



Análisis de las emisiones de CO₂ en España en el último septenio

MEMORIA PRESENTADA POR:

[Borja Micó Moltó]

GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESA

Convocatoria de defensa: Julio 2018



Tabla de contenido

1. INTRODUCCIÓN Y OBJETIVOS	4
1.1 Introducción y motivación	5
1.2 Hipótesis	6
1.3 Objetivos	7
2. METODOLOGÍA Y TÉCNICAS	8
2.1 Metodología	9
2.2 Técnicas	10
2.2.1 Análisis Univariante	10
2.2.1.1 Resumen de Estadísticos	10
2.2.1.2 Gráfico de Caja y Bigotes	12
2.2.2 Análisis Bivariante	12
2.2.2.1 Gráfico X-Y	12
2.2.2.2 Multicolinealidad	12
2.2.3 Análisis multivariante	14
2.2.3.1 Significatividad	14
2.2.3.2 Normalidad de residuos	15
2.2.3.3 Heterocedasticidad	16
2.2.3.4 Autocorrelación	16
2.2.3.5 Puntos influyentes	18
3. CALCULOS, RESULTADOS Y EXPLOTACIÓN	19
3.1 ANALISIS UNIVARIANTE	21
3.2 BIVARIANTE	35
3.2.1 MULTICOLINEALIDAD	41
3.3 Multivariante	45
3.3.1 Ajuste del modelo	45
3.3.2 Normalidad de residuos	46
3.3.3 Histograma de residuos	47
3.3.4 Contraste de normalidad	47
3.3.5 Heterocedasticidad	48
3.3.6 Autocorrelación	48
3.3.7 Puntos influyentes	50
3.3.8 Puntos anómalos	51
3.3.9 Solución autocorrelación	51
3.4 Predicción	53
4. CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN	58



4.1 Conclusiones	59
4.2 Futuras líneas de investigación	60
5. Bibliografía	62
6. ANEXO	63
6.1 Índice de tablas	66
6.2 Índice de ilustraciones	67



1. *INTRODUCCIÓN Y OBJETIVOS*

1.1 Introducció i motivació

Una de les majors causes perjudicials de contaminació per al planeta és a lo que se li llama al calentament global o efecte invernadero, i uno de les majors responsables de esto és el Diòxid de carboni (CO_2 segun su composició química). El ser humano contribueix molt a la producció del CO_2 mitjançant se accions, i esto no ajuda para nada al planeta i sin darse cuenta se esta perjudicando, destruyendo su hogar, la Tierra.

La producció de diòxid de carboni aumenta año tras año desde la revolució industrial, especialmente en los años en los que se empezó a abusar de la combustión de los derivados del petróleo, donde más del 75% de las emisiones de CO_2 provocadas por el hombre proceden de los combustibles fósiles, en los que incluimos el gas natural, el carbón y el petróleo. [1]



Il·lustració 1.1 Emisiones CO_2 .

La elecció per la realizació de este proyecto sobre las emisiones de CO_2 fue por la motivación de conocer a modo personal y hacer conocer un tema importantísimo hoy en día para todos, ya que afecta gravemente al planeta tierra, y que a través de los años se han ido desarrollando nuevas fuentes de energía donde el consumo de ellas provoca unas emisiones de CO_2 menores que otras fuentes de energía tradicionales, aunque hoy en día siguen siendo de uso minoritario. Los seres humanos están avocados a evolucionar o a destruir el planeta con sus propias acciones. Por lo tanto, la sociedad debe empezar a concienciarse ya que están en sus manos el futuro del planeta y sus habitantes.

1.2 Hipòtesis

1. Explicar la variable emisiones de CO₂ en España a partir de otras variables explicativas.
2. Comprobar si la variable consumo de energías renovables afecta directamente al modelo.
3. Comprobar si la variable consumo de productos petrolíferos influye directamente con las emisiones de CO₂ en España.
4. Comprobar si la variable consumo de energía eléctrica provoca un aumento en las emisiones de CO₂ en España.

1.3 Objectivos

El principal objetivo de este trabajo es la de analizar y estudiar la conducta de la variable explicada, emisiones de CO₂ en España, y ver si existe alguna relación con las variables explicativas, consumo de carbón, consumo de gases derivados del carbón, consumo de productos petrolíferos, consumo de gas natural y manufacturados, consumo de electricidad y consumo de energías renovables.

Donde para lograrlo se han determinados los siguientes subobjetivos para alcanzarlo:

- Elaborar un análisis descriptivo que pueda ser definido a través de las variables las emisiones de CO₂.
- Señalar mediante el estudio de multicolinealidad si existe relación entre las variables explicativas.
- Realizar el ajuste y la validación de un modelo econométrico con las variables explicativas deducidas del análisis de multicolinealidad relacionadas a las emisiones de CO₂.
- Confeccionar predicciones con el modelo definitivo para observar sus futuras condiciones.



2. *METODOLOGÍA Y TÉCNICAS*

2.1 Metodología

La realización de este proyecto se va a desglosar de la siguiente manera.

1. Se va a proceder a la búsqueda de los datos.
2. Se realizará el análisis univariante donde se calcularán todos sus parámetros.
3. Se efectuará el análisis bivariante donde se analizará cada correlación.
4. Se realizará un modelo de regresión, donde se procederá con un ajuste del modelo y su pertinente validación.
5. Finalmente se efectuarán las predicciones.

2.2 Tècniques

2.2.1 Anàlisi Univariante

El anàlisi univariante analiza las directrices de cada variable independientemente, por lo que se examinará el resumen de estadísticos.

2.2.1.1 Resumen de Estadísticos

Media aritmética, también llamada promedio o media, de un conjunto finito de números es el valor característico de una serie de datos cuantitativos, objeto de estudio que parte del principio de la esperanza matemática o valor esperado, se obtiene a partir de la suma de todos sus valores dividida entre el número de sumandos. [2]

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

Mediana, se denomina mediana en estadística al valor que se encuentra en el lugar central de todos los datos de un estudio cuando éstos están ordenados de menor a mayor. al valor que se encuentra en el lugar central de todos los datos de un estudio cuando éstos están ordenados de menor a mayor. El símbolo de la mediana se representa por M_e . La mediana es por tanto el número central de un grupo de números ordenados por su tamaño. [3]

$$Me = \begin{cases} X_{\frac{n+1}{2}} & \text{si } n \text{ es impar} \\ \left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right) * \frac{1}{2} & \text{si } n \text{ es par} \end{cases}$$

Rango, es el intervalo entre el valor máximo y el valor mínimo; por ello, comparte unidades con los datos. Permite obtener una idea de la dispersión de los datos, cuanto mayor es el rango, más dispersos están los datos (sin considerar la afectación de los valores extremos). [4]

$$R = X_{Max} - X_{Min}$$

1er cuartil (Q1), es el valor mayor que el 25% de los valores de la distribución.

2do cuartil (Q2), es la mediana de la distribución 50%, es el valor de la variable que ocupa el lugar central en un conjunto de datos ordenados.

3er cuartil (Q3), es el valor que sobrepasa al 75% de los valores de la distribución.

Rango Inter-cuartil, es la distancia entre el primer 1er cuartil y el 3er cuartil (Q3-Q1); de esta manera, abarca el 50% central de los datos. [5]



Ilustración 2.1 Los cuartiles separan los datos en cuatro segmentos con igual número de datos.

Varianza, es aquella medida de dispersión que ostenta una variable aleatoria respecto a su esperanza. La varianza se relaciona con la desviación típica o desviación estándar. [6]

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

Desviación típica, es la raíz cuadrada de la varianza de la variable. Sirve principalmente para conocer la desviación que presentan los datos en su distribución respecto a la media aritmética de dicha distribución. Es decir, nos muestra una visión más acorde con la realidad en el momento de tomar las decisiones. [7]

$$\sqrt{\sigma^2}$$

Mínimo, es el dato menor de la variable.

Máximo, es el dato mayor de la variable.

2.2.1.2 Gráfico de Caja y Bigotes

El gráfico de caja y bigotes son una presentación visual que describe varias características importantes, al mismo tiempo, tales como la dispersión y simetría.

Para su realización se representan los tres cuartiles y los valores mínimo y máximo de los datos, sobre un rectángulo, alineado horizontal o verticalmente. [8]

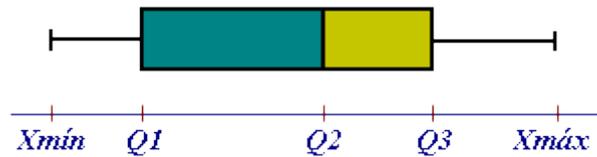


Ilustración 2.2 Caja y Bigotes.

2.2.2 Análisis Bivariante

Las técnicas estadísticas bivariantes permiten el análisis conjunto de dos características de los individuos de una población con el propósito de detectar posibles relaciones entre ellas. La naturaleza (nominal, ordinal o numérica) de las características objeto de estudio determinará las herramientas más adecuadas para su análisis. [9]

2.2.2.1 Gráfico X-Y

Los datos son representados por puntos delimitados por los ejes X e Y, dando la posibilidad de unir los puntos, donde nos indica si existe algún tipo de relación entre estas variables. [10]

2.2.2.2 Multicolinealidad

El proceso o término de multicolinealidad en Econometría es una situación en la que se presenta una fuerte correlación entre variables explicativas del modelo. La correlación ha de ser fuerte, ya que siempre existirá correlación entre dos variables explicativas en un modelo, es decir, la no correlación de dos variables es un proceso idílico, que sólo se podría encontrar en condiciones de laboratorio. [11]

Estas son las siguientes pruebas para identificar si existe multicolinealidad o no:

➤ Matriz de correlación.

La matriz de correlación muestra los valores de correlación de Pearson, que miden el grado de relación lineal entre cada par de elementos o variables. Los valores de correlación se pueden ubicar entre -1 y +1.

Las variables con valores de correlación mayores que 0,7 se consideran altamente correlacionadas por lo tanto indicará problemas de multicolinealidad. [12]

➤ Inversa de la matriz de correlación.

Por el contrario que en la matriz de correlación, solo se tendrá en cuenta la diagonal principal de la matriz, donde si algún valor es superior a 10 estará indicando que tiene problemas de multicolinealidad.

➤ Índice de acondicionamiento.

La ultima prueba que se debería hacer es el índice de acondicionamiento donde se obtendrá este valor mediante la siguiente formula:

$$I.C = \sqrt{\frac{Autovalor_{MAX}}{Autovalor_{MIN}}}$$

Si $0 \leq I.C. < 10 \rightarrow$ No existe problema de multicolinealidad.

Si $10 \leq I.C. < 20 \rightarrow$ Existe multicolinealidad moderada.

Si $20 \leq I.C. \rightarrow$ Existe multicolinealidad moderada.

Una vez realizado esto, para solucionar esta correlación de estas variables explicativas será eliminar una variable, que con ayuda de la matriz de correlación inversa nos indicará cual será la candidata para eliminar, todos los valores que superen a 10 serán las candidatas, pero se escogerá el valor mayor. Si los valores de varios candidatos son muy similares se realizaría dos modelos diferentes donde se realizaría una regresión múltiple eliminando las variables candidatas, donde el modelo con mayor R^2 sería con el que se deba continuar y la variable problemática sería la candidata para eliminar.

2.2.3 Anàlisis multivariante

El Anàlisis Multivariante es el conjunto de métodos estadísticos cuya finalidad es analizar simultáneamente conjuntos de datos multivariantes en el sentido de que hay varias variables medidas para cada individuo u objeto estudiado.

Su razón de ser radica en un mejor entendimiento del fenómeno objeto de estudio obteniendo información que los métodos estadísticos univariantes y bivariantes son incapaces de conseguir. [13]

2.2.3.1 Significatividad

El nivel de significación es comúnmente representado por el símbolo griego α (alfa). Son comunes los niveles de significación del 0,05, 0,01 y 0,001. Si un contraste de hipótesis proporciona un valor p inferior a α , la hipótesis nula es rechazada, siendo tal resultado denominado 'estadísticamente significativo'. Cuanto menor sea el nivel de significación, más fuerte será la evidencia de que un hecho no se debe a una mera coincidencia (al azar). [14]

Se comprobará la significatividad tanto del modelo como de los otros parámetros.

- Contraste modelo:

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

$$H_1: \text{Algún } \beta_i \neq 0$$

- Contraste de uno de los parámetros:

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0$$

2.2.3.2 Normalidad de residuos

La hipótesis de normalidad afirma que los errores del modelo siguen una distribución normal. Esta hipótesis se contrasta a partir de los residuos estandarizados.

Uno de los gráficos con los que se puede observar la normalidad de los datos es el histograma.

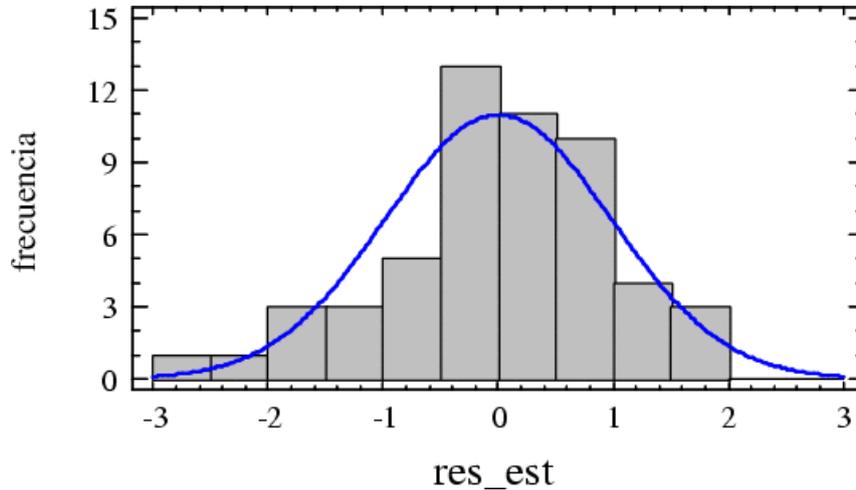


Ilustración 2.3 Histograma.

