

RESUMEN

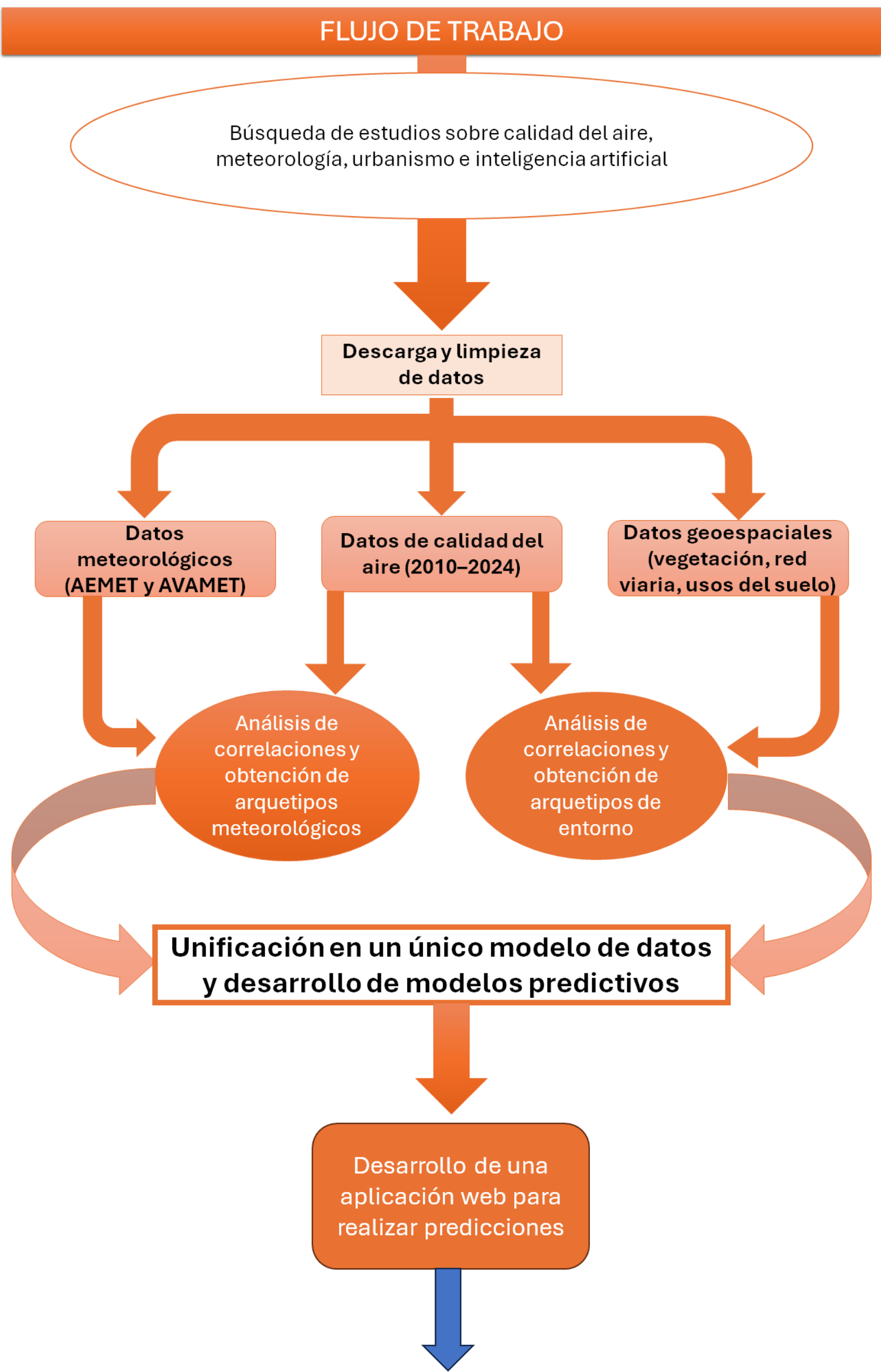
Este proyecto analiza la influencia de las condiciones meteorológicas y la ubicación de las estaciones de medición en la calidad del aire en la ciudad de Valencia. La contaminación atmosférica es un problema clave en entornos urbanos, con efectos en la salud pública y el medio ambiente. Comprender cómo la meteorología y la localización de las estaciones afectan la medición de la contaminación es fundamental para mejorar las estrategias de gestión ambiental.

Para ello, se recopilan y analizan datos de calidad del aire y variables climáticas desde 2010 hasta 2024, con el objetivo de identificar patrones que expliquen la variabilidad de los niveles de contaminación. Además, se desarrolla un modelo predictivo basado en Machine Learning que permita anticipar episodios de contaminación en función de las condiciones meteorológicas y el entorno de cada estación.

La primera parte del estudio se centra en la caracterización de las estaciones mediante la creación de arquetipos meteorológicos y el análisis de su entorno urbano. En la segunda parte, se aplican técnicas de modelado para evaluar la relación entre los datos de calidad del aire y las variables identificadas. Los resultados de este proyecto proporcionarán herramientas para optimizar la planificación urbana y mejorar la gestión de la contaminación en Valencia.



METODOLOGIA



CONCLUSIONES

Este trabajo demuestra que la calidad del aire urbano está condicionada por una interacción compleja entre factores meteorológicos y características del entorno urbano inmediato. Se identificaron correlaciones significativas, como la relación negativa entre temperatura mínima y NOx ($r \approx -0.47$) y la positiva entre viento máximo y ozono ($r \approx 0.53$). Además, variables urbanas como el índice de vegetación (NDVI) o la densidad edificatoria mostraron un impacto claro sobre la concentración de contaminantes. A partir del análisis se definieron tres arquetipos meteorológicos (estival, transicional, invernal) y tres urbanos (urbano denso, mixto verde-residencial, y periurbano natural), que explican la variabilidad espacial y estacional de la contaminación. Los modelos predictivos basados en Random Forest alcanzaron buena precisión ($R^2 = 0.58$ para O_3), y han sido integrados en una herramienta web funcional para predicción en tiempo real. Estos resultados refuerzan la necesidad de enfoques integrados y una mejora en la calidad y estandarización de los datos ambientales para apoyar decisiones urbanas y de salud pública.

Referencias:

Zhang, X., Wang, Y., Hao, J., & Wu, Y. (2022). Forecasting urban air quality using deep learning and meteorological data. Environmental Pollution, 295, 118654. <https://doi.org/10.1016/j.envpol.2021.118654>

Instituto ITACA. (n.d.). PREDATICS – Predicción de Datos de Contaminación mediante Inteligencia Artificial en Ciudades Sostenibles. Instituto ITACA, Universidad Politécnica de Valencia. Disponible en: <https://www.itaca.upv.es/predatics/>

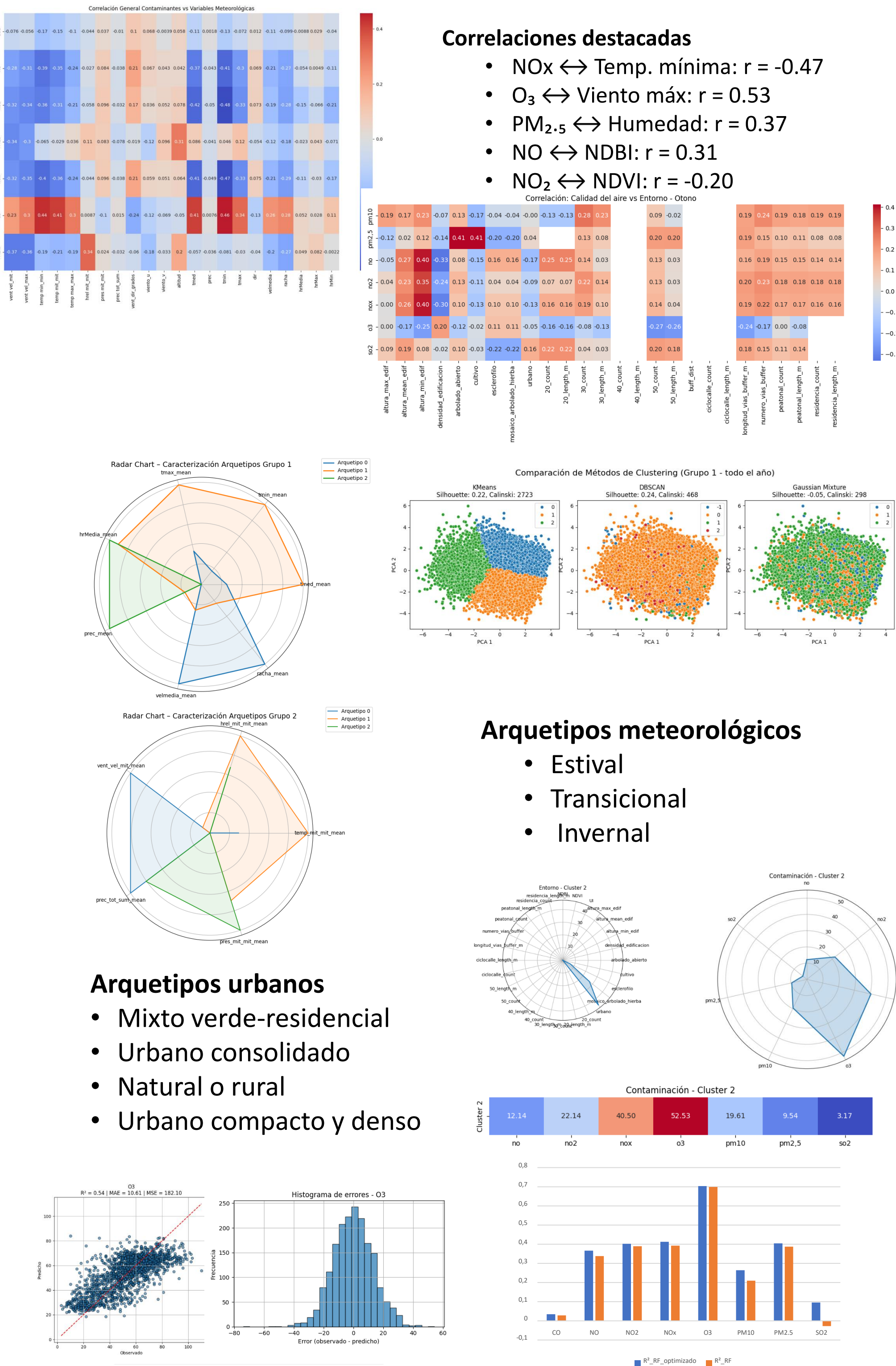
World Health Organization (WHO). (2021). WHO Global Air Quality Guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. Disponible en: <https://www.who.int/publications/i/item/9789240034228>

Dai, Z., Liu, D., Yu, K., Cao, L., & Jiang, Y. (2020). Meteorological variables and synoptic patterns associated with air pollution in Eastern China during 2013–2018. International Journal of Environmental Research and Public Health, 17(7), 2528. <https://doi.org/10.3390/ijerph17072528>

OBJETIVOS

- El objetivo principal de este Trabajo Fin de Máster es analizar la influencia de las condiciones meteorológicas y de la ubicación física de las estaciones de medición sobre la calidad del aire en la ciudad de Valencia, con el fin de contribuir a la mejora de los modelos predictivos desarrollados en el marco del proyecto PREDATICS. Para alcanzar este objetivo general, se plantean los siguientes objetivos específicos:
- Obtención y manejo de un gran volumen de datos de diferentes fuentes tanto de las estaciones de calidad del aire y datos geoespaciales que ayuden a definir su entorno, como datos de las estaciones meteorológicas.
 - Búsqueda de relaciones entre variables meteorológicas y concentraciones de contaminantes, identificando patrones recurrentes y anomalías relevantes.
 - Definición de arquetipos meteorológicos, es decir, situaciones atmosféricas típicas (por ejemplo, días con estabilidad atmosférica, alta presión, lluvias, viento de levante, etc.) y estudiar cómo se comportan los niveles de contaminantes en cada uno de ellos.
 - Caracterización del entorno de cada estación de calidad del aire utilizando datos geoespaciales de tráfico, infraestructuras y vegetación. Y encontrar relaciones entre las variables del entorno y los datos de calidad del aire.
 - Definición de arquetipos del entorno próximo a las estaciones.
 - Desarrollo de modelos predictivos de calidad del aire que integren variables meteorológicas y del entorno urbano, para estimar concentraciones de contaminantes atmosféricos.
 - Desarrollo de una herramienta interactiva para la predicción de contaminantes atmosféricos mediante modelos de aprendizaje automático, con el modelo con mejores resultados, permitir al usuario meter parámetros y predecir el valor del contaminante con su porcentaje de fiabilidad.

RESULTADOS



- Modelos predictivos**
- R^2 O_3 : 0.58
 - R^2 NOx: 0.41
 - Mejor modelo: Random Forest (con hiperparámetros)

Predicción para O3: 77.78 µg/m³

Predicción para SO2: 3.21 µg/m³

Fiabilidad del modelo (R^2): 0.619

Fiabilidad del modelo (R^2): 0.064