



Escola Tècnica Superior d'Enginyeria Informàtica

Universitat Politècnica de València

Análisis exploratorio de la calidad del aire

Proyecto de la asignatura

Proyecto I, Comprensión de datos

Grado en Ciencia de Datos

Autores:

Sheila Díaz Giménez
Alejandro Hervás Castillo
Laiqian Ji
Carles Navarro Esteve
Álvaro Prado Expósito

Tutor:

Francisco Pedroche Sánchez

Curso 2022 - 2023

Resumen

El objetivo principal de este estudio es analizar la calidad del aire en 20 ciudades españolas durante el periodo comprendido entre 2009 y 2020 utilizando técnicas de análisis estadístico. Los datos han sido obtenidos de un histórico presente en la web del Ministerio para la Transición Ecológica y el Reto Demográfico, recogidos en la normativa vigente (Directiva Europea 2008/50/CE, relativa a la calidad del aire y una atmósfera más limpia). Una vez recogidos los datos, se va a utilizar el Índice de Calidad del Aire (ICA) para evaluar el estudio en las diferentes ciudades y la evolución temporal de los contaminantes. Esta temática es de gran importancia en la actualidad debido a que la exposición de contaminantes del aire puede tener efectos negativos en la salud humana y en el medio ambiente. Por lo tanto, un análisis detallado de la calidad del aire puede proporcionar información valiosa para los responsables de la toma de decisiones en la implementación de medidas que reduzcan la contaminación en nuestras ciudades. Además, este tema está directamente ligado con los Objetivos de Desarrollo Sostenible (ODS) establecidos por la ONU, en particular el ODS 3 de “Salud y bienestar”, ODS 13 “Acción por el clima” y el ODS 11 de “Ciudades y comunidades sostenibles”, lo que reafirma aún más la importancia de su elección para el proyecto.

La evaluación de la calidad del aire se realiza para los contaminantes, con una mayor presencia en la atmósfera: dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂), óxido de nitrógeno (NOx), partículas en suspensión (PM10 y PM2,5), benceno (C₆H₆), monóxido de carbono (CO) y ozono (O₃). Cinco de ellos son utilizados para el cálculo del Índice Nacional de Calidad del Aire.

PALABRAS CLAVE:

Índice de calidad del aire, medio ambiente, valores límite, contaminación

Resum

L'objectiu principal d'aquest estudi és analitzar la qualitat de l'aire en 20 ciutats espanyoles durant el període comprès entre 2009 i 2020 utilitzant tècniques d'anàlisi estadística. Les dades han estat obtingudes d'un històric present en el web del Ministeri per a la Transició Ecològica i el Repte Demogràfic (MITECO), recollits en la normativa vigent (Directiva Europea 2008/50/CE, relativa a la qualitat de l'aire i una atmosfera més neta). Una vegada recollides les dades, s'utilitzarà l'Índex de Qualitat de l'Aire (ICA) per avaluar l'estudi en les diferents ciutats i l'evolució temporal dels contaminants. Aquesta temàtica és de gran importància en l'actualitat a causa que l'exposició de contaminants de l'aire pot tenir efectes negatius en la salut humana i en el medi ambient. Per tant, una anàlisi detallada de la qualitat de l'aire pot proporcionar informació valiosa per als responsables de la presa de decisions en la implementació de mesures que redueixin la contaminació en les nostres ciutats. A més, aquest tema està directament lligat amb els Objectius de Desenvolupament Sostenible (ODS) establerts per l'ONU, en particular l'ODS 3 de "Salut i benestar", l'ODS 13 "Acció pel clima" i l'ODS 11 de "Ciutats i comunitats sostenibles", el que reafirma encara més la importància de la seva elecció per al projecte.

L'avaluació de la qualitat de l'aire es realitza pels contaminants, que han demostrat una major presència en l'atmosfera: diòxid de sofre (SO₂), diòxid de nitrogen (NO₂), òxid de nitrogen (NO_x), partícules en suspensió (PM10 i PM2,5), benzè (C₆H₆), monòxid de carboni (CO) i ozó (O₃). Cinc d'ells són utilitzats per al càlcul de l'Índex Nacional de Qualitat de l'Aire.

PARAULES CLAU:

Índex de qualitat de l'aire, medi ambient, valors límit, contaminació

Abstract

The main objective of this study is to analyze the air quality in 20 Spanish cities from 2009 to 2020 using statistical analysis techniques. The data have been obtained from the Ministry for the Ecological Transition and Demographic Challenge (MITECO) website, collected according to the current normative (European Directive 2008/50/CE, on air quality and cleaner atmosphere). Once the data is collected, Air Quality Index (AQI) will be used to evaluate the study in different cities and assess temporal evolution of the pollutants. Currently, this topic holds great importance as exposure to air pollutants can have adverse effects on human health and the environment. Therefore, a detailed analysis can provide valuable information for decision-makers in implementing measures to reduce pollution in our cities. Furthermore, this subject is directly linked to the United Nations Sustainable Development Goals (SDGs), particularly SDG 3 "Health and well-being", SDG 13 "Climate action" and SDG 11 "Sustainable cities and communities". This further emphasizes the significance of choosing this theme for the project.

The assessment of air quality focuses on the pollutants which are known to be most prevalent in the atmosphere: sulphur dioxide (SO₂), nitrogen dioxide (NO₂), nitrogen oxide (NO_x), particulate matter (PM10 and PM2.5), benzene (C₆H₆), carbon monoxide (CO) and ozone (O₃). Five of these pollutants are used to calculate the National Air Quality Index.

KEYWORDS:

Air quality index, environment, limit values, pollution

Tabla de contenidos

1. Introducción.....	5
1.1 Estado del arte.....	5
1.1.1 Zonas de bajas emisiones.....	5
1.1.2 Cambio climático.....	6
1.1.3 Ministerio para la transición ecológica y el reto demográfico.....	6
1.2 Contaminantes atmosféricos.....	7
1.3 Índice Nacional de Calidad del Aire.....	7
1.4 Motivación.....	8
2. Objetivos.....	8
2.1 Generales.....	8
2.2 Específicos.....	8
3. Material y metodología.....	9
3.1 Metodología.....	9
3.1.1 Muestra.....	9
3.1.2 Limitaciones.....	9
3.1.3 Procedimientos.....	10
3.1.4 Plan de trabajo.....	11
3.2 Valores límites.....	12
3.3 Lugares de medición.....	12
4. Resultados.....	13
4.1 Estudio de la evolución de la concentración.....	13
4.1.1 Monóxido de carbono.....	14
4.1.2 Monóxido de nitrógeno.....	16
4.1.3. Óxidos de nitrógeno total.....	17
4.1.5 Dióxido de azufre.....	21
4.1.6 Dióxido de Nitrógeno.....	23
4.1.7 Material particulado de 2.5 µg.....	25
4.1.8 Material particulado de 10 µg.....	26
4.1.9 O ₃	27
4.2 Correlaciones.....	29
4.2.1 Correlación entre compuestos.....	30
4.2.2 Población y Tráfico.....	32
4.2.3 Distancia al mar y Altitud.....	34
4.2.4 Días de la semana y mes del año.....	35
4.2.5 Nivel del ICA.....	37
4.2.6 Regresión lineal y predicción.....	38
5. Conclusiones.....	39
Lista de figuras.....	42
Bibliografía.....	44

1. Introducción

1.1 Estado del arte

La calidad del aire es un tema crucial en la agenda política y social de las sociedades modernas. El aire es un recurso vital que está en constante contacto con las personas y puede afectar directamente a la salud y bienestar de estas, por lo que ha sido objeto de preocupación durante décadas. A medida que la población mundial crece y a pesar de los esfuerzos realizados para mejorar la calidad del aire, la contaminación atmosférica sigue siendo uno de los mayores desafíos a los que se enfrenta el país. La pandemia COVID-19 ha tenido un impacto significativo con el cierre de fábricas y la disminución del tráfico vehicular. Sin embargo, con la reapertura de la economía, los niveles de contaminación han vuelto a aumentar.

En los últimos años, se han producido importantes avances en la investigación sobre la calidad del aire y sus efectos en la salud humana. La evidencia científica ha demostrado que la exposición a la contaminación del aire puede aumentar el riesgo de enfermedades respiratorias, cardiovasculares y cáncer, entre otros. Un ejemplo de las repercusiones que está teniendo esta situación en España es la publicación de un estudio de Ecologistas en Acción sobre la calidad del aire en España durante el año 2022. Este estudio expone que, durante el año 2020, la cifra de personas que respiraron aire contaminado por encima de los niveles recomendados por la OMS alcanza los 42 millones de personas, aproximadamente el 88 por ciento de la población española a esa fecha. Además, según un estudio del Instituto de Salud Global de Barcelona: "La reducción de la contaminación atmosférica a las concentraciones medidas más bajas podría evitar 124 729 (83 332-166 535) muertes al año por exposición a PM2.5 y 79 435 (0-215 165) muertes al año por exposición a NO₂".

La normativa española vigente sobre calidad del aire se compone de dos normas principales: la Ley 34/2007 [1] y el Real Decreto 102/2011 [2]. La Ley 34/2007 del 15 de noviembre de 2007 tiene como objetivo alcanzar niveles óptimos de calidad del aire para prevenir efectos negativos en la salud humana y el medio ambiente. Por otro lado, el Real Decreto 102/2011 de 28 de enero transpone al ordenamiento jurídico español el contenido de la Directiva 2008/50/CE [3] y la Directiva 2004/107/CE [4] y busca evitar los efectos nocivos de las sustancias en la salud y en el medio ambiente. Ambas normas han sido modificadas en años posteriores para incorporar nuevos requisitos y normas, como la aprobación del Índice Nacional de Calidad del Aire (ICA) en el Real Decreto 39/2017 [5], que permite informar de una manera clara sobre la calidad del aire que se respira en cada momento. En enero de 2023, se aprobó el Real Decreto 34/2023 [6] para incorporar el Plan Marco de Acción a corto plazo en caso de episodios de alta contaminación.

1.1.1 Zonas de bajas emisiones

Las zonas de bajas emisiones (ZBE) son áreas dentro de la ciudad -o en algunos casos, en la ciudad entera- en las que el acceso a determinados vehículos está restringido. Es una medida pensada para mejorar la calidad del aire mediante la reducción de la circulación de los vehículos más contaminantes y el fomento del uso de aquellos más sostenibles o el transporte público. En estas zonas, únicamente se permite la entrada a los vehículos que cumplen con los mejores estándares de emisiones. Por ejemplo, en Madrid, la zona de

bajas emisiones restringe la entrada a los vehículos que no sean de residentes o 0 emisiones, aunque este área es reducida.

La creación de zonas de bajas emisiones es una medida útil para luchar contra la contaminación atmosférica, pero por sí solas no son suficientes. Para lograr un cambio real, las ZBE han de estar acompañadas de otras medidas que fomenten la disminución del uso del vehículo privado como uno de los principales focos de emisión de gases contaminantes.

1.1.2 Cambio climático

Por definición, el cambio climático es la variación global del clima de la Tierra. Esta variación se atribuye a causas naturales y a la acción del hombre, y se produce sobre todos los fenómenos climáticos: temperatura, precipitaciones, nubosidad, etc, a muy diversas escalas de tiempo. El calentamiento global es una de las principales consecuencias del cambio climático. Desde la era preindustrial, la temperatura promedio de la Tierra ha aumentado y se espera que lo siga haciendo en las próximas décadas. Este aumento de la temperatura está provocando el derretimiento de los glaciares, la acidificación de los océanos, el aumento del nivel del mar y la alteración de los patrones climáticos, entre otras graves consecuencias.

A nivel internacional, se han llevado a cabo varios acuerdos y compromisos para abordar el cambio climático. El más importante es el Acuerdo de París, firmado en 2015 por casi todos los países del mundo, que establece como objetivo limitar el aumento de la temperatura global a menos de 2 °C por encima de los niveles preindustriales, así como un esfuerzo añadido para que este límite se sitúe en 1,5 °C, situación mucho más preferible. Además, numerosos países están adoptando políticas y estrategias para reducir la emisión de gases de efecto invernadero. Estas políticas implican el fomento de las energías renovables, la conservación de la energía, la electrificación del transporte y la salvaguarda de los bosques y otros depósitos de carbono. Además, varias empresas e instituciones se han comprometido a mitigar su huella de carbono.

1.1.3 Ministerio para la transición ecológica y el reto demográfico

La medición de la calidad del aire en España se lleva a cabo por medio de redes de estaciones de medición distribuidas por todo el territorio, gestionadas por el Ministerio para la Transición Ecológica y el Reto Demográfico [7] y las comunidades autónomas, como el RVVCCA [8]. Respecto al primero, el Ministerio para la Transición Ecológica y el Reto Demográfico es un ministerio del Gobierno de España que se encarga de coordinar y llevar a cabo las políticas medioambientales y de lucha contra el cambio climático en el país. MITECO se formó en 2018 a partir de la división del anterior Ministerio de Agricultura, Alimentación y Medio Ambiente, con el objetivo de dar mayor relevancia y prioridad a los temas relacionados con la transición ecológica y la sostenibilidad. Su propósito es proteger y mejorar el medio ambiente, fomentar el desarrollo sostenible y la economía circular, y combatir el cambio climático. Para ello, el Ministerio trabaja en la elaboración de políticas y estrategias que promuevan la reducción de emisiones de gases de efecto invernadero, la conservación de la biodiversidad, la gestión sostenible de los recursos naturales y la adaptación al cambio climático.

1.2 Contaminantes atmosféricos

La contaminación atmosférica se define como la presencia de pequeñas partículas o productos que suponen un riesgo para la salud de las personas y los demás seres vivos en la atmósfera. En gran medida, las actividades humanas como la quema de combustibles fósiles, la industria, la agricultura, el transporte y la generación de energía han supuesto fuentes principales de contaminación del aire, aunque también han habido sucesos naturales puntuales no relacionados con la actividad humana que han contribuido a dicha contaminación, como es el caso de las erupciones volcánicas. Es importante tener en cuenta que para que un agente sea contaminante no es preciso que sea diferente a los componentes naturales de esta, si no que simplemente no se encuentre en su proporción natural. Es decir, un contaminante no lo es por sí mismo, sino por encontrarse en una cantidad que no es natural en su variabilidad.

La atmósfera sufre una mezcla o contaminación de numerosos elementos. Estos agentes nocivos se pueden dividir en contaminantes primarios y secundarios. Los primarios son aquellos que tienen origen directo desde una fuente o foco. En cambio, los secundarios se forman a partir de la reacción o combinación de los anteriores.

1.3 Índice Nacional de Calidad del Aire

El ICA (Índice de Calidad del Aire) es un indicador que se utiliza para medir la calidad del aire en una determinada zona geográfica. Este índice se calcula a partir de la concentración de los principales contaminantes atmosféricos, como el dióxido de azufre (SO₂), el dióxido de nitrógeno (NO₂), el ozono (O₃), las partículas en suspensión (PM10 y PM2.5) y el monóxido de carbono (CO).

Para calcular el ICA, se miden las concentraciones de cada uno de los contaminantes mencionados anteriormente y se comparan con unos valores límite establecidos por la normativa. Estos **valores límite** se conocen como valores umbral y, en este caso, tomaremos los valores umbral que se establecen en el Índice Nacional de Calidad del Aire, que se aprobó en el BOE (Boletín Oficial del Estado), 10/09/2020, núm. 242, Sec. I., Pág. 75835 [9]. A partir de las concentraciones de los diferentes contaminantes y de los valores umbral, se calcula el ICA para cada uno de ellos. El valor final del ICA se obtiene tomando el valor más alto de entre los diferentes valores calculados para cada contaminante. Es decir, en el caso de que un solo contaminante alcance el máximo nivel de contaminación posible, el ICA de la zona será el peor posible, independientemente del resto de contaminantes. Por ejemplo, si el valor correspondiente al ICA del ozono tiene un nivel extremadamente desfavorable, independientemente del nivel de los otros contaminantes, el ICA de ese día será extremadamente desfavorable.

SO ₂		PM2,5		PM10		O ₃		NO ₂		CATEGORÍA DEL ÍNDICE
0	100	0	10	0	20	0	50	0	40	BUENA
101	200	11	20	21	40	51	100	41	90	RAZONABLEMENTE BUENA
201	350	21	25	41	50	101	130	91	120	REGULAR
351	500	26	50	51	100	131	240	121	230	DESFAVORABLE
501	750	51	75	101	150	241	380	231	340	MUY DESFAVORABLE
751-1250		76-800		151-1200		381-800		341-1000		EXTREMADAMENTE DESFAVORABLE

Figura 1. Tabla de valores umbral del Índice de Calidad del Aire. Fuente: [6]

1.4 Motivación

Inicialmente, para el proyecto se barajaron diversas opciones, tales como la moda rápida o un análisis de la comida que se ingiere. Después de evaluar todas las opciones, se decidió centrar el proyecto en el estudio de la calidad del aire en España. Esta temática es de gran importancia en la actualidad debido a que la exposición de contaminantes del aire puede tener efectos negativos en la salud humana y en el medio ambiente. Por lo tanto, un análisis detallado de la calidad del aire puede proporcionar información valiosa para los responsables de la toma de decisiones en la implementación de medidas que reduzcan la contaminación en nuestras ciudades. Con el análisis, se espera llegar a todos los residentes de España para que entiendan cuál es nuestra situación actual respecto a la contaminación atmosférica, y en caso de ser negativa, concienciar para poder mejorarla dando razones y soluciones.

2. Objetivos

2.1 Generales

El objetivo general es ofrecer un análisis estadístico, descriptivo y objetivo de la calidad del aire en las ciudades estudiadas bajo la normativa vigente. También realizar inferencia y predecir el posible ICA en las ciudades estudiadas.

2.2 Específicos

1. Estudiar la evolución de la composición del aire desde 2009 hasta 2020 en las ciudades estudiadas, indicando si la composición es “Buena”, “Razonablemente buena”, “Regular”, “Desfavorable”, “Muy desfavorable” o “Extremadamente desfavorable” según el ICA (índice de calidad del aire).