Se diría que existe normalidad de residuos cuando los datos del histograma formen una campana de Gauss como la de la Ilustración 2.3.

Otro de los gráficos que se puede utilizar para identificar la normalidad de los residuos es con el gráfico probabilístico.

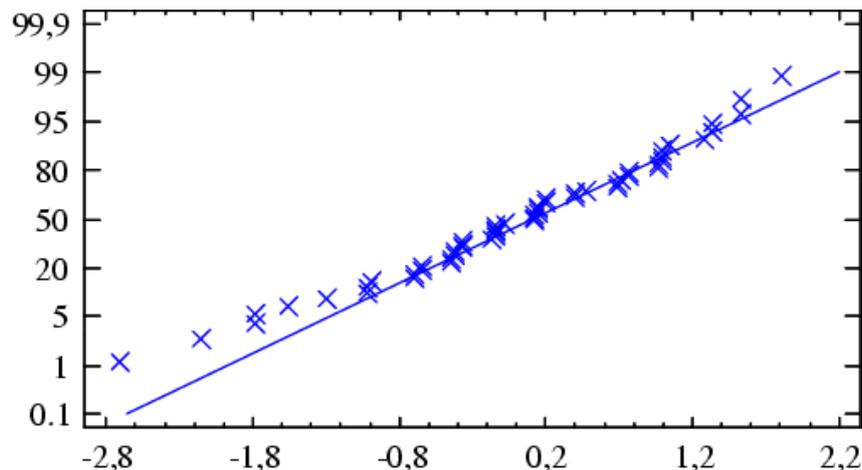


Ilustración 2.4 Gráfico de probabilidad normal.

Analizando la gráfica se puede identificar si existe normalidad de residuos o no, observando la dispersión de las cruces en el gráfico se puede decir que tiene normalidad de residuos ya que los datos están alrededor de la línea de normalidad, en

caso contrario estarían los valores muy dispersos de la línea de normalidad y se concluiría que hay problemas de normalidad de residuos.

2.2.3.3 Heterocedasticidad

En estadística se denomina heterocedasticidad cuando los errores no son constantes a lo largo de toda la muestra. El término es contrario a homocedasticidad.

En otras palabras, en los modelos de regresión lineales se dice que hay elasticidad cuando la varianza de los errores no es igual en todas las observaciones realizadas. En caso de que haya heterocedasticidad no se cumple uno de los requisitos básicos de las hipótesis de los modelos lineales. [15]

Para detectar problemas de heterocedasticidad en un modelo existen dos métodos, por contraste de hipótesis y por análisis gráfico.

2.2.3.4 Autocorrelación

En el análisis multivariante también se comprobará si existe problema de autocorrelación utilizando el contraste de autocorrelación de 1º orden y los gráficos de FAS y FAP.

Test Durbin-Watson

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

Si se identifica que el P-valor es mayor que $\alpha = 0,05$ no existiría autocorrelación ya que la hipótesis nula se aceptaría.

Otro de los métodos para identificar la existencia de autocorrelación es el análisis de los gráficos FAS y FAP.

Autocorrelaciones Estimadas para RESIDUOS

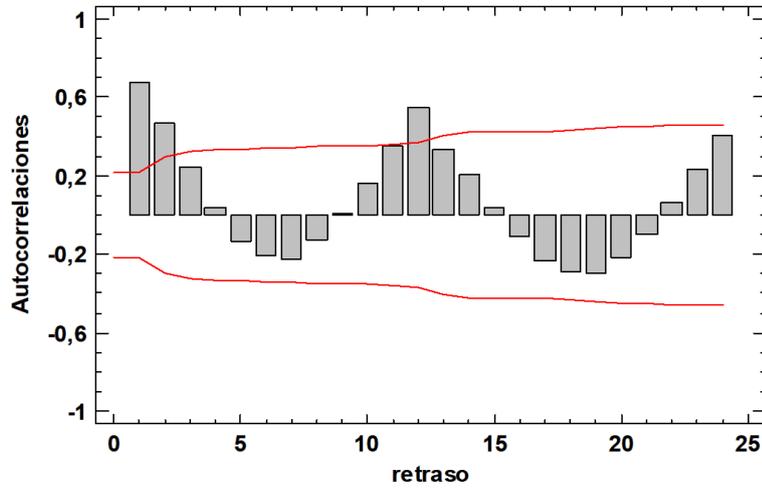


Ilustración 2.5 Función de autocorrelación simple.

En el gráfico de función de autocorrelación simple mide la relación lineal que existe entre el residuo en el instante de tiempo y el residuo en el instante. Para que no haya problema de autocorrelación ningún factor de autocorrelación debe sobrepasar las líneas que delimitan los límites.

Autocorrelaciones Parciales Estimadas para RESIDUOS

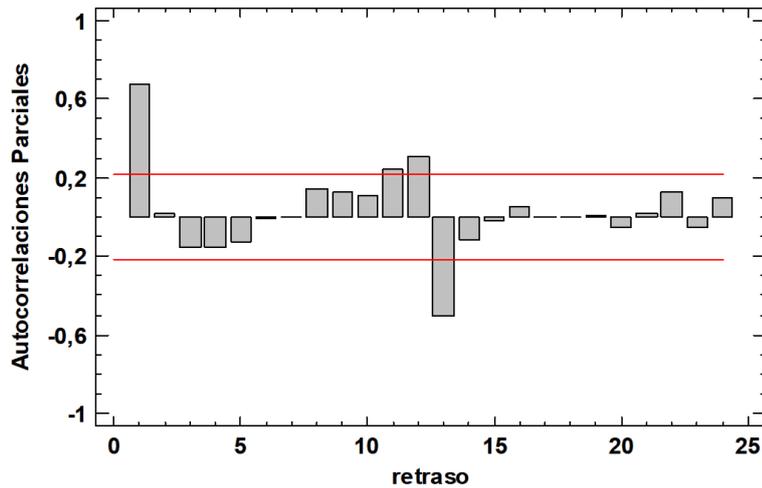


Ilustración 2.6 Función de autocorrelación parciales.

Y en el gráfico de función de autocorrelación parciales mide la relación directa que existe entre el residuo tiempo y el residuo instante. Para que no haya problema de autocorrelación ningún factor de autocorrelación debe sobrepasar las líneas que delimitan los límites.

2.2.3.5 Puntos influyentes.

Los valores atípicos o anómalos son aquellas que su comportamiento es distinto del resto de valores.

Los puntos influyentes son aquellos valores que tienen un comportamiento diferente al resto e influyen de tal modo que hace afectan al resultado.

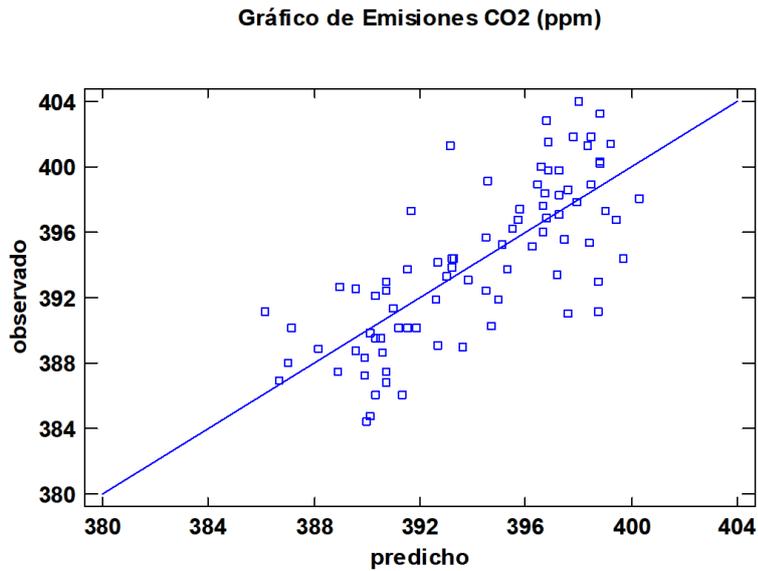


Ilustración 2.7 Puntos influyente.



3. *CALCULOS, RESULTADOS Y EXPLOTACIÓN*



A continuación, se va a indicar todas las variables que se van a estudiar en el modelo teórico, ya que la tabla a sido imposible ponerla ya que debido a que el estudio de las variables es mensual durante un septenio a sido imposible plasmarla entera. Los datos de las variables han sido extraídos en la base de datos del instituto nacional de estadística a nivel estatal.

Variables:

- Emisiones CO₂ en España.
- Consumo de carbón en España.
- Consumo de gases derivados del carbón en España.
- Consumo de productos petrolíferos en España.
- Consumo de gas natural y derivados en España.
- Consumo de electricidad en España.
- Consumo de energía renovable en España.

3.1 ANALISIS UNIVARIANTE

El primero de los análisis a realizar va a ser el Análisis Univariante. El análisis univariante se utiliza para estudiar el comportamiento de las variables de forma individual. Por lo que se va a calcular el resumen de estadísticos para cada variable con sus respectivas gráficas de caja y bigotes, histograma, y serie de tiempo desde enero del 2009, hasta diciembre del 2015 en España.

Análisis variable emisión CO₂.

Esta variable muestra la cantidad en ppm de CO₂ que se emite en España a la atmosfera.

Tabla 3.1 Resumen estadísticos para emisiones de CO₂.

Recuento	84
Promedio	394,12
Mediana	393,795
Varianza	23,711
Desviación Estándar	4,86939
Mínimo	384,43
Máximo	403,96
Rango	19,53
Cuartil Inferior	390,17
Cuartil Superior	397,93
Sesgo	0,0472308
Curtosis	-0,902564

Gráfico de Caja y Bigotes

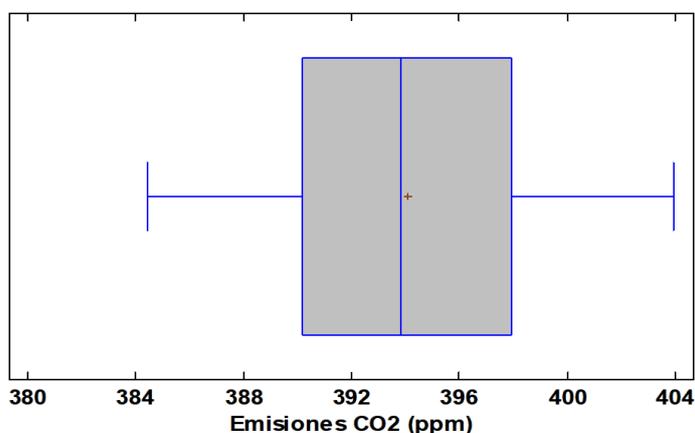


Ilustración 3.1 Gráfico de caja y bigotes para emisiones de CO₂.

En el gráfico de Caja y Bigotes no se observa a primera vista ningún valor atípico donde pueda afectar fuertemente al resultado.

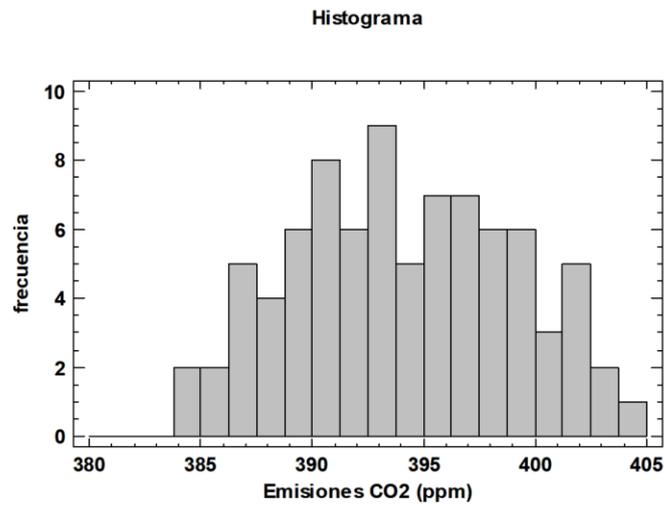


Ilustración 3.2 Histograma para emisiones de CO₂.

En este histograma se observa que los datos se ajustan a una distribución normal ya que se observa una distribución formando una curva conocida como la campana de Gauss.

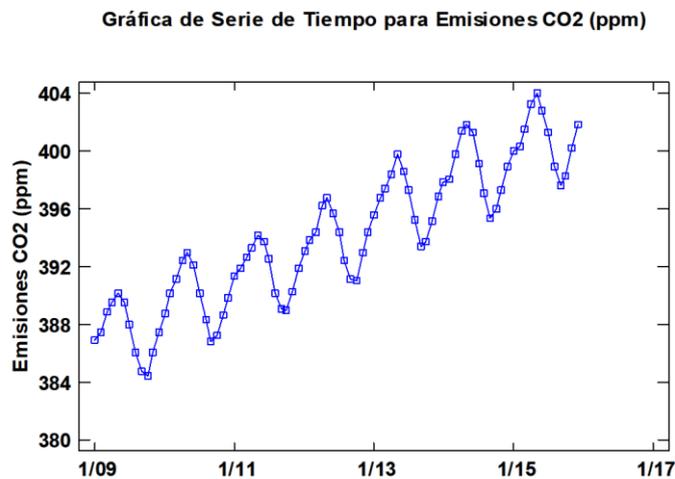


Ilustración 3.3 Gráfica de serie de tiempo para emisiones de CO₂.

