



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)

INGENIERO INDUSTRIAL

ANÁLISIS DE ALGUNOS INDICADORES DE SOSTENIBILIDAD EN ESPAÑA

Autor: Eva Marina González Isla

Director: Pedro Linares Llamas

Madrid

Septiembre 2018

AUTORIZACIÓN PARA LA DIGITALIZACIÓN, DEPÓSITO Y DIVULGACIÓN EN RED DE PROYECTOS FIN DE GRADO, FIN DE MÁSTER, TESINAS O MEMORIAS DE BACHILLERATO

1º. Declaración de la autoría y acreditación de la misma.

El autor D. Eva Marina González Isla DECLARA ser el titular de los derechos de propiedad intelectual de la obra: Análisis de , que ésta es una obra original, y que ostenta la condición de autor en el sentido que otorga la Ley de Propiedad Intelectual.

2º. Objeto y fines de la cesión.

Con el fin de dar la máxima difusión a la obra citada a través del Repositorio institucional de la Universidad, el autor **CEDE** a la Universidad Pontificia Comillas, de forma gratuita y no exclusiva, por el máximo plazo legal y con ámbito universal, los derechos de digitalización, de archivo, de reproducción, de distribución y de comunicación pública, incluido el derecho de puesta a disposición electrónica, tal y como se describen en la Ley de Propiedad Intelectual. El derecho de transformación se cede a los únicos efectos de lo dispuesto en la letra a) del apartado siguiente.

3º. Condiciones de la cesión y acceso

Sin perjuicio de la titularidad de la obra, que sigue correspondiendo a su autor, la cesión de derechos contemplada en esta licencia habilita para:

- a) Transformarla con el fin de adaptarla a cualquier tecnología que permita incorporarla a internet y hacerla accesible; incorporar metadatos para realizar el registro de la obra e incorporar “marcas de agua” o cualquier otro sistema de seguridad o de protección.
- b) Reproducir la en un soporte digital para su incorporación a una base de datos electrónica, incluyendo el derecho de reproducir y almacenar la obra en servidores, a los efectos de garantizar su seguridad, conservación y preservar el formato.
- c) Comunicarla, por defecto, a través de un archivo institucional abierto, accesible de modo libre y gratuito a través de internet.
- d) Cualquier otra forma de acceso (restringido, embargado, cerrado) deberá solicitarse expresamente y obedecer a causas justificadas.
- e) Asignar por defecto a estos trabajos una licencia Creative Commons.
- f) Asignar por defecto a estos trabajos un HANDLE (URL *persistente*).

4º. Derechos del autor.

El autor, en tanto que titular de una obra tiene derecho a:

- a) Que la Universidad identifique claramente su nombre como autor de la misma
- b) Comunicar y dar publicidad a la obra en la versión que ceda y en otras posteriores a través de cualquier medio.
- c) Solicitar la retirada de la obra del repositorio por causa justificada.
- d) Recibir notificación fehaciente de cualquier reclamación que puedan formular terceras personas en relación con la obra y, en particular, de reclamaciones relativas a los derechos de propiedad intelectual sobre ella.

5º. Deberes del autor.

El autor se compromete a:

- a) Garantizar que el compromiso que adquiere mediante el presente escrito no infringe ningún derecho de terceros, ya sean de propiedad industrial, intelectual o cualquier otro.
- b) Garantizar que el contenido de las obras no atenta contra los derechos al honor, a la intimidad y a la imagen de terceros.
- c) Asumir toda reclamación o responsabilidad, incluyendo las indemnizaciones por daños, que pudieran ejercitarse contra la Universidad por terceros que vieran infringidos sus derechos e intereses a causa de la cesión.

- d) Asumir la responsabilidad en el caso de que las instituciones fueran condenadas por infracción de derechos derivada de las obras objeto de la cesión.

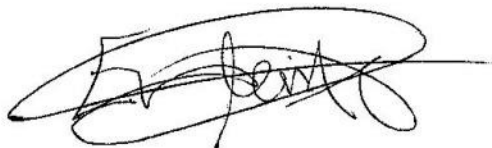
6º. Fines y funcionamiento del Repositorio Institucional.

La obra se pondrá a disposición de los usuarios para que hagan de ella un uso justo y respetuoso con los derechos del autor, según lo permitido por la legislación aplicable, y con fines de estudio, investigación, o cualquier otro fin lícito. Con dicha finalidad, la Universidad asume los siguientes deberes y se reserva las siguientes facultades:

- La Universidad informará a los usuarios del archivo sobre los usos permitidos, y no garantiza ni asume responsabilidad alguna por otras formas en que los usuarios hagan un uso posterior de las obras no conforme con la legislación vigente. El uso posterior, más allá de la copia privada, requerirá que se cite la fuente y se reconozca la autoría, que no se obtenga beneficio comercial, y que no se realicen obras derivadas.
- La Universidad no revisará el contenido de las obras, que en todo caso permanecerá bajo la responsabilidad exclusiva del autor y no estará obligada a ejercitar acciones legales en nombre del autor en el supuesto de infracciones a derechos de propiedad intelectual derivados del depósito y archivo de las obras. El autor renuncia a cualquier reclamación frente a la Universidad por las formas no ajustadas a la legislación vigente en que los usuarios hagan uso de las obras.
- La Universidad adoptará las medidas necesarias para la preservación de la obra en un futuro.
- La Universidad se reserva la facultad de retirar la obra, previa notificación al autor, en supuestos suficientemente justificados, o en caso de reclamaciones de terceros.

Madrid, a 03 de Septiembre de 2018

ACEPTA



Fdo Eva Marina González Isla

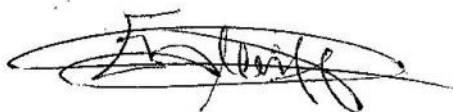
Motivos para solicitar el acceso restringido, cerrado o embargado del trabajo en el Repositorio Institucional:

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título
ANÁLISIS DE ALGUNOS INDICADORES DE SOSTENIBILIDAD EN
ESPAÑA en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas
en el

curso académico 2017/2018 es de mi autoría, original e inédito y
no ha sido presentado con anterioridad a otros efectos. El Proyecto no es
plagio de otro, ni total ni parcialmente y la información que ha sido tomada
de otros documentos está debidamente referenciada.

Fdo.: González Isla, Eva Marina

Fecha: 03/09/2018



Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO



Fdo.: Linares Llamas, Pedro

Fecha: 04/09/2018



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)

INGENIERO INDUSTRIAL

ANÁLISIS DE ALGUNOS INDICADORES DE SOSTENIBILIDAD EN ESPAÑA

Autor: Eva Marina González Isla

Director: Pedro Linares Llamas

Madrid

Septiembre 2018

ANÁLISIS DE ALGUNOS INDICADORES DE SOSTENIBILIDAD DE ESPAÑA

Autor: González Isla, Eva Marina.

Director: Linares Llamas, Pedro.

Entidad colaboradora: Cátedra BP de Energía y Sostenibilidad de la Universidad Pontificia de Comillas.

RESUMEN DEL PROYECTO

Introducción y planteamiento del problema

La Cátedra BP de Energía y Sostenibilidad de la Universidad Pontificia de Comillas es una entidad que surge de un “Acuerdo Marco de Colaboración” entre la Universidad y BP España en 2002, y tiene como finalidad “promover estudios, acciones formativas y divulgación” sobre la consecución de un modelo de desarrollo sostenible en España [CBPEsf].

Una de las herramientas que utiliza la Cátedra con el fin de “divulgar y crear conciencia social en temas de energía y desarrollo sostenible” son los Observatorios. Éstos son informes publicados normalmente con carácter anual en los que se analiza la evolución energética durante ese periodo de tiempo, para determinar si ha sido beneficiosa o no desde el punto de vista del desarrollo sostenible. Su contenido incluye la elaboración y análisis de algunos indicadores energéticos contruidos a partir de diversas fuentes de información, comparando la situación española con la de otras potencias de parecida situación económica [CBPEsf].

Cambios en los indicadores pueden producirse por reformas en las regulaciones, como la directiva de eficiencia energética de la UE; presiones en el ámbito político, como el precio del petróleo o el gas, y por último, variabilidad aleatoria, como la presente en las condiciones meteorológicas.

La existencia de variabilidad aleatoria limita en cierta medida el alcance de los análisis realizados en los Observatorios, por lo que en este proyecto se tratará de eliminar parte de esa variabilidad aleatoria de algunos indicadores de sostenibilidad. Las áreas de trabajo abordadas en este proyecto son la producción de energías renovables y el consumo de demanda térmica, concretamente, se ha decidido estudiar la producción de energía hidráulica y el consumo energético de uso térmico tanto del sector residencial como del sector servicios en un año típico en España.

Para evaluar el impacto en los indicadores se ha de utilizar el modelo de simulación del sector energético español de la Cátedra, introduciendo en él los datos corregidos de la demanda térmica y la producción hidroeléctrica.

Metodología

Para eliminar la componente aleatoria de la producción hidroeléctrica anual y de la demanda térmica de los sectores residencial y servicios se han seguido procedimientos distintos, ya que la producción hidroeléctrica depende esencialmente de las precipitaciones, que son un suceso aleatorio e independiente, mientras que la demanda térmica depende de factores económicos, tecnológicos, demográficos, y por supuesto meteorológicos.

En el caso de la producción eléctrica, la metodología para obtener el valor típico de España fue encontrar su distribución de probabilidad para posteriormente hallar el valor típico con el cálculo de la esperanza. Para determinar la distribución de probabilidad se ha supuesto que la variable aleatoria en cuestión proviene de una familia de probabilidad conocida, de la que se estimaron los parámetros mediante el método de máxima verosimilitud. Por su presencia en la naturaleza, se escogieron como distribuciones posibles la normal, la log-normal, la Weibull y la gamma. Una vez estimados todos los parámetros se determinó la distribución que mejor se ajustaba a los datos mediante contrastes de bondad de ajuste entre la muestra y las distribuciones teóricas.

En el caso de la demanda térmica, puesto que depende de distintos factores, se emplearon modelos de regresión lineal múltiple y se obtuvo la serie corregida sustrayendo la componente climática de dicho modelo. Para evitar la aparición de coeficientes despreciables frente al resto en la elaboración de los modelos como consecuencia de la distinta naturaleza de las variables explicativas, se utilizaron valores normalizados de las variables explicativas según su media y desviación típica, como muestra la ecuación (a). Posteriormente se determinaron los coeficientes empleando el método de máxima verosimilitud.

$$Z_i = \frac{X_i - \bar{X}}{\sigma(X)} \quad (a)$$

Tras ello, se evaluó el poder explicativo del modelo mediante contrastes de hipótesis sobre el modelo, como el F-test del modelo; sobre los parámetros, a través de t-tests individuales, y finalmente se estudió la correlación entre la variable dependiente y las explicativas mediante el coeficiente de correlación y el coeficiente de correlación ajustado [CARO13]. De los modelos obtenidos se escogió el más adecuado, es decir, aquel que poseía una mayor correlación y que podía ser aceptado con un nivel de significación menor.

Resultados

Para representar la cantidad típica de energía hidráulica que puede producirse en España se utilizó el producible hidráulico, del cual se adquirió una muestra de Red Eléctrica de España [REE_sf].

En la Ilustración 1 se hallan representadas todas las posibles distribuciones de densidad de probabilidad superpuestas con el histograma de los datos de la muestra. Las distribuciones fueron ajustadas con la función *fitdist* de Matlab.

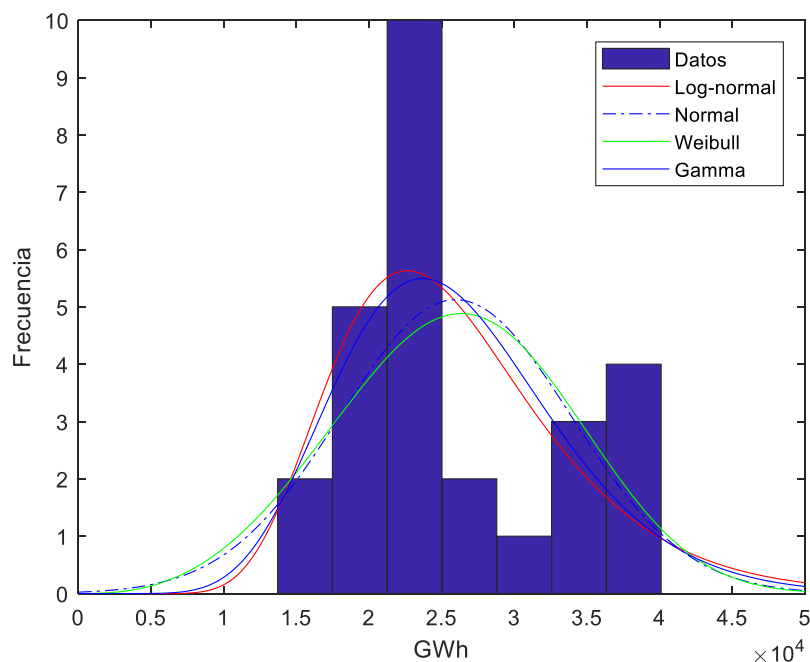


Ilustración 1 : Histograma de datos superpuesto con las posibles funciones de probabilidad. Elaboración propia utilizando Matlab.

La muestra resultó no ser lo suficientemente grande como para realizar el contraste de bondad de ajuste Chi-cuadrado. Los resultados del test de Kolmogorov-Smirnov tampoco fueron concluyentes para rechazar ninguna de las distribuciones, aunque de acuerdo al p-valor obtenido, la más apropiada era la log-normal. Finalmente, la única distribución que no pudo ser rechazada en el contraste de Anderson-Darling fue también la log-normal. Por lo tanto, el valor esperado del producible hidráulico español que se obtuvo es de 26.004 GWh, que es la esperanza de la distribución log-normal obtenida.