2. Elaborar una página web cuyos contenidos serán: memoria del proyecto, contexto y situación, mapa donde visualizar la calidad del aire a lo largo del tiempo, vídeo del proyecto, código escrito en *Jupyter Notebook* e información del equipo.
3. Determinar la tendencia de los contaminantes.
4. Contabilizar las ocasiones donde el nivel de calidad del aire ha alcanzado niveles nocivos.
5. Obtener la relación entre contaminantes y correlaciones entre estos y otros factores como la distancia al mar, la altura sobre el nivel del mar y la cantidad de vehículos matriculados en la provincia, entre otros.
6. Realizar una predicción de la evolución de los contaminantes en las ciudades estudiadas.

3. Material y metodología

3.1 Metodología

3.1.1 Muestra

La muestra utilizada ha sido extraída de la base de datos del Ministerio de Transición Ecológica y el Reto Demográfico. Anualmente realizan un estudio sobre la calidad del aire y disponen de un mapa interactivo que muestra la calidad del aire día por día. Los datasets contienen datos desde el año 2000 hasta el 2021 de 9 contaminantes y 5 metales, repartidos en alrededor de 600 estaciones de medición, que a su vez están repartidas en todo el territorio nacional, incluyendo península e islas. La muestra que hemos recabado va desde el año 2009 hasta el año 2021, juntando los datasets correspondientes a cada año. Posteriormente hemos realizado un tratamiento de datos, especificado en la sección posterior, para un análisis más sencillo, seleccionando previamente los datos referentes a un total de 20 estaciones. Estas estaciones están repartidas en toda España y se ha tenido en cuenta que engloben diferentes tipos de terreno, i.e. que haya estaciones costeras, rurales, etc.

3.1.2 Limitaciones

Durante la realización del proyecto, se han tomado diversas limitaciones en términos de capacidades técnicas de los integrantes del equipo, así como relacionadas a la fuente de los datos y herramientas utilizadas.

En primer lugar, los datos recopilados contienen datos ausentes, provocando que haya pequeños espacios e intervalos donde no contemplamos ningún dato. En algunos casos, la cantidad de datos ausentes es total entre 2009 y 2021, posiblemente debido a que determinadas estaciones no realizan la medición de todos los compuestos mencionados. Esto ha provocado de forma directa la consideración de reducir la amplitud del estudio, pues ante la falta de datos, pueden aparecer problemas de magnitud considerable en términos de precisión del estudio. Posteriormente, el costoso proceso de filtrado de datos ha provocado que se redujera el coste temporal puesto que se estaba extendiendo demasiado el tiempo definido para la limpieza y filtrado de los datos.

3.1.3 Procedimientos

A lo largo de todo el proyecto se han realizado diversos procedimientos y tratamientos de datos, ya sea para la limpieza, filtrado o análisis. Para la gran mayoría de ellos, se ha empleado *Python* para facilitar la tarea debido a la repetitividad de la tarea en cuestión.

En primer lugar, para la obtención de la muestra, se han filtrado los datasets de la web del Ministerio mediante un pequeño programa de *Python* con la librería *pandas*. Este programa obtiene la fila correspondiente a la provincia y municipios especificados y los guarda en un dataframe, que una vez finalizado el filtrado se guarda en un fichero de *Excel*. De aquí se han obtenido los datos correspondientes a los 9 compuestos en las 20 estaciones en los años establecidos.

```
# Limpieza de datos
print("Iniciando limpieza...")
for anyo in range(2009, 2021): #Iteración sobre todos los años
    path = "./Datos " + str(anyo) + "/Horarios/" #Dirección relativa de los ficheros de datos
    print("Limiando año " + str(anyo))
    for filename in totalDatos[anyo]: #Iteración sobre los compuestos de un año
        df_inicial = pd.read_excel(path + filename + ".xlsx") #Obtengo el df del compuesto del año correspondiente
        aux = filename.split("_")
        compuesto = aux[0]
        for lugar, condiciones in lugares.items():
            try:
                df = filtrarDatos(condiciones, df_inicial) #Filtro el df
                if not df.empty: #Si no está vacío, lo guardo y añado una tupla al resumen
                    if anyo in resumen:
                        resumen[anyo].append((lugar, compuesto, True))
                    else:
                        resumen[anyo] = [(lugar, compuesto, True)]
                print("Introduciendo datos de {lugar} en {compuesto}_{anyo}.xlsx")
                writer = pd.ExcelWriter(compuesto+"_"+str(anyo)+".xlsx", engine='openpyxl', mode='a', if_sheet_exists='replace')
                #Tenemos un archivo "compuesto_anyo.xlsx" para cada compuesto y su año, posteriormente los pegamos en un archivo denominado "compuesto.xlsx" con los datos de cada año en una hoja.
                #Se utiliza esto para evitar sobrecargar el objeto writer con muchos datos en su variable book
                writer.book = load_workbook(compuesto+"_"+str(anyo)+".xlsx")
                writer.sheets = {ws.title: ws for ws in writer.book.worksheets}
                df.to_excel(writer, sheet_name=str(anyo), startrow=writer.sheets[str(anyo)].max_row, index=False)
                writer.close()
                writer.handles = None
                print("Datos introducidos")
                print("\n")
            except: #Si está vacío, no introduzco nada
                print("El df está vacío")
                print("No se introdujo nada de {lugar} en {compuesto}_{anyo}.xlsx")
                print("\n")
                if anyo in resumen:
                    resumen[anyo].append((lugar, compuesto+"_"+str(anyo)+".xlsx", False))
                else:
                    resumen[anyo] = [(lugar, compuesto+"_"+str(anyo)+".xlsx", False)]
            except Exception as e:
                print("No se pudo introducir los datos de {lugar} en {compuesto}_{anyo}.xlsx")
                print("\n")
                print(f"El error es:\n{e}")
                if anyo in resumen:
                    resumen[anyo].append((lugar, compuesto+"_"+str(anyo)+".xlsx", False))
                else:
                    resumen[anyo] = [(lugar, compuesto+"_"+str(anyo)+".xlsx", False)]
print("Limpieza completada")
```

Figura 2. Fragmento del programa de filtrado de datos

Los ficheros de datos resultantes de este programa se combinan en el fichero “DatosFinales.xlsx”, donde se encuentran todos los datos ordenados, incluyendo los espacios en blanco causados por la no medición de ellos. En este fichero, se contienen, en una misma hoja, los datos correspondientes a cada estación y a cada compuesto medido por él, de forma que es más fácil de trabajar con los datos.

Posteriormente, se han agrupado las 20 estaciones por el tipo de área en el que se encuentra. A partir de las 5 agrupaciones dadas, se ha realizado un programa para obtener la evolución de la concentración media de cada tipo de área, pues la otra alternativa era realizar el análisis estación por estación, un proceso prácticamente 4 veces más tedioso. Así, quedan un total de 45 gráficas (9 contaminantes por 5 tipos de estación). Para estas visualizaciones se ha utilizado también la librería *plotly.graph_objs*.

```

for chem in chems:
    for tipo, lista in dic.items():
        layout=go.Layout(title=tipo+"/"+ " "+chem)
        fig=go.Figure(layout=layout)
        for place in lista:
            valores=dfMM.loc[:, place+"/"+ " "+chem]
            trace=go.Scatter(x=fechas, y=valores, mode='lines+markers', name=place)
            fig.add_trace(trace)
    #fig.write_image("C:/Users/prado/OneDrive/Dокументos/PROYECTO/GRÁFICOS/"+tipo+"-"+chem+".jpg")
fig.show()

```

Figura 3. Fragmento del programa de creación de visualizaciones

Paralelamente, se ha realizado otro programa para calcular el ICA, cuyo cálculo se ha mencionado con anterioridad. Para esto, se obtuvo el ICA correspondiente a cada contaminante y después se seleccionó, para cada estación y día, el nivel más alto de ICA.

```

#Now, we are going to create a new column that shows the AQI for each place
#We obtain the value by picking the worst level of AQI of the place

aqi_values = {
    1 : "Buena",
    2 : "Razonablemente buena",
    3 : "Regular",
    4 : "Desfavorable",
    5 : "Muy desfavorable",
    6 : "Extremadamente desfavorable",
    0 : "Dato no válido"
}

chemicals = ["SO2", "NO2", "PM10", "PM25", "O3"]
unique_places = [place[:-5].strip() for place in so2_places]
place_chemicals = {}
for place in unique_places:
    place_chemicals[place] = []
    for chemical in chemicals:
        place_chemicals[place].append(place+"/"+ " "+chemical + " ICA")

for place, lista in place_chemicals.items():
    df[place + " ICA"] = df[[lista[0], lista[1], lista[2], lista[3], lista[4]]].max(axis=1).map(aqi_values)
    df[place + " ICA num"] = df[[lista[0], lista[1], lista[2], lista[3], lista[4]]].max(axis=1)

```

Figura 4. Fragmento del programa de medición del ICA

3.1.4 Plan de trabajo

En primer lugar, se han asignado una serie de roles principales teniendo en cuenta las habilidades de cada integrante del equipo. Independientemente de ello, durante la praxis los integrantes también han realizado tareas que no corresponden a su rol debido a una mayor carga de trabajo en otros roles. Por ejemplo, la persona encargada de la redacción, puede haber complementado a otra persona en el rol de análisis en la parte inicial del proyecto.

La distribución de los roles fue el siguiente:

- Alejandro Hervás: programador
- Álvaro Prado: Investigador
- Carles Navarro: programador.
- Sheila Díaz: Artista e investigadora.
- Laiqian Ji: coordinador/portavoz y programador.

Posteriormente, se ha realizado una planificación estimando el tiempo que tardará en ejecutarse cada tarea. Esto queda reflejado en una hoja de cálculo de Google Sheets en la que se han aplicado diferentes funciones para formatearla. Se dividieron las tareas en tres secciones principales: análisis, entregables y actividades de control.

Las primeras corresponden a las tareas inherentes al estudio de los datos, su limpieza y el análisis de correlación. En segundo lugar, encontramos las tareas relacionadas con los elementos que se deben entregar al concluir el trabajo, estos son la página web, el vídeo, la presentación y el informe. Finalmente, las actividades de control se refieren a las

presentaciones realizadas en clase a lo largo del período del estudio, estas están definidas previamente por el profesorado, y a la presentación final, donde se deben entregar todos los elementos creados.

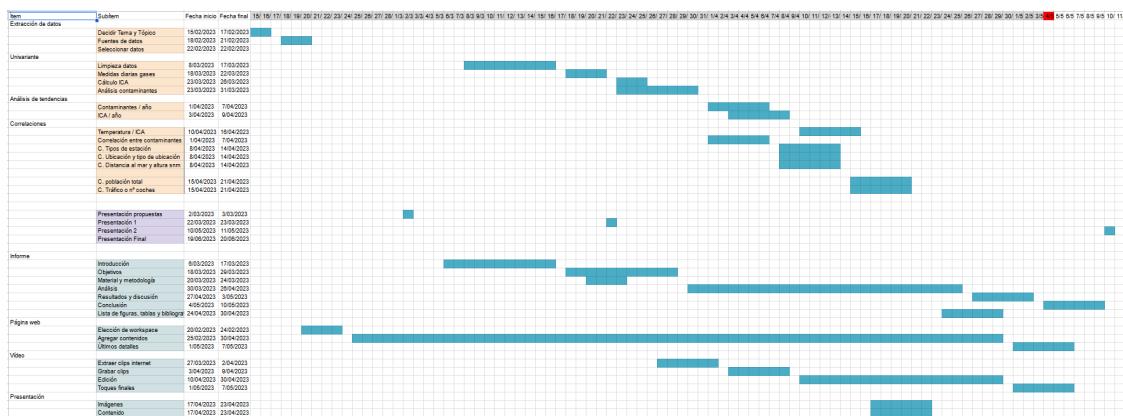


Figura 5. Diagrama de Gantt [9]

3.2 Valores límites

En la Directiva 2008/50/CE del Parlamento Europeo y del Consejo, de 21 de Mayo de 2008, relativo a la calidad del aire ambiente y a un atmósfera más limpia en Europa, se instauran una serie de valores límite de todos los contaminantes presentes en el aire siguiendo una serie de criterios científicos con el fin de evitar, prevenir y reducir los efectos nocivos que causan estos sobre el ser humano y el medio ambiente. Todas las zonas deben instaurar una política para asegurarse que los datos se encuentran por debajo de los valores límite.

3.3 Lugares de medición

Se han seleccionado un total de 20 estaciones repartidas en todo el territorio nacional, comprendiendo diferentes tipos de clima, ubicación y población, de forma que haya estaciones de diversos tipos.

El formato que siguen las estaciones para su identificación es (Provincia, Municipio, Estación), pudiendo consultar todos los códigos correspondientes a cada estación en el archivo “Metadata.xlsx”. Estos códigos son idénticos a los presentados en la base de datos ofrecida por MITECO y se pueden visualizar los códigos de todas las estaciones en el libro de datos “metadata”, que aparece en todos los ficheros descargables.

Se ha realizado la clasificación según el tipo de ubicación de la estación observada:

- **Estaciones urbanas costeras** si se encuentran cerca de una urbe de tamaño considerable y, además, se encuentran en zonas costeras o cercanas a ellas
- **Islas** si son estaciones pertenecientes a las Islas Canarias y las Islas Baleares.
- **Rurales** si son estaciones que se encuentran en pequeñas poblaciones de interior, en este caso, únicamente encontramos una estación de Cáceres, Toril.
- **Urbanos de interior** a aquellas estaciones que se encuentran cerca o dentro de una urbe de tamaño considerable en zonas de interior.
- **Industriales** a aquellas estaciones que se ubican cercanas a zonas industriales o polígonos.

Tipo de ubicación	Estaciones (Provincia, Municipio, Estación)
Urbanos costeros	1. Barcelona, L'Eixample (8, 19, 43) 2. Valencia, Paterna (46, 190, 5) 3. Málaga (29, 67, 6) 4. Bilbao, Dirección de Salud (48, 20, 6) 5. Cádiz (11, 12, 2) 6. Gijón (33, 24, 25) 7. Almería (4, 13, 4)
Islas	1. Palma (7, 40, 2) 2. Tenerife (38, 38, 10)
Rurales	1. Cáceres, Toril (10, 182, 1)
Urbanos de interior	1. Madrid, Plaza de España (28, 79, 4) 2. Albacete (2, 3, 1) 3. Burgos, Plaza lavaderos (9, 59, 3) 4. Zaragoza, Roger de la Flor (50, 297, 29) 5. Sevilla, La ranilla (41, 91, 10) 6. Valladolid, Arco de Ladrillo II(47, 186, 30) 7. Badajoz (6, 15, 1)
Industriales	1. Igualada, La Masuca (Barcelona) (8, 102, 5) 2. El Grao de Castellón (12, 40, 10) 3. Vigo, Estación 1, Avda Citröen (36, 57, 21)

Figura 6. Tabla de estaciones. Fuente: Elaboración propia a partir de los datos

4. Resultados

4.1 Estudio de la evolución de la concentración

En primer lugar, se han obtenido las gráficas de tendencias. Cada una de las gráficas corresponde a un compuesto y a un grupo de estaciones a la vez, de forma que hay 9 grupos de 5 gráficas ya que hay 9 compuestos que se analizan para los grupos de 5 estaciones. Dichas gráficas se han realizado con un programa mencionado previamente, utilizando la biblioteca plotly. A partir de estas gráficas, se han respondido, en primer lugar, a una serie de cuestiones en relación a las tendencias observadas en cada una de ellas. Dichas cuestiones son las siguientes:

1. Del compuesto analizado, ¿existen datos faltantes o se cuenta con los datos de todos los años? ¿Tiene relación este compuesto con el cálculo del Índice Nacional de Calidad del Aire?
2. Comentar brevemente los parámetros muestrales
3. ¿Qué estaciones están mostrando una tendencia creciente y/o decreciente en los últimos años (directriz del movimiento)? Comentar también la ciclicidad de los datos (estacionalidad).
4. ¿Existe relación entre las tendencias de los compuestos?

Previamente, con ayuda del programa mencionado anteriormente y haciendo uso de la librería *scipy*, se ha realizado el test Shapiro-Wilk para analizar la normalidad de las variables. Como resultado, se ha obtenido que únicamente 57 de las 180 variables seguían una distribución normal. Por ello, únicamente se adjuntan las gráficas que presentan estaciones normalizadas y una cantidad de datos considerable, evitando aquellas que presentan muy pocos datos y tengan muchos ausentes.

4.1.1 Monóxido de carbono

Urbanos Costeros/ CO

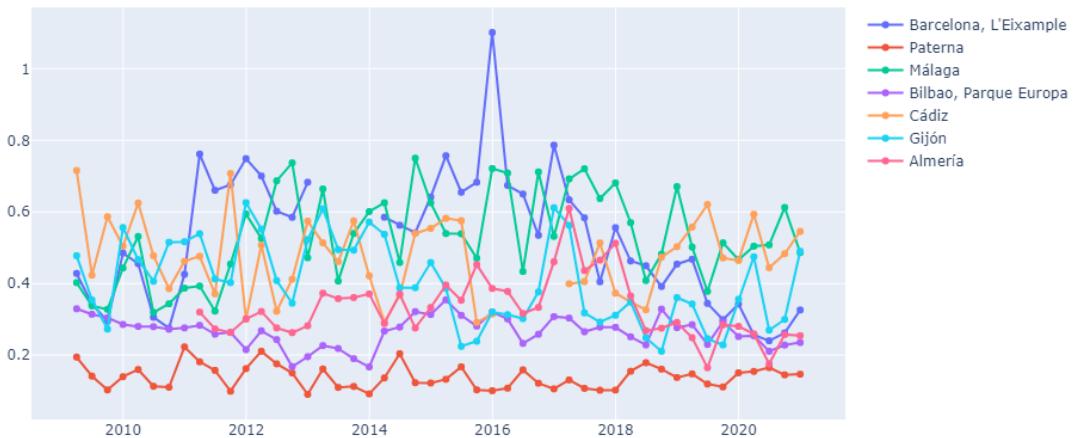


Figura 7. Evolución de la concentración media Urbanos Costeros CO.