Como se puede observar en el gráfico de serie de tiempo para emisiones de CO₂, el aumento de emisión ha sido progresivo siendo siempre el pico más alto en los meses de mayo y los picos más bajos en los meses de septiembre y octubre. También se puede identificar en esta grafica una tendencia positiva con variaciones estacionales.

Análisis variable consumo carbón.

Esta variable indica la cantidad en miles de Tep que se consume de carbón en todo el territorio español.

Tabla 3.2 Resumen Estadístico para consumo de carbón.

Recuento	84
Promedio	110,917
Mediana	107,5
Varianza	360,391
Desviación Estándar	18,984
Coefficiente de Variación	17,1155%
Mínimo	79,0
Máximo	175,0
Rango	96,0
Cuartil Inferior	97,0
Cuartil Superior	124,0
Sesgo	0,785082
Curtosis	0,922853

Gráfico de Caja y Bigotes

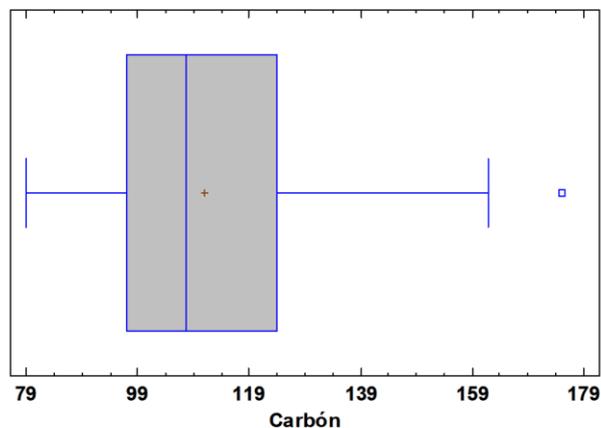
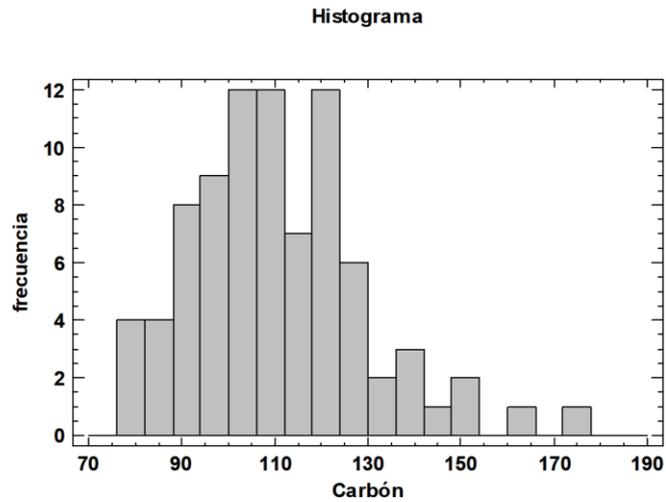


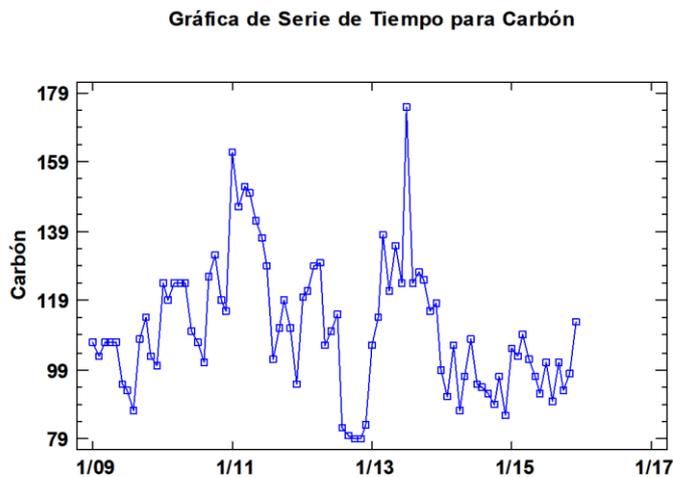
Ilustración 3.4 Gráfica de caja y bigotes para consumo de carbón.

En el gráfico de Cajas y Bigotes de la variable consumo de carbón se identifica un punto atípico el cual podría repercutir fuertemente en el resultado. El dato atípico lo encontramos en Julio del 2013.



Il·lustració 3.5 Histograma para consumo de carbón.

En el histograma de la variable consumo del carbón se puede decir que se observa una distribución normal de los datos, ya que se puede observar una leve curva conocida como la campana de Gauss la cual indica normalidad en los datos.



Il·lustració 3.6 Gráficas de serie de tiempo para consumo de carbón.

En este gráfico se observa que el consumo del carbón en los últimos años ha sido muy variante, donde en los últimos 2, 3 años ha disminuido su consumo en comparación de los años anteriores. Al igual que se puede observar en la (Ilustración 3.4) se indica claramente que en julio del 2013 fue el mes de máximo consumo durante el sexenio.

Análisis variable consumo gases derivados del carbón.

Esta variable muestra la cantidad en miles de Tep que se consume de gases derivados del carbón en todo el territorio español.

Tabla 3.3 Resumen Estadístico para Gases derivados del carbón.

Recuento	84
Promedio	20,881
Mediana	20,0
Varianza	25,311
Desviación Estándar	5,031
Mínimo	8,0
Máximo	32,0
Rango	24,0
Cuartil Inferior	18,0
Cuartil Superior	25,0
Sesgo	0,0851397
Curtosis	-0,307109

Gráfico de Caja y Bigotes

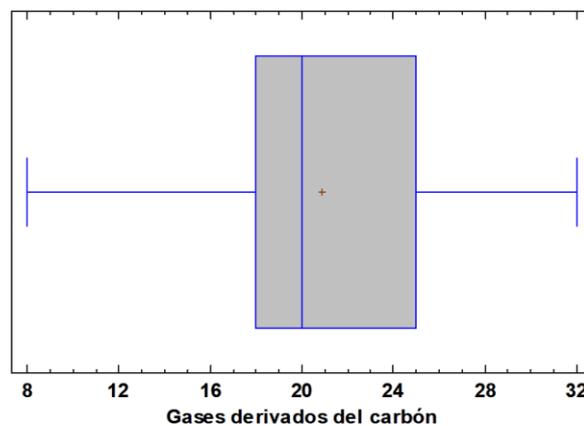


Ilustración 3.7 Gráfico de caja y bigote para Gases derivados del carbón.

En este gráfico no se identifica ningún valor alejado de la distribución de los otros valores, por lo cual, no afectarán gravemente al resultado, pero lo que sí que se puede observar es una asimetría hacia la izquierda.

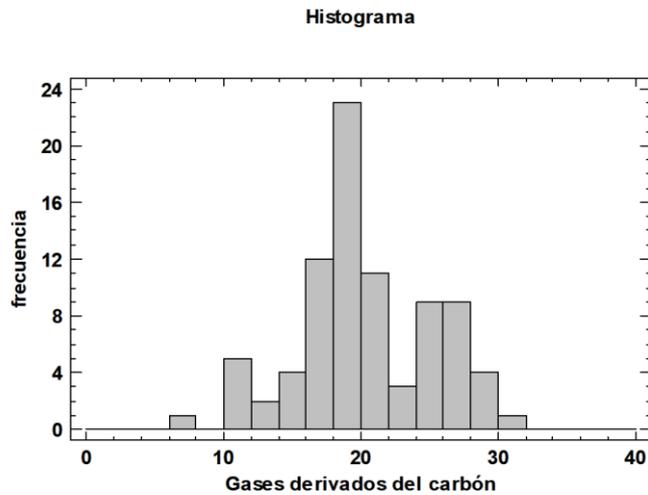


Ilustración 3.8 Histograma para Gases derivados del carbón.

En este histograma se puede decir que los datos siguen una distribución normal ya que los datos forman una campana de Gauss.

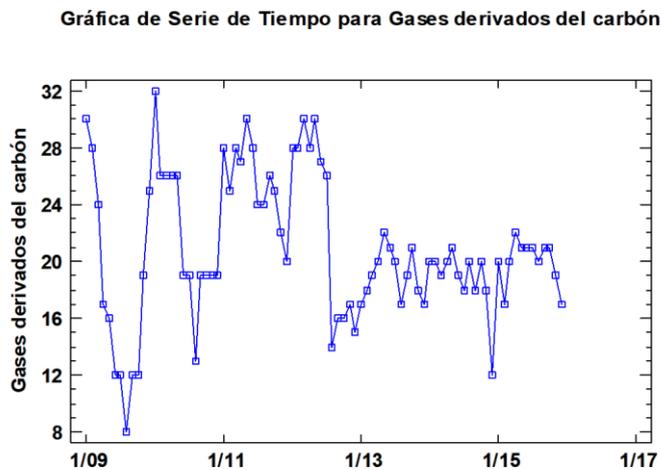


Ilustración 3.9 Gráfico de serie de tiempo para Gases derivados del carbón.

Se puede observar en este gráfico de la variable del consumo de gases derivados del carbón, su disminución a partir de la mitad del año 2012 donde se puede observar una disminución en su consumo, siendo los meses de invierno los de mayor consumo y los de verano los de menor, presentando así variaciones estacionales.

Análisis variable consumo productos petrolíferos.

Esta Variables mide en miles de Tep el consumo de productos petrolíferos en todo el territorio español.

Tabla 3.4 Resumen estadístico para productos petrolíferos.

Recuento	84
Promedio	3946,6
Mediana	3856,5
Varianza	173994,
Desviación Estándar	417,126
Mínimo	3268,0
Máximo	4908,0
Rango	1640,0
Cuartil Inferior	3600,5
Cuartil Superior	4316,0
Rango Intercuartílico	715,5
Sesgo	0,369118
Curtosis	-1,03953

Gráfico de Caja y Bigotes

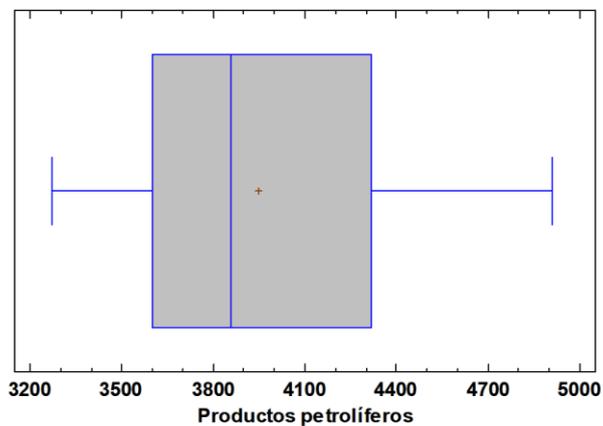


Ilustración 3.10 Gráfico de caja y bigotes para productos petrolíferos.

En el gráfico de Caja y Bigotes de esta variable no se identifica ningún valor atípico que pueda afectar al resultado, aunque lo que sí que muestra una asimetría hacia la derecha.

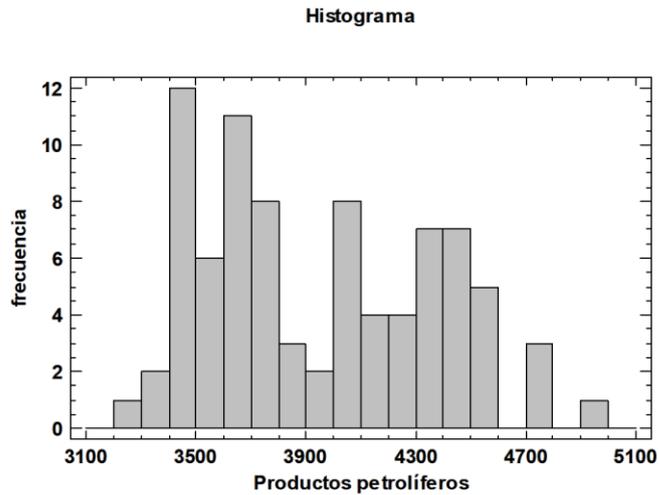


Ilustración 3.11 Histograma para productos petrolíferos.

En el histograma de la variable de consumo de productos petrolíferos también se identifica en principio que los datos se ajustan a una distribución normal formando una leve curva conocida como la campana de Gauss.

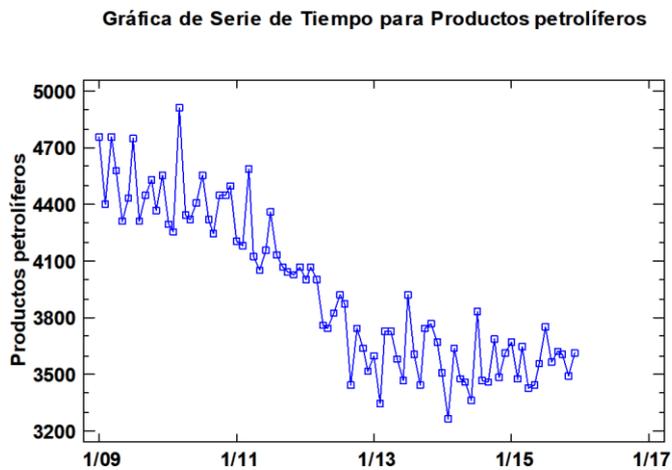


Ilustración 3.12 Gráfica serie de tiempo para productos petrolíferos.

Sobre esta gráfica se puede destacar que el consumo de productos petrolíferos ha ido disminuyendo paulatinamente durante el paso de los años. También se puede identificar en esta gráfica que tiene una tendencia negativa.

Análisis variable consumo gas natural y manufacturado.