Respecto al consumo de calefacción del sector residencial, el modelo más apropiado obtenido era significativo al 5% y poseía un coeficiente de correlación de 0,86 y un 0,83 de correlación ajustada. Las variables que lo componían eran el PIB a precios constantes [BANCsf], el ODEX residencial, un índice de eficiencia energética determinado por ODYSSEE [ODYS17], y los HDD (por sus siglas en inglés), que representan el número de días en los que la temperatura interior (o temperatura de confort) difiere de la temperatura exterior y en cuantos grados.

Para el cálculo de los HDD se utilizaron las temperaturas máximas y mínimas diarias en las capitales de provincia [AEME18] y se escogió la temperatura de confort de 18°C. El procedimiento para el cálculo de un día se describe en la Ilustración 2. Para obtener resultados anuales se sumaron los resultados para todos los días del año y se multiplicaron por el porcentaje de población en cada provincia [INE_18].

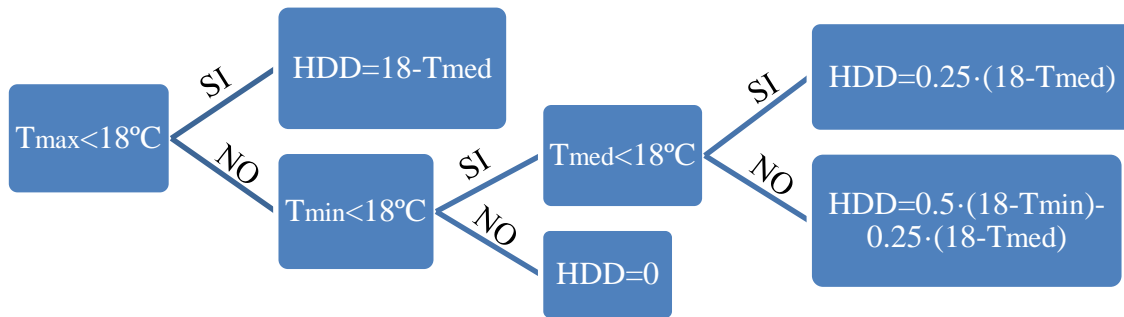


Ilustración 2 : Árbol de decisión en el cálculo de HDD. Elaboración propia a partir de [LINA18].

Los resultados de la serie corregida se muestran en el Gráfico 1. Se observan pequeñas diferencias entre las series, llegando a ser del 5% para los años más extremos.

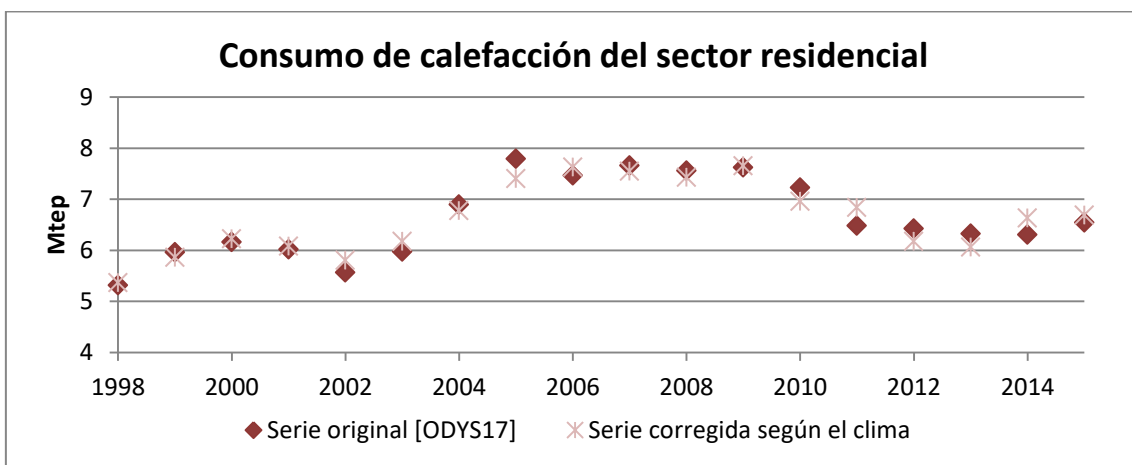


Gráfico 1 : Serie corregida según el clima del consumo de calefacción del sector residencial. Elaboración propia.

En el caso de consumo de calefacción del sector servicios, los HDD resultaron inútiles como variables de los modelos de regresión. Se obtuvieron niveles de significación elevados e incluso coeficientes negativos en repetidas ocasiones. Las variables del mejor modelo fueron la temperatura media en invierno y el PIB a precios constantes [BANCsf]. Este modelo era significativo al 1%, y poseía un coeficiente de correlación de 0.78 y de correlación ajustada de 0.75.

La serie corregida se presenta en el Gráfico 2. Se observa que las diferencias son mayores que en el caso del sector residencial, llegando a porcentajes de 12% en los años más extremos.

Finalmente, no fue posible obtener modelos dependientes de la temperatura exterior para el caso del consumo de aire acondicionado de ambos sectores.

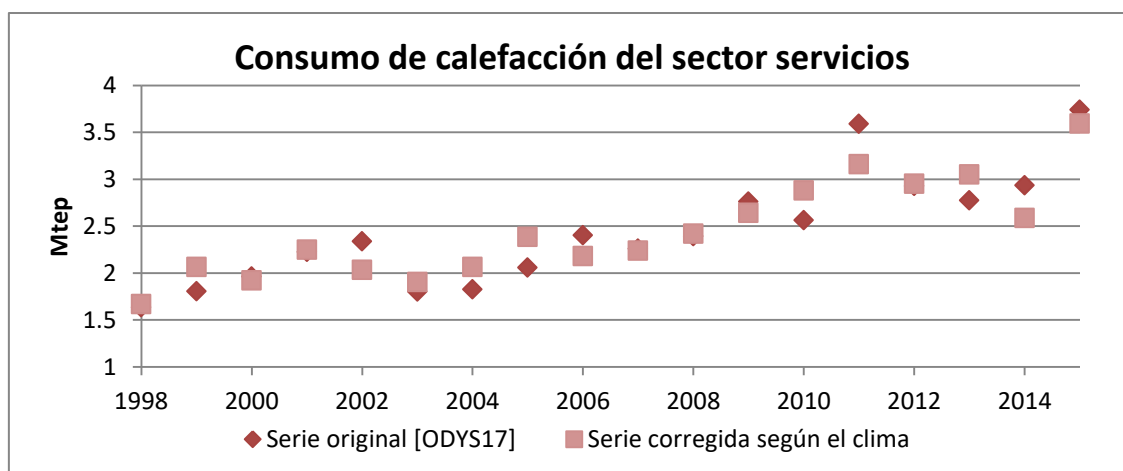


Gráfico 2 : Resultados de la serie corregida según el clima del consumo de calefacción del sector servicios. Elaboración propia.

Conclusiones

Respecto a la energía hidráulica, el producible hidráulico en España durante un año típico es de 26.004 GWh.

Por otro lado, el consumo corregido de calefacción según el clima puede obtenerse sustrayendo a las series originales en mtep el producto de la variable climática determinada en el modelo de regresión lineal por su coeficiente de regresión. En el caso del sector residencial se debe multiplicar 0,2164 por los HDD normalizados y en el caso del sector servicios el coeficiente vale 0,2315 y la variable climática es la temperatura media en invierno normalizada.

Los resultados de la serie corregida del consumo de calefacción del sector residencial se consideran más sólidos que los del sector servicios por la diferencia de verosimilitud de los HDD frente a la temperatura media en invierno para cuantificar el efecto de la temperatura exterior en el uso de calefacción. Además, la tendencia de ambas series corregidas respecto a los datos originales resulta opuesta la mayoría de los años. Por ello, existen motivos para poner en duda la relación entre el consumo de calefacción del sector servicios y la temperatura exterior.

Por último, respecto al consumo de aire acondicionado, la falta de modelos explicativos tanto del sector residencial como del sector servicios puede ser consecuencia de que el mercado de aire acondicionado continúa creciendo.

Referencias

- [CBPEsf] Página web de la Cátedra BP de Energía y Sostenibilidad, s.f.
- [AEME18] Asociación Española de Meteorología (AEMET), 2018.
- [INE_18] Instituto Nacional de Estadística (INE), 2018.
- [BANCsf] Banco Mundial, “PIB (UMN a precios constantes)”, s.f.
- [ODYS17] ODYSSEE, “ODYSSEE database – july 2017”, 2017.
- [CARO13] Caro R., J. Reneses, “Estadística. No soy mala... me han dibujado así”. Thomson Reuters, Pamplona 2013.

- [LINA18] Linares P., Comentario sobre el método de cálculo de los HDD y los CDD, 2018.
- [REE_sf] Red Eléctrica de España, s.f.

ABSTRACT

Introduction and problem description

The BP Chair on Energy and Sustainability aims “to promote studies, training and spreading actions” about the achievement of a sustainable development model in Spain. The Chair was created in 2002, after the signature of a “Framework Agreement of Collaboration” between the Comillas Pontifical University and BP Spain [BPCEud].

Annual reports are published by the Observatory of the Chair “to spread and develop social awareness about Energy and Sustainable development”. The information issued in these reports includes the description of the Spanish energy situation alongside with some relevant indicators built from a great variety of sources. In the reports, the yearly evolution of these indicators is analysed, comparing the Spanish situation to others’ within our economic area and determining whether the changes are beneficial from a sustainable development perspective [BPCEud].

Changes in the indicators’ trends may occur for several reasons. Some examples could be updates in the EU directives, for instance in energy efficiency. Moreover, international politics might impact oil and gas prices; however, the indicators are also subjected to random variability, especially regarding climate conditions.

The scope of this project is to eliminate the random variability of some indicators, since the climate randomness restricts to the extent of the Observatory annual reports. The project focuses on the following areas: renewable energy production and thermal demand. More specifically, it aims to determine the hydropower production and the thermal demand of the residential and services sectors that occur in a typical climate year in Spain.

To evaluate the impact on the indicators, the climatic corrected data has to be inserted in the Spanish energy sector simulation model belonging to the Chair.

Methodology

Distinct approaches were undertaken to determine the hydropower production in Spain and the thermal demand from the residential and services sectors during a typical year. The main reason was that the hydroelectric energy produced depends essentially on the rain, which is an independent and random occurrence, while the thermal demand depends on economic, demographic, technological and climatological factors.

To obtain the Spanish hydropower production on a representative year its probability density distribution function was determined to get its expected value. It was supposed that the hydraulic production belonged to a known probability distribution among the following: normal, log-normal, Weibull and gamma, which have been chosen because of their occurrences in nature. Then, the parameters of all distributions were established using the maximum likelihood method. Finally, some goodness of fit hypothesis tests were performed to choose the best option.

To compute the thermal demand with climatic corrections, a multiple regression model was built and then the corrected data was estimated by subtracting the climatic component of the model. The explanatory variables were normalized using their mean and standard deviation according to equation (a) to avoid negligible coefficients in the linear models due to differences between the variables.

$$Z_i = \frac{X_i - \bar{X}}{\sigma(X)} \quad (a)$$

Then, the coefficients of the linear model were calculated using the maximum likelihood method. Afterwards, the linear model explanatory power was quantified. For this, a number of tools were used, such as an F-test on the whole model and t-tests on the individual parameters. It is important to measure the correlation between the dependent variable and the explanatory variables, for which the correlation coefficient and the adjusted correlation coefficient were used. Finally, the model with the highest statistical significance and the best correlation was accepted [CARO13].

Results

To determine the hydropower production in a typical year, the probability distribution function of the hydraulic production capacity was studied using a sample from the Spanish national grid [REE_ud].

The distribution functions with their estimated parameters alongside with the data histogram are represented in Figure 1. All the parameters from the probability density distribution functions were estimated using the *fitdistst* command from Matlab.

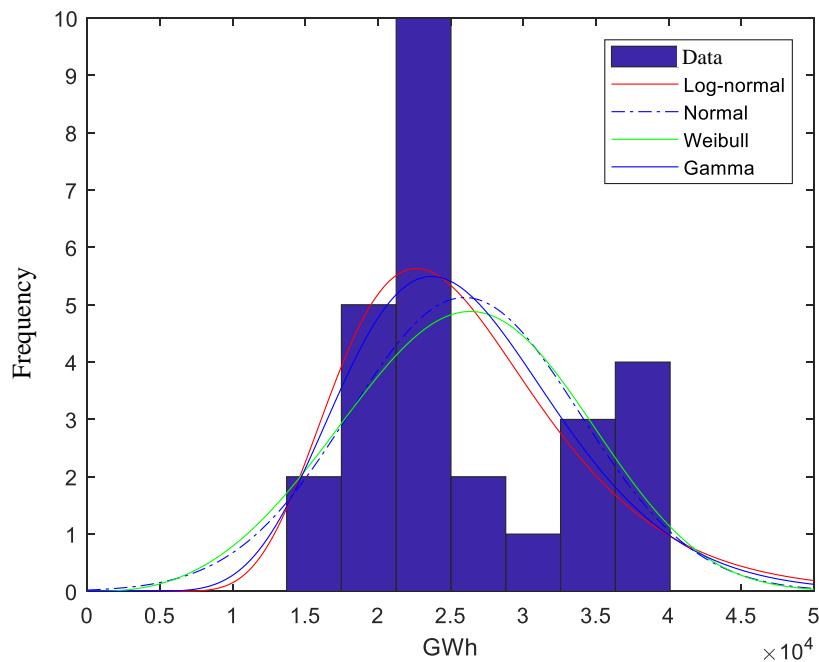


Figure 1: Superimposition of the data histogram and the different probability density distribution functions computed. Elaborated by the author using Matlab.

Regarding the goodness of fit hypothesis contrasting, the data sample was not big enough to perform the Chi-squared goodness of fit test. The execution of the Kolmogorov-Smirnov test was not conclusive either, since no probability distribution function was rejected; however, according to this test p-value, the best solution was the log-normal function. Finally, the completion of the Anderson-Darling statistician allowed the rejection of all the functions except the log-normal. Therefore, the typical Spanish hydraulic production capacity resulted to be 26,004 GWh, which is the expected value obtained by this log-normal distribution obtained.

