



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

*Análisis de  
los principales gases  
contaminantes en  
España  
mediante modelos de  
regresión.*

---

**MEMORIA PRESENTADA POR: Zaira Casanova García**

**TUTOR: Jorge Jordán Núñez**

GRADO DE ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

Convocatoria de defensa: Junio – 2017



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

# **ANÁLISIS DE LOS PRINCIPALES GASES CONTAMINANTES EN ESPAÑA MEDIANTE MODELOS DE REGRESIÓN.**



Memoria presentada por: **Zaira Casanova García**  
Tutor: **Jorge Jordán Núñez**  
**Grado de Administración y Dirección de empresas**

Convocatoria de defensa: **junio de 2017**



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

**AGRADECIMIENTOS:**

- A mi director de proyecto, Jorge Jordán por toda su ayuda y facilidades para realizar este análisis.
- A mis padres, por hacer posible que haya podido tener estudios superiores.
- A toda mi familia y mi pareja por siempre tener su apoyo.



## Índice:

1. Introducción y Objetivos: .....	7
Motivación: .....	8
Hipótesis:.....	9
Objetivos: .....	9
Definición de variables: .....	10
2. Metodología y Técnicas:.....	15
Estructura del proceso: .....	16
Preparación de base de datos:.....	18
Análisis univariante: .....	20
Análisis Bivariante: .....	23
Análisis multivariable: .....	25
3. Cálculos, Resultados, y Explotación: .....	29
Análisis Univariable: .....	30
Análisis Bivariante: .....	44
Multicolinealidad: .....	50
Análisis multivariable: .....	72
• Dióxido de azufre (SO <sub>2</sub> ): .....	72
○ Modelo teórico propuesto:.....	72
○ Interpretación de los parámetros y unidades físicas: .....	72
○ Primera estimación del modelo ajustado: .....	73
○ Análisis de significatividad:.....	73
○ Análisis de normalidad: .....	75
▪ Gráfico probabilístico: .....	75
▪ Histograma de residuos: .....	75
▪ Test de normalidad: .....	76
Pruebas de Normalidad para RESIDUOS .....	76
○ Análisis de heterocedasticidad:.....	76
○ Autocorrelación:.....	77
▪ Test de Durbin – Watson:.....	77
▪ Gráfico de Durbin – Watson:.....	78



## CAMPUS D'ALCOI

▪ FAS y FAP:.....	78
○ Puntos influyentes:.....	79
○ Puntos anómalos:.....	80
○ Resolución de problemas:.....	80
• Dióxido de carbono (CO <sub>2</sub> ):.....	81
○ Modelo teórico propuesto:.....	81
○ Interpretación de los parámetros y unidades físicas:.....	81
○ Primera estimación del modelo ajustado:.....	81
○ Análisis de significatividad:.....	81
○ Análisis de normalidad:.....	83
▪ Gráfico probabilístico:.....	83
▪ Histograma de residuos:.....	84
▪ Test de normalidad:.....	84
○ Análisis de heterocedasticidad:.....	85
○ Autocorrelación:.....	86
▪ Test de Durbin – Watson:.....	86
▪ Gráfico de Durbin – Watson:.....	86
▪ FAS y FAP:.....	86
○ Puntos influyentes:.....	87
○ Puntos anómalos:.....	88
○ Resolución de problemas:.....	88
• Óxidos de Nitrógeno (NO <sub>x</sub> ):.....	93
○ Modelo teórico propuesto:.....	93
○ Interpretación de los parámetros y unidades físicas:.....	93
○ Primera estimación del modelo ajustado:.....	93
○ Análisis de significatividad:.....	94
○ Análisis de normalidad:.....	96
▪ Gráfico probabilístico:.....	96
▪ Histograma de residuos:.....	96
▪ Test de normalidad:.....	97
○ Análisis de heterocedasticidad:.....	97



## CAMPUS D'ALCOI

○ Autocorrelación:.....	99
▪ Test de Durbin – Watson:.....	99
▪ Gráfico de Durbin – Watson:.....	99
▪ FAS y FAP:.....	99
○ Puntos influyentes:.....	100
○ Puntos anómalos:.....	101
○ Resolución de problemas:.....	101
4. Predicciones: .....	103
5. Comparación del Dióxido de carbono en las provincias Españolas: .....	110
6. Conclusiones: .....	114
7. Líneas futuras de investigación .....	117
8. Anexos:.....	119
Noticias:.....	120
Variables explicadas:.....	139
Variables explicativas:.....	141
9. Bibliografía: .....	144



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

# 1. Introducción y Objetivos:



## Motivación:

La contaminación del aire es cualquier cambio en sus compuestos naturales por los que la existencia de cualquier elemento que sea perjudicial para el ser humano, animales y plantas. Esta puede producirse a partir de gases, cenizas volcánicas, humo de incendios no provocados, polvo, polen, esporas de plantas, hongos y bacterias, aunque la contaminación más grave es la derivada por las actividades del ser humano en actividades industriales y comerciales, los motores de los vehículos...

Los contaminantes generados por las actividades humanas son:

- Actividades industriales: óxidos de azufre, nitrógeno y plomo metálico.
- Actividades domésticas: óxidos de azufre y nitrógeno.
- Transporte: óxidos de nitrógeno y plomo, y óxidos de azufre. [[EcuRed](#)]

Las consecuencias de estas actividades son muy peligrosas tanto para el ser humano como para el medio ambiente. Para el ser humano las más importante son las enfermedades respiratorias (asma, bronquitis, cáncer pulmonar...) y las dermatológicas (manchas, cáncer de piel...) y a su vez a las plantas les dificulta la realización de la fotosíntesis. [[Causas y consecuencias de...](#)]

Se decide realizar un análisis de los principales gases contaminantes en España mediante modelos de regresión porque se detecta en varias noticias que en la actualidad existe gran preocupación por la calidad del aire. La principal noticia por la cual se piensa que puede ser un buen tema de investigación es la publicada el 12 de enero de 2017 en la que se argumentan las consecuencias de la siguiente medida: *"El 29 de diciembre, Carmena activó por primera vez en España la prohibición de circular a la mitad de las matrículas ante los elevados niveles de NO<sub>2</sub>"*. [[La Razón](#)]





UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

### **Hipòtesis:**

La hipótesis consiste en que se puede modelar los diferentes gases contaminantes de cada una de las provincias españolas, y a partir de estos modelos poder observar cómo afectan las diversas variables explicativas a estos.

### **Objetivos:**

- Investigar cuales son los principales gases contaminantes de las provincias españolas.
- Obtener variables explicativas con las cuales poder modelizar y explicar los gases contaminantes de España.
- Conseguir modelos válidos de los tres gases propuestos: Dióxido de azufre, Dióxido de carbono y Óxidos de nitrógeno.
- Realizar predicciones a partir de un modelo válido con las cuales poder explicar cómo se comportan los diferentes gases, realizando variaciones en las variables explicativas.



## Definición de variables:

A continuación, se expresan las variables definidas para estos modelos:

- **Variables explicadas:**

- **Dióxido de azufre (SO<sub>2</sub>):**

“Gas incoloro, irritante, no inflamable y con un olor penetrante que consiste en un átomo de azufre y dos de oxígeno”. [[Fundación Crana](#)]

El dióxido de azufre se produce en procesos industriales de la quema de combustibles fósiles que contienen azufre a altas temperaturas. Para intentar controlar los niveles de este gas, el gobierno decreta unos valores que no se deben superar los cuales se presentan en la siguiente tabla.

Tabla 1.1. Valores Dióxido de azufre

Valor legislado	Valor límite	Periodo
Valor límite horario (VLH) para la protección de la salud humana	350 µg/m <sup>3</sup>	Valor medio en 1h. No debe superarse en más de 24 ocasiones por año civil.
Valor límite (VLD) para la protección de la salud humana.	125 µg/m <sup>3</sup>	Valor medio en 24h. No debe superarse en más de 3 ocasiones por año civil
Valor límite para la protección del ecosistema.	20 µg/m <sup>3</sup>	Año civil y periodo invernal (1 de octubre del año anterior a 31 de marzo del año en curso)

El dióxido de azufre está regulado porque es muy perjudicial para la salud y para el ecosistema ya que, aunque la actividad industrial esté lejos de la población puede causar muchos problemas en la salud en cuanto a inflamación del sistema respiratorio, insuficiencia pulmonar, dolor de cabeza. Por otra parte, en cuanto al medio ambiente crea daños en la vegetación, disminución de la clorofila, reducción de la fotosíntesis de las plantas y pérdida de varias especies.

Los valores que establece la Organización Mundial de la Salud son:

SO<sub>2</sub>      —————>      20 µg/m<sup>3</sup>, media de 24 horas  
                 —————>      500 µg/m<sup>3</sup>, media de 10 minutos

[[Gobierno de España. Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente](#)]



## CAMPUS D'ALCOI

- **Dióxido de carbono (CO<sub>2</sub>):**

El dióxido de carbono es un gas formado por dos átomos de oxígeno y uno de carbono. Normalmente lo podemos encontrar en la naturaleza en forma de gas, pero cuando la temperatura es baja y hay una presión elevada se vuelve líquido e incluso puede llegar a ser sólido. [[Conciencia cero](#)]

La aparición de este gas es la quema, pero en este caso de sectores no industriales. El ejemplo más claro es el tráfico, ya que se realiza quema de combustible, aunque este cada vez va bajando más porque en los automóviles se utilizan cada vez más los catalizadores para retener estos gases.

El dióxido de carbono es perjudicial para el ser humano ya que provoca que el tiempo que tarda el oxígeno hasta llegar a la sangre sea lento y puede crear el deterioro de algunos órganos, así como creación de cataratas, dolores de cabeza, daños en el sistema nervioso. Todos estos efectos se pueden dar tanto en el ser humano como en la fauna.

También tiene efecto en el clima porque este favorece a la creación del efecto invernadero. [[Gobierno de España. Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente](#)]

- **Óxidos de nitrógeno (NO<sub>x</sub>):**

Los óxidos de nitrógeno (NO<sub>x</sub>) están formados por el monóxido de nitrógeno (NO) y el dióxido de nitrógeno (NO<sub>2</sub>).

Estos gases se originan por combustión y tienen relación con el tráfico, fundamentalmente con los automóviles de motor Diesel. Además, también están relacionados con las actividades industriales a altas temperatura.

En el transcurso de los años los niveles de estos gases han disminuido porque existen menos industrias energéticas y los valores más altos se encuentran en las grandes ciudades y carreteras de tráfico muy denso.

Existen unos valores legislados para el Dióxido de nitrógeno y el Monóxido de nitrógeno:



Tabla 1.2. Valores Óxidos de Nitrógeno.

Valor legislado	Valor límite	Periodo
Valor límite horario (VLH) para la protección de la salud humana.	200 $\mu\text{g}/\text{m}^3$	Valor medio en 1 h No debe superarse en más de 18 ocasiones por año civil
Valor límite anual (VLA) para la protección de la salud humana	40 $\mu\text{g}/\text{m}^3$	Año civil
Valor límite de $\text{NO}_x$ para la protección de los ecosistemas.	30 $\mu\text{g}/\text{m}^3$	Año civil

Los valores que recomienda no sobrepasar la Organización Mundial de la salud son los siguientes:

$\text{NO}_2$        $\longrightarrow$       40  $\mu\text{g}/\text{m}^3$ , media anual  
                  $\longrightarrow$       200  $\mu\text{g}/\text{m}^3$ , media de una hora

[\[Gobierno de España. Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente\]](#)

- **Variables explicativas:**

- **Temperatura (TEMP):**

Variable que indica el valor en grados que se encuentran en la atmósfera. La magnitud con la cual está expresada es: grados centígrados ( $^{\circ}\text{C}$ ).

La temperatura tiene una gran importancia en la calidad el aire porque a más grados centígrados más fácilmente se disiparán los diferentes gases.

- **Precipitaciones (LLU):**

Cantidad media de agua caída en cada una de las provincias españolas. Esta variable se expresa en milímetros.

Según dice textualmente la noticia el periódico La Razón: “Dicen los expertos que la contaminación no desaparece, que simplemente se traslada o se diluye cuando llueve”.

[\[La Razón\]](#)

- **Horas de sol (SOL):**

Número de horas medias anuales entre la salida y la puesta de sol.



## CAMPUS D'ALCOI

Según cita el periódico la razón: “los niveles de polución dependen de la meteorología y de su capacidad de condensarla o disiparla.” [[La Razón](#)]

- **Viento (VI):**

Velocidad del aire, cuya magnitud utilizada es kilómetros/hora.

Existe una disminución de la concentración de contaminantes en aire con el aumento de la velocidad del viento. [[El confidencial](#)]

- **Población (P):**

Cantidad de personas que residen en cada una de las provincias españolas. La población se mide en miles de personas.

Esta variable afecta a la contaminación ya que, a más población se existirá mayor consumo de componentes que afectan a la calidad del aire.

- **Ocupados (O):**

La tasa de ocupación agrupa el total de personas que se encuentran trabajando durante el ejercicio 2014. Se mide en miles de personas.

A mayor número de personas ocupadas mayor uso de vehículos y, por tanto, mayor contaminación.

- **Empresas (E):**

Número de centros de trabajo en cada una de las provincias españolas.

Cuantas más empresas existan mayor cantidad de gases contaminantes serán expulsados a la atmósfera.

- **Permisos de conducir (PC):**

Número de permisos de conducir en vigor de cada provincia en el año 2014. Esta variable es importante ya que cuantos más permisos expedidos de conducir puede incrementar el número de vehículos que expulsan gases contaminantes. Se mide en unidades.



- **Vehículos (VE):**

Los vehículos el medio más utilizado para el transporte de personas. Aunque estos sean de gran ayuda para la población expulsan gran cantidad de gases contaminantes. Se mide en unidades.

- **Autobuses (AU):**

Los autobuses son uno de los transportes públicos en las ciudades. Cuantos más autobuses existan en una población y mejor combinación para acceder a todas las direcciones menor será el uso de los “vehículos” y esto afectará a que no se desprendan tantos gases a la atmósfera. Se mide en unidades.

Todos los valores de las variables son del año 2014 para que todas estas sean homogéneas.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

## **2. Metodología y Técnicas:**



En este apartado se representa de forma teórica las herramientas estadísticas utilizadas para la formulación de los diferentes modelos.

### Estructura del proceso:

A continuación, se realiza la estructura básica realizada en este análisis para obtener un modelo aceptable y válido.

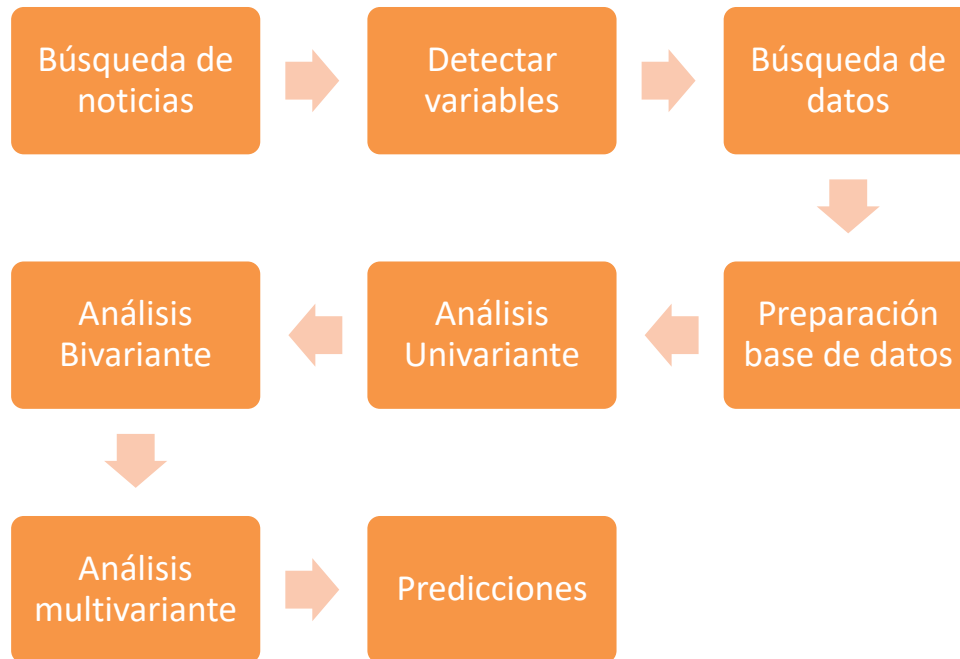


Figura 2.1. Proceso análisis

Se ha de decir que la fase de búsqueda de datos fue muy complicada, ya que cada estación mide diferentes gases. Además, al principio no se encontraba ninguna base de datos la cuál diera información por provincias, así que se debía mirar provincia por provincia y en bases de datos por días. Al final, tras muchas horas de búsqueda se detectó que en el Ministerio de agricultura y pesca, alimentación y medio ambiente existían los datos de todos los contaminantes y sólo se debía homogenizar los datos y agrupar por provincias.





En la figura 2.2 se realice un desglose de cada uno de los apartados de la estructura básica.

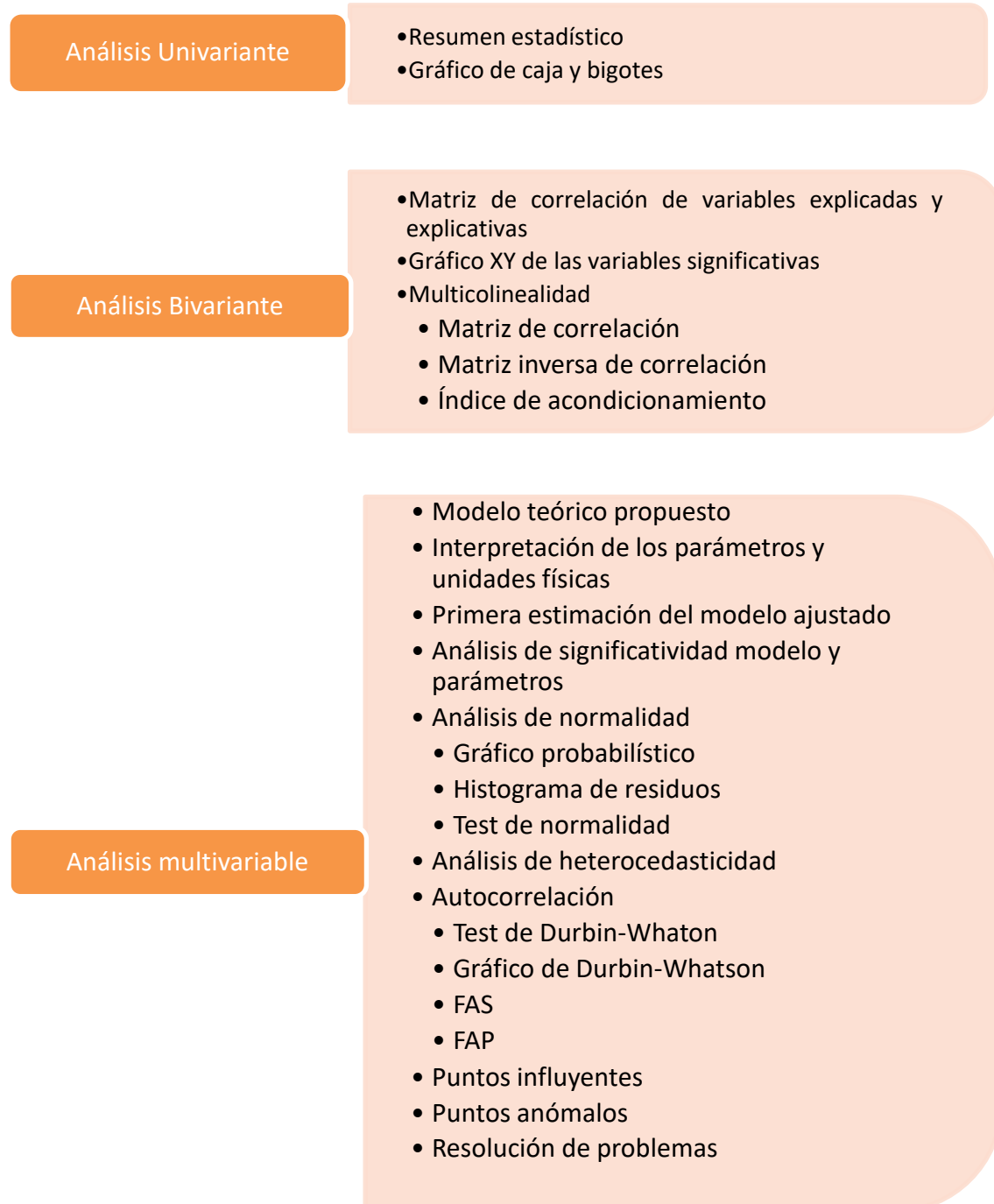


Figura 2.2. Resumen análisis



## Preparación de base de datos:

Para empezar a realizar este análisis se necesitaba saber cuáles eran las variables más importantes en la contaminación ambiental. Tras elegir estas variables que más relación tenían con este tema se debía realizar la búsqueda de datos.

Al principio se empezó a buscar los valores de las variables explicativas, ya que se pensaba que se podía encontrar la información más rápidamente. Se sacan directamente todas las variables sin necesidad de realizar operaciones, menos en el caso de vehículos que se debía restar la cantidad de autobuses porque se creía oportuno tenerla como una variable independiente.

A continuación, se intentaba encontrar los valores de los gases más importantes, pero se encontraba el problema que parecía que no existían datos agrupados. Se debían realizar las búsquedas de las provincias por separado, pero sólo se encontraban datos diarios no extrapolables, por ello se debían agrupar los datos uno a uno y manualmente.

Después de realizar muchas búsquedas y pensar que la recopilación de datos se hacía insostenible se llega de forma fortuita a la página del ministerio de agricultura y pesca, alimentación y medio ambiente. Allí se encuentran los datos históricos de la mayoría de gases y metales distribuidos por provincias desde 2001 a 2015. A partir de estos se visualizan todos para intentar homogeneizar todos los datos y saber que variables encontrábamos para todas las provincias. Se decide utilizar los valores del año 2014 y los gases en los que se encontraban todos los datos eran Dióxido de Azufre ( $\text{SO}_2$ ) y Óxidos de nitrógeno ( $\text{NO}_x$ ).

Para poder homogeneizar los valores de estos dos gases se debe filtrar por provincias y realizar una media de todas las estaciones. Se contaba con más de 143,000 estaciones y los datos estaban distribuidos por horas.

Al realizar las operaciones mencionadas anteriormente se podía decir que ya teníamos los suficientes datos para realizar el modelo de regresión, pero se decide buscar los datos del Dióxido de carbono ( $\text{CO}_2$ ) porque siempre que hablamos de contaminación se habla de este gas.

Tras horas de búsqueda y pensar que no se lograría encontrar este gas se localiza una la página de la empresa Certicalia en la que se indican los valores de Dióxido de carbono ( $\text{CO}_2$ ) pero en una magnitud energética y se corrobora con una llamada telefónica de que año son los valores. A partir de estos datos se necesitaba cambiar la unidad de medida de este gas a  $\text{KT}/\text{CO}_2$ .

Para realizar este cambio de magnitud se realizan los diferentes pasos:

- Multiplicamos las unidades de  $\text{KWh}/\text{m}^2$  por un factor de conversión el cual es 0,248654
- El valor hallado se multiplica por la superficie en metros cuadrados de cada una de las provincias españolas
- Se multiplica por 365 para que las unidades se conviertan en  $\text{KgCO}_2/\text{año}$
- Por último, como los datos son muy elevados se decide pasar la unidad a  $\text{KTCO}_2/\text{año}$



$$\frac{\text{KWh}}{\text{m}^2} \xrightarrow[\text{FC}^*]{\times 0,248656} \frac{\text{Kg CO}_2}{\text{m}^2 \times \text{año}} \xrightarrow{\times \text{Sup (m}^2\text{)}} \frac{\text{Kg CO}_2}{\text{año}}$$

$$\text{FC}^* \rightarrow \frac{1\text{KWh}}{\text{m}^2 \times \text{año}} = \frac{0,248656 \text{ Kg CO}_2}{\text{m}^2}$$

Después de realizar todo este proceso se podía decir que ya se había construido la base de datos que se utilizaría para realizar el modelo de regresión, pero se detecta que faltan varios datos para las provincias de Palencia y Guadalajara no existen datos suficientes con lo cual se prefiere eliminar estas dos provincias y no eliminar toda la variable de que no existe el dato.

Al realizar todo el proceso redactado anteriormente por fin se había obtenido una base de datos con la suficiente información para realizar el modelo de regresión, pero se piensa que es más lógico realizar tres modelos para observar cómo se comportan los tres gases obtenidos.



## Análisis univariante:

El análisis univariante realiza el estudio de cada una de las variables de forma independiente.

- **Promedio o media aritmética:**

Valor obtenido de la suma de datos entre la cantidad de datos.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

[Newbold, P; L. Carlson, W; Thorne]

- **Mediana:**

Valor que se encuentra el lugar central de las observaciones.

- n, número impar: lugar central.

$$Me = x_{(n+1)/2}$$

- n, número par: media de los dos valores centrales.

$$Me = \frac{\left( \frac{x_n}{2} + \frac{x_{n/2+1}}{2} \right)}{2}$$

[[Wikipedia](#)]

- **Varianza muestral:**

Suma de las diferencias entre los valores y el valor medio elevado al cuadrado entre la población.

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}$$



- **Desviación estándar:**

Raíz cuadrada de la varianza. Esta es una medida de dispersión que muestra cuál es la dispersión de los valores de la muestra.

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}}$$

[Newbold, P; L. Carlson, W; Thorne, B.]

- **Mínimo:**

Menor valor de la población

- **Máximo:**

Mayor valor de la población

- **Rango:**

Diferencia entre el mayor y menor valor.

$$\text{Rango} = x_n - x_1$$

- **Cuartil:**

Partes en las cuales se divide la población (añadir fórmula)

- Q1 Inferior: 25%
- Q3 Superior: 75%

- **Sesgo:**

Medida de simetría de la población.

$$\text{Sesgo} = \frac{1}{n} \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{S^3}$$

[Newbold, P; L. Carlson, W; Thorne, B.]



- **Curtosis:**

Mide como de concentrados se encuentran los valores de la muestra. [\[Wikipedia\]](#)

$$Curtosis = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{N \sigma^4} \quad [\text{Universo fórmulas}]$$

- **Gráfico de caja y bigotes:**

También llamado boxplots o box and whiskers. Es este se detalla visualmente las características del análisis univariante como los cuartiles, el valor máximo y mínimo.

En el centro de la caja se encuentra la mediana y la muestra la relación entre los cuartiles. Por otra parte, en los bigotes se encuentra el valor mínimo y máximo. Si existen datos no homogéneos estarían fuera de los bigotes y a estos los denominaríamos puntos anómalos.

[\[Liceo Marta Donoso Espejo\]](#)

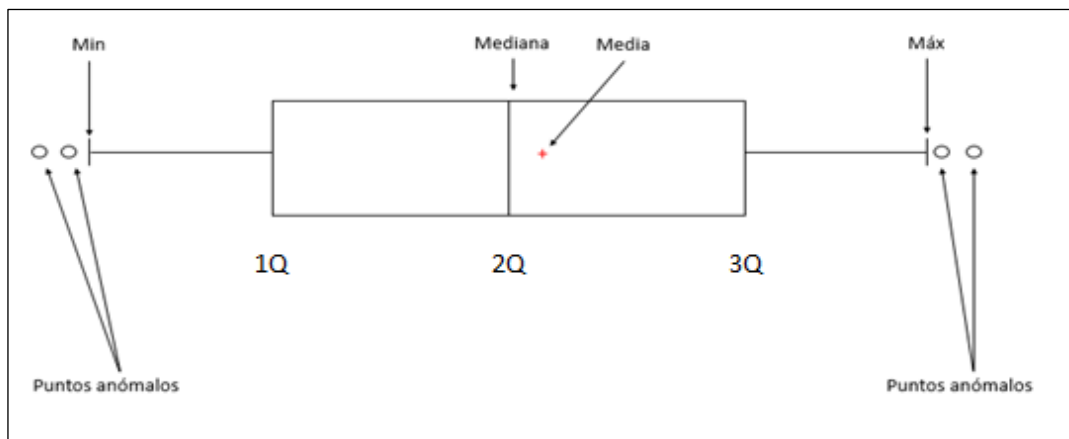


Gráfico 2.1. Gráfico de Caja y Bigotes



## Análisis Bivariante:

El análisis bivariante se utilizar para poder analizar dos variables conjuntamente. [[Uva](#)]

- **Matriz de correlación:**

Matriz en la cual se observan los coeficientes de correlación entre las variables.

- **Gráfico XY:**

Gráfico de puntos en el cual se representan los datos de la variable explicada contra la variable explicativa.

- **Multicolinealidad:**

La multicolinealidad produce errores en la estimación de los parámetros de un modelo de regresión, además de afectar al cálculo del P-valor. Existe problema de multicolinealidad cuando varias variables tienen valores parecidos o están relacionados.

- **Matriz de correlación:**

Se visualiza los coeficientes de correlación de cada una de las variables. Existe problema de multicolinealidad cuando al observar los valores inferiores de la diagonal principal de la matriz de correlación en valor absoluto, existe alguno superior a 0,7. (diapositivas clase)

- **Matriz inversa de correlación:**

Relaciona una variable explicada con todas las demás. Existe problema de multicolinealidad cuando al analizar los valores de la diagonal principal en la matriz inversa de correlación, alguno sea superior a 10.

- **Índice de acondicionamiento:**

Muestra la relación entre todas las variables a la vez. La fórmula de este es:

$$I.C. = \sqrt{\frac{\text{Autovalor}_{max}}{\text{Autovalor}_{min}}}$$

$10 \leq I.C. \leq 10$

no existe multicolinealidad

$10 \leq I.C. \leq 20$

multicolinealidad moderada

$20 \leq I.C.$

existe multicolinealidad elevada



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

## CAMPUS D'ALCOI

Para que no exista problema de multicolinealidad debe estar libre de esta en los tres métodos. Si por el caso contrario una de estas tiene problema de multicolinealidad se debe eliminar una de las variables a partir de uno de los siguientes criterios:

- 1º criterio: Al tener la matriz inversa se elimina el valor más grande de la diagonal principal salvo que la diferencia entre el valor más grande y el siguiente más grande sea menor a 1, ya que no se podría usar este criterio.
- 2º criterio: Realizar dos regresiones múltiples en el que en cada modelo faltará una de esas variables problemáticas, y nos quedaríamos con el modelo con el R-cuadrado más alto.





## Análisis multivariable:

- **Modelo teórico propuesto:**

Fórmula en la cual se observa de forma teórica la relación entre la variable explicada y las variables explicativas, además de la perturbación.

- **Análisis de normalidad:**

- Gráfico probabilístico:

El gráfico probabilístico y de probabilidad normal se utiliza para saber si los residuos son normales o no. Si estos son normales se observarán los puntos encima de la línea de normalidad y si por el contrario no lo son, existirá más dispersión entre los puntos y la línea. [[Wikipedia](#)]

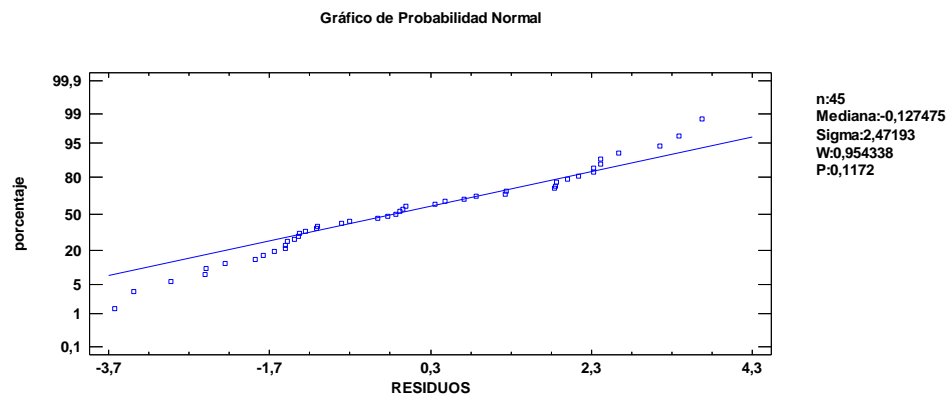


Gráfico 2.2. Distribución normal

- **Histograma de residuos:**

Gráfico donde se representan las frecuencias de cada uno de los valores. Este sirve para obtener de forma general la distribución de la muestra. Con este se puede observar el grado de normalidad de los residuos; los cuales son normales si el histograma presenta forma de campana de Gauss. Si por el contrario si los residuos no forman campana de Gauss podemos decir que los residuos puede que no sean normales.

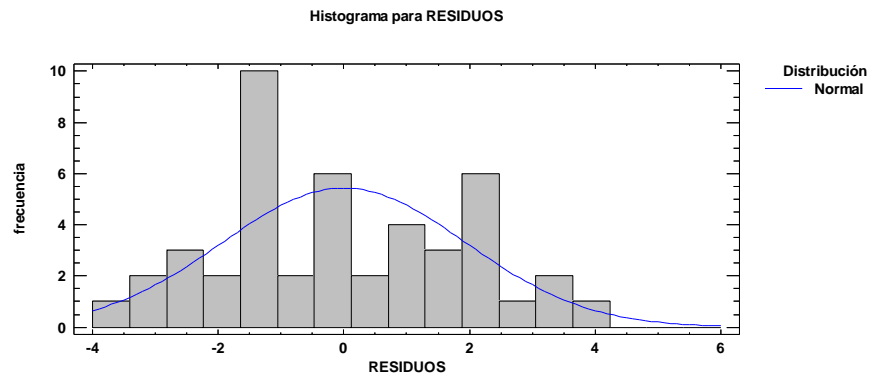


Gráfico 2.3. Histograma

○ **Test de normalidad:**

Al igual que los gráficos anteriores, el test de normalidad expresa si los residuos son o no normales. Si el P-valor obtenido es mayor a  $\alpha$  se aceptará la hipótesis nula y concluirá que los residuos son normales, si por el contrario el P-valor es menor se rechazará la hipótesis nula y los residuos no serán normales.

$$\left. \begin{array}{l} H_0: \text{Los residuos son normales} \\ H_1: \text{Los residuos no son normales} \end{array} \right\} \alpha: 0,05$$

• **Análisis de heterocedasticidad:**

Existe problema de heterocedasticidad cuando la varianza de la perturbación no es constante. Este problema se detecta cuando al realizar la regresión múltiple de los residuos al cuadrado y las variables explicativas, alguna de estas su P-valor es menor a  $\alpha$ .

$$\left. \begin{array}{l} H_0: \sigma^2: \text{Constante} \\ H_0: \sigma^2: \text{No es constante} \end{array} \right\} \alpha: 0,05$$



- **Autocorrelación:**

Existe autocorrelación cuando los valores de los residuos están relacionados en algún momento del tiempo.

- **Test de Durbin-Whatson:** Muestra si existe autocorrelación de primer orden en el modelo.

$$\left. \begin{array}{l} H_0: \rho_1 = 0 \text{ no hay autocorrelación} \\ H_1: \rho_1 \neq 0 \text{ hay autocorrelación} \end{array} \right\} \alpha=0,05$$

- **Gráfico de Durbin-Whatson:**

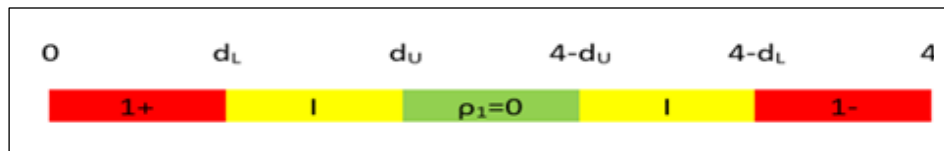


Gráfico 2.4. Gráfico de Durbin Whatson.

Al ubicar el estadístico de Durbin-Whatson en su gráfico puede existir autocorrelación de primer orden positiva y negativa, incertidumbre o no existir autocorrelación de primer orden.

- **FAS y FAP:**

Autocorrelaciones Estimadas para RESIDUOS

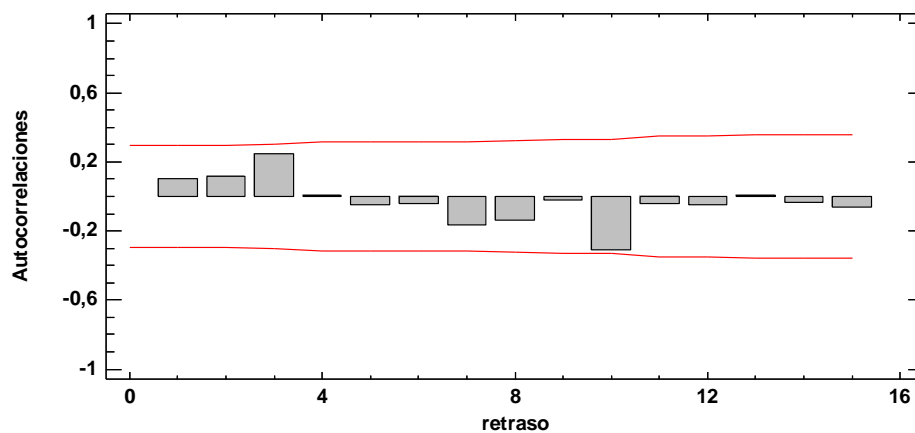


Gráfico 2.5. FAS



Autocorrelaciones Parciales Estimadas para RESIDUOS

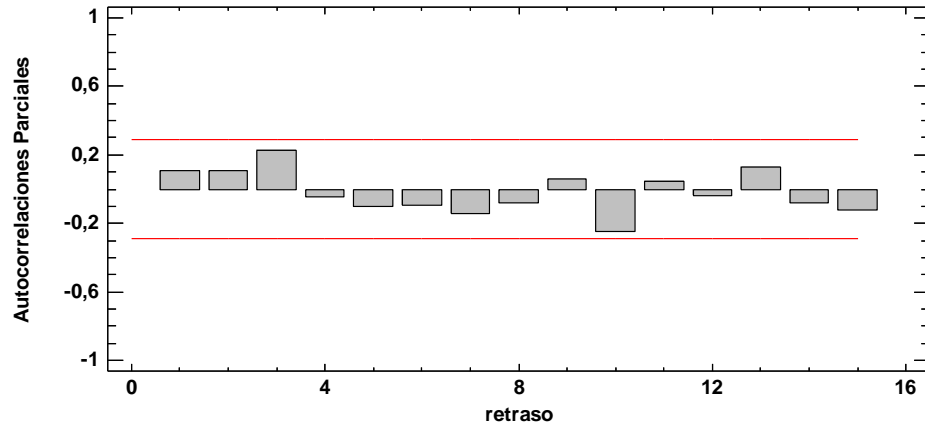


Gráfico 2.6. FAB

Si al observar los gráficos del FAS y el FAP no se observa ningún factor de correlación que supere los límites establecidos, por esto se podría decir que no existe problema de autocorrelación.

- **Puntos influyentes:** Valores del modelo que son importantes a la hora de realizar el ajuste del modelo, modificando sus parámetros.
- **Puntos anómalos:** Valores los cuales no son homogéneos al resto.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

### **3. Cálculos, Resultados, y Explotación:**



### Análisis Univariable:

Para realizar este tipo de análisis se utilizará una tabla de valores estadísticos y los gráficos de Dispersión y de Caja y bigotes.

**Dióxido de azufre (SO<sub>2</sub>):** El SO<sub>2</sub> se mide en µg/m<sup>3</sup>

Tabla 3.1. Resumen estadístico dióxido de azufre

Estadísticos	Valor
Promedio	4,319
Mediana	3,845
Desviación Estándar	2,249
Mínimo	0,957
Máximo	9,74
Rango	8,783
Cuartil Inferior	2,713
Cuartil Superior	5,66
Sesgo	0,624
Curtosis	-0,262

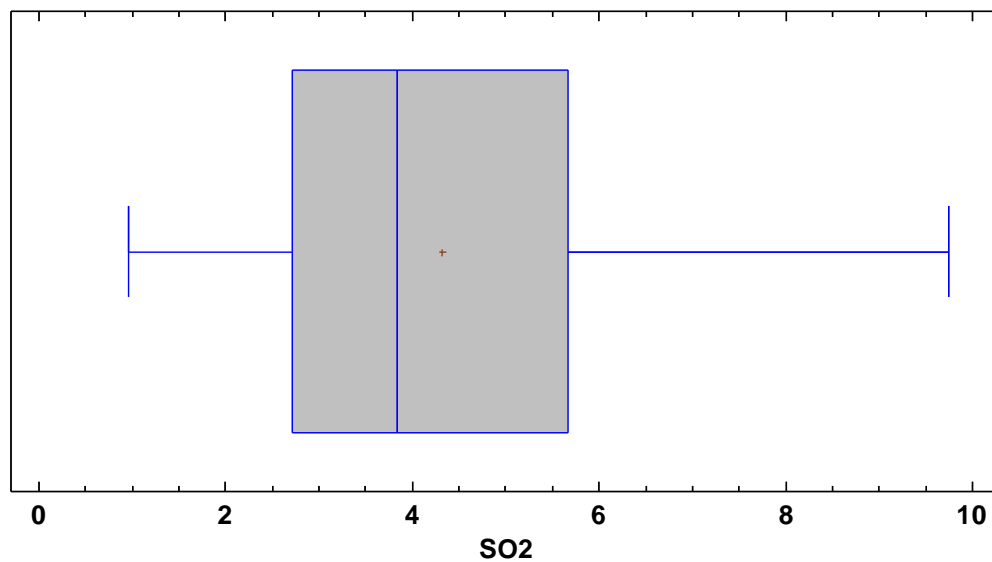


Gráfico 3. 1. Caja y Bigotes dióxido de azufre

El Dióxido de Azufre que se encuentra en la atmósfera en las provincias españolas en un valor de 3,855 µg/m<sup>3</sup>.



El mínimo se encuentra en Zamora con sólo 0.957  $\mu\text{g}/\text{m}^3$  y el valor máximo en la provincia de Cádiz con 9,740  $\mu\text{g}/\text{m}^3$ .

La desviación típica de los valores es de 2,249  $\mu\text{g}/\text{m}^3$ . Al realizar la comparación con la media de esta variable podemos decir que existe poca dispersión, ya que el valor es menor.

En este caso no se detecta ningún punto anómalo.

**Dióxido de carbono (CO<sub>2</sub>):** Este se mide en kilotoneladas al año.

Tabla 3.2. Resumen estadístico dióxido de carbono

Estadísticos	Valor
Promedio	141325
Mediana	126504
Desviación Estándar	82609
Mínimo	23919,4
Máximo	368214
Rango	344294
Cuartil Inferior	73698,5
Cuartil Superior	201301
Sesgo	0,698
Curtosis	-0,114

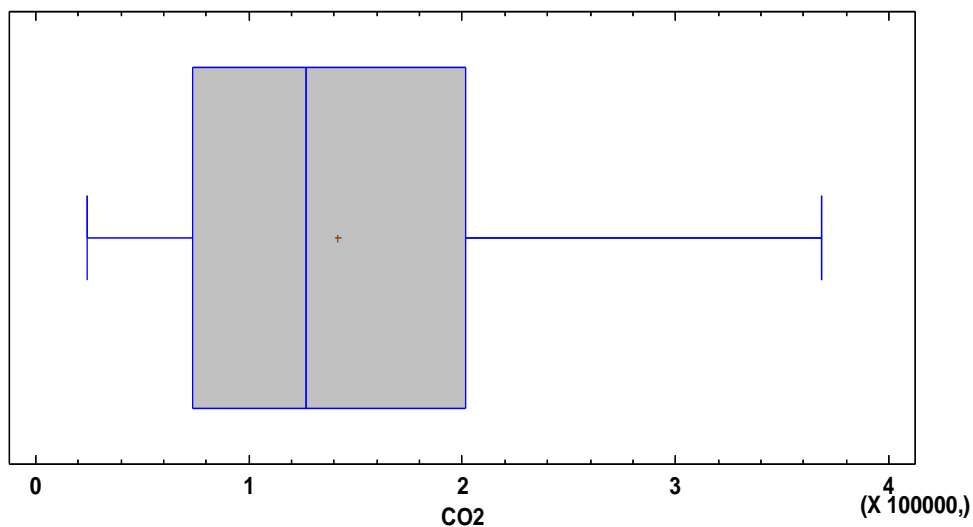


Gráfico 3.2. Caja y Bigotes dióxido de azufre



El promedio de dióxido de carbono en las provincias españolas es de 141.325 Kilo Toneladas al año.

El nivel mínimo se encuentra en la provincia de Guipúzcoa con un valor 23919,4 KT/año y el valor máximo en la provincia de Cáceres con un valor 368214 KT/año.

La desviación típica de las KT/año indica que los valores están poco dispersos, ya que el valor promedio es inferior a esta.

No existen puntos anómalos.

**Óxidos de nitrógeno (NO<sub>x</sub>):** Los óxidos de nitrógeno se miden en  $\mu\text{g}/\text{m}^3$

Tabla 3.3. Resumen estadístico óxidos de nitrógeno

Estadísticos	Valor
Promedio	24,152
Mediana	22,051
Desviación Estándar	10,212
Mínimo	6,312
Máximo	51,513
Rango	45,200
Cuartil Inferior	16,036
Cuartil Superior	31,544
Sesgo	0,609
Curtosis	-0,125

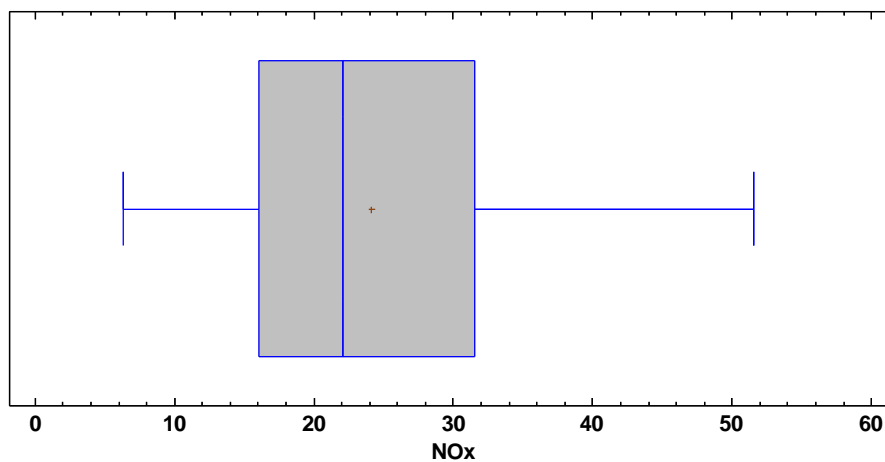


Gráfico 3.3. Caja y Bigotes óxidos de nitrógeno





El promedio o valor medio de óxidos de nitrógeno es de  $24,152 \mu\text{g}/\text{m}^3$ . El valor mínimo se encuentra en la provincia de Badajoz con  $6,312 \mu\text{g}/\text{m}^3$  y el máximo en Madrid con  $51,513$ .

En cuanto la desviación estándar asciende a  $10,2119 \mu\text{g}/\text{m}^3$ , al este valor ser la mitad a la media se puede observar que los datos están poco dispersos.

No se observan puntos anómalos.

**Temperatura (TEMP):** Expresada en grados centígrados

Tabla 3.4. Resumen estadístico temperatura

Estadísticos	Valor
Promedio	16,186
Mediana	16,075
Desviación Estándar	2,719
Mínimo	11,325
Máximo	21,750
Rango	10,425
Cuartil Inferior	14,079
Cuartil Superior	18,600
Sesgo	0,102
Curtosis	-0,766

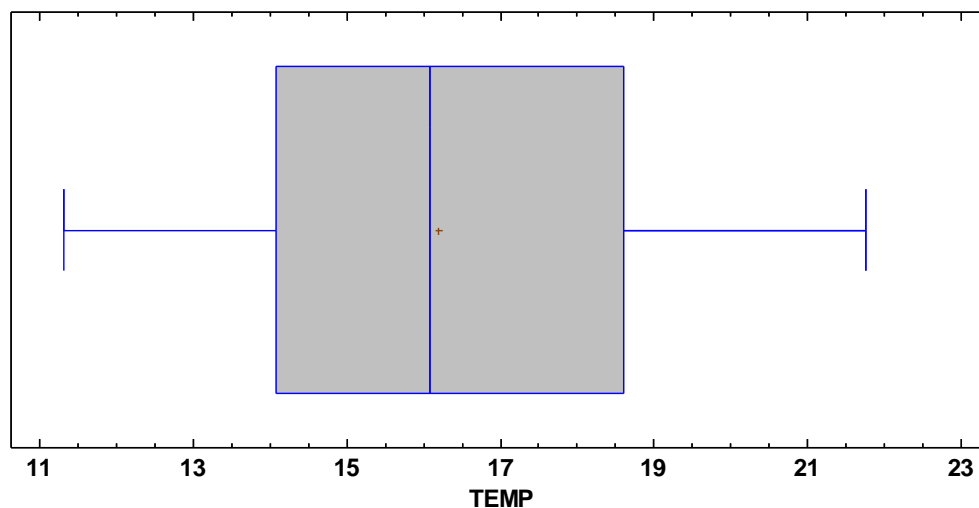


Gráfico 3.4. Caja y Bigotes temperatura



La temperatura media en las provincias españolas es de 16,186°C.

Se encuentra la temperatura más baja en el norte de España, en la provincia de León con una temperatura de 11,325 °C.

Por otra parte, la temperatura más alta se localiza en la provincia de Las palmas con un valor de 21,75 °C.

En cuanto a la desviación típica se observa que existe muy poca dispersión, ya que este valor es menor a la temperatura promedio.

No se observan puntos anómalos.

**Precipitaciones (LLU):** Se expresa en milímetros

Tabla 3.5. Resumen estadístico precipitaciones

Estadísticos	Valor
Promedio	49,530
Mediana	40,713
Desviación Estándar	30,761
Mínimo	10,550
Máximo	154,183
Rango	143,633
Cuartil Inferior	30,808
Cuartil Superior	53,171
Sesgo	1,595
Curtosis	2,971

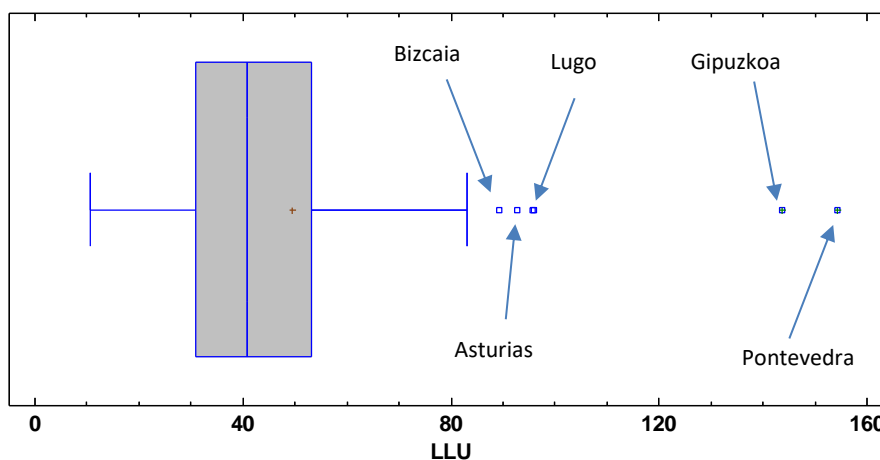


Gráfico 3.5. Caja y Bigotes precipitaciones



En cuanto a la cantidad media de lluvias en las provincias españolas se obtiene un valor de 49,5295mm.

La cantidad de lluvia obtiene un valor máximo en Pontevedra (154,183mm) y un valor mínimo en Almería (10,550mm)

Se encuentran varios puntos anómalos situados en Bizcaia, Asturias, Lugo, Gipuzkoa y Pontevedra.

**Horas de sol (SOL):** Expresado en horas medias de sol al mes

Tabla 3.6. Resumen estadístico horas de sol

Estadísticos	Valor
Promedio	215,886
Mediana	224,446
Desviación Estándar	41,937
Mínimo	122,600
Máximo	274,875
Rango	152,275
Cuartil Inferior	174,145
Cuartil Superior	249,417
Sesgo	-0,584
Curtosis	-0,767

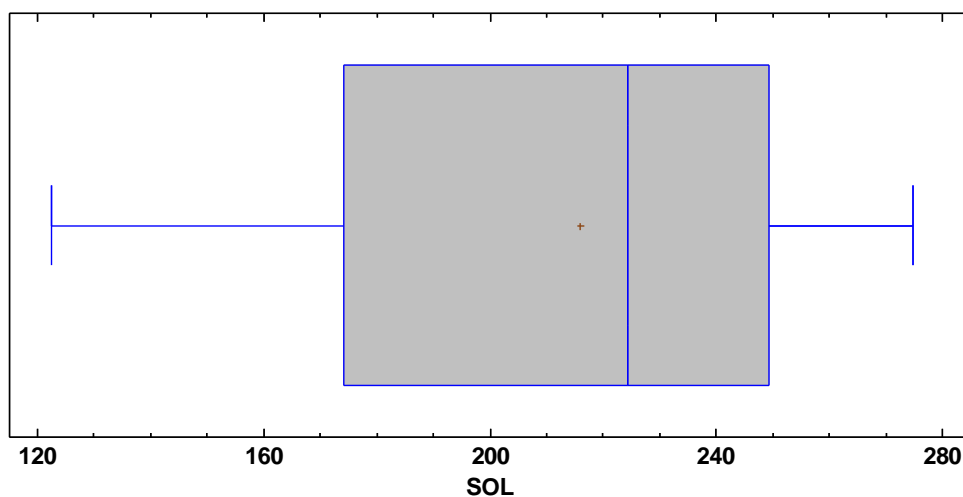


Gráfico 3.6. Caja y bigotes horas de sol



## CAMPUS D'ALCOI

El valor medio de horas de sol en las provincias españolas es de 215,886 horas.

El mínimo de horas de sol se encuentra en la provincia de Bizcaia (122,6 horas) y el máximo en la provincia de Sevilla (274,875 horas).

Existe poca dispersión, ya que el valor del promedio de las precipitaciones es mucho más elevado que el de la desviación estándar.

No se encuentra puntos anómalos.

### **Viento (VI):** Expresado en kilómetros hora

Tabla 3.7. Resumen estadístico viento

Estadísticos	Valor
Promedio	43,984
Mediana	44,054
Desviación Estándar	10,684
Mínimo	21,526
Máximo	64,429
Rango	42,903
Cuartil Inferior	36,708
Cuartil Superior	49,619
Sesgo	-0,072
Curtosis	-0,639

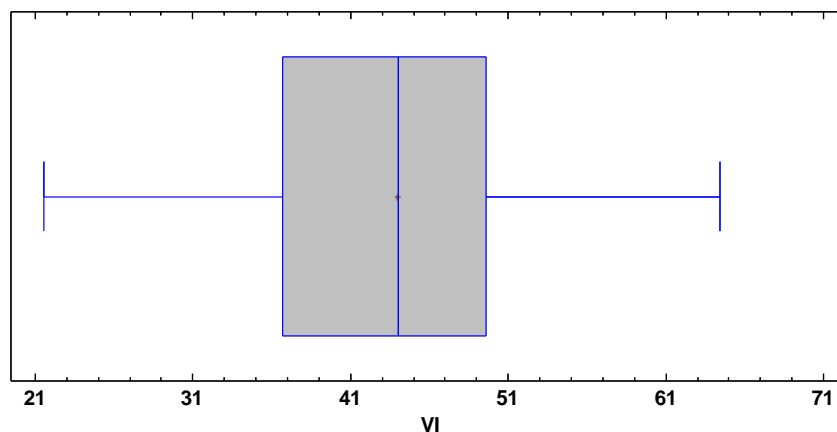


Gráfico 3.7. Caja y Bigotes viento



En cuanto la velocidad del viento, tiene un promedio de 43,9838km/h. El valor mínimo se encuentra en la provincia de Barcelona (21,5256km/h) y el valor máximo en La Coruña (64,4286km/h).

Los kilómetros horas medios son superiores a la desviación, por lo que no existe demasiada dispersión.

No existen puntos anómalos.

**Población (P):** Expresado en miles de personas

Tabla 3.8. Resumen estadístico población

Estadísticos	Valor
Promedio	962059
Mediana	674981
Desviación Estándar	1180390
Mínimo	92221
Máximo	6454440
Rango	6362220
Cuartil Inferior	354824
Cuartil Superior	1080890
Sesgo	3,518
Curtosis	13,678

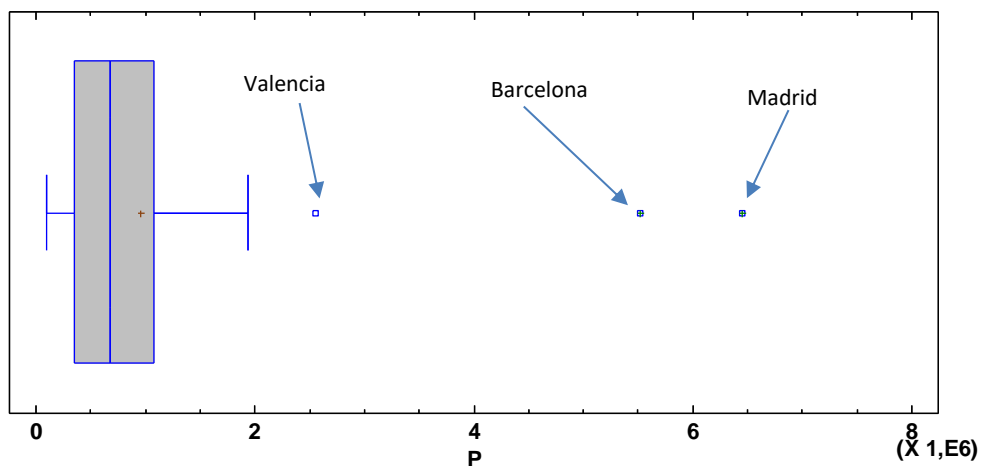


Gráfico 3.8. Caja y Bigotes población

El promedio de población de las provincias españolas es de 962059 miles de personas. La provincia con menor población es Soria con 92221 miles de personas y el máximo se encuentra en la provincia de Madrid con 6454440 miles de personas.



En este caso la desviación típica es mayor al promedio con lo cual se puede decir que la población en las provincias españolas existe mucha dispersión.

Se encuentran puntos anómalos en las provincias de Valencia, Barcelona y Madrid.

**Ocupados (O):** Se mide en miles de personas.

Tabla 3.9. Resumen estadístico ocupados

Estadísticos	Valor
Promedio	356,696
Mediana	227,000
Desviación Estándar	484,380
Mínimo	36,7
Máximo	2715
Rango	2678,3
Cuartil Inferior	134,8
Cuartil Superior	378,1
Sesgo	3,841
Curtosis	15,936

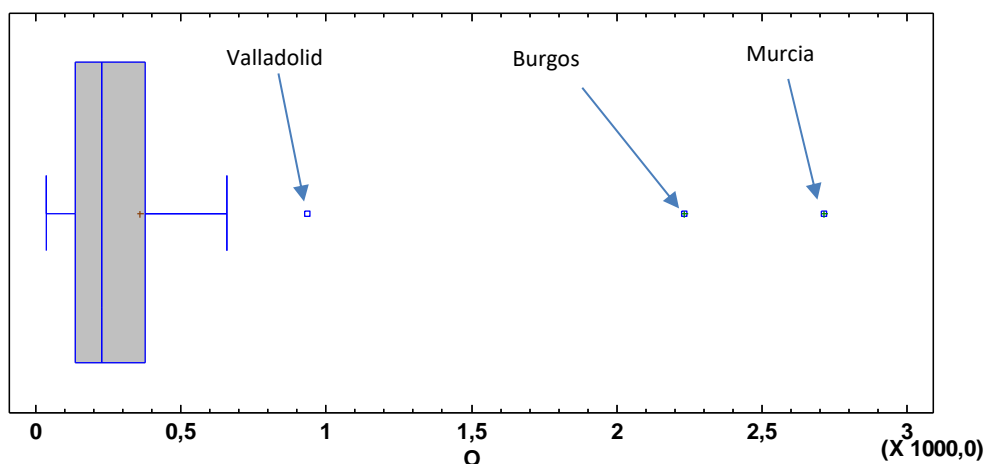


Gráfico 3.9. Caja y Bigotes ocupados

El valor medio de miles de ocupados en las provincias españolas es de 356,696. La provincia con menor número de ocupados es Tarragona con 36,7 millones de personas. El mayor número de ocupados corresponde a Murcia con 2715 miles de personas.



Existe escasa dispersión, ya que el valor de la desviación típica es mayor al valor del promedio.  
Existen puntos anómalos en las provincias de Valladolid, Burgos y Murcia.

**Empresas (E):** Número de centros de trabajo

Tabla 3.10. Resumen estadístico empresas

Estadísticos	Valor
Promedio	64350
Mediana	39080
Desviación Estándar	90745
Mínimo	5794
Máximo	494509
Rango	488715
Cuartil Inferior	23617
Cuartil Superior	65933
Sesgo	3,862
Curtosis	15,807

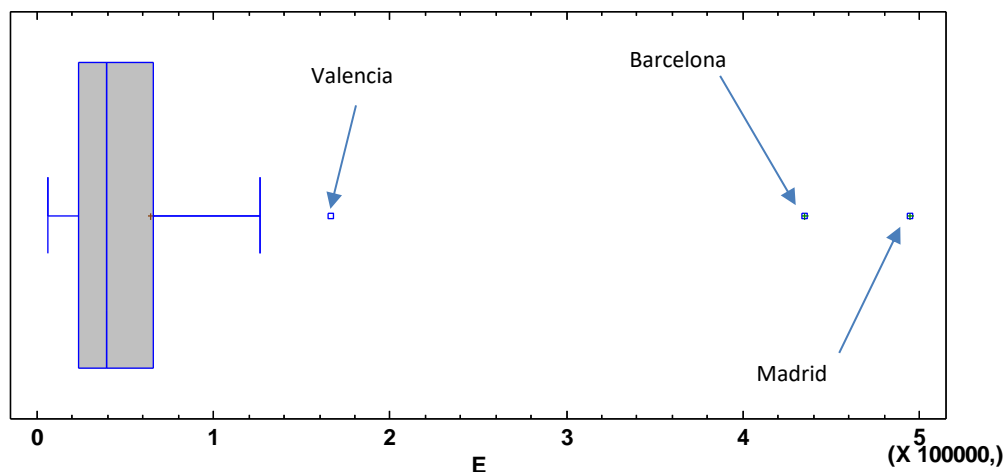


Gráfico 3.10. Caja y Bigotes empresas

El valor medio de empresas en España es de 64350. El valor máximo se encuentra en la provincia de Soria con un valor de 5794 empresas, por el contrario, el valor máximo se encuentra en Madrid con 494509 empresas.

Existe gran dispersión ya que el valor medio es menor que el valor de la desviación típica.  
Se encuentran puntos anómalos en Valencia, Barcelona y Madrid.



**Permisos de conducir (PC):** Se mide en unidades.

Tabla 3.11. Resumen estadístico permisos de conducir.

Estadísticos	Valor
Promedio	64350
Mediana	39080
Desviación Estándar	90745
Mínimo	5794
Máximo	494509
Rango	488715
Cuartil Inferior	23617,5
Cuartil Superior	65933,5
Sesgo	3,862
Curtosis	15,807

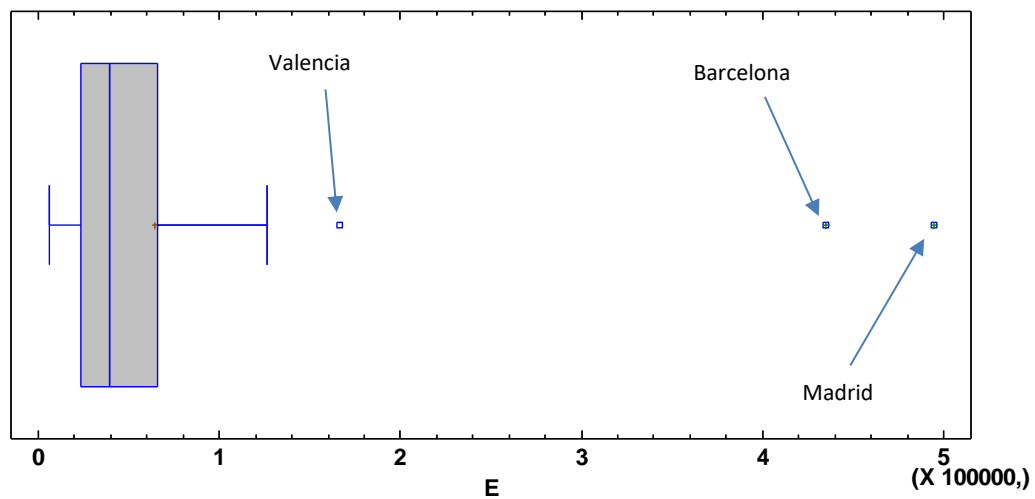


Gráfico 3.11. Caja y Bigotes permisos de conducir

Los permisos de conducir medios expedidos en las provincias españolas ascienden a 13539,5. El valor mínimo se encuentra en la provincia de Soria y el valor máximo en la provincia de Madrid. La dispersión es elevada, ya que el valor de la desviación típica es mayor al promedio.

Se encuentran puntos anómalos en Valencia, Barcelona y Madrid.





**Vehículos (VE):** Se mide en unidades.

Tabla 3.12. Resumen estadístico vehículos

Estadísticos	Valor
Promedio	635436
Mediana	436484
Desviación Estándar	757012
Mínimo	70816
Máximo	4,22E+06
Rango	4,15E+06
Cuartil Inferior	254398
Cuartil Superior	696213
Sesgo	3,54126
Curtosis	13,989

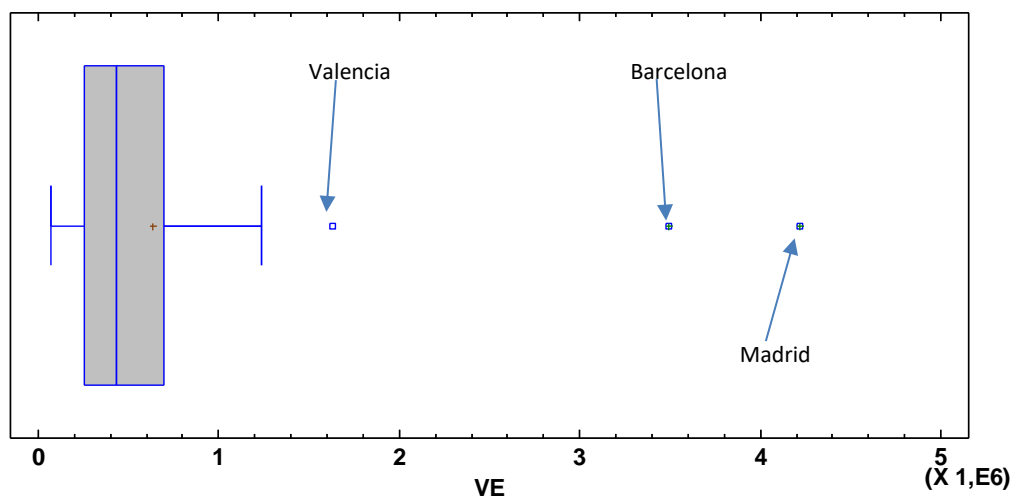


Gráfico 3.12. Caja y Bigotes vehículos

El valor medio de vehículos en las provincias es de 635436. La provincia en la que menos número de vehículos existe es Soria con 70816 vehículos y la provincia con más vehículos es Madrid con 4220250 vehículos.

La dispersión es muy baja, ya que el valor medio es mejor que el de la desviación típica. Existen puntos anómalos en Valencia, Barcelona y Madrid.



**Autobuses (AU):** Se mide en unidades.

Tabla 3.13. Resumen estadístico autobuses

Estadísticos	Valor
Promedio	1233,29
Mediana	711
Desviación Estándar	1646,51
Mínimo	100
Máximo	1,02E+04
Rango	1,01E+04
Cuartil Inferior	482
Cuartil Superior	1467,5
Sesgo	4,10761
Curtosis	20,0112

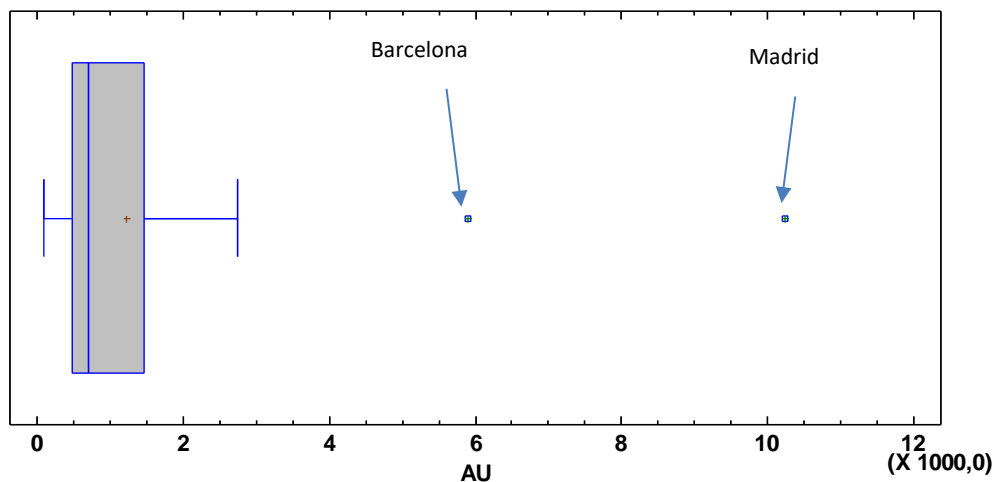


Gráfico 3.13. Caja y Bigotes autobuses

El promedio de autobuses en las provincias españolas es de 1233,29. El número mínimo de autobuses se encuentra en Soria (100 autobuses) y el número máximo se encuentra en Madrid (10232 autobuses).

La dispersión es pequeña, ya que el valor medio es menor que el de la desviación típica. Existen puntos anómalos en Barcelona y Madrid.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

A continuació, se realitzarà el anàlisis bivariante, el cual estudia datos homogéneos. Por este motivo se eliminan los tres puntos anómalos detectados en el anàlisis bivariante, estos son: Barcelona, Madrid y Valencia.



## Análisis Bivariante:

Se va a realizar el análisis bivariante de tres modelos, uno con cada uno de los gases y se van a relacionar con todas las variables explicativas.

Para realizarlo se ha calculado la matriz de correlación con todas las variables explicadas y explicativas.

Tabla 3.14. Matriz de correlación variables explicadas y explicativas

	TEMP	LLU	SOL	VI	P	O	E	PC	VE	AU
SO <sub>2</sub>	0,380	-0,097	0,139	-0,260	<b>0,441</b>	0,160	0,347	<b>0,449</b>	<b>0,400</b>	0,332
CO <sub>2</sub>	-0,274	-0,303	0,272	-0,014	<b>-0,425</b>	0,107	<b>-0,461</b>	-0,398	<b>-0,436</b>	<b>-0,496</b>
NO <sub>x</sub>	0,019	0,340	-0,175	-0,042	0,378	0,218	0,374	0,317	0,328	0,238

A partir de la matriz anterior se va a representar la relación entre los gases y las variables explicativas que tienen mayor correlación.

### Variable Dióxido de azufre (SO<sub>2</sub>) frente a las variables x:

- Dióxido de azufre (SO<sub>2</sub>) frente a la variable población (P):

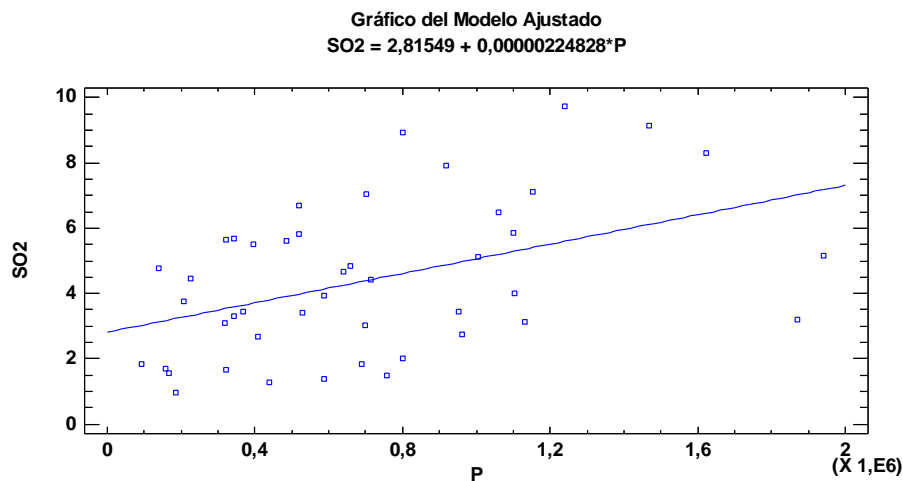


Gráfico 3.14. Dióxido de azufre frente población

R-cuadrada = 19,512%

Al analizar en el gráfico 3.14, el ajuste del modelo es muy bajo. Además, presenta forma cónica con lo cual podría tener un problema de heterocedasticidad.

Como se observa en la matriz de correlación la variable Dióxido de azufre frente a la población tiene una relación directa, ya que el R obtenido es de 0,441.



- Dióxido de azufre (SO<sub>2</sub>) frente variable Permisos de conducir (PC):

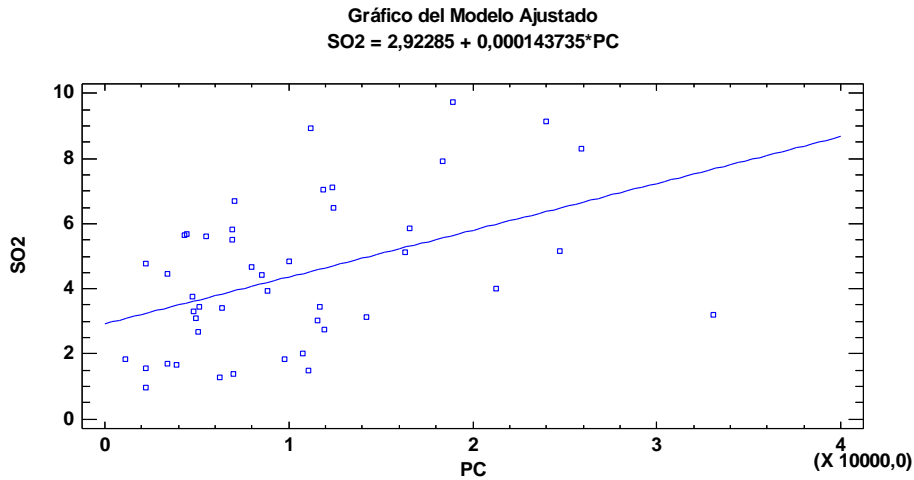


Gráfico 3.15. Dióxido de azufre frente permisos de conducir

R-cuadrada = 20,117%

Al prestar atención al gráfico del modelo ajustado de la variable dióxido de azufre (SO<sub>2</sub>) frente a los permisos de conducir (PC) se puede decir que existe un ajuste muy bajo. Asimismo, se observa forma cónica por ello podría existir un problema de heterocedasticidad.

Al analizar a la matriz de correlación la variable explicada frente a la variable explicativa tiene una relación directa, ya que el R obtenido es de 0,449.

- Dióxido de azufre (SO<sub>2</sub>) frente vehículos (V):

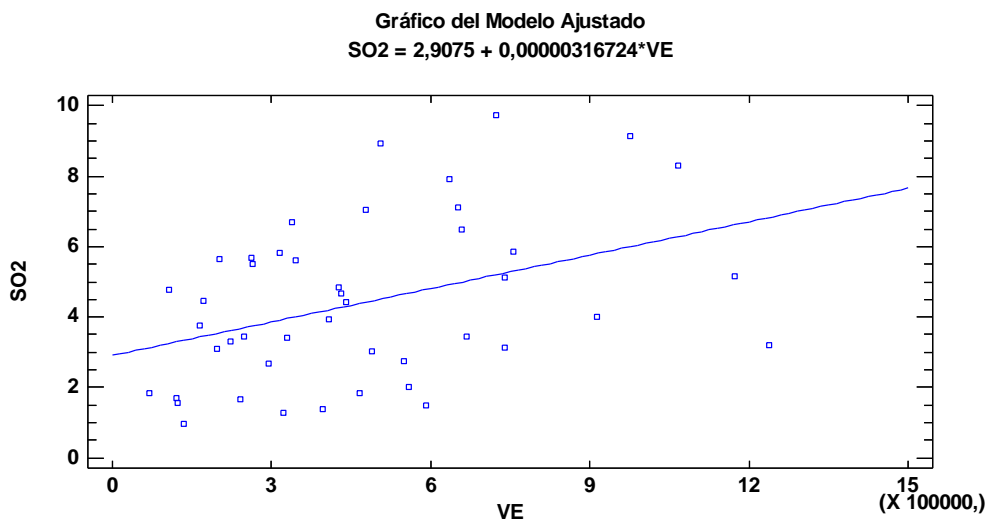


Gráfico 3.16. Dióxido de azufre frente vehículos

R-cuadrada = 16,029%



Al analizar el gráfico 3.16 el ajuste del modelo es muy bajo. También presenta forma cónica por lo cual podría hallarse un problema de heterocedasticidad.

Observado la matriz de correlación la variable Dióxido de azufre (SO<sub>2</sub>) frente a los vehículos (V) tiene una relación directa, ya que el R conseguido es de 0,400.

### Variable Dióxido de carbono (CO<sub>2</sub>) frente a las variables x:

- Dióxido de carbono (CO<sub>2</sub>) frente a la población (P):

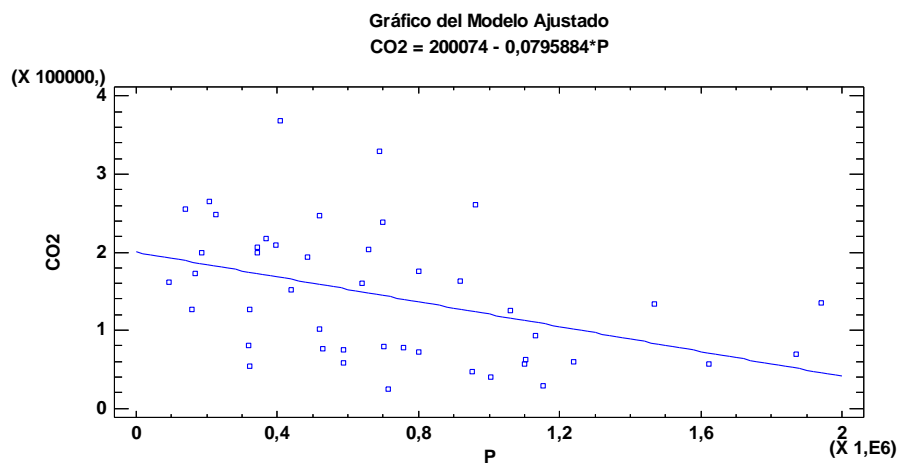


Gráfico 3.17. Dióxido de carbono frente población

R-cuadrada = 18,091%

Al analizar el gráfico del modelo ajustado de la variable dióxido de carbono (CO<sub>2</sub>) frente a la población se observa que existe un ajuste muy bajo.

Al prestar atención a la matriz de correlación encontramos una relación directa, ya que el R obtenido es de 0,425.



CAMPUS D'ALCOI

- CO<sub>2</sub> frente a la variable número de empresas (E):

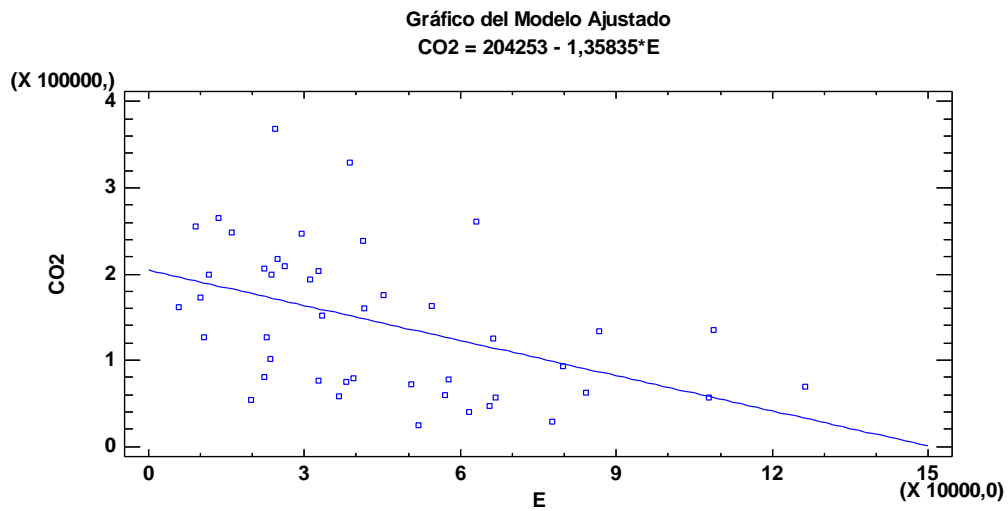


Gráfico 3.18. Dióxido de carbono frente empresas

R-cuadrada = 21,2139%

Al prestar atención al gráfico del modelo ajustado de la variable dióxido de carbono (CO<sub>2</sub>) frente al número de empresas (E) se observa que existe un ajuste muy bajo.

Al analizar la matriz de correlación la variable explicada frente a la variable explicativa tiene una relación directa, ya que el R obtenido es de 0,449.



- Dióxido de carbono (CO<sub>2</sub>) frente a número de vehículos (V):

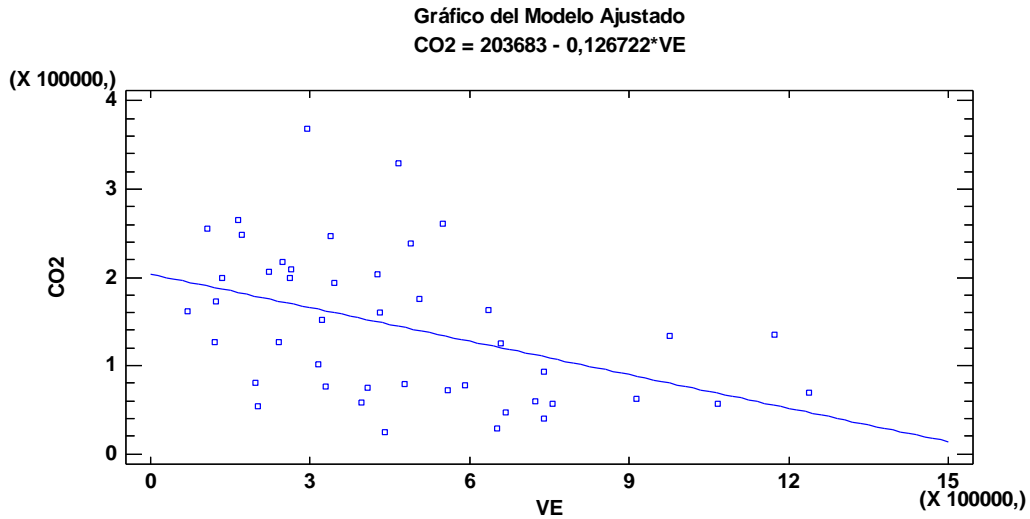


Gráfico 3.19. Dióxido de carbono frente número de vehículos

R-cuadrada = 18,984%

Como se observa en el gráfico 3.19 el ajuste del modelo es muy bajo. Además, presenta una línea curva con lo cual este modelo podría tener un problema de falta de linealidad.

En la matriz de correlación la variable Dióxido de carbono (CO<sub>2</sub>) frente a la población tiene una relación directa, ya que el R obtenido es de -0,436.





- Dióxido de carbono (CO<sub>2</sub>) frente a número de autobuses (AU):

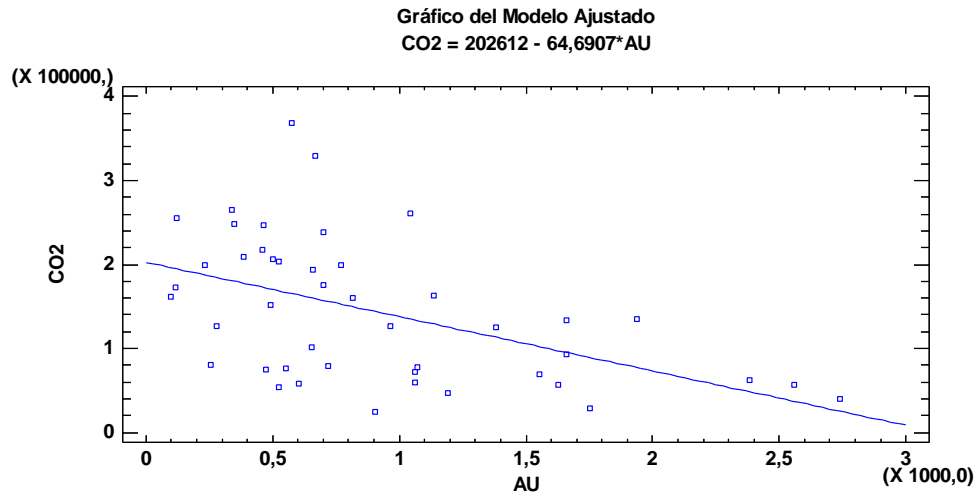


Gráfico 3.20. Dióxido de carbono frente autobuses

R-cuadrada = 24,641%

Al observar el gráfico de modelo de ajuste entre la variable explicada y explicativa se llega a la conclusión que existe un ajuste muy bajo.

La relación entre el Dióxido de carbono y el número de autobuses es directa.



## Multicolinealidad:

Existen problemas de multicolinealidad cuando cuando dos o más variables explicativas están relacionadas entre sí y toman valores semejantes o proporcionales.

Para averiguar si existe este tipo de problema se realizarán tres métodos de análisis: matriz de correlación, matriz inversa de correlación e índice de acondicionamiento.

En primer lugar, se realiza la matriz de correlación:

Tabla 3.15. Matriz de correlación

	TEMP	LLU	SOL	VI	P	O	E	PC	VE	AU
TEMP	1,000	-0,345	0,502	-0,266	0,664	-0,136	0,604	0,703	0,685	0,639
LLU	<b>-0,345</b>	1,000	-0,711	0,183	-0,014	-0,113	0,038	-0,169	-0,029	0,053
SOL	<b>0,502</b>	<b>-0,711</b>	1,000	-0,227	0,186	0,080	0,104	0,317	0,198	-0,044
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1,000	-0,048	0,068	-0,001	-0,127	-0,078	-0,218
P	<b>0,664</b>	<b>-0,014</b>	<b>0,186</b>	<b>-0,048</b>	1,000	0,016	0,978	0,960	0,987	0,793
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	<b>0,016</b>	1,000	0,009	0,046	0,015	-0,033
E	<b>0,604</b>	<b>0,038</b>	<b>0,104</b>	<b>-0,001</b>	<b>0,978</b>	<b>0,009</b>	1,000	0,945	0,982	0,812
PC	<b>0,703</b>	<b>-0,169</b>	<b>0,317</b>	<b>-0,127</b>	<b>0,960</b>	<b>0,046</b>	<b>0,945</b>	1,000	0,974	0,768
VE	<b>0,685</b>	<b>-0,029</b>	<b>0,198</b>	<b>-0,078</b>	<b>0,987</b>	<b>0,015</b>	<b>0,982</b>	<b>0,974</b>	1,000	0,832
AU	<b>0,639</b>	<b>0,053</b>	<b>-0,044</b>	<b>-0,218</b>	<b>0,793</b>	<b>-0,033</b>	<b>0,812</b>	<b>0,768</b>	<b>0,832</b>	1,000

Como se puede observar en la matriz de correlación, por debajo de la diagonal principal existen  $|R_{ij}|$  superiores a 0,7. Esto indica que existe un problema de multicolinealidad.

A partir de la matriz de correlación se realiza la matriz inversa de correlación.

Tabla 3.16. Matriz inversa de correlación

TEMP	<b>3,942</b>	0,292	-1,504	-0,106	-2,268	0,653	3,979	0,389	-2,782	-2,018
LLU	0,292	<b>2,793</b>	1,719	0,283	0,296	0,118	0,753	4,774	-7,165	1,253
SOL	-1,504	1,719	<b>3,909</b>	0,382	1,392	-0,203	2,960	-0,163	-6,169	2,870
VI	-0,106	0,283	0,382	<b>1,462</b>	0,435	-0,134	-2,207	1,968	-1,046	1,189
P	-2,268	0,296	1,392	0,435	<b>46,953</b>	-0,393	-11,661	3,792	-42,369	6,167
O	0,653	0,118	-0,203	-0,134	-0,393	<b>1,142</b>	0,868	-0,536	-0,118	-0,306
E	3,979	0,753	2,960	-2,207	-11,661	0,868	<b>49,704</b>	0,001	-40,581	-0,247
PC	0,389	4,774	-0,163	1,968	3,792	-0,536	0,001	<b>40,095</b>	-47,202	5,357
VE	-2,782	-7,165	-6,169	-1,046	-42,369	-0,118	-40,581	-47,202	<b>144,866</b>	-16,080
AU	-2,018	1,253	2,870	1,189	6,167	-0,306	-0,247	5,357	-16,080	<b>7,174</b>



Al visualizar la diagonal principal existen valores superiores a 10 con lo cual existe un problema de multicolinealidad. Las variables en las que se encuentran problemas son: población (P), empresas (E), permisos de conducir (PC) y vehículos (VE).

A continuación, se va a realizar el índice de acondicionamiento.

Tabla 3.17. Autovalores matriz de correlación

Número	Autovalor
1	5,246
2	1,947
3	1,124
4	0,868
5	0,349
6	0,304
7	0,113
8	0,027
9	0,016
10	0,005

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{5,246}{0,005}} = 31,295$$

El valor del índice de condicionamiento es mayor a 10 con lo cual existe problema de multicolinealidad muy elevada.

Para la eliminación de variables existen dos criterios:

- 1º criterio: Al tener la matriz inversa se debe eliminar el valor más grande de la diagonal principal salvo que la diferencia entre el valor más grande y el siguiente más grande sea menor a 1 ya que no se podría usar este criterio.
- 2º criterio: Realizar dos regresiones múltiples en el que en cada modelo faltará una de esas variables problemáticas y nos quedaríamos con el R-cuadrado más alto.



A partir de las tres pruebas se decide eliminar la variable vehículos ya que, siguiendo el primer criterio debemos eliminar el valor más grande la diagonal principal salvo que la diferencia entre el mayor y el siguiente sea menor a 1. Después de eliminar esta variable se sigue teniendo problemas de multicolinealidad. Por este motivo se deben volver a realizar las tres pruebas anteriores.

De nuevo se revuelve a realizar la matriz de correlación.

Tabla 3.18. Matriz de correlación

	TEMP	LLU	SOL	VI	P	O	E	PC	AU
TEMP	1	-0,345	0,502	-0,266	0,664	-0,136	0,604	0,703	0,639
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,014	-0,113	0,038	-0,169	0,053
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,186	0,080	0,104	0,317	-0,044
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	-0,048	0,068	-0,001	-0,127	-0,218
P	<b>0,664</b>	<b>-0,014</b>	<b>0,186</b>	<b>-0,048</b>	1	0,016	0,978	0,960	0,793
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	<b>0,016</b>	1	0,009	0,046	-0,033
E	<b>0,604</b>	<b>0,038</b>	<b>0,104</b>	<b>-0,001</b>	<b>0,978</b>	<b>0,009</b>	1	0,945	0,812
PC	<b>0,703</b>	<b>-0,169</b>	<b>0,317</b>	<b>-0,127</b>	<b>0,960</b>	<b>0,046</b>	<b>0,945</b>	1	0,768
AU	<b>0,639</b>	<b>0,053</b>	<b>-0,044</b>	<b>-0,218</b>	<b>0,793</b>	<b>-0,033</b>	<b>0,812</b>	<b>0,768</b>	1

Al realizar el análisis de la matriz de correlación se detecta que valor valores situados por debajo de la diagonal principal en valor absoluto son mayores a 0,70. Por este motivo se puede decir que existe problema de multicolinealidad.

A continuación, se realiza la matriz inversa de correlación.

Tabla 3.19. Matriz inversa de correlación

TEMP	<b>3,889</b>	0,154	-1,622	-0,126	-3,082	0,651	3,199	-0,518	-2,327
LLU	0,154	<b>2,439</b>	1,414	0,231	-1,799	0,112	-1,254	2,440	0,457
SOL	-1,622	1,414	<b>3,646</b>	0,337	-0,412	-0,208	1,232	-2,173	2,185
VI	-0,126	0,231	0,337	<b>1,454</b>	0,129	-0,135	-2,500	1,627	1,072
P	-3,082	-1,799	-0,412	0,129	<b>34,562</b>	-0,428	-23,530	-10,013	1,463
O	0,651	0,112	-0,208	-0,135	-0,428	<b>1,142</b>	0,835	-0,575	-0,319
E	3,199	-1,254	1,232	-2,500	-23,530	0,835	<b>38,336</b>	-13,221	-4,752
PC	-0,518	2,440	-2,173	1,627	-10,013	-0,575	-13,221	<b>24,716</b>	0,118
AU	-2,327	0,457	2,185	1,072	1,463	-0,319	-4,752	0,118	<b>5,389</b>

Al observar los valores obtenidos en la diagonal principal existen valores mayores a 10. Esto nos hace saber que existe un problema de multicolinealidad.



Después, se efectúa el último de los tres análisis que es el índice de acondicionamiento

Tabla 3.20. Autovalores de la matriz de correlación

Número	Autovalor
1	4,297
2	1,918
3	1,118
4	0,868
5	0,342
6	0,300
7	0,113
8	0,027
9	0,016

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{4,297}{0,016}} = 16,217$$

El resultado del índice de correlación es mayor a 10 por lo que podemos decir que existe un problema de multicolinealidad moderada.

Como al realizar las tres pruebas se encuentra problema de multicolinealidad se debe eliminar una de las variables. Al dirigirse a la matriz inversa de correlación se detecta que el valor más alto de la diagonal es el de la variable empresas (E) y como la diferencia con el valor más próximo es mayor a uno se puede eliminar la variable a partir del primer criterio.

Se debe realizar los mismos análisis para tratar de eliminar el problema de multicolinealidad.

Primero se realiza la matriz de correlación eliminando la variable empresas como se ha comentado anteriormente.

Tabla 3.21. Matriz de correlación

	TEMP	LLU	SOL	VI	P	O	PC	AU
TEMP	1	-0,345	0,502	-0,266	0,664	-0,136	0,703	0,639
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,014	-0,113	-0,169	0,053
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,186	0,080	0,317	-0,044
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	-0,048	0,068	-0,127	-0,218
P	<b>0,664</b>	<b>-0,014</b>	<b>0,186</b>	<b>-0,048</b>	1	0,016	0,960	0,793
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	<b>0,016</b>	1	0,046	-0,033
PC	<b>0,703</b>	<b>-0,169</b>	<b>0,317</b>	<b>-0,127</b>	<b>0,960</b>	<b>0,046</b>	1	0,768
AU	<b>0,639</b>	<b>0,053</b>	<b>-0,044</b>	<b>-0,218</b>	<b>0,793</b>	<b>-0,033</b>	<b>0,768</b>	1



Al analizar los valores por debajo de la diagonal existen varios con valor mayor a 0,7. Estos valores indican que existe problema de multicolinealidad.

Posteriormente se realizará el segundo análisis que consiste en la matriz inversa de correlación.

Tabla 3.22. Matriz inversa de correlación

TEMP	<b>3,622</b>	0,259	-1,725	0,083	-1,118	0,581	0,585	-1,931
LLU	0,259	<b>2,398</b>	1,454	0,149	-2,569	0,139	2,007	0,302
SOL	-1,725	1,454	<b>3,607</b>	0,417	0,344	-0,235	-1,748	2,338
VI	0,083	0,149	0,417	<b>1,291</b>	-1,406	-0,081	0,765	0,763
P	-1,118	-2,569	0,344	-1,406	<b>20,119</b>	0,084	-18,128	-1,453
O	0,581	0,139	-0,235	-0,081	0,084	<b>1,124</b>	-0,287	-0,216
PC	0,585	2,007	-1,748	0,765	-18,128	-0,287	<b>20,156</b>	-1,521
AU	-1,931	0,302	2,338	0,763	-1,453	-0,216	-1,521	<b>4,800</b>

Al observar la diagonal principal hay dos valores mayores a 10 con lo cual existe problemas de multicolinealidad.

Por último, se realiza el análisis del índice de acondicionamiento.

Tabla 3.23. Autovalores de la matriz de correlación

Número	Autovalor
1	3,496
2	1,794
3	1,097
4	0,863
5	0,335
6	0,275
7	0,113
8	0,026

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{3,496}{0,026}} = 11,635$$

Al realizar la fórmula del valor de acondicionamiento mayor a 10, con lo cual existe un problema de multicolinealidad moderada.

Para decidir que variable va a ser eliminada hay que observar la diagonal principal de la matriz inversa.



En este caso entre el mayor valor (permisos de conducir: 20,156) y el siguiente (población: 20,119) existe una diferencia menor a uno por lo que para eliminar una variable se debe utilizar el segundo criterio de eliminación de parámetros.

Para ello se realiza dos regresiones múltiples por cada uno de los gases establecidos ( $\text{SO}_2$ ,  $\text{CO}_2$  y  $\text{NO}_x$ )

Primero se realiza con la variable explicada dióxido de azufre ( $\text{SO}_2$ ):

- Regresión múltiple de dióxido de azufre ( $\text{SO}_2$ ) sin la variable población (P): R-cuadrada = 31,119%
- Regresión múltiple de dióxido de azufre ( $\text{SO}_2$ ) sin la variable permisos de conducir (PC): R-cuadrada = 32,473%

Después de realizar dos regresiones múltiples de decide continuar eliminar la variable Permisos de conducir, ya que este modelo obtiene un R-cuadrado mayor.

A continuación, se va a realizar el mismo estudio con la variable dióxido de carbono ( $\text{CO}_2$ ):

- Regresión múltiple de dióxido de carbono ( $\text{CO}_2$ ) sin la variable población (P): R-cuadrada = 38,535%
- Regresión múltiple de dióxido de carbono ( $\text{CO}_2$ ) sin la variable permisos de conducir (PC): R-cuadrado = 36,711%

Para le modelo de Dióxido de carbono se elimina la variable población.

Por último, se realizará la misma operación para los óxidos de nitrógeno ( $\text{NO}_x$ ).

- Regresión múltiple de óxidos de nitrógeno ( $\text{NO}_x$ ) sin la variable población (P): R-cuadrado = 35,989%
- Regresión múltiple de óxidos de nitrógeno ( $\text{NO}_x$ ) sin la variable permisos de conducir (PC): R-cuadrada = 37,938%

Para esta eliminamos la variable permisos de conducir, ya que al realizar la regresión múltiple de óxidos de nitrógeno sin la variable permisos de conducir su R-cuadrado es mayo.



A continuación, se realiza una tabla resumen de las variables que continúan en cada modelo:

Tabla 3.24. Cuadro resumen variables

SO <sub>2</sub>	CO <sub>2</sub>	NO <sub>x</sub>
Temperatura (TEMP)	Temperatura (TEMP)	Temperatura (TEMP)
Precipitaciones (LLU)	Precipitaciones (LLU)	Precipitaciones (LLU)
Horas de sol (SOL)	Horas de sol (SOL)	Horas de sol (SOL)
Viento (VI)	Viento (VI)	Viento (VI)
Población (P)	Ocupados (O)	Ocupados (O)
Ocupados (O)	Permisos de conducir (PC)	Población (P)
Autobuses (AU)	Autobuses (AU)	Autobuses (AU)

A partir de este momento el único cambio que se va a realizar es que las pruebas se harán respectivamente para cada uno de los gases ya que, para cada uno de ellos se ha eliminado una variable diferente.

- **Dióxido de azufre (SO<sub>2</sub>):**

Para eliminar el problema de multicolinealidad se van a realizar las tres pruebas, la primera de ellas es la matriz de correlación.

Tabla 3.25. Matriz de correlación

	TEMP	LLU	SOL	VI	P	O	AU
TEMP	1	-0,345	0,502	-0,266	0,664	-0,136	0,639
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,014	-0,113	0,053
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,186	0,080	-0,044
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	-0,048	0,068	-0,218
P	<b>0,664</b>	<b>-0,014</b>	<b>0,186</b>	<b>-0,048</b>	1	0,016	0,793
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	<b>0,016</b>	1	-0,033
AU	<b>0,639</b>	<b>0,053</b>	<b>-0,044</b>	<b>-0,218</b>	<b>0,793</b>	<b>-0,033</b>	1

Al observar los valores situados por debajo de la diagonal principal existen varios mayores de 0,7 en valor absoluto. Por este motivo se puede decir que existe problema de multicolinealidad.





A continuación, se realiza la matriz inversa de correlación.

Tabla 3.26. Matriz inversa de correlación

TEMP	<b>3,605</b>	0,200	-1,674	0,060	-0,591	0,589	-1,886
LLU	0,200	<b>2,198</b>	1,628	0,073	-0,763	0,168	0,454
SOL	-1,674	1,628	<b>3,455</b>	0,484	-1,229	-0,260	2,206
VI	0,060	0,073	0,484	<b>1,262</b>	-0,718	-0,070	0,820
P	-0,591	-0,763	-1,229	-0,718	<b>3,815</b>	-0,174	-2,821
O	0,589	0,168	-0,260	-0,070	-0,174	<b>1,120</b>	-0,237
AU	-1,886	0,454	2,206	0,820	-2,821	-0,237	<b>4,685</b>

Al analizar la diagonal principal no se observan valores mayores a 10, con lo cual no existe problema de multicolinealidad.

Por último, se realiza el índice de acondicionamiento.

Tabla 3.27. Autovalores de la matriz de correlación

Número	Autovalor
1	2,691
2	1,737
3	1,062
4	0,859
5	0,330
6	0,216
7	0,106

$$IC = \frac{\sqrt{\text{Autovalor máximo}}}{\sqrt{\text{Autovalor mínimo}}} = \frac{\sqrt{2,691}}{\sqrt{0,106}} = 5,034$$

El valor del índice de acondicionamiento no presenta problema de multicolinealidad.

Aunque la matriz inversa de correlación y el índice de acondicionamiento no presenten problemas de multicolinealidad se debe eliminar una de las variables porque la matriz de correlación si presenta el problema.



CAMPUS D'ALCOI

Al dirigirse a la matriz inversa de correlación la diferencia entre el mayor valor y el siguiente es menor a 1. Por este motivo se debe utilizar el segundo criterio para intentar eliminar el problema de multicolinealidad.

- Regresión múltiple de dióxido de azufre (SO<sub>2</sub>) sin la variable autobuses (AU): R-cuadrada = 29,128%
- Regresión múltiple de dióxido de azufre (SO<sub>2</sub>) sin la variable población (P): R-cuadrada = 23,462%

Se decide eliminar la variable autobuses, ya que al realizar la regresión múltiple del modelo eliminando esta variable tiene mayor R-cuadrado.

Al eliminar esta variable se deben realizar de nuevo las tres pruebas.

Primero se realiza la matriz de correlación:

Tabla 3.28. Matriz de correlación

	TEMP	LLU	SOL	VI	P	O
TEMP	1	-0,345	0,502	-0,266	0,664	-0,136
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,014	-0,113
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,186	0,080
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	-0,048	0,068
P	<b>0,664</b>	<b>-0,014</b>	<b>0,186</b>	<b>-0,048</b>	1	0,016
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	<b>0,016</b>	1

Al observar los valores por bajo de la diagonal principal, existe uno de ellos mayor a 0,7 en valor absoluto. Por este motivo se puede decir que existe un problema de multicolinealidad.

A continuación, se realiza la matriz inversa de correlación.

Tabla 3.29. Matriz inversa de correlación

TEMP	<b>2,845</b>	0,383	-0,786	0,391	-1,727	0,494
LLU	0,383	<b>2,154</b>	1,414	-0,006	-0,490	0,191
SOL	-0,786	1,414	<b>2,417</b>	0,097	0,100	-0,148
VI	0,391	-0,006	0,097	<b>1,118</b>	-0,224	-0,028
P	-1,727	-0,490	0,100	-0,224	<b>2,116</b>	-0,317
O	0,494	0,191	-0,148	-0,028	-0,317	<b>1,108</b>

Al analizar la diagonal principal no existen valores mayores a 10, con lo cual no existe problema de multicolinealidad.



Por último, se realiza el índice de acondicionamiento.

Tabla 3.30. Autovalores de la matriz de correlación

Número	Autovalor
1	2,356
2	1,292
3	1,045
4	0,826
5	0,266
6	0,216

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{2,356}{0,216}} = 3,305$$

El valor del índice de acondicionamiento es menor a 10 con lo cual, no existe problema de multicolinealidad.

Como en la matriz de correlación existe problema de multicolinealidad se debe eliminar una de las variables.

Al dirigirse la matriz inversa de correlación se observa que la diferencia entre el mayor valor y el siguiente es menor a uno, por ese motivo se debe utilizar el segundo criterio para eliminar el problema de multicolinealidad.

- Regresión múltiple de dióxido de azufre (SO<sub>2</sub>) sin la variable temperatura (TEMP): R-cuadrado = 28,354%
- Regresión múltiple de dióxido de azufre (SO<sub>2</sub>) sin la variable horas de sol (SOL): R-cuadrada = 28,665%

Se decide eliminar la variable horas de sol ya que, al realizar la regresión múltiple sin la variable horas de sol (SOL) obtenemos mayor R-cuadro.

De nuevo se vuelve a realizar las tres pruebas para intentar eliminar el problema de multicolinealidad.

La primera prueba es la matriz de correlación.

Tabla 3.31. Matriz de correlación

	TEMP	LLU	VI	P	O
TEMP	1	-0,345	-0,266	0,664	-0,136
LLU	<b>-0,345</b>	1	0,183	-0,014	-0,113
VI	<b>-0,266</b>	<b>0,183</b>	1	-0,048	0,068
P	<b>0,664</b>	<b>-0,014</b>	<b>-0,048</b>	1	0,016
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,068</b>	<b>0,016</b>	1



## CAMPUS D'ALCOI

Al observar los valores de bajo de la diagonal principal no existen valores mayores a 0,7 en valor absoluto. Por este motivo se puede decir que no existe problema de multicolinealidad.

A continuación, se realiza la matriz inversa de correlación.

Tabla 3.32. Matriz inversa de correlación

TEMP	<b>2,590</b>	0,843	0,422	-1,695	0,446
LLU	0,843	<b>1,326</b>	-0,063	-0,548	0,278
VI	0,422	-0,063	<b>1,114</b>	-0,228	-0,022
P	-1,695	-0,548	-0,228	<b>2,112</b>	-0,311
O	0,446	0,278	-0,022	-0,311	<b>1,098</b>

En la diagonal principal de la matriz de correlación no existe ningún valor mayor a 10, con lo cual no existe problema de multicolinealidad.

Por último, se realiza el índice de acondicionamiento.

Tabla 3.33. Autovalores de la matriz de correlación

Número	Autovalor
1	1,854
2	1,113
3	1,044
4	0,768
5	0,220

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{1,854}{0,220}} = 2,902$$

El índice de acondicionamiento no presenta problema de multicolinealidad.

A partir de estas últimas tres pruebas hemos comprobado que el modelo dióxido de azufre no presenta problemas de multicolinealidad. Las variables explicativas que permanecen en este modelo son: temperatura (TEMP), precipitaciones (LLU), viento (VI), población (P) y ocupados (O).



• **Dióxido de carbono (CO<sub>2</sub>):**

En este apartado se van a realizar las tres pruebas para eliminar el problema de multicolinealidad en el modelo de dióxido de carbono.

La primera prueba es la matriz de correlación.

Tabla 3.34. Matriz de correlación

	TEMP	LLU	SOL	VI	O	PC	AU
TEMP	1	-0,345	0,502	-0,266	-0,136	0,703	0,639
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,113	-0,169	0,053
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,080	0,317	-0,044
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	0,068	-0,127	-0,218
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	1	0,046	-0,033
PC	<b>0,703</b>	<b>-0,169</b>	<b>0,317</b>	<b>-0,127</b>	<b>0,046</b>	1	0,768
AU	<b>0,639</b>	<b>0,053</b>	<b>-0,044</b>	<b>-0,218</b>	<b>-0,033</b>	<b>0,768</b>	1,000

Al analizar los valores de bajo de la diagonal principal se encuentran varios mayores a 0,7 en valor absoluto. Por este motivo se puede decir que existe un problema de multicolinealidad.

A continuación, se realiza la matriz inversa de correlación.

Tabla 3.35. Matriz inversa de correlación

TEMP	<b>3,560</b>	0,116	-1,706	0,004	0,586	-0,422	-2,011
LLU	0,116	<b>2,070</b>	1,498	-0,030	0,150	-0,307	0,117
SOL	-1,706	1,498	<b>3,601</b>	0,441	-0,237	-1,439	2,363
VI	0,004	-0,030	0,441	<b>1,193</b>	-0,075	-0,501	0,661
O	0,586	0,150	-0,237	-0,075	<b>1,123</b>	-0,211	-0,210
PC	-0,422	-0,307	-1,439	-0,501	-0,211	<b>3,822</b>	-2,830
AU	-2,011	0,117	2,363	0,661	-0,210	-2,830	<b>4,695</b>

Los valores de la diagonal principal son menores a 10 con lo cual no existe problema de multicolinealidad.



Por último, se realiza el índice de acondicionamiento.

Tabla 3.36. Autovalores de la matriz de correlación

Número	Autovalor
1	2,810
2	1,623
3	1,056
4	0,852
5	0,324
6	0,231
7	0,105

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{2,810}{0,105}} = 5,182$$

El valor del índice de acondicionamiento es menor a 10. Por este no existe problema de multicolinealidad.

Se debe eliminar una de las variables ya que existe el problema de multicolinealidad en la matriz de correlación. Al dirigirse a la matriz inversa de correlación la diferencia entre los dos valores mayores es menor a uno, con lo cual debe utilizarse el segundo el segundo criterio.

- Regresión múltiple de dióxido de carbono (CO<sub>2</sub>) sin la variable autobuses (AU): R-cuadrada = 38,407%
- Regresión múltiple de dióxido de carbono (CO<sub>2</sub>) sin la variable permisos de conducir (PC): R-cuadrada = 35,937%

Se decide eliminar la variable autobuses porque al realizar la regresión múltiple de este modelo obtiene mayor R-cuadrado.

A continuación, se repiten las tres pruebas eliminando la variable autobuses. La primera prueba es la matriz de correlación.

Tabla 3.37. Matriz de correlación

	TEMP	LLU	SOL	VI	O	PC
TEMP	1	-0,345	0,502	-0,266	-0,136	0,703
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,113	-0,169
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,080	0,317
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	0,068	-0,127
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	1	0,046
PC	<b>0,703</b>	<b>-0,169</b>	<b>0,317</b>	<b>-0,127</b>	<b>0,046</b>	1



CAMPUS D'ALCOI

Al analizar los valores por bajo de la diagonal principal existe uno mayor a 0,7 en valor absoluto. Por ese motivo existe un problema de multicolinealidad.

Después, se realiza la matriz inversa de correlación.

Tabla 3.38. Matriz inversa de correlación

TEMP	<b>2,698</b>	0,166	-0,694	0,288	0,496	-1,634
LLU	0,166	<b>2,067</b>	1,439	-0,047	0,155	-0,237
SOL	-0,694	1,439	<b>2,412</b>	0,109	-0,131	-0,014
VI	0,288	-0,047	0,109	<b>1,100</b>	-0,045	-0,103
O	0,496	0,155	-0,131	-0,045	<b>1,114</b>	-0,338
PC	-1,634	-0,237	-0,014	-0,103	-0,338	<b>2,115</b>

Los valores de la matriz principal no presentan problema de multicolinealidad.

Por último, se realiza el índice de condicionamiento.

Tabla 3.39. Autovalores de la matriz de correlación

Número	Autovalor
1	2,499
2	1,181
3	0,992
4	0,827
5	0,272
6	0,229

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{2,499}{0,229}} = 3,302$$

El valor del índice de condicionamiento es menor a 10, por este motivo se puede decir que no existe problema de multicolinealidad.

Se debe eliminar una de las variables que da problemas. Al observar la matriz inversa de correlación la diferencia de los valores mayores es menor a 1 con lo cual se utiliza el segundo criterio.

- Regresión múltiple de dióxido de carbono (CO<sub>2</sub>) sin la variable temperatura (TEMP): R-cuadrada = 35,556%
- Regresión múltiple de dióxido de carbono (CO<sub>2</sub>) sin la variable horas de sol (SOL): R-cuadrada = 31,783%



CAMPUS D'ALCOI

Se decide eliminar la variable temperatura ya que al realizar la regresión consigue mayor R-cuadrado.

A continuación, se realizan de nuevos las tres pruebas para saber si se ha eliminado el problema de multicolinealidad.

La primera de las pruebas es la matriz de correlación.

Tabla 3.40. Matriz de correlación

	LLU	SOL	VI	O	PC
LLU	1,000	-0,711	0,183	-0,113	-0,169
SOL	<b>-0,711</b>	1,000	-0,227	0,080	0,317
VI	<b>0,183</b>	<b>-0,227</b>	1,000	0,068	-0,127
O	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	1,000	0,046
PC	<b>-0,169</b>	<b>0,317</b>	<b>-0,127</b>	<b>0,046</b>	1,000

Al analizar los valores por debajo de la diagonal principal existe uno mayor a 0,7 en valor absoluto. Por este motivo podemos decir que existe problemas de multicolinealidad.

A continuación, se realiza la matriz inversa de correlación.

Tabla 3.41. Matriz inversa de correlación

LLU	<b>2,057</b>	1,482	-0,064	0,125	-0,136
SOL	1,482	<b>2,234</b>	0,183	-0,004	-0,434
VI	-0,064	0,183	<b>1,069</b>	-0,098	0,071
O	0,125	-0,004	-0,098	<b>1,023</b>	-0,037
PC	-0,136	-0,434	0,071	-0,037	<b>1,125</b>

Los valores de la matriz principal son menores a 10 por ese motivo se puede decir que no existe problema de multicolinealidad.

Por último, se realiza el índice de condicionamiento.

Tabla 3.42. Autovalores de la matriz de correlación

Número	Autovalor
1	1,970
2	1,062
3	0,876
4	0,822
5	0,270





CAMPUS D'ALCOI

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{1,970}{0,270}} = 2,700$$

El índice de acondicionamiento es menor a 10 con lo cual no existe problema de multicolinealidad.

Aunque la matriz inversa de correlación y el índice de condicionamiento no presenten problemas de multicolinealidad se debe eliminar una de las variables ya que la matriz de correlación si presenta el problema.

La diferencia entre los valores más altos de la diagonal principal es menor a uno, por tanto, para realizar la eliminación de variable se debe utilizar el segundo criterio.

- Regresión múltiple de dióxido de carbono (CO<sub>2</sub>) sin la variable horas de sol (SOL): R-cuadrado = 30,728%
- Regresión múltiple de dióxido de carbono (CO<sub>2</sub>) sin la variable precipitaciones (LLU): R-cuadrado = 34,433%

Se decide eliminar la variable precipitaciones (LLU), ya que al realizar la regresión múltiple sin esta se obtiene mayor R-cuadrado.

A continuación, se vuelve a realizar las tres pruebas sin la variable precipitaciones.

La primera prueba es la matriz de correlación.

Tabla 3.43. Matriz de correlación

	SOL	VI	O	PC
SOL	1	-0,227	0,080	0,317
VI	<b>-0,227</b>	1	0,068	-0,127
O	<b>0,080</b>	<b>0,068</b>	1	0,046
PC	<b>0,317</b>	<b>-0,127</b>	<b>0,046</b>	1

Al analizar los valores de bajo de la diagonal principal no existe ningún valor mayor a 0,7 en valor absoluto. Por este motivo se puede decir que no existe problema de multicolinealidad.

A continuación, se realiza la matriz inversa de correlación.

Tabla 3.44. Matriz inversa de correlación

SOL	<b>1,166</b>	0,229	-0,094	-0,336
VI	0,229	<b>1,067</b>	-0,094	0,067
O	-0,094	-0,094	<b>1,015</b>	-0,029
PC	-0,336	0,067	-0,029	<b>1,116</b>

Al observar la matriz principal no existe problema de multicolinealidad.



Por último, se realiza el índice de condicionamiento:

Tabla 3.45. Autovalores de la matriz de correlación

Número	Autovalor
1	1,461
2	1,051
3	0,834
4	0,653

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{1,461}{0,653}} = 1,495$$

El valor del índice de acondicionamiento es menor a 10 con lo cual no existe problema de multicolinealidad.

Después de realizar estas tres pruebas se ha comprobado que ha desaparecido el problema de multicolinealidad.

Por tanto, el modelo de dióxido de carbono (CO<sub>2</sub>) mantendrá las variables horas de sol (SOL), viento (VI), ocupados (O) y permisos de conducir (PC).



• **Óxidos de nitrógeno (NO<sub>x</sub>):**

Se van a realizar las tres pruebas para intentar eliminar el problema de multicolinealidad en el modelo de óxidos de nitrógeno (NO<sub>x</sub>).

La primera prueba es la matriz de correlación.

Tabla 3.46. Matriz de correlación

	TEMP	LLU	SOL	VI	P	O	AU
TEMP	1	-0,345	0,502	-0,266	0,664	-0,136	0,639
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,014	-0,113	0,053
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,186	0,080	-0,044
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	-0,048	0,068	-0,218
P	<b>0,664</b>	<b>-0,014</b>	<b>0,186</b>	<b>-0,048</b>	1	0,016	0,793
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	<b>0,016</b>	1	-0,033
AU	<b>0,639</b>	<b>0,053</b>	<b>-0,044</b>	<b>-0,218</b>	<b>0,793</b>	<b>-0,033</b>	1

Al observar los valores de bajo de la diagonal principal existen varios mayores a 0,7 en valor absoluto. Por este motivo se puede decir que existe problema de multicolinealidad.

A continuación, se realiza la matriz inversa de correlación.

Tabla 3.47. Matriz inversa de correlación

TEMP	<b>3,605</b>	0,200	-1,674	0,060	-0,591	0,589	-1,886
LLU	0,200	<b>2,198</b>	1,628	0,073	-0,763	0,168	0,454
SOL	-1,674	1,628	<b>3,455</b>	0,484	-1,229	-0,260	2,206
VI	0,060	0,073	0,484	<b>1,262</b>	-0,718	-0,070	0,820
P	-0,591	-0,763	-1,229	-0,718	<b>3,815</b>	-0,174	-2,821
O	0,589	0,168	-0,260	-0,070	-0,174	<b>1,120</b>	-0,237
AU	-1,886	0,454	2,206	0,820	-2,821	-0,237	<b>4,685</b>

Al analizar la diagonal principal de la matriz inversa de correlación no existe problema de multicolinealidad.



Por último, se realiza el índice de condicionamiento.

Tabla 3.48. Autovalores de la matriz de correlación.

Número	Autovalor
1	2,691
2	1,737
3	1,062
4	0,859
5	0,330
6	0,216
7	0,106

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{2,691}{0,106}} = 5,034$$

El índice de correlación no presenta problema de multicolinealidad.

Como en la matriz de correlación existe problema de multicolinealidad se debe eliminar una de las variables. Al observar los valores de la matriz inversa de correlación la diferencia del mayor valor y el siguiente es menor a uno con lo cual, se debe eliminar una variable a partir del segundo criterio.

- Regresión múltiple de óxidos de nitrógeno (NO<sub>x</sub>) sin la variable autobuses (AU): R-cuadrada = 35,975%
- Regresión múltiple de óxidos de nitrógeno (NO<sub>x</sub>) sin la variable población (P): R-cuadrada = 24,833%

Se decide eliminar la variable autobuses, ya que al realizar la regresión múltiple sin esta variable obtiene mayor R-cuadrado.

A continuación, se vuelven a realizar los tres métodos de análisis.

El primer método es la matriz de correlación.

Tabla 3.49. Matriz de correlación

	TEMP	LLU	SOL	VI	P	O
TEMP	1	-0,345	0,502	-0,266	0,664	-0,136
LLU	<b>-0,345</b>	1	-0,711	0,183	-0,014	-0,113
SOL	<b>0,502</b>	<b>-0,711</b>	1	-0,227	0,186	0,080
VI	<b>-0,266</b>	<b>0,183</b>	<b>-0,227</b>	1	-0,048	0,068
P	<b>0,664</b>	<b>-0,014</b>	<b>0,186</b>	<b>-0,048</b>	1	0,016
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,080</b>	<b>0,068</b>	<b>0,016</b>	1



CAMPUS D'ALCOI

Al analizar los valores por bajo de la matriz principal se observa uno de ellos mayor a 0,7 en valor absoluto con lo cual existe problema de multicolinealidad.

Posteriormente se realiza la matriz inversa de correlación.

Tabla 3.50. Matriz inversa de correlación

TEMP	<b>2,845</b>	0,383	-0,786	0,391	-1,727	0,494
LLU	0,383	<b>2,154</b>	1,414	-0,006	-0,490	0,191
SOL	-0,786	1,414	<b>2,417</b>	0,097	0,100	-0,148
VI	0,391	-0,006	0,097	<b>1,118</b>	-0,224	-0,028
P	-1,727	-0,490	0,100	-0,224	<b>2,116</b>	-0,317
O	0,494	0,191	-0,148	-0,028	-0,317	<b>1,108</b>

Al observar la diagonal principal de la matriz inversa de correlación se puede decir que no existe problema de multicolinealidad.

Por último, se realiza el índice de acondicionamiento.

Tabla 3.51. Autovalores de la matriz de correlación

Número	Autovalor
1	2,356
2	1,292
3	1,045
4	0,826
5	0,266
6	0,216

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{2,356}{0,216}} = 3,305$$

El valor el índice de acondicionamiento es menor a 10 con lo cual no existe problema de multicolinealidad.

Como el primer método tiene problema de multicolinealidad se debe eliminar una de las variables problemáticas. Al dirigirse a la diagonal principal de la matriz inversa, la diferencia entre los dos valores mayores es menor a uno con lo cual se realizarse la eliminación de una variable a partir del segundo criterio.

- Regresión múltiple de óxidos de nitrógeno (NO<sub>x</sub>) sin la variable temperatura (TEMP): R-cuadrada = 33,926%
- Regresión múltiple de óxidos de nitrógeno (NO<sub>x</sub>) sin la variable horas de sol (SOL): R-cuadrada = 35,890%



CAMPUS D'ALCOI

Se decide eliminar la variable horas de sol, ya que al realizar la regresión múltiple eliminando esta variable se obtiene mayor R-cuadrado.

De nuevo se vuelven a realizar los métodos para intentar eliminar el problema de multicolinealidad.

El primer método a realizar es la matriz de correlación.

Tabla 3.52. Matriz de correlación

	TEMP	LLU	VI	P	O
TEMP	1	-0,345	-0,266	0,664	-0,136
LLU	<b>-0,345</b>	1	0,183	-0,014	-0,113
VI	<b>-0,266</b>	<b>0,183</b>	1	-0,048	0,068
P	<b>0,664</b>	<b>-0,014</b>	<b>-0,048</b>	1	0,016
O	<b>-0,136</b>	<b>-0,113</b>	<b>0,068</b>	<b>0,016</b>	1

Al analizar los valores inferiores a la diagonal principal no existen ninguno mayor a 0,7 en valor absoluto. Por este motivo se puede decir que no existe problema de multicolinealidad.

Posteriormente, se realiza la matriz inversa de correlación.

Tabla 3.53. Matriz inversa de correlación

TEMP	<b>2,590</b>	0,843	0,422	-1,695	0,446
LLU	0,843	<b>1,326</b>	-0,063	-0,548	0,278
VI	0,422	-0,063	<b>1,114</b>	-0,228	-0,022
P	-1,695	-0,548	-0,228	<b>2,112</b>	-0,311
O	0,446	0,278	-0,022	-0,311	<b>1,098</b>

Al observar los valores de la diagonal principal no se observan problema de multicolinealidad.

Por último, se realiza el índice de acondicionamiento.

Tabla 3.54. Autovalores de la matriz de correlación

Número	Autovalor
1	1,854
2	1,113
3	1,044
4	0,768
5	0,220



CAMPUS D'ALCOI

$$IC = \sqrt{\frac{\text{Autovalor máximo}}{\text{Autovalor mínimo}}} = \sqrt{\frac{1,854}{0,220}} = 2,902$$

El valor del índice de acondicionamiento es menor a 10, con lo cual no existe problema de multicolinealidad.

Al realizar los tres métodos se ha eliminado el problema de multicolinealidad. El modelo de óxidos de nitrógeno (NO<sub>x</sub>) permanecerá con las variables temperatura (TEMP), lluvia (LLU), viento (VI), población (P) y ocupados (O).

Después de eliminar el problema de multicolinealidad de los tres gases se va a realizar un cuadro resumen donde observar que variables explicativas que permanecen en cada uno de ellos.

Tabla 3.55. Resumen de variables sin multicolinealidad

<b>Dióxido de azufre (SO<sub>2</sub>)</b>	<b>Dióxido de carbono (CO<sub>2</sub>)</b>	<b>Óxidos de nitrógeno (NO<sub>x</sub>)</b>
Temperatura (TEMP)	Horas de sol (SOL)	Temperatura (TEMP)
Precipitaciones (LLU)	Viento (VI)	Lluvia (LLU)
Viento (VI)	Ocupados (O)	Viento (VI)
Población (P)	Permisos conducir (PC)	Población (P)
Ocupados (O)		Ocupados (O)



## Análisis multivariable:

Se va a realizar el estudio multivariante de los tres modelos propuestos: dióxido de azufre (SO<sub>2</sub>), dióxido de carbono (CO<sub>2</sub>) y óxidos de nitrógeno (NO<sub>x</sub>).

- **Dióxido de azufre (SO<sub>2</sub>):**

- **Modelo teórico propuesto:**

$$SO_2 = \beta_0 + \beta_1 \text{TEMP} + \beta_2 \text{LLU} + \beta_3 \text{VI} + \beta_4 \text{P} + \beta_5 \text{O} + U$$

- **Interpretación de los parámetros y unidades físicas:**

$\beta_0$  = Valor medio de dióxido de azufre (SO<sub>2</sub>) cuando las variables explicativas valen 0.

$$\frac{\mu\text{g}}{\text{m}^3}$$

$\beta_1$  = Incremento medio de dióxido de azufre (SO<sub>2</sub>) cuando aumentamos un grado centígrado la temperatura (TEMP) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu\text{g}}{\text{m}^3 \times ^\circ\text{C}}$$

$\beta_2$  = Incremento medio de dióxido de azufre (SO<sub>2</sub>) cuando aumentamos un milímetro las precipitaciones (LLU) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu\text{g}}{\text{m}^3 \times \frac{\text{mm}}{\text{mes}}}$$

$\beta_3$  = Incremento medio de dióxido de azufre (SO<sub>2</sub>) cuando aumentamos un kilómetro / hora el viento (VI) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu\text{g}}{\text{m}^3 \times \frac{\text{km}}{\text{hora}}}$$

$\beta_4$  = Incremento medio de dióxido de azufre (SO<sub>2</sub>) cuando aumentamos una unidad la población (P) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu\text{g}}{\text{m}^3 \times \text{miles de personas}}$$

$\beta_5$  = Incremento medio de dióxido de azufre (SO<sub>2</sub>) cuando aumentamos un el número de ocupados (O) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu\text{g}}{\text{m}^3 \times \text{miles de ocupados}}$$





○ **Primera estimación del modelo ajustado:**

$$SO_2 = 3,47181 + 0,095101*TEMP + 0,000698311*LLU - 0,0498212*VI + 0,00000179301*P + 0,000856404*O \longrightarrow R\text{-cuadrado: } 28,665\%$$

○ **Análisis de significatividad:**

Tabla 3.56. Valores significatividad variables explicativas

Parámetro	Estimación	Error Estándar	Estadístico T	Valor-P
CONSTANTE	3,47181	3,42326	1,01418	0,3167
TEMP	0,095101	0,181671	0,52348	<b>0,6036</b>
LLU	0,000698311	0,0113882	0,0613189	0,9514
VI	-0,0498212	0,0313155	-1,59095	0,1197
P	0,00000179301	0,00000100037	1,79236	0,0808
O	0,000856404	0,000652868	1,31176	0,1973

**Análisis de Varianza**

Tabla 3.57. Valor significatividad del modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	66,6942	5	13,3388	3,13	<b>0,0179</b>
Residuo	165,973	39	4,25571		
Total (Corr.)	232,667	44			

▪ Contraste significación global (modelo):

$$\left. \begin{array}{l} H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0 \\ H_1: \text{Algún } \beta_i \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,0179 < \alpha$ : Se rechaza la hipótesis nula y consideramos que el modelo sí que es significativo.

▪ Contraste significación individual (constante):

$$\left. \begin{array}{l} H_0: \beta_0 = 0 \\ H_1: \beta_0 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,3167 > \alpha$ . Se acepta la hipótesis nula y concluimos que la constante no es significativa.



CAMPUS D'ALCOI

- Contraste significación individual (temperatura):  
 $H_0: \beta_1 = 0$   
 $H_1: \beta_1 \neq 0$   
}  $\alpha=0,05$   
P-valor 0,6036 >  $\alpha$ . Se acepta la hipótesis nula y concluimos que la variable temperatura (TEMP) no resulta significativa.
- Contraste significación individual (precipitaciones):  
 $H_0: \beta_2 = 0$   
 $H_1: \beta_2 \neq 0$   
}  $\alpha=0,05$   
P-valor 0,9514 >  $\alpha$ . Se acepta la hipótesis nula y concluimos que la variable precipitaciones (LLU) no resulta significativa.
- Contraste significación individual (viento):  
 $H_0: \beta_3 = 0$   
 $H_1: \beta_3 \neq 0$   
}  $\alpha=0,05$   
P-valor 0,1197 >  $\alpha$ . Se acepta la hipótesis nula y que la variable viento (VI) no resulta significativa.
- Contraste significación individual (población):  
 $H_0: \beta_4 = 0$   
 $H_1: \beta_4 \neq 0$   
}  $\alpha=0,05$   
P-valor 0,0808 >  $\alpha$ . Se acepta la hipótesis nula y que la variable población (P) no resulta significativa.
- Contraste significación individual (ocupados):  
 $H_0: \beta_5 = 0$   
 $H_1: \beta_5 \neq 0$   
}  $\alpha=0,05$   
P-valor 0,1973 >  $\alpha$ . Se acepta la hipótesis nula y que la variable ocupados (O) no resulta significativa.



- **Análisis de normalidad:**
  - **Gráfico probabilístico:**

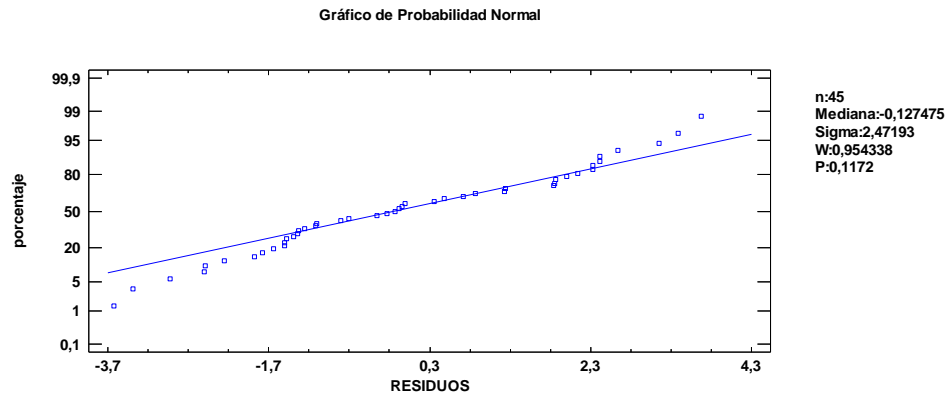


Gráfico 3.21. Gráfico probabilístico, dióxido de azufre

La mayoría de los residuos están muy cerca de la línea de tendencia, aunque existen algunas excepciones. Se puede decir que los residuos parecen tener una distribución normal.

- **Histograma de residuos:**

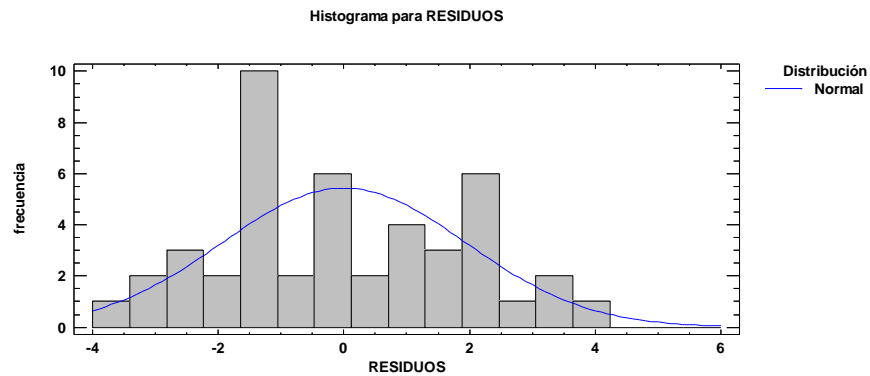


Gráfico 3.22. Histograma dióxido de azufre

En el histograma no se aprecia la forma de Campana de Gauss ya que la mayoría de residuos se encuentran en el centro, por esto se parece que no tienen un comportamiento normal.



▪ **Test de normalidad:**

**Pruebas de Normalidad para RESIDUOS**

Tabla 3.58. Pruebas de normalidad

Prueba	Estadístico	Valor-P
Chi-Cuadrado	13,0	0,602298
Estadístico W de Shapiro-Wilk	0,954338	0,117173
Valor-Z para asimetría	0,263795	0,791933
Valor-Z para curtosis	-2,14905	<b>0,03163</b>

$$\left. \begin{array}{l} H_0: \text{Distribución normal} \\ H_1: \neq \text{Distribución normal} \end{array} \right\} \alpha: 0,05$$

0,03163 <  $\alpha$ . Se rechaza la hipótesis nula, por tanto, el test concluye que los residuos no son normales.

○ **Análisis de heterocedasticidad:**

$$\text{Residuos}^2 = \beta_0 + \beta_1 \text{TEMP} + \beta_2 \text{LLU} + \beta_3 \text{VI} + \beta_4 \text{P} + \beta_5 \text{O} + U$$

Tabla 3.59. Estadísticos residuos al cuadrado.

Parámetro	Estimación	Error		Valor-P
		Estándar	T	
CONSTANTE	10,0847	6,06722	1,66216	0,1045
TEMP	-0,186872	0,321985	-0,580374	<b>0,5650</b>
LLU	-0,00745941	0,0201839	-0,369572	0,7137
VI	-0,0886864	0,0555021	-1,59789	0,1181
P	0,00000210731	0,00000177301	1,18855	0,2418
O	-0,00156122	0,00115711	-1,34924	0,1850

▪ Contraste de hipótesis (temperatura):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,5650 >  $\alpha$ . Se acepta la hipótesis nula. La variable temperatura no presenta problema de heterocedasticidad.



CAMPUS D'ALCOI

- Contraste de hipótesis (precipitaciones)

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{No constante} \end{array} \right\} \alpha = 0,05$$

P-valor:  $0,7137 > \alpha$ . Se acepta la hipótesis nula. La variable precipitaciones no presenta problema de heterocedasticidad.

- Contraste de hipótesis (viento):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor:  $0,1181 > \alpha$ . Se acepta la hipótesis nula. La variable viento no presenta problema de heterocedasticidad.

- Contraste de hipótesis (Población):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor:  $0,2418 > \alpha$ . Se acepta la hipótesis nula. No existe problema de heterocedasticidad.

- Contraste de hipótesis (Ocupados):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor:  $0,1850 > \alpha$ . Se acepta la hipótesis nula. No existe problema de heterocedasticidad.

○ **Autocorrelación:**

- **Test de Durbin - Watson:**

Este test sólo va a comprobar si existe autocorrelación de primer orden

Estadístico Durbin-Watson =  $2,14149$  ( $P=0,7329$ )

$$\left. \begin{array}{l} H_0: \rho_1 = 0 \text{ no hay autocorrelación} \\ H_1: \rho_1 \neq 0 \text{ hay autocorrelación} \end{array} \right\} \alpha = 0,05$$

P-valor  $0,7329 > \alpha$ . se acepta la hipótesis nula y se concluye que el modelo no tiene autocorrelación de primer orden.



▪ **Gráfico de Durbin - Watson:**

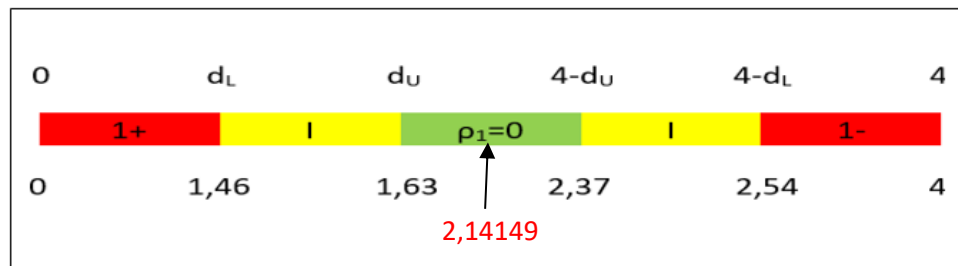


Gráfico 3.23. Gráfico de Durbin-Whatson

Al ubicar el estadístico de Durbin-Whatson en su gráfico se observa que no existe autocorrelación de primer orden.

▪ **FAS y FAP:**

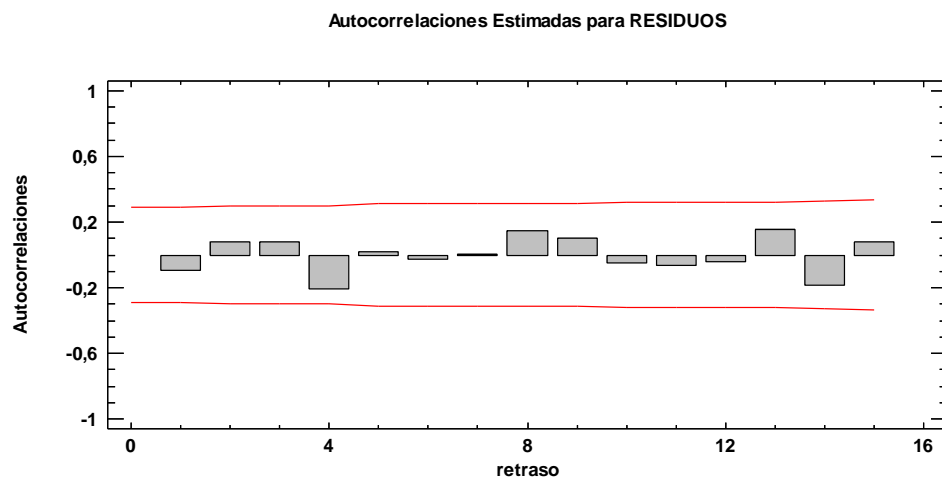


Gráfico 3.24. FAS



Autocorrelaciones Parciales Estimadas para RESIDUOS

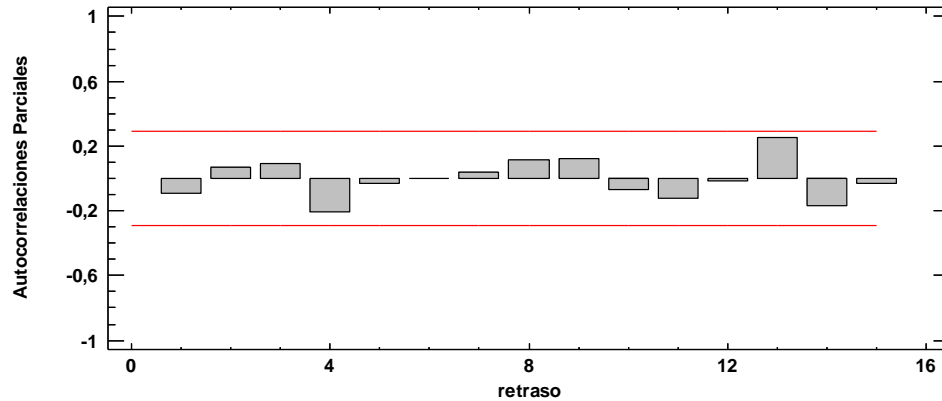


Gráfico 3.25. FAB

Al observar los gráficos del FAS y el FAP no se observa ningún factor de correlación que supere los límites establecidos, por esto se podría decir que no existe problema de autocorrelación.

○ **Puntos influyentes:**

Tabla 3.60. Puntos influyentes

Fila	Provincia	Influencia	Distancia de Mahalanobis	DFITS
3	Alicante	0,268979	14,8445	-1,21193
4	Almería	0,189587	9,08209	0,843026
27	Murcia	0,633253	73,2696	0,903666

Influencia media de un solo punto = 0,133333

▪ Puntos influyes a priori:

$$n_{ii} > 2 \bar{n} \longrightarrow n_{ii} > 0.2666666$$

$0,268979 < 2 \bar{n}$ : Alicante no es un punto influyente a priori.

$0,189587 < 2 \bar{n}$ : Almería no es un punto influyente a priori.

$0,633253 < 2 \bar{n}$ : Murcia no es un punto influyente a priori.

▪ Puntos influyentes a posteriori:

$$|DFITS| \geq 2 \sqrt{\bar{n}} \longrightarrow |DFITS| \geq 0,703$$

➤  $1,21193 \geq 2 \sqrt{\bar{n}}$  : Alicante es un punto influyente a posteriori.

➤  $0,843026 \geq 2 \sqrt{\bar{n}}$  : Almería es un punto influyente a posteriori.

➤  $0,903666 \geq 2 \sqrt{\bar{n}}$  : Murcia es un punto influyente a posteriori.



## CAMPUS D'ALCOI

- **Puntos anómalos:**

Para este modelo no existen puntos anómalos.

- **Resolución de problemas:**

Se procede a eliminar las variables no significativas con el método de selección hacia delante que consiste en introducir las variables según su significatividad.

Tabla 3.61. Estadísticos significatividad variables

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
CONSTANTE	2,81549	0,580234	4,85233	0,0000
P	0,00000224828	6,96344E-7	3,22869	0,0024

### Análisis de Varianza

Tabla 3.62. Estadísticos significatividad modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	45,399	1	45,399	10,42	0,0024
Residuo	187,268	43	4,35507		
Total (Corr.)	232,667	44			

R-cuadrada = 19,5124 por ciento

R-cuadrado (ajustado para g.l.) = 17,6406 por ciento

$$SO_2 = 2,81549 + 0,00000224828 * P$$

Al realizar el método de selección hacia delante solamente queda la variable población como significativa. Este modelo obtiene un R-cuadrado ajustado del 17,6406%.

Podemos apreciar que la cantidad de dióxido de azufre está relacionado directamente con el tamaño de la población.





• **Dióxido de carbono (CO<sub>2</sub>):**

○ **Modelo teórico propuesto:**

$$CO_2 = \beta_0 + \beta_1 SOL + \beta_2 VI + \beta_3 O + \beta_4 PC + U$$

○ **Interpretación de los parámetros y unidades físicas:**

$\beta_0$  = Valor medio de dióxido de carbono (CO<sub>2</sub>) cuando las variables explicativas valen 0.

$$\frac{Kt CO_2}{año}$$

$\beta_1$  = Incremento medio de dióxido de carbono (CO<sub>2</sub>) cuando aumentamos una hora de sol (SOL) y el resto de las variables explicativas pertenecen constantes.

$$\frac{Kt CO_2}{año \times hora}$$

$\beta_2$  = Incremento medio de dióxido de carbono (CO<sub>2</sub>) cuando aumentamos un kilómetro/hora la velocidad del viento (VI) y el resto de las variables explicativas pertenecen constantes.

$$\frac{Kt CO_2}{año \times \frac{km}{hora}}$$

$\beta_3$  = Incremento medio de dióxido de carbono (CO<sub>2</sub>) cuando aumentamos una unidad los miles de parados (P) y el resto de las variables explicativas pertenecen constantes.

$$\frac{Kt CO_2}{año \times miles \ de \ personas}$$

$\beta_4$  = Incremento medio de dióxido de carbono (CO<sub>2</sub>) cuando aumentamos una unidad los permisos de conducir y el resto de las variables explicativas pertenecen constantes.

$$\frac{Kt CO_2}{año \times unidades}$$

○ **Primera estimación del modelo ajustado:**

$$CO_2 = 12342,5 + 865,988 * SOL + 83,3558 * VI + 16,2857 * O - 6,36246 * PC$$

○ **Análisis de significatividad:**

Tabla 3.63. Estadísticos significatividad variables

Parámetro	Estimación	Error Estándar	Estadístico T	Valor-P
CONSTANTE	12342,5	81868,9	0,15076	0,8809
SOL	865,988	273,42	3,16724	0,0029
VI	83,3558	1066,48	0,07816	0,9381
O	16,2857	21,8441	0,745539	0,4603
PC	-6,36246	1,5938	-3,992	0,0003



Análisis de Varianza

Tabla 3.64. Análisis significatividad modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	1,08283E11	4	2,70707E10	5,25	0,0017
Residuo	2,06194E11	40	5,15486E9		
Total (Corr.)	3,14477E11	44			

- Contraste significación global (modelo):

$$\left. \begin{array}{l} H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_1: \text{Algún } \beta_i \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,0017 < \alpha$ : Se rechaza la hipótesis nula y se considera que el modelo sí que es significativo.

- Contraste significación individual (constante):

$$\left. \begin{array}{l} H_0: \beta_0 = 0 \\ H_1: \beta_0 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,8809 > \alpha$ . Se acepta la hipótesis nula y se concluye que la constante no es significativa.

- Contraste significación individual (horas de sol):

$$\left. \begin{array}{l} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,0029 > \alpha$ . Se rechaza la hipótesis nula y se concluye que la variable horas de sol resulta significativa.

- Contraste significación individual (viento):

$$\left. \begin{array}{l} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,9381 > \alpha$ . Se acepta la hipótesis nula y se concluye que la variable viento no resulta significativa.



- Contraste significación individual (ocupados):

$$\left. \begin{array}{l} H_0: \beta_3 = 0 \\ H_1: \beta_3 \neq 0 \end{array} \right\} \alpha=0,05$$

P-valor 0,4603 >  $\alpha$ . Se acepta la hipótesis nula y se concluye que la variable ocupados no resulta significativa.

- Contraste significación individual (permisos de conducir):

$$\left. \begin{array}{l} H_0: \beta_4 = 0 \\ H_1: \beta_4 \neq 0 \end{array} \right\} \alpha=0,05$$

P-valor 0,0003 >  $\alpha$ . Se rechaza la hipótesis nula y se concluye que la variable permisos de conducir resulta significativa.

○ **Análisis de normalidad:**

- **Gráfico probabilístico:**

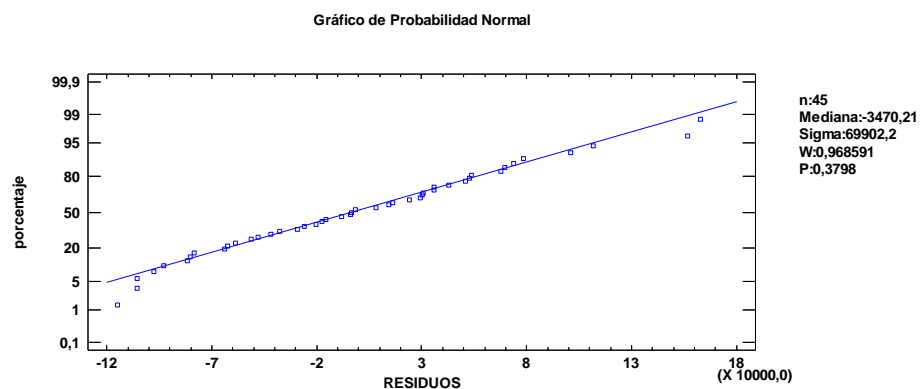


Gráfico 3.26. Gráfico probabilidad normal

La mayoría de los residuos están muy cerca de la línea de normalidad, aunque existen algunas excepciones. Se puede decir que los residuos parecen tener una distribución normal.



▪ **Histograma de residuos:**

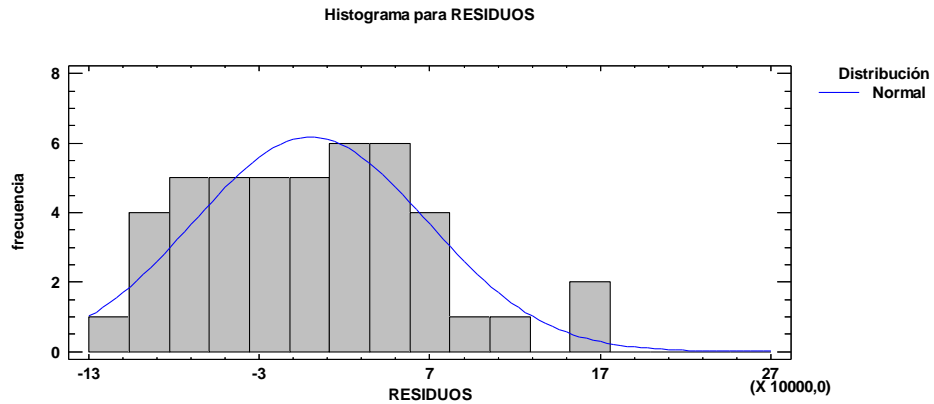


Gráfico 3.27. Histograma de residuos

En el histograma se aprecia la forma de Campana de Gauss ya que la mayoría de residuos se encuentran en el centro, por esto parece que los residuos tienen un comportamiento normal.

▪ **Test de normalidad:**

Para realizar el test de normalidad se utiliza la prueba con menor p-valor.

Tabla 3.65. Test de normalidad

Prueba	Estadístico	Valor-P
Chi-Cuadrado	8,2	0,915514
Estadístico W de Shapiro-Wilk	0,968591	0,379791
Valor-Z para asimetría	0,692775	0,488448
Valor-Z para curtosis	-0,254906	0,798792

$H_0$ : Distribución normal  
 $H_1$ :  $\neq$  Distribución normal

}  $\alpha$ : 0,05

$0,379791 > \alpha$ . Se acepta la hipótesis nula, por tanto, el test concluye que los residuos son normales.



○ **Análisis de heterocedasticidad:**

$$Residuos^2 = \beta_0 + \beta_1 SOL + \beta_2 VI + \beta_3 O + \beta_4 PC + U$$

Tabla 3.66. Estadísticos residuos al cuadrado

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
CONSTANTE	-8,19499E9	6,36089E9	-1,28834	0,2050
SOL	5,00485E7	2,12437E7	2,35592	<b>0,0235</b>
VI	1,14811E8	8,28611E7	1,38558	0,1736
O	-1,79601E6	1,6972E6	-1,05822	0,2963
PC	-243807,	123832,	-1,96885	0,0559

- Contraste de hipótesis (horas de sol):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,0235 <  $\alpha$ . Se rechaza la hipótesis nula. La variable horas de sol presenta problema de heterocedasticidad.

- Contraste de hipótesis (viento):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0.05$$

P-valor: 0,1736 >  $\alpha$ . Se acepta la hipótesis nula. La variable temperatura no presenta problema de heterocedasticidad.

- Contraste de hipótesis (ocupados):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,2963 >  $\alpha$ . Se acepta la hipótesis nula. La variable temperatura no presenta problema de heterocedasticidad.

- Contraste de hipótesis (permisos de conducir):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,0559 >  $\alpha$ . Se acepta la hipótesis nula. La variable temperatura no presenta problema de heterocedasticidad.



○ **Autocorrelación:**

▪ **Test de Durbin - Watson:**

Estadístico Durbin-Watson = 2,12477 (P=0,6777)

$H_0: \rho_1 = 0$  no hay autocorrelación }  
 $H_1: \rho_1 \neq 0$  hay autocorrelación }  $\alpha=0,05$

P-valor 0,6777 >  $\alpha$ . se acepta la hipótesis nula y se concluye que el modelo no tiene autocorrelación de primer orden.

▪ **Gráfico de Durbin - Watson:**

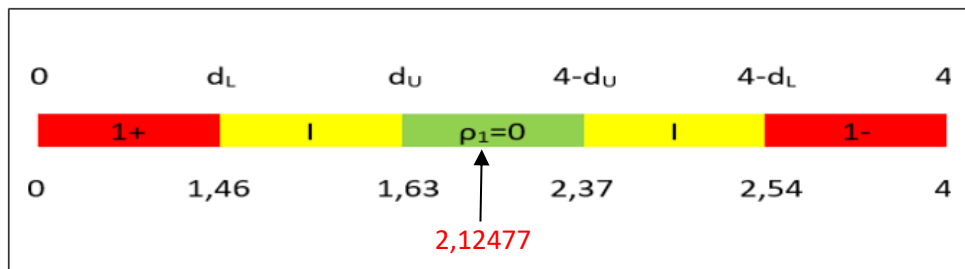


Gráfico 3.28. Gráfico de Durbin - Watson

Al ubicar el estadístico de Durbin-Watson en su gráfico se observa que no existe autocorrelación de primer orden.

▪ **FAS y FAP:**

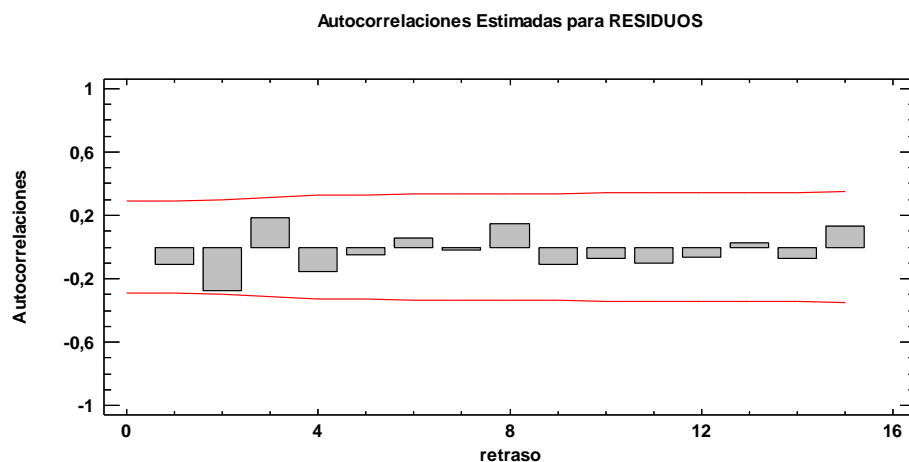


Gráfico 3.29. FAS

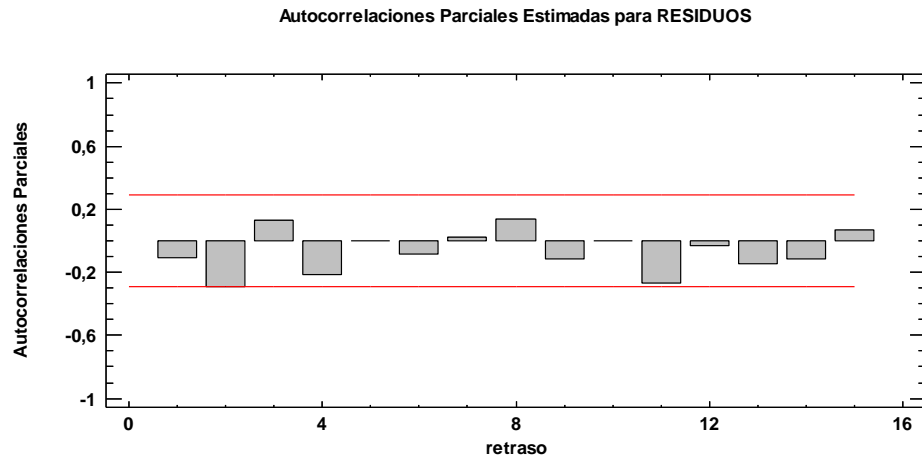


Gráfico 3.30. FAP

Al observar los gráficos del FAS y el FAP no se observa ningún factor de correlación que supere los límites establecidos, por esto se podría decir que no existe problema de autocorrelación.

○ **Puntos influyentes:**

Tabla 3.67. Puntos influyentes

Fila	Provincia	Influencia	Distancia de Mahalanobis	DFITS
8	Burgos	0,368646	24,1303	0,40042
27	Murcia	0,596078	62,4791	-0,0467257

Influencia media de un solo punto = 0,111111

- Puntos influyes a priori:  
 $n_{ii} > 2 \bar{n} \longrightarrow n_{ii} > 0,666$ 
  - $0,368646 < 2 \bar{n}$ : Burgos no es un punto influyente a priori.
  - $0,596078 < 2 \bar{n}$ : Murcia no es un punto influyente a priori.
  
- Puntos influyentes a posteriori:  
 $|DFITS| \geq 2 \sqrt{\bar{n}} \longrightarrow |DFITS| \geq 0,703$ 
  - $0,40042 \leq 2 \sqrt{\bar{n}}$ : Burgos no es un punto influyente a posteriori.
  - $0,0467257 \geq 2 \sqrt{\bar{n}}$ : Murcia no es un punto influyente a posteriori.



○ **Puntos anómalos:**

Tabla 3.68. Puntos anómalos

			Y		Residuo
Fila		Y	Predicha	Residuo	Estudentizado
6	Badajoz	329379,	166830,	162549,	2,47
9	Cáceres	368214,	211664,	156550,	2,39

Badajoz y Cáceres son considerados puntos anómalos.

○ **Resolución de problemas:**

Resolución de la heterocedasticidad.

$$\frac{CO_2}{SOL^c} = \frac{\beta_0}{SOL^c} + \beta_1 \frac{SOL}{SOL^c} + \beta_2 \frac{VI}{SOL^c} + \beta_3 \frac{O}{SOL^c} + \beta_4 \frac{PC}{SOL^c}$$

$$c = \frac{h}{2}$$

$$Residuos^2 = \beta_0 + \beta_1 SOL^h + U$$

Se realizan varias regresiones elevando la variable horas de sol (SOL) a h y nos quedaremos con la que tenga el R-cuadrado más alto. Este valor de h toma valores enteros positivos, enteros negativos y fraccionados entre 0 y 1.

Tabla 3.69. Regresiones elevando la variable horas de sol (SOL) a h

h	1	2	-1	-2	1/2	1/3	1/4	1/5
R <sup>2</sup>	4,40722	4,18352	4,56894	4,48885	4,48629	4,50722	4,51664	4,52194

h	1/6	1/7	1/8	1/9	1/10	1/11	1/12	1/13
R <sup>2</sup>	4,52533	4,52768	4,52941	4,53073	4,53177	4,53261	4,53331	4,53389

h	1/14	1/15	1/16	1/17	1/18	1/19	1/20	1/21
R <sup>2</sup>	4,53439	4,53481	4,53519	4,53551	4,5358	4,53606	4,53629	4,5365

h	1/22	1/23	1/24	1/25	1/26	1/27	1/28	1/29
R <sup>2</sup>	4,53669	4,53686	4,53702	4,53717	4,5373	4,5373	4,53754	4,53764

h	1/30	1/31	1/32	1/33	1/34	1/35	1/36	1/37
R <sup>2</sup>	4,53774	4,53783	4,53792	4,538	4,53808	4,53815	4,53822	4,53828





CAMPUS D'ALCOI

h	1/38	1/39	1/40	1/41	1/42	1/43	1/44	1/45
R <sup>2</sup>	4,53834	4,5384	4,53845	4,53851	4,53856	4,5386	4,53865	4,53869

h	1/46	1/47	1/48	1/49	1/50	1/51	1/52	1/53
R <sup>2</sup>	4,53873	4,53877	4,53881	4,53884	4,53888	4,53891	4,53894	4,53898

h	1/54	1/55	1/56	1/57	1/58	1/59	1/60	1/61
R <sup>2</sup>	4,539	4,53903	4,53906	4,53909	4,53911	4,53914	4,53916	4,53918

h	1/62	1/63	1/64	1/65	1/66	1/67	1/68	1/69
R <sup>2</sup>	4,53921	4,53923	4,53925	4,53927	4,53929	4,53931	4,53933	4,53934

h	1/70	1/71	1/72	1/73	1/74	1/75	1/76	1/77
R <sup>2</sup>	4,53936	4,53938	4,53939	4,53941	4,53943	4,53944	4,53946	4,53947

h	1/78	1/79	1/80	1/81	1/82	1/83	1/84	1/85
R <sup>2</sup>	4,53948	4,5395	4,53951	4,53952	4,53954	4,53955	4,53956	4,53957

h	1/86	1/87	1/88	1/89	1/90	1/91	1/92	1/93
R <sup>2</sup>	4,53958	4,5396	4,53961	4,53962	4,53963	4,53964	4,53965	4,53966

h	1/94	1/95	1/96	1/97	1/98	1/99	1/100	1/101
R <sup>2</sup>	4,53967	4,53968	4,53969	4,5397	4,5397	4,53971	4,53972	4,53973

h	1/102	1/103	1/104	1/105	1/106	1/107	1/108	1/109
R <sup>2</sup>	4,53974	4,53975	4,53975	4,53976	4,53977	4,53978	4,53978	4,53979

h	1/110	1/111	1/112	1/113	1/114	1/115	1/150	1/200
R <sup>2</sup>	4,5398	4,5398	4,53981	4,53982	4,53982	4,53983	4,54	4,54014

h	1/250	1/300	1/400	1/600	1/800	1/1000	1/1500	1/3000
R <sup>2</sup>	4,54022	4,54028	4,54034	4,54041	4,54045	4,54047	4,5405	4,54052

h	1/6000	1/9000	1/10000
R <sup>2</sup>	4,54054	4,54054	4,54054

Se escoge el valor h con mayor R – cuadrado, el cual es -1

A continuación, se calcula el valor c:

$$c = \frac{-1}{2} = -0,5$$



El modelo se queda de la siguiente forma:

$$\frac{CO_2}{SOL^{-0,5}} = \frac{\beta_0}{SOL^{-0,5}} + \beta_1 \frac{SOL}{SOL^{-0,5}} + \beta_2 \frac{VI}{SOL^{-0,5}} + \beta_3 \frac{O}{SOL^{-0,5}} + \beta_4 \frac{PC}{SOL^{-0,5}}$$

Tabla 3.70. Estadísticos significatividad variables

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
1/(SOL <sup>-0,5</sup> )	14603,2	87709,9	0,166494	0,8686
SOL/(SOL <sup>-0,5</sup> )	857,424	304,91	2,81205	0,0076
VI/(SOL <sup>-0,5</sup> )	60,5931	1102,2	0,0549749	0,9564
O/(SOL <sup>-0,5</sup> )	15,7842	21,565	0,731935	0,4685
PC/(SOL <sup>-0,5</sup> )	-6,28324	1,57592	-3,98703	0,0003

### Análisis de Varianza

Tabla 3.71. Estadísticos significatividad modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	2,36936E14	5	4,73871E13	40,37	0,0000
Residuo	4,69558E13	40	1,1739E12		
Total	2,83892E14	45			

$$CO_2/(SOL^{-0,5}) = 14603,2 * 1/(SOL^{-0,5}) + 857,424 * SOL/(SOL^{-0,5}) + 60,5931 * VI/(SOL^{-0,5}) + 15,7842 * O/(SOL^{-0,5}) - 6,28324 * PC/(SOL^{-0,5})$$

A continuación, se debe realizar de nuevo el test de hipótesis para comprobar que se ha eliminado el problema de heterocedasticidad.

$$Residuos^2 = \frac{\beta_0}{SOL^{-0,5}} + \beta_1 \frac{SOL}{SOL^{-0,5}} + \beta_2 \frac{VI}{SOL^{-0,5}} + \beta_3 \frac{O}{SOL^{-0,5}} + \beta_4 \frac{PC}{SOL^{-0,5}} + \beta_z$$

Tabla 3.72. Estadísticos de residuos al cuadrado de las variables

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
CONSTANTE	4,67187E12	1,46243E13	0,319459	0,7511
1/(SOL <sup>-0,5</sup> )	-6,50582E11	1,5573E12	-0,417763	0,6784
SOL/(SOL <sup>-0,5</sup> )	1,73937E9	2,64105E9	0,65859	0,5140
VI/(SOL <sup>-0,5</sup> )	1,64988E9	1,43707E9	1,14809	0,2579
O/(SOL <sup>-0,5</sup> )	-2,60009E7	2,7562E7	-0,94336	0,3513
PC/(SOL <sup>-0,5</sup> )	-4,25971E6	2,19652E6	-1,9393	0,0597



Análisis de Varianza

Tabla 3.73. Estadísticos residuos al cuadrado del modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	2,08987E25	5	4,17973E24	2,20	0,0735
Residuo	7,39902E25	39	1,89719E24		
Total (Corr.)	9,48889E25	44			

- Contraste de hipótesis (horas de sol):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,5140 >  $\alpha$ . Se rechaza la hipótesis nula. La variable horas de sol no presenta problema de heterocedasticidad.

- Contraste de hipótesis (viento):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,2579 >  $\alpha$ . Se rechaza la hipótesis nula. La variable viento no presenta problema de heterocedasticidad.

- Contraste de hipótesis (viento):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,3513 >  $\alpha$ . Se rechaza la hipótesis nula. La variable ocupados no presenta problema de heterocedasticidad.

- Contraste de hipótesis (permisos de conducir):

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P-valor: 0,0597 >  $\alpha$ . Se rechaza la hipótesis nula. La variable ocupados no presenta problema de heterocedasticidad.

Por último, después de solucionar el problema de heterocedasticidad se debe eliminar los puntos anómalos (Badajoz y Cáceres) y eliminar las variables no significativas mediante la prueba selección paso a paso hacia delante.



Tabla 3.74. Estadísticos significatividad de las variables

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
SOL/(SOL <sup>-0,5</sup> )	875,575	78,4446	11,1617	0,0000
PC/(SOL <sup>-0,5</sup> )	-5,45834	1,34049	-4,0719	0,0002

### Análisis de Varianza

Tabla 3.75. Estadísticos significatividad del modelo.

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	1,88372E14	2	9,41862E13	112,23	0,0000
Residuo	3,44096E13	41	8,39259E11		
Total	2,22782E14	43			

R-cuadrada = 84,5546 por ciento

R-cuadrado (ajustado para g.l.) = 84,1779 por ciento

$$CO_2/(SOL^{-0,5}) = 875,575 * SOL/(SOL^{-0,5}) - 5,45834 * PC/(SOL^{-0,5})$$

Para explicar la variable Dióxido de carbono utilizamos las variables horas de sol y permisos de conducir.

Según este modelo a más horas de sol mayores serán los valores de dióxido de carbono y por el contrario a más permisos de conducir menos CO<sub>2</sub>. Aunque la fórmula exprese estos valores, se piensa que no son del todo lógicos, ya que según las noticias encontradas antes de realizar el análisis el dióxido de carbono, está relacionado directamente con los automóviles de motor diésel. Se observa en la matriz de correlación (tabla 3.15) los vehículos tienen una relación directa muy elevada con los permisos de conducir expedidos ( $R_{ij}=0,974$ ).



• **Óxidos de Nitrógeno (NO<sub>x</sub>):**

○ **Modelo teórico propuesto:**

$$NO_x = \beta_0 + \beta_1 TEMP + \beta_2 LLU + \beta_3 VI + \beta_4 P + \beta_5 O + U$$

○ **Interpretación de los parámetros y unidades físicas:**

$\beta_0$  = Valor medio de óxidos de nitrógeno (NO<sub>x</sub>) cuando las variables explicativas valen 0.

$$\frac{\mu g}{m^3}$$

$\beta_1$  = Incremento medio de óxidos de nitrógeno (NO<sub>x</sub>) cuando aumentamos un grado centígrado la temperatura (TEMP) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu g}{m^3 \times ^\circ C}$$

$\beta_2$  = Incremento medio de óxidos de nitrógeno (NO<sub>x</sub>) cuando aumentamos un milímetro las precipitaciones (LLU) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu g}{m^3 \times \frac{mm}{mes}}$$

$\beta_3$  = Incremento medio de óxidos de nitrógeno (NO<sub>x</sub>) cuando aumentamos un kilómetro / hora el viento (VI) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu g}{m^3 \times \frac{km}{hora}}$$

$\beta_4$  = Incremento medio de óxidos de nitrógeno (NO<sub>x</sub>) cuando aumentamos una unidad la población (P) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu g}{m^3 \times miles \ de \ personas}$$

$\beta_5$  = Incremento medio de óxidos de nitrógeno (NO<sub>x</sub>) cuando aumentamos un el número de ocupados (O) y el resto de las variables explicativas pertenecen constantes.

$$\frac{\mu g}{m^3 \times miles \ de \ ocupados}$$

○ **Primera estimación del modelo ajustado:**

$$NO_x = 28,2273 - 0,791607*TEMP + 0,0987383*LLU - 0,139547*VI + 0,0000111226*P + + 0,00434276*O$$



○ **Análisis de significatividad:**

Tabla 3.76. Estadísticos significatividad variables

Parámetro	Estimación	Error		Estadístico T	Valor-P
		Estándar			
CONSTANTE	28,2273	13,5712		2,07994	0,0442
TEMP	-0,791607	0,720216		-1,09912	0,2785
LLU	0,0987383	0,0451474		2,18702	0,0348
VI	-0,139547	0,124147		-1,12405	0,2679
P	0,0000111226	0,00000396586		2,80458	0,0078
O	0,00434276	0,00258823		1,67789	0,1014

**Análisis de Varianza**

Tabla 3.77. Estadísticos significatividad modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	1460,32	5	292,065	4,37	0,0030
Residuo	2608,52	39	66,885		
Total (Corr.)	4068,84	44			

R-cuadrada = 35,8904%

- Contraste significación global (modelo):

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_1: \text{Algún } \beta_i \neq 0$$

$$\left. \begin{array}{l} H_0 \\ H_1 \end{array} \right\} \alpha = 0,05$$

P-valor 0,0030 <  $\alpha$ : Se rechaza la hipótesis nula y se considera que el modelo sí que es significativo.

- Contraste significación individual (constante):

$$H_0: \beta_0 = 0$$

$$H_1: \beta_0 \neq 0$$

$$\left. \begin{array}{l} H_0 \\ H_1 \end{array} \right\} \alpha = 0,05$$

P-valor 0,0442 >  $\alpha$ . Se rechaza la hipótesis nula y se concluye que la constante es significativa.



CAMPUS D'ALCOI

- Contraste significación individual (temperatura):  
$$\left. \begin{array}{l} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,2785 > \alpha$ . Se acepta la hipótesis nula y se concluye que la variable temperatura no resulta significativa.
- Contraste significación individual (lluvia):  
$$\left. \begin{array}{l} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,0348 > \alpha$ . Se acepta la hipótesis nula y se concluye que la variable lluvia no resulta significativa.
- Contraste significación individual (viento):  
$$\left. \begin{array}{l} H_0: \beta_3 = 0 \\ H_1: \beta_3 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,2679 > \alpha$ . Se acepta la hipótesis nula y se concluye que la variable viento no resulta significativa.
- Contraste significación individual (población):  
$$\left. \begin{array}{l} H_0: \beta_4 = 0 \\ H_1: \beta_4 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,0078 < \alpha$ . Se rechaza la hipótesis nula y se concluye que la variable población resulta significativa.
- Contraste significación individual (ocupados)  
$$\left. \begin{array}{l} H_0: \beta_5 = 0 \\ H_1: \beta_5 \neq 0 \end{array} \right\} \alpha = 0,05$$

P-valor  $0,1014 > \alpha$ . Se rechaza la hipótesis nula y se concluye que la variable ocupados no resulta significativa.



- **Análisis de normalidad:**
  - **Gráfico probabilístico:**

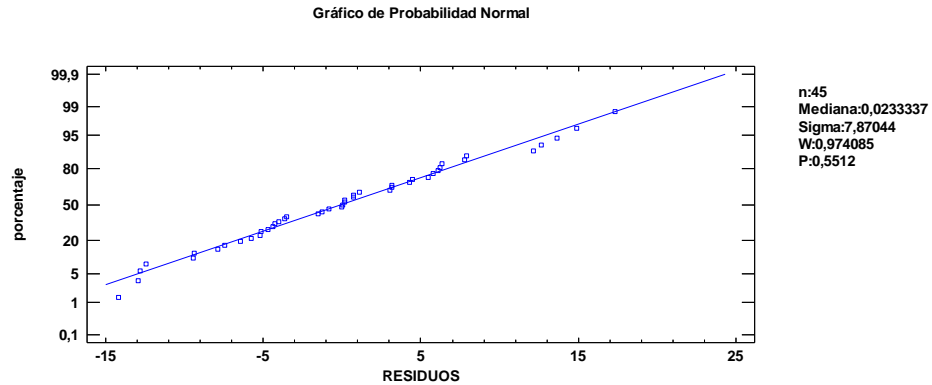


Gráfico 3.31. Gráfico probabilidad normal

La mayoría de los residuos están muy cerca de la línea de normalidad. Se puede decir que los residuos parecen tener una distribución normal

- **Histograma de residuos:**

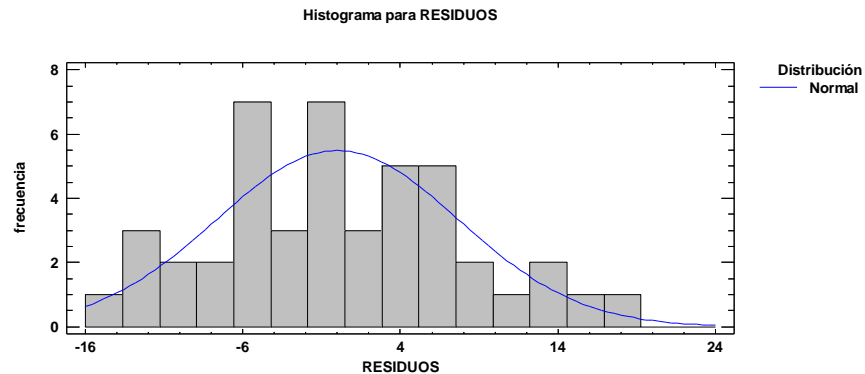


Gráfico 3.32. Histograma

En el histograma se aprecia la forma de Campana de Gauss ya que la mayoría de residuos se encuentran en el centro, por esto se parece que los residuos tienen un comportamiento normal.





▪ **Test de normalidad:**

Para realizar el test de normalidad se utiliza la prueba con menor p-valor.

Tabla 3.78. Test de normalidad

Prueba	Estadístico	Valor-P
Chi-Cuadrado	13,8	0,540749
Estadístico W de Shapiro-Wilk	0,974085	0,551199
Valor-Z para asimetría	0,475776	0,634231
Valor-Z para curtosis	-0,412904	0,679673

$H_0$ : Distribución normal  
 $H_1$ :  $\neq$  Distribución normal

}  $\alpha$ : 0,05

0,540749 > 0,05. Se acepta la hipótesis nula, por tanto, el test concluye que los residuos son normales.

○ **Análisis de heterocedasticidad:**

$$\text{Residuos}^2 = \beta_0 + \beta_1 \text{TEMP} + \beta_2 \text{LLU} + \beta_3 \text{VI} + \beta_4 \text{P} + \beta_5 \text{P} + U$$

Tabla 3.79. Estadísticos residuos al cuadrado de las variables

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
CONSTANTE	113,069	124,825	0,905825	0,3706
TEMP	-6,46587	6,6244	-0,976069	0,3350
LLU	0,116694	0,415257	0,281017	0,7802
VI	0,456148	1,14188	0,399471	0,6917
P	0,0000314659	0,0000364772	0,86262	0,3936
O	0,00175889	0,023806	0,0738842	0,9415

**Análisis de Varianza**

Tabla 3.80. Estadísticos residuos al cuadrado del modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	13787,2	5	2757,44	0,49	0,7836
Residuo	220679,	39	5658,43		
Total (Corr.)	234466,	44			



CAMPUS D'ALCOI

- Contraste de hipótesis (Temperatura):  
 $H_0: \sigma^2 = \text{Constante}$   
 $H_1: \sigma^2 \neq \text{Constante}$  }  $\alpha = 0,05$   
P-valor:  $0,3350 > \alpha$ . Se acepta la hipótesis nula. La variable temperatura no presenta problema de heterocedasticidad.
- Contraste de hipótesis (precipitaciones)  
 $H_0: \sigma^2 = \text{Constante}$   
 $H_1: \sigma^2 \neq \text{Constante}$  }  $\alpha = 0,05$   
P-valor:  $0,7802 > \alpha$ . Se acepta la hipótesis nula. La variable precipitaciones no presenta problema de heterocedasticidad.
- Contraste de hipótesis (Viento):  
 $H_0: \sigma^2 = \text{Constante}$   
 $H_1: \sigma^2 \neq \text{Constante}$  }  $\alpha = 0,05$   
P-valor:  $0,6917 > \alpha$ . Se acepta la hipótesis nula. La variable viento no presenta problema de heterocedasticidad.
- Contraste de hipótesis (Población):  
 $H_0: \sigma^2 = \text{Constante}$   
 $H_1: \sigma^2 \neq \text{Constante}$  }  $\alpha = 0,05$   
P-valor:  $0,3936 > \alpha$ . Se acepta la hipótesis nula. La variable población no presenta problema de heterocedasticidad.
- Contraste de hipótesis (Ocupados):  
 $H_0: \sigma^2 = \text{Constante}$   
 $H_1: \sigma^2 \neq \text{Constante}$  }  $\alpha = 0,05$   
P-valor:  $0,9415 > \alpha$ . Se acepta la hipótesis nula. La variable población no presenta problema de heterocedasticidad.



○ **Autocorrelación:**

▪ **Test de Durbin - Watson:**

Estadístico Durbin-Watson = (P=0,2065)

$H_0: \rho_1 = 0$  no hay autocorrelación  
 $H_1: \rho_1 \neq 0$  hay autocorrelación

$\alpha=0.05$

P-valor 0,2065 >  $\alpha$ . se acepta la hipótesis nula y se concluye que el modelo no tiene autocorrelación de primer orden.

▪ **Gráfico de Durbin - Watson:**

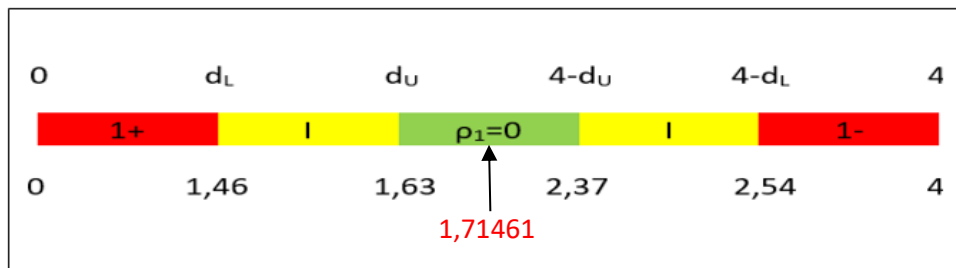


Gráfico 3.33. Gráfico de Durbin - Watson

Al ubicar el estadístico de Durbin-Watson en su gráfico se observa que no existe autocorrelación de primer orden.

▪ **FAS y FAP:**

Autocorrelaciones Estimadas para RESIDUOS

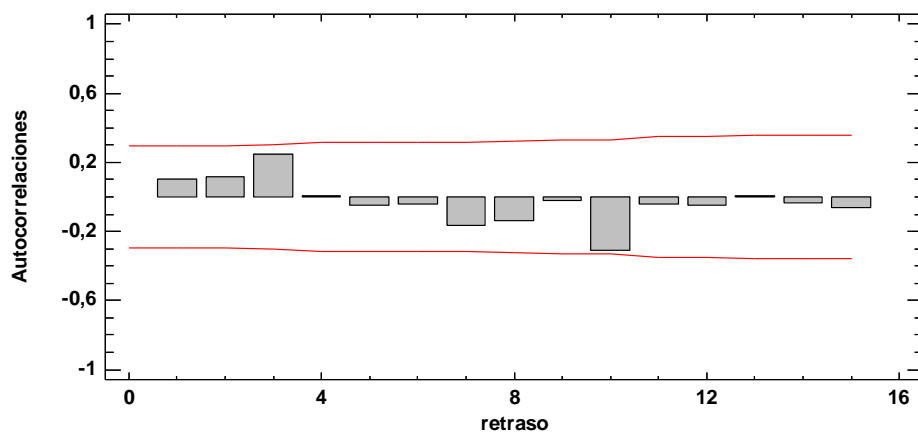


Gráfico 3.34. FAS



Autocorrelaciones Parciales Estimadas para RESIDUOS

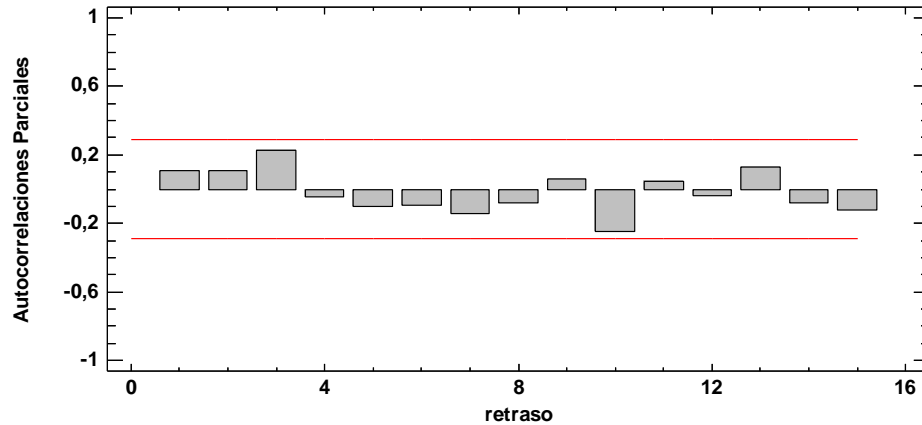


Gráfico 3.35. FAP

Al observar los gráficos del FAS y el FAP no se observa ningún factor de correlación que supere los límites establecidos, por esto se podría decir que no existe problema de autocorrelación.

○ **Puntos influyentes:**

Tabla 3.81. Puntos influyentes

Fila	Provincia	Influencia	Distancia de Mahalanobis	DFITS
8	Burgos	0,375466	24,874	-1,14266
14	Coruña	0,159523	7,18412	-0,856245
17	Granada	0,111719	4,43081	0,844374
18	Gipuzkoa	0,242373	12,7789	0,991234
27	Murcia	0,633253	73,2696	0,28715
32	Pontevedra	0,291077	16,6781	-0,735511
42	Valladolid	0,123766	5,0964	0,758643

Influencia media de un solo punto = 0,133333

▪ Puntos influyes a priori:

$$n_{ii} > 2 \bar{n} \longrightarrow n_{ii} > 0,267$$

- 0,375466 > 2  $\bar{n}$ : Burgos es un punto influyente a priori.
- 0,159523 < 2  $\bar{n}$ : Coruña no es un punto influyente a priori.
- 0,111719 < 2  $\bar{n}$ : Granada no es un punto influyente a priori.
- 0,242373 < 2  $\bar{n}$ : Gipuzkoa no es un punto influyente a priori.
- 0,633253 > 2  $\bar{n}$ : Murcia es un punto influyente a priori.
- 0,291077 > 2  $\bar{n}$ : Pontevedra es un punto influyente a priori.
- 0,123766 < 2  $\bar{n}$ : Valladolid no es un punto influyente a priori.



- Puntos influyentes a posteriori:  
 $|DFITS| \geq 2 \sqrt{\bar{n}} \longrightarrow |DFITS| \geq 0,7302$ 
  - $1,14266 \geq 2 \sqrt{\bar{n}}$  : Burgos es un punto influyente a posteriori
  - $0,856245 \geq 2 \sqrt{\bar{n}}$  : Coruña es un punto influyente a posteriori.
  - $0,844374 \geq 2 \sqrt{\bar{n}}$  : Granada es un punto influyente a posteriori.
  - $0,991234 \geq 2 \sqrt{\bar{n}}$  : Gipuzkoa es un punto influyente a posteriori.
  - $0,28715 \leq 2 \sqrt{\bar{n}}$  : Murcia no es un punto influyente a posteriori.
  - $0,735511 \geq 2 \sqrt{\bar{n}}$  : Pontevedra es un punto influyente a posteriori.
  - $0,758643 \geq 2 \sqrt{\bar{n}}$  : Valladolid es un punto influyente a posteriori.

○ **Puntos anómalos:**

Tabla 3.82. Puntos anómalos

			Y		Residuo
Fila	Provincia	Y	Predicha	Residuo	Estudentizado
17	Granada	43,9741	26,6308	17,3434	2,38
42	Valladolid	37,2943	22,4164	14,8779	2,02

Granada y Valladolid son puntos anómalos.

○ **Resolución de problemas:**

A continuación, se propone el modelo para óxidos de nitrógeno eliminando las variables no significativas a partir del procedimiento de ajuste de la selección paso a paso hacia delante y los puntos anómalos.

Tabla 3.83. Estadísticos significatividad de las variables

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
CONSTANTE	11,0507	2,82694	3,90907	0,0003
LLU	0,121426	0,0362834	3,3466	<b>0,0018</b>
P	0,00000798881	0,00000251837	3,17222	0,0029

**Análisis de Varianza**

Tabla 3.84. Estadísticos significatividad del modelo

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	1186,73	2	593,364	10,51	<b>0,0002</b>
Residuo	2258,62	40	56,4655		
Total (Corr.)	3445,35	42			

R-cuadrada = 34,4444 por ciento

R-cuadrado (ajustado para g.l.) = 31,1666 por ciento

Error estándar del est. = 7,51435

Error absoluto medio = 5,83209



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

## CAMPUS D'ALCOI

Estadístico Durbin-Watson = 1,99694 (P=0,5148)

Autocorrelación de residuos en retraso 1 = -0,0496241

$NO_x = 11,0507 + 0,121426*LLU + 0,00000798881*P$

Para explicar la variable explicativa se utiliza las precipitaciones y la población de cada una de las provincias españolas.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

## 4. Predicciones:



Después de realizar el análisis multivariante se obtienen los siguientes tres modelos de regresión:

- $SO_2 = 2,81549 + 0,00000224828 * P \longrightarrow R^2 = 19,5124\%$
- $CO_2 / (SOL^{-0,5}) = 875,575 * SOL / (SOL^{-0,5}) - 5,45834 * PC / (SOL^{-0,5}) \longrightarrow R^2 = 84,5546\%$
- $NO_x = 11,0507 + 0,121426 * LLU + 0,00000798881 * P \longrightarrow R^2 = 34,4444\%$

Estos tres modelos son totalmente válidos, pero sólo el de dióxido de carbono tiene un R-cuadrado aceptable.

Para este se van a realizar las correspondientes predicciones para la provincia de Alicante.

En primer lugar, se realiza un aumento de las horas de sol a 300 y se mantienen los permisos de conducir constantes. A partir de estos cambios se obtiene la predicción puntual y el intervalo de confianza de la media.

- Predicción puntual: 1422290 KT/año de  $CO_2$
- Intervalo de confianza de la media:  $[459425 \leq CO_2 \leq 2385150]_{95\%}$

Como nuestro modelo está dividido entre horas de sol elevado a -0,5 se despeja la predicción puntual y el intervalo de confianza.

- Predicción puntual:  
 $\frac{CO_2}{SOL^{-0,5}} = 1422290$

$$\frac{CO_2}{300^{-0,5}} = 1422290$$

$$CO_2 = 1422290 \times 300^{-0,5}$$

$$CO_2 = 82115,9514 \text{ KT/año de } CO_2$$

- Intervalo de confianza de la media:

$$\frac{CO_2}{SOL^{-0,5}} = 459425$$

$$\frac{CO_2}{SOL^{-0,5}} = 2385150$$

$$\frac{CO_2}{300^{-0,5}} = 459425$$

$$\frac{CO_2}{300^{-0,5}} = 2385150$$

$$CO_2 = 459425 \times 300^{-0,5}$$

$$CO_2 = 2385150 \times 300^{-0,5}$$

$$CO_2 = 26524,9147$$

$$CO_2 = 137706,699$$

$$[24524,9147 \leq CO_2 \leq 137706,699]_{95\%}$$

Se comprueba que al aumentar las horas de sol de la provincia de Alicante aumenta el dióxido de carbono.





272,8 horas de sol → 300 horas de sol  
 33079 permisos de conducir → 33079 permisos de conducir  
 69711,54807 KT/año de CO<sub>2</sub> → 82115,9514 KT/año de CO<sub>2</sub>

A continuación, se realiza una disminución de las horas de sol a 220 y de nuevo se mantienen los permisos de conducir constantes.

- Predicción puntual: 179024KT/año de CO<sub>2</sub>
- Intervalo de confianza de la media:  $[-754714 \leq CO_2 \leq 1112760]_{95\%}$

De nuevo se debe despejar el CO<sub>2</sub>.

- Predicción puntual:

$$\frac{CO_2}{SOL^{-0,5}} = 179024$$

$$\frac{CO_2}{220^{-0,5}} = 179024$$

$$CO_2 = 179024 \times 220^{-0,5}$$

$$CO_2 = 12069,7956$$

- Intervalo de confianza de la media:

$$\frac{CO_2}{SOL^{-0,5}} = -754714$$

$$\frac{CO_2}{220^{-0,5}} = -754714$$

$$CO_2 = -754714 \times 220^{-0,5}$$

$$CO_2 = -50882,8075$$

$$\frac{CO_2}{SOL^{-0,5}} = 1112760$$

$$\frac{CO_2}{220^{-0,5}} = 1112760$$

$$CO_2 = 1112760 \times 220^{-0,5}$$

$$CO_2 = 75022,2639$$

$$[-50882,8075 \leq CO_2 \leq 75022,2639]_{95\%}$$

Co<sub>2</sub> no puede ser negativo.

$$[0 \leq CO_2 \leq 75022,2639]_{95\%}$$



Al bajar las horas de sol a 220 y mantener constante los permisos de conducir se reducen considerablemente las kilotoneladas de dióxido de carbono.

272,8 horas de sol	—————→	300 horas de sol
33079 permisos de conducir	—————→	33079 permisos de conducir
69711,54807 KT/año de CO <sub>2</sub>	—————→	12069,7956 KT/año de CO <sub>2</sub>

De nuevo, se vuelven a realizar las predicciones, pero en este caso las horas de sol se mantienen constantes y los permisos de conducir aumentan a 34500.

- Predicción puntual: 834815 KT/año de CO<sub>2</sub>
- Intervalo de confianza de la media:  $[-179854 \leq CO_2 \leq 1849480]_{95\%}$

A continuación, se despejan el dióxido de carbono.

- Predicción puntual:

$$\frac{CO_2}{SOL^{-0,5}} = 834815$$

$$\frac{CO_2}{272,8^{-0,5}} = 834815$$

$$CO_2 = 834815 \times 272,8^{-0,5}$$

$$CO_2 = 50543,81988$$

- Intervalo de confianza de la media:

$$\frac{CO_2}{SOL^{-0,5}} = -179854$$

$$\frac{CO_2}{SOL^{-0,5}} = 1849480$$

$$\frac{CO_2}{272,8^{-0,5}} = -179854$$

$$\frac{CO_2}{272,8^{-0,5}} = 1849480$$

$$CO_2 = -179854 \times 272,8^{-0,5}$$

$$CO_2 = 1849480 \times 272,8^{-0,5}$$

$$CO_2 = -10889,24873$$

$$CO_2 = 111976,6463$$

$$[-10889,24873 \leq CO_2 \leq 111976,6463]_{95\%}$$

El dióxido de carbono no puede ser negativo con lo cual el intervalo de confianza quedaría de la siguiente forma:

$$[0 \leq CO_2 \leq 111976,6463]_{95\%}$$



Al mantener constantes las horas de sol y aumentar los permisos de conducir disminuyen las kilotoneladas de dióxido de carbono.

272,8 horas de sol → 272,8 horas de sol  
 33079 permisos de conducir → 34500 permisos de conducir  
 69711,54807 KT/año de CO<sub>2</sub> → 48198,0665 KT/año de CO<sub>2</sub>

Después, se siguen manteniendo las horas de sol constantes, pero por el contrario que anteriormente se disminuyen los permisos de conducir a 32500.

- Predicción puntual: 1015120 KT/año de CO<sub>2</sub>
- Intervalo de confianza de la media: [82678,7 ≤ CO<sub>2</sub> ≤ 1947570]<sub>95%</sub>

Se debe realizar el despeje para poder comparar los valores:

- Predicción puntual:

$$\frac{CO_2}{SOL^{-0,5}} = 1015120$$

$$\frac{CO_2}{272,8^{-0,5}} = 1015120$$

$$CO_2 = 1015120 \times 272,8^{-0,5}$$

$$CO_2 = 61460,337438$$

- Intervalo de confianza de la media:

$$\frac{CO_2}{SOL^{-0,5}} = 82678,7$$

$$\frac{CO_2}{SOL^{-0,5}} = 1947570$$

$$\frac{CO_2}{272,8^{-0,5}} = 82678,7$$

$$\frac{CO_2}{272,8^{-0,5}} = 1947570$$

$$CO_2 = 82678,7 \times 272,8^{-0,5}$$

$$CO_2 = 1947570 \times 272,8^{-0,5}$$

$$CO_2 = 5005,776514$$

$$CO_2 = 117915,499$$

$$[5005,776514 \leq CO_2 \leq 117915,499]_{95\%}$$

Al disminuir los permisos de conducir se observa que disminuyen las kilotoneladas de dióxido de carbono.

272,8 horas de sol → 272,8 horas de sol  
 33079 permisos de conducir → 32500 permisos de conducir  
 69711,54807 KT/año de CO<sub>2</sub> → 58607,9805 KT/año de CO<sub>2</sub>



CAMPUS D'ALCOI

A continuación, se aumentan las horas de sol a 300 y se disminuyen los permisos de conducir 32500.

- Predicción puntual: 1477030 KT/año de CO<sub>2</sub>
- Intervalo de confianza de la media:  $[538217 \leq CO_2 \leq 2415840]_{95\%}$

Como en los casos anteriores se debe despejar el modelo para poder compararlo.

- Predicción puntual

$$\frac{CO_2}{SOL^{-0,5}} = 1477030$$

$$\frac{CO_2}{300^{-0,5}} = 1477030$$

$$CO_2 = 1477030 \times 300^{-0,5}$$

$$CO_2 = 85276,3668$$

- Intervalo de confianza de la media:

$$\frac{CO_2}{SOL^{-0,5}} = 538217$$

$$\frac{CO_2}{300^{-0,5}} = 538217$$

$$CO_2 = 538217 \times 300^{-0,5}$$

$$CO_2 = 31073,973$$

$$\frac{CO_2}{SOL^{-0,5}} = 2415840$$

$$\frac{CO_2}{300^{-0,5}} = 2415840$$

$$CO_2 = 2415840 \times 300^{-0,5}$$

$$CO_2 = 1816402,56$$

$$[31073,973 \leq CO_2 \leq 1816402,56]_{95\%}$$

Al aumentar las horas de sol y disminuir los permisos de conducir aumentan considerablemente las kilotoneladas de dióxido de carbono.

272,8 horas de sol	—————▶	300 horas de sol
33079 permisos de conducir	—————▶	32500 permisos de conducir
69711,54807 KT/año de CO <sub>2</sub>	—————▶	85276,3668 KT/año de CO <sub>2</sub>

Por último, se decide disminuir las horas de sol a 220 y aumentar los permisos de conducir a 34500.



CAMPUS D'ALCOI

- Predicción puntual: 63979,7 KT/año de CO<sub>2</sub>
- Intervalo de confianza de la media:  $[-924395 \leq CO_2 \leq 1052350]_{95\%}$

Se debe despejar para poder comparar las kilotoneladas de dióxido de carbono.

- Predicción puntual:

$$\frac{CO_2}{SOL^{-0,5}} = 63979,7$$

$$\frac{CO_2}{220^{-0,5}} = 63979,7$$

$$CO_2 = 63979,7 \times 220^{-0,5}$$

$$CO_2 = 4313,510494$$

- Intervalo de confianza de la media:

$$\frac{CO_2}{SOL^{-0,5}} = -924395$$

$$\frac{CO_2}{220^{-0,5}} = -924395$$

$$CO_2 = -924395 \times 220^{-0,5}$$

$$CO_2 = -62322,69819$$

$$\frac{CO_2}{SOL^{-0,5}} = 1052350$$

$$\frac{CO_2}{220^{-0,5}} = 1052350$$

$$CO_2 = 1052350 \times 220^{-0,5}$$

$$CO_2 = 70949,42253$$

$$[-62322,69819 \leq CO_2 \leq 70949,42253]_{95\%}$$

El dióxido de carbono no puede ser negativo:

$$[0 \leq CO_2 \leq 70949,42253]_{95\%}$$

Al disminuir y aumentar las horas de sol existe una disminución considerable de las kilotoneladas de dióxido de carbono al año.

272,8 horas de sol	→	220 horas de sol
33079 permisos de conducir	→	34500 permisos de conducir
69711,54807 KT/año de CO <sub>2</sub>	→	3693,8697 KT/año de CO <sub>2</sub>



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

## **5. Comparación del Dióxido de carbono en las provincias Españolas:**



A continuación, se realiza una comparación de niveles de Dióxido de Carbono en todas las provincias españolas. En este se observa que provincias se encuentran por encima, por debajo y en niveles medios de dióxido de carbono.

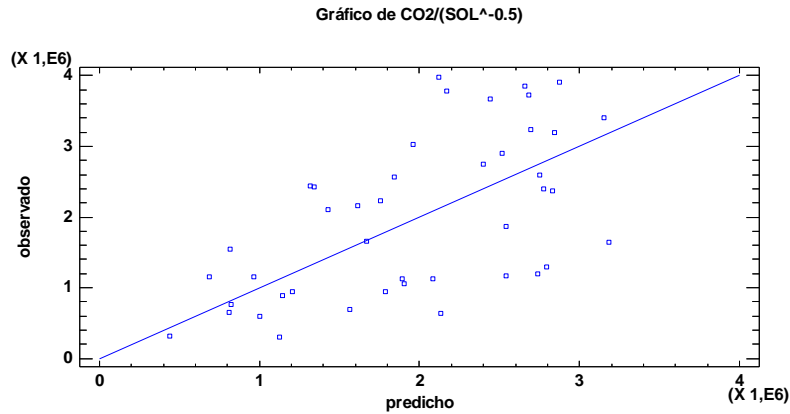


Gráfico 5.1. Gráfico valores observados contra valores predichos

En primer lugar, se realizan las predicciones del modelo para todas las provincias y se hayan los valores teóricos para poder comparar estos con los valores reales de Dióxido de Carbono (CO<sub>2</sub>)

A continuación, se realiza la diferencia entre ambos valores y se haya la desviación típica de los valores teóricos para saber cuál es la desviación respecto al valor medio y se divide entre dos.

Por último, a partir de la desviación típica y la diferencia entre en valor real y valor teórico se puede observar que valores se encuentran por encima, por debajo o en la media de Dióxido de Carbono.

- Si la diferencia del valor real y valor teórico  $>$  desviación típica entre dos: Por encima de la media.

$$\text{Diferencia valor real y valor teórico} > 22029,52537$$

- Si la diferencia del valor real y valor teórico  $<$  menos la desviación típica entre dos: Por debajo de la media.

$$\text{Diferencia valor real y valor teórico} < -22029,52537$$

- Si diferencia del valor real y valor teórico se encuentra entre los valores positivos y negativos de la desviación típica entre dos: En la media.

$$-22029,52537 < \text{diferencia valor real y valor teórico} < 22029,52537$$

Tabla 5.2. Predicción provincias españolas



Provincias	CO <sub>2</sub> real	CO <sub>2</sub> teórico	Diferencia	Por encima	Por debajo	En la media
Araba/Álava	53236,38	121614,6158	-68378,2358	0	1	0
Albacete	209323,5024	193708,3844	15615,11799	0	0	1
Alicante/Alacant	69711,54807	58300,17095	11411,37712	0	0	1
Almería	79012,10224	170494,4245	-91482,32225	0	1	0
Ávila	173082,736	183945,3514	-10862,61543	0	0	1
Balears, Illes	62905,08678	80591,53358	-17686,4468	0	0	1
Burgos	217664,2194	141141,2887	76522,93072	1	0	0
Cádiz	60042,63483	113504,0005	-53461,36564	0	1	0
Castellón/Castelló	75312,87617	172608,5566	-97295,68045	0	1	0
Ciudad Real	246663,5819	181315,1547	65348,42718	1	0	0
Córdoba	175400,2348	153250,7822	22149,45258	1	0	0
Coruña, A	93140,19339	55714,50627	37425,68712	1	0	0
Cuenca	265231,0457	152032,7402	113198,3055	1	0	0
Girona	77060,08326	128807,6348	-51747,55149	0	1	0
Granada	162556,7778	116814,5437	45742,23406	1	0	0
Gipuzkoa	23919,43209	89882,34312	-65962,91103	0	1	0
Huelva	100976,8543	195195,6748	-94218,8205	0	1	0
Huesca	248061,3889	178828,8981	69232,49081	1	0	0
Jaén	203120,5058	168704,6953	34415,8105	1	0	0
León	193614,5107	107267,8804	86346,63028	1	0	0
Lleida	151663,5068	180734,8913	-29071,38454	0	1	0
Rioja, La	80046,1134	147643,2435	-67597,13014	0	1	0
Lugo	199453,4524	107384,1583	92069,29408	1	0	0
Málaga	57149,23102	72978,81698	-15829,58596	0	0	1
Murcia	133433,1184	99350,36141	34082,75699	1	0	0
Navarra	159475,6267	108500,8263	50974,80041	1	0	0
Ourense	126556,133	128047,0767	-1490,943745	0	0	1
Asturias	125358,4603	65843,98102	59514,47928	1	0	0
Palmas, Las	57241,43952	62291,82229	-5050,382765	0	0	1
Pontevedra	47118,82563	78556,38541	-31437,55978	0	1	0
Salamanca	206736,0281	183733,4055	23002,62261	1	0	0
Santa Cruz de Tenerife	40125,52518	133649,4973	-93523,97208	0	1	0
Cantabria	58157,33617	72261,94418	-14104,60801	0	0	1
Segovia	126451,6177	172285,9374	-45834,31968	0	1	0
Sevilla	134332,7358	105786,9886	28545,74716	1	0	0





Soria	161957,2027	186972,6775	-25015,47485	0	1	0
Tarragona	72084,07144	129901,2372	-57817,16573	0	1	0
Teruel	254661,6362	169498,5491	85163,08706	1	0	0
Toledo	238601,3667	164614,8226	73986,54413	1	0	0
Valladolid	76860,7817	167150,8006	-90290,01894	0	1	0
Bizkaia	29000,51634	39705,61837	-10705,10203	0	0	1
Zamora	199480,5036	173021,5165	26458,9871	1	0	0
Zaragoza	260651,2463	138984,0651	121667,1812	1	0	0

Desviación típica valor teórico: 44059,05074

Desviación típica valor teórico entre dos: 22029,52537

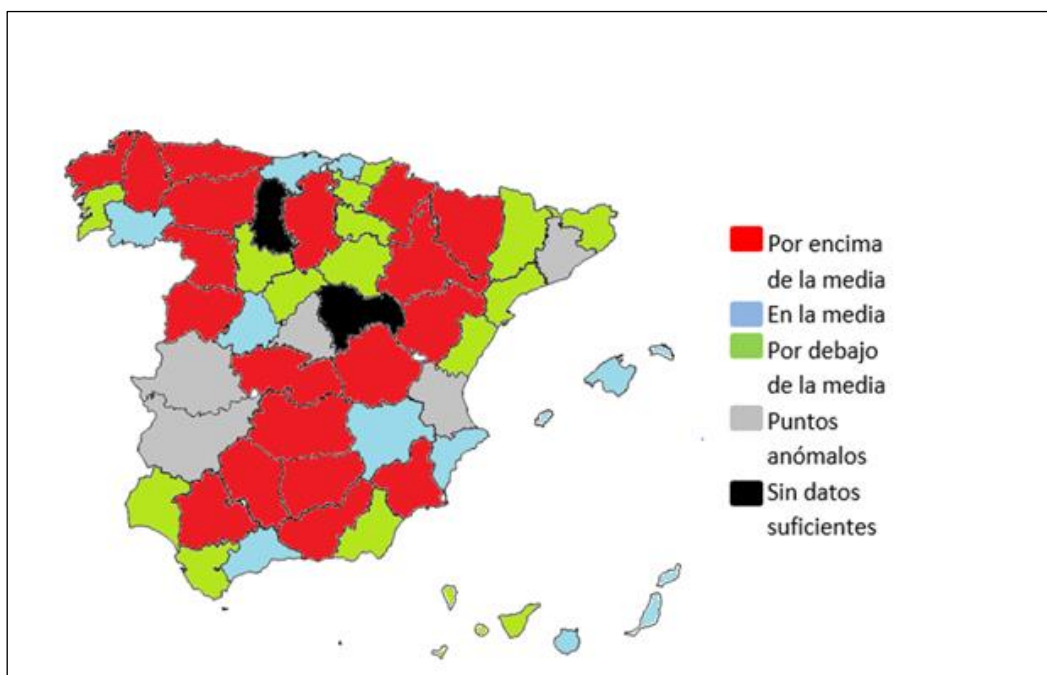


Gráfico 5.1. Distribución Dióxido de carbono en las provincias españolas



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

## 6. Conclusiones:



## CAMPUS D'ALCOI

- Antes de iniciar los cálculos, resultados y explotación se eliminaron las provincias de Guadalajara y Lérida por no disponer de la suficiente información.
- Se realizó el análisis univariante de cada una de las variables explicadas y explicativas. En este análisis se localizaron tres puntos anómalos comunes a la mayoría de variables, estos eran: Barcelona, Madrid y Valencia.
- Se decidió eliminar los puntos anómalos mencionados anteriormente ya que, no eran datos homogéneos con el resto.
- Después, se realizó el análisis de multicolinealidad en el cuál existían problemas. Las variables con las cuales no existía problema de multicolinealidad para cada uno de los diferentes gases serían:
  - Dióxido de azufre: temperatura, precipitaciones, viento, población y ocupados.
  - Dióxido de carbono: horas de sol, viento, ocupados, permisos de conducir.
  - Óxidos de nitrógeno: temperatura, lluvia, viento, población y ocupados.
- Luego, se hizo el análisis multivariante de los tres modelos:
  - Dióxido de azufre:
    - No presentaba problemas, aunque la mayoría de las variables no eran significativas.
    - Al realizar el método de selección hacia delante se quedó la constante y la variable población.
    - Obtuvo un R-cuadrado muy bajo de un 19,514%
    - El modelo obtenido fue:  $SO_2 = 2,81549 + 0,00000224828 * P$
    - Se pudo apreciar que la cantidad de dióxido de azufre estaba relacionado directamente con el tamaño de la población.
  - Dióxido de carbono:
    - Existía problema de heterocedasticidad.
    - Al eliminar el problema mencionado anteriormente se obtuvo el siguiente modelo:
$$\frac{CO_2}{SOL^{-0,5}} = \frac{\beta_0}{SOL^{-0,5}} + \beta_1 \frac{SOL}{SOL^{-0,5}} + \beta_2 \frac{VI}{SOL^{-0,5}} + \beta_3 \frac{O}{SOL^{-0,5}} + \beta_4 \frac{PC}{SOL^{-0,5}}$$
    - Se obtuvo un R-cuadrado de 84,556%
    - El modelo resultante para el dióxido de carbono fue:
$$CO_2 / (SOL^{-0,5}) = 875,575 * SOL / (SOL^{-0,5}) - 5,45834 * PC / (SOL^{-0,5})$$
    - Para explicar la variable Dióxido de carbono se utilizó las variables horas de sol y permisos de conducir. Según este modelo a más horas de sol mayores serán los valores de dióxido de carbono y por el contrario a más permisos de conducir menos  $CO_2$ .
  - Óxidos de nitrógeno:
    - No se detectaron problemas, aunque existen variables no significativas.
    - El R-cuadrado obtenido fue del 34,444%
    - El modelo obtenido fue:  $NO_x = 11,0507 + 0,121426 * LLU + 0,00000798881 * P$



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

- Para explicar la variable explicativa se utiliza las precipitaciones y la población de cada una de las provincias españolas.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

## **7. Líneas futuras de investigación**



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

Con el fin de que algún alumno pudiera seguir con el análisis de los principales gases contaminantes en España a mediante modelos de regresión se recomendaría intentar ampliar la cantidad de gases y añadir los principales metales en suspensión. Además, se podrían cambiar las variables explicativas para analizar que otros elementos pueden estar relacionados con el comportamiento de estos gases.

Por otra parte, sería interesante sustituir la eliminación de parámetros utilizada para solucionar el problema de multicolinealidad por componentes principales.

Por último, comprobar si existen problemas utilizando otras estructuras (semilogarítmica, logarítmica, inversa, logarítmica inversa) para intentar sacar otros modelos válidos y significativos.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

## **8. Anexos:**



**Noticias:**

## **Las restricciones no evitan el suspenso en contaminación**

La capital rebasó en nueve de las estaciones de medición el valor límite anual de NO<sub>2</sub> marcado por la legislación europea.

12 de enero de 2017. 00:59h

Marta Palacio



El 29 de diciembre, Carmena activó por primera vez en España la prohibición de circular a la mitad de las matrículas ante los elevados niveles de NO<sub>2</sub>

Alberto R. Roldán

Dicen los expertos que la contaminación no desaparece, que simplemente se traslada o se diluye cuando llueve y que, por tanto, los niveles de polución dependen de la meteorología y de su capacidad de condensarla o disiparla. El Gobierno de Manuela Carmena sin embargo se ha propuesto reducir los niveles de emisiones de dióxido de nitrógeno (NO<sub>2</sub>) de la ciudad poniendo los vehículos en su punto de mira. Este gas amarillento, responsable junto a las partículas en suspensión de la conocida como «boina» de Madrid, tiene su origen en la combustión de los motores de los vehículos diésel y en la industria eléctrica, pero el Ayuntamiento se ha propuesto rebajarlo a costa de una reducción del tráfico de los vehículos privados en el centro de la ciudad. Para ello, el pasado mes de febrero Carmena endureció el protocolo anticontaminación, que se activa cuando se superan los niveles recomendados.





La alcaldesa rebajó el nivel de preaviso a los 180 mg/m<sup>3</sup> (cuando el nivel fijado por la UE y la OMS es de 200 mg/m<sup>3</sup>) y aumentó las restricciones en caso de superarlo. Debido a este endurecimiento, el pasado 29 de diciembre el Ayuntamiento prohibió por primera vez la circulación de las matrículas pares dentro del perímetro de la M-30, además de reducir la velocidad en la M-30 y prohibir aparcar en toda la almendra central a los no residentes durante varios días.

A pesar de que Carmena sacó toda su artillería contra la contaminación, los niveles de polución volvieron a rebasarse, según ha quedado constatado en un informe que Ecologistas en Acción publicó ayer sobre la contaminación en 2016; un informe que deja en evidencia la efectividad de las restricciones que Carmena aprobó en febrero.

Según este balance, la calidad del aire en Madrid mejoró en 2016 respecto al año anterior, que fue especialmente malo para Madrid, «porque las condiciones meteorológicas del 2015 fueron particularmente adversas (...) y porque la meteorología del 2016 fue en general menos adversa, a pesar del episodio de alta contaminación por NO<sub>2</sub> sufrido en los últimos días del año».

El estudio analiza los registros que dan hora a hora las 24 estaciones de medición de contaminantes ambientales que tiene el Ayuntamiento de Madrid. En los tres principales contaminantes ambientales, el dióxido de nitrógeno, el ozono troposférico y las partículas en suspensión, las situaciones de superación fueron «graves» teniendo en cuenta los niveles establecidos por la UE y la OMS. En el caso del dióxido de nitrógeno (gas tóxico que se concentra en la ciudad, principalmente en invierno), nueve de las 24 estaciones superaron el límite medio anual de contaminación marcado por la UE (40 microgramos por metro cúbico), y cuatro de ellas rebasaron el valor límite horario, lo que implica que hubo más de 18 horas en 2016 en las que los valores medios de esa estación durante 60 minutos superaron los 200 microgramos por metro cúbico. Los ecologistas pidieron ayer que se intensifiquen las restricciones al tráfico rodado y recordaron que precisamente este contaminante es el que han falseado los fabricantes de coches, lo que implica que muchos vehículos nuevos que circulan emitan más NO<sub>2</sub> de lo permitido por normativa. [\[La razón\]](#)



## Málaga y la Costa tienen uno de los niveles de contaminación más bajos de España

El tráfico es el causante de casi el 50% de las emisiones, aunque los valores máximos sólo se superaron de manera puntual el año pasado

IGNACIO LILLO | MÁLAGA

Aunque el aumento del tráfico urbano no ayuda, lo cierto es que la provincia tiene el nivel de contaminación ambiental más bajo de Andalucía, y uno de los menores de España. Así lo reflejan las estadísticas de la Consejería de Medio Ambiente de la Junta, referidas a 2015, que tiene una red de medidores distribuidos por todo el territorio y que ofrecen información en tiempo real. Durante el año pasado sólo hubo un contaminante concreto, el ozono, que marcó valores superiores a los permitidos en momentos puntuales de tres días concretos, en los meses de julio y agosto.

En el caso del dióxido de nitrógeno (NO<sub>2</sub>) y de las partículas PM10 (menores de 10 micrómetros), el informe autonómico, al que tuvo acceso SUR, advierte de que en el caso de Málaga «no se ha sobrepasado la referencia legal pero puede existir un problema potencial». Sobre las PM10 hay que hacer la salvedad de que se aplican factores correctores por situaciones naturales, como es la entrada de la calima o polvo en suspensión procedente de África. De lo contrario, se superarían con creces los límites para este contaminante.

El resto de las provincias andaluzas –no digamos ciudades como Madrid, donde los episodios de alerta por contaminación son cada vez más frecuentes– sufren más problemas ambientales que Málaga. La peor situación general se vive en Huelva, por el polo químico; mientras que en núcleos urbanos destacan Córdoba, Granada y varias zonas rurales, como la de Villanueva del Arzobispo (Jaén). De hecho, según los técnicos de Medio Ambiente, a nivel nacional la Costa del Sol es uno de los sitios que ofrecen un aire menos contaminado, sólo por detrás de zonas como Galicia y Levante.

### **Lo peor, el tráfico**

La Junta gestiona de forma directa seis estaciones de medición en la capital (Juan XXIII, El Atabal, Carranque y Campanillas); Marbella y Campillos. A éstas, se unen las dos que tienen Gas Natural, para su central térmica del PTA; y la cementera de La Araña.



Aunque la situación no es mala, los expertos advierten de que el tráfico es el causante de casi el 50% de las emisiones en la provincia. Los coches producen problemas de todo tipo para el medio ambiente: monóxido de carbono, partículas, óxido de nitrógeno, etc. Realmente, la contaminación urbana es casi la única que padece la provincia, por la ausencia de industria, a excepción de la central térmica de ciclo combinado (que casi todo lo que emite es vapor de agua); y la cementera de La Araña, que produce emisiones fundamentalmente de partículas. Tampoco hay apenas consumo de calefacción. A ello, se une la ubicación costera de la mayoría de la población, sometida a continuos vientos de levante y de poniente, que, junto con el mar, hacen las veces de sumideros de los gases.

Las partículas están consideradas como el mayor problema de contaminación atmosférica en esta zona, y está provocado sobre todo por la calima africana. Si el valor máximo permitido es de 50 micrómetros por metro cúbico, en esta zona se alcanzan los 55 algunos días. En cambio, las mediciones oficiales permiten deducir de esa carga general la que está provocada por el polvo de las intrusiones saharianas. Después de este cálculo, la provincia se queda en 48 micrómetros. Los valores máximos sólo se han superado de manera puntual.

### Principales contaminantes

El segundo es el ozono, que tiene dos características: la buena, que filtra los rayos ultravioleta; y la mala, que es que, a una concentración por encima de la normal, distorsiona la formación de células en el organismo, así como la fotosíntesis de las plantas. Su producción está directamente relacionada con los gases de escape de los vehículos. En Málaga sólo se superó el valor máximo permitido de manera puntual, un máximo de tres días al año. El límite legal se mide por una media cada ocho horas. Con todo, su presencia se mantiene con frecuencia muy cerca del máximo y por eso hay que vigilarlo. En Sevilla, la alerta por este motivo saltó hasta en cuatro ocasiones el año pasado. Los expertos autonómicos señalan que la única forma de control es la reducción del uso del vehículo privado, especialmente en las zonas más congestionadas de las ciudades.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

## CAMPUS D'ALCOI

La peor situación se produce en verano y durante los episodios de terral, por varios motivos: hay más estabilidad atmosférica, no llueve y las intrusiones saharianas llegan en esa época; además del calor, que también contribuye negativamente.

La Estrategia Andaluza de Calidad del Aire fija para la provincia la necesidad de reducir el tráfico privado en las grandes ciudades, como vía para controlar las emisiones de partículas y el óxido de nitrógeno. A ello, se unen las aportaciones del transporte y de la extracción como las canteras de áridos. Como conclusión, cabe señalar que el aire en Málaga tiene una buena calidad, pero no se puede bajar la guardia. [[Diario Sur](#)]



## La lluvia puede con la contaminación

Las precipitaciones, ausentes en los últimos veinte días, permitieron bajar los niveles de partículas suspendidas en el aire

La Fade cree «injusto» el trato a las empresas, mientras Equo y Ciudadanos llevan al Congreso iniciativas para mejorar la calidad del aire

MARCO MENÉNDEZ

11 enero 2017 01:22

Que la Dirección General de Prevención y Control Ambiental prevea la posibilidad de restringir la actividad industrial en casos de niveles de contaminación muy altos, fue calificado ayer por el director general de la Federación Asturiana de Empresarios (Fade), Alberto González, de «desproporcionado» e «injusto». Apuntó que «da la sensación de que las únicas que contaminan son las empresas, cuando son precisamente las que tienen una mayor preparación de cara a las emisiones y las que mayores controles pasan a la hora de realizar su actividad normal».

González aseguró que «todas las empresas potencialmente contaminantes y las que son consideradas como tal, están sometidas a férreos controles administrativos, que están documentados en la correspondiente autorización administrativa, que es muy complicada de obtener y que se revisa periódicamente por las administraciones públicas». Por ello, considera que «cargar toda la culpa de lo que sucede en Asturias y en España sobre las empresas nos parece desproporcionado e injusto».

Alberto González considera que hay otros factores que influyen en los altos niveles de contaminación: «Es curioso que en puntos donde mayores cotas de contaminación hay no están próximos a las empresas. Eso indica que son varios los factores que están actuando, como el tráfico y las probables emisiones de calefacciones, que tuvieron una actividad frenética estos días y que no son todo lo modernas que debieran ni están controladas».

De cualquier modo, el director general de la Fade apunta que, si hay alguna empresa que no cumpla con los niveles de emisiones «que se atenga a las consecuencias, pero hay otros factores que no se están atendiendo con claridad». Lo que sí ha resultado efectiva contra la contaminación, excepto para el benceno en Trubia y Llaranes, fue la lluvia del día de ayer. Su reaparición tras veinte días de buen tiempo permitió reducir la presencia de partículas en suspensión en el aire. La jornada comenzó con altos niveles de partículas en Avilés, Lugones, La Felguera, Sama y Gijón, pero anoche solo La Felguera tenía una calidad del aire regular.

Por otro lado, la mala calidad del aire ha llegado al Congreso de los Diputados. El líder de Ciudadanos, Albert Rivera, presentará un acuerdo para luchar contra la polución que cuente con



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

## CAMPUS D'ALCOI

el consenso del Gobierno, el Parlamento, las autonomías y los ayuntamientos, «dado que es un problema que no entiende de fronteras, de administraciones ni de ideologías». Más allá fue Equo, pues la diputada de Unidos Podemos Rosa Martínez puso como ejemplos a Madrid y Gijón en una batería de preguntas que presentó al Gobierno de Mariano Rajoy. En concreto, se interesa por la revisión de la relación de contaminantes, los objetivos de calidad del aire, el nuevo Plan sobre la Calidad del Aire, la eliminación de las deducciones por inversiones medioambientales del impuesto sobre sociedades y el intercambio de información entre autonomías. [[El comercio](#)]





# No es solo Madrid: otras ciudades españolas que sufren y combaten la contaminación

Ninguna iguala los niveles de la capital, pero localidades como Barcelona, Granada o Avilés también han sido noticia recientemente por su lucha contra la contaminación

## AUTOR

ROCÍO P. BENAVENTE

Cuando se habla de ciudades contaminadas en España, **Madrid se convierte en protagonista indiscutible**. La gran densidad del tráfico los periodos de estabilidad climática y la falta de corrientes de aire constantes forman sobre la capital esa ya tristemente conocida boina gris y obligan de vez en cuando al ayuntamiento a tomar **medidas de emergencia** como limitar la velocidad del tráfico o el número de coches que pueden acceder al centro de la ciudad.

Pero Madrid **no es la única ciudad** que sufre y combate la contaminación del aire en nuestro país. Aunque en la mayoría de los casos el problema no se acerca en cuanto a gravedad al que sufre la capital, hay otras ciudades que lidian con él con mediciones continuas, planes de emergencia y estrategias a medio y largo plazo.

## Barcelona

Barcelona, igual que Madrid y que otras grandes áreas metropolitanas europeas supera habitualmente los límites establecidos por la UE en cuanto a contaminación del aire. El contaminante más abundante es el dióxido de nitrógeno, procedente en su mayoría de los vehículos diésel.

Esa es la principal fuente de contaminación atmosférica de Barcelona: el tráfico rodado, especialmente los coches con motor diésel y los más antiguos. También contribuyen, aunque en menor medida, las actividades industriales (especialmente las obras) y las nubes de polvo subsahariano que llegan periódicamente a través del Mediterráneo. La orografía en la que se enclava Barcelona, entre la cordillera de Collserola, el puerto, Montjuïc y el río Besòs, provoca que la contaminación se acumule durante días, especialmente en periodos anticiclónicos.

Según recoge la web del ayuntamiento barcelonés, la ciudad sufre entre tres y cinco episodios de contaminación especialmente agudos al año. Es la Generalitat la que los declara y la que pone en marcha las medidas pertinentes: se reduce la velocidad máxima permitida en los accesos a Barcelona, se establecen límites en las emisiones de las grandes industrias y se establecen campañas de comunicación en paneles de tráfico y otros soportes.



A finales de 2016, Barcelona anunciaba algunos **planes que pensaba poner en marcha** para atajar el problema de la contaminación a medio y largo plazo. A partir de abril de este año, los coches de gasolina matriculados antes del año 2000 y los diésel anteriores al 2006 no podrán circular durante los episodios de contaminación, y a partir de 2020 la prohibición podría ser permanente. Además, una particular versión del 'plan renove' incentivaría la sustitución de los vehículos contaminantes con abonos de transporte público gratuitos familiares durante un año.

### Avilés

Los habitantes de Avilés también padecen a causa de la contaminación. La ciudad obtuvo malos resultados en la **evaluación de 2016 de la calidad** del aire que recoge la web del Observatorio de Salud en Asturias. Además, el ayuntamiento se veía obligado a **decretar la prealerta por contaminación** el pasado mes de noviembre después de que se superasen los límites diarios de partículas en suspensión en varias estaciones de la red de mediciones de contaminación atmosférica durante tres días seguidos.

Cuando se habla de ciudades contaminadas en España, **Madrid se convierte en protagonista indiscutible**. La gran densidad del tráfico, los periodos de estabilidad climática y la falta de corrientes de aire constantes forman sobre la capital esa ya tristemente conocida boina gris y obligan de vez en cuando al ayuntamiento a tomar **medidas de emergencia** como limitar la velocidad del tráfico o el número de coches que pueden acceder al centro de la ciudad.



Parche contra la contaminación en Madrid: así la combaten otros países

### SERGIO FERRER

El problema del tráfico y sus consecuencias contra la salud no es exclusivo de la capital española. Otras ciudades del resto del mundo han introducido medidas similares en momentos de necesidad





Pero Madrid **no es la única ciudad** que sufre y combate la contaminación del aire en nuestro país. Aunque en la mayoría de los casos el problema no se acerca en cuanto a gravedad al que sufre la capital, hay otras ciudades que lidian con él con mediciones continuas, planes de emergencia y estrategias a medio y largo plazo.

### **Barcelona**

Barcelona, igual que Madrid y que otras grandes áreas metropolitanas europeas supera habitualmente los límites establecidos por la UE en cuanto a contaminación del aire. El contaminante más abundante es el dióxido de nitrógeno, procedente en su mayoría de los vehículos diésel.

Esa es la principal fuente de contaminación atmosférica de Barcelona: el tráfico rodado, especialmente los coches con motor diésel y los más antiguos. También contribuyen, aunque en menor medida, las actividades industriales (especialmente las obras) y las nubes de polvo subsahariano que llegan periódicamente a través del Mediterráneo. La orografía en la que se enclava Barcelona, entre la cordillera de Collserola, el puerto, Montjuïc y el río Besòs, provoca que la contaminación se acumule durante días, especialmente en periodos anticiclónicos.



Según recoge la web del ayuntamiento barcelonés, la ciudad sufre entre tres y cinco episodios de contaminación especialmente agudos al año. Es la Generalitat la que los declara y la que pone en marcha las medidas pertinentes: se reduce la velocidad máxima permitida en los accesos a Barcelona, se establecen límites en las emisiones de las grandes industrias y se establecen campañas de comunicación en paneles de tráfico y otros soportes.

A finales de 2016, Barcelona anunciaba algunos **planes que pensaba poner en marcha** para atajar el problema de la contaminación a medio y largo plazo. A partir de abril de este año, los coches de gasolina matriculados antes del año 2000 y los diésel anteriores al 2006 no podrán



circular durante los episodios de contaminación, y a partir de 2020 la prohibición podría ser permanente. Además, una particular versión del 'plan renove' incentivaría la sustitución de los vehículos contaminantes con abonos de transporte público gratuitos familiares durante un año.

### **Avilés**

Los habitantes de Avilés también padecen a causa de la contaminación. La ciudad obtuvo malos resultados en la evaluación de 2016 de la calidad del aire que recoge la web del Observatorio de Salud en Asturias. Además, el ayuntamiento se veía obligado a decretar la prealerta por contaminación el pasado mes de noviembre después de que se superasen los límites diarios de partículas en suspensión en varias estaciones de la red de mediciones de contaminación atmosférica durante tres días seguidos.

En concreto, varios medidores registraron niveles de partículas superiores a los límites establecidos. La acumulación y persistencia de contaminación se debió a la estabilidad meteorológica. Normalmente, el aire tanto en horizontal por la acción del viento como en vertical por las diferencias de temperatura. El aire caliente está abajo y se enfría a medida que sube. En otoño e invierno, cuando se producen anticiclones que pueden durar varios días, estos procesos se detienen. El pasado mes de noviembre, a esto se sumó otro fenómeno natural que aumentó la cantidad de partículas en el aire: las nubes de polvo sahariano.

La Consejería de Medio Ambiente de Asturias declaró el nivel de prealerta en la zona de Avilés. El protocolo activado consiste en informar a la población de la zona afectada, en el control de las emisiones provenientes de la industria y de las actividades portuarias, limitar la circulación de vehículos de más de 3.500 kilos, especialmente los que transporten material polvoriento, baldear las carreteras, limitar el encendido de calefacciones con combustibles sólidos y prohibir las quemas agrícolas.

### **Granada**

Granada es otra ciudad donde los problemas de contaminación ambiental son recurrentes en los últimos años. Sus estaciones de medición registran periódicamente niveles superiores a los permitidos de dióxido de nitrógeno que, si bien siguen siendo inferiores a los de Madrid y Barcelona, indican que los granadinos conviven con un riesgo cotidiano para su salud.

La principal fuente de contaminación en Granada es el tráfico, seguida de las calefacciones. Según este artículo de 'El Ideal', se estima que el 40% de las comunidades de propietarios utilizan calderas de gasóleo. Eso, unido a las escasas precipitaciones anuales y la falta de corrientes de aire, provoca la concentración de contaminantes y dificulta su dispersión.

A finales de 2016 el ayuntamiento descartó medidas de restricción al tráfico, y declaró su intención de extender las líneas de transporte público para favorecer que los granadinos dejen el coche en casa, así como un "plan de concienciación que permita el cambio de las calderas en las comunidades de vecinos".



## Valencia

A finales de 2016, Joan Ribó, alcalde de Valencia, anunciaba que la Concejalía de Medio Ambiente estaba **trabajando en un protocolo de emergencia** por contaminación ambiental a pesar de que en la ciudad no se dan episodios relevantes de **superación de los límites establecidos**. Un año antes, en septiembre de 2015, Valencia sí presentaba **concentraciones de contaminación que superaban la normativa** y los niveles saludables.



La causa principal, igual que en Madrid y Barcelona, era el tráfico, unido a una situación anticiclónica que dificultaba la dispersión de las partículas contaminantes. Habitualmente, cuando los niveles de contaminación son menores, el aire proveniente del mar dispersa las partículas y por lo tanto los problemas de contaminación no son frecuentes.

A pesar de ello, durante la alcaldía de Rita Barberá los medidores se alejaron del centro de la ciudad, donde se concentra la mayor parte del tráfico, y se colocaron en zonas más verdes y espacios abiertos donde las concentraciones de partículas contaminantes son siempre menores. La oposición **pidió a la Comisión Europea** que investigase los niveles de contaminación reales que soportaban los valencianos, asegurando que eran mayores de lo que revelaban las cifras oficiales

## La Línea de la Concepción

En 2014, la Organización Mundial de la Salud publicaba un informe en el que analizaba los **niveles de partículas pequeñas en suspensión de miles de ciudades** en el mundo, entre ellas 46 españolas. De estas últimas, La Línea de la Concepción, en Cádiz, **aparecía en primer lugar**. Las causas de la menor calidad del aire eran las fábricas y el intenso tráfico de vehículos para cruzar a Gibraltar.

La Junta de Andalucía cuenta con planes de mejora de la calidad del aire aprobados para determinadas áreas de la comunidad donde esta cuestión es especialmente problemática. Entre ellas se encuentra la Bahía de Algeciras, que engloba la Línea de la Concepción. En ese plan se incluyen medidas como mejorar la cobertura de líneas de bus urbano y renovar su flota,



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

fomentar los desplazamientos a pie y el uso del carril bici, limpiar las vías de circulación para evitar que la circulación levante polvo y fomentar la conducción eficiente. [El confidencial]



## Madrid desactiva el lunes todas las restricciones al tráfico por contaminación

A lo largo de este domingo han continuado las restricciones de velocidad y la recomendación de no aparcar en la zona de estacionamiento regulado (SER)



AUTOR EC

01.01.2017 – 11:37 H.

El Ayuntamiento de Madrid ha comunicado este domingo que, debido a una "leve" mejora de los niveles de concentración de dióxido de nitrógeno (NO<sub>2</sub>), **desactiva a medianoche del domingo todas las restricciones al tráfico** contempladas en el protocolo contra la contaminación. Eso sí, a lo largo del día 1 se han mantenido **las restricciones de velocidad** y la recomendación de no aparcar en la zona de estacionamiento regulado (SER), según recoge el Ayuntamiento en un comunicado.

El sábado, el Consistorio informó de que los niveles de contaminación por dióxido de nitrógeno continúan siendo elevados -por encima de los 200 microgramos por centímetro cúbico-, por lo que decidió elevar este domingo las restricciones al tráfico, al pasar del denominado Escenario 1 al Escenario 2. Dado que el Servicio de Estacionamiento Regulado no funciona al ser festivo, el Ayuntamiento **aconseja no aparcar en la zona centro** y recomienda para los desplazamientos por la ciudad la utilización del transporte público.

La ciudad de Madrid redujo en 2016 un 10% los valores medios de dióxido de nitrógeno (NO<sub>2</sub>), mientras que los niveles de ozono troposférico (O<sub>3</sub>) u "ozono malo" han sido similares a 2015, según los datos de la red municipal de calidad del aire.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

## CAMPUS D'ALCOI

La meteorología, **con menos periodos estacionales y olas de calor que otros años**, y en menor medida el protocolo municipal anticontaminación y la sensibilización ciudadana, entre otros factores, contribuyeron a esa mejora. No obstante, la capital incumplió por séptimo año consecutivo la normativa europea referente a los límites para la protección humana de NO<sub>2</sub>, un gas procedente principalmente de los tubos de escape de los vehículos con motor diésel que irrita las vías respiratorias.

De las 24 estaciones de medición distribuidas por cinco zonas de la ciudad, **nueve rebasaron en 2016 el límite establecido en 40 microgramos** por metro cúbico de concentración media anual, frente a las trece del año anterior. En 2014 fueron seis.

Con datos actualizados hasta el 26 de diciembre, en el conjunto de las estaciones de medición se habían registrado 182 picos de más de 200 microgramos de media en una hora, mientras que en 2015 fueron 554. [\[El confidencial\]](#)



## Zaragoza no tiene protocolo al tener escasa contaminación

El clima y el tipo de movilidad hace que esté muy lejos de los niveles peligrosos

### EL PERIÓDICO

30/12/2016



El Ayuntamiento de Zaragoza no tiene establecido un protocolo de medidas concretas para actuar ante altos índices de contaminación porque no se llega a los parámetros máximos de 200 microgramos por metro cúbico de dióxido de nitrógeno (NO<sub>2</sub>) en el aire establecidos por la Organización Mundial de la Salud (OMS) y la Unión Europea (UE). De hecho, la capital aragonesa queda muy lejos de esos niveles mínimos exigidos para la alerta, puesto que en la ciudad es una excepción alcanzar los 100.

Fuentes municipales indicaron a Efe que tampoco se llega en la ciudad al valor máximo de partículas de suspensión en el aire, establecido en 50, ya que el valor más alto de la media diaria se registra en 35. Por ello, insistieron en que el consistorio no tiene establecido ningún protocolo para actuar contra la contaminación, si bien apuntaron que se acatará el requerimiento realizado por el Defensor del Pueblo, institución a la que remitirán toda la información que solicita.

Zaragoza forma parte, junto a Avilés, La Coruña, Barcelona, Bilbao, Ciudad Real, Granada, León, Madrid, Murcia, Santander, Sevilla, Toledo y Valencia, de los catorce municipios sobre los que el Defensor del Pueblo ha iniciado una actuación de oficio para conocer las medidas que contemplan sus ayuntamientos contra la contaminación atmosférica.



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

## CAMPUS D'ALCOI

Desde el consistorio zaragozano añadieron que, aunque por el momento no se ha detectado la necesidad de actuar, existen medidas de prevención y mejora de la calidad del aire a largo plazo que tienen como finalidad «ir a más» y que son «en positivo», como las que intentan fomentar los desplazamientos en bicicleta o el uso del tranvía.

### **USO DEL TRANVÍA Y LA BICI**

En este sentido destacaron que uno de los principales motivos por los que han descendido un 18 % los niveles de CO<sub>2</sub> desde el 2005 hasta ahora se debe que el transporte ha cambiado, con la implantación del tranvía y del uso de la bicicleta, así como por la mayor eficiencia de los sistemas térmicos, como la rehabilitación energética, el uso del gas ciudad o la creación de barrios eficientes.

A estos factores también se le une un entorno rural y el hecho de que haya muchos días de viento que dispersa las partículas. [\[El diario\]](#)





## Disminuir la contaminación es proteger la salud

Aunque los responsables municipales lo hayan obviado hasta el presente, hace tiempo que se conoce que la contaminación producida por los vehículos de motor es un serio problema de salud pública que influye negativamente sobre la morbilidad y mortalidad de la población, y las evidencias científicas no cesan de incrementarse

Marciano Sánchez Bayle - Asociación para la Defensa de la Sanidad Pública de Madrid

17/01/2017 - 20:43h



El Ayuntamiento de Madrid parece que por fin se ha tomado en serio la contaminación que sufre crónicamente la ciudad y ha comenzado a limitar la circulación por el centro de la ciudad, después de más de 20 años dedicados a todo lo contrario. Las obras faraónicas que nos han endeudado a la ciudadanía han hecho todo lo posible para facilitar el acceso fácil de los vehículos particulares al centro de Madrid, consolidando una cultura de atasco, polución y despilfarro económico y energético.

Lógicamente va a costar revertir esta situación y no solo porque la señora Aguirre tenga más dificultades para dejar el coche en doble fila en la Gran Vía, o "se quede encerrada en su casa", sino sobre todo porque muchas personas han organizado su movilidad urbana en torno al vehículo privado y les va a costar adaptarse a la nueva situación. Ahí es donde hace falta mucha pedagogía e información por parte del Ayuntamiento y, por supuesto, mejorar de una manera significativa el transporte público.

Aunque los responsables municipales lo hayan obviado hasta el presente, hace tiempo que se conoce que la contaminación producida por los vehículos de motor es un serio problema de salud pública que influye negativamente sobre la morbilidad y mortalidad de la población, y las evidencias científicas no cesan de incrementarse.



Ahora mismo se sabe que los niveles de los contaminantes ambientales influyen en la salud desde el estado embrionario, produciendo bajo peso al nacer, prematuridad, microcefalia, y alteraciones en el epigenoma, etc; posteriormente, durante la infancia, producen aumento de los problemas respiratorios y agravamiento de los mismos, asma, disminución de la función y el desarrollo pulmonar, inicio de arteriosclerosis y alteraciones del neurodesarrollo y el comportamiento, y en la edad adulta se han evidenciado declive de la función pulmonar, aumento del cáncer de pulmón, diabetes tipo 2, arteriosclerosis, eventos coronarios, accidentes cerebrovasculares, y un largo etc de patologías.

Se ha calculado que los efectos de la contaminación sobre la salud producirían un coste de entre 243 y 712 millones de euros anuales en la UE. Por poner solo un modesto ejemplo, en Madrid alrededor del 9% de los ingresos pediátricos podrían evitarse si no se superasen los límites establecidos de los contaminantes ambientales.

Y, además, es importante tener en cuenta que, aunque se han establecido topes de alerta para los distintos contaminantes, su actuación sobre la salud se produce con niveles inferiores y es creciente según se incrementan los valores de los mismos, y que su efecto es mucho mayor en la infancia y en las personas mayores. O sea, que garantizar un medio ambiente saludable es también velar por la salud de nuestros niños y del conjunto de las personas en el medio y largo plazo.

Obviamente cuando se producen epidemias de virus respiratorios –ahora la de gripe–, la contaminación empeora notablemente la situación, favoreciendo más ingresos, más complicaciones y más mortalidad.

Por todo eso ya era hora de que alguien desde nuestras administraciones públicas se tomase en serio este importante problema. Aunque es obvio que ha habido problemas en la aplicación de las medidas que se han tomado, se trata de un primer paso, que habrá que continuar, perfilándolo mejor, y sobre todo pensando en el medio plazo, en la prevención. porque no es razonable la situación de Madrid con unos niveles de contaminantes que siempre andan rozando los valores límites establecidos.

Hay que avanzar planes hacia entornos sin contaminación o con una contaminación mínima, pero también hay que tener en cuenta que para solucionarlo de manera definitiva es crucial que la ciudadanía tome conciencia del problema y colabore activamente. [\[El periódico de Aragón\]](#)



**Variables explicadas:**

PROVINCIA	SO <sub>2</sub>	CO <sub>2</sub>	NO <sub>x</sub>
Araba/Álava	5,636221417	53236,38	30,67549607
Albacete	5,502132827	209323,5024	16,11842981
Alicante/Alacant	3,192159948	69711,54807	22,23525995
Almería	7,04340012	79012,10224	20,31781892
Ávila	1,542179224	173082,736	14,00428611
Badajoz	1,83991338	329379,4215	6,312260664
Balears, Illes	4,015896812	62905,08678	15,76970337
Barcelona	2,518807883	78637,5147	15,76970337
Burgos	3,457057472	217664,2194	21,37126327
Cáceres	2,673268337	368213,8112	8,047855016
Cádiz	9,740070211	60042,63483	22,89824434
Castellón/Castelló	3,938860623	75312,87617	15,95300675
Ciudad Real	6,70205018	246663,5819	19,70663353
Córdoba	8,918968565	175400,2348	27,71263983
Coruña, A	3,122569483	93140,19339	14,80688685
Cuenca	3,751419405	265231,0457	32,46196501
Girona	1,480561478	77060,08326	29,45492612
Granada	7,934299419	162556,7778	43,97413605
Gipuzkoa	4,422581845	23919,43209	44,62588669
Huelva	5,812423207	100976,8543	16,21581422
Huesca	4,452105947	248061,3889	22,15216626
Jaén	4,85527922	203120,5058	27,55849889
León	5,602198325	193614,5107	18,63560334
Lleida	1,260832581	151663,5068	14,1005123
Rioja, La	3,088582586	80046,1134	15,8889587
Lugo	5,684659087	199453,4524	12,57818982
Madrid	3,740182117	104409,345	51,5126945
Málaga	8,299530497	57149,23102	36,1820603
Murcia	9,124376213	133433,1184	37,1447719
Navarra	4,682101806	159475,6267	24,76056702
Ourense	1,678537851	126556,133	32,41263814
Asturias	6,497337988	125358,4603	37,3487425
Palmas, Las	5,87147881	57241,43952	16,37377364
Pontevedra	3,453941983	47118,82563	29,0294643



Salamanca	3,291871521	206736,0281	19,48479441
Santa Cruz de Tenerife	5,10502627	40125,52518	19,22865284
Cantabria	1,398098489	58157,33617	26,62645328
Segovia	1,69445093	126451,6177	17,77795262
Sevilla	5,163344844	134332,7358	33,86782761
Soria	1,846700864	161957,2027	26,59926968
Tarragona	2,009126616	72084,07144	19,9233504
Teruel	4,78465651	254661,6362	11,59691374
Toledo	3,03915246	238601,3667	21,94979122
Valencia/València	3,219207779	116330,9137	26,10655196
Valladolid	3,409757695	76860,7817	37,29434549
Bizkaia	7,129967436	29000,51634	38,60248158
Zamora	0,957013608	199480,5036	11,08609935
Zaragoza	2,752623548	260651,2463	35,01810243



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

**Variables explicativas:**

PROVINCIA	TEMP	LLU	SOL	VI	P	O	E	PC	VE	AU
Araba/Álava	12,79166667	76,65	166,0583333	33,25	321.932	135,9	19.652	4357	201656	523
Albacete	15,575	23,40833333	264,45	57,8	396.987	658,7	26.184	6932	264817	384
Alicante/Alacant	19,23333333	12,71666667	272,8	45,50746269	1.868.438	228,4	126.389	33079	1237801	1551
Almería	19,84166667	10,55	268,6083333	57,22222222	701.688	133,7	39.530	11852	478614	721
Ávila	12,15833333	33,625	223,85	48,16666667	167.015	376,3	10.046	2208	123776	118
Badajoz	17,825	38,24166667	244,5416667	47,85714286	690.929	55,8	38.631	9773	466546	670
Balears, Illes	18,525	41,285	224,641667	40,47058824	1.103.442	216,8	84.270	21270	914219	2383
Barcelona	17,56666667	43,21666667	170,85	21,52564103	5.523.784	482,9	434.914	73523	3489947	5902
Burgos	11,425	47,13333333	193,1166667	49,2	366.900	2.234,5	24.761	5120	247667	458
Cáceres	16,66666667	51,93333333	255,0416667	40,71428571	408.703	449,6	24.298	5070	295358	573
Cádiz	18,725	53,03333333	247,5	31,66666667	1.240.175	148,7	56.944	18907	724204	1061
Castellón/Castelló	18,75833333	23,9	252,2583333	49,36842105	587.508	137,1	38.084	8842	407598	471
Ciudad Real	16,58333333	28,53333333	250,9	40,25	519.613	339,0	29.344	7029	337884	462
Córdoba	18,55	50,16666667	244,75	27,375	799.402	225,6	45.261	11184	504827	701
Coruña, A	14,7375	80,5541667	152,454167	64,42857143	1.132.735	208,9	79.854	14248	739089	1660
Cuenca	14,35833333	45,38333333	203,4363636	39,75	207.449	162,8	13.352	4780	163730	341
Girona	15,725	77,29166667	216,0666667	29,6440678	756.156	247,5	57.658	11061	592336	1069
Granada	16,08333333	30,90833333	247,9333333	28,76923077	919.455	430,9	54.404	18370	635614	1134
Gipuzkoa	14,40833333	143,775	156,05	55,375	715.148	282,0	52.042	8565	441604	904
Huelva	18,65	46,19166667	266,2916667	37,8	519.229	303,4	23.532	6955	314725	652



Huesca	15,49166667	49,7	225,55	48,5	224.909	285,4	15.975	3418	170833	346
Jaén	17,55	37,625	255,025	29,18181818	659.033	102,6	32.704	10001	427741	523
León	11,325	53,14166667	156,9666667	48,53333333	484.694	169,0	31.134	5527	346139	658
Lleida	16,325	35,55833333	245,3	34,54166667	438.001	88,4	33.354	6237	323495	493
Rioja, La	14,86666667	35,575	199,4083333	61	319.002	197,7	22.314	4938	197174	256
Lugo	12,90833333	95,74166667	150,2916667	49,66666667	342.748	161,9	23.703	4435	261129	771
Madrid	16,06666667	34,775	224,25	47,64705882	6.454.440	183,4	494.509	78928	4220253	10232
Málaga	19,86666667	31,09166667	244,8416667	42,23809524	1.621.968	125,4	107.851	25905	1067495	1627
Murcia	20,15	15,54166667	262,9666667	42,04545455	1.466.818	2.715,0	86.782	23981	976142	1658
Navarra	13,98333333	72,16666667	173,8166667	35,66666667	640.790	538,5	41.582	8004	431363	817
Ourense	15,69166667	83,08333333	170,375	37,75	322.293	530,1	22.663	3871	241113	964
Asturias	14,06666667	92,66666667	152,7333333	56,78947368	1.061.756	262,5	66.342	12437	658733	1384
Palmas, Las	21,75	10,74166667	174,4727273	27	1.100.027	58,3	66.754	16575	756237	2561
Pontevedra	15	154,1833333	162,4833333	49,57142857	950.919	383,5	65.525	11672	668222	1192
Salamanca	12,65	26,36666667	239,7916667	41,5	342.459	336,8	22.303	4804	223532	501
Santa Cruz de Tenerife	21,70833333	35,38333333	254,3	25,28571429	1.004.788	128,2	61.764	16307	740025	2738
Cantabria	15,63333333	95,96666667	126,275	63,16666667	588.656	123,8	36.698	7017	395513	602
Segovia	12,99166667	40,14166667	218,0083333	42,6	159.303	355,3	10.664	3407	119903	280
Sevilla	19,54166667	52,875	274,875	48,1	1.941.355	67,1	108.704	24712	1172561	1939
Soria	11,65833333	50,025	220,525	40	92.221	628,8	5.794	1120	70816	100
Tarragona	18,65833333	53,2	215,2583333	57,35294118	800.962	36,7	50.639	10731	559183	1060
Teruel	13,78333333	24,16666667	207,525	48,14285714	140.365	309,6	9.007	2236	106960	119



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

Toledo	16,75833333	32,53333333	260,2166667	53	699.136	51,0	41.370	11583	489387	701
Valencia/València	19,51666667	24,06666667	227,6083333	46,28169014	2.548.898	242,6	166.382	35969	1630002	2358
Valladolid	13,61666667	36,15	230,5833333	61,2	529.157	933,4	32.790	6365	329291	553
Bizkaia	16,09166667	89,275	122,6	40,5	1.151.905	206,7	77.551	12392	651305	1751
Zamora	14,09166667	30,70833333	211,6666667	35,25	185.432	61,3	11.662	2255	134846	234
Zaragoza	17,01666667	26,5	233,20	52,57142857	960.111	379,9	63.132	11945	549541	1042



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

CAMPUS D'ALCOI

## **9. Bibliografía:**





CAMPUS D'ALCOI

- Andalucía web [http://www.andalucia-web.net/clima\\_cadiz.htm](http://www.andalucia-web.net/clima_cadiz.htm) 16/02/2017 21 de enero de 2017
- Apuntes de econometría del curso 2015. Profesores Bárbara Micó y Jorge Jordán.
- Causas y consecuencias de...  
<http://causas-consecuencias.com/causas-de-contaminacion-atmosferica/>, 29 de abril de 2017
- Certicalia, <https://www.certicalia.com/blog/emisiones-globales-por-comunidad-provincia>, 1 de marzo de 2017
- Climate data, <http://www.climatedata.eu/climate.php?loc=spxx0050&lang=es> 16 de febrero 2017
- Conciencia Cero, <http://www.concienciaeco.com/2012/02/20/que-es-el-co2/>, 6 de mayo de 2017
- DGT, <http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/parque-vehiculos/tablas-estadisticas/> 21 de enero de 17
- DGT, <http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/permisos-conduccion/> 21 de enero 2017
- Diario Sur, <http://www.diariosur.es/malaga-capital/201612/28/malaga-cost-tienen-niveles-20161227215548.html> ,14 de enero de 2017
- EcuRed, [https://www.ecured.cu/Contaminaci%C3%B3n\\_del\\_aire](https://www.ecured.cu/Contaminaci%C3%B3n_del_aire) 29 de abril de 2017
- El comercio,  
<http://www.elcomercio.es/asturias/201701/11/lluvia-puede-contaminacion-20170111001528-v.html> , 14 de enero de 2017
- El confidencial, [http://www.elconfidencial.com/espana/madrid/2017-01-01/madrid-contaminacion-restricciones-lunes-2-enero\\_1311083/](http://www.elconfidencial.com/espana/madrid/2017-01-01/madrid-contaminacion-restricciones-lunes-2-enero_1311083/) 17 de enero de 2017
- El confidencial, [http://www.elconfidencial.com/tecnologia/2017-01-15/no-es-solo-madrid-las-otras-ciudades-que-sufren-y-combaten-la-contaminacion\\_1316261/](http://www.elconfidencial.com/tecnologia/2017-01-15/no-es-solo-madrid-las-otras-ciudades-que-sufren-y-combaten-la-contaminacion_1316261/) 17 de enero de 2017
- El diario, [http://www.eldiario.es/tribunaabierta/Disminuir-contaminacion-proteger-salud\\_6\\_602699758.html](http://www.eldiario.es/tribunaabierta/Disminuir-contaminacion-proteger-salud_6_602699758.html) , 19 de enero de 2017
- El periódico de Aragón,  
[http://www.elperiodicodearagon.com/noticias/aragon/zaragoza-no-tiene-protocolo-tener-escasa-contaminacion\\_1170182.html](http://www.elperiodicodearagon.com/noticias/aragon/zaragoza-no-tiene-protocolo-tener-escasa-contaminacion_1170182.html) , 17 de enero de 2017
- Fundación Crana,  
[http://www.crana.org/es/contaminacion/mas-informacion\\_3/diaxido-azufre-so2](http://www.crana.org/es/contaminacion/mas-informacion_3/diaxido-azufre-so2), 6 de mayo de 2017)
- Gobierno de España, Ministerio de Agricultura y Pesca, Alimentación y medio Ambiente,  
<http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/salud/dioxido-azufre.aspx> , 6 de mayo de 2017
- Gobierno de España, Ministerio de Agricultura y Pesca, Alimentación y medio Ambiente,  
<http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/salud/monoxido-carbono.aspx> , 7 de mayo de 2017



CAMPUS D'ALCOI

- INE, <http://www.ine.es/jaxi/Tabla.htm?path=/t38/p604/a2000/l0/&file=1900001.px&L=0> 21 de enero 2017
- INE, <http://www.ine.es/jaxiT3/Tabla.htm?t=3991&L=0> 21 de enero de 2017
- INE, <http://www.ine.es/jaxiT3/Tabla.htm?t=2852> , 21 de enero 2017
- INE, <http://www.ine.es/jaxiT3/Tabla.htm?t=3996> 21 de enero 2017
- INE <http://www.ine.es/jaxiT3/Tabla.htm?t=2852> 21 de enero 2017
- Mapama,  
[http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/evaluacion-datos/datos/Datos\\_2001\\_2014.aspx](http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/evaluacion-datos/datos/Datos_2001_2014.aspx), 1 de marzo de 2017
- La razón, <http://www.larazon.es/economia/las-restricciones-no-evitan-el-suspenso-en-contaminacion-MP14287290>, 13 de enero de 2017
- Liceo Marta Donoso Espejo,  
[http://instmat.utsalca.cl/tem/sitiolmde/primer/guiasliceo/recuperacion/Diagrama de Caja y Bigotes-2.pdf](http://instmat.utsalca.cl/tem/sitiolmde/primer/guiasliceo/recuperacion/Diagrama_de_Caja_y_Bigotes-2.pdf) 9 de mayo de 2017
- Newbold, P; L. Carlson, W; Thorne, B. 2008. Estadística para administración y economía. Pearson Educación, S.A., Madrid, España
- Universo fórmulas <http://www.universoformulas.com/estadistica/descriptiva/curtosis/> 8 de mayo de 2017
- Uva, <http://www5.uva.es/estadmed/datos/bivariante/bivar.htm>, 10 de mayo de 2017
- Wikipedia, [https://es.wikipedia.org/wiki/Mediana\\_\(estad%C3%ADstica\)](https://es.wikipedia.org/wiki/Mediana_(estad%C3%ADstica)) 8 de mayo de 2017
- Wikipedia <https://es.wikipedia.org/wiki/Curtosis> 8 de mayo de 2017
- Wikipedia [https://es.wikipedia.org/wiki/Gr%C3%A1fico de probabilidad normal](https://es.wikipedia.org/wiki/Gr%C3%A1fico_de_probabilidad_normal), 10 de mayo de 2017