Esta variable indica en miles de Tep del consumo de gas natural y manufacturado en todo el territorio español.

Tabla 3.5 Resumen Estadístico para gas natural y manufacturado.

Recuento	84
Promedio	1216,11
Mediana	1118,0
Varianza	84504,1
Desviación Estándar	290,696
Mínimo	775,0
Máximo	1793,0
Rango	1018,0
Cuartil Inferior	972,5
Cuartil Superior	1478,5
Sesgo	0,400522
Curtosis	-1,17843

Gráfico de Caja y Bigotes

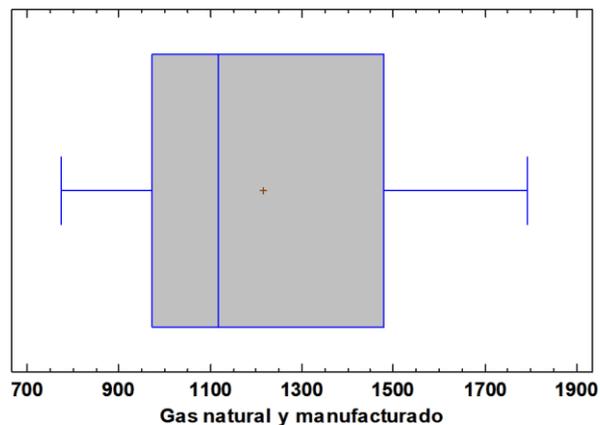


Ilustración 3.13 Gráfico de caja y bigotes para gas natural y manufacturado.

En el gráfico de Cajas y bigotes de esta variable no se detecta valores atípicos, pero si que presenta una gran asimetría hacia la derecha.

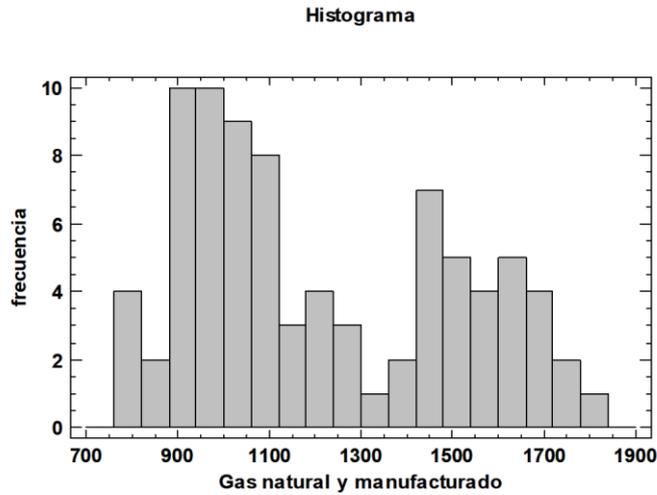


Ilustración 3.14 Histograma para gas natural y manufacturado.

En este histograma de la variable del consumo de gas natural y manufacturado se observan dos campanas de Gauss por lo que en principio indica normalidad en su distribución.

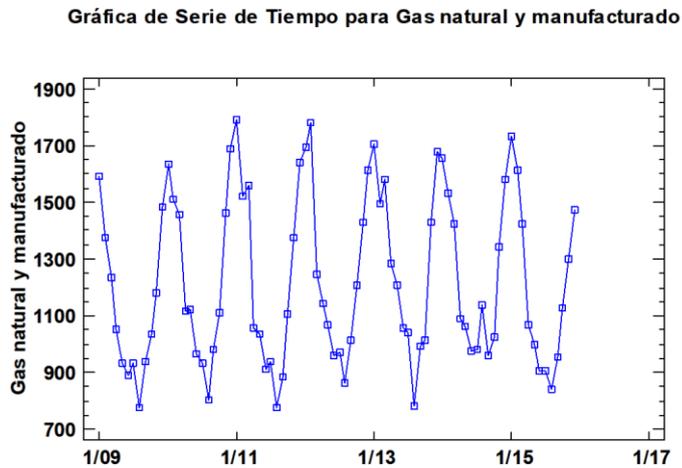


Ilustración 3.15 Gráfica de serie de tiempo para gas natural y manufacturado.

En este gráfico se puede observar que el consumo de gas natural y manufacturado ha sido similar durante el transcurso de los años, su consumo se ha visto más destacado en los meses de invierno siendo en la mayoría de los casos en enero como el mes donde mayor consumo se ha registrado, y en los meses de verano el consumo más bajo, presentando así variaciones estacionales.

Análisis variable consumo electricidad

Esta variable muestra en miles de Tep del consumo eléctrico en todo el territorio español.

Tabla 3.6 Resumen estadístico para electricidad.

Recuento	84
Promedio	1725,35
Mediana	1708,0
Varianza	15379,0
Desviación Estándar	124,012
Coefficiente de Variación	7,18767%
Mínimo	1505,0
Máximo	2035,0
Rango	530,0
Cuartil Inferior	1635,0
Cuartil Superior	1790,0
Sesgo	0,608613
Curtosis	-0,0515538

Gráfico de Caja y Bigotes

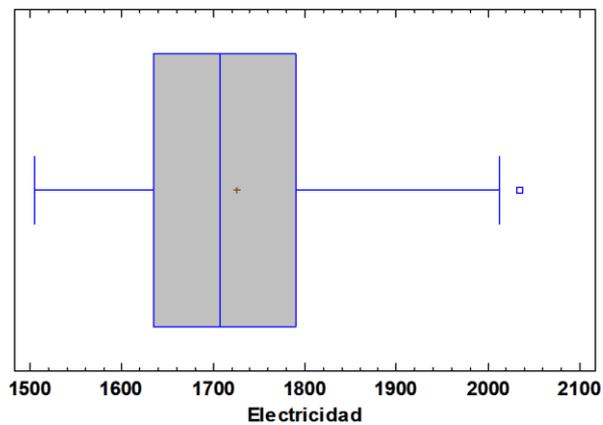
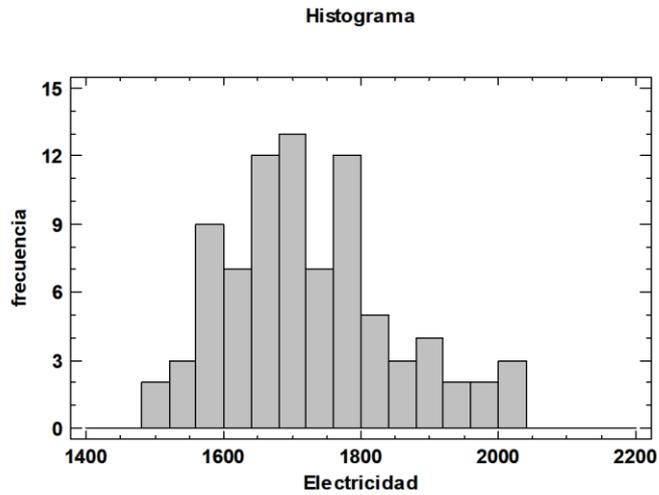


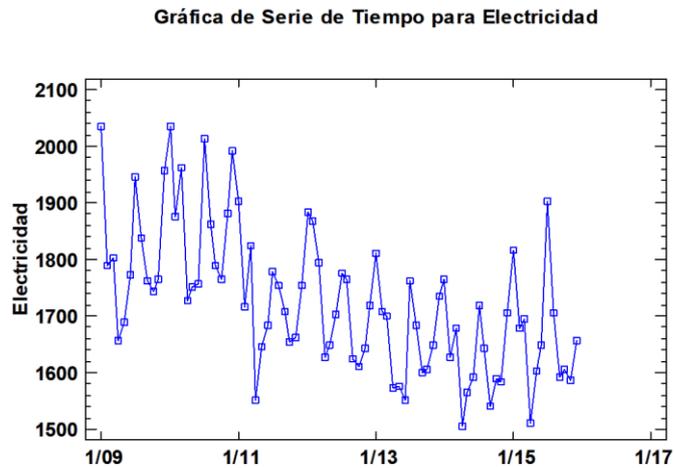
Ilustración 3.16 Gráfico de caja y bigote para electricidad.

En la variable consumo de electricidad observando la gráfica de Caja y bigotes, se identifica un punto atípico, el cual puede afectar en el resultado, sin embargo, no muestra una gran asimetría. Este dato anómalo se puede localizar en enero del 2009.



Il·lustració 3.18 Histograma para electricidad.

En el histograma para el consumo de electricidad se puede decir que los datos se ajustan a una distribución normal en principio, ya que los datos forman una leve curva conocida como la campana de Gauss indicando normalidad en ella.



Il·lustració 3.19 Gràfica de serie de tiempo para electricidad.

En el gráfico se observa una disminución tras el paso de los años en el consumo eléctrico siendo el mes de enero y julio los de mayor consumo habitualmente, y abril el mes de menor consumo eléctrico. Se puede decir que este gráfico presenta una tendencia negativa con variaciones estacionales.

Análisis variable consumo energías renovables.

Esta variable mide en miles de Tep el consumo de energías renovables en todo el territorio español.

Tabla 3.7 Resumen estadístico para energías renovables.

Recuento	84
Promedio	457,44
Mediana	447,5
Varianza	4319,04
Desviación Estándar	65,7194
Mínimo	321,0
Máximo	633,0
Rango	312,0
Cuartil Inferior	403,5
Cuartil Superior	492,0
Sesgo	0,580954
Curtosis	0,127074

Gráfico de Caja y Bigotes

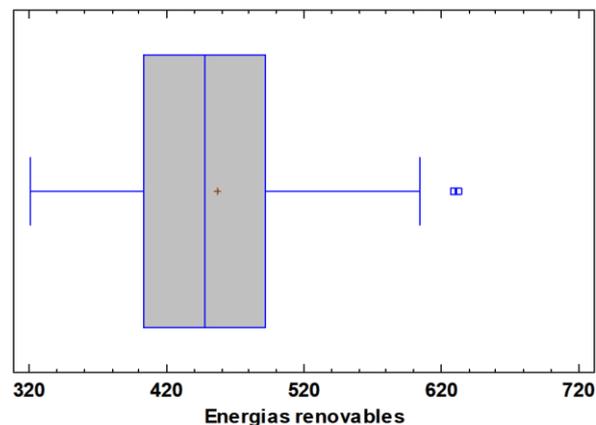


Ilustración 3.20 Gráfico de caja y bigotes para energías renovables.

En este gráfico se puede identificar varios puntos atípicos los cuales pueden afectar bruscamente al resultado, por otro lado, no muestra una gran asimetría. Estos dos puntos atípicos se localizan en enero y diciembre del 2012 respectivamente.

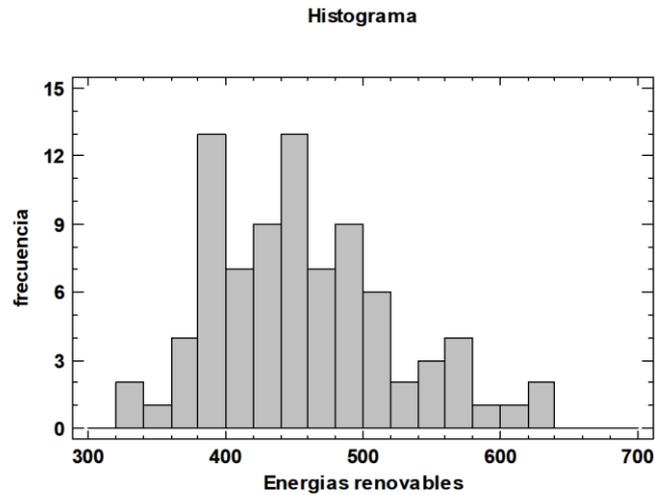


Ilustración 3.21 Histograma para energías renovables.

En este histograma de la variable consumo de energías renovables se puede decir que en principio los datos se ajustan a una distribución normal ya que se identifica una campana de Gauss lo cual indica en principio una distribución normal de los valores.

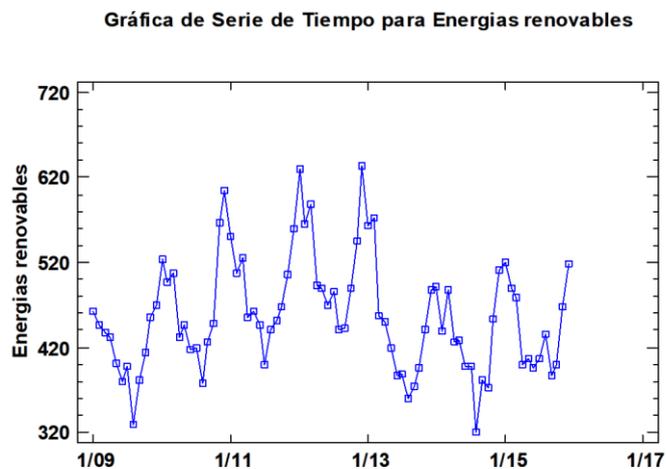


Ilustración 3.22 Gráfica de serie de tiempo para energías renovables.

En este gráfico se identifica una involución del consumo de energías renovables tras el paso de los años, siendo el último año el 2012 el último año de gran consumo de energías renovables habiendo mayor consumo en los meses de invierno, y menor consumo en verano, es decir, que este gráfico indica variaciones estacionales.

3.2 BIVARIANTE

Las técnicas estadísticas bivariantes permiten el análisis conjunto de dos características de los individuos de una población con el propósito de detectar posibles relaciones entre ellas.

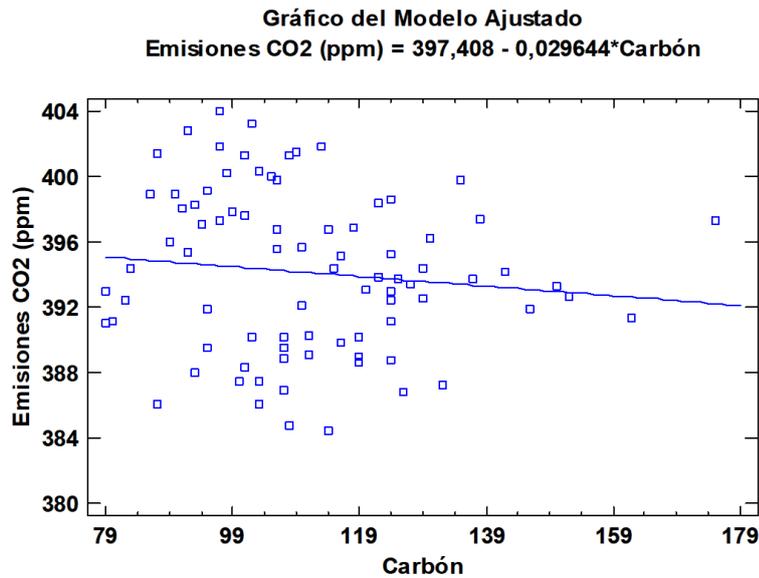


Ilustración 3.23 Gráfico del modelo ajustado.

Coefficiente de Correlación = -0,115571

Con este coeficiente de correlación y el gráfico del modelo ajustado se puede decir que existe una relación baja entre las emisiones de CO₂ y el consumo de carbón con una tendencia negativa.

Gráfico del Modelo Ajustado
Emisiones CO2 (ppm) = 393,958 + 0,00771857*Gases derivados del carbón

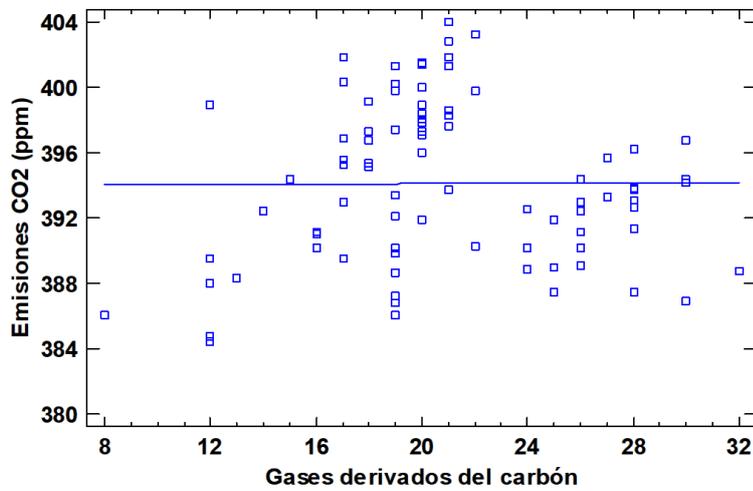


Ilustración 3.24 Gráfico del modelo ajustado.

Coeficiente de Correlación = 0,00797473

Con este coeficiente de correlación y el gráfico del modelo ajustado se observa que hay una relación muy baja entre la variable de emisiones de CO₂ y la variable de consumo de gases derivados del carbón.

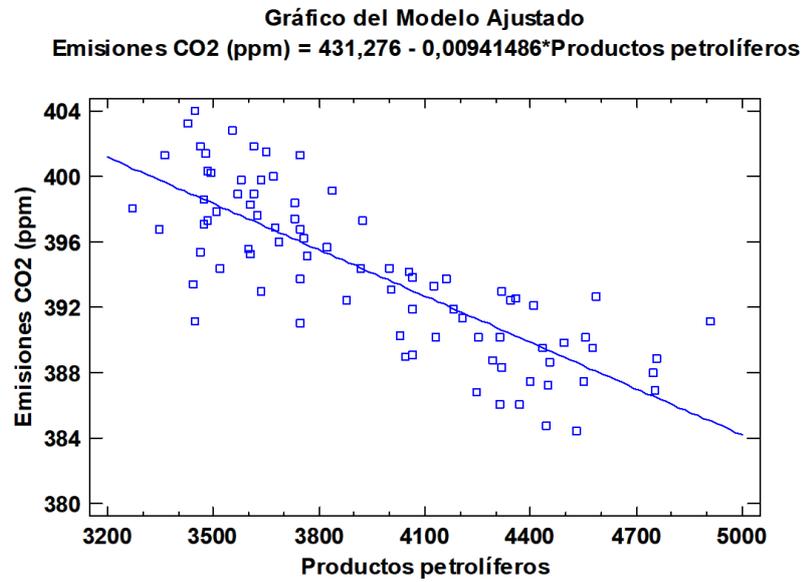


Ilustración 3.25 Gráfico del modelo ajustado.

Coefficiente de Correlación = -0,806503

Con este coeficiente de correlación y el gráfico del modelo ajustado se puede observar que las emisiones de CO₂ y el consumo de productos petrolíferos tienen una relación elevada con tendencia negativa.

Gráfico del Modelo Ajustado
Emissiones CO2 (ppm) = 392,754 + 0,00112271*Gas natural y manufacturado

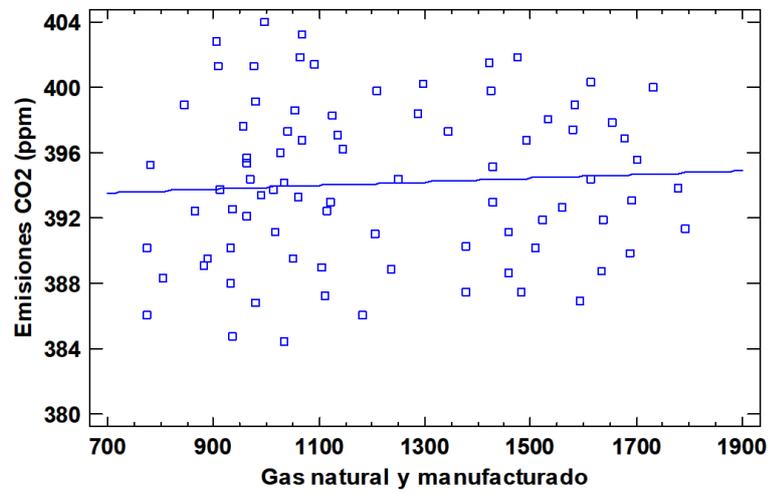


Ilustración 3.26 Gráfico del modelo ajustado.

Coeficiente de Correlación = **0,0670242**

Con este coeficiente de correlación y el gráfico del modelo ajustado se puede decir que existe una relación directa baja entre las emisiones de CO₂ y el consumo de gas natural y manufacturados.

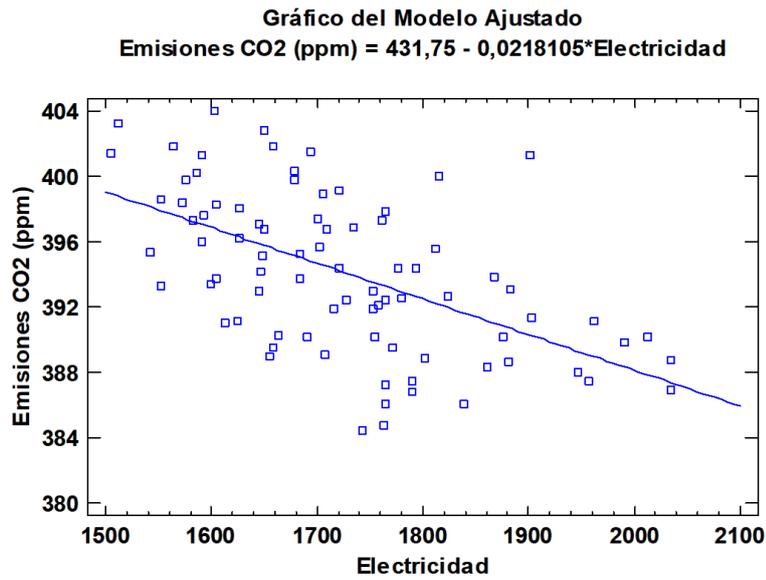


Ilustración 3.27 Gráfico del modelo ajustado.

Coeficiente de Correlación = -0,555462

Con este coeficiente de correlación y el gráfico del modelo ajustado se puede decir que existe una relación indirecta moderada entre las emisiones de CO₂ y el consumo de electricidad cuya tendencia es negativa.

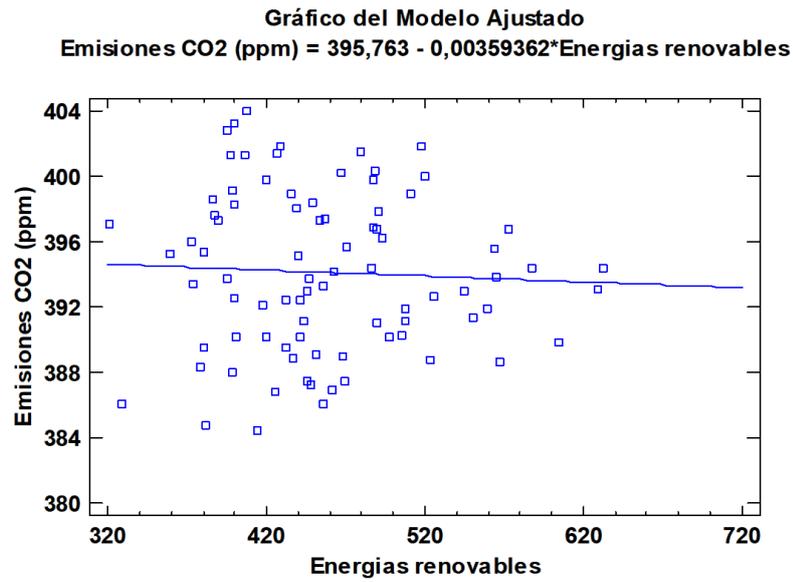


Ilustración 3.28 Gráfico del modelo ajustado.

Coeficiente de Correlación = -0,048501

Con este coeficiente de correlación y el gráfico del modelo ajustado se puede decir que existe una relación indirecta bastante baja entre las emisiones de CO₂ y el consumo de energías renovables.

3.2.1 MULTICOLINEALIDAD

La multicolinealidad es una situación en la que se presenta una fuerte correlación entre variables explicativas del modelo. La correlación ha de ser fuerte, ya que siempre existirá correlación entre dos variables explicativas en un modelo.

1º Análisis matriz de correlación.

Tabla 3.8 Matriz de correlación.

	Carbón	Gases derivados del carbón	Productos petrolíferos	Gas natural y manufacturado	Electricidad	Energías renovables
Carbón	1,00	0,50	0,29	0,15	0,15	0,14
Gases derivados del carbón	0,50	1,00	0,17	0,21	0,16	0,32
Productos petrolíferos	0,29	0,17	1,00	-0,03	0,70	0,06
Gas natural y manufacturado	0,15	0,21	-0,03	1,00	0,31	0,77
Electricidad	0,15	0,16	0,70	0,31	1,00	0,34
Energías renovables	0,14	0,32	0,06	0,77	0,34	1,00

Analizando esta tabla se observan una relación elevada entre la variable consumo de electricidad y la variable consumo de productos petrolíferos, y también entre la variable consumo de energías renovables y consumo de gas natural y manufacturado, observando problemas de multicolinealidad.

2º Matriz inversa de correlación.

Tabla 3.9 Matriz inversa de correlación.

Carbón	1,48	-0,69	-0,53	-0,32	0,29	0,18
Gases derivados del carbón	-0,69	1,48	-0,03	0,18	0,01	-0,51
Productos petrolíferos	-0,53	-0,03	2,54	0,71	-1,94	0,05
Gas natural y manufacturado	-0,32	0,18	0,71	2,69	-0,66	-1,89
Electricidad	0,29	0,01	-1,94	-0,66	2,62	-0,31
Energías renovables	0,18	-0,51	0,05	-1,89	-0,31	2,69

Analizando esta tabla se puede decir que en los coeficientes de la línea diagonal de esta no supera el valor 10, por lo tanto, se puede decir que en principio no se observa problemas de multicolinealidad.

3º Índice de acondicionamiento.

Tabla 3.10 Índice de acondicionamiento.

IC: 3,20

Eigenvalor
2,43
1,46
1,19
0,49
0,24

Observando el resultado del índice de acondicionamiento se puede decir que no se observa problemas de multicolinealidad ya que no es superior al valor 10 relacionando.

Solución multicolinealidad

1ºAplicar componentes principales.

Tabla 3.11 Análisis de componentes principales.

Componente Número	Eigenvalor	Porcentaje de Varianza	Porcentaje Acumulado
1	2,43214	40,536	40,536
2	1,46332	24,389	64,924
3	1,18577	19,763	84,687
4	0,493217	8,220	92,908
5	0,237145	3,952	96,860
6	0,188401	3,140	100,000

Modelo teórico propuesto para eliminar la multicolinealidad.

0,343733*Carbón + 0,380739*Gases derivados del carbón + 0,349602*
*Productos petrolíferos + 0,432728*Gas natural y manufacturado +
0,453656*Electricidad + +0,470752*Energías renovables

Tabla 3.12 *Tabla de pesos de los componentes*

	<i>Componente 1</i>	<i>Componente 2</i>	<i>Componente 3</i>
Carbón	0,343733	-0,207899	-0,597165
Gases derivados del carbón	0,380739	-0,0207376	-0,581403
Productos petrolíferos	0,349602	-0,608671	0,23655
Gas natural y manufacturado	0,432728	0,514299	0,170965
Electricidad	0,453656	-0,327061	0,447024
Energías renovables	0,470752	0,463028	0,142651

Como se puede observar en esta tabla, de 6 variables pasa a ser de 3 componentes principales que explican el 84'687% de los datos.