With regard to the residential heating consumption, the most appropriate model was accepted with a 5% significance level and it held a 0.86 value for the correlation coefficient and a 0.83 value for the adjusted correlation coefficient. The explanatory variables for this model were the GDP in constant prices [WORLud], the ODEX index, which is a technology efficiency index computed by ODYSSEE [ODYS17]; and the Heating Degree Days (HDD), which quantify the difference between the outside temperature and a defined comfort temperature along the year.

The HDD for one day were computed following the steps shown in Figure 2 using the daily maximum and minimum temperatures in the Spanish province capitals [AEME18]. The comfort temperature was defined as 18°C and the year amount was obtained by computing the weighted addition according to the provinces population [INE_ud].

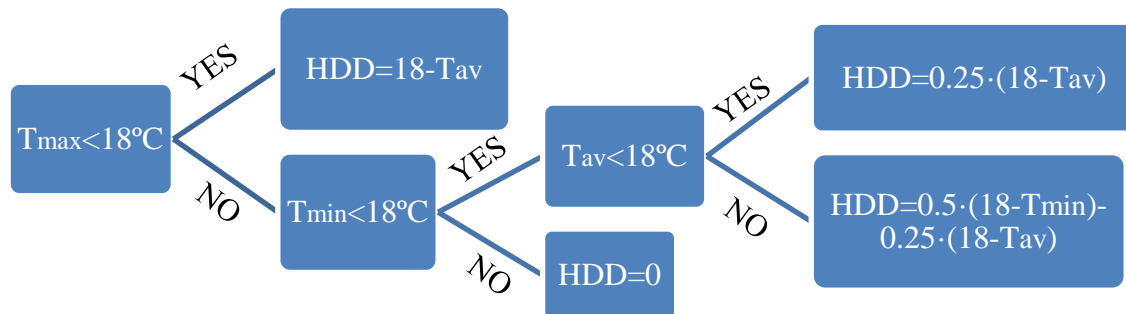


Figure 2: Decision tree for HDD calculation. Elaborated by the author from [LINA18].

The results for the residential heating consumption series with climatic corrections are shown in Figure 3. It can be observed that the differences between the corrected series and the original are within the 5% margin.

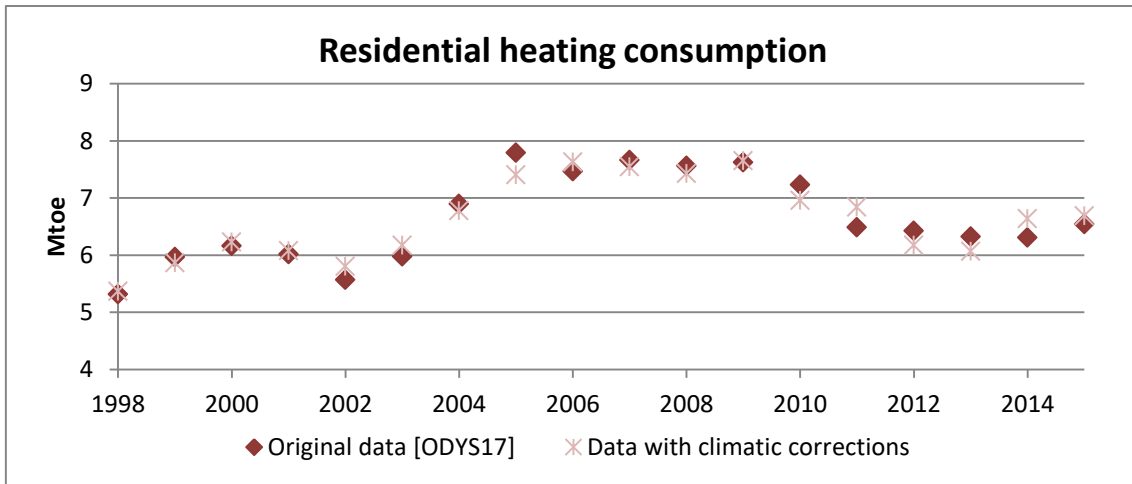


Figure 3 : Results for the corrected residential heating consumption. Elaborated by the author.

With respect to the service sector heating consumption, the HDD were not explanatory; negative coefficients and high individual p-values were obtained for all cases. The chosen model was composed by the GDP in constant prices [WORLud] and by the winter average temperature[AEME18]. This model was accepted with a significance level of 1%, a correlation coefficient of 0.78 and an adjusted correlation coefficient of 0.75.

The service sector heating consumption with climatic corrections is shown in Figure 4. The differences between both series are higher than in the residential sector, reaching the amount of 12% for some years.

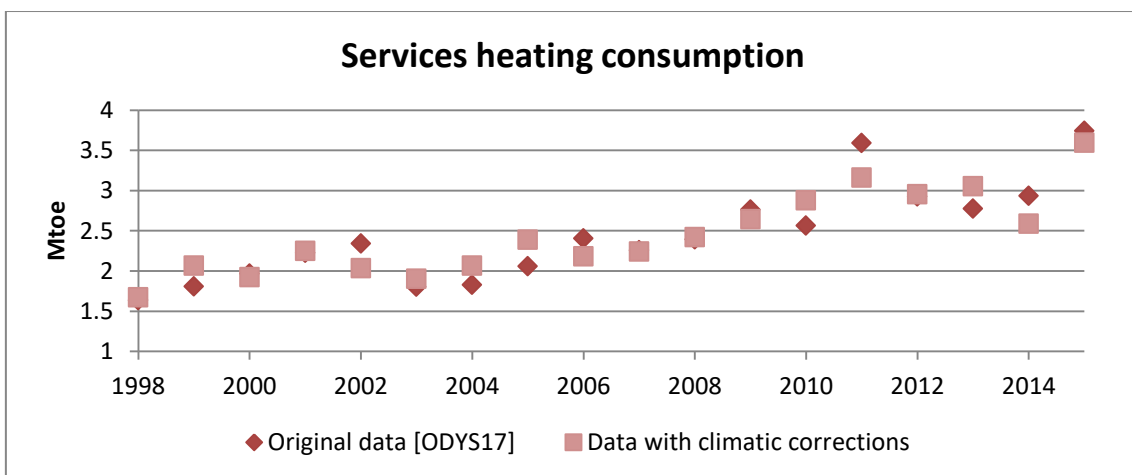


Figure 4 : Results for the corrected services heating consumption. Elaborated by the author.

Finally, no satisfactory model including any climate variable was built for residential or services air conditioning consumption.

Conclusions

Regarding hydropower generation, the hydraulic production capacity of Spain during atypical year is 26,004 GWh.

The corrected Spanish heating demand regarding climate can be obtained subtracting the climate variable determined in the regression model and multiplying I by its regression coefficient from the original series in mtoe. For the residential sector, the coefficient values 0.2164 and the variable corresponds to the normalized HDD and for the services sector, the coefficient is 0.2315 and the variable is the normalized average temperature in winter.

The robustness of the corrected residential heating consumption series is considered higher than the one of the corrected services heating consumption because the HDD are seen as a more reliable variable to quantify the effect of the outside temperature in heating consumption. Moreover, for most of the years, the errors of the two corrected series are opposite when compared with the original data. Therefore, the relationship between the outside temperature and the services heating consumption could be inexistent. If it was the case, Spain could be saving energy for heating purposes, which may mean losing energy efficiency.

Regarding the air conditioning consumption, the lack of multilinear regression models including a climatic variable for both sectors can be a consequence of the growing state of the air conditioning market.

References

- [AEME18] Spanish Meteorology Association (AEMET), 2018.
- [BPCEud] BP Chair of Energy and Sustainability web page, undated.
- [CARO13] Caro R., J. Reneses, “Estadística. No soy mala... me han dibujado así”. Thomson Reuters, Pamplona 2013.
- [INE_18] Spanish National Statistics Institute (INE), 2018.
- [LINA18] Linares P., Comment on HDD and CDD calculation method, 2018.
- [ODYS17] ODYSSEE, “ODYSSEE database – july 2017”, 2017.
- [REE_ud] Spanish National Electricity Network (REE), undated.
- [WORLud] World Bank, “GDP (MU in constant prices)”, undated.

Contenido

1.	Introducción y planteamiento del proyecto.	35
1.1.	La Cátedra BP de Energía y Sostenibilidad.	35
1.2.	Observatorios de Energía y Desarrollo Sostenible	35
1.2.1.	Indicadores energéticos incluidos en los Observatorios	35
1.2.2.	Análisis de los indicadores.	36
1.3.	Descripción del problema.	37
1.4.	Planteamiento del proyecto.	37
2.	Metodología	39
2.1.	Análisis de variables aleatorias independientes	39
2.1.1.	Obtención de posibles distribuciones de probabilidad	39
2.1.2.	Elección de la distribución más apropiada	39
2.1.3.	Obtención del valor típico	40
2.2.	Análisis de variables aleatorias dependientes	40
2.2.1.	Elección de variables	41
2.2.2.	Estimación de los coeficientes.	41
2.2.3.	Evaluación del modelo de regresión lineal.	41
2.2.4.	Elección del modelo más apropiado	43
2.2.5.	Obtención de la serie temporal corregida	43
3.	Estudio de las variables meteorológicas: producible hidráulico.	44
3.1.	Producible hidráulico español.	44
3.1.1.	Estimación de parámetros de las funciones de probabilidad.	45
3.1.2.	Bondad de ajuste: pruebas no paramétricas	51
3.1.3.	Conclusiones	54
4.	Análisis de variables aleatorias dependientes: consumo energético de uso térmico	55
4.1.	Variables del modelo lineal	55
4.1.1.	Variables dependientes	55
4.1.2.	Variables explicativas	57
4.2.	Consumo de calefacción del sector residencial	67
4.2.1.	Modelos de regresión lineal de calefacción del sector residencial	67
4.2.2.	Resultados del modelo de regresión para el consumo de calefacción en el sector residencial.	77

4.3.	Consumo de calefacción del sector servicios	79
4.3.1.	Modelos de regresión del consumo de calefacción del sector servicios ..	79
4.3.2.	Resultados de los modelos de consumo de calefacción del sector servicios	89
4.4.	Consumo de aire acondicionado del sector residencial	90
4.4.1.	Modelos del consumo de aire acondicionado del sector residencial	90
4.4.2.	Resultados de los modelos de aire acondicionado del sector residencial .	96
4.5.	Consumo de aire acondicionado del sector servicios	97
4.5.1.	Modelos del consumo de aire acondicionado del sector servicios	97
4.5.2.	Resultados de los modelos de aire acondicionado del sector servicios ..	102
4.6.	Datos corregidos según las condiciones climáticas de ODYSSEE	103
4.6.1.	Modelos del consumo de calefacción corregido del sector residencial ..	103
4.6.2.	Resultados del consumo de calefacción corregido del sector residencial	108
4.6.3.	Modelos del consumo de calefacción corregido del sector servicios	109
4.6.4.	Resultados del consumo de calefacción corregido del sector servicios ..	113
4.7.	Conclusiones sobre la demanda térmica en España	113
4.7.1.	Conclusiones sobre el consumo de calefacción.....	113
4.7.2.	Conclusiones sobre el uso de aire acondicionado	114
5.	Resultados.....	115
6.	Limitaciones y futuras consideraciones	116
7.	Conclusiones	117
8.	Bibliografía	118
9.	Anexos.....	121
9.1.	Anexo 1: Lista de indicadores empleados en los Observatorios.	121
9.2.	Anexo 2: Fuentes de información utilizadas en la elaboración de indicadores.	122

1. Introducción y planteamiento del proyecto.

1.1. La Cátedra BP de Energía y Sostenibilidad.

La Cátedra BP de Energía y Sostenibilidad de la Universidad Pontificia de Comillas tiene como finalidad “promover estudios, acciones formativas y divulgación” sobre la consecución de un modelo de desarrollo sostenible en España [CBPEsf] (1). Fue creada en diciembre de 2002 tras la firma de un “Acuerdo Marco de Colaboración” entre la Universidad y BP España [CBPEsf] (1).

La Cátedra posee ciertas herramientas con el fin de divulgar información, como las publicaciones de seminarios y tesis, así como del DES y del GREDS, aunque la más importante son los Observatorios de Energía y Desarrollo Sostenible. Éstos últimos además tratan de estimular un debate sobre las consecuencias del desarrollo sostenible para la energía [CBPEsf] (2).

1.2. Observatorios de Energía y Desarrollo Sostenible

Los Observatorios son informes que la Cátedra realiza con el fin de “divulgar y crear conciencia social en temas de energía y desarrollo sostenible en la sociedad española, estimulando un necesario debate en nuestro país sobre las implicaciones de la energía sobre la sostenibilidad” [CBPEsf] (3).

Estos informes se publican tanto con carácter anual como con carácter excepcional, y analizan la evolución energética durante un determinado periodo de tiempo (generalmente un año) para determinar si ha sido beneficiosa o no desde la perspectiva del desarrollo sostenible. La Cátedra sólo realiza informes sobre la situación española, comparándola con la de otros países con parecida situación económica [CBPEsf] (3).

En los Observatorios se resumen para cada año los eventos más importantes en el ámbito de la energía y la sostenibilidad, como regulaciones o acuerdos tanto a nivel nacional como internacional. También incluyen algunas dificultades presentes en España en materia de energía y sostenibilidad, como puede ser la pobreza energética [CBPE17] (1), así como posibles soluciones, con objeto de crear discusiones en torno a los problemas de España.

En los Observatorios también se incluye la elaboración y el análisis de múltiples indicadores energéticos contruidos a partir de diversas fuentes de información. La evolución del conjunto de indicadores representa el progreso de la situación energética española. Estos datos son analizados y valorados por expertos externos así como por colaboradores de la Cátedra [CBPEsf] (3).