Respecto a las estaciones estudiadas en el compuesto CO, se poseen datos de prácticamente todos los años estudiados, habiendo pocos datos faltantes como en el año 2013 en la estación de L'Eixample, en el año 2016 en la estación de Cádiz y en el intervalo desde 2009 hasta 2010 en Almería. Para el cálculo del Índice Nacional de Calidad del Aire, no se tiene en cuenta este compuesto, por lo que no habrá complicaciones en ese aspecto. Las tendencias relativas a este compuesto son relativamente parecidas, aumentando y disminuyendo en las mismas zonas del año. Sin embargo, la variabilidad que presenta cada estación es diferente, como podemos ver en la Figura 7, cada estación varía en mayor o menor magnitud.

	Barcelona, L'Eixample CO	Paterna CO	Málaga CO	Bilbao, Parque Europa CO	Cádiz CO	Gijón CO	Almería CO
count	4018,000	4383,000	4383,000	4383,000	3713,000	4346,000	3557,000
mean	0,524	0,139	0,530	0,266	0,474	0,405	0,330
std	0,263	0,065	0,208	0,083	0,199	0,229	0,118
min	0,000	0,000	0,083	0,046	0,007	0,066	0,018
25%	0,321	0,100	0,376	0,210	0,326	0,263	0,246
50%	0,478	0,113	0,500	0,260	0,449	0,368	0,316
75%	0,671	0,175	0,667	0,313	0,603	0,495	0,387
max	2,233	0,508	1,347	0,758	1,238	7,145	0,978

Figura 8. Parámetros muestrales Urbanos Costeros CO.

Revisando los parámetros y comparando con la Figura 8, se observa que Paterna es la población que presenta menor presencia del monóxido de carbono, seguida por la estación del Parque Europa en Bilbao, que presenta una presencia un 93% mayor que Paterna. También se observa que, tanto Barcelona como Málaga cuentan con la mayor presencia de

este compuesto. Respecto a las desviaciones, estas son mayores cuanto más presencia presentan.

Urbanos Interior/ CO

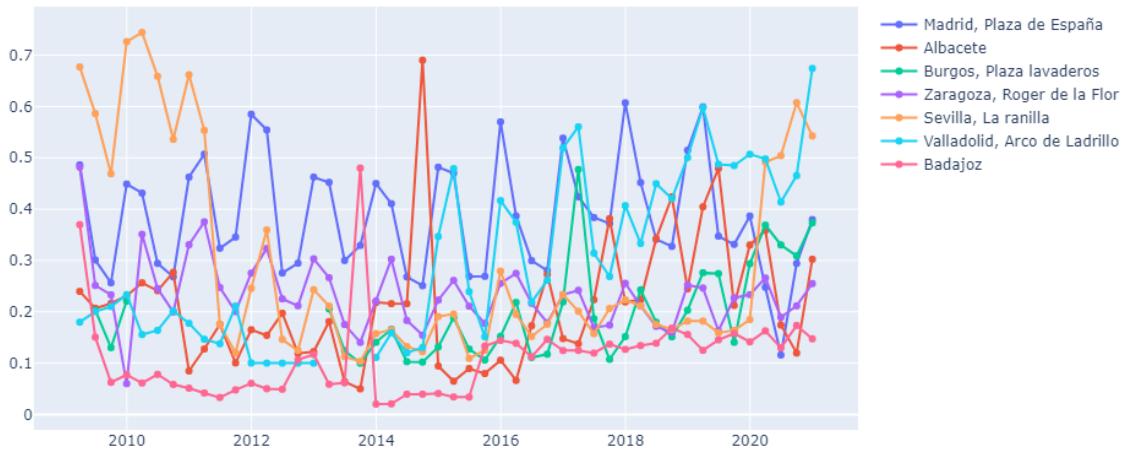


Figura 9. Evolución de la concentración media Urbanos Interior CO.

En la Figura 19 se observa una situación parecida a la mencionada anteriormente, se contemplan esos picos y fondos en temporadas parecidas, aunque con gran grado de variabilidad. Además, tras observar ambas gráficas de tendencias, se puede pensar que las tendencias están siendo constantes y en algunas ocasiones descendentes, por lo que se puede pensar en que las medidas tomadas pueden estar surgiendo efecto.

Barcelon	Madrid, Plaza de España CO	Albacete CO	Burgos, Plaza lavaderos CO	Zaragoza, Roger de la Flor CO	Sevilla, La ranilla CO	Valladolid, Arco de Ladrillo CO	Badajoz CO
count	4383,000	4383,000	3056,000	4261,000	4136,000	4014,000	4204,000
mean	0,384	0,213	0,199	0,237	0,290	0,298	0,109
std	0,186	0,153	0,126	0,109	0,225	0,209	0,112
min	0,100	0,040	0,100	0,009	0,100	0,100	0,000
25%	0,265	0,094	0,100	0,163	0,139	0,113	0,042
50%	0,333	0,197	0,146	0,216	0,180	0,225	0,080
75%	0,463	0,276	0,254	0,285	0,412	0,438	0,148
max	1,475	1,056	0,958	1,178	1,386	1,575	2,550

Figura 10. Parámetros muestrales Urbanos interior CO.

Analizando los parámetros muestrales y comparándolos con el visual, se puede visualizar como aquellas estaciones con menor media suelen tener menos volatilidad, pero no es una norma que sigan todas las estaciones. Por ejemplo, la dispersión está bastante agravada en muchas de las estaciones, en especial en Badajoz (donde la desviación típica es mayor que la media) y más débil en Madrid. Sin embargo, la variabilidad no está achacada al nivel promedio, pues como se ha observado, los coeficientes de variación no siguen una correlación con la media. Es decir, tiene más que ver con la estación en sí, que con la media de esa estación.

	Igualada, La Masuca (Barcelona) CO	El Grao de Castellón CO	Vigo, Estación 1, Avda Cítröen CO
count	3993,000	3927,000	4024,000
mean	0,334	0,176	0,189
std	0,144	0,104	0,132
min	0,200	0,000	0,000
25%	0,225	0,104	0,100
50%	0,300	0,142	0,140
75%	0,396	0,213	0,242
max	2,279	1,600	1,981

Figura 11. Parámetros muestrales Industriales CO.

4.1.2 Monóxido de nitrógeno

Respecto al análisis del compuesto óxido de nitrógeno, se cuenta con casi la totalidad de los datos, a expensas de situaciones concretas como el año 2010 en Cádiz, y el año 2013 en Paterna. Para el cálculo del Índice de Calidad del Aire, no se tiene en cuenta este compuesto, por lo que no habrá complicaciones en ese aspecto.

Urbanos Costeros/ NO

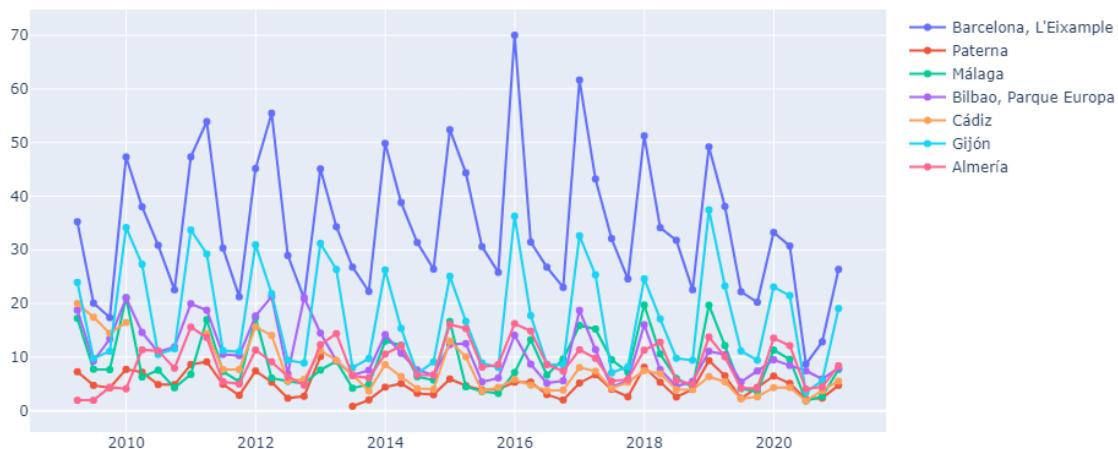


Figura 12. Evolución de la concentración media Urbanos Costeros NO.

Este compuesto presenta una relación más notoria que el primer compuesto, como se contempla en las Figuras 12 y 13. Vemos que la ciclicidad es prácticamente total y los movimientos son similares, teniendo en cuenta que los niveles son mayores o menores. La tendencia es totalmente constante, no han crecido ni disminuido los niveles. No obstante, sí se observa que las estaciones relativas a ciudades más grandes como Barcelona y Madrid presentan concentraciones elevadas, sugiriendo que el NO aumenta según el tamaño de la población.

Urbanos Interior/ NO

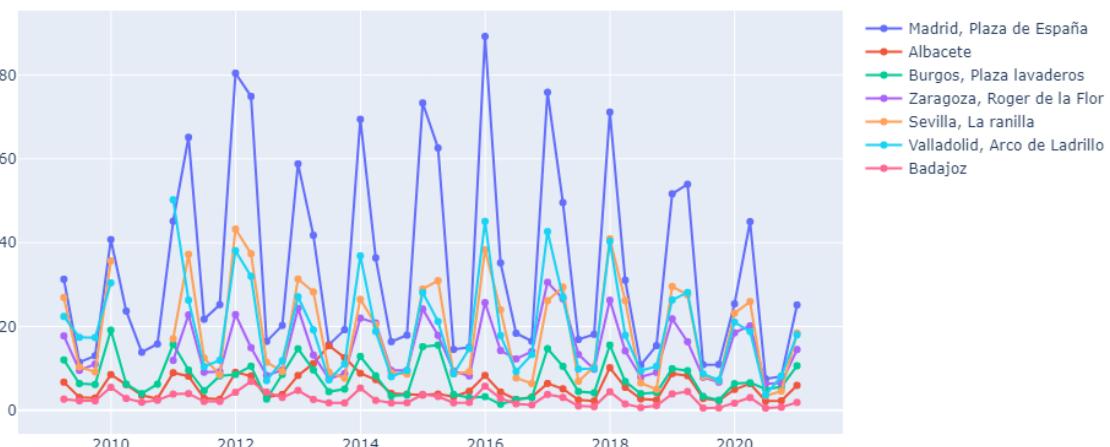


Figura 13. Evolución de la concentración media Urbanos Interior NO.

Analizando los parámetros muestrales de las estaciones y comparándolos con los gráficos obtenidos, se encontró que las poblaciones más grandes presentan mayor presencia de monóxido de nitrógeno. En comparación con Paterna, la presencia de contaminantes en Barcelona es aproximadamente un 86,3% mayor, poniendo en contexto la diferencia de niveles.

	Barcelona, L'Eixample NO	Paterna NO	Málaga NO	Bilbao, Parque Europa NO	Cádiz NO	Gijón NO	Almería NO
count	4355,000	3786,000	4181,000	4159,000	3893,000	4339,000	4333,000
mean	34,109	4,912	9,037	10,688	7,377	17,351	8,887
std	27,875	4,590	11,437	10,719	7,405	18,594	8,219
min	1,000	0,000	0,000	0,000	0,167	1,000	0,542
25%	16,667	2,250	3,125	3,833	2,958	6,375	4,125
50%	27,000	3,375	5,125	7,176	4,750	10,609	6,417
75%	42,000	5,739	9,625	13,563	9,000	20,333	10,250
max	290,167	50,010	118,400	85,000	88,333	144,125	80,875

Figura 14. Parámetros muestrales Urbano Costeros NO.

Una situación curiosa es la respectiva a las estaciones industriales (Figura 15), que presentan una concentración prácticamente pareja en las tres estaciones que pertenecen a esta categoría. Coincidir en los ciclos y en los niveles, sugiriendo que es posible que el NO puede verse influenciado fuertemente por el hecho de pertenecer a una zona industrial, cosa que parece lógica pues estas zonas emiten gases de combustión, donde es fácil hallar este compuesto. Por otro lado, cabe mencionar que la tendencia es ligeramente decreciente durante los últimos años.

Industriales/ NO



Figura 15. Evolución de la concentración media Industriales NO.

4.1.3. Óxidos de nitrógeno total

De forma similar al anterior compuesto, pero en grado menor, las tendencias y ciclos son similares entre las diferentes estaciones, tal y como observamos en las Figuras 16 y 17, correspondientes a zonas urbanas costeras e islas. Obviamente, teniendo en cuenta que existen diferencias en cuanto al nivel promedio de los diferentes contaminantes. Al igual que en otros casos previamente mencionados, los picos también son en épocas invernales, coincidiendo con las épocas de mayor emisión de gases de combustión.

Por otro lado, la ausencia de datos es muy baja en este compuesto, además, no se contabiliza para el ICA por lo que no traerá problemas en ese aspecto.

Urbanos Costeros/ NOx



Figura 16. Evolución de la concentración media Urbanos Costeros NOx.

Islas/ NOx

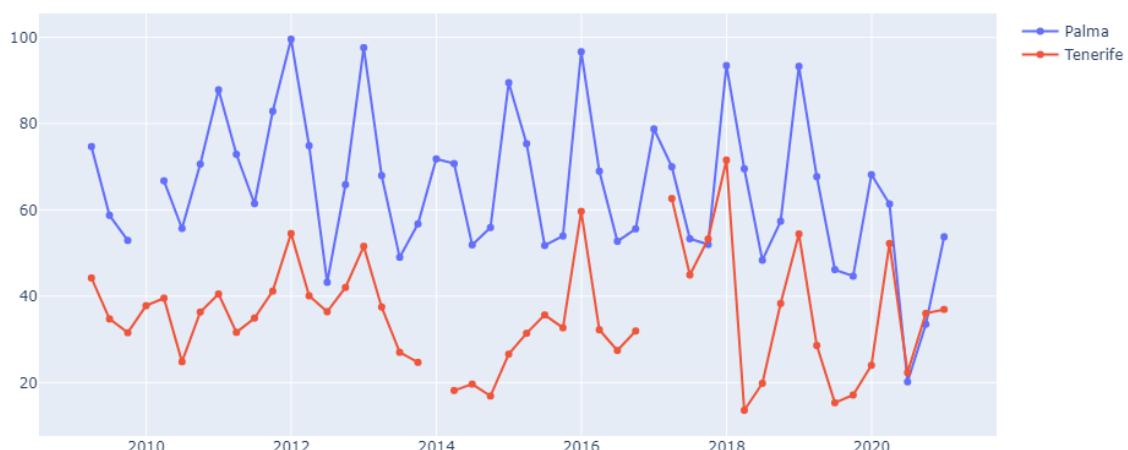


Figura 17. Evolución de la concentración media Islas NOx.

Revisando los parámetros y comparando con la gráfica, se observa que algunas estaciones como Barcelona presentan mayor presencia de NOx (con una media de 107,305 µg/m³ y un valor máximo de 553,792 µg/m³, y una mediana de 95,667 µg/m³). En comparación con Paterna, la presencia de contaminantes en Barcelona es aproximadamente un 78,30% mayor, mostrando otra vez esa gran diferencia. Se puede observar en el gráfico que Barcelona queda muy separada del resto.

	Barcelona, L'Eixample NOx	Paterna NOx	Málaga NOx	Bilbao, Parque Europa NOx	Cádiz NOx	Gijón NOx	Almería NOx
count	4377,000	3504,000	4228,000		4181,000	4220,000	4345,000
mean	107,305	23,235	42,049		43,934	27,933	56,982
std	58,074	16,211	29,713		26,494	19,061	41,074
min	0,000	1,045	4,792		0,000	3,077	5,483
25%	70,000	11,870	22,000		25,333	15,012	30,813
50%	95,667	18,822	33,438		37,125	22,833	44,167
75%	130,909	29,583	52,139		55,125	35,172	69,750
max	553,792	128,933	225,792		209,125	184,792	355,087

Figura 18. Parámetros muestrales Urbanos Costeros NOx.

Mencionamos en este caso a la estación de Toril, en Cáceres, por su evolución muy favorable a la baja, indicando que se está reduciendo la contaminación por óxidos nitrosos en la zona. Sin embargo, esto puede ser causado por la disminución de la población y tráfico en esta zona, que se ha visto gravemente afectada por dichos problemas durante los últimos años.

Rural/ NOx

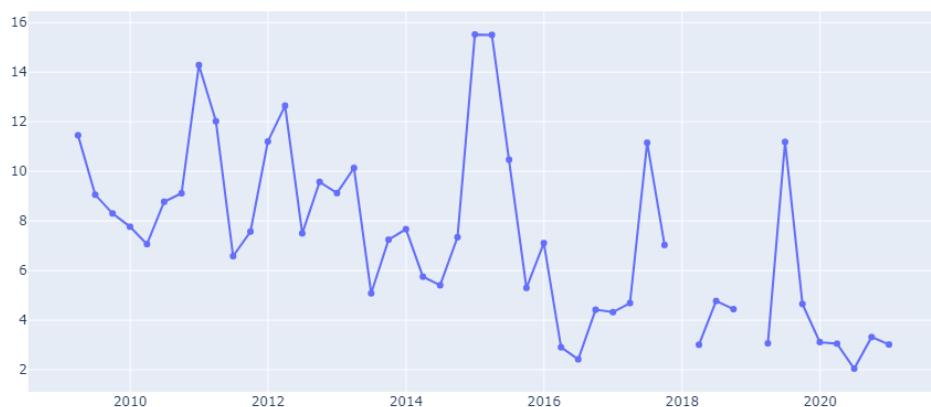


Figura 19. Evolución de la concentración media Rural NOx.