2º Método factorial

Tabla 3.13 *Coeficiente del factor.*

	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>
Carbón	0,0337194	0,137858	0,74795
Gases derivados del carbón	0,197601	0,070959	0,651568
Productos petrolíferos	-0,10366	0,942746	0,209467
Gas natural y manufacturado	0,846053	0,0523933	0,106274
Electricidad	0,292105	0,766135	0,0587709
Energías renovables	0,880006	0,107501	0,165203

Tabla 3.14 *Análisis de factores.*

<i>Factor Número</i>	<i>Eigenvalor</i>	<i>Porcentaje de Varianza</i>	<i>Porcentaje Acumulado</i>
1	2,14695	52,904	52,904
2	1,47749	36,407	89,311
3	0,43379	10,689	100,000
4	0,0	0,000	100,000
5	0,0	0,000	100,000
6	0,0	0,000	100,000

De 6 variables pasa a ser de 2 factores principales que explican el 89'311% de los datos.



Conclusión

Tras analizar las 2 posibles opciones para solucionar la multicolinealidad se ha considerado que la alternativa más beneficiosa es la elección del método factorial ya que con dos factores se obtiene un 89'311% de los datos, frente con un 84'687% de los datos aplicando los componentes principales.

3.3 Multivariante

3.3.1 Ajuste del modelo

Emisiones CO₂ (ppm) = 394,12 - 2,16969*FSCORE_1 + 0,810442*FSCORE_2

R-cuadrada = 57,7185 por ciento

Prueba del análisis

Tabla 3.15 Prueba del análisis.

Parámetro	Estimación	Error Estándar	Estadístico T	Valor-P
CONSTANTE	394,12	0,349709	1126,99	0,0000
FSCORE_1	-2,16969	0,206511	-10,5064	0,0000
FSCORE_2	0,810442	0,204428	3,96443	0,0002

Contraste β_0

$H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0$ $\alpha = 0'05$

P-valor = 0'000 < α ► Se rechaza la hipótesis nula y se concluye que la constante es significativa.

Contraste β_1

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$ $\alpha = 0'05$

P-valor = 0'0000 < α ► Se rechaza la hipótesis nula y se concluye que el factor 1 es significativo.

Contraste β_2

$H_0: \beta_2 = 0$

$H_1: \beta_2 \neq 0$ $\alpha = 0'05$

P-valor = 0'0002 < α ► Se rechaza la hipótesis nula y se concluye que el factor 2 es significativo.

Tabla 3.16 Análisis de varianza.

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	1135,91	2	567,954	55,29	0,0000
Residuo	832,104	81	10,2729		
Total (Corr.)	1968,01	83			

Contraste modelo

$H_0: \beta_1 = \beta_2 = 0$

$H_1: \text{Algún } \beta_i \neq 0 \quad \alpha = 0,05$

P-valor = 0,000 < α ► Se rechaza la hipótesis nula y se concluye que el modelo es significativo.

3.3.2 Normalidad de residuos

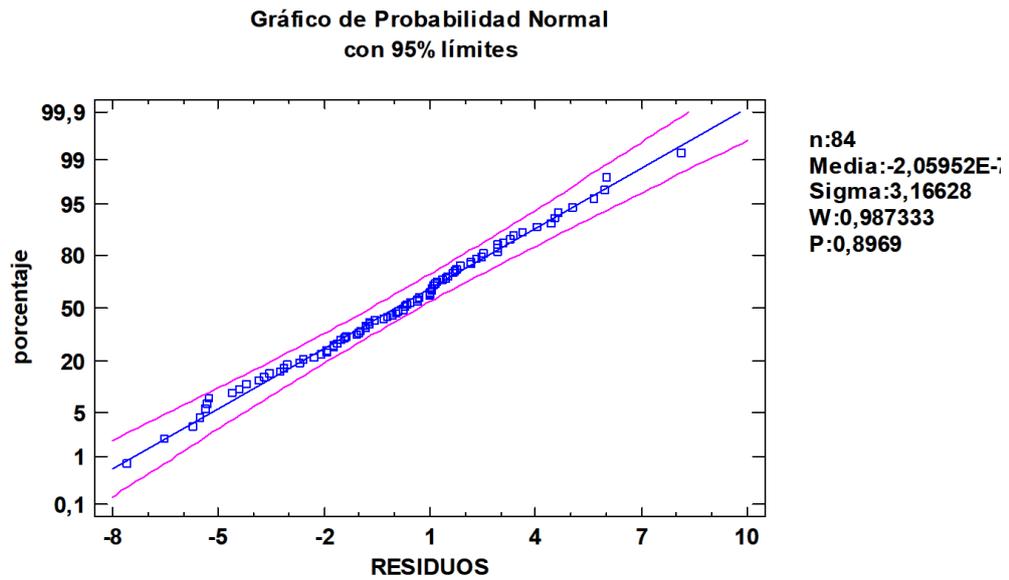


Ilustración 3.29 Gráfico de probabilidad normal.

Tras analizar este gráfico se puede observar que la distribución de la variable es normal ya que los puntos quedan cerca de la línea de normalidad.

3.3.3 Histograma de residuos

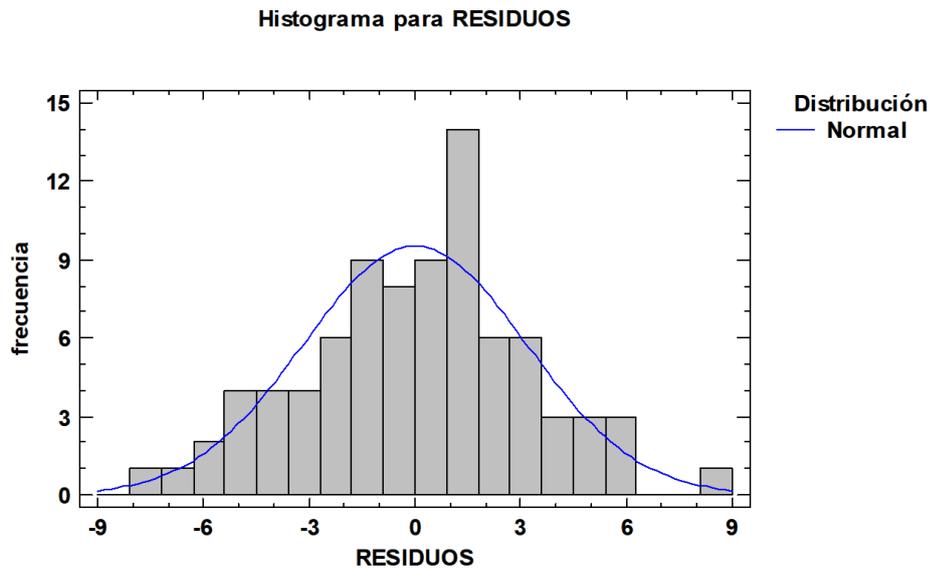


Ilustración 3.30 Histograma para residuos.

Tras observar el histograma se puede interpretar que los datos tienen una distribución normal ya que se distribuyen como una campana de Gauss significando normalidad en su distribución.

3.3.4 Contraste de normalidad.

Tabla 3.17 Pruebas de normalidad para residuos

Prueba	Estadístico	Valor-P
Chi-Cuadrado	10,7381	0,952734
Estadístico W de Shapiro-Wilk	0,987333	0,896863
Valor-Z para asimetría	0,201716	0,840134
Valor-Z para curtosis	-0,125227	0,900339

H_0 : Los residuos son normales.

H_1 : Los residuos no son normales. $\alpha = 0,05$

P-valor: $0,840134 > 0,05$

Se acepta la hipótesis nula, es decir, que los residuos son normales.

3.3.5 Heterocedasticidad

$$\text{RESIDUOS}^2 = 9,90601 - 0,795514 * \text{FSCORE}_1 - 1,02897 * \text{FSCORE}_2$$

Tabla 3.18 Heterocedasticidad.

Parámetro	Estimación	Error Estándar	Estadístico T	Valor-P
CONSTANTE	9,90601	1,43796	6,88895	0,0000
FSCORE_1	-0,795514	0,849145	-0,936841	0,3516
FSCORE_2	-1,02897	0,840583	-1,22412	0,2245

$H_0: \sigma^2 = \text{constante}$

$H_1: \sigma^2 \neq \text{constante} \quad \alpha = 0'05$

F1=P-valor: 0,3516 > α : Se acepta la hipótesis nula

F2=P-valor: 0,2245 > α : Se acepta la hipótesis nula

Por lo tanto, se concluye que no hay problema de heterocedasticidad.

3.3.6 Autocorrelación

Estadístico Durbin-Watson = 0,630218 (P=0,0000)

$H_0: \rho_1 = 0$

$H_1: \rho_1 \neq 0 \quad \alpha = 0'05$

P-valor = 0,0000 < α : Se rechaza la hipótesis nula

Se concluye que existe autocorrelación de 1º orden.

Función de autocorrelación simple

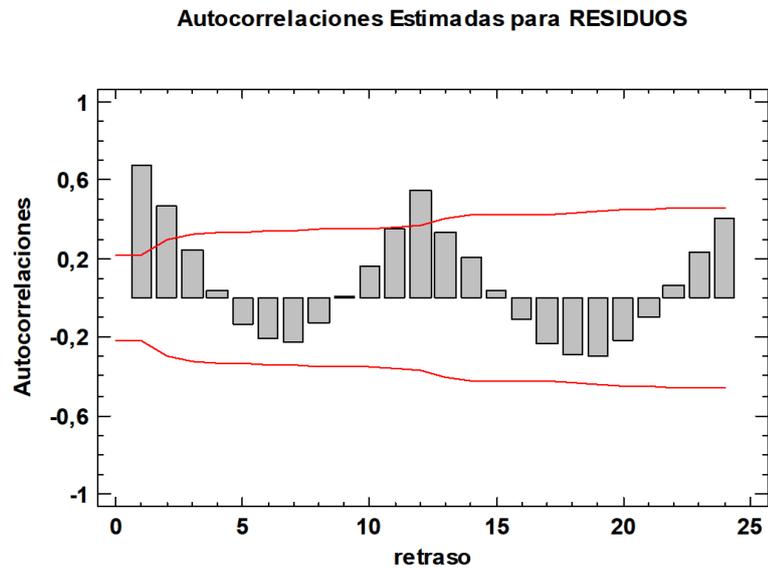


Ilustración 3.31 Autocorrelación estimada para residuos.

Tras analizar este gráfico se identifica problemas de autocorrelación de primer orden positiva, de segundo orden positiva y de doceava orden positiva, por lo tanto, se concluye que existe problemas de autocorrelación, además se observa que tiene un comportamiento similar al de una serie temporal.

3.3.7 Puntos influyentes

-A priori:

Influencia media de un solo punto = 0,0357143

Tabla 3.19 Puntos a priori.

Fila	Mes	Influencia	
8	2009M08	0,0794268	$>2*\bar{n}$
9	2009M09	0,0486437	$<2*\bar{n}$
15	2010M03	0,080156	$>2*\bar{n}$
45	2012M09	0,0369448	$<2*\bar{n}$
48	2012M12	0,0717849	$>2*\bar{n}$

Tras observar la tabla se puede identificar que los meses de agosto del 2009, marzo del 2010 y diciembre del 2012 afectan a la pendiente del modelo.

-A posteriori.

Influencia media de un solo punto = 0,0357143

Tabla 3.20 Puntos a posteriori.

Fila	Mes	DFITS		
8	2009M08	-0,403222	$>2*\sqrt{\bar{n}}$	influyente a posteriori
9	2009M09	-0,391339	$>2*\sqrt{\bar{n}}$	influyente a posteriori
15	2010M03	0,487915	$>2*\sqrt{\bar{n}}$	influyente a posteriori
45	2012M09	-0,488771	$>2*\sqrt{\bar{n}}$	influyente a posteriori
48	2012M12	-0,491226	$>2*\sqrt{\bar{n}}$	influyente a posteriori

Tras analizar la tabla se puede identificar que los meses de agosto y septiembre del 2009, el mes de marzo del 2010, y los meses de septiembre y diciembre del 2012 afectan a la constante del modelo.

3.3.8 Puntos anómalos

Tabla 3.21 Puntos anómalos.

Fila	Mes	Residuo Estudentizado	
45	2012M09	-2,5	Punto anómalo
46	2012M10	-2,11	Punto anómalo
79	2015M07	2,66	Punto anómalo

Como se observa en la tabla, el mes de septiembre y octubre del 2012, y julio del 2015 son puntos anómalos ya que son mayores a 2.

3.3.9 Solución autocorrelación

Tras probar los 5 métodos de solución de autocorrelación no han sido suficiente para encontrar su corrección, por lo que se ha optado por desordenar los datos en Excel utilizando la función aleatoria, y a continuación hemos recalculado y actualizado la autocorrelación

Estadístico Durbin-Watson = 1,8974 (P=**0,3264**)

$H_0: \rho_1 = 0$

$H_1: \rho_1 \neq 0 \quad \alpha = 0'05$

P-valor = 0,3265 < α : Se acepta la hipótesis nula

Se concluye que no existe autocorrelación de 1º orden.

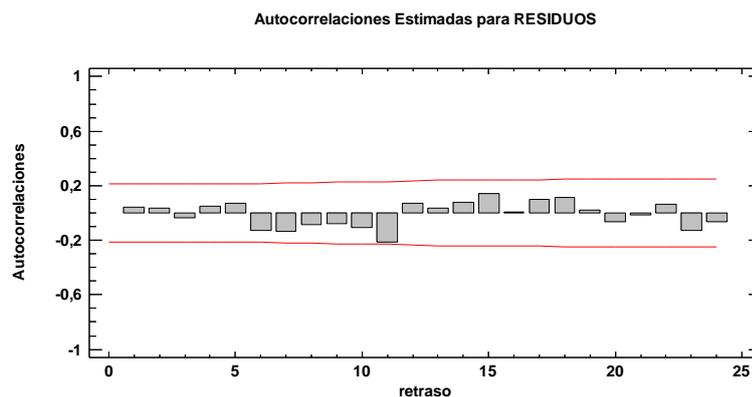


Ilustración 3.32 Autocorrelación estimada para residuos.

Tras analizar (Ilustración 3.32) se puede observar que el problema de autocorrelación a desaparecido.

Modelo final

$$\text{Emisiones CO}_2 \text{ (ppm)} = 394,12 - 2,16969 \cdot \text{FSCORE}_1 + 0,810442 \cdot \text{FSCORE}_2$$

Interpretación de cómo afecta las variables explicativas a la variable de emisiones.

▲ F1 → ▼ Emisiones CO₂.

▲ F2 → ▲ Emisiones CO₂.

▲ Carbón → ▲ F1 → ▼ Emisiones CO₂.

▲ Gases Carbón → ▲ F2 → ▲ Emisiones CO₂.

▲ Productos petrolíferos → ▲ F1 → ▼ Emisiones CO₂.

▲ Gas natural → ▲ F2 → ▲ Emisiones CO₂.

▲ Electricidad → ▲ F1 → ▼ Emisiones CO₂.

▲ Energía renovable → ▲ F2 → ▲ Emisiones CO₂.

Interpretando estos resultados se podría decir que para disminuir las emisiones de CO₂ debería disminuir el consumo de gases de carbono, gas natural y energías renovables, y por el contrario, aumentar el consumo de las fuentes de energía como el carbón, los productos petrolíferos y la electricidad, ya que el modelo indica que al utilizar estas fuentes disminuyen las emisiones de CO₂. No obstante, se identifica un comportamiento ilógico ya que se observa que a mayor consumo de energías renovables mayores emisiones de CO₂, como por el aumento del consumo del carbón un descenso de las emisiones de CO₂.

3.4 Predicción

Condición de las variables x.
La predicción puntal es 394,57 Tap.
Predicción por intervalo de confianza.

$$[388,154 \leq \text{Emisiones de CO}_2 \leq 400,987]_{95\%}$$

Tras la predicción se va a realizar el análisis de emisiones por series temporales.

Gráfica de Serie de Tiempo para Emisiones CO₂ (ppm)

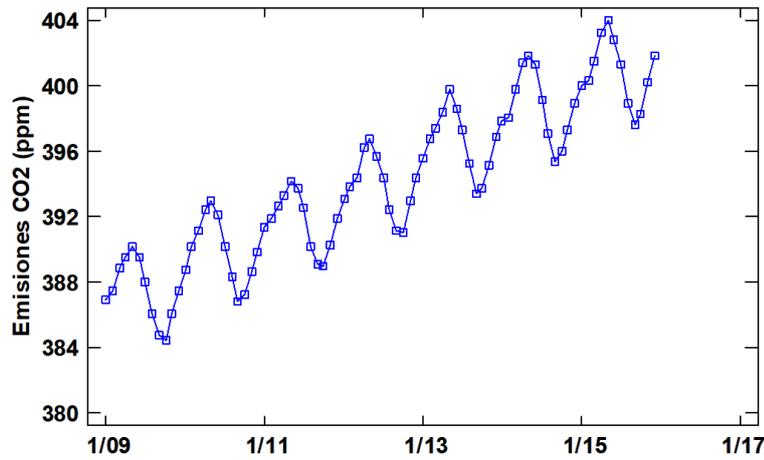


Ilustración 3.33 Gráfica de serie de tiempo para emisiones CO₂.

Como se puede observar en esta gráfica de serie de tiempo para emisiones de CO₂ tiene una tendencia positiva a medida que va pasando el tiempo, con unos puntos máximos en los meses de mayo y con puntos mínimos en los meses de octubre y noviembre, por lo cual, se identifica un comportamiento estacional. También se puede decir que consta de una varianza constante ya que los datos se encuentran entre el mismo rango.

Autocorrelaciones Estimadas para Emisiones CO₂ (ppm)

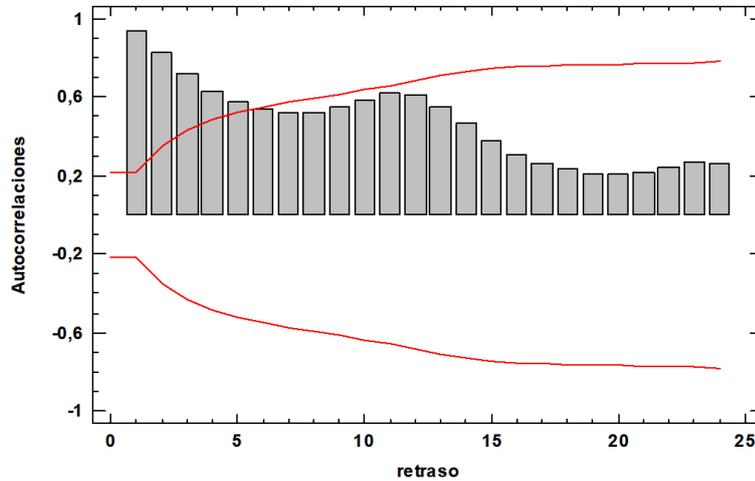


Ilustración 3.34 Autocorrelación emitidas para emisiones CO₂.

Analizando este gráfico de autocorrelaciones estimadas para emisiones de CO₂ se identifica un problema de autocorrelación hasta de 5^o orden positivo, donde los valores de autocorrelación van bajando en forma sinusoidal hace cero. Este comportamiento se asocia a la serie estacional.

Se concluye que la serie no es estacionaria ya que tiene tendencia positiva y variaciones estacionales.

Para hacer la serie estacionaria se ha aplicado una diferenciación no estacional y una diferenciación estacional.

Gráfica de Serie de Tiempo para ajuste de Emisiones CO₂ (ppm)

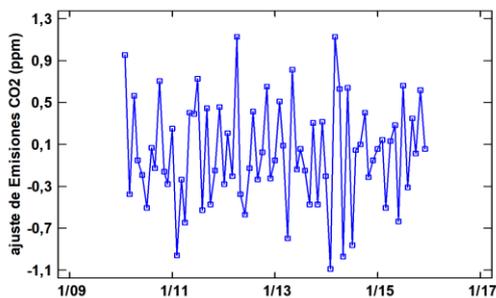


Ilustración 3.35 Gráfico serie de tiempo.

Autocorrelaciones Estimadas para ajuste de Emisiones CO₂ (ppm)

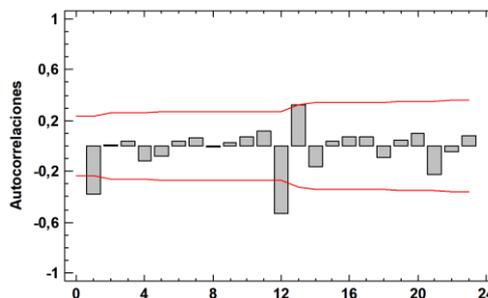


Ilustración 3.36 Autocorrelación estimada.

Tras aplicar estas órdenes la serie ya es estacionaria por lo tanto se propone un modelo ARIMA.

$$\text{ARIMA } (0,1,1) \times (0,1,1)_{12}$$

Validación modelo ARIMA

Para la validación del modelo ARIMA se debe constatar 4 pasos.

1º Significatividad del modelo.

Tabla 3.22 Resumen de modelo ARIMA.

Parámetro	Estimado	Error Estd.	T	Valor-P
MA(1)	0,479598	0,108563	4,41768	0,000036
SMA(1)	0,819548	0,0453925	18,0547	0,000000

Como el P-Valor es menor que α el modelo es significativo.

2º Gráfico de residuos.

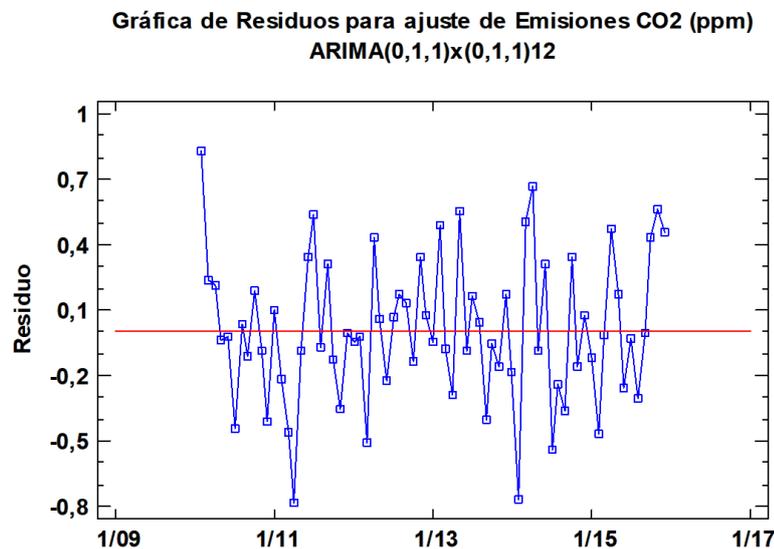


Ilustración 3.37 Gráfica de residuos para ajuste de emisiones CO₂.

Tras analizar la gráfica de residuos para ajuste de emisiones de CO₂ se comprueba que los residuos tienen media nula, no obstante, se detecta un problema de homocedasticidad ya que en los residuos no se observa forma cónica.

3º Comprobar la normalidad de residuos.

Tabla 3.23 Pruebas de normalidad para residuos.

Prueba	Estadístico	Valor-P
Chi-Cuadrado	23,3521	0,177419
Estadístico W de Shapiro-Wilk	0,982529	0,739817
Valor-Z para asimetría	0,127678	0,898399
Valor-Z para curtosis	0,0549694	0,956157

P-valor: 0,177419 > α : Se concluye que los residuos son normales.

4º Comprobar que el modelo ARIMA no tiene autocorrelación.

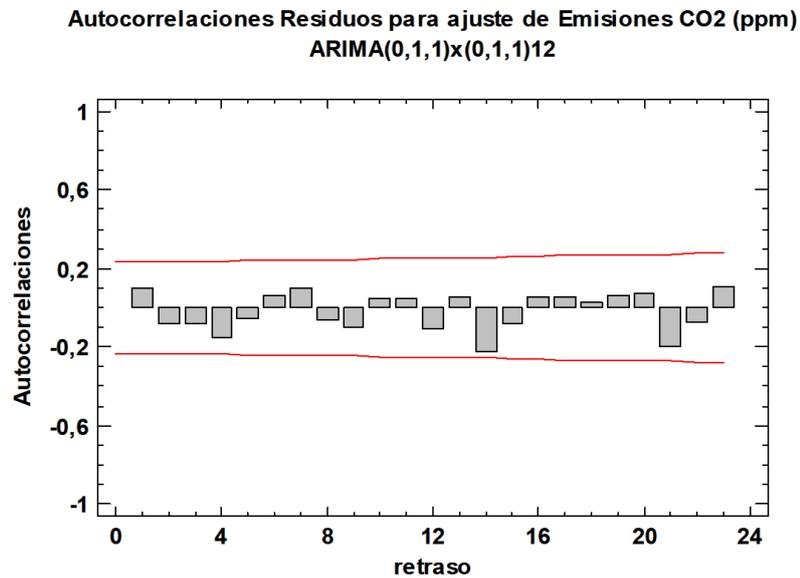


Ilustración 3.38 Autocorrelaciones residuos para ajuste de emisiones CO₂.

Se observa en la (Ilustración 3. 38) que ningún dato supera los límites, por lo tanto, se puede decir que no hay problemas de correlación.

Tras comprobar que los 4 pasos de validación han sido superados se concluye que el modelo ARIMA es válido.

Predicción modelo ARIMA

Tabla 3.24 Predicción modelo ARIMA.

Periodo	Pronóstico	Límite en 95,0% Inferior	Límite en 95,0% Superior
1/16	402,831	402,127	403,535
2/16	403,469	402,675	404,264
3/16	404,553	403,679	405,428
4/16	405,901	404,952	406,849
5/16	406,623	405,606	407,64
6/16	405,741	404,66	406,822
7/16	404,175	403,033	405,316
8/16	402,034	400,835	403,233
9/16	400,615	399,362	401,869
10/16	400,92	399,614	402,226
11/16	402,498	401,141	403,854
12/16	404,036	402,631	405,441

Gráfica de Secuencia en Tiempo para Emisiones CO2 (ppm)
ARIMA(0,1,1)x(0,1,1)12

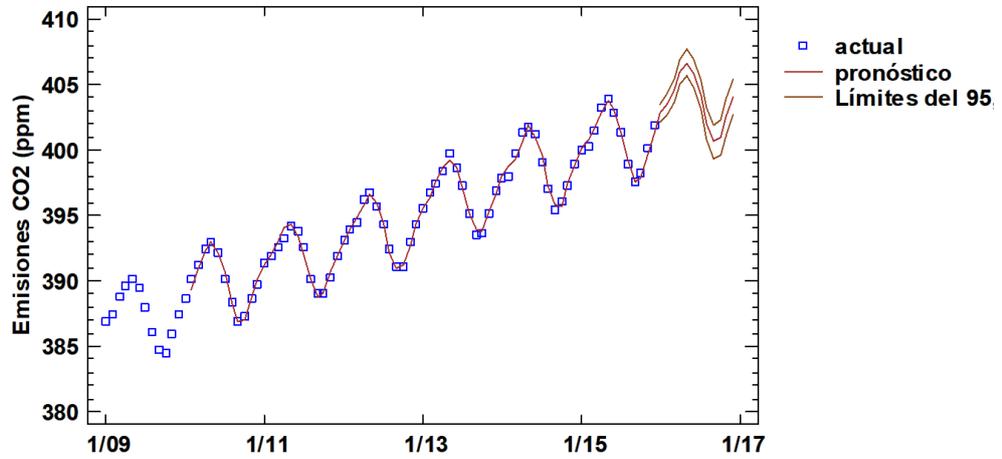


Ilustración 3.39 Gráfica de secuencia en tiempo para emisiones CO₂.