1.2.1. Indicadores energéticos incluidos en los Observatorios

Los indicadores energéticos utilizados en los Observatorios pueden ser de carácter nacional o sectorial, cubriendo por ejemplo las áreas de producción eléctrica, transporte o servicios. Pueden clasificarse en las siguientes categorías [CBPE17] (1):

- 1) Indicadores sobre la energía primaria: en su mayoría comparan la situación en España con la de la UE-15 y alguna otra, mostrando el progreso a lo largo de los últimos años.
- 2) Indicadores sobre emisiones: estos indicadores dan cuenta del progreso anual de emisiones de CO₂ y de otros gases de efecto invernadero, comparando la situación en España, en la UE-15, en la OCDE y en el resto del mundo.
- 3) Indicadores sobre precios: en este caso se observa el progreso en los últimos años de varios precios, como es el caso de los carburantes, la electricidad o el de emisiones de CO₂ entre otros.
- 4) Indicadores sobre la movilidad de mercancías o viajeros en España:
- 5) Diagramas de Sankey: éstos representan gráficamente la relación entre algunos indicadores, así como los diversos flujos presentes en el año de estudio en la actividad económica española, diferenciando entre los tipos de orígenes y usos finales, así como si éstos se sitúan dentro o fuera de nuestras fronteras.

Todos los indicadores se elaboran a partir de datos procedentes de múltiples fuentes tanto de carácter nacional, como puede ser el Instituto Nacional de Estadística, como de carácter europeo o mundial, tales como Eurostat o el Banco Mundial (ver *Anexo 2: Fuentes de información utilizadas en la elaboración de indicadores*).

1.2.2. Análisis de los indicadores.

La mayoría de los indicadores se presenta en forma de gráfico para facilitar que se observe su evolución a lo largo de los últimos años. Se discute el progreso del año en cuestión en comparación al año anterior y a la tendencia del indicador.

Determinadas reformas en regulaciones o leyes pueden llevar a cambios en la situación energética, y por lo tanto, en los valores de los indicadores. Por ejemplo, los objetivos de desarrollo sostenible de la Unión Europea influyen en la instalación y operación de energías renovables [EUCO14]. Así mismo, las distintas directivas de la UE influyen en las políticas energéticas locales, como la directiva en eficiencia energética, basada en el paquete de energía y sostenibilidad para 2020 [PCUE12].

Por otro lado, las presiones en el ámbito político pueden influir en determinadas variables, como el precio del petróleo o el gas. El precio del gas en España determina a su vez el coste de la electricidad, y por tanto la utilización de recursos de gas natural para la producción eléctrica.

Así mismo, los indicadores utilizados en los Observatorios también pueden experimentar variabilidad aleatoria. Ésta se debe a fluctuaciones en variables ajenas y externas a todo lo anterior, tales como las condiciones meteorológicas.

Por lo tanto, la variación de los indicadores en el año en cuestión respecto a los anteriores puede deberse a factores legales, políticos, económicos, mejoras en la tecnología o a hechos puramente aleatorios.

1.3. Descripción del problema.

La existencia de la variabilidad aleatoria limita el alcance del análisis realizado en los Observatorios.

La producción de energías renovables, y con ella la importación de combustibles fósiles, depende en cierta medida de las condiciones meteorológicas: la radiación solar para las centrales tanto fotovoltaicas como solares térmicas, el perfil de velocidades del viento para la producción eólica y el agua disponible en las presas, procedente principalmente de las precipitaciones, para las centrales hidroeléctricas.

De la misma forma, el perfil de temperaturas unido a otros factores, como el aislamiento térmico, la eficiencia en los aparatos o la población, influyen en el consumo de energía de uso térmico tanto en invierno como en verano.

Por lo tanto, la presencia de aleatoriedad en los datos supone que no exista una relación directa entre las mejoras en los indicadores de un determinado año y el progreso en la situación energética española hacia la sostenibilidad.

1.4. Planteamiento del proyecto.

El objetivo del presente Trabajo de Fin de Master es eliminar la componente aleatoria de las variaciones de algunos indicadores empleados en la elaboración de los observatorios de la Cátedra BP de Energía y Sostenibilidad. La variabilidad aleatoria, siempre que no represente un evento singular, puede determinarse dentro de un margen de error establecido para su posterior eliminación.

Para juzgar el progreso de utilización de energías renovables pueden obtenerse los valores típicos de tales condiciones en España y deducir cuánta energía se habría producido y cuántos combustibles fósiles se habrían importado en esas condiciones. En el presente proyecto se ha decidido estudiar la energía hidráulica producible en un año corriente en España, cuyo producible hidráulico depende esencialmente de las precipitaciones.

Por otro lado, pueden extraerse conclusiones sobre el consumo energético y la eficiencia energética en España. Para este proyecto se persigue determinar el consumo energético de uso térmico, tanto en invierno como en verano, de los sectores residencial y servicios. En este caso, la componente aleatoria dependerá fundamentalmente de la temperatura exterior en las distintas localidades en invierno y en verano.

La metodología a emplear incluye principios de inferencia estadística y tratamiento de datos para obtener las series corregidas, y posteriormente, el modelo de simulación del sector energético español para evaluar el impacto dichas series en los indicadores. Cambios en la energía hidráulica consumida así como en el consumo de energía de uso térmico modificaría las importaciones de combustibles fósiles y con ello el consumo total de energía primaria en España. Esto podría traducirse en diferencias más o menos significativas en los siguientes indicadores:

- *Consumo de energía primaria per cápita.*
- *Intensidad energética primaria.* Se calcula mediante el cociente entre la energía primaria utilizada y el PIB y representa la energía primaria necesaria para producir una unidad monetaria.
- *Energía primaria por PIB: Intensidad energética.* De manera análoga al anterior, muestra el valor medio de energía en su forma final de consumo necesario para producir una unidad de riqueza.
- *Consumo de energía primaria en España.* Este indicador pone de manifiesto el total de energía primaria consumido en España y su repartición entre las distintas fuentes de energía.
- *Composición de energías renovables en la energía primaria.* Es un gráfico circular que representa, en porcentaje, las fuentes de energía renovables que han contribuido a la generación de energía renovable en España en el último año.
- *Emisiones de CO₂ per cápita.*
- *Emisiones de CO₂ por PIB: Intensidad de emisiones.* Representan, de manera análoga a la intensidad energética, las emisiones necesarias para producir una unidad de riqueza.
- *Precios de la electricidad.* Este indicador compara la evolución del precio de un kilowatio hora sin impuestos para España y la UE tanto en el sector doméstico como en el industrial.
- *Diagrama de fuentes de energía, usos y transformaciones.* En este diagrama se representan las distintas fuentes de energía y las transformaciones que sufren hasta llegar a su uso final por sectores.
- *Diagrama de origen de las emisiones de CO₂.* En él se representan las distintas fuentes de energía emisoras de CO₂, nacionales e importadas, y su evolución hacia pérdidas o emisiones de los distintos sectores.
- *Diagrama de flujos económicos.* Este diagrama muestra importaciones de energía primaria y su transformación en usos finales, pérdidas, impuestos o exportaciones.
- *Diagrama de externalidades.* En este caso se muestran algunas fuentes de energía y su contribución neta a cada sector más las externalidades que generan, tales como emisiones de otros gases contaminantes.
- *Diagrama de balance exegético.* En él se distinguen las fuentes de energía nacionales e importadas y su transformación en pérdidas o trabajo útil para los distintos sectores.

2. Metodología

Para eliminar la componente aleatoria de los indicadores analizados en los Observatorios de Energía y Sostenibilidad de la Cátedra pueden emplearse diferentes métodos de inferencia estadística y análisis de series temporales.

Concretamente, los experimentos aleatorios con que se ha trabajado son de dos tipos: independientes de otros eventos aleatorios, como las condiciones meteorológicas, y dependientes, como los valores en los consumos energéticos. Para trabajar con experimentos aleatorios es necesario conocer los principios de la inferencia estadística y de los contrastes de hipótesis.

2.1. Análisis de variables aleatorias independientes

En el análisis de variables aleatorias independientes, el objetivo es obtener el valor típico de dicha variable aleatoria para un año típico en España. Para ello, mediante inferencia estadística se determinan posibles distribuciones de probabilidad y parámetros de la variable aleatoria. Más adelante, se evalúan y comparan los resultados obtenidos mediante contrastes de hipótesis. Finalmente, se obtiene el valor típico con la distribución de probabilidad más plausible.

2.1.1. Obtención de posibles distribuciones de probabilidad

La inferencia estadística se ha utilizado para determinar varias funciones de probabilidad verosímiles para el conjunto de datos de que se dispone.

Para poder obtener las distribuciones de probabilidad se ha supuesto que las muestras pertenecen a familias de distribución de probabilidad conocidas. Para este proyecto se decidió utilizar las siguientes distribuciones: normal, log-normal, Weibull y gamma, por su presencia en la naturaleza y por su uso para caracterizar fenómenos meteorológicos.

Entonces, a partir de las muestras aleatorias y de la familia de distribución de probabilidad, se pueden determinar los valores de los parámetros que definen tales distribuciones (estadísticos) mediante el método de máxima verosimilitud.

El método de máxima verosimilitud se basa en la función de probabilidad de la población, permitiendo obtener parámetros que no están relacionados con ningún momento muestral. Este método permite escoger como estimador de un parámetro el estadístico que haga más probable obtener la muestra en concreto [CARO13].

La inferencia estadística no es infalible, y posee ciertas limitaciones: los estimadores obtenidos de una muestra sólo son una aproximación de los valores poblacionales, ya que dependen de la muestra en cuestión [CARO13].

2.1.2. Elección de la distribución más apropiada

Los contrastes de hipótesis o test de hipótesis se han empleado con el fin de juzgar si los estimadores y las poblaciones cumplen con las características obtenidas de la inferencia estadística.

Algunos test que favorecen la comparación de la distribución de probabilidad muestral a otras conocidas son los siguientes: el test Chi-cuadrado bondad de ajuste, el test de Kolmogorov-Smirnov y el estadístico de Anderson Darling. En todos ellos, la hipótesis nula será la igualdad entre la función de distribución de probabilidad empírica y las distribuciones normal, log-normal, Weibull y gamma obtenidas.

El criterio de elección de la distribución más apropiada será aquella que no haya sido rechazada por ningún contraste de hipótesis y cuyo p-valor asociado para todos ellos sea el más elevado.

Una limitación de los contrastes de hipótesis es que si la hipótesis nula no puede rechazarse, estadísticamente no se ha demostrado nada [HURCSf].

2.1.3. Obtención del valor típico

Una vez que se ha determinado la distribución de probabilidad y sus parámetros, los valores típicos pueden obtenerse a través del cálculo de la esperanza o valor esperado. La esperanza se obtiene por medio de la integral:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_x(x) \cdot dx \quad (1)$$

Donde x representa la variable aleatoria en cuestión, $E(X)$ su esperanza y $f_x(x)$ la función de distribución de probabilidad de la variable aleatoria x . Puede observarse que, para cada distribución de probabilidad, el cálculo de la esperanza es distinto.

2.2. Análisis de variables aleatorias dependientes

En el análisis de variables aleatorias dependientes, el objetivo es realizar predicciones sobre dichas variables cuando sólo cambia el factor meteorológico. Para ello, pueden utilizarse modelos de regresión. El proceso a seguir en este caso pasa por construir modelos de regresión lineal que relacionen la variable dependiente con las explicativas, determinar el poder explicativo y la validez de cada uno, escoger uno para que represente el comportamiento y finalmente obtener los datos de la variable dependiente sin tener en cuenta la variable climática.

Puesto que en este estudio hay varias variables explicativas, se han utilizado modelos de regresión lineal múltiple, cuya estructura se representa en la ecuación:

$$Y = \beta_0 + \beta_1 \cdot X_1 + \dots + \beta_K \cdot X_K + E \quad (2)$$

Donde Y es la variable aleatoria que se desea explicar, X_i representa cada una de las variables explicativas, K es el número total de variables, β_i representa la contribución marginal de cada variable explicativa y E representa un término de error aleatorio.

Para poder establecer los modelos de regresión lineal se han tomado las siguientes hipótesis [CARO13]:

- 1) La relación entre las variables explicativas y la dependiente es lineal.

- 2) El término del error representa otros factores no expresados por ninguna de las variables explicativas.
- 3) El término del error es una variable aleatoria que sigue una distribución normal de media cero y varianza constante.
- 4) Las variables deterministas son independientes entre sí.
- 5) Las observaciones de la variable dependiente han de ser independientes entre sí.

Sin embargo, los modelos de regresión lineal pueden tener limitaciones. Si existe una fuerte correlación entre las variables explicativas (éstas deberían ser independientes), el modelo puede resultar arbitrario. Así mismo, cuando el tamaño muestral no es lo suficientemente grande, es difícil obtener p valores individuales lo suficientemente bajos.

2.2.1. Elección de variables

En primer lugar, se han de elegir las variables más apropiadas para cuantificar los efectos climáticos, tecnológicos, económicos y demográficos.

2.2.2. Estimación de los coeficientes.

El ajuste del modelo se realiza utilizando el método de máxima verosimilitud, cuya solución coincide con el método de mínimos cuadrados para el caso de error con distribución normal [CARO13]. En este método se minimiza el error del modelo en función de la estimación de los parámetros.

El resultado es conocido para el caso del modelo de regresión múltiple, y se representa en la ecuación $\hat{\beta} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot Y$ (3), donde $\hat{\beta}$ es un vector columna de dimensión $K+1$ que contiene las estimaciones de los parámetros $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_K$ y \mathbf{X} es la llamada matriz de diseño, representada en la ecuación $\mathbf{X} =$

$$\begin{pmatrix} 1 & x_{11} & \dots & x_{k1} \\ 1 & x_{12} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots \\ 1 & x_{1n} & \dots & x_{kn} \end{pmatrix} \quad (4).$$

$$\hat{\beta} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot Y \quad (3)$$

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \dots & x_{k1} \\ 1 & x_{12} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots \\ 1 & x_{1n} & \dots & x_{kn} \end{pmatrix} \quad (4)$$

2.2.3. Evaluación del modelo de regresión lineal.

Para comprobar el poder explicativo del modelo, pueden emplearse diferentes contrastes de hipótesis:

- Contraste de regresión o F-test del modelo: este contraste se utiliza para mostrar que la mayor parte de la variabilidad de Y es explicada por la de X. Las hipótesis pueden enumerarse del siguiente modo:

H₀ : Todos los coeficientes del modelo de regresión lineal múltiple son igual a cero.