4.1.4 Benceno

En el caso del benceno, la ausencia de datos ha sido mayor, provocando una menor cantidad de gráficas coherentes. En varios tipos de estación, no se ha podido obtener la media trimestral que se ha estado utilizando hasta ahora. Por ese motivo, los resultados son mucho menores a los de otros compuestos. Sin embargo, este no es el único caso, pues el del material particulado es peor. Aún así, dicha ausencia

Urbanos Costeros/ C6H6

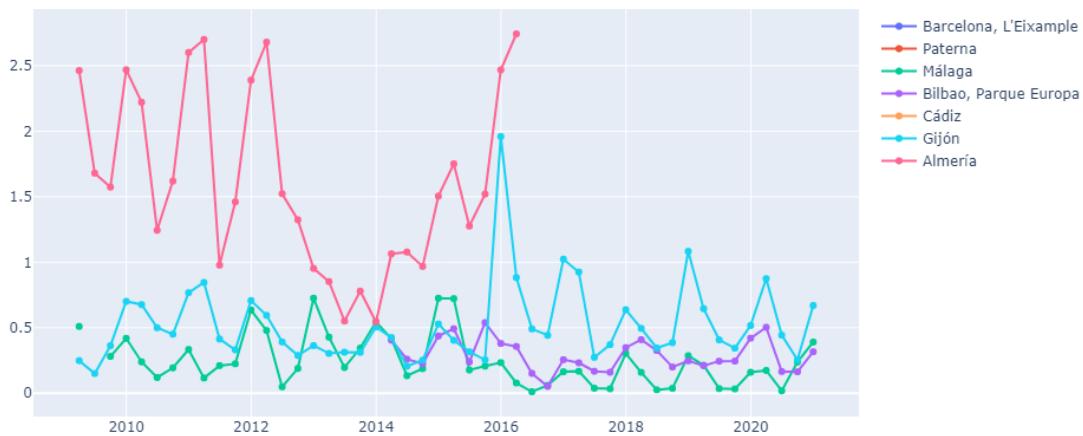


Figura 20. Evolución de la concentración media Urbanos Costeros C6H6.

De las ciudades costeras, sólo se tienen datos del compuesto C6H6 en Gijón, Málaga, Bilbao y Almería (en Almería se interrumpen en 2016), no se han podido extraer los datos relativos a las otras estaciones. Las observaciones más interesantes corresponden al alto nivel promedio de benceno contenido en el aire de Almería, que supera por 3 veces a la segunda estación que más nivel de benceno contiene -Gijón-.

Por otro lado, las tendencias respetan lo visto hasta ahora, realizando aumentos en temporada invernal y disminuyendo en temporadas estivales. Son, por lo general, constantes.

Urbanos Interior/ C6H6

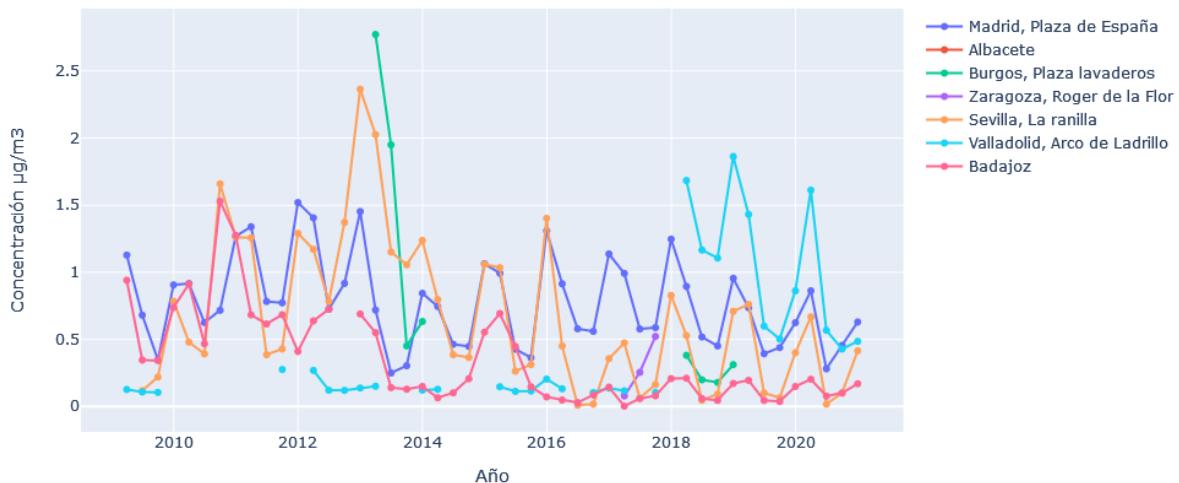


Figura 21. Evolución de la concentración media Urbanos Interior C6H6.

Respecto a las estaciones urbanas de interior, descontando la ausencia de bastantes datos, se contempla una situación similar a la anterior, con una ciclicidad generalizada y con movimientos similares. Sin embargo, cabe destacar que las tendencias de la mayoría de las estaciones presentan movimientos descendentes, lo cual es una buena señal.

4.1.5 Dióxido de azufre

El SO₂ es un compuesto especialmente relevante para nuestro estudio, ya que es uno de los 5 compuestos utilizados para medir el Índice de Calidad del Aire. Por lo tanto, se prioriza que haya una cantidad considerable de datos.

Urbanos Costeros/ SO₂

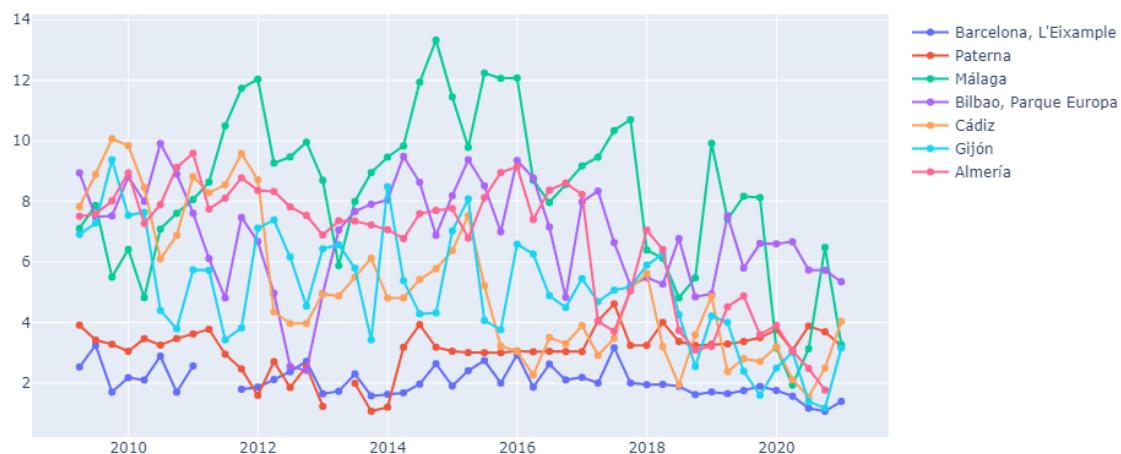


Figura 22. Evolución de la concentración media Urbanos Costeros SO₂.

En las estaciones urbanas costeras, hay gran variedad de casos respecto a los movimientos, aunque parece que la mayoría son descendentes. Por lo general, se respeta la ciclicidad, pero también existen estaciones que presentan movimientos prácticamente constantes. Además, los picos no parecen darse justamente en épocas invernales. Destaca Málaga, Bilbao y Almería como las estaciones con mayor nivel promedio de dióxido de azufre, esto se puede explicar al considerar que en el sur de España se encuentran dos grandes yacimientos de azufre.

Rural/ SO₂

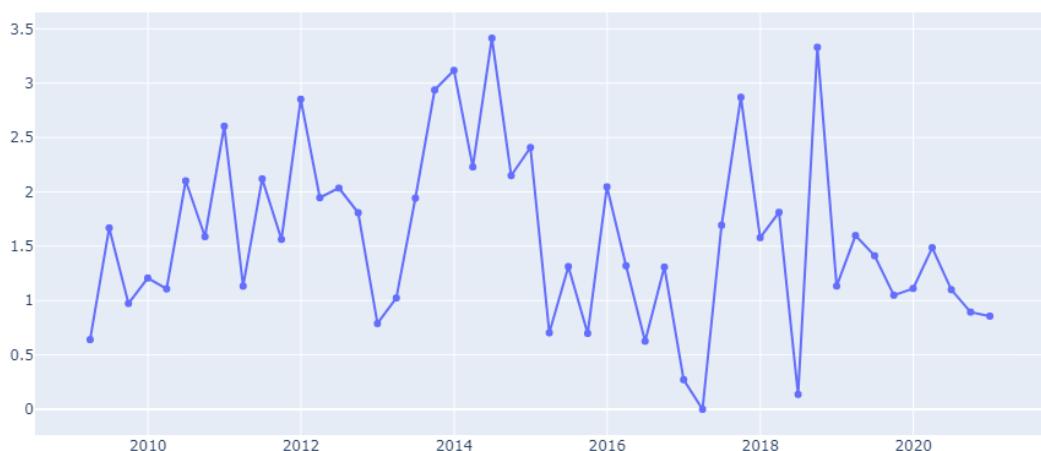


Figura 23. Evolución de la concentración media Rural SO₂.

Observando la gráfica de la estación rural, se puede notar que la media es notablemente más baja que en las estaciones vistas anteriormente, situándose al nivel de Barcelona y Valencia. Respecto a la dispersión, vemos que tiene una desviación típica relativamente alta

y un coeficiente de variación de 0,822. Analizando el gráfico, parece haber una tendencia constante. Además se visualizan picos en los valores, aunque no se puede afirmar que siguen algún patrón.

Urbanos Interior/ SO2



Figura 24. Evolución de la concentración media Urbanos Interior SO2.

Las estaciones urbanas de interior presentan un caso similar al de los costeros, los movimientos son bastante aleatorios y siguen un patrón complejo. Aunque sí es verdad que la mayoría se concentran en valores inferiores a 10 µg/m³. En este caso, Madrid se sitúa también por encima del nivel de las otras estaciones, afectado también por unos primeros años con concentraciones más elevadas.

Observando el gráfico, se visualiza que Madrid tiene una clara tendencia descendente. El resto de estaciones siguen una tendencia constante. Hay picos en todas las estaciones pero no se puede afirmar que siguen algún patrón, como se ha mencionado antes.

Industriales/ SO2

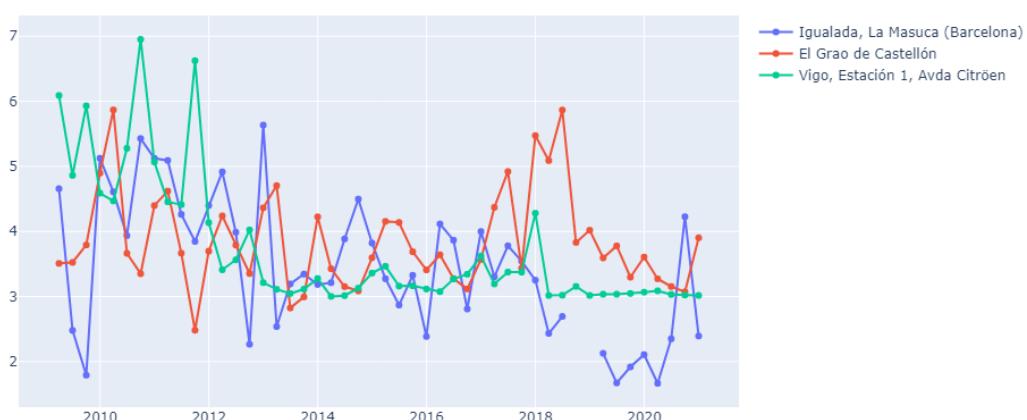


Figura 25. Evolución de la concentración media Industriales SO2.

Respecto a los datos en la estaciones industriales del compuesto SO2 se contemplan en las tres estaciones unas medias muy parecidas, siendo la más alta la de Castellón y la más baja la de Igualada. Al analizar la dispersión, los datos de Igualada son lo más dispersos, con un coeficiente de variación de 0,7. Castellón y Vigo tienen coeficientes de variación de 0,435 y 0,431 respectivamente. Analizando el gráfico, tanto en Igualada como Castellón se

sigue una tendencia constante, mientras que en Vigo se sigue una tendencia descendente. En las tres estaciones, hay picos periódicos que suelen coincidir con los meses de invierno.

4.1.6 Dióxido de Nitrógeno

Urbanos Costeros/ NO2



Figura 26. Evolución de la concentración media Urbanos Costeros NO2.

Respecto a los datos de las estaciones costeras del compuesto NO2 se aprecia que la media de Barcelona ($50 \mu\text{g}/\text{m}^3$) destaca notablemente sobre el resto, que toman valores entre 15 y $30 \mu\text{g}/\text{m}^3$. En cuanto a los movimientos, vemos que hay una clara diferencia entre Barcelona y el resto. En Barcelona, los valores siguen una tendencia ascendente hasta 2012, donde se inicia una leve tendencia descendente que se acrecienta en 2020. En el resto de estaciones, se sigue una leve tendencia descendente durante todos los años estudiados. Todas las estaciones tienen picos periódicos en los mismos meses del año, que coinciden con el invierno.

Islas/ NO2

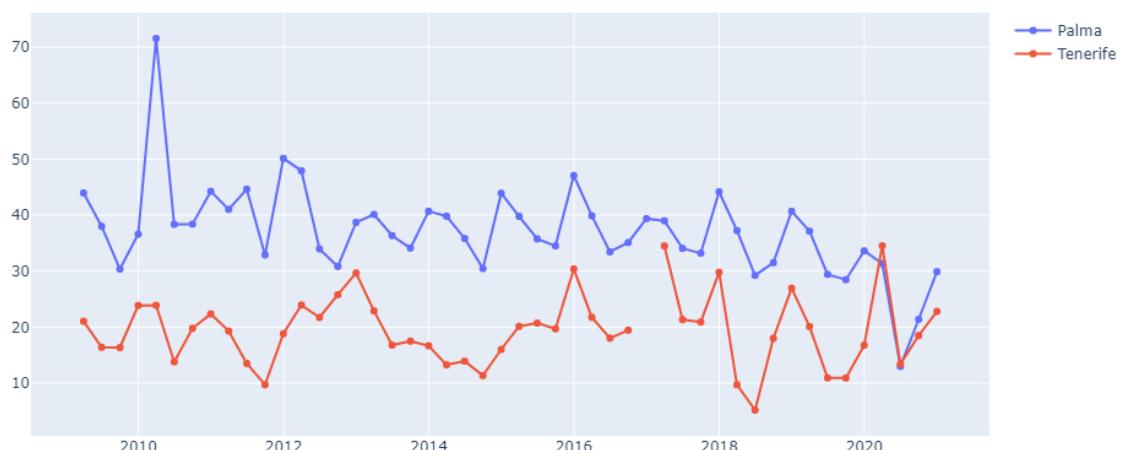


Figura 27. Evolución de la concentración media Islas NO2.

En cuanto a los datos en las estaciones isleñas de NO2, se contempla que la media de Palma ($36 \mu\text{g}/\text{m}^3$), es considerablemente superior que la de Tenerife ($19 \mu\text{g}/\text{m}^3$). En

En cuanto a los movimientos, se aprecia que en Palma se sigue una leve tendencia descendente, mientras que en Tenerife se sigue una tendencia constante. Ambas estaciones tienen picos periódicos en los meses de invierno.

Rural/ NO₂



Figura 28. Evolución de la concentración media Rural NO₂.

Referenciando los datos en Toril de NO₂, se contempla que su media ($5 \mu\text{g}/\text{m}^3$) es notablemente inferior a la de las estaciones medidas anteriormente. En cuanto a su evolución, se aprecia una clara tendencia descendente, interrumpida por un pico a finales de 2017. En cuanto a la ciclicidad, se observa que la variable presenta constantes picos en su transcurso, pero con una tendencia decreciente que también se refleja en estos, indicando una disminución considerable de su nivel de dióxido de nitrógeno.

Urbanos Interior/ NO₂

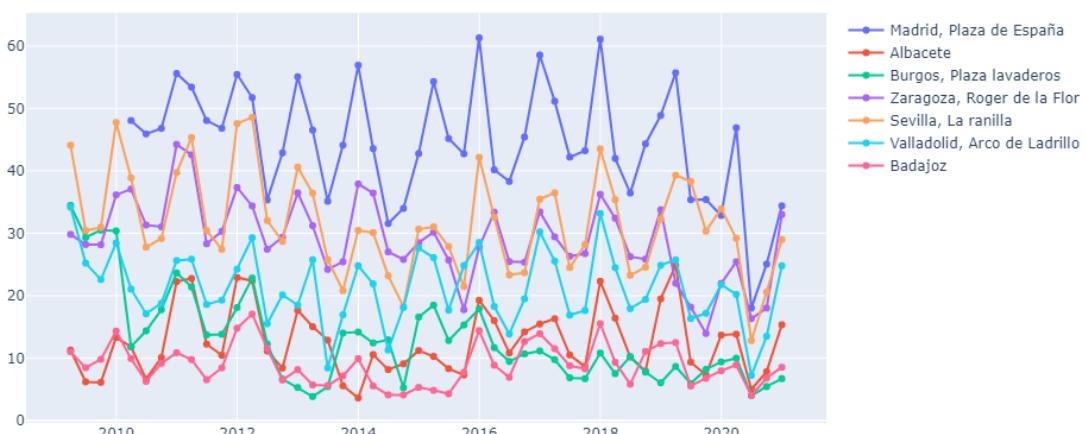


Figura 29. Evolución de la concentración media Urbanos Interior NO₂.

Respecto a los datos en las estaciones urbanas de interior, se aprecia que la media de Madrid es notablemente superior que el resto, seguida por las de Sevilla y Zaragoza. En cuanto a la evolución, se contempla que todas las estaciones siguen una tendencia constante, con picos periódicos en los meses de invierno. La perspectiva general es uniforme, pero contando con la diferencia de nivel correspondiente a cada estación.

4.1.7 Material particulado de 2.5 µg

El material particulado de 2.5 µg también forma parte de la medición del Índice de Calidad del aire y es uno de los contaminantes más preocupantes, pues según el Instituto de Salud Global de Barcelona, su reducción evitaría un gran número de muertes al año. Desgraciadamente, cuenta con bastantes datos faltantes, posiblemente debido a que las estaciones no presentan datos de este contaminante.

Urbanos Interior/ PM25



Figura 30. Evolución de la concentración media Urbanos Interior PM2.5.

La primera impresión es que, al igual que las otras variables, sigue cierta ciclicidad aunque no es tan clara, no se puede afirmar la acumulación de este en una época determinada del año. Además, la tendencia es totalmente descendente, implicando una reducción de la presencia promedio de este contaminante. Algo curioso es que, de las estaciones estudiadas, es Albacete la que más nivel tiene respecto a este contaminante, sin contar con las ausencias de Madrid y Sevilla, entre otros.

Industriales/ PM25

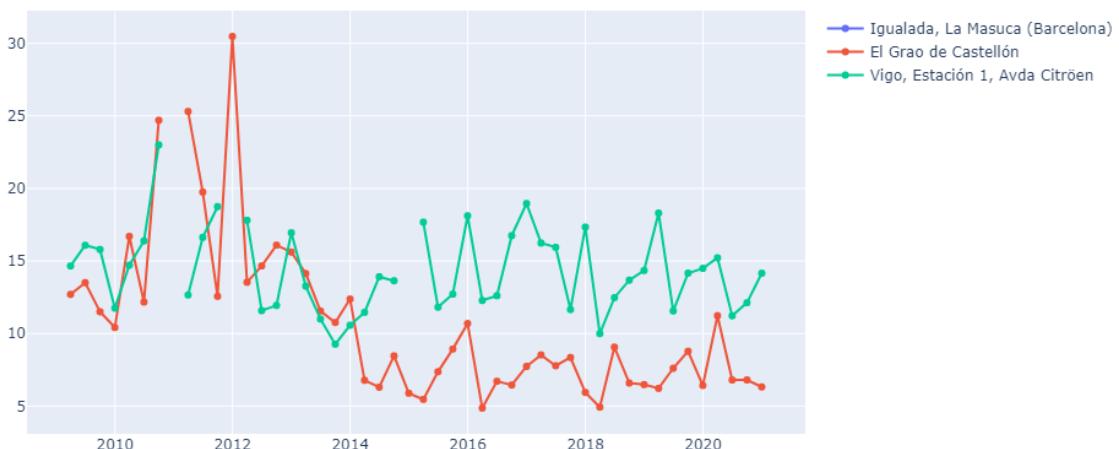


Figura 31. Evolución de la concentración media Industriales PM2.5.

Además de las estaciones urbanas de interior, se cuenta con datos de las estaciones industriales (Figura 31), pero no hay más respecto a costeras, islas ni rurales. Por lo que

respecta a este tipo de estación, se observa una tendencia constante en Vigo y una claramente descendente en el Grao de Castellón. Dicha tendencia se manifiesta abruptamente a partir del año 2012. Sí presenta cierta ciclicidad, pero en espacios de tiempo irregulares, aún así, parecen ser parejos en ambas estaciones.

4.1.8 Material particulado de 10 µg

El material particulado de 10 µg también forma parte de la medición del Índice de Calidad del aire. A diferencia del material particulado de 2.5 µg, presenta una cantidad de datos algo y con muchos menos faltantes.

Urbanos Interior/ PM10

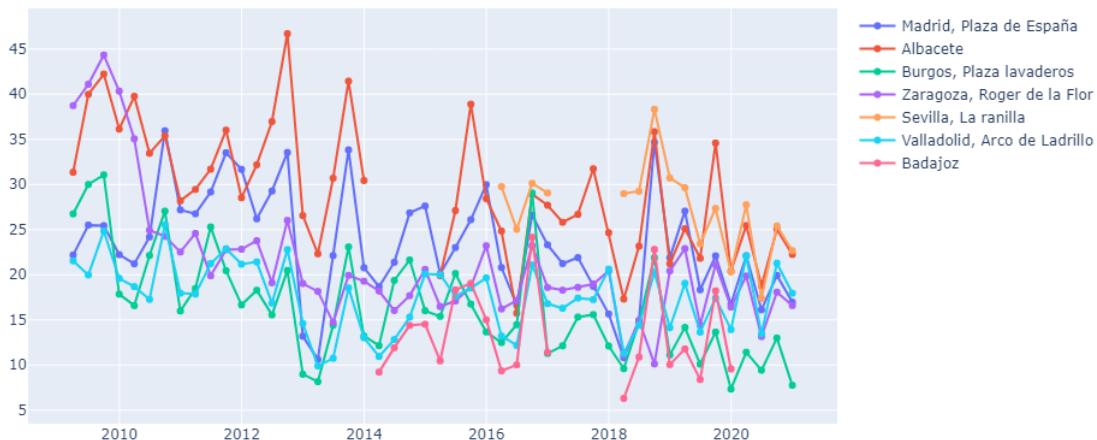


Figura 32. Evolución de la concentración media Urbanos Interior PM10.

La gráfica observada en la Figura 32 muestra un nivel de este contaminante muy uniforme, que se expresa con ciertos picos de período irregular. El movimiento de todas las estaciones es al unísono, indicando que la presencia de este contaminante posiblemente no dependa de la estación en la que se encuentre sino que se contempla a una escala mayor. Afortunadamente, su tendencia parece disminuir. Algo curioso es que la estación con una mayor presencia de estas partículas es Albacete y no Madrid, reforzando la idea de que no depende de la estación, esto se observa tanto en el Evolución de la concentración media (Figura 34) y la tabla con los parámetros muestrales (Figura 33)

	Madrid, Plaza de España PM10	Albacete PM10	Burgos, Plaza lavaderos PM10	Zaragoza, Roger de la Flor PM10	Sevilla, La ranilla PM10	Valladolid, Arco de Ladrillo PM10	Badajoz PM10
count	4319,000	3819,000	4247,000	4279,000	1423,000	4344,000	1778,000
mean	23,313	29,396	16,599	21,301	27,278	17,601	13,389
std	13,414	13,185	10,931	12,236	12,176	10,287	10,587
min	3,167	4,333	1,000	1,935	4,196	2,391	0,000
25%	13,783	20,708	8,522	12,931	20,183	10,129	6,504
50%	20,749	27,458	13,833	18,340	25,967	15,417	10,549
75%	29,771	35,584	22,250	25,887	32,393	22,167	17,017
max	227,833	186,125	178,857	95,455	224,612	116,583	115,806

Figura 33. Parámetros muestrales Urbanos Interior PM10.

Industriales/ PM10

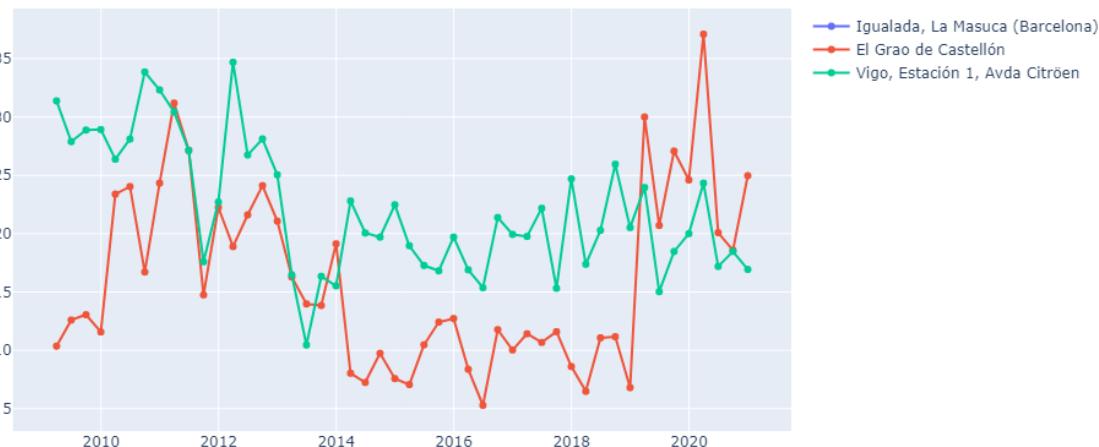


Figura 34. Evolución de la concentración media Industriales PM10.

Volviendo a estaciones industriales, el movimiento es similar al del material particulado de 2.5 µg, a excepción de la subida de los últimos años en el Grao de Castellón. Por lo demás, presenta unas características similares a las observadas en el punto anterior. Cabe señalar que el movimiento es un poco caótico y no parece seguir ningún patrón.

4.1.9 O3

Finalmente, el ozono es un compuesto que también se contabiliza en el ICA y del cual se dispone una gran cantidad de datos, por lo que se pueden examinar con mayor libertad los datos.

Rural/ O3



Figura 35. Evolución de la concentración media Rural O3.

La Figura 35 ilustra el movimiento del promedio trimestral del ozono en la estación rural de Cáceres. Esta refleja un movimiento que se va a observar repetidamente en el caso del ozono, pues marcan una tendencia, por lo general, constante y con unos ciclos muy claros

que van a la inversa de todos los contaminantes previamente estudiados. Es decir, marcan máximos locales en temporada estival y mínimos en épocas invernales.

Industriales/ O3

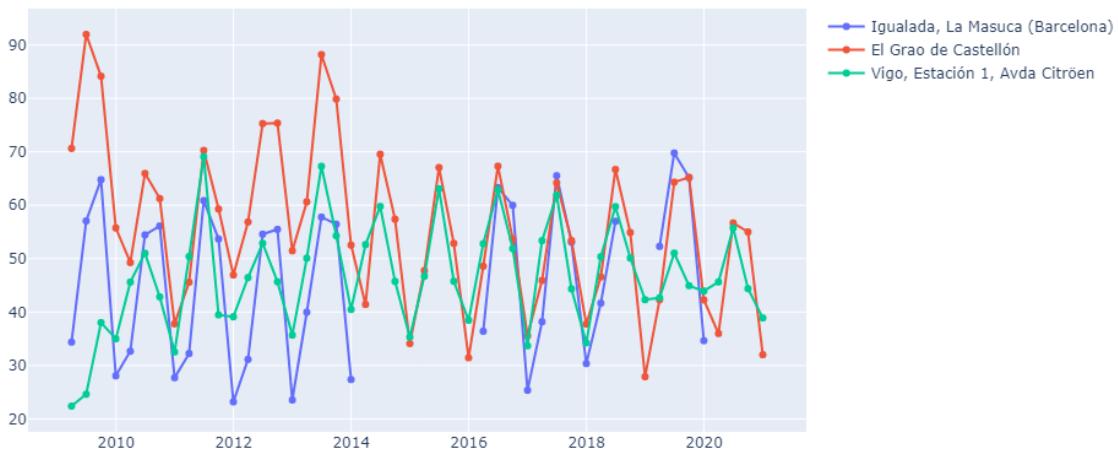


Figura 36. Evolución de la concentración media Industriales O3.

Como se ha mencionado previamente, los movimientos son muy parecidos en todas las estaciones, ya sean industriales (Figura 36), urbanas de interior (Figura 37) o de costa (Figura 38).

Urbanos Interior/ O3

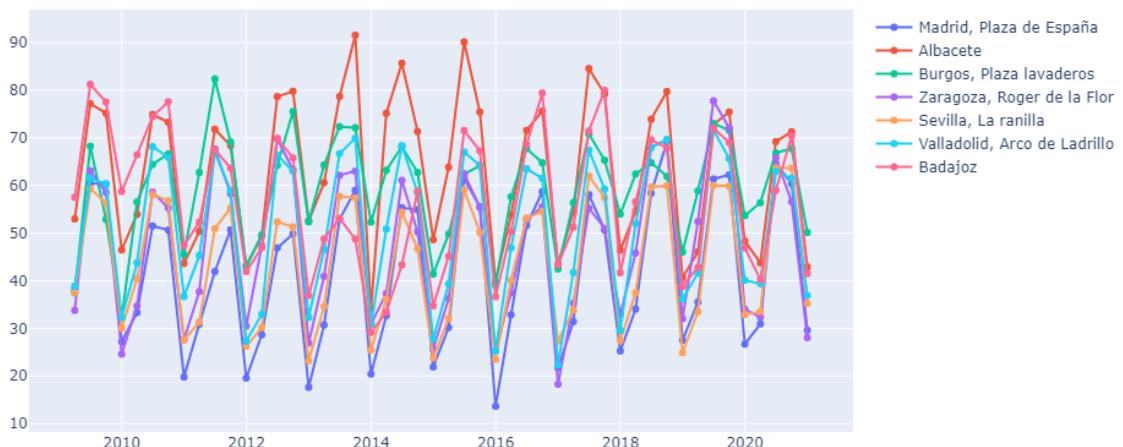


Figura 37. Evolución de la concentración media Urbanos Interior O3.

Los movimientos son claramente uniformes, constantes y periódicos. Curiosamente, este compuesto tampoco depende de la estación en la que se mida, pues ciudades como Albacete, Cádiz y Paterna son las que más presencia contabilizan dentro de sus categorías correspondientes.

Urbanos Costeros/ O3

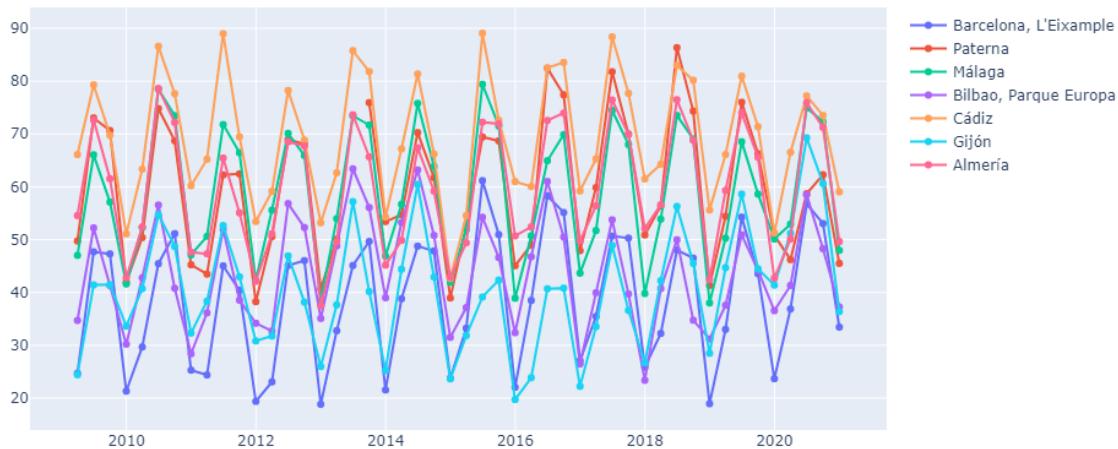


Figura 38. Evolución de la concentración media Urbanos Costeros O3.

4.2 Correlaciones

Para la realización del análisis multivariante y las correlaciones, se han seguido diversos métodos, principalmente se ha realizado el cálculo del coeficiente de correlación lineal de Pearson. Para escalabilidad de este proceso, se han realizado diversas matrices de correlaciones, que muestran, de forma más visual, el grado de correlación entre variables. Para ello, se ha hecho uso de diferentes librerías de *Python*, entre ellas *Seaborn*, *Pandas*, *NumPy* y *Matplotlib*, que han permitido, entre otras cosas, el cálculo y la graficación de las matrices. Además, se han extraído otros ficheros de datos externos para tener más muestras, o poblaciones, para comparar.

Por otro lado, se ha realizado un programa que para el cálculo de las medias anuales de cada compuesto por estación, esto permitirá comparar fácilmente con otros ficheros de datos, puesto que gran parte de ellos se encuentran en un formato temporal anual.

Fecha	Barcelona, L'Eixample/ C6H6	Barcelona, L'Eixample/ CO	Barcelona, L'Eixample/ NO	Barcelona, L'Eixample/ NOx	Barcelona, L'Eixample/ NO2
2009		0,387	29,755	108,527	40,985
2010		0,365	34,638	118,188	41,421
2011		0,712	37,599	123,156	40,517
2012		0,643	37,614	119,048	60,718
2013		1,235	33,337	108,808	56,814
2014		0,583	37,318	109,790	52,408
2015		0,800	43,066	117,975	56,450
2016		0,661	35,761	107,835	52,074
2017		0,544	37,779	115,015	59,364
2018		0,439	34,434	104,864	53,908
2019		0,363	28,455	92,133	49,434
2020		0,271	19,646	62,249	35,250

Figura 39. Tabla de medias anuales de compuestos por estación. Fuente: Elaboración propia a partir de los datos

En primer lugar, se han planteado diferentes hipótesis de posibles correlaciones, siendo algunas de ellas la población del municipio o provincia, el parque de vehículos de la zona, la distancia y altura sobre el mar, y la presencia de unos compuestos, que podrían provocar el aumento de otros.

A continuación, se evalúan las correlaciones recientemente mencionadas y se confirman o, en su defecto, se rechazan dichas hipótesis.

4.2.1 Correlación entre compuestos

En primer lugar, se considera la posible relación entre los propios compuestos. Esto es que para unos valores determinados para el nivel de un compuesto, se presenten otros valores para el nivel de otro compuesto, y esta relación puede ser tanto directa, como inversa, así como fuerte o débil.

Para comparar todos los compuestos, se ha realizado una matriz de correlaciones lineales para cada estación, mostrando la correlación entre pares de compuestos. Los resultados son diversos como se pueden observar en las Figuras 40 y 41.

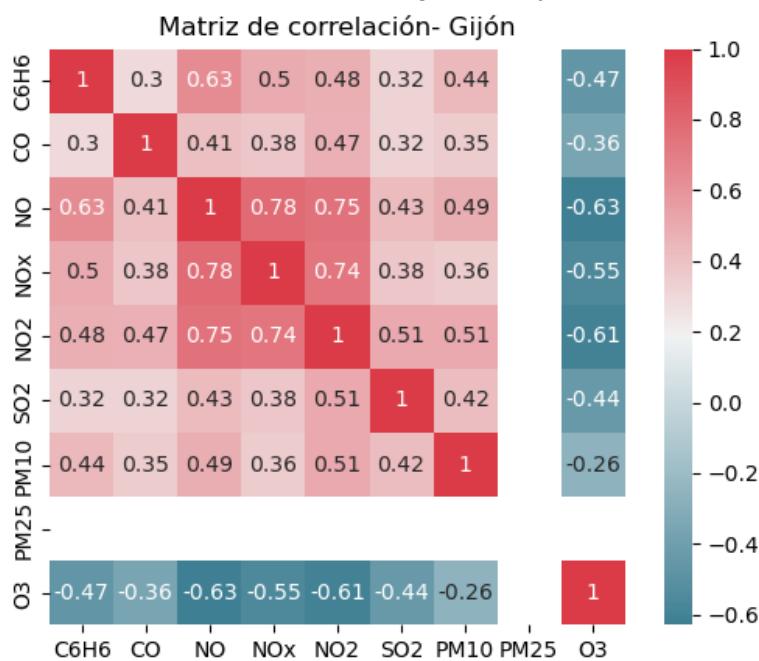


Figura 40. Correlaciones entre compuestos - Gijón.

Como se observa en la Figura 40, la correlación puede llegar a ser media-alta entre los diferentes compuestos (a excepción de PM2.5, en este caso por ausencia de datos). Esto induce a pensar que si hay presencia de un compuesto a niveles elevados, los niveles de otros compuestos también será elevada.

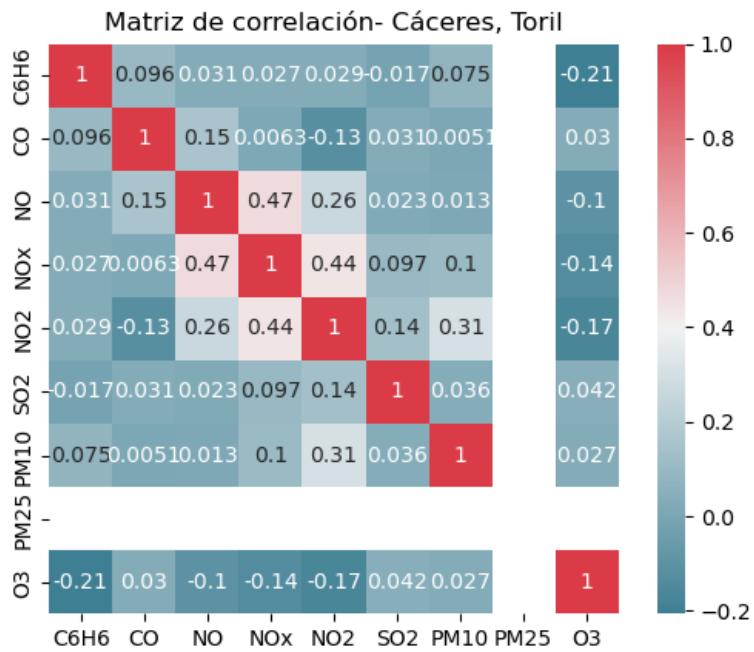


Figura 41. Correlaciones entre compuestos - Cáceres.

Por otro lado, en la Figura 41 se observa que la correlación es prácticamente nula entre los compuestos (valores entre -0.2 y 0.2). Aún así, parece haber correlación entre NO, NOx y NO2, cosa que parece lógica pues la combinación entre NO y NO2 es el NOx.

En ambas Figuras se puede observar que la correlación entre los diferentes compuestos y O3 es inversa, posiblemente indicando que a mayor presencia media de todos los compuestos, el nivel de O3 decrece. Para tratar de analizarlo, se han realizado otras gráficas tratando de ver cómo evolucionan los compuestos a lo largo del tiempo en una misma estación.

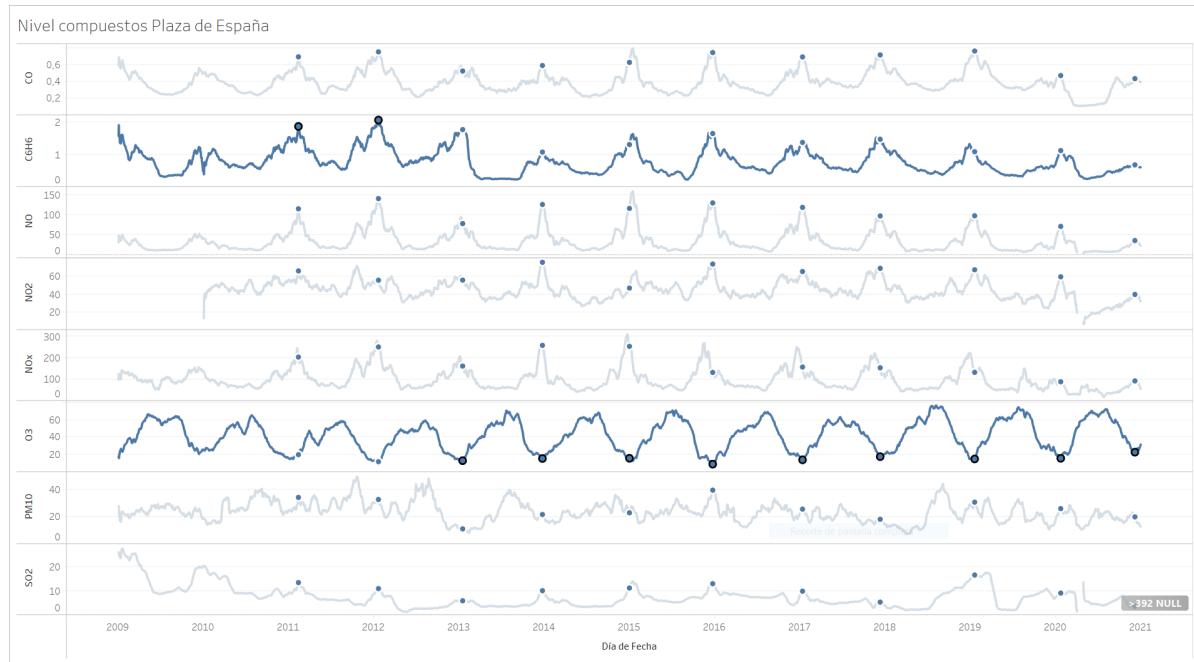


Figura 42. Nivel compuestos Plaza de España, Madrid.

En la Figura 46, los días de menor nivel de O₃, los niveles de los otros componentes marcan máximos locales. Esto no solo ocurre en esta estación pues en las otras estaciones también ocurre los mismos. Además, estas zonas de mínimos de ozono suelen darse a finales y comienzo de año, coincidiendo con los meses invernales. Esta situación puede darse debido a las emisiones generadas por la calefacción y la bajada de las temperaturas, entre otras cosas.

Para comparar la correlación media entre los compuestos (Figura 43), se ha graficado una matriz de correlaciones que promedia todas las correlaciones para mostrar de qué forma se correlacionan. El resultado muestra que las principales relaciones son la relación inversa del ozono con el resto de compuestos, visible en la última fila y columna de la figura siguiente y la relación directa y medianamente alta entre los compuestos nitrosos y el dióxido de azufre, dando valores superiores al 0,5. Esto indica que entre estos últimos 4 compuestos existe una fuerte relación, que se consolida mediante una presencia conjunta que crece y decrece de forma similar. Por otro lado, parece que el material particulado también tiene una correlación considerable con el resto de compuestos -a excepción del CO-. ya que presenta coeficientes del rango medio-alto. Por lo que respecta al resto de compuestos, no hay indicios de una muestra más relevante.

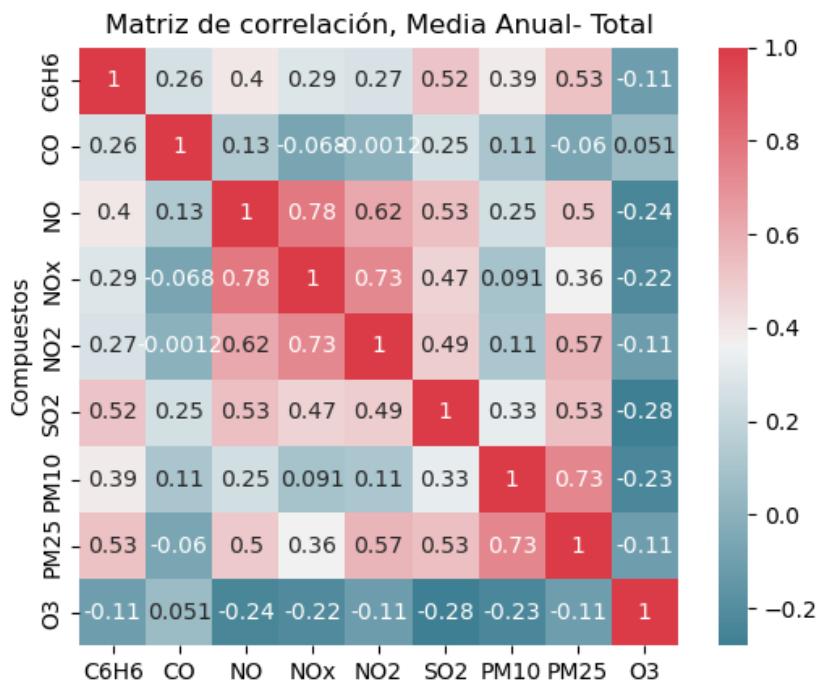


Figura 43. Matriz de correlación promediada por compuesto.

4.2.2 Población y Tráfico

Se plantea la posibilidad de que haya relación entre la población del municipio o provincia en la que se encuentra cada una de las estaciones, el tráfico de la provincia y las mediciones de las propias estaciones. Para calcular ambas correlaciones, se han extraído los datos correspondientes a la población de los municipios en los que se encuentran las estaciones a partir de la base de datos del Instituto Nacional de Estadística bajo el título “Cifras oficiales de población de los municipios españoles en aplicación de la Ley de Bases del Régimen Local (Art.17)”. Por otro lado, los datos del Parque Nacional de Vehículos se han extraído de las bases de datos del Ministerio de Fomento, bajo el título de “Parque

nacional de vehículos por comunidad autónoma, provincia, tipo de vehículo y tipo de carburante”.

Dado que solo se necesitan datos correspondientes a las estaciones del fichero de datos final, se han seleccionado a mano, tanto de los ficheros de población como el del Parque Nacional de Vehículos, y se han bajado los correspondientes archivos. Esto fichero se han

Fecha	Barcelona, L	Paterna	Málaga	Bilbao, Parq	Cádiz	Gijón	Almería	Palma	Tenerife	Cáceres, Toi
2009	1621537	814208	568305	354860	126766	277554	188810	401270	222417	179
2010	1619337	809267	568507	353187	125826	277198	190013	404681	222643	178
2011	1615448	798033	568030	352700	124892	277559	190349	405318	222271	176
2012	1620943	797028	567433	351629	123948	277733	191443	407648	206965	183
2013	1611822	792303	568479	349356	122990	275274	192697	398162	206593	189
2014	1602386	786424	566913	346574	121739	275735	193351	399093	205279	167
2015	1604555	786189	569130	345141	120468	274290	194203	400578	203811	170
2016	1608746	790201	569009	345122	118919	273422	194515	402949	203585	169
2017	1620809	787808	569002	345110	118048	272365	195389	406492	203692	170
2018	1620343	791413	571026	345821	116979	271843	196851	409661	204856	164
2019	1636762	794288	574654	346843	116027	271780	198533	416065	207312	148
2020	1664182	800215	578460	350184	115439	271717	201322	422587	209194	153

reorganizado y estructurado por años -los mismos que los estudiados-, dando lugar a los ficheros que se puedes observar a continuación. El fichero correspondiente a la población se comparte bajo el nombre “Poblacion.xlsx”, el del parque de vehículo lo hace bajo el nombre “TRAFICO.xlsx”.

Figura 44. Tabla de Población por municipio - “Poblacion.xlsx”.

A continuación, se ha desarrollado un código de *Python* que se ha guardado en el fichero “Correlaciones_Trafico_Poblacion.xlsx”. Este se encarga de la realización de una matriz de correlación en base a las medias anuales de cada compuesto, la población de cada municipio y el parque nacional de vehículos.

```

1 file = "DatosFinales.xlsx"
2 inhabitants = "Poblacion.xlsx"
3 traffic = "TRAFICO.xlsx"
4 sheet = "Completo"
5
6 df = pd.read_excel(file, sheet_name=sheet)
7 df_inhabitants = pd.read_excel(inhabitants)
8 df_traffic = pd.read_excel(traffic)

1 dfMA=df.groupby(pd.Grouper(key='Fecha', freq='Y')).mean()

1 places=["Barcelona, L'Eixample", "Paterna", "Málaga", "Bilbao, Parque Europa", "Cádiz", "Gijón", "Almería", "Palma", "Tenerife"]
2 chems=["C6H6", "CO", "NO", "NOx", "NO2", "SO2", "PM10", "PM25", "O3"]
< >

1 #Printing the correlation matrixes
2
3 for place in places:
4     df_aux=pd.DataFrame()
5     for chem in chems:
6         df_aux[place+ "/" +chem]=df[place+ "/" +chem]
7
8     df_aux2 = pd.concat([df_aux, df_inhabitants[place]], axis=1)
9     df_aux2 = df_aux2.rename(columns={place: 'Población ' + place})
10    df_aux2 = pd.concat([df_aux2, df_traffic[place]], axis=1)
11    df_aux2 = df_aux2.rename(columns={place: 'Tráfico ' + place})
12
13    corr_matrix=df_aux2.corr()
14    cmap = sns.diverging_palette(220, 10, as_cmap=True)
15    sns.heatmap(corr_matrix, annot=True, cmap=cmap, square=True)
16    plt.title(f'Matriz de correlación Población - Estación {place}')
17    plt.show()
```

Figura 45. Código para las matrices de correlaciones de Tráfico y Población.

El resultado de este programa son las matrices de correlaciones que nos han permitido comparar las correlaciones lineales entre las diferentes variables. Obviamente, se han comparado los pares de una misma estación, pues la comparación de datos entre diferentes subconjuntos de datos no es procedente. Las matrices resultantes presentan gran variedad

de valores para las correlaciones, mostrando que no es un factor totalmente determinante en todas las estaciones, pero aún así, sí llega a afectar en algunas estaciones y de forma relativamente considerable, tal y como vemos en las siguientes figuras.

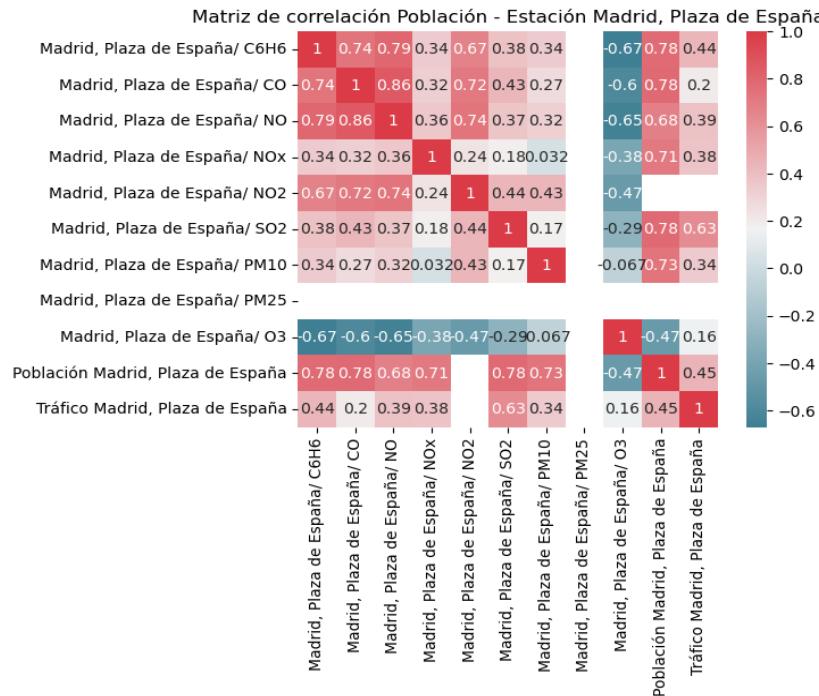


Figura 46. Matriz de Correlaciones de Tráfico y Población en Plaza de España..

Por ejemplo, observando la Figura 46, se contempla que la correlación de los compuestos con la población es relativamente alta y directa, a su vez, se correlaciona inversamente con O₃, tal y como hemos mencionado en la sección “Correlación entre compuestos”. Omitiendo esa información, se contempla que a mayor población y tráfico -especialmente población- hay mayor presencia de los compuestos.

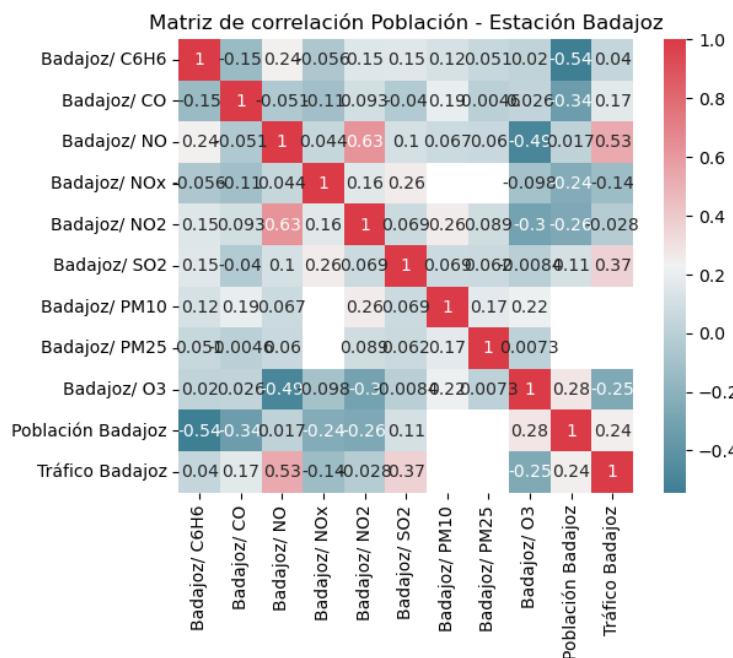
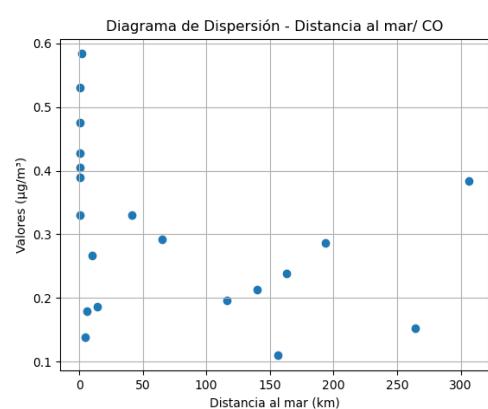
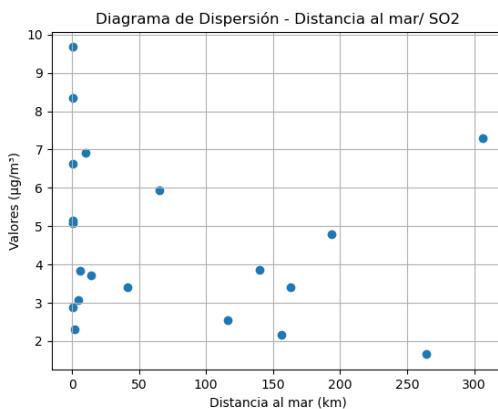


Figura 47. Matriz de Correlaciones de Tráfico y Población en Badajoz..

4.2.3 Distancia al mar y Altitud

Se considera la posibilidad de que exista relación entre la distancia al mar y la altitud de la estación seleccionada con los valores medidos. Para calcular ambas correlaciones, se han extraído los datos geográficos de cada estación mediante el uso de herramientas como Google Earth. Puesto que los datos son valores constantes, no pueden ser usados para calcular un coeficiente de correlación. Por lo tanto, para comprobar si existen las correlaciones se han de realizar gráficos de dispersión que comparan en cada compuesto si existen las relaciones estudiadas. Con este fin, se ha desarrollado un código en *Python* guardado en el notebook “dispersion.ipynb” que mediante el uso de las librerías *pandas* y *matplotlib*, crea los gráficos de dispersión necesarios para comprobar si existe relación.



Figuras 48 y 49. Gráficos de dispersión - Distancia al mar SO₂ y CO.

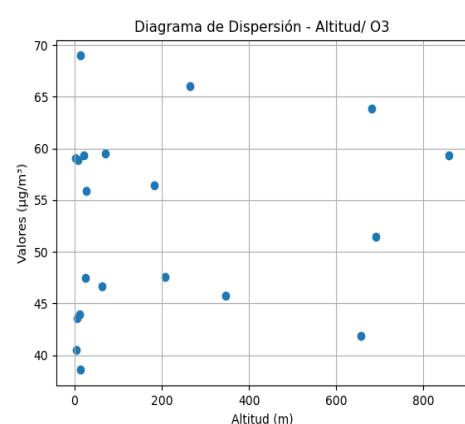
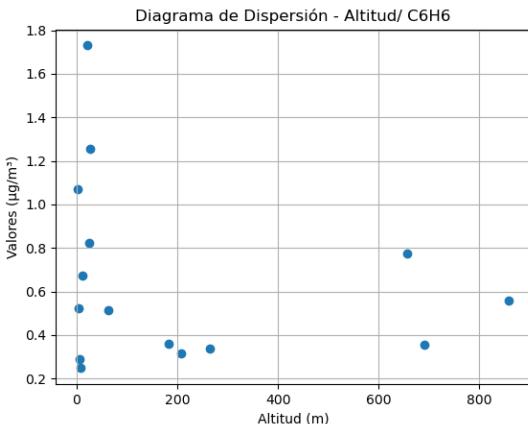


Figura 50 y 51. Gráficos de dispersión - Altura sobre el nivel del mar C₆H₆ y O₃.

Analizando los gráficos, no parece que exista ningún tipo de relación entre la distancia al mar o la altitud con los valores de los compuestos contaminantes. Este resultado se puede deber a qué la mayoría de estaciones escogidas se encuentran cerca del litoral, por lo que su distancia al mar y altura son muy bajas. Sin embargo, no parece haber relación alguna, rechazando la hipótesis de que la altura y la distancia al mar influyen en la presencia de compuestos contaminantes.

4.2.4 Días de la semana y mes del año

Por último, se ha realizado una medición del nivel de cada contaminante por días y meses del año y se ha graficado mediante un mapa de calor, cuyas filas corresponden al mes o día

de la semana y las columnas a diferentes estaciones de medición. Dichas gráficas se han realizado mediante *Tableau* y permiten analizar si existe una correlación entre los días de la semana y los meses del año con el nivel de los contaminantes.

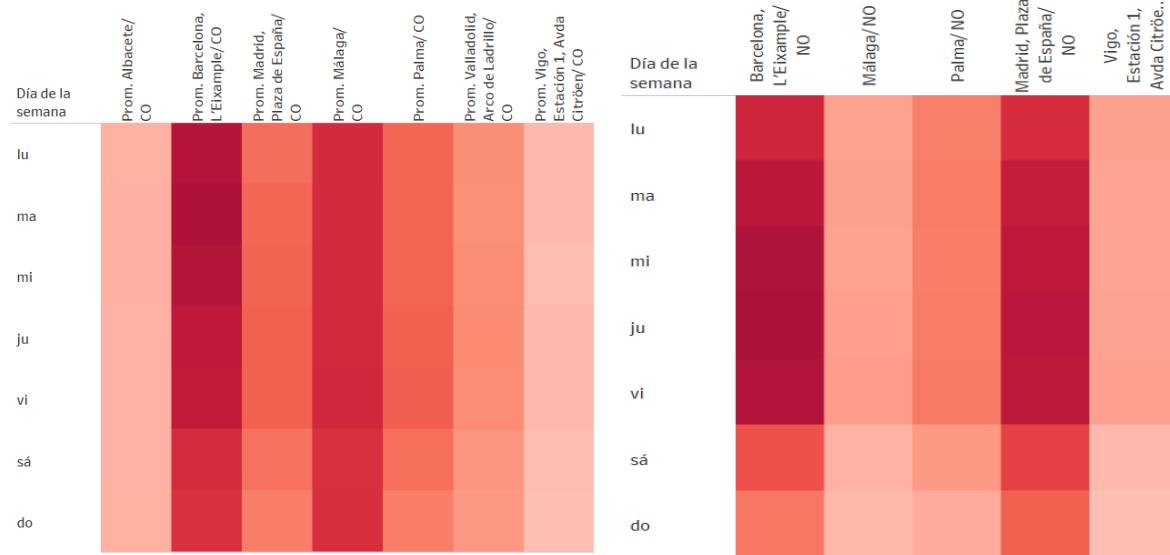


Figura 52 y 53. Mapa de calor - día de la semana CO y NO.

De estas gráficas se ha podido extraer que sí existe cierta relación con el día de la semana y el nivel de los contaminantes, pues si se contemplan detenidamente las gráficas, se observa que los últimos días laborales de la semana son aquellos con un mayor nivel promedio de dichos contaminantes, especialmente jueves y viernes.

Aunque esta situación ocurre con todos los compuestos, el ozono es el único que no lo cumple. Anteriormente, se ha mencionado que sigue una relación inversa entre el ozono y el resto de contaminantes y se expresa de igual manera a lo largo de las semanas. De esta forma, los fines de semana son los días cuya presencia de ozono es mayor, en comparación con el resto de la semana.

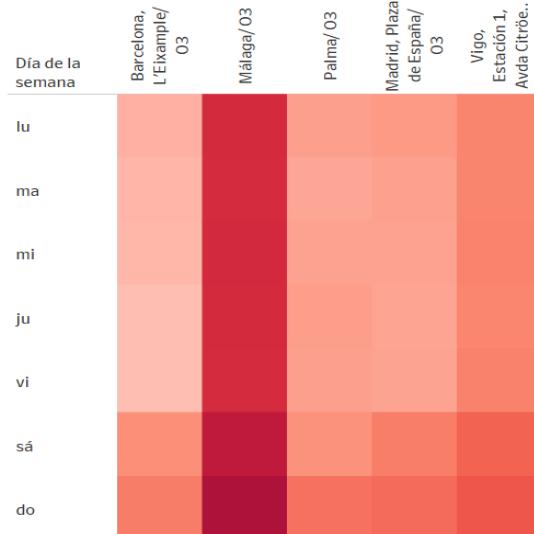


Figura 54. Mapa de calor - día de la semana O3.

En el caso de los meses del año, también se percibe cierta relación. Como se observa en las Figuras 55 y 56, los meses con los niveles más altos de contaminantes son los

invernales, como se ha podido observar. Por otro lado, el ozono también decrece durante dichos meses y aumenta en la época estival.

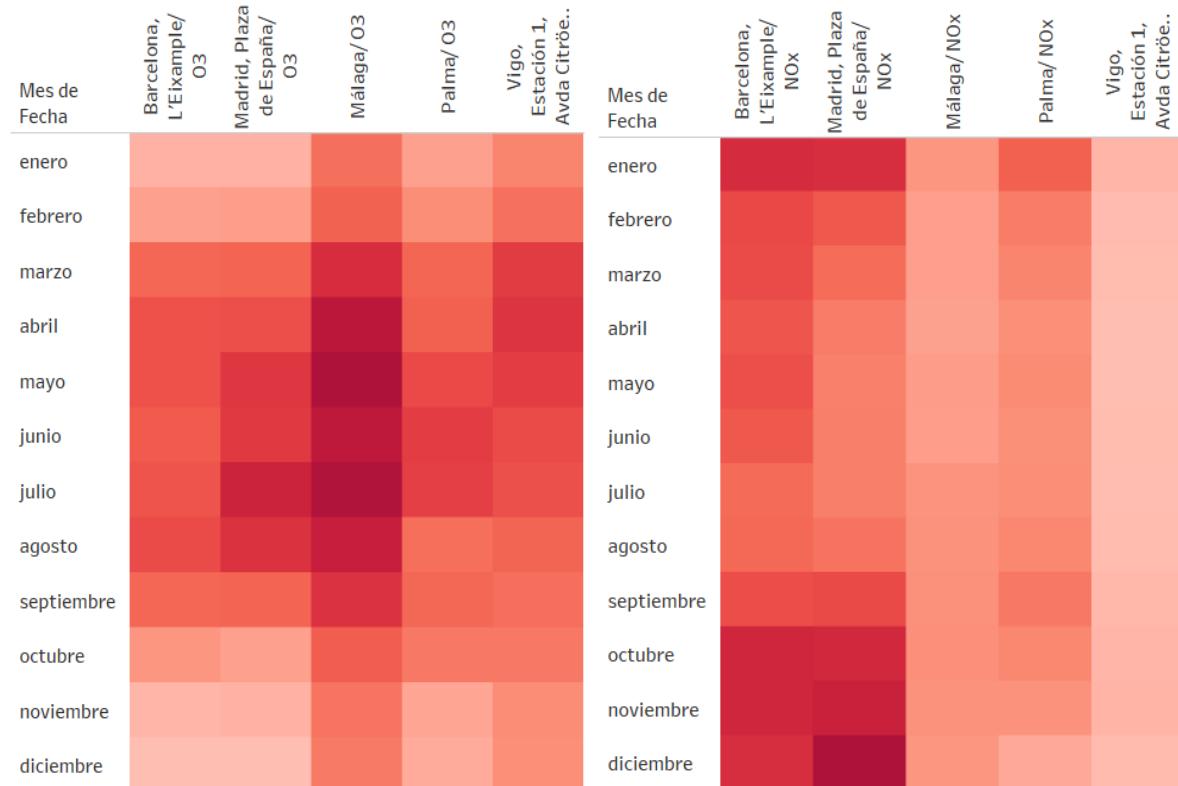


Figura 55 y 56. Mapa de calor - mes del año O3 y NOx.

4.2.5 Nivel del ICA

Con el fin de averiguar qué estaciones sobrepasan los niveles límite del ICA más frecuentemente, se ha programado un código que cuenta cada día en el que el nivel del ICA de la estación es superior a 3, es decir, que es desfavorable, muy desfavorable o extremadamente desfavorable.

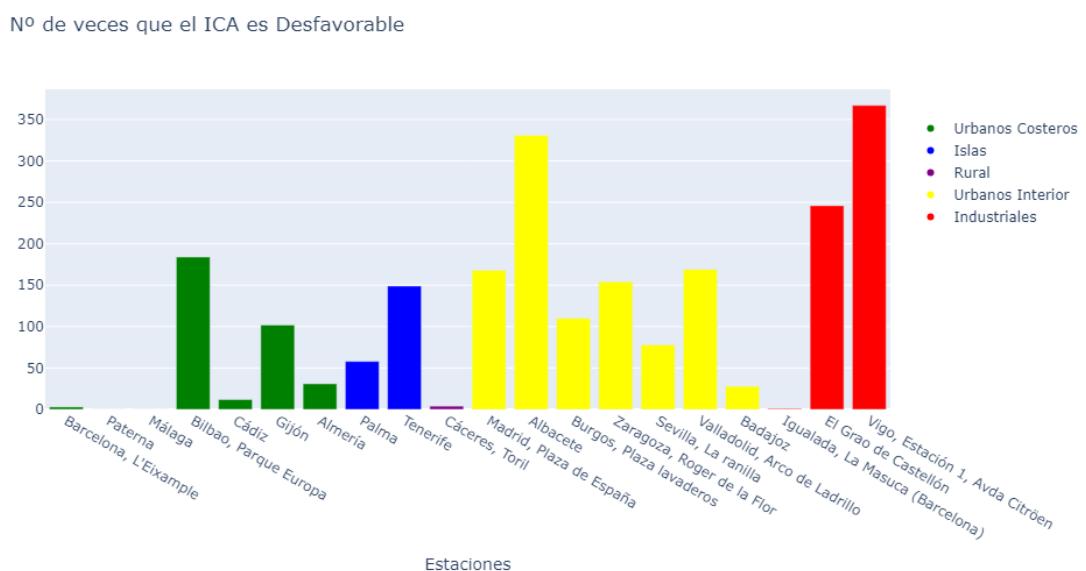


Figura 57. Gráfico de Barras de las veces que se supera el Nivel Desfavorable de ICA.

Analizando el gráfico, se aprecia a simple vista que las estaciones urbanas de interior e industriales destacan sobre el resto. El resultado de Albacete es sorprendente, su ICA es desfavorable o peor el doble de días que Madrid u otras estaciones con mucha más población y actividad industrial. Las estaciones industriales de Vigo y Castellón superan muy frecuentemente los niveles desfavorables de ICA.

Por otra parte, se contempla que las ciudades cercanas al mar, como Barcelona, Paterna y Málaga, rara vez superan los límites establecidos del ICA. Sorprenden Bilbao, Gijón y Tenerife, con resultados parecidos a las estaciones de interior. En Toril casi nunca la calidad del aire es desfavorable.

4.2.6 Regresión lineal y predicción

En primer lugar, se han realizado las gráficas de regresión lineal de cada variable. A cada estación le corresponden 9 gráficas, una para cada compuesto, por lo que en total hay 180 gráficas. Las gráficas mencionadas se han realizado haciendo uso de las librerías *matplotlib* y *sklearn*. Además, se han calculado las pendientes de la recta de regresión (P) y la intersección con el eje Y (I). De la totalidad de las gráficas, se han seleccionado aquellas con más información de cada uno de los compuestos involucrados en el cálculo del ICA y se han comentado.

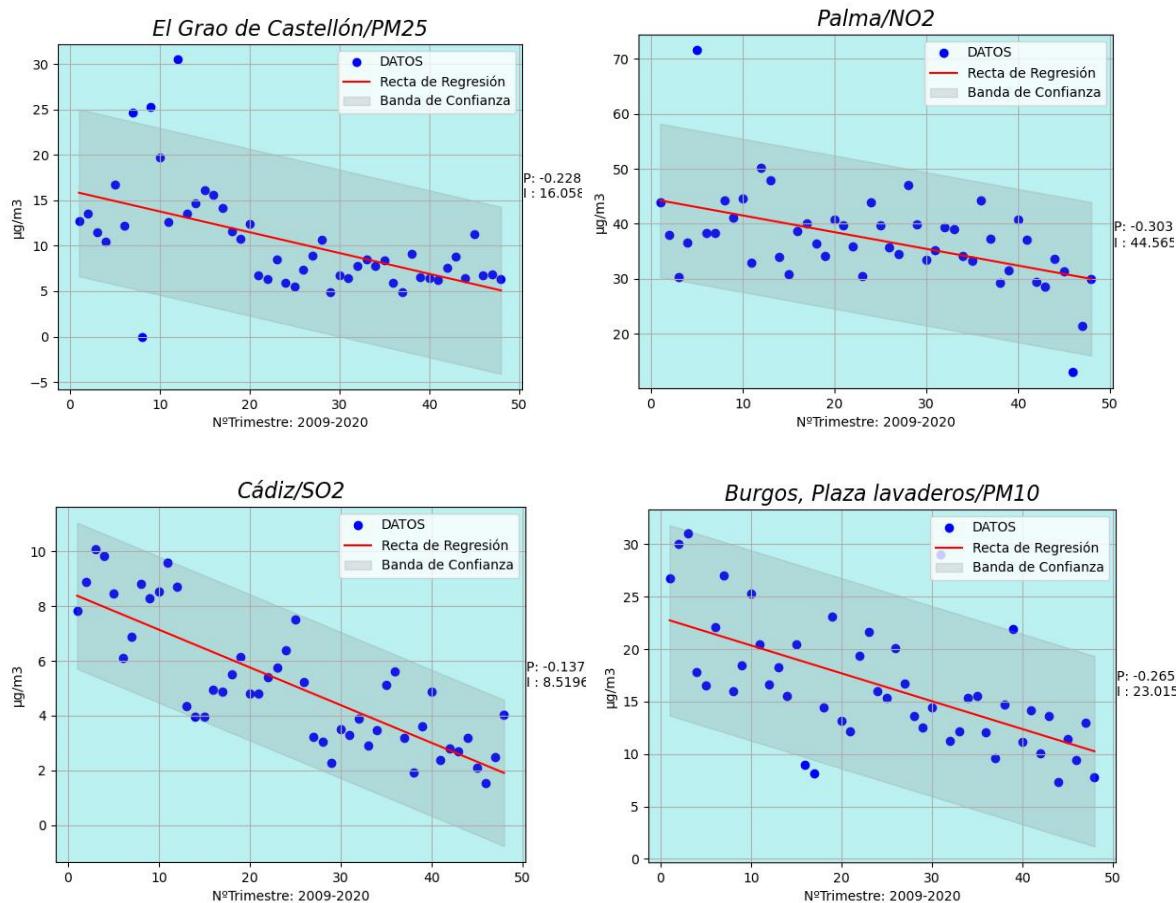


Figura 58-61. Regresión lineal de PM2.5, NO2, SO2 y PM10

Al analizar las 4 gráficas y apoyándonos en la pendiente de sus respectivas rectas, se observa que todas presentan una clara tendencia descendente. Lo que indica que, con el paso del tiempo, la concentración de los compuestos en cada una de las estaciones

disminuirá, siendo una disminución más suave en Palma y Castelló y más notable en Burgos y Cádiz.

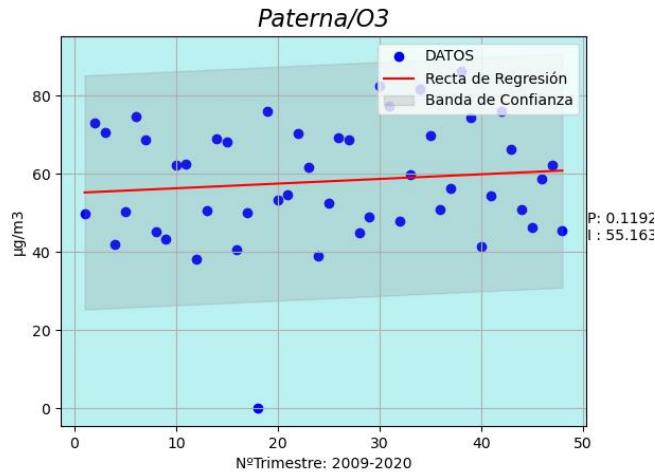


Figura 62. Regresión lineal de O₃

Por último, al observar la gráfica del compuesto O₃ en Paterna, junto al valor de su pendiente, se indica una muy ligera tendencia ascendente. Por ello, con el paso del tiempo, es posible que la concentración del compuesto aumente, aunque sea un aumento ínfimo.

5. Conclusiones

En base a todos los resultados examinados en los puntos anteriores, se ha realizado una discusión y una puesta en común de las observaciones, que se van a exponer resumidamente a continuación.

- En primer lugar, respecto al movimiento de la mayoría de los contaminantes, la tendencia está siendo descendente, lo que desemboca en cierta mejora de la calidad del aire. Esto indica que las mejoras que se están realizando están teniendo resultado. Al contrastar con las medias anuales, el movimiento es el mismo (Figura 63).

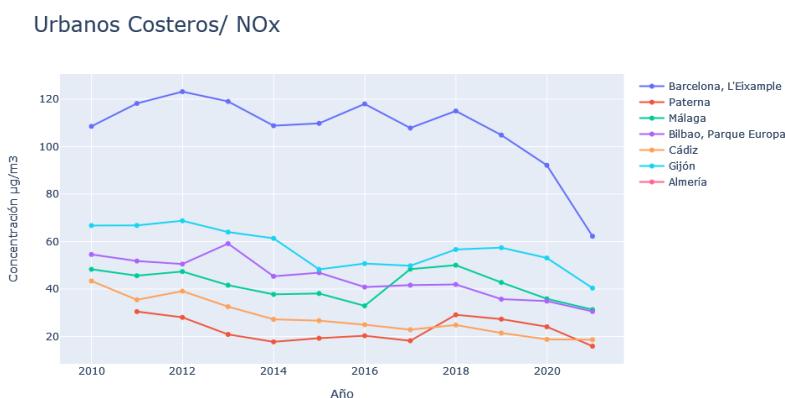


Figura 63. Evolución de la concentración media anual - Urbanos Costeros NOx.

- Respecto a la **ciclicidad de los datos** y con apoyo de la lectura de las noticias: podemos apreciar que las festividades anuales influyen considerablemente en los niveles de concentración de los contaminantes. Por ejemplo, el uso de fuegos artificiales en fechas señaladas [17] causa un exceso de emisiones de PM2.5,

equiparable a las emisiones de este compuesto producidas por dos meses de tráfico. En consecuencia, países como Alemania y EEUU han decidido restringir el uso de fuegos artificiales en sus festividades. Otro ejemplo de este fenómeno es el consumo excesivo de energía durante la Navidad [18]. Históricamente, el abuso de luces navideñas ha generado emisiones de contaminantes excesivas, además de incrementar la contaminación lumínica. Sin embargo, en los últimos años la contaminación atmosférica relacionada con la Navidad ha decrecido notablemente gracias a la imposición de las luces LED, aunque sigue siendo un problema.

- Por lo que respecta a las **correlaciones** analizadas, se detecta que existen muchos factores que influyen en la contaminación del aire. Por ese motivo, los factores elegidos no logran explicar de forma completa el nivel de contaminación, esto se ve reflejado en que las correlaciones no presentan una intensidad muy elevada. En primer lugar, **la presencia de determinados compuestos sí que se correlaciona con la presencia de otros compuestos**. Esto explica que, si hay contaminación del aire, esta es debido a una presencia elevada de varios compuestos, no de un único compuesto. Es decir, si en una estación hay una presencia elevada de moléculas particuladas y provocan que el ICA se eleve a niveles regulares, probablemente también hay una presencia elevada de otros compuestos como pueden ser los óxidos de nitrógeno.
- Por otro lado, **la población y el tráfico de la estación a analizar sí influye**, en cierta medida, en la contaminación, pero no son los únicos factores que lo hacen. Al analizar las correlaciones, existen algunas estaciones que presentan un coeficiente con una intensidad media-alta con el nivel de contaminación, esto significa que en algunas estaciones determinadas, la contaminación se explica, en cierta medida, por la población y el total de tráfico. Sin embargo, algunas estaciones presentan un coeficiente de correlación mucho menor, esto puede indicar que ambos factores no son especialmente determinantes en tales lugares, dando cabida a que en estas estaciones haya otros factores más influyentes en el nivel de contaminación del aire.
- En tercer lugar, **la altura sobre el nivel del mar y la distancia al mar** -u océano, en su defecto-, no son factores que se relacionan mucho con la contaminación. Las correlaciones observadas son prácticamente nulas, descartando la hipótesis de que ambos factores influyen en el nivel de contaminación.
- Seguidamente, se analiza la distribución de la contaminación en función **del día de la semana y el mes del año**, tratando de observar en qué días y meses la contaminación es más elevada. Para ello, se han empleado los mapas de calor observados en las Figuras 57-61. Se extrae de ellos que existen días y meses concretos en que el nivel es más elevado, esto también se corresponde a la ciclicidad de las concentraciones medias, pues ya se ha mencionado que, en meses invernales, todos los compuestos menos el ozono presentan concentraciones más elevadas. Se confirma, de nuevo, al observar los gráficos correspondientes a los meses, donde se contempla un promedio más elevado. En cuanto a los días de la semana con mayor contaminación, las concentraciones son más elevadas los días más cercanos a los fines de semana, sin contar sábado y domingo. Esta situación posiblemente se corresponda a los movimientos realizados por vehículos para ir al trabajo. Asimismo, el ozono se manifiesta con una mayor concentración los fines de semana, es decir, sábado y domingo, reafirmando de nuevo la correlación inversa entre los compuestos y el ozono.

- Los **niveles de ICA** tienen valores desfavorables, muy desfavorables o extremadamente desfavorables con más frecuencia en las estaciones industriales o urbanas de interior. Esto se debe a que la actividad industrial emana una gran cantidad de compuestos contaminantes al aire, y que las ciudades de interior no se ven beneficiadas por las corrientes de aire provenientes del mar para deshacerse de los contaminantes. Las estaciones cercanas al mar y las rurales presentan niveles desfavorables de calidad del aire con mucha menos frecuencia. Además, se contabiliza que, en total, se ha estado expuesto a niveles de calidad del aire dañinos para la salud alrededor de
- Finalmente, las líneas de **regresión** expresan valores negativos para la pendiente de las rectas, reafirmando la conclusión de que los niveles de concentración de los diferentes compuestos están disminuyendo. Como hemos visto antes, la mayoría de compuestos es proporcionalmente inversa al O₃, y así lo muestra también esa recta de regresión positiva.

Lista de figuras

- Figura 1. Tabla de valores umbral del Índice de Calidad del Aire
Figura 2. Fragmento del programa de filtrado de datos
Figura 3. Fragmento del programa de creación de visualizaciones
Figura 4. Fragmento del programa de medición del ICA
Figura 5. Diagrama de Gantt
Figura 6. Tabla de estaciones
Figura 7. Evolución de la concentración media Urbanos Costeros CO
Figura 8. Parámetros muestrales Urbanos Costeros CO
Figura 9. Evolución de la concentración media Urbanos Interior CO
Figura 10. Parámetros muestrales Urbanos interior CO
Figura 11. Parámetros muestrales Industriales CO
Figura 12. Evolución de la concentración media Urbanos Costeros NO
Figura 13. Evolución de la concentración media Urbanos Interior NO
Figura 14. Parámetros muestrales Urbano Costeros NO
Figura 15. Evolución de la concentración media Industriales NO
Figura 16. Evolución de la concentración media Urbanos Costeros NOx
Figura 17. Evolución de la concentración media Islas NOx
Figura 18. Parámetros muestrales Urbanos Costeros NOx
Figura 19. Evolución de la concentración media Rural NOx
Figura 20. Evolución de la concentración media Urbanos Costeros C6H6
Figura 21. Evolución de la concentración media Urbanos Interior C6H6
Figura 22. Evolución de la concentración media Urbanos Costeros SO2
Figura 23. Evolución de la concentración media Rural SO2
Figura 24. Evolución de la concentración media Urbanos Interior SO2
Figura 25. Evolución de la concentración media Industriales SO2
Figura 26. Evolución de la concentración media Urbanos Costeros NO2
Figura 27. Evolución de la concentración media Islas NO2
Figura 28. Evolución de la concentración media Rural NO2
Figura 29. Evolución de la concentración media Urbanos Interior NO2
Figura 30. Evolución de la concentración media Urbanos Interior PM2.5
Figura 31. Evolución de la concentración media Industriales PM2.5
Figura 32. Evolución de la concentración media Urbanos Interior PM10
Figura 33. Parámetros muestrales Urbanos Interior PM10
Figura 34. Evolución de la concentración media Industriales PM10
Figura 35. Evolución de la concentración media Rural O3
Figura 36. Evolución de la concentración media Industriales O3
Figura 37. Evolución de la concentración media Urbanos Interior O3
Figura 38. Evolución de la concentración media Urbanos Costeros O3
Figura 39. Tabla de medias anuales de compuestos por estación
Figura 40. Correlaciones entre compuestos - Gijón
Figura 41. Correlaciones entre compuestos - Cáceres
Figura 42. Nivel compuestos Plaza de España, Madrid
Figura 43. Matriz de correlación promediada por compuesto
Figura 44. Tabla de Población por municipio - "Poblacion.xlsx"
Figura 45. Código para las matrices de correlaciones de Tráfico y Población
Figura 46. Matriz de Correlaciones de Tráfico y Población en Plaza de España

- Figura 47. Matriz de Correlaciones de Tráfico y Población en Badajoz
Figuras 48 y 49. Gráficos de dispersión - Distancia al mar SO₂ y CO
Figura 50 y 51. Gráficos de dispersión - Altura sobre el nivel del mar C₆H₆ y O₃
Figura 52 y 53. Mapa de calor - día de la semana CO y NO
Figura 54. Mapa de calor - día de la semana O₃
Figura 55 y 56. Mapa de calor - mes del año O₃ y NO_x
Figura 57. Gráfico de Barras de las veces que se supera el Nivel Desfavorable de ICA
Figura 58-61. Regresión lineal de PM_{2.5}, NO₂, SO₂ y PM₁₀
Figura 62. Regresión lineal de O₃

Bibliografía

[1] Ley 34/2007, de 15 de noviembre, de calidad del aire y protección de la atmósfera. *BOE*, núm. 275, de 16/11/2007, pp. 46962 a 46987.

<https://www.boe.es/boe/dias/2007/11/16/pdfs/A46962-46987.pdf>

[2] Real Decreto 102/2011, de 28 de enero, relativo a la mejora de la calidad del aire. *BOE*, núm. 25, de 29/01/2011.

<https://www.boe.es/buscar/pdf/2011/BOE-A-2011-1645-consolidado.pdf>

[3] Directiva 2008/50/CE del Parlamento Europeo y del Consejo, de 21 de mayo de 2008, relativa a la calidad del aire ambiente y a una atmósfera más limpia en Europa. *DOUE* núm. 152, de 11/6/2008, páginas 1 a 44.

<https://www.boe.es/doue/2008/152/L00001-00044.pdf>

[4] Directiva 2004/107/CE del Parlamento Europeo y del Consejo, de 15 de diciembre de 2004, relativa al arsénico, el cadmio, el mercurio, el níquel y los hidrocarburos aromáticos policíclicos en el aire ambiente. *DOUE* núm. 23, 26/1/2005, páginas 3 a 16.

<https://www.boe.es/doue/2005/023/L00003-00016.pdf>

[5] Real Decreto 39/2017, de 27 de enero, por el que se modifica el Real Decreto 102/2011, de 28 de enero, relativo a la mejora de la calidad del aire. *BOE* núm. 24, de 28/1/2017, páginas 6918 a 6930.

<https://www.boe.es/eli/es/rd/2017/01/27/39/dof/spa/pdf>

[6] Real Decreto 34/2023, de 24 de enero, por el que se modifican el Real Decreto 102/2011, de 28 de enero, relativo a la mejora de la calidad del aire; el Reglamento de emisiones industriales y de desarrollo de la Ley 16/2002, de 1 de julio, de prevención y control integrados de la contaminación, aprobado mediante el Real Decreto 815/2013, de 18 de octubre; y el Real Decreto 208/2022, de 22 de marzo, sobre las garantías financieras en materia de residuos. *BOE* núm. 21, de 25/1/2023, páginas 10326 a 10348.

<https://www.boe.es/boe/dias/2023/01/25/pdfs/BOE-A-2023-2026.pdf>

[7] MITECO-Ministerio para la transición ecológica y el reto demográfico, Gobierno de España.

Histórico de la calidad del aire en España. Desde: 22/02/2023.

https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/evaluacion-datos/datos/Historico_calidad_aire.aspx

[8] Descripción de la RVVCCA - Calidad Ambiental - Generalitat Valenciana. (s. f.). Calidad Ambiental.

<https://agroambient.gva.es/es/web/calidad-ambiental/descripcion-de-la-rvvcca>

[9] Resolución de 2 de septiembre de 2020, de la Dirección General de Calidad y Evaluación Ambiental, por la que se modifica el Anexo de la Orden TEC/351/2019, de 18 de marzo, por la que se aprueba el Índice Nacional de Calidad del Aire. *BOE* núm. 242, 10/09/2020, Sec. I., Pág. 75835.

<https://www.boe.es/boe/dias/2020/09/10/pdfs/BOE-A-2020-10426.pdf>

[10] INE - Instituto Nacional de Estadística. Cifras oficiales de población de los municipios españoles en aplicación de la Ley de Bases del Régimen Local (Art.17). 09/03/2023.

<https://www.ine.es/dynt3/inebase/es/index.htm?padre=525>

[11] FOMENTO- Ministerio de transporte, movilidad y agenda urbana, Gobierno de España. Consulta OTLE. Observatorio del Transporte y la Logística en España: Parque nacional de vehículos por comunidad autónoma, provincia, tipo de vehículo y tipo de carburante. Fecha de acceso 09/03/2023.

<https://apps.fomento.gob.es/bdotle/visorBDpop.aspx?i=396>

[12] MITECO-Ministerio para la transición ecológica y el reto demográfico, Gobierno de España. PRTR España.Registro Estatal de Emisiones y Fuentes Contaminantes. Fecha de acceso 09/03/2023.

<https://prtr-es.es/>

[13] M. Planelles & F. Khan, La gran revisión científica sobre la crisis climática: "La ventana para asegurar un futuro habitable se cierra". *El País*. 20 de marzo de 2023.
<https://elpais.com/clima-y-medio-ambiente/2023-03-20/la-gran-revision-cientifica-sobre-la-crisis-climatica-la-ventana-para-asegurar-un-futuro-habitable-se-cierra.html>

[14] J. Serrano Rodriguez, Representación en la calidad del aire y posibles alternativas en la reducción de agentes contaminantes. TFG. Universidad de Valladolid .2020.
<http://uvadoc.uva.es/handle/10324/44629>

[15] A. Albors Trabés, Estudio de la calidad del aire en la provincia de Valencia y establecimiento de sus niveles de fondo. TFM. Universitat Politècnica de València. 2021.

<http://hdl.handle.net/10251/176855>

[16] Diagrama de Gantt
[https://docs.google.com/spreadsheets/d/1o9G8wxIL0jLUe9-JalLtySm2w63Xq5wL/edit?usp=sharing&ouid=10620677781600130772&rtpof=true&sd=true](https://docs.google.com/spreadsheets/d/1o9G8wxIL0jLUe9-JalLtySm2w63Xq5wL/edit?usp=ssharing&ouid=10620677781600130772&rtpof=true&sd=true)

[17] García, T., & García, T. (2021, 13 noviembre). Los fuegos artificiales también calientan el planeta. *El País*.

<https://elpais.com/planeta-futuro/seres-urbanos/2021-11-13/los-fuegos-artificiales-tambien-calientan-el-planeta.html>

[18] Ambiental, A. P. (2022). ¿Cómo afecta la Navidad en el medio ambiente? *Ambientum Portal Lider Medioambiente*.

<https://www.ambientum.com/ambientum/energia/como-afecta-la-navidad-en-el-medio-ambiente.asp>