Como se observa en la (Tabla X y la ilustración Y) año tras año el comportamiento de las emisiones de CO₂ son similares, incluso su predicción sigue el mismo comportamiento estacional con tendencia positiva.



4. *CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN*

4.1 Conclusiones

- Existe una elevada multicolinealidad entre las variables consumo de electricidad y consumo de productos petrolíferos. También existe gran correlación entre las variables consumo de energías renovables y consumo de gas y manufacturados.
- Como solución al problema de multicolinealidad se ha optado por el método factorial, y con 2 factores se explican el 89'31% de los datos.
- Se logra un modelo válido con un R^2 de 57,72% el cual parecía bajo, pero sirve para ver el efecto de las variables explicativas sobre la dependiente.
- En la validación del modelo se ha tenido que solucionar los problemas de autocorrelación.
- No se ha encontrado ningún dato anómalo candidato a eliminar.
- Con el modelo ARIMA de series temporales se consigue una buena herramienta para la predicción del CO_2 .



4.2 Futuras líneas de investigación

Seguir con el mismo proyecto actualizando los datos, ya que el trabajo esta realizado con los datos entre los meses del 2009 hasta el 2015, pudiendo comprobar las predicciones y realizar nuevas futuras predicciones.

También se podría probar a realizar distintas estructuras con el modelo de regresión lineal como la estructura logarítmica, semilogarítmica, inversa, logarítmico-inversa.



5. *BIBLIOGRAFIA*

5. Bibliografía

- [1] <https://remicaopinion.es/novedades/fuentes-de-energia/> 28/02/2018
- [2] https://es.wikipedia.org/wiki/Media_aritm%C3%A9tica 04/06/2018
- [3] <https://es.plusmaths.com/la-mediana-en-estadistica.html> 04/06/2018
- [4] [https://es.wikipedia.org/wiki/Rango_\(estad%C3%ADstica\)](https://es.wikipedia.org/wiki/Rango_(estad%C3%ADstica)) 04/06/2018
- [5] <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/graphs/how-to/boxplot/interpret-the-results/quartiles/> 06/06/2018
- [6] <https://www.definicionabc.com/general/varianza.php> 04/06/2018
- [7] <https://okdiario.com/curiosidades/2017/07/13/desviacion-tipica-1133576> 04/06/2018
- [8] <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/graphs/how-to/boxplot/interpret-the-results/key-results/> 09/03/2018
- [9] <http://www5.uva.es/estadmed/datos/bivariante/bivar.htm> 25/03/2018
- [10] http://www.ite.educacion.es/formacion/materiales/180/cd/m4_10/grficos_xy_dispersin.html 07/06/2018
- [11] <https://es.wikipedia.org/wiki/Multicolinealidad> 07/06/2018
- [12] https://es.wikipedia.org/wiki/Matriz_de_correlaci%C3%B3n 07/06/2018
- [13] <https://es.slideshare.net/federicodonneysg/analisis-multivariado-8829215>
09/06/2018
- [14] https://es.wikipedia.org/wiki/Significaci%C3%B3n_estad%C3%ADstica 09/06/2018
- [15] <http://economipedia.com/definiciones/heterocedasticidad.html> 10/06/2018

6. ANEXO

MES	Emisiones CO ₂ (ppm)	Carbón	Gases derivados del carbón	Productos petrolíferos	Gas natural y manufacturado	Electricidad	Energías renovables
2009M01	386,94	107	30	4754	1594	2034	462
2009M02	387,48	103	28	4401	1377	1790	446
2009M03	388,82	107	24	4758	1236	1802	437
2009M04	389,55	107	17	4574	1049	1658	432
2009M05	390,14	107	16	4314	933	1690	401
2009M06	389,48	95	12	4434	888	1772	380
2009M07	388,03	93	12	4746	933	1946	398
2009M08	386,11	87	8	4313	775	1839	329
2009M09	384,74	108	12	4446	937	1763	382
2009M10	384,43	114	12	4532	1033	1742	414
2009M11	386,02	103	19	4369	1181	1764	456
2009M12	387,42	100	25	4551	1482	1957	469
2010M01	388,71	124	32	4293	1634	2035	523
2010M02	390,2	119	26	4251	1508	1876	497
2010M03	391,17	124	26	4908	1458	1962	508
2010M04	392,46	124	26	4342	1116	1727	432
2010M05	393	124	26	4320	1120	1752	446
2010M06	392,15	110	19	4407	964	1758	418
2010M07	390,2	107	19	4555	931	2012	420
2010M08	388,35	101	13	4318	805	1861	378
2010M09	386,85	126	19	4248	981	1790	426
2010M10	387,24	132	19	4448	1111	1764	448
2010M11	388,67	119	19	4452	1460	1881	567
2010M12	389,79	116	19	4496	1688	1991	604
2011M01	391,33	162	28	4206	1793	1903	550
2011M02	391,86	146	25	4179	1522	1715	508
2011M03	392,6	152	28	4584	1560	1823	526
2011M04	393,25	150	27	4123	1059	1552	456
2011M05	394,19	142	30	4053	1035	1647	463
2011M06	393,73	137	28	4159	912	1684	447
2011M07	392,51	129	24	4357	936	1779	400
2011M08	390,13	102	24	4130	775	1755	441
2011M09	389,08	111	26	4064	883	1707	451
2011M10	389	119	25	4045	1103	1655	468
2011M11	390,28	111	22	4029	1377	1663	505



CAMPUS D'ALCOI

2011M12	391,86	95	20	4064	1638	1753	559
2012M01	393,12	120	28	4005	1692	1883	629
2012M02	393,86	122	28	4067	1778	1868	565
2012M03	394,4	129	30	4001	1248	1794	588
2012M04	396,18	130	28	3758	1144	1626	493
2012M05	396,74	106	30	3744	1068	1649	490
2012M06	395,71	110	27	3822	962	1702	470
2012M07	394,36	115	26	3920	969	1777	486
2012M08	392,39	82	14	3877	865	1764	441
2012M09	391,11	80	16	3445	1015	1624	443
2012M10	391,05	79	16	3746	1207	1612	490
2012M11	392,98	79	17	3636	1427	1644	545
2012M12	394,34	83	15	3519	1613	1720	633
2013M01	395,55	106	17	3598	1703	1811	564
2013M02	396,8	114	18	3346	1492	1709	573
2013M03	397,43	138	19	3731	1580	1700	457
2013M04	398,41	122	20	3729	1286	1572	449
2013M05	399,78	135	22	3579	1210	1576	420
2013M06	398,61	124	21	3472	1055	1552	386
2013M07	397,32	175	20	3921	1041	1761	389
2013M08	395,2	124	17	3603	780	1684	359
2013M09	393,45	127	19	3440	990	1600	374
2013M10	393,7	125	21	3746	1012	1605	395
2013M11	395,16	116	18	3765	1428	1648	440
2013M12	396,84	118	17	3674	1677	1734	487
2014M01	397,85	99	20	3508	1654	1764	491
2014M02	398,01	91	20	3268	1534	1626	439
2014M03	399,77	106	19	3635	1426	1678	487
2014M04	401,38	87	20	3478	1091	1505	427
2014M05	401,78	97	21	3463	1065	1564	429
2014M06	401,25	108	19	3363	976	1591	397
2014M07	399,1	95	18	3836	981	1720	398
2014M08	397,03	94	20	3471	1136	1644	321
2014M09	395,38	92	18	3464	962	1542	381
2014M10	396,03	89	20	3687	1026	1590	373
2014M11	397,28	97	18	3481	1345	1583	454
2014M12	398,91	86	12	3612	1583	1705	511
2015M01	399,98	105	20	3672	1733	1816	520
2015M02	400,28	103	17	3480	1615	1678	489
2015M03	401,54	109	20	3649	1423	1694	479
2015M04	403,28	102	22	3426	1067	1512	400
2015M05	403,96	97	21	3446	996	1602	407
2015M06	402,8	92	21	3554	905	1649	395
2015M07	401,31	101	21	3748	908	1902	406
2015M08	398,93	90	20	3569	843	1705	436



CAMPUS D'ALCOI

2015M09	397,63	101	21	3622	957	1593	387
2015M10	398,29	93	21	3604	1125	1605	400
2015M11	400,16	98	19	3495	1298	1586	467
2015M12	401,85	113	17	3616	1475	1658	518

6.1 Índice de tablas

Tabla 3.1 Resumen estadísticos para emisiones de CO₂.

Tabla 3.2 Resumen Estadístico para consumo de carbón.

Tabla 3.3 Resumen Estadístico para Gases derivados del carbón.

Tabla 3.4 Resumen estadístico para productos petrolíferos.

Tabla 3.5 Resumen Estadístico para gas natural y manufacturado.

Tabla 3.6 Resumen estadístico para electricidad.

Tabla 3.7 Resumen estadístico para energías renovables.

Tabla 3.8 Matriz de correlación.

Tabla 3.9 Matriz inversa de correlación.

Tabla 3.10 Índice de acondicionamiento.

Tabla 3.11 Análisis de componentes principales.

Tabla 3.12 Tabla de pesos de los componentes

Tabla 3.13 Coeficiente del factor.

Tabla 3.14 Análisis de factores.

Tabla 3.15 Prueba del análisis.

Tabla 3.16 Análisis de varianza.

Tabla 3.17 Pruebas de normalidad para residuos

Tabla 3.18 Heterocedasticidad.

Tabla 3.19 Puntos a priori.

Tabla 3.20 Puntos a posteriori.

Tabla 3.21 Puntos anómalos.

Tabla 3.22 Resumen de modelo ARIMA.

Tabla 3.23 Pruebas de normalidad para residuos.

Tabla 3.24 Predicción modelo ARIMA.

6.2 Índice de ilustraciones

Ilustración 1.1 Emisiones CO₂.

Ilustración 2.1 Los cuartiles separan los datos en cuatro segmentos con igual número de datos.

Ilustración 2.2 Caja y Bigotes.

Ilustración 2.3 Histograma.

Ilustración 2.4 Gráfico de probabilidad normal.

Ilustración 2.5 Función de autocorrelación simple.

Ilustración 2.6 Función de autocorrelación parciales.

Ilustración 2.7 Puntos influyente.

Ilustración 3.1 Gráfico de caja y bigotes para emisiones de CO₂.

Ilustración 3.2 Histograma para emisiones de CO₂.

Ilustración 3.3 Gráfica de serie de tiempo para emisiones de CO₂.

Ilustración 3.4 Gráfica de caja y bigotes para consumo de carbón.

Ilustración 3.5 Histograma para consumo de carbón.

Ilustración 3.6 Gráficas de serie de tiempo para consumo de carbón.

Ilustración 3.7 Gráfico de caja y bigote para Gases derivados del carbón.

Ilustración 3.8 Histograma para Gases derivados del carbón.

Ilustración 3.9 Gráfico de serie de tiempo para Gases derivados del carbón.

Ilustración 3.10 Gráfico de caja y bigotes para productos petrolíferos.

Ilustración 3.11 Histograma para productos petrolíferos.

Ilustración 3.12 Gráfica serie de tiempo para productos petrolíferos.

Ilustración 3.13 Gráfico de caja y bigotes para gas natural y manufacturado.

Ilustración 3.14 Histograma para gas natural y manufacturado.

Ilustración 3.15 Gráfica de serie de tiempo para gas natural y manufacturado.

Ilustración 3.16 Gráfico de caja y bigote para electricidad.

Ilustración 3.18 Histograma para electricidad.

Ilustración 3.19 Gráfica de serie de tiempo para electricidad.

Ilustración 3.20 Gráfico de caja y bigotes para energías renovables.

Ilustración 3.21 Histograma para energías renovables.

Ilustración 3.22 Gráfica de serie de tiempo para energías renovables.

Ilustración 3.23 Gráfico del modelo ajustado.

Ilustración 3.24 Gráfico del modelo ajustado.

Ilustración 3.25 Gráfico del modelo ajustado.



Ilustración 3.26 Gráfico del modelo ajustado.

Ilustración 3.27 Gráfico del modelo ajustado.

Ilustración 3.28 Gráfico del modelo ajustado.

Ilustración 3.29 Gráfico de probabilidad normal.

Ilustración 3.30 Histograma para residuos.

Ilustración 3.31 Autocorrelación estimada para residuos.

Ilustración 3.32 Autocorrelación estimada para residuos.

Ilustración 3.33 Gráfica de serie de tiempo para emisiones CO₂.

Ilustración 3.34 Autocorrelación emitidas para emisiones CO₂.

Ilustración 3.35 Gráfico serie de tiempo.

Ilustración 3.36 Autocorrelación estimada.

Ilustración 3.37 Gráfica de residuos para ajuste de emisiones CO₂.

Ilustración 3.38 Autocorrelaciones residuos para ajuste de emisiones CO₂.

Ilustración 3.39 Gráfica de secuencia en tiempo para emisiones CO₂.