H₁ : Algún coeficiente es distinto de cero.

Para realizar este contraste se utiliza el estadístico F, definido en la ecuación

$$F = \frac{\sum(y_i - \hat{y}_i)^2 / K}{\sum(e_i)^2 / n - K - 1} \quad (5), \text{ el cual se distribuye según}$$

una F de Snedecor con K y n-K-1 grados de libertad en caso de ser H₀ cierta. K representa el número de variables, n el número de observaciones, y_i los datos de la variable dependiente, \hat{y}_i los estimadores de la variable dependiente según el modelo, y finalmente e_i los errores del modelo en cada observación.

$$F = \frac{\sum(y_i - \hat{y}_i)^2 / K}{\sum(e_i)^2 / n - K - 1} \quad (5)$$

Si el p-valor de este test es menor que el nivel de significación escogido, entonces el modelo es explicativo.

- Contrastes de hipótesis sobre los parámetros o t-test de los parámetros: en este contraste de hipótesis se plantea el poder explicativo de una variable en concreto y no del modelo. Sus hipótesis se enumeran a continuación:

H₀: El valor del coeficiente de una determinada variable en particular es cero.

H₁: El valor de dicho coeficiente es distinto de cero.

Para este contraste de hipótesis se define el estadístico t en la ecuación $t =$

$$\frac{\hat{\beta}_i - \beta_i}{S_{RES} \cdot \sqrt{q_{ii}}} \quad (6), \text{ el cual sigue una distribución t de}$$

Student con n-p grados de libertad si el error sigue una distribución normal.

$$t = \frac{\hat{\beta}_i - \beta_i}{S_{RES} \cdot \sqrt{q_{ii}}} \quad (6)$$

Si el p-valor de dicho test es inferior al nivel de significación escogido, entonces el coeficiente es distinto de cero y la variable es significativa.

Con los contrastes de hipótesis anteriores puede detectarse la multicolinealidad puede detectarse si el modelo resulta explicativo (p-valor del F-test bajo) pero los t-test individuales de cada parámetro dan altos p-valores.

Además de los contrastes de hipótesis anteriores, se dispone de las siguientes herramientas para medir la correlación entre la variable dependiente y las independientes:

- El coeficiente de correlación R^2 mide cómo de buena es la relación entre las variables explicativas y la que se desea explicar. Resulta del cociente entre la desviación cuadrática del modelo de la variable aleatoria con sus valores reales entre la varianza de la variable aleatoria.

$$R^2 = 1 - \frac{s_{res}^2}{s_y^2} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} \quad (7)$$

Cuanto mayor sea el valor de R^2 , mayor será la correlación entre la variable que se desea explicar y las variables explicativas.

- El coeficiente $\overline{R^2}$ (también llamado R^2 ajustado) mide la correlación penalizando un mayor número de variables que explican el modelo. Se calcula mediante la siguiente fórmula:

$$\overline{R^2} = 1 - (1 - R^2) \cdot \frac{n-1}{n-p} \quad (8)$$

Cuanto mayor es el valor de $\overline{R^2}$, mayor es la correlación entre la variable que se desea explicar y las variables explicativas.

Se ha de comprobar además que los residuos siguen una distribución normal de media cero, varianza constante y que son independientes entre sí.

2.2.4. Elección del modelo más apropiado

Las características del modelo más apropiado se enumeran a continuación:

- Se han considerado modelos aptos aquellos que pueden aceptarse a un 10 % de significación en los dos contrastes de hipótesis, siendo preferible aquellos aceptables a un 5%. El motivo es el tamaño muestral, el cual al no ser muy grande puede dar valores de p-valor altos.
- Se considera que existe correlación cuando los coeficientes R^2 y $\overline{R^2}$ son mayores que 0.6 y próximos entre sí, aunque valores en torno a 0.8 o mayores son más apropiados.

En caso de existir varios modelos similares, premiará aquel con mayor correlación frente a un mayor nivel de significación.

2.2.5. Obtención de la serie temporal corregida

Una vez se ha escogido el modelo de regresión más apropiado, puede pasarse a calcular el valor corregido según la componente climática del modelo de regresión. El cálculo se realiza simplemente sustrayendo dicha componente de la variable dependiente como sugiere la ecuación $Y^* = Y - \hat{\beta}_1 \cdot X_1$ (9), donde Y^* representa la componente corregida, Y los datos originales, $\hat{\beta}_1$ el estimador del coeficiente de la contribución marginal de las condiciones climáticas y finalmente X_1 el valor de dichas condiciones.

$$Y^* = Y - \hat{\beta}_1 \cdot X_1 \quad (9)$$

3. Estudio de las variables meteorológicas: producible hidráulico.

Para determinar la producción hidroeléctrica en España durante un año típico, la variable aleatoria que se ha decidido analizar es el producible hidráulico. Según Red Eléctrica de España (en [REE_sf]), el producible hidráulico es la “cantidad máxima de energía eléctrica que teóricamente se podría producir [...] durante un determinado período de tiempo y una vez deducidas las detracciones de agua realizadas para riego o para otros usos distintos de la producción de energía eléctrica”.

El producible hidráulico está muy relacionado con la lluvia, que es un suceso meteorológico, y por ende aleatorio. Se trata por lo tanto de una variable aleatoria continua con espacio muestral positivo. El estudio de esta variable se detalla en el siguiente apartado.

3.1. Producible hidráulico español.

El objetivo de esta sección es obtener el valor de producible hidráulico en España para un año típico. Para ello, en este apartado se ha estudiado la distribución de probabilidad del producible hidráulico español. La muestra de que se dispone [REE_17] recoge datos desde 1991 hasta 2017, y se hallan representados en el Gráfico 1.

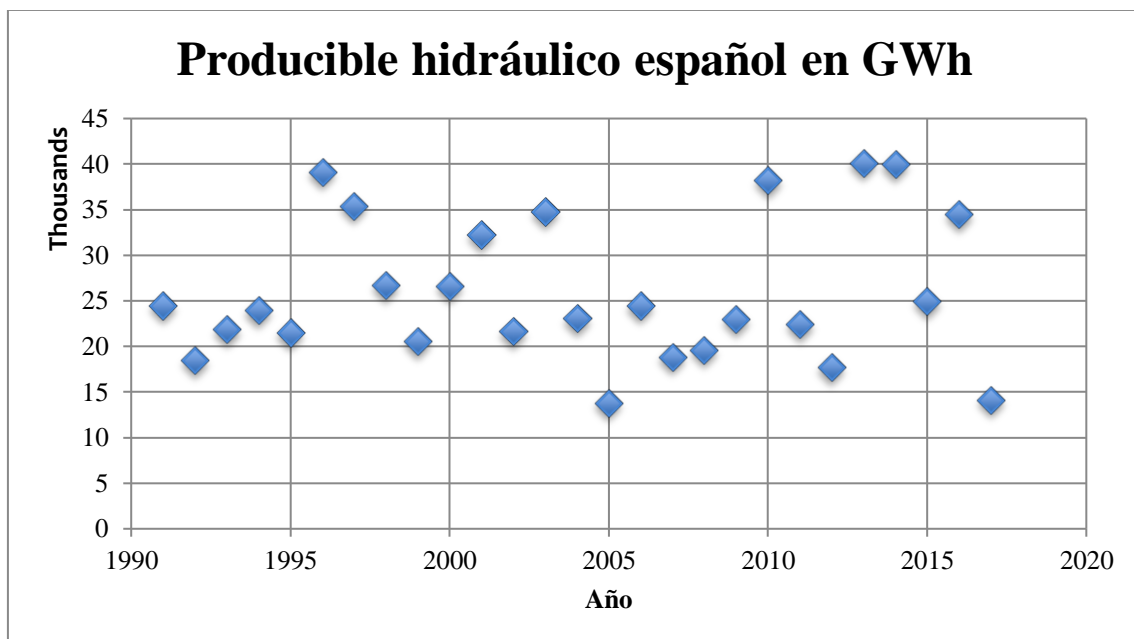


Gráfico 1: Producible hidráulico español en GWh entre los años 1991 y 2017. Elaboración propia a partir de datos de [REE_17]

El Gráfico 2 representa la función de distribución empírica del total de producible hidráulico español. En el siguiente apartado se tratará de encontrar funciones de probabilidad plausibles con el fin de caracterizar el producible hidráulico español.

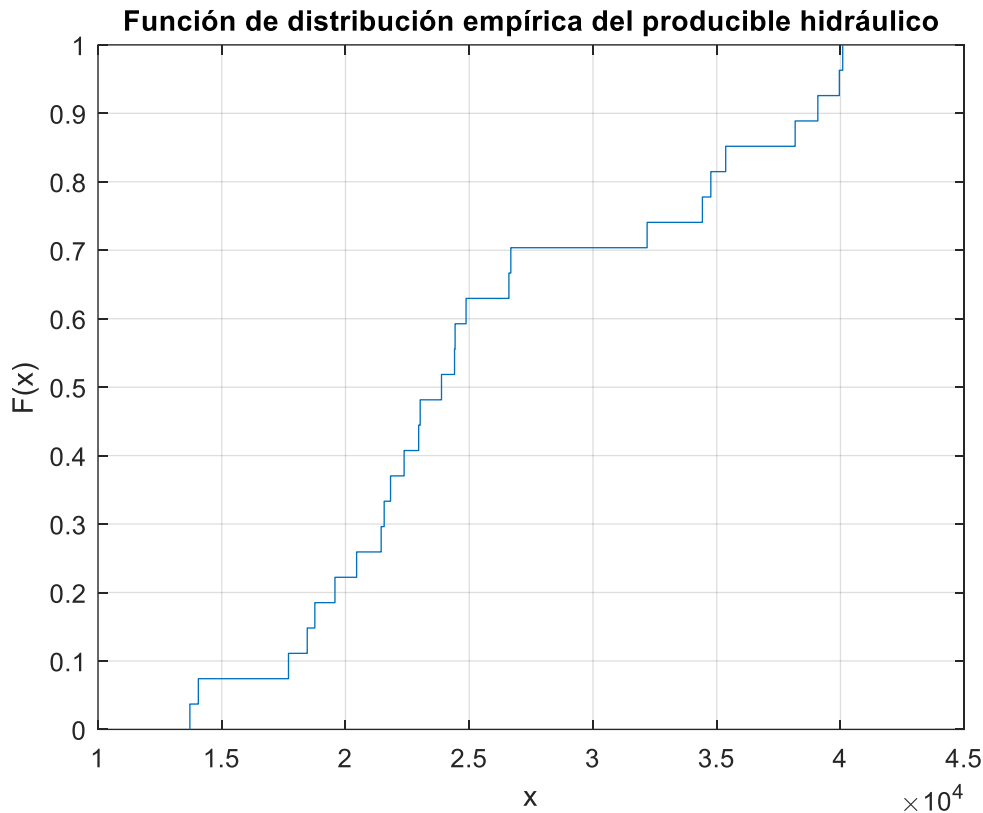


Gráfico 2 : Distribución de probabilidad del producible hidráulico (en GWh). Elaboración propia utilizando Matlab.

3.1.1. Estimación de parámetros de las funciones de probabilidad.

Para determinar posibles funciones de probabilidad se han elegido las cuatro familias de distribuciones de probabilidad que se han considerado más verosímiles para representar el producible hidráulico: la distribución normal, la log-normal, la de Weibull y la gamma.

Los resultados numéricos de las estimaciones presentes en los siguientes apartados se han obtenido utilizando la función *fitdist* de MATLAB, el cual emplea el algoritmo de máxima verosimilitud específico para cada familia de distribución de probabilidad excepto en dos casos [MATHsf]:

- En la distribución normal con datos sin censura, la estimación de la desviación típica se obtiene empleando la raíz cuadrada del estimador de la varianza insesgado.
- En la distribución log-normal con datos sin censura, la estimación del parámetro de forma sigma cuadrado es la raíz cuadrada del estimador insesgado de la variación del registro de datos.

El Gráfico 3 representa el histograma de la muestra con las curvas de densidad de probabilidad anteriores superpuestas. La muestra es asimétrica a la derecha (según la definición encontrada en [CARO13]),

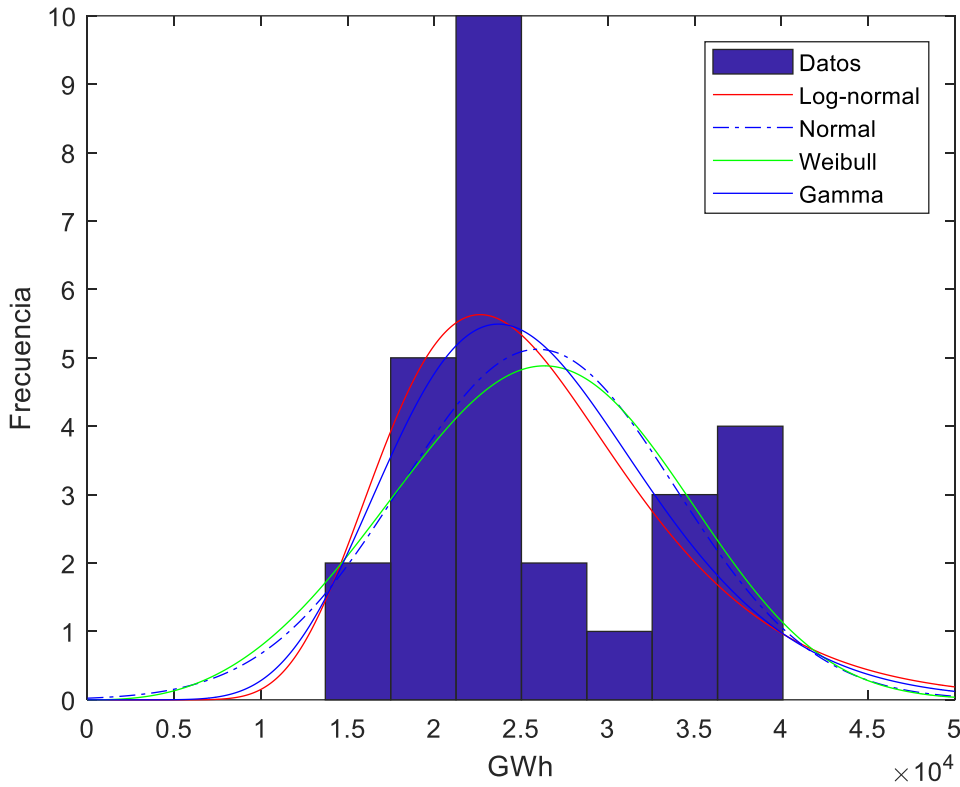


Gráfico 3: Distintas funciones de probabilidad superpuestas con el histograma de los datos. Fuente: elaboración propia utilizando Matlab.

Puede observarse que todas las curvas tienen formas parecidas y que tratan de ajustarse al histograma de los datos. La distribución log-normal también presenta una asimetría hacia la derecha, al igual que la gamma, mientras que la distribución normal es simétrica y la de Weibull presenta la asimetría hacia el otro lado.

3.1.1.1. Distribución normal

La distribución normal se escogió por ser la distribución de probabilidad más importante y por su amplia presencia en la naturaleza. Bajo esta distribución de probabilidad, los datos quedan definidos por su media y su desviación estándar. Su función de densidad se escribe siguiendo la ecuación $X \sim N(\mu, \sigma) \rightarrow f(x; \mu, \sigma) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot$

$$e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2} \quad (10):$$

$$X \sim N(\mu, \sigma) \rightarrow f(x; \mu, \sigma) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2} \quad (10)$$

Los valores estimados para los parámetros se muestran en la Tabla 1. En la columna del P-valor se muestra el correspondiente al del contraste de hipótesis del test t-student para la media y del test Chi-cuadrado para la varianza.

Análisis de algunos indicadores de sostenibilidad en España

El test t-student se utiliza para evaluar la verosimilitud de un valor para la media de una distribución normal cuando la varianza es desconocida. El contraste de hipótesis se enumera del siguiente modo:

H_0 : La media de la distribución normal tiene valor μ_0 .

H_1 : La media de la distribución normal tiene otro valor distinto de μ_0 .

El test Chi-cuadrado es un contraste sobre la varianza de una distribución normal con media desconocida. Las hipótesis se enumeran a continuación:

H_0 : La varianza de la distribución normal tiene valor σ_0 .

H_1 : La varianza de la distribución normal tiene otro valor distinto de σ_0 .

	Estimación	Intervalo de confianza al 95%	P-valor
μ	25954.9	[22821.5 , 29088.3]	1
σ	7920.9	[6237.91 , 10855.2]	0.9262

Tabla 1: Resultados de los parámetros para la distribución normal. Fuente: elaboración propia.

De la Tabla 1 puede deducirse que las hipótesis nulas para los test t-student para la media y Chi-cuadrado para la varianza no pueden rechazarse, ya que el p-valor asociado a ambos es superior a los niveles de significación convencionales.

En el Gráfico 4 se muestra el ajuste de los datos de la muestra a una distribución normal. Puede observarse que los datos no se acercan en exceso a la línea de tendencia de la distribución normal.

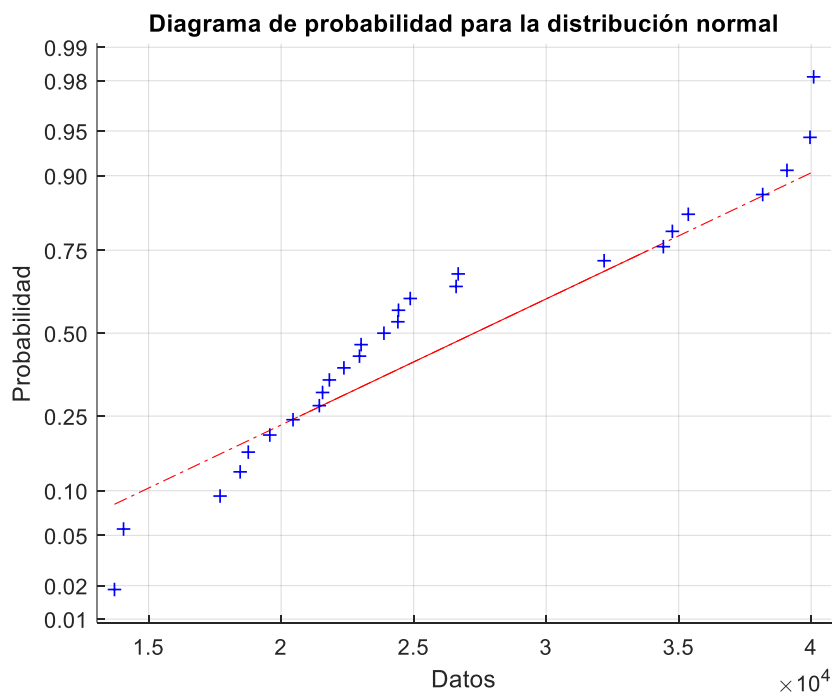


Gráfico 4: Ajuste de la muestra a una distribución normal. Elaboración propia utilizando Matlab.

3.1.1.2. Distribución log-normal

La distribución log-normal se incluyó por su relación con la distribución normal, y por el hecho de que su espacio muestral está compuesto por números positivos. Con este tipo de distribución, al igual que con la distribución normal, los datos quedan definidos por dos parámetros: μ y σ . La distribución tiende a la siguiente función densidad de probabilidad:

$$X \sim LN(\mu, \sigma) \rightarrow f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} \text{ para } x > 0 \quad (11)$$

Los resultados de la estimación de parámetros según el método de máxima verosimilitud se hayan en la Tabla 2.

	Estimación	Intervalo de confianza al 95%
μ	10.1198	[9.99948,10.24]
σ	0.30407	[0.23946,0.416708]

Tabla 2: Resultados de los parámetros de la log-normal. Fuente: elaboración propia.

El Gráfico 5 muestra el ajuste de los datos de la muestra a una distribución log-normal. Puede observarse, cualitativamente, que la muestra está más próxima de esta curva que de la curva de la distribución normal.

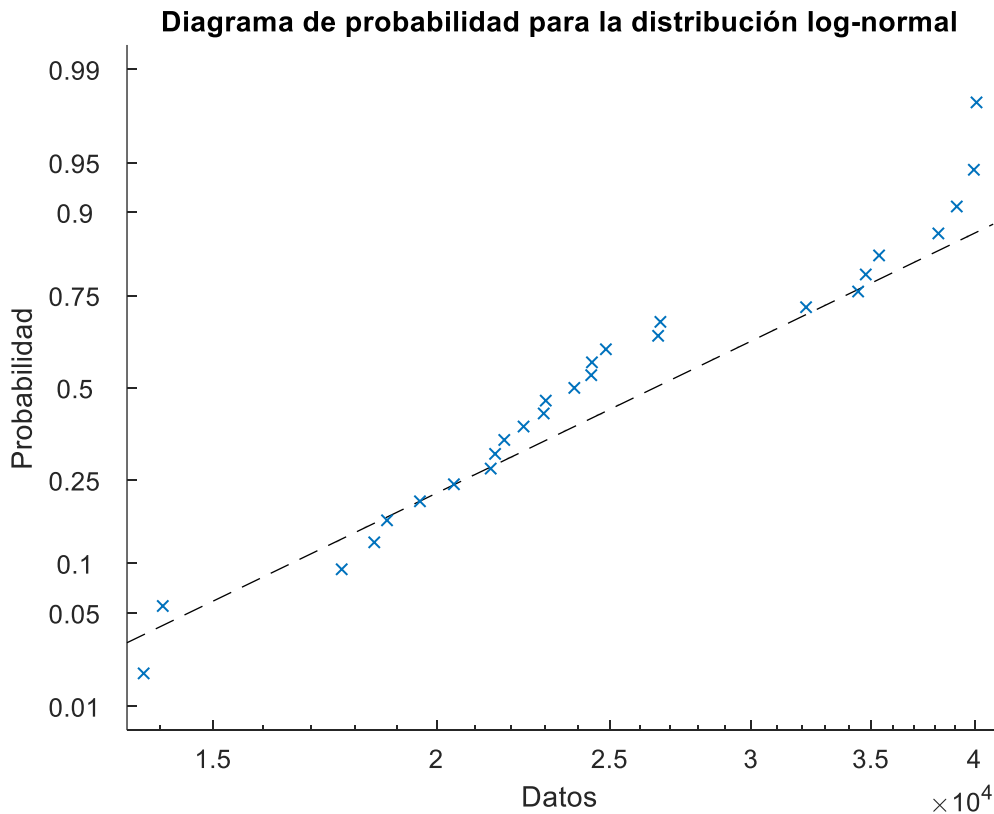


Gráfico 5: Ajuste de la muestra a una distribución log-normal. Elaboración propia utilizando Matlab.

3.1.1.3. Distribución Weibull

La distribución de Weibull se consideró adecuada debido a las múltiples formas que puede adoptar según los valores de sus parámetros. Además de ello, suele utilizarse para caracterizar fenómenos meteorológicos y su espacio muestral resulta positivo si se utiliza la variante estandarizada [CARO13].

La función de densidad de probabilidad de la distribución de Weibull se detalla en la

$$\text{ecuación } X \sim W(\lambda, k) \rightarrow f(x) = \frac{k}{x} \left(\frac{x}{\lambda}\right)^{k-1} \cdot e^{-\left(\frac{x}{\lambda}\right)^k} \text{ para } k, \lambda, x > 0 \quad (12):$$

$$X \sim W(\lambda, k) \rightarrow f(x) = \frac{k}{x} \left(\frac{x}{\lambda}\right)^{k-1} \cdot e^{-\left(\frac{x}{\lambda}\right)^k} \text{ para } k, \lambda, x > 0 \quad (12)$$

Donde k es un parámetro de forma y λ es un parámetro de escala. Utilizando un parámetro de escala de mayor valor se obtienen funciones más suaves, en cuanto al parámetro de forma, destacan los siguientes casos [CARO13]:

- Si su valor es mayor que 3 pero menor que 4, su forma se aproxima a la de la distribución normal.
- Para el valor unidad se corresponde con la distribución exponencial.
- Para valores entre medias, se obtiene una distribución unimodal asimétrica a la derecha.

En la Tabla 3 se recogen las estimaciones de los parámetros de esta distribución a partir de los datos de la muestra.

	Estimación	Intervalo de confianza al 95%
λ	28833.4	[25804.8,32217.4]
k	3.6037	[2.70155,4.80712]

Tabla 3: Parámetros de la distribución de Weibull. Fuente: elaboración propia.

Finalmente, el ajuste de los datos a este tipo de distribuciones puede observarse en el Gráfico 6, en el cual el ajuste es muy pobre en la cola izquierda.

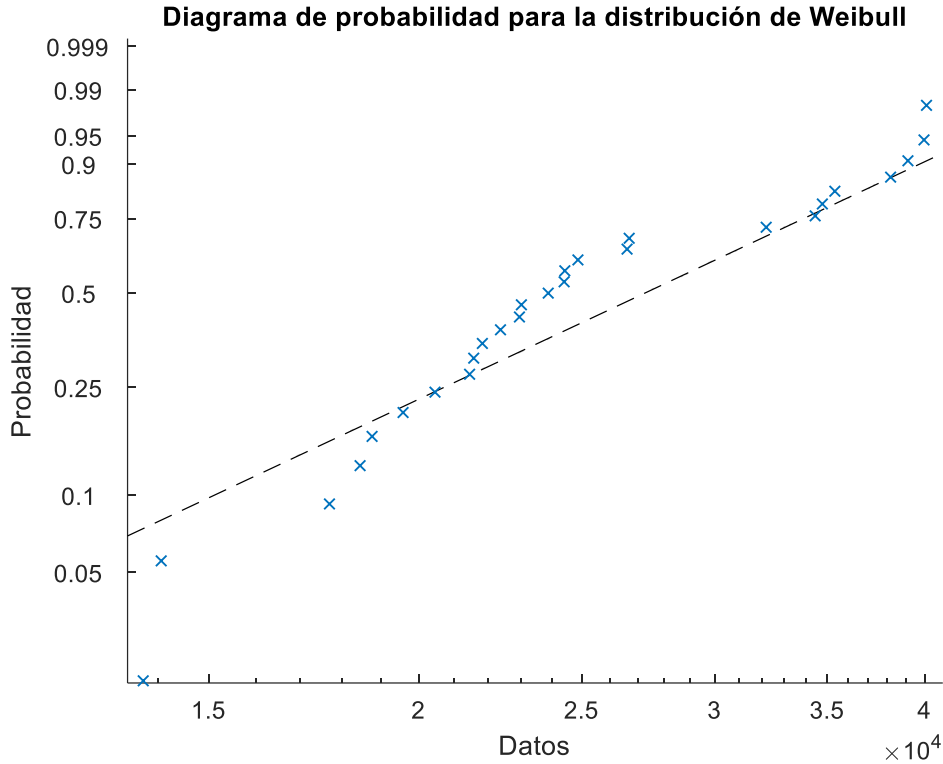


Gráfico 6 : Ajuste de la muestra a la distribución Weibull. Elaboración propia utilizando Matlab.

3.1.1.4. Distribución Gamma

La distribución Gamma también se escogió por su presencia en la naturaleza, ya que suele utilizarse para caracterizar fenómenos meteorológicos. Su función de densidad sigue la forma expresada a continuación [CARO13].

$$X \sim \Gamma(\alpha, \beta) \rightarrow f(x) = x^{\alpha-1} \cdot e^{-x/\beta} \cdot \frac{1}{\beta^\alpha \Gamma(\alpha)}; \quad x, \alpha, \beta > 0 \quad (13)$$

Donde β es un parámetro de escala y α es un parámetro de forma. La representación bidimensional de la distribución gamma puede parecerse a la distribución de Weibull para los distintos valores sus parámetros [CARO13]:

- Un factor de escala mayor proporciona una función más suave.
- Valores altos del parámetro de forma aproximan la distribución a una normal.
- Para valores de α mayores que 1 se tiene una distribución unimodal asimétrica a la derecha.
- En el caso $\alpha=1$ se tiene la distribución exponencial.

En la siguiente tabla se muestran los estimadores obtenidos para la distribución gamma:

	Estimación	Intervalo de confianza al 95%
α	11.4374	[6.76018,19.3507]
β	2269.3	[1325.69,3884.56]

Tabla 4: Estimación de los parámetros de la distribución Gamma. Fuente: elaboración propia.

3.1.2. Bondad de ajuste: pruebas no paramétricas

El objetivo de esta sección es determinar la función de distribución de probabilidad del apartado anterior más apropiada para representar el producible hidráulico español. Para ello, se han utilizado contrastes de hipótesis no paramétricos para determinar la bondad de ajuste de las distribuciones de probabilidad anteriores a los datos.

3.1.2.1. El test Chi-cuadrado bondad de ajuste.

Este test trata de comprobar si una determinada muestra sigue una distribución de probabilidad conocida. Las hipótesis propias del mismo se enumeran a continuación:

H₀: El conjunto de datos sigue una distribución de probabilidad concreta conocida.

H₁: El conjunto de datos no sigue esa distribución de probabilidad.

El conjunto de datos en cuestión ha de consistir en un número total de observaciones n agrupadas en k clases, y cada clase tendrá asociada una frecuencia observada ($O_1 \dots O_i \dots O_k$). El siguiente paso es decidir si estas frecuencias observadas están o no en concordancia con las frecuencias esperadas ($E_1 \dots E_i \dots E_k$), es decir, si el número de resultados observados coincide aproximadamente al número esperado según la distribución [MONGsf].

	Frecuencia observada	Frecuencia esperada
Clase 1	O_1	E_1
Clase i	O_i	E_i
Clase K	O_k	E_k
Total	n	N

Tabla 5: Características de la muestra para usar un contraste de hipótesis Chi-cuadrado de bondad de ajuste. Fuente: [MONGsf].

El contraste de hipótesis que se utiliza para comprobarlo utiliza la distribución Chi-cuadrado, cuyo estadístico de contraste es [MONGsf]:

$$\chi^{2*} = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (14)$$

En este contraste de hipótesis, cuanto menor sea el valor del estadístico, más coherentes son las observaciones obtenidas con los valores esperados. La hipótesis nula suele rechazarse cuando el estadístico es mayor que un determinado valor crítico [MONGsf].

Las condiciones para realizar este contraste de hipótesis son las siguientes [MONGsf]:

- El valor del estadístico se podrá aproximar por una distribución Chi cuadrado cuando el tamaño muestral n sea grande ($n > 30$) y todas las frecuencias esperadas sean iguales o mayores a cinco.
- Las observaciones han de ser obtenidas mediante muestreo aleatorio a partir de una población particionada en categorías.

3.1.2.2. *El test de Kolmogorov-Smirnov*

El test de Kolmogorov-Smirnov compara las funciones de distribución empírica de la muestra y la que se desea contrastar y decide si los datos muestrales provienen de dicha distribución [KYSB10].

H₀: El conjunto de datos sigue una distribución de probabilidad concreta conocida.

H₁: El conjunto de datos no sigue esa distribución de probabilidad.

El estadístico que se emplea en este test de hipótesis es el máximo de la distancia entre la función empírica de distribución de los datos ($F_e(x)$) y la de la hipótesis nula ($F(x)$) [KYSB10].

$$D = \max_x |F_e(x) - F(x)| \quad (15)$$

El valor obtenido de D ha de compararse con determinados valores establecidos de acuerdo al tamaño muestral y al nivel de significación. Si D es mayor que su correspondiente en la tabla, se ha de rechazar la hipótesis nula.

3.1.2.3. *El estadístico de Anderson Darling.*

El estadístico de Anderson Darling mide el grado de seguimiento de los datos a una distribución específica. Un valor más bajo de este estadístico significa un mejor ajuste de los datos a dicha distribución. Las hipótesis de este contraste de hipótesis son las siguientes [MINIsf]:

H₀: El conjunto de datos sigue la distribución especificada.

H₁: El conjunto de datos no sigue la distribución especificada.

El estadístico se basa, al igual que el de Kolmogorov-Smirnov en la distancia entre los datos y la verdadera curva de distribución. Sin embargo, en este caso no se toma el máximo, sino que se pondera cuadráticamente dando un mayor peso a las colas [MINIsf].

Este estadístico también puede utilizarse para comparar el ajuste de los datos a varias distribuciones y comprobar cuál es la mejor. Sin embargo, en este caso ha de evitarse comparar el valor del estadístico directamente, ya que se distribuyen de manera diferente según la distribución especificada. Lo que se recomienda es utilizar datos adicionales, tales como las gráficas de distribución o el p-valor del estadístico [MINIsf]. Valores p más bajos evidencian que los datos no siguen una cierta distribución.

3.1.2.4. *Método a utilizar.*

Puesto que el tamaño muestral no es suficiente para aplicar el test Chi-cuadrado de bondad de ajuste, se utilizará el contraste de hipótesis de Kolmogorov-Smirnov con el objetivo de probar si el conjunto de datos podría provenir de las distintas funciones de

probabilidad expresadas arriba. Esto permitirá descartar las muestras de manera individual.

De las muestras restantes, para elegir la más apropiada, se empleará el estadístico de Anderson-Darling y su p-valor asociado.

Los contrastes de hipótesis en cuestión se realizaron con el programa MATLAB. Los resultados pueden observarse en el siguiente apartado.

3.1.2.5. Resultados

La Tabla 7 muestra los resultados del test de Kolmogorov-Smirnov sobre el ajuste de las funciones optimizadas por Matlab para el producible hidráulico. En todos los casos no puede rechazarse la hipótesis nula, sin embargo, p-valores altos indican una mayor verosimilitud de una determinada distribución. Entonces, según el test de Kolmogorov-Smirnov, la distribución de probabilidad más adecuada es la distribución log-normal.

Distribución	Valores	Resultado ($\alpha=5\%$)	P-valor
Normal	N(25954'9, 7920'9)	No se rechaza la hipótesis nula	0,2838
Weibull	We(28833'4,3'6)	No se rechaza la hipótesis nula	0,2747
Log-normal	LogN(10'1198, 0'3041)	No se rechaza la hipótesis nula	0,7266
Gamma	$\Gamma(11'43, 2269'3)$	No se rechaza la hipótesis nula	0,5531

Tabla 6: Resultados del test de Kolmogorov-Smirnov para las distribuciones anteriores.

La Tabla 8 muestra los resultados del p-valor del estadístico de Anderson-Darling (excepto para el caso de la distribución gamma, donde no fue posible). De ellos puede deducirse que la única distribución estadística plausible para representar el producible hidráulico español es otra vez la log-normal.

Distribución	Valores	Resultado ($\alpha=5\%$)	P-valor
Normal	N(25954'9, 7920'9)	Se rechaza la hipótesis nula	0,0149
Weibull	We(28833'4,3'6)	Se rechaza la hipótesis nula	0,0164
Log-normal	LN(10'1198, 0'3041)	No se rechaza la hipótesis nula	0,1846

Tabla 7: Resultados del estadístico de Anderson-Darling para algunas de las distribuciones anteriores.

Una limitación es que el tamaño muestral no es lo suficientemente grande para aplicar el test de bondad de ajuste chi-cuadrado, y es posible que con una muestra de datos mayor, los resultados de los parámetros y puede que de la distribución podrían ser distintos. Además, al realizar contrastes de hipótesis, estadísticamente no se demuestra nada si no se rechaza la hipótesis nula.

3.1.3. Conclusiones

Según el contraste de hipótesis de Kolmogorov-Smirnov, todas las son adecuadas para representar el producible hidráulico anual español, pero el p-valor más alto en este test es el asociado a la distribución log-normal. Por otro lado, el p-valor del estadístico de Anderson-Darling corrobora los resultados del test de Kolmogorov-Smirnov, por lo que se deduce que la distribución más apropiada para la representación de los datos es la log normal descrita en el apartado 2.

$$X \sim LN(10,1198; 0,3041) \quad (16)$$

Por lo tanto, el valor típico de producible hidráulico en España obtenido corresponde a la esperanza de la distribución log-normal descrita en apartados anteriores. La esperanza de una distribución log-normal se calcula del siguiente modo:

$$X \sim LN(\mu, \sigma) \rightarrow E(X) = \int_0^{\infty} f_x(x) \cdot x \cdot dx = e^{\mu + \frac{1}{2} \cdot \sigma^2} \quad (17)$$

Donde la variable aleatoria x es el producible hidráulico, $f_x(x)$ es la distribución de probabilidad de la ecuación $X \sim LN(\mu, \sigma) \rightarrow f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}$ para $x > 0$ (11), y los parámetros μ y σ tienen los valores señalados en la Tabla 2.

El resultado del producible hidráulico típico en España es por lo tanto de 26.004,61 GWh.

Este resultado sería más robusto si se dispusiese de una muestra de mayor tamaño.

4. Análisis de variables aleatorias dependientes: consumo energético de uso térmico

El objetivo de esta parte es determinar cuál es el consumo energético de uso térmico en España para un año normal. El consumo energético se ha diseminado en cuatro partes: consumo de calefacción y de aire acondicionado de los sectores residencial y servicios.

Para determinar cuál es ese consumo se ha considerado que depende de varios factores, tales como la población, la riqueza, la eficiencia técnica y por supuesto la temperatura exterior. En un primer momento se han escogido variables que puedan cuantificar el efecto de dichos factores, para posteriormente obtener un modelo de regresión lineal apropiado y a partir de él calcular el consumo térmico de un año normal.

4.1. Variables del modelo lineal

4.1.1. Variables dependientes

En esta sección se estudian cualitativamente las variables dependientes, comprobando que pueden ser utilizadas como variables dependientes de modelos de regresión. Éstas se definen como variables aleatorias cuyo valor oscila alrededor de un valor medio según el valor de las variables independientes [CARO13].

Todas las series de datos fueron extraídas de la base de datos de ODYSSEE [ODYS17].

4.1.1.1. Consumo de calefacción del sector residencial

Los datos de demanda térmica del sector residencial se representan en el Gráfico 7. Puede decirse que esta variable experimenta variaciones año tras año con tendencia ascendente, por lo que es probable que dichos cambios puedan explicarse mediante la regresión.

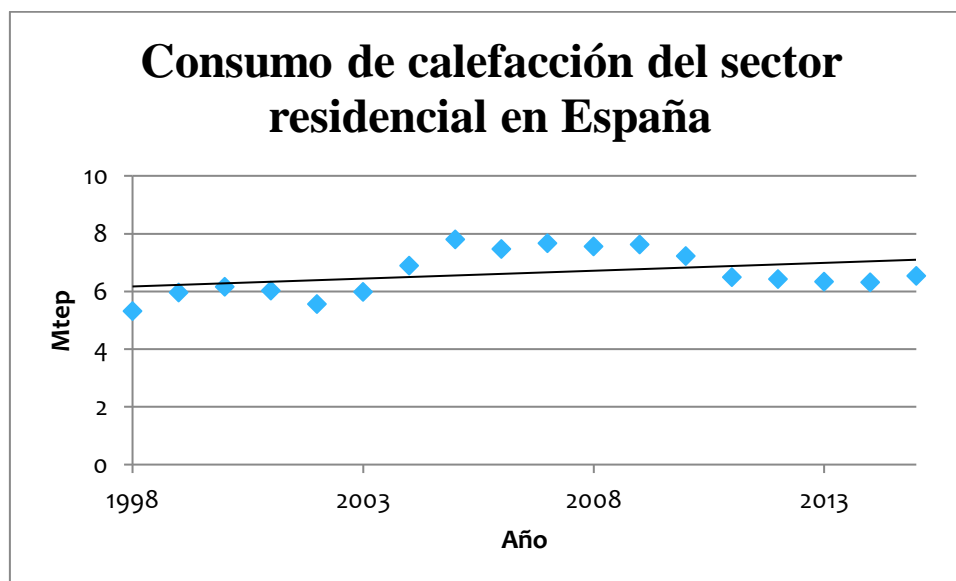


Gráfico 7: Consumo de calefacción del sector residencial en España. Fuente: ODYSSEE.

4.1.1.2. Consumo de calefacción del sector servicios

Los datos de demanda térmica del sector servicios se representan en el Gráfico 8. En este caso, la tendencia creciente es más acusada que en el sector residencial, y sus oscilaciones parecen más aleatorias. De cualquier manera, estos datos parecen poder explicarse igualmente a través de la regresión lineal.

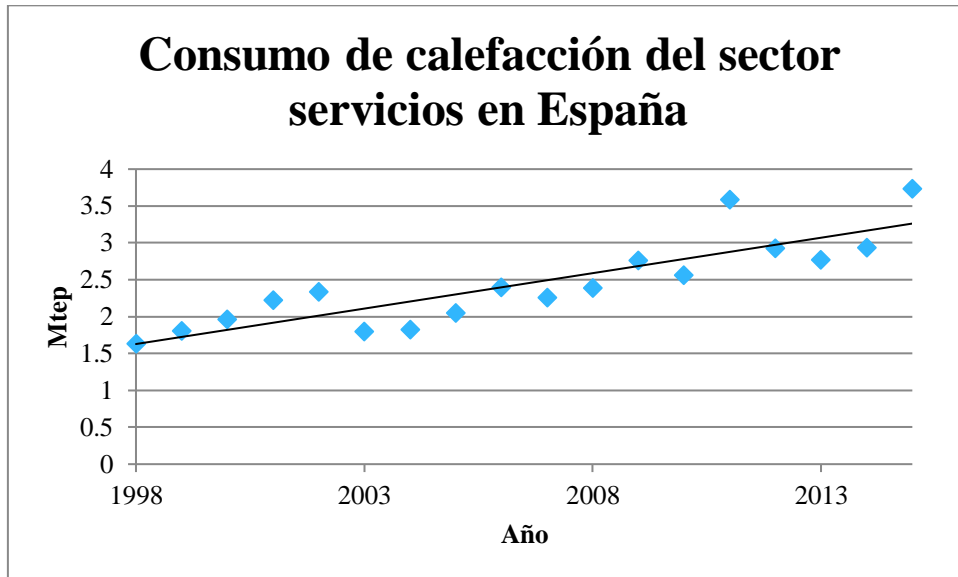


Gráfico 8 : Consumo de calefacción del sector servicios en España. Fuente: ODYSSEE.

4.1.1.3. Consumo de aire acondicionado del sector residencial

En el Gráfico 9 se representa, en megatoneladas de petróleo equivalentes, la evolución del consumo de aire acondicionado del sector residencial español desde 1998 hasta 2015. Se observa una gran pendiente en los primeros años y estabilidad en los últimos.

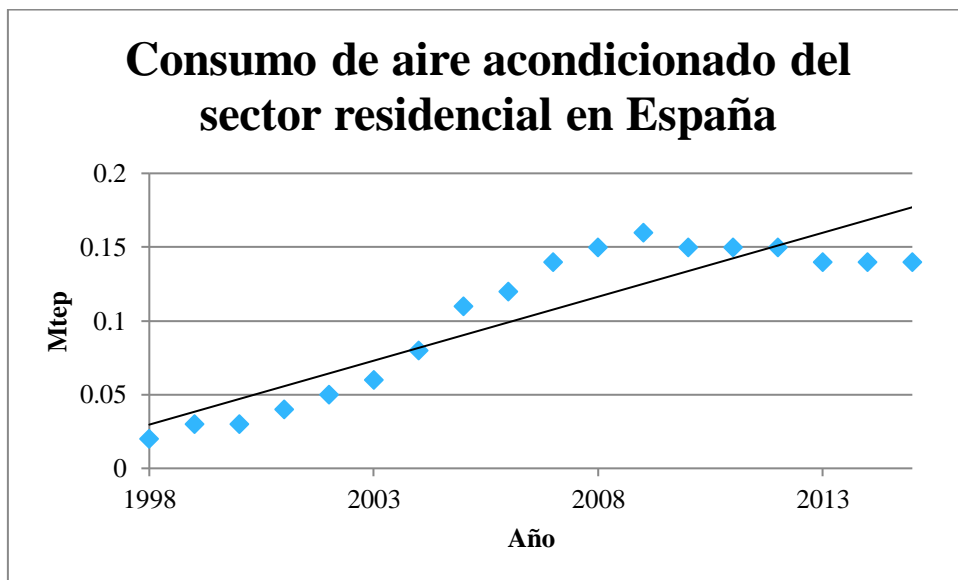


Gráfico 9 : Consumo de aire acondicionado del sector residencial en España. Fuente: ODYSSEE.

Puede deducirse que es apropiado su estudio en los modelos de regresión lineal.

4.1.1.4. Consumo de aire acondicionado del sector servicios

En el Gráfico 10 se muestra el consumo de aire acondicionado del sector servicios español desde el año 1998 al 2015. Puede observarse su tendencia ascendente, al igual que sus fluctuaciones. Estas características pueden indicar que los datos son apropiados para su representación en un modelo de regresión lineal.

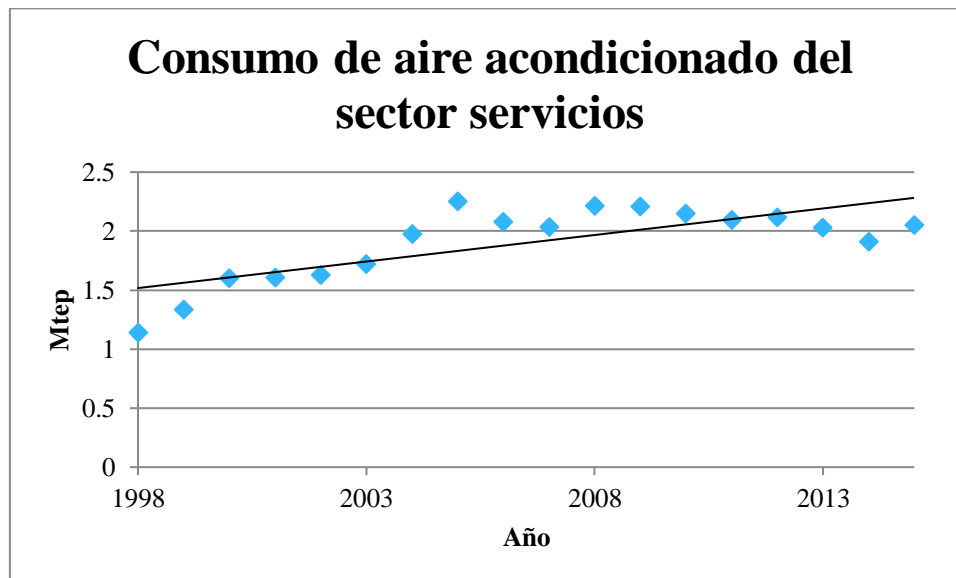


Gráfico 10: Consumo de aire acondicionado del sector servicios en España. Fuente: ODYSSEE.

4.1.1.5. Otros datos de consumo de calefacción

Además de los datos anteriores, también se dispone de los consumos de calefacción de los sectores residencial y servicios corregidos por las condiciones meteorológicas estimado por ODYSSEE. Se calcula siguiendo la ecuación [ODYSsf] (2):

$$H = H_n \cdot \frac{HDD}{HDD_n} \quad (18)$$

Donde H es el consumo real de calefacción en un determinado año, H_n el consumo que hubiera tenido lugar si ese año hubiese sido un año normal, HDD los días en grados de calefacción de ese año y HDD_n los días en grados de calefacción de un año normal, definido por ODYSSEE como la media a lo largo de veinte años o más de los HDD [ODYSsf].

4.1.2. Variables explicativas

En este apartado se pretende obtener variables explicativas que cuantifiquen los factores meteorológico, tecnológico, población y riqueza.

4.1.2.1. Cuantificación de las condiciones climáticas

Las condiciones climáticas tienen un determinado impacto en el uso de calefacción, en particular la temperatura exterior. Dos variables que pueden describir el impacto de la temperatura exterior en el confort en el interior son los días en grados de calefacción (HDD, por sus siglas en inglés) y los días en grados de refrigeración (CDD, también por sus siglas en inglés). Estas variables representan el número de días en los que la

temperatura interior (o temperatura de confort) difiere de la temperatura exterior y en cuantos grados. Existen varios métodos para su cálculo, pero los dos utilizados se explican en los apartados 4.1.2.1.1 y 4.1.2.1.2.

El cálculo de los HDD y CDD se llevó a cabo a partir de la temperatura máxima y mínima diaria en las capitales de provincia [AEME18] y de la población en cada provincia [INE_18], además se determinaron las siguientes variables:

- Temperatura media nacional.
- Temperatura media nacional ponderada según la población.
- Temperatura media nacional en los meses invernales, definidos como meses en los cuales los HDD calculados según el método I son distintos de cero.
- Temperatura media nacional en invierno desagregada según la población.
- Temperatura media nacional en los meses estivales, definidos como meses en los cuales los HDD calculados según el método I son distintos de cero.
- Temperatura media nacional en los meses estivales desagregada según la población.

4.1.2.1.1. Método clásico para el cálculo de HDD y CDD

El método clásico (en adelante, método I) es el más aceptado ([ODYSsf] (2), [INVEsf], [DUTTSf]) y su cálculo para un día está representado en la ecuación $HDD = T_{base}^{inv} - T_{media}$ (19) para los HDD y en la En ellas se distingue entre dos temperaturas de confort: una para invierno y otra para verano, y la temperatura media corresponde a la temperatura media diaria.

$$HDD = T_{base}^{inv} - T_{media} \quad (19)$$

$$CDD = T_{media} - T_{base}^{ver} \quad (20)$$

Para obtener el total en un año se han de sumar los HDD o los CDD relativos a cada día para cada provincia y posteriormente los de cada provincia para cada año. La temperatura base se elige de acuerdo a la temperatura de confort en el interior de los hogares. Esta temperatura se ha escogido de 18°C en invierno y de 24°C en verano.

No obstante, este cálculo es más exacto si en lugar de utilizarse la temperatura media diaria se midiese la temperatura en intervalos más cortos para posteriormente obtener la suma ponderada según la duración de esos intervalos [INVEsf].

4.1.2.1.2. Método refinado para el cálculo de HDD y CDD

Este método (en adelante, método II) es más sofisticado que el anterior, ya que distingue entre tres casos diferentes para el cálculo de ambas variables. Los casos considerados para los HDD pueden consultarse en la Ilustración 1, y su cálculo se detalla a continuación [LINA18]:

- i. En primer lugar, si la temperatura máxima de un día es menor que 18°C (la temperatura base escogida para invierno), se considera que en tal día sólo

existirá necesidad de calefacción. Por ello, se considera un HDD en su totalidad y se calcula de acuerdo a la ecuación $HDD = T_{base}^{inv} - T_{media}$ (19).

- ii. En segundo lugar, si la temperatura máxima es mayor que 18 °C y la mínima es menor, pero la media es mayor, entonces se utiliza la ecuación $HDD = 0.25 \cdot (T_{base}^{inv} - T_{media})$ (21).

$$HDD = 0.25 \cdot (T_{base}^{inv} - T_{media}) \quad (21)$$

- iii. En tercer lugar, si la temperatura máxima es mayor que la temperatura base y la mínima es menor, pero la media es inferior, entonces el cálculo a realizar es el siguiente:

$$HDD = 0.5 \cdot (T_{base}^{inv} - T_{min}) - 0.25 \cdot (T_{base}^{inv} - T_{media}) \quad (22)$$

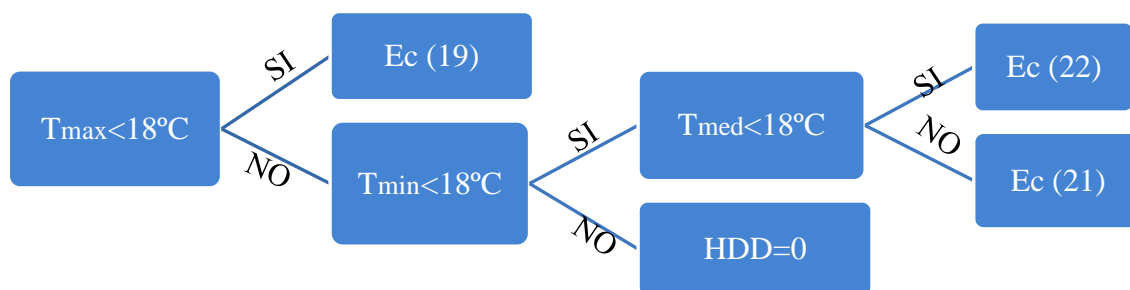


Ilustración 1 : Árbol de decisión para el cálculo de los HDD siguiendo el método II. Elaboración propia a partir de [LINA18].

De manera análoga, los casos que se distinguen para el cálculo de los CDD pueden consultarse en la Ilustración 2 y su cálculo se detalla en los siguientes puntos [LINA18]:

- i. Primero, si la temperatura mínima es menor que la temperatura base de verano, entonces se considera que el día sólo posee su componente de CDD, la cual se calcula siguiendo la ecuación $CDD = T_{media} - T_{base}^{ver}$ (20).
- ii. En otro caso, si la temperatura máxima es mayor que 24°C, la temperatura mínima menor pero la media resulta inferior, se sigue la siguiente ecuación.

$$CDD = 0.5 \cdot (T_{max} - T_{base}^{ver}) - 0.25 \cdot (T_{base}^{ver} - T_{med}) \quad (23)$$

- iii. Por último, si la temperaturas media y mínima de tal día son inferiores a la temperatura base, el cálculo que ha de realizarse es el siguiente:

$$CDD = 0.25 \cdot (T_{max} - T_{base}^{ver}) \quad (24)$$

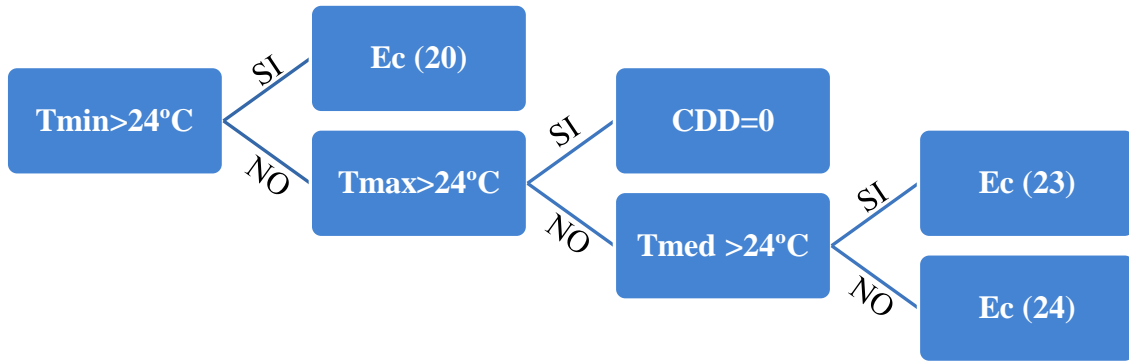


Ilustración 2 : Árbol de decisión para el cálculo de los CDD siguiendo el método II. Elaboración propia a partir de [LINA18].

Utilizar los HDD desagregados según la población puede resultar útil para analizar el consumo energético en España, ya que una temperatura más severa en una provincia con poca población no tendrá el mismo efecto en el consumo total si la provincia es de las más pobladas.

Para su cálculo, los HDD totales en cada provincia se multiplicaron por el porcentaje de la población española residente en tal provincia durante cada año (extraídos de [INE_sf]), y posteriormente se sumaron para cada año. Los resultados pueden observarse en el Gráfico 12, y el método de cálculo en la ecuación $DD^j = \sum_p^i DD_i^j \cdot \frac{P_i}{P_T}$ (25), donde DD^j hace referencia al total de HDD o CDