su distribución, tendencia central, dispersión y forma de la variable, proporcionando una base para entender su comportamiento y orientar los análisis posteriores.

Tabla 4.2. Estadísticos descriptivos del volumen neto de desvíos

Estadístico	Valor
Mínimo	-5315.600
Percentil 25	-374.300
Media	123.338
Mediana	130.900
Percentil 75	627.200
Máximo	5507.100
Asimetría	0.019
Curtosis	1.659
Coefficiente de variación	7.024
Desviación típica	866.366

El volumen neto de los desvíos presenta un amplio rango de valores, con un mínimo de -5315.60 y un máximo de 5507.10, lo cual indica una alta dispersión y la existencia de desviaciones extremas tanto negativas como positivas. La media (123.34) y la mediana (130.90) están relativamente cercanas, lo cual, junto al valor muy bajo de asimetría (0.019), sugiere que la distribución es casi simétrica, aunque con una ligera inclinación hacia valores más

altos. La variabilidad es considerable, como lo reflejan el coeficiente de variación (7.02) y la desviación típica (866.37). Con respecto a la curtosis, el valor calculado (1.66) corresponde al exceso de curtosis respecto a una distribución normal, lo que indica una forma leptocúrtica, con mayor concentración de datos cercanos a la media y colas más pesadas. Finalmente, los percentiles 25 (-374.30) y 75 (627.20) definen un intervalo central de dispersión moderada, en el que se concentra el 50 % de los datos.

Análisis gráfico

Se presentan a continuación tres visualizaciones clave que permitirán obtener información relacionada con la distribución de la variable de estudio, si es estacionaria o no, y si presenta algún tipo de sesgo. Cabe destacar que los estadísticos descriptivos fueron capaces de proveer cierta información. Por lo tanto, se espera que las visualizaciones complementen dichos análisis.

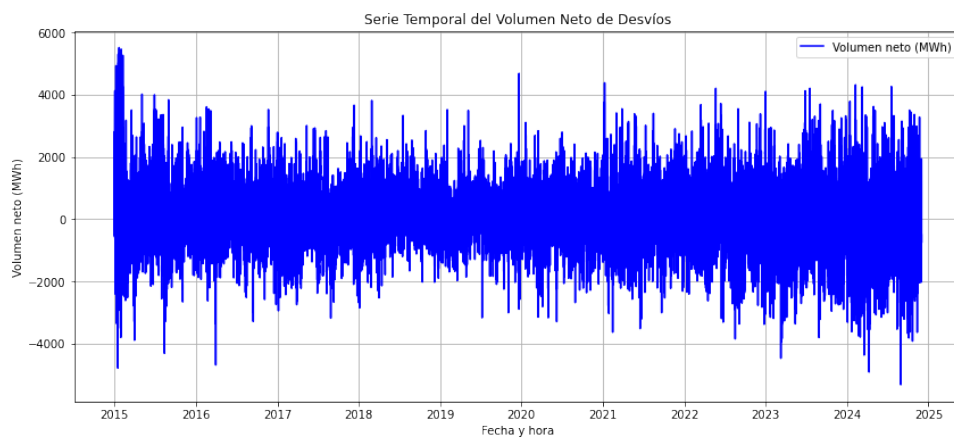


Figura 4.1. Serie temporal del volumen neto de desvíos.
Fuente: Cálculos propios.

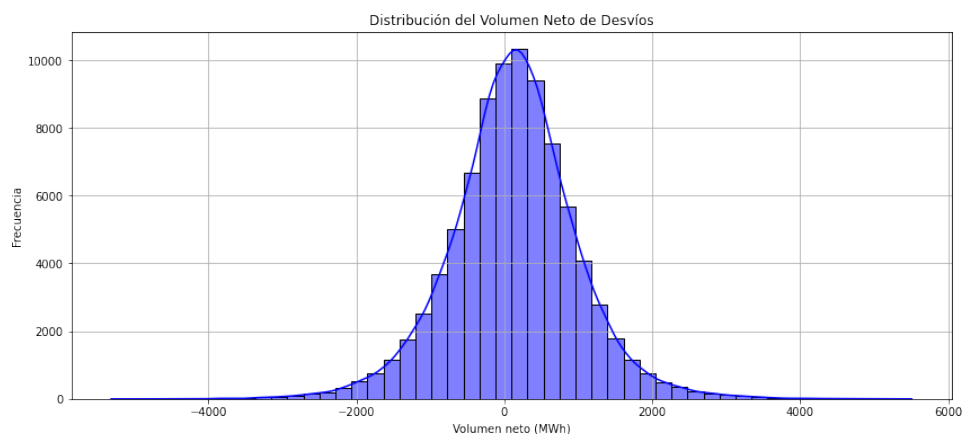


Figura 4.2. Histograma con densidad del volumen neto de los desvíos.
Fuente: Cálculos propios.

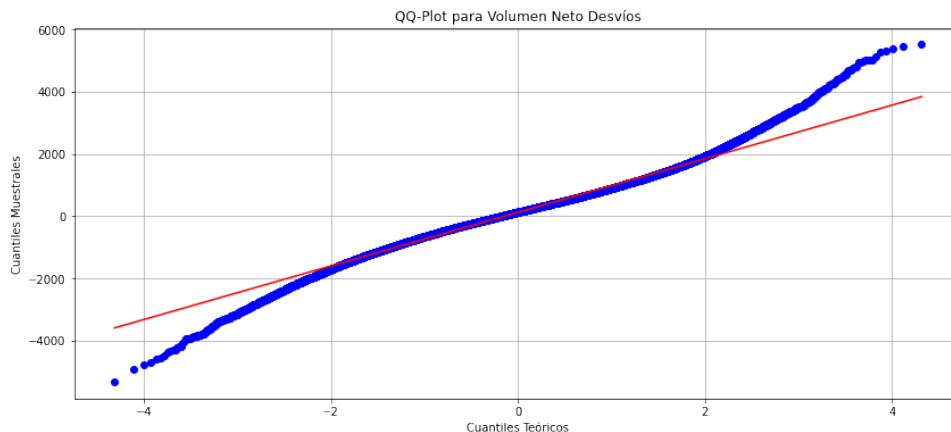


Figura 4.3. Gráfico Q-Q del volumen neto de los desvíos.

Fuente: Cálculos propios.

Se puede apreciar que la gráfica de serie temporal corrobora visualmente la alta dispersión y el amplio rango de valores, mostrando fluctuaciones constantes y la presencia de desviaciones extremas a lo largo del tiempo, sin un patrón estacional o tendencia clara aparente. Por otra parte, el histograma complementa el análisis previo en los estadísticos descriptivos de la casi simetría de la distribución y evidencia su forma leptocúrtica. Finalmente, el gráfico Q-Q refuerza esta conclusión al mostrar una clara desviación de los puntos respecto a la línea de normalidad en los extremos, lo que subraya la naturaleza no normal de la distribución y la existencia de colas pesadas.

ACF y PACF

La exploración de los correlogramas es esencial para diagnosticar autocorrelación en la serie de desvíos. La *función de autocorrelación* (ACF) muestra la correlación entre un valor y sus rezagos, incluidos los efectos indirectos que pasan por rezagos intermedios, mientras que la *función de autocorrelación parcial* (PACF) aísla únicamente la correlación directa de cada rezago, controlando esas influencias intermedias (Hyndman & Athanasopoulos, 2018).

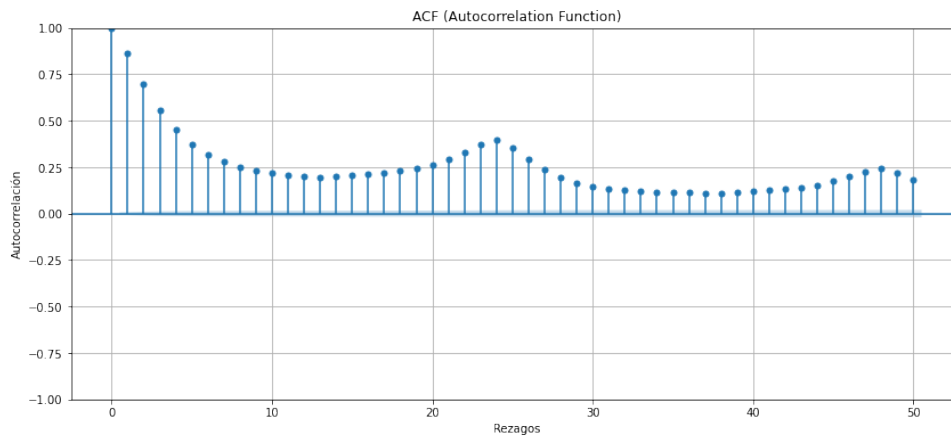


Figura 4.4. Función de autocorrelación (ACF).
Fuente: Cálculos propios.

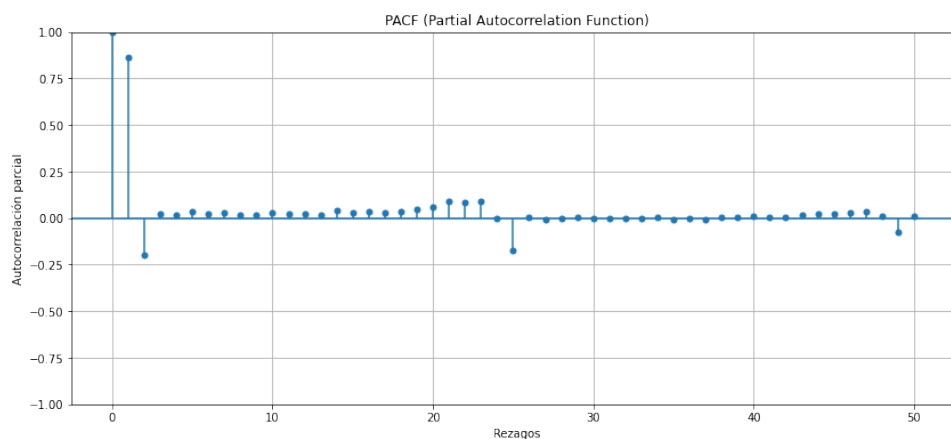


Figura 4.5. Función de autocorrelación parcial (PACF).
Fuente: Cálculos propios.

Se puede apreciar que ambos gráficos revelan indicios de autocorrelación en la serie. En el correlograma ACF se observa un decaimiento progresivo que se mantiene por encima de los límites de confianza hasta aproximadamente el rezago seis, lo cual indica memoria prolongada en la serie (efectos indirectos). En el PACF, únicamente el primer rezago supera el umbral de significación, mientras que los siguientes caen rápidamente al nivel de ruido. Este patrón de decaimiento lento en el ACF, junto con un único pico en el PACF, sugiere que un modelo autorregresivo de orden uno ($AR(1)$) podría capturar adecuadamente la dependencia lineal de corto plazo en el volumen neto de los desvíos.

Contraste de hipótesis

Es de vital importancia corroborar de manera rigurosa los análisis y conclusiones obtenidos anteriormente. Para ello, se llevaron a cabo una serie de pruebas estadísticas que permiten

determinar, con un nivel de significación adecuado, si los supuestos antes mencionados son verídicos o no.

A continuación, se presenta la prueba de Dickey–Fuller aumentada, la cual permite determinar si la serie temporal es estacionaria. Posteriormente, se realiza una prueba para evaluar si la serie presenta dependencia temporal; dicha prueba se denomina Ljung–Box. Por último, se verifica si la distribución empírica de la serie es similar a una distribución normal. Para ello, se empleó la prueba de Shapiro–Wilk.

ADF (Dickey–Fuller aumentada)

Hipótesis:

- H_0 : La serie tiene raíz unitaria (no estacionaria).
- H_1 : La serie es estacionaria.

Tabla 4.3. Resultados del test ADF (Dickey–Fuller aumentada) para $\alpha = 0,05$

Contraste	Estadístico	Región crítica	Nivel de significación	p-valor
ADF	$\tau = -26,10$	$\tau < -1,95$	0,05	$p = 0,000$

Conclusión: Con un nivel de significación del 5 %, se rechaza H_0 . Por lo tanto, se puede concluir que la serie temporal del volumen neto de los desvíos en el intervalo 2015–2024 es estacionaria.

Test de Ljung–Box

Hipótesis:

- H_0 : No existe autocorrelación significativa hasta el lag 10.
- H_1 : Existe autocorrelación significativa en al menos un lag ≤ 10 .

Tabla 4.4. Resultados del test de Ljung–Box (lag=10) para $\alpha = 0,05$

Contraste	Estadístico	Región crítica	Nivel de significación	p-valor
Ljung–Box	$\chi^2 = 193894,72$	$\chi^2 > 18,31$	0,05	$p = 0,000$

Conclusión: Con un nivel de significación del 5 %, se rechaza H_0 . Por lo tanto, se confirma que la serie presenta autocorrelación.

Shapiro–Wilk

Hipótesis:

- H_0 : La distribución de la variable es normal.
- H_1 : La distribución no es normal.

Tabla 4.5. Resultados del test de Shapiro–Wilk para $\alpha = 0,05$

Contraste	Estadístico	Región crítica	Nivel de significación	p-valor
Shapiro–Wilk	$W = 0,89$	$W < 0,975$	0,05	$p = 0,001$

Conclusión: Con un nivel de significación del 5 %, se rechaza H_0 . Se observa que la distribución del volumen neto de los desvíos no se ajusta a una normal.

Análisis a diferentes frecuencias temporales

Como complemento al análisis univariante de la serie horaria, se agruparon los datos de volumen neto de desvíos en frecuencias mayores (diaria, semanal, mensual, trimestral y anual). Esto permite detectar patrones de variación a distintas escalas temporales, como tendencias a largo plazo, estacionalidades o ciclos que no siempre resultan evidentes en la serie horaria.

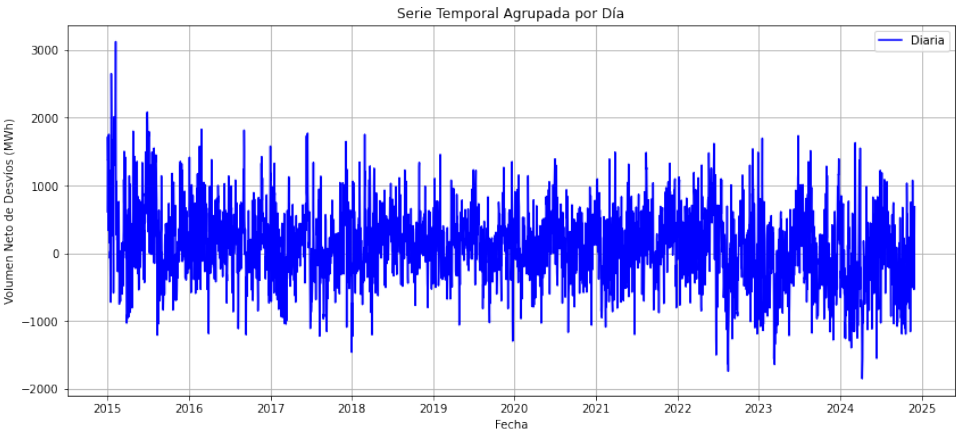


Figura 4.6. Serie diaria del volumen neto de desvíos.
Fuente: Cálculos propios.

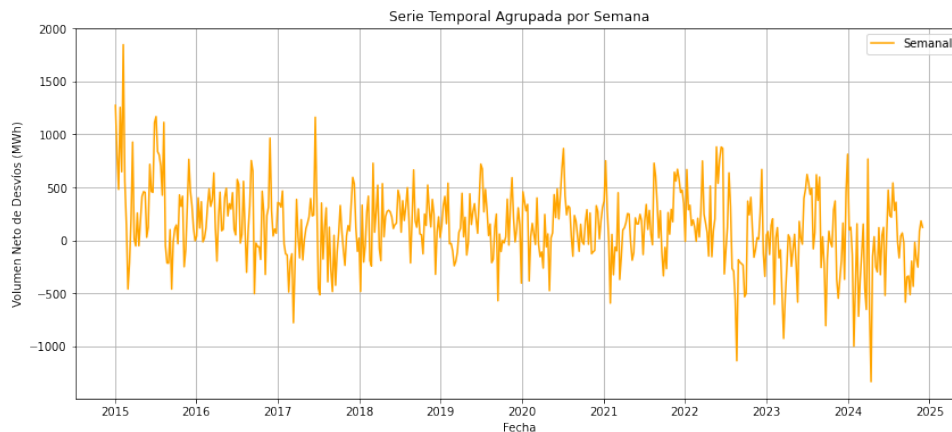


Figura 4.7. Serie semanal del volumen neto de desvíos.
Fuente: Cálculos propios.

Se puede observar que existe una leve tendencia bajista a través del tiempo. Es decir, que durante los años 2015 hasta 2018 la energía consumida parece ser mayor que la programada en promedio. Posteriormente, durante los años 2019 hasta 2022, parece que la volatilidad de la serie disminuye, con valores cercanos a cero y pocos valores atípicos. Finalmente, durante el periodo de 2022 hasta 2024, hay patrones que indican que la energía consumida es menor que la programada, debido a la mayor cantidad de picos negativos.

Para las frecuencias mensual, trimestral, semestral y anual vease el Anexo A.

Evolución mensual año a año

Para profundizar en el comportamiento estacional a nivel anual, se generaron gráficos de la serie mensual del volumen neto de desvíos desglosados por año. Esto permite comparar patrones de variación intraanuales y detectar posibles cambios en la estacionalidad a lo largo del periodo de estudio.

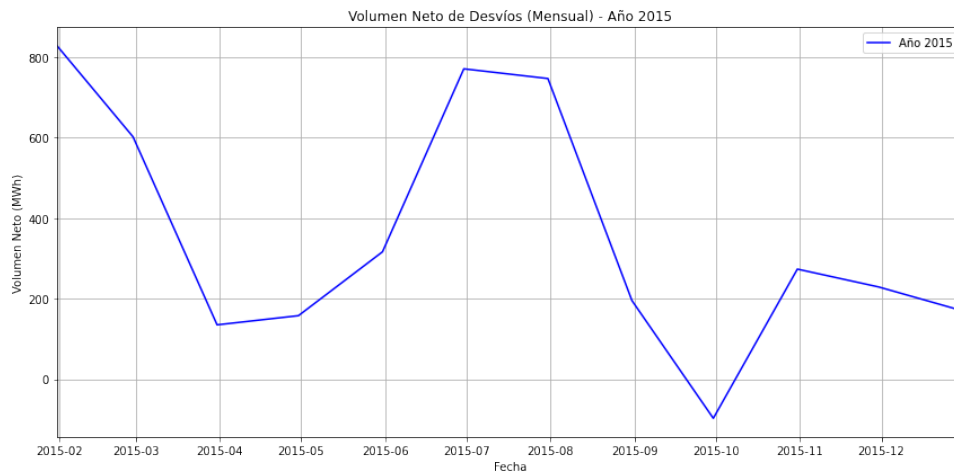


Figura 4.8. Serie mensual del volumen neto de desvíos (2015).
Fuente: Cálculos propios.

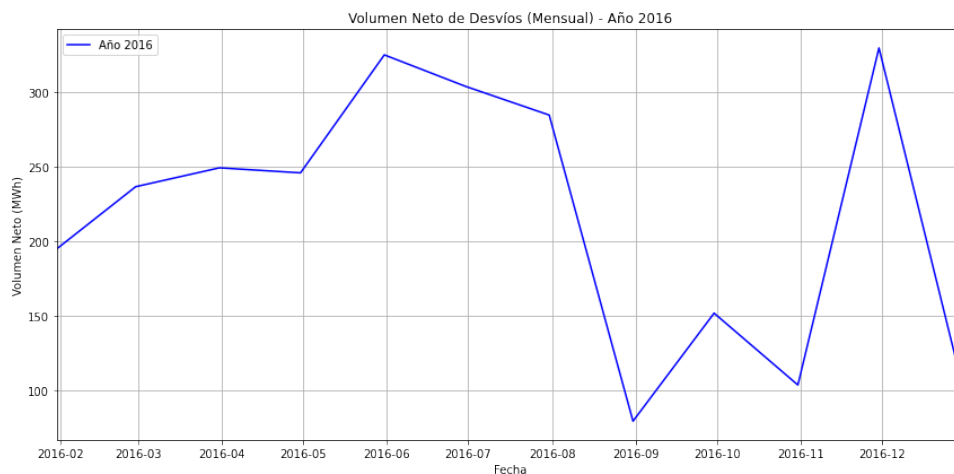


Figura 4.9. Serie mensual del volumen neto de desvíos (2016).
Fuente: Cálculos propios.

Se puede observar que, durante los meses de abril a septiembre en los años 2015 y 2016, el volumen neto de los desvíos presenta valores mensuales promedio superiores al del resto de los meses. Esto indica que existen indicios de estacionalidad en la serie, por lo que generar variables como la estación del año o el trimestre podría ayudar a los modelos de aprendizaje automático a capturar dichos patrones y proporcionar pronósticos más precisos.

Los gráficos correspondientes a los años 2017–2024 se incluyen en el Anexo B.

Patrones intrasemanales mediante mapas de calor

Para visualizar de manera compacta la variación media del volumen neto de desvíos a lo largo de la semana y el día, se emplean mapas de calor que muestran el promedio histórico

para cada combinación de hora y día de la semana. Esta representación facilita la detección de franjas horarias críticas y diferencias entre jornadas laborales y fines de semana.

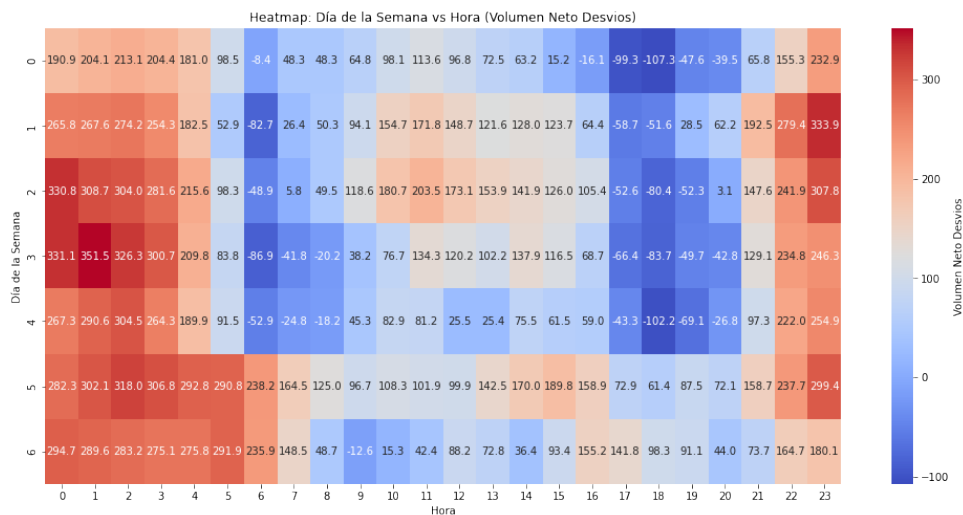


Figura 4.10. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día.

Fuente: Cálculos propios.

Se puede apreciar que existen patrones en las horas de los días de la semana, ya que el mapa de calor refleja una tonalidad roja en las horas comprendidas entre las 22:00 y las 4:00, de lunes a viernes. Por otra parte, durante los fines de semana, este intervalo se prolonga hasta las 7:00. Sin embargo, en las horas laborales o durante las horas del día, el desvío tiende a ser muy próximo a cero, en comparación con los intervalos antes mencionados, con la excepción del intervalo de 17:00 a 20:00, donde parece que, en promedio, la energía consumida fue menor que la programada.

Por lo tanto, se puede considerar la captura de estos patrones como variables para futuros modelos predictivos.

Los mapas de calor desglosados mes a mes se presentan en el Anexo C, proporcionando una visión más detallada de cómo estos patrones intrasemanales pueden variar a lo largo del año.

4.4.2. Análisis Multivariante

Relaciones entre variables mediante matriz de correlación

Para explorar la intensidad y dirección de la asociación lineal entre el volumen neto de desvíos y las variables exógenas (precio spot, previsión eólica, temperatura, etc.), se construyó una matriz de correlación. Este mapa de calor facilita la identificación de factores con mayor influencia y posibles multicolinealidades que deban considerarse en el modelado.

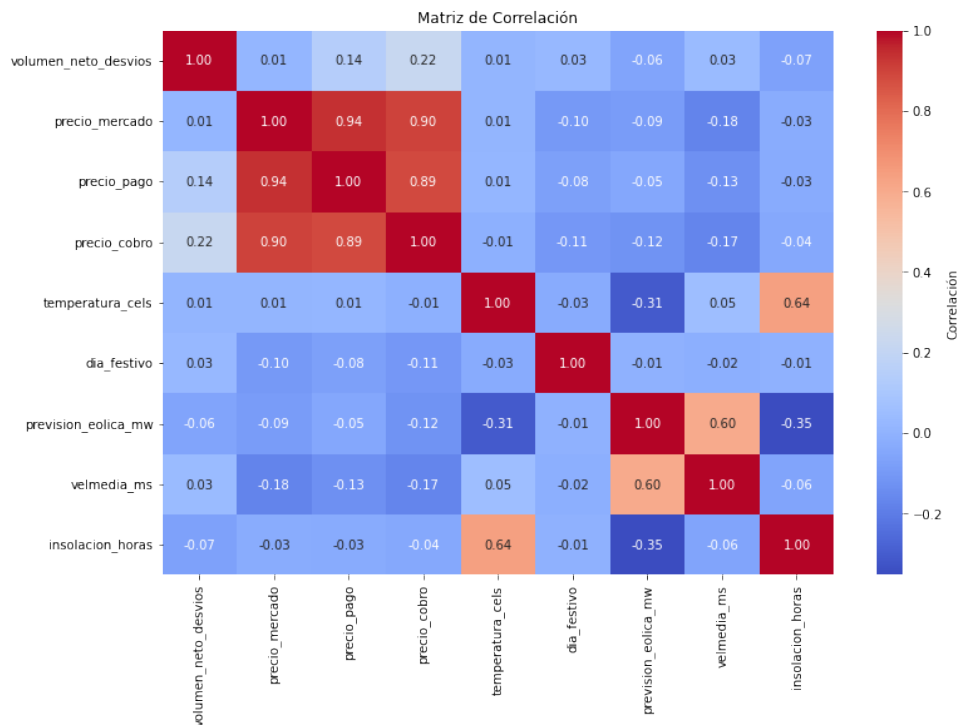


Figura 4.11. Matriz de correlación entre el volumen neto de desvíos y variables exógenas.
Fuente: Cálculos propios.

Se puede observar que existe una alta correlación entre las variables explicativas: precio del mercado, precio de pago y precio de cobro. Esto es esperable, dado que representan diferentes conceptos dentro del mismo mercado. Se evalúa eliminar el precio de cobro y pago en los futuros modelos, ya que el precio de mercado podría explicar suficientemente el comportamiento del conjunto. Además, el precio de cobro y pago se determina posteriormente al desvío; por lo tanto, en un modelo predictivo no sería posible emplearlo como variable explicativa, ya que se requieren variables que aporten información previa al suceso.

Por otra parte, existen otras correlaciones evidentes, como la de insolación horaria con la temperatura en grados Celsius, y la de previsión eólica con la velocidad media del viento. No obstante, dichas variables serán puestas a prueba en un modelo predictivo de aprendizaje automático, con el fin de evaluar cuáles resultan más significativas para el rendimiento del modelo.

Finalmente, se observa que el volumen neto de los desvíos no presenta una fuerte correlación con ninguna de las variables explicativas del proyecto. Sin embargo, las variables precio de cobro, insolación horaria y previsión eólica parecen ser las más importantes. No obstante, es necesario realizar más pruebas y evaluar su nivel de significación mediante técnicas más robustas.

Importancia de características con CatBoost

Con el objetivo de entender qué factores explicativos contribuyen en mayor medida a la predicción del volumen neto de desvíos, se entrenó un modelo `CatBoostRegressor` con los principales hiperparámetros optimizados y posteriormente se calculó la importancia de cada variable. Este análisis permite identificar qué rezagos, variables meteorológicas o de calendario resultan más determinantes y guiar refinamientos futuros en la selección de *features*.

Tabla 4.6. Hiperparámetros del modelo CatBoost

Parámetro	Valor
iterations	500
learning_rate	0.05
depth	6
loss_function	RMSE
eval_metric	RMSE
random_seed	42

Tabla 4.7. Importancia de características obtenida de CatBoost (%)

Feature	Importancia
volumen_netos_desvios_lag1	76.41
hora	9.32
anio	3.49
insolacion_horas	1.89
temperatura_cels	1.81
grupo_hora	1.45
precio_mercado	1.24
semana	1.12
prevision_eolica_mw	0.74
mes	0.71
velmedia_ms	0.39
dia_semana	0.29
tipo_clima	0.25
estacion	0.22
trimestre	0.22
festivo	0.18
dia	0.15
semestre	0.11

Análisis de resultados: La importancia de características revela que el rezago de un periodo horario (`volumen_netos_desvios_lag1`) domina con un 76 %, lo cual confirma que la depen-

dencia inmediata es la principal fuente de información para el modelo. A continuación, la hora y el año aportan casi un 13 % en conjunto, indicando patrones intradía y de largo plazo. Variables climáticas como `insolacion_horas` y `temperatura_cels` tienen un peso moderado (alrededor de 1.8–1.9 %), mientras que los indicadores de calendario (`grupo_hora`, `festivo`, `trimestre`, `semestre`) suman menos del 3 %, lo que sugiere que las variaciones estacionales y festivas tienen un efecto residual. Con base en estos resultados, podría valorarse descartar o agrupar variables de baja importancia en modelos posteriores.

Resultados del modelo CatBoost

Tabla 4.8. Mejores hiperparámetros encontrados por GridSearchCV

Parámetro	Valor
<code>bagging_temperature</code>	0
<code>depth</code>	4
<code>iterations</code>	300
<code>l2_leaf_reg</code>	3
<code>learning_rate</code>	0.05

Tabla 4.9. Desempeño en validación cruzada (5 folds)

Métrica	Valor
$RMSE_{CV}$	398.53 MWh

Tabla 4.10. Accuracy de signo en entrenamiento+validación

Métrica	Valor
Accuracy de signo	0.844

Análisis de resultados:

El procedimiento de validación cruzada (5 folds) arrojó un RMSE medio de 398.53 MWh con la configuración óptima de hiperparámetros. Además, el modelo acertó la dirección del desvío (positivo o negativo) en un 84.4 % de los casos sobre el conjunto de entrenamiento y validación. Estos resultados justifican el uso de los parámetros seleccionados y muestran un buen equilibrio entre error absoluto y habilidad para predecir la dirección del desequilibrio.

Resultados del modelo Prophet

Tabla 4.11. Mejores hiperparámetros de Prophet según $RMSE_{CV}$

Parámetro	Valor
changepoint_prior_scale	0.01
seasonality_prior_scale	10.0

Tabla 4.12. Desempeño de Prophet en validación cruzada (5 folds)

Métrica	Valor
$RMSE_{CV}$	408.66 MWh

Tabla 4.13. Accuracy de signo de Prophet en entrenamiento+validación

Métrica	Valor
Accuracy de signo	0.842

Análisis de resultados:

Con los parámetros óptimos, Prophet alcanzó un RMSE medio de 408.66 MWh en la validación cruzada, mostrando un desempeño ligeramente inferior al de CatBoost. La accuracy de signo del 84.2 % indica que Prophet también captura de forma razonable la dirección del desvío, aunque con mayor error en magnitud. Estos resultados refuerzan la necesidad de combinar ambos modelos en un enfoque híbrido para mejorar la precisión global.

Resultados del modelo híbrido

Tabla 4.14. Desempeño del modelo híbrido

Métrica	Valor
RMSE	402.24 MWh
Accuracy de signo	0.8437

Comparativo de modelos

Tabla 4.15. Comparación de RMSE y accuracy entre modelos

Modelo	RMSE (MWh)	Accuracy de signo
Dummy (lag 1h)	424.26	0.8410
Prophet	408.66	0.8420
CatBoost	398.53	0.8440
Híbrido	402.24	0.8437

Análisis comparativo:

El modelo *Dummy* basado en el valor de la hora anterior sirve como referencia mínima. Prophet reduce ligeramente el error frente a ese baseline, pero CatBoost ofrece la mayor reducción de RMSE (398.5 MWh) y la mejor precisión de signo. El modelo híbrido combina ambos enfoques: consigue un RMSE intermedio (402.2 MWh) y mantiene una accuracy de tendencia cercana a la de CatBoost (84.37 %), demostrando que la combinación ponderada aporta estabilidad en la predicción de dirección sin sacrificar demasiado la precisión en magnitud.

Resultados del modelo CatBoost multi-salida

Tabla 4.16. Hiperparámetros finales del modelo multi-salida

Parámetro	Valor
Profundidad (model__estimator__depth)	6
Iteraciones (model__estimator__iterations)	200
Regularización L2 (model__estimator__l2_leaf_reg)	3

Tabla 4.17. Desempeño continuo por horizonte (RMSE)

Horizonte	RMSE (MWh)
+1 hora	438.61
+2 horas	489.38
+3 horas	517.49
+4 horas	541.13
+5 horas	641.82
+6–+12 h	845.10–1183.08
+13–+24 h	771.27–1264.38
Promedio	952.04

Tabla 4.18. Precisión de signo por horizonte

Horizonte	Accuracy
+1 hora	0.78
+2 horas	0.73
+3→+5 h	0.77–0.64
+6→+12 h	0.62–0.64
+13→+24 h	0.58–0.69
Promedio	0.66

Tabla 4.19. Comparativo RMSE y accuracy de signo: multi-salida vs. modelos dummy

Modelo	RMSE (MWh)	Accuracy de signo
Dummy (24h lag)	917.77	0.628
Dummy (168h lag)	1086.20	0.556
CatBoost multi-salida (prom.)	952.04	0.660

Análisis de resultados:

El modelo multi-salida muestra un error creciente con el horizonte de pronóstico, partiendo de un RMSE de 438.6 MWh en la hora +1 hasta valores superiores a 1 200 MWh a partir de la hora +11, con un RMSE promedio de 952.0 MWh. La precisión de signo también decae levemente con el tiempo, situándose en un 78 % al primer horizonte y un 66 % global. Frente a los modelos dummy, el multi-salida ofrece un balance intermedio: mejora considerablemente al dummy semanal y mejora ligeramente al dummy diario en RMSE, mientras mantiene una precisión de dirección superior. Esto confirma que incorporar múltiples horizontes en una única regresión aporta beneficios frente a estrategias ingenuas de lag.

Simulación económica: impacto de la estrategia predictiva

Tabla 4.20. Costes de penalización de la comercializadora

Escenario	Coste total (EUR)
Sin estrategia (compra fija)	2 157 079.32
Con estrategia predictiva	2 052 189.84
Ahorro absoluto	104 889.48
Ahorro porcentual	4.86 %

Análisis de resultados:

La simulación muestra que, empleando la estrategia basada en las predicciones de volumen neto de desvíos, la comercializadora habría reducido sus costes de penalización de 2 157

079,32 EUR a 2 052 189,84 EUR. Esto equivale a un ahorro absoluto de 104 889,48 EUR y a una reducción del 4,86 % en el total de penalizaciones. Estos resultados ilustran el valor práctico de incorporar el modelo en la operativa diaria, contribuyendo a optimizar la posición de compra y a disminuir significativamente los gastos asociados al desequilibrio del sistema eléctrico.

Capítulo 5. Discusión

En esta sección reflexionamos de forma crítica sobre los principales resultados, las decisiones metodológicas más relevantes, las limitaciones encontradas y el impacto potencial de este trabajo.

5.1 Interpretación de los hallazgos

El modelo CatBoost simple obtuvo el mejor ajuste cuantitativo ($RMSE \approx 398$ MWh, accuracy de signo 84,4 %), mientras que Prophet ofreció un desempeño algo inferior ($RMSE \approx 409$ MWh, accuracy 84,2 %). El enfoque híbrido equilibró ambas metodologías, manteniendo una accuracy comparable (84,4 %) y un RMSE intermedio (402 MWh). El modelo multisalida, si bien permite pronosticar 24 horas de forma directa, muestra un aumento de error al crecer el horizonte ($RMSE$ promedio ≈ 952 MWh, accuracy global 66 %). La simulación económica ejemplifica el valor práctico: una reducción del 4,86 % en costes de penalización, equivalente a más de 100 000 EUR.

5.2 Decisiones metodológicas y adaptaciones

Durante el desarrollo se realizaron varias adaptaciones para garantizar robustez y eficiencia:

- Validación cruzada temporal estratificada para evitar filtrado de información futura.
- Ajuste manual de hiperparámetros en Prophet para controlar la flexibilidad de tendencia y estacionalidad.
- Reducción de dimensionalidad con PCA antes de CatBoost multisalida, para contener tiempos de entrenamiento.
- Normalización del ajuste en la simulación (división del desvío por 100) para cambios de compra coherentes.

5.3 Limitaciones

A pesar del éxito de los modelos, el estudio presenta limitaciones:

- Calidad y latencia de datos públicos (ESIOS, AEMET), con posibles sesgos o huecos.
- Pérdida de precisión en horizontes largos ($h > 6$ h) en el modelo multisalida.
- Simplificaciones en la simulación económica (distribución normal de desvíos reales).
- Carga computacional elevada en grid searches y modelos multisalida.

5.4 Impacto práctico

Los resultados demuestran que incorporar predicciones de desvíos en la estrategia de compra ofrece un ahorro económico tangible y mejora la estabilidad operativa. La metodología es escalable e integrable en herramientas de decisión en tiempo real, beneficiando a las comercializadoras y al sistema eléctrico.

Capítulo 6. Conclusiones

6.1 Conclusiones del trabajo

En este proyecto se ha desarrollado y comparado un conjunto de modelos de predicción del volumen neto de desvíos en el sistema eléctrico español, incluyendo Prophet, CatBoost y un enfoque híbrido, así como un modelo multisalida y dos estrategias dummy de referencia. Los resultados muestran que:

- CatBoost alcanza el RMSE más bajo (≈ 398 MWh) y una accuracy de signo del 84,4 %, superando a Prophet (RMSE ≈ 409 MWh, 84,2 %) y a las estrategias dummy.
- El modelo híbrido combina ambos enfoques, manteniendo una accuracy similar con un RMSE intermedio (≈ 402 MWh).
- El modelo multisalida permite pronosticar 24 horas de forma directa, aunque con un incremento de error en horizontes largos (RMSE promedio ≈ 952 MWh, accuracy global 66 %).
- La simulación económica evidencia un ahorro del 4,86 % en costes de penalización (más de 100 000 EUR) al aplicar la estrategia predictiva frente a una compra fija.

Estos hallazgos confirman que incorporar predicciones de desvíos en la planificación de compras contribuye a reducir significativamente los costes operativos y a mejorar la estabilidad del sistema eléctrico.

6.2 Conclusiones personales

Durante el desarrollo de este trabajo he profundizado en técnicas de series temporales, validación cruzada temporal y modelos de aprendizaje automático, lo que me ha permitido mejorar mi habilidad para manejar grandes volúmenes de datos y ajustar modelos complejos. Aprendí la importancia de validar cuidadosamente los supuestos estadísticos (estacionariedad, autocorrelación, normalidad) antes de modelar y de combinar enfoques para aprovechar las fortalezas de cada uno.

Este proyecto me ha enseñado a iterar sobre la metodología, adaptándola según los resultados intermedios y las limitaciones de los datos. Además, me ha dado una visión práctica de cómo la ciencia de datos puede generar ahorros reales y aportar valor en el sector energético. En el futuro, me gustaría explorar arquitecturas avanzadas (por ejemplo, redes neuronales para multisalida) y desarrollar herramientas en tiempo real para facilitar la adopción de estas estrategias por parte de las empresas.

Capítulo 7. Futuras líneas de trabajo

A partir de los resultados obtenidos y de las experiencias durante el desarrollo, se identifican las siguientes oportunidades para enriquecer y ampliar este proyecto:

- **Arquitectura de datos y despliegue automatizado:** Diseñar un flujo de ingesta continua que, mediante APIs y técnicas de web scraping, extraiga automáticamente los datos de ESIOS, AEMET y demás fuentes. Estos datos se almacenarían en una base de datos relacional o de series temporales, y un proceso programado (por ejemplo, con `cron` o Airflow) los transformaría y volcaría al modelo. De esta forma, el sistema generaría predicciones periódicas sin intervención manual, simulando un entorno de producción real en una compañía de energía.
- **Comparativa con modelos de deep learning:** Con recursos de cómputo más potentes (GPU), entrenar y evaluar arquitecturas basadas en redes neuronales recurrentes (LSTM, GRU) o modelos híbridos (por ejemplo, Transformer) para forecasting multisalida. Estos experimentos permitirían comparar su precisión y eficiencia frente a Prophet y CatBoost, especialmente en horizontes largos.
- **Evaluación de modelos en tiempo real:** Integrar un servicio web (REST API) que, a partir de consultas en tiempo real, devuelva las predicciones de desvíos y estimaciones de costes, facilitando su adopción por parte de las comercializadoras. Esto incluiría un panel de control con visualizaciones dinámicas y alertas automáticas ante posibles desequilibrios críticos.
- **Ampliación de variables exógenas:** Incorporar datos adicionales (por ejemplo, precios de mercados internacionales, generación distribuida, indicadores económicos o demanda sectorial) para explorar su impacto en la precisión del modelo y enriquecer el análisis de correlaciones.

Referencias bibliográficas

- Al Mamun, A., Sohel, M., Mohammad, N., Sunny, M. S. H., Dipta, D. R., & Hossain, E. (2017). *A Comprehensive Review of the Load Forecasting Techniques Using Single and Hybrid Predictive Models*. IEEE Access. Recuperado de <https://doi.org/10.1109/ACCESS.2017.DoiNumber>
- Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). *A note on the validity of using cross-validation for time series*. Computational Statistics & Data Analysis, 120, 70–83. Recuperado de <https://doi.org/10.1016/j.csda.2017.10.003>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley.
- Caro, E., & Juan, J. (2020). *Short-Term Load Forecasting for Spanish Insular Electric Systems*. Energies, 13(3645). Recuperado de <https://doi.org/10.3390/en13143645>
- Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. En Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794).
- Clemen, R. T. (1989). *Combining forecasts: A review and annotated bibliography*. International Journal of Forecasting, 5(4), 559–583. Recuperado de [https://doi.org/10.1016/0169-2070\(89\)90032-7](https://doi.org/10.1016/0169-2070(89)90032-7)
- Comisión Europea. (2023). *Plan de implementación España*. Recuperado de https://energy.ec.europa.eu/system/files/2023-11/01_PlandeImplementacion_ES_0.pdf
- Comisión Nacional de los Mercados y la Competencia. (2019). *Informe de supervisión del mercado peninsular 2023*. Recuperado de https://www.cnmc.es/sites/default/files/3291860_0.pdf
- Comisión Nacional de los Mercados y la Competencia. (2023). *Informe de supervisión del mercado peninsular 2023*. Recuperado de <https://www.cnmc.es/sites/default/files/5779264.pdf>
- Conde, F. (2016). *Automatización y Optimización de las Actividades de una Comercializadora* (TFM). Universidad de Sevilla. Recuperado de <https://biblus.us.es/bibing/proyectos/abreproy/71130/fichero/TFM-1130-CONDE.pdf>
- Dickey, D. A., & Fuller, W. A. (1979). *Distribution of the estimators for autoregressive time series with a unit root*. Journal of the American Statistical Association, 74(366a),

427–431. Recuperado de <https://doi.org/10.1080/01621459.1979.10482531>

- Endesa. (s.f.). *El mercado eléctrico*. Fundación Endesa. Recuperado de <https://www.fundacionendesa.org/es/educacion/endesa-educa/recursos/el-mercado-electrico>
- Filgueira Fernández, G. (2024). *Predicción de Demanda de Gas y Electricidad Utilizando Técnicas Estadísticas Avanzadas e Inteligencia Artificial* (TFG). Universidad Politécnica de Madrid. Recuperado de https://oa.upm.es/82673/1/TFG_GUILLERMO_FILGUEIRA_FERNANDEZ.pdf
- Friedman, J. H. (2001). *Greedy function approximation: a gradient boosting machine*. *Annals of Statistics*, 29(5), 1189–1232.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts. Recuperado de <https://OTexts.com/fpp2>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts.
- López García, M. (2016). *Modelos de predicción de demanda eléctrica utilizando técnicas de inteligencia artificial. Aplicación al mercado eléctrico español* (Tesis doctoral). Universidad Miguel Hernández de Elche. Recuperado de <https://dialnet.unirioja.es/servlet/tesis?codigo=64987>
- Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions*. En *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 4768–4777).
- MagnusCMD. (s.f.). *Desvíos en el sistema eléctrico: La visión del consumidor*. Recuperado de https://magnuscmd.com/es/desvios_en_el_sistema_electrico_la_vision_del_consumidor/
- Meteo For Energy. (2022). *Análisis sobre los costes de desvío de las energías renovables (Junio 2022)*. Recuperado de <https://www.meteoforenergy.com/meteonew/galeria/2022/Meteoblog/Costes%20de%20desv%20a%20de%20las%20energ%20as%20renovables%20-%20Junio%202022.pdf>
- Miele, E. S., Ludwig, N., & Corsini, A. (2023). *Multi-Horizon Wind Power Forecasting Using Multi-Modal Spatio-Temporal Neural Networks*. *Energies*, 16(8), 3522. Recuperado de <https://doi.org/10.3390/en16083522>

- OMIE (Operador del Mercado Ibérico de Energía). (s.f.). *Mercado intradiario continuo*. Recuperado de https://www.omie.es/es/mercados_y_productos/mercado_intradiario_continuo
- Pearson, K. (1895). *Note on regression and inheritance in the case of two parents*. Proceedings of the Royal Society of London, 58, 240–242.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). *CatBoost: unbiased boosting with categorical features*. arXiv preprint. Recuperado de <https://arxiv.org/abs/1706.09516>
- Red Eléctrica de España. (2023). *Informe del sistema eléctrico español 2023*. Recuperado de https://www.ree.es/sites/default/files/07_SALA_PRENSA/Documentos/2024/NP_CyL.pdf
- Red Eléctrica de España. (2024). *Resumen Servicios de ajuste*. Recuperado de <https://www.sistemaelectrico\%ree.es/informe\%del\%sistema\%electrico/mercados/servicios\%ajuste/resumen\%servicios\%ajuste>
- Red Eléctrica de España. (2025, 18 marzo). *Informe del sistema eléctrico español 2024*. Recuperado de https://www.ree.es/sites/default/files/07_SALA_PRENSA/Documentos/2025/NP_Madrid.pdf
- Red Eléctrica de España. (s.f.). *El sistema eléctrico español*. Recuperado de <https://www.ree.es/es/operacion/sistema-electrico>
- Sanz Muñoz, J. (2023). *Predicción de precios de la electricidad con neural ODE* (Trabajo Fin de Máster). Universidad Pontificia Comillas. Recuperado de <https://repositorio.comillas.edu/jspui/bitstream/11531/75263/1/TFM\%20-%20Sanz\%20Munoz\%20Jaime.pdf>
- Shapiro, S. S., & Wilk, M. B. (1965). *An analysis of variance test for normality (complete samples)*. Biometrika, 52(3/4), 591–611. Recuperado de <https://doi.org/10.1093/biomet/52.3-4.591>
- Taylor, S. J., & Letham, B. (2018). *Forecasting at scale*. The American Statistician, 72(1), 37–45. Recuperado de <https://doi.org/10.1080/00031305.2017.1380080>
- Trull Domínguez, J. (s.f.). *Predicción a corto plazo de la demanda horaria de energía eléctrica en España mediante modelos optimizados de Holt–Winters múltiple* (TFM). Universidad de Sevilla. Recuperado de https://biblus.us.es/bibing/proyectos/abreproy/72713/descargar_fichero/TFM-2713\%2BJÃ¼c
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.

Capítulo 8. ANEXOS

Anexo A

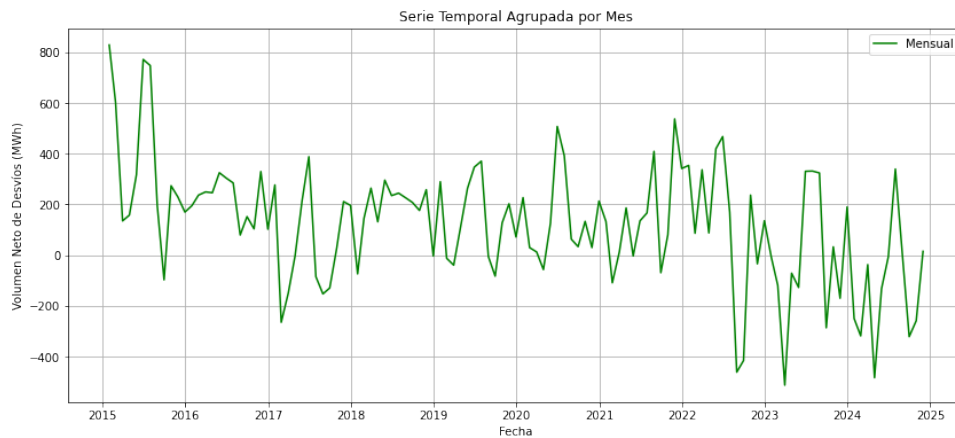


Figura 8.1. Serie mensual del volumen neto de desvíos.
Fuente: Cálculos propios.

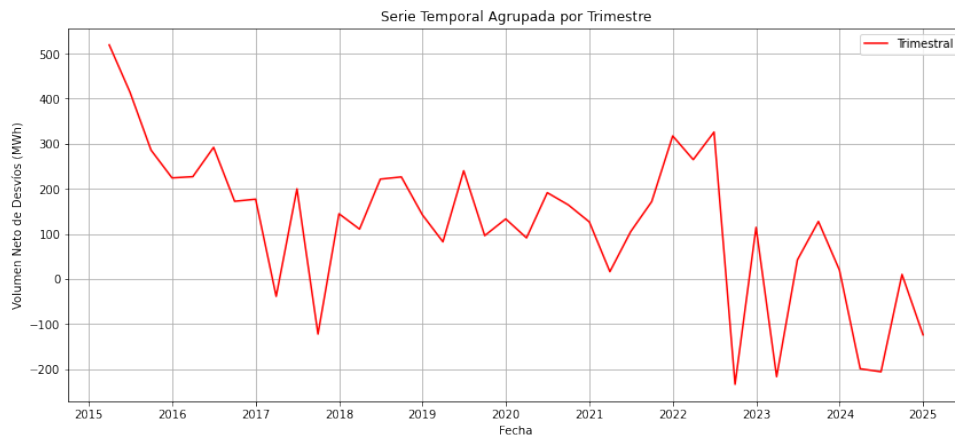


Figura 8.2. Serie trimestral del volumen neto de desvíos.
Fuente: Cálculos propios.

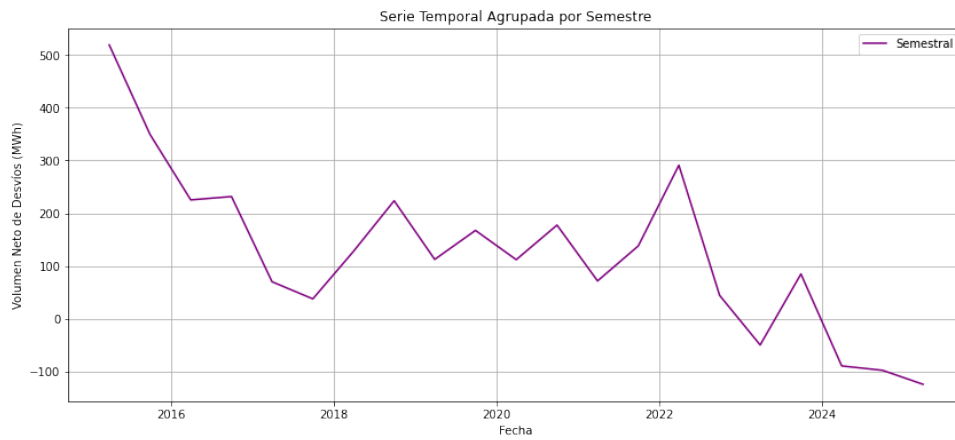


Figura 8.3. Serie semestral del volumen neto de desvíos.
Fuente: Cálculos propios.

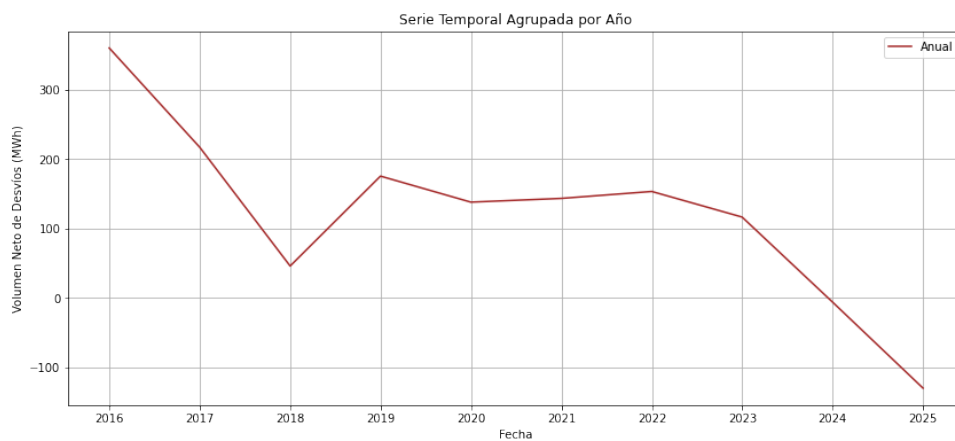


Figura 8.4. Serie anual del volumen neto de desvíos.
Fuente: Cálculos propios.

Anexo B

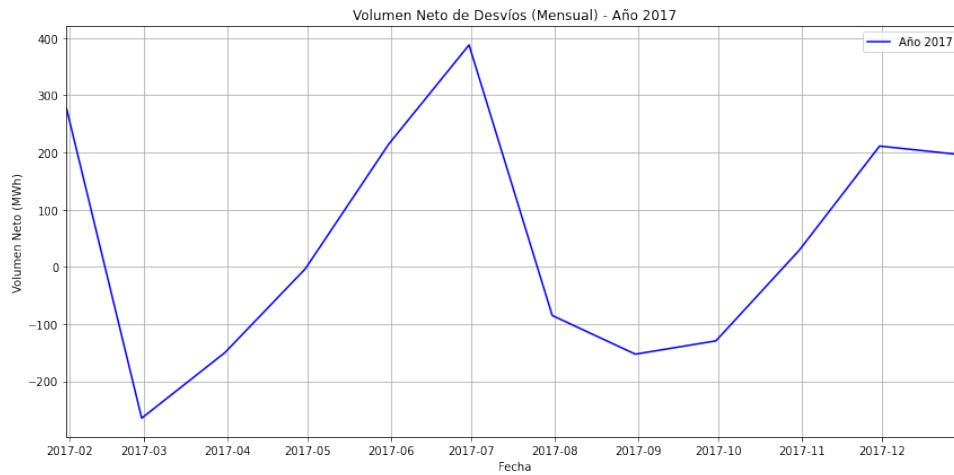


Figura 8.5. Serie mensual del volumen neto de desvíos (2017).
Fuente: Cálculos propios.

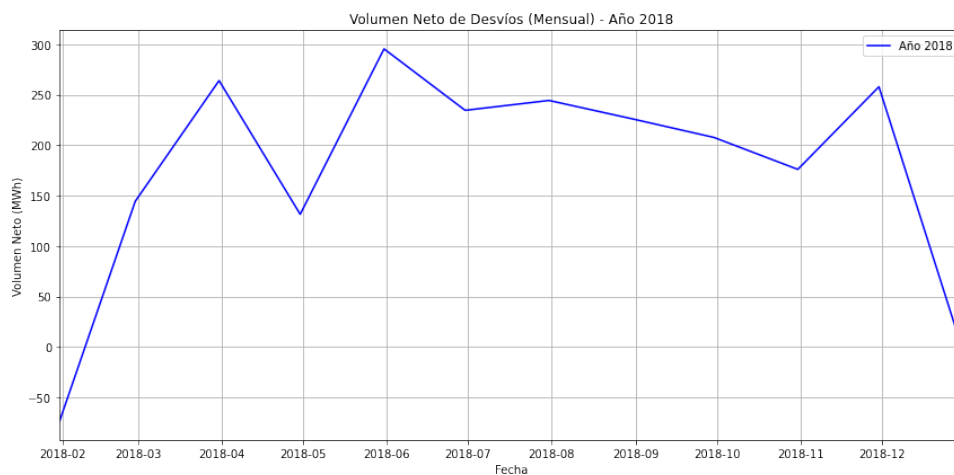


Figura 8.6. Serie mensual del volumen neto de desvíos (2018).
Fuente: Cálculos propios.

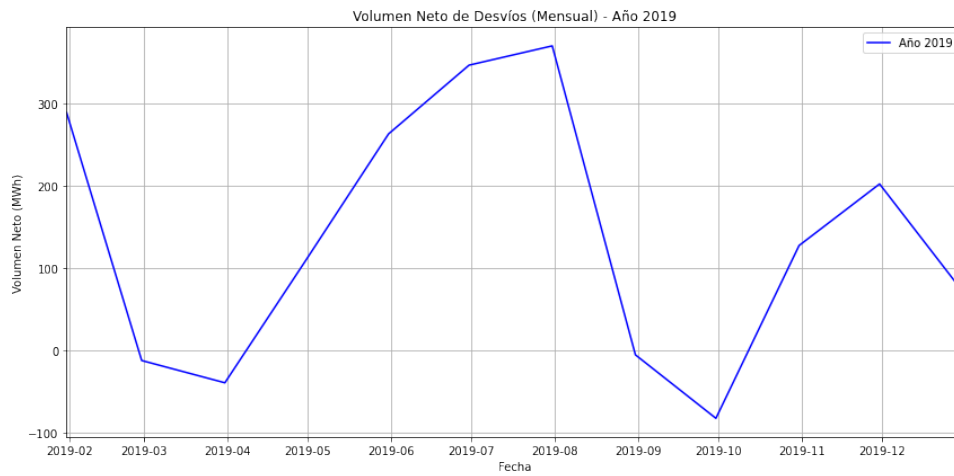


Figura 8.7. Serie mensual del volumen neto de desvíos (2019).
Fuente: Cálculos propios.

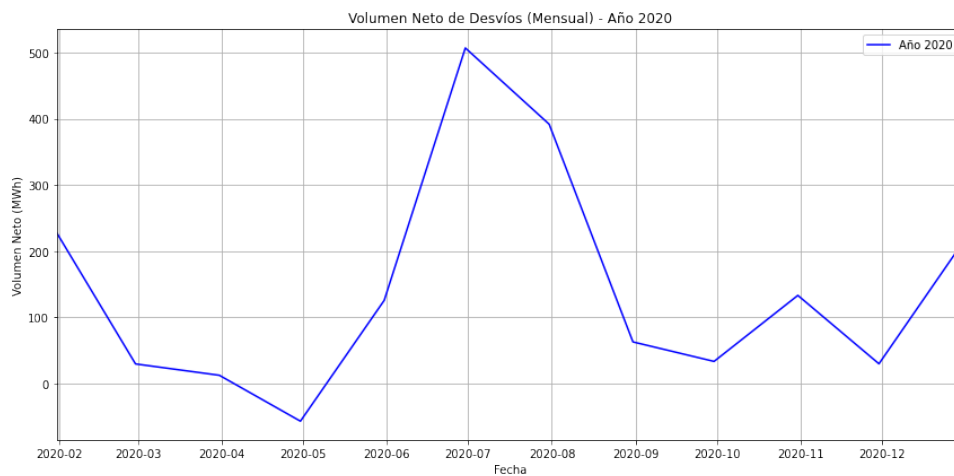


Figura 8.8. Serie mensual del volumen neto de desvíos (2020).
Fuente: Cálculos propios.

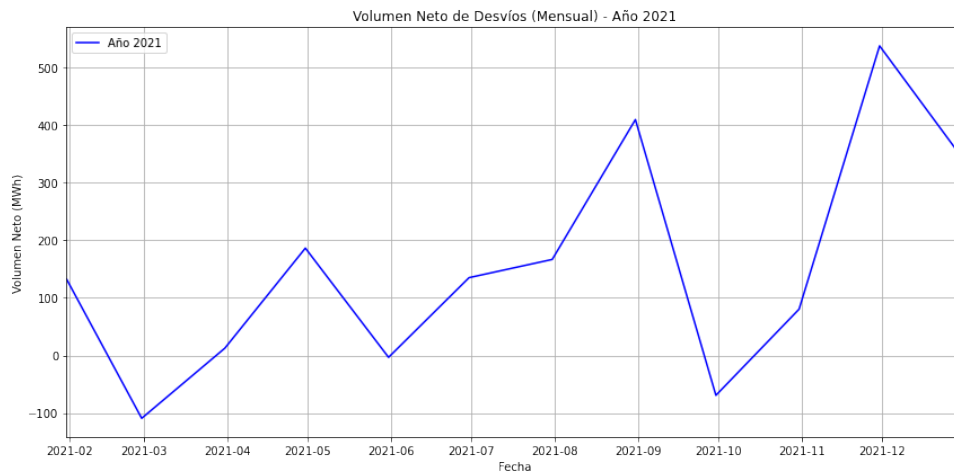


Figura 8.9. Serie mensual del volumen neto de desvíos (2021).
Fuente: Cálculos propios.

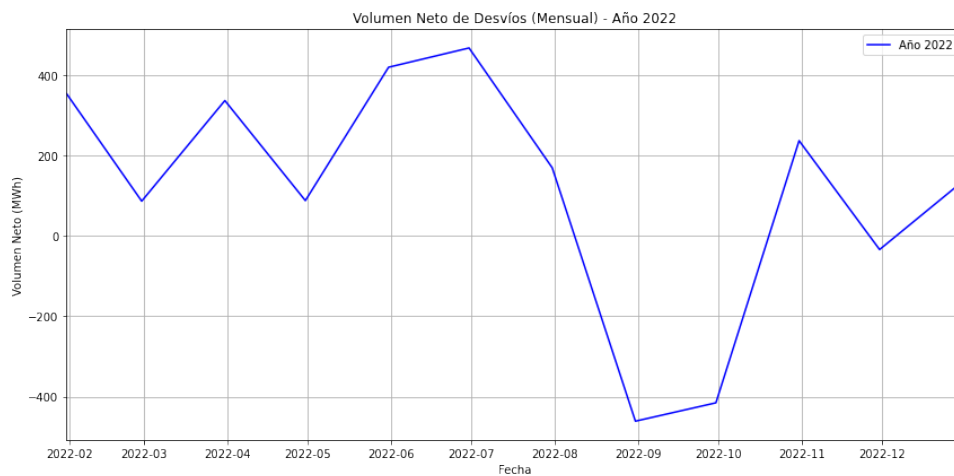


Figura 8.10. Serie mensual del volumen neto de desvíos (2022).
Fuente: Cálculos propios.

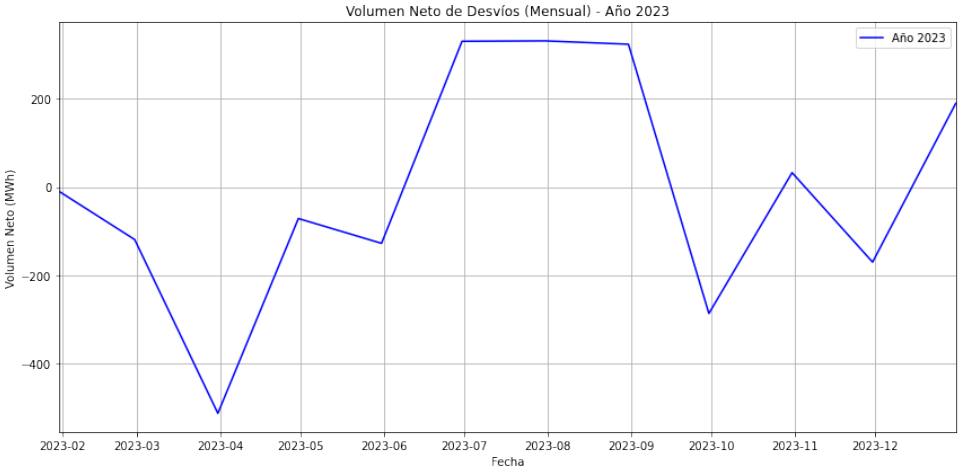


Figura 8.11. Serie mensual del volumen neto de desvíos (2023).
Fuente: Cálculos propios.

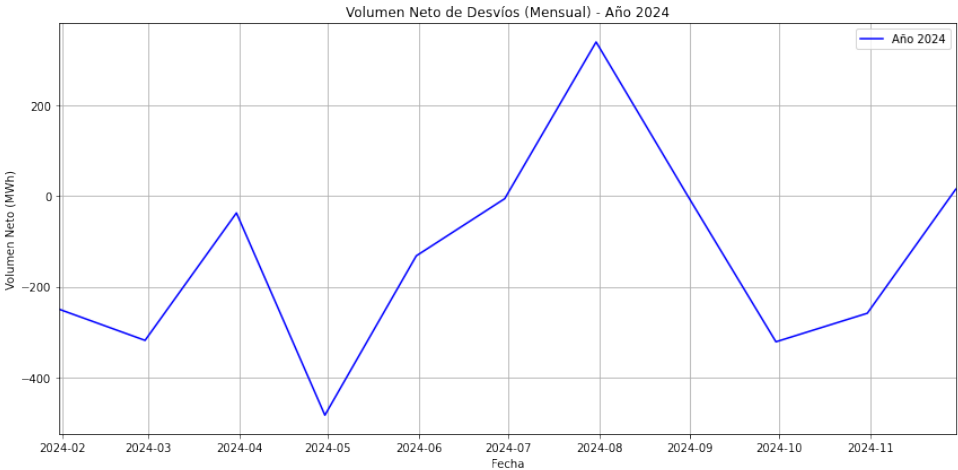


Figura 8.12. Serie mensual del volumen neto de desvíos (2024).
Fuente: Cálculos propios.

Anexo C

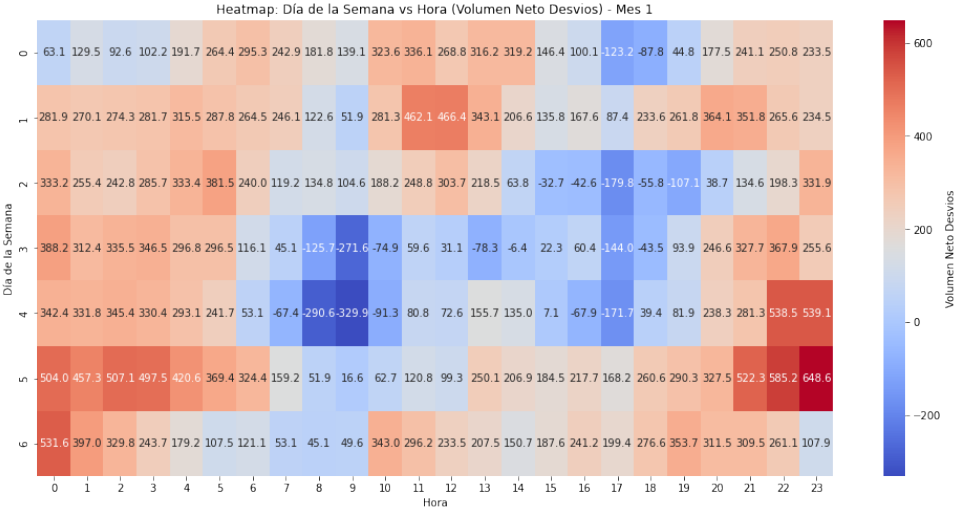


Figura 8.13. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Enero).
Fuente: Cálculos propios.

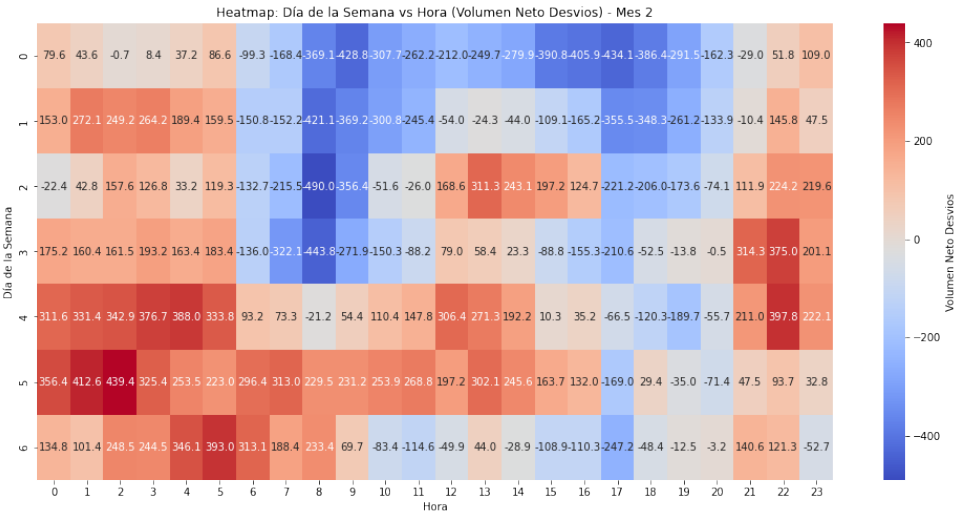


Figura 8.14. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Febrero).
Fuente: Cálculos propios.

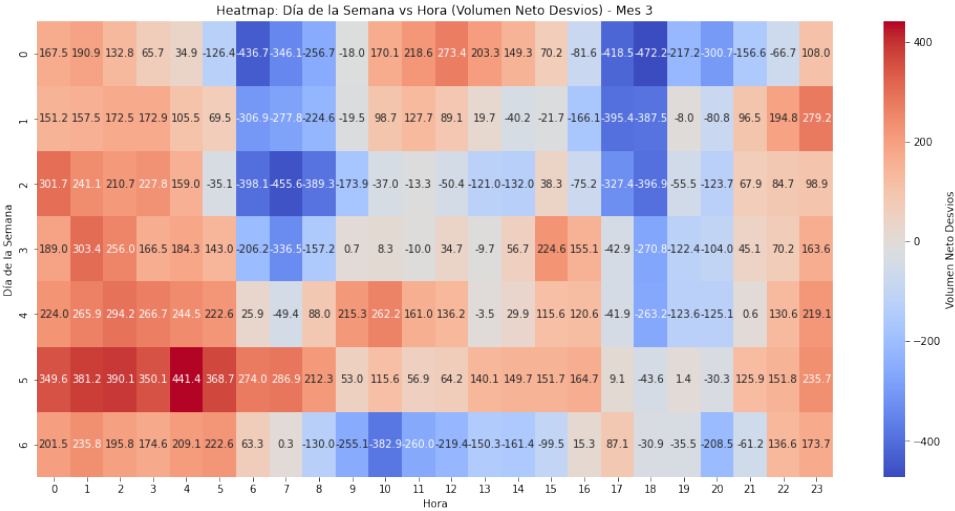


Figura 8.15. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Marzo).
Fuente: Cálculos propios.

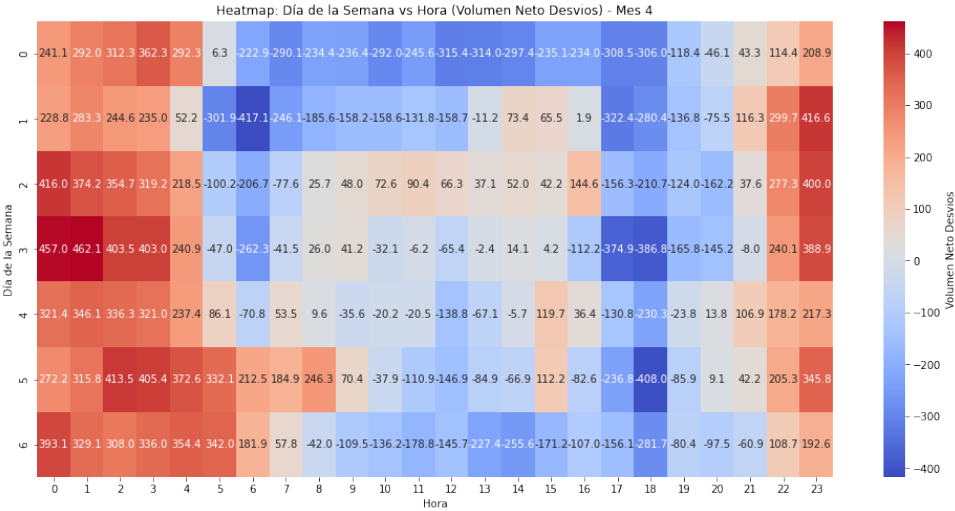


Figura 8.16. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Abril).
Fuente: Cálculos propios.

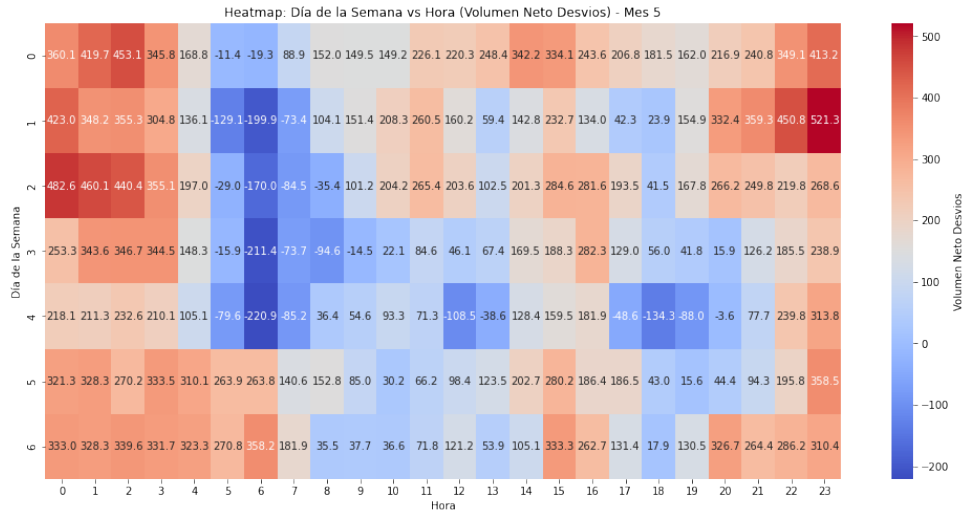


Figura 8.17. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Mayo).
Fuente: Cálculos propios.

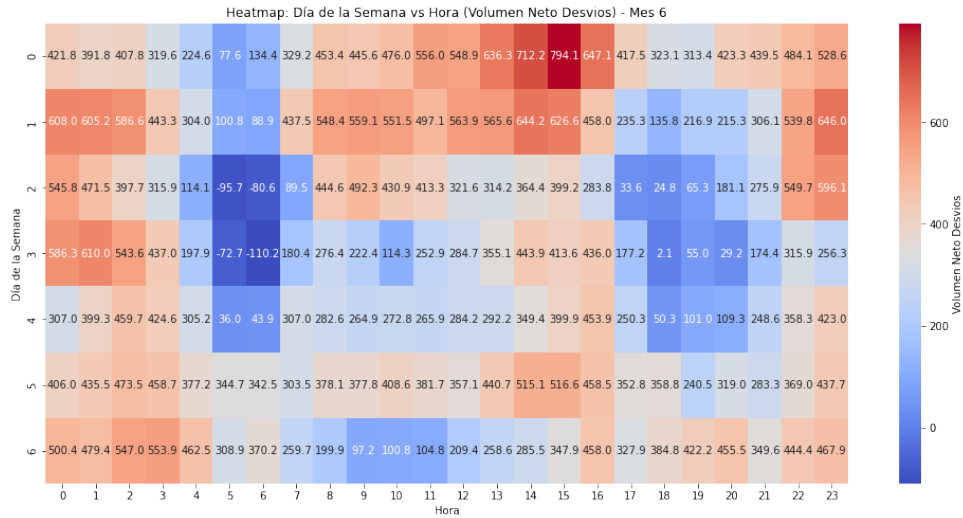


Figura 8.18. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Junio).
Fuente: Cálculos propios.

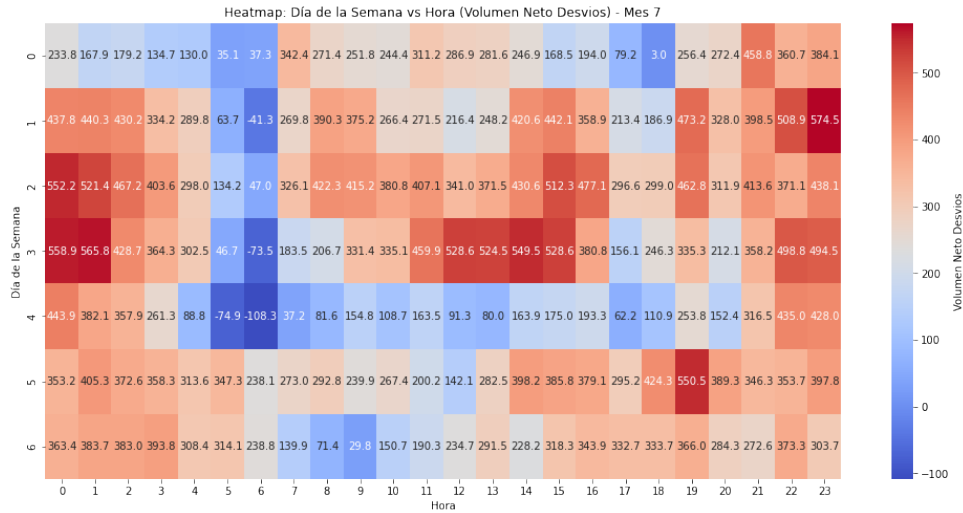


Figura 8.19. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Julio).
Fuente: Cálculos propios.

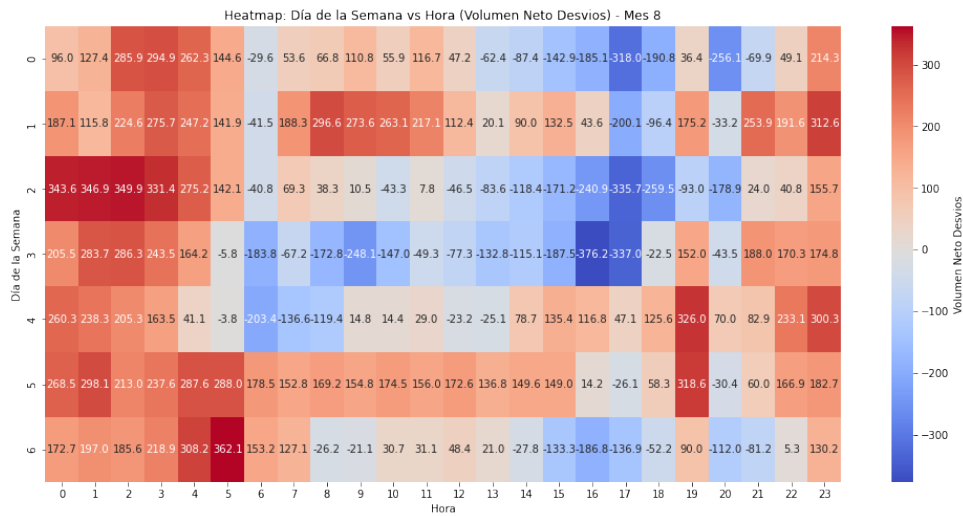


Figura 8.20. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Agosto).
Fuente: Cálculos propios.

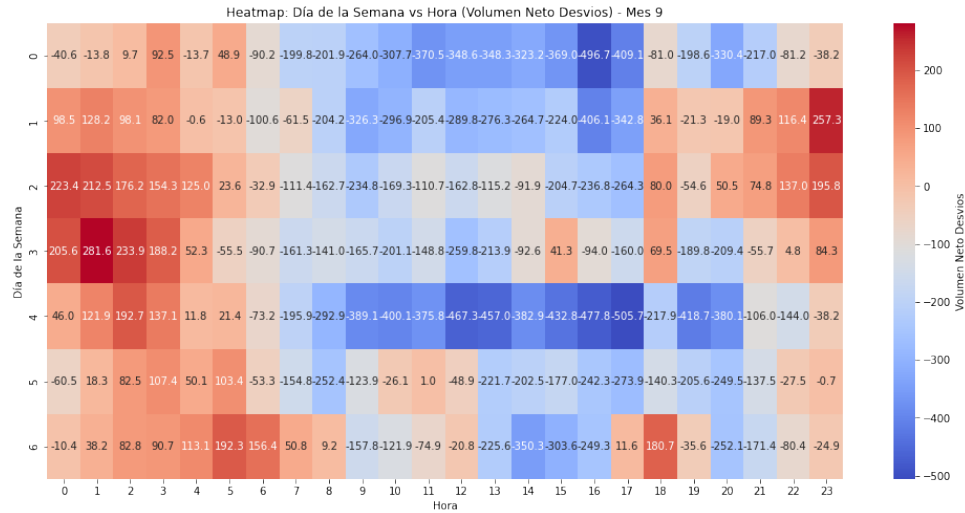


Figura 8.21. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Septiembre).
Fuente: Cálculos propios.

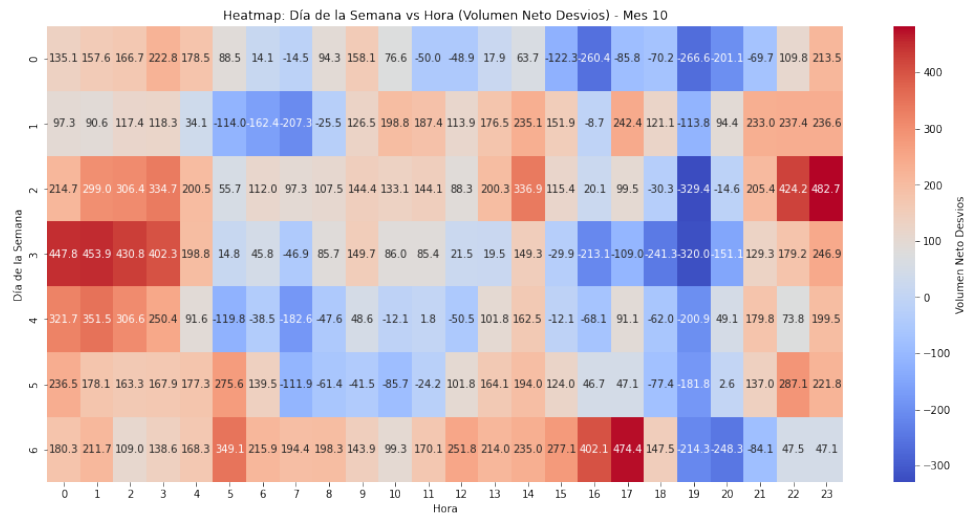


Figura 8.22. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Octubre).
Fuente: Cálculos propios.

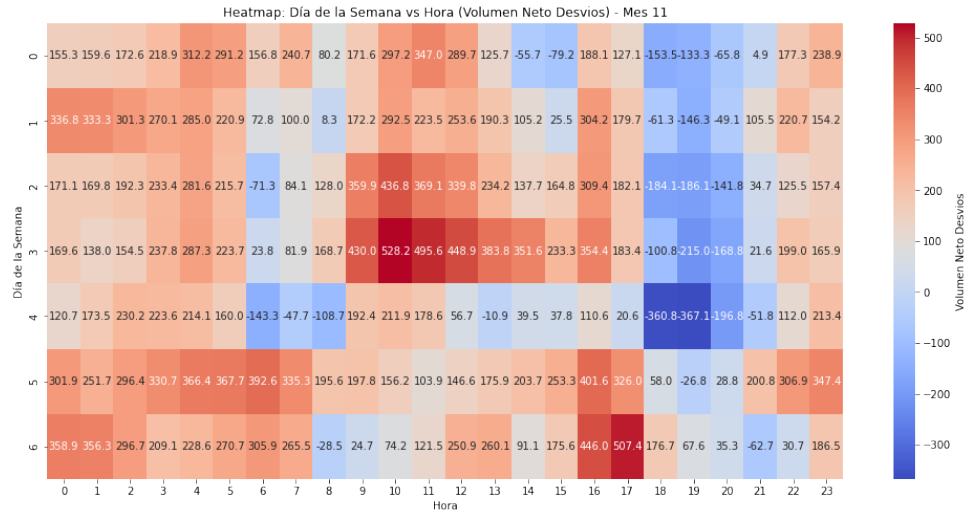


Figura 8.23. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Noviembre).
Fuente: Cálculos propios.

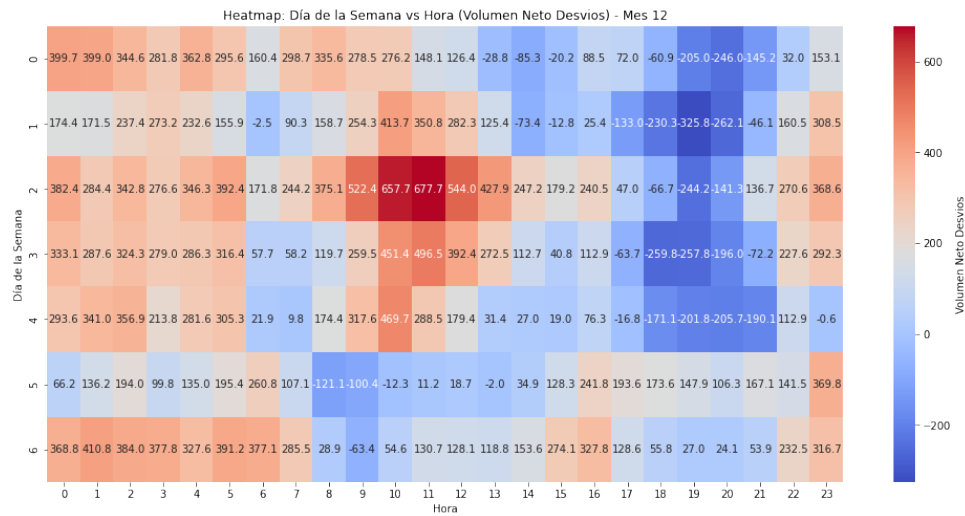


Figura 8.24. Mapa de calor del volumen neto de desvíos promedio por día de la semana y hora del día (Diciembre).
Fuente: Cálculos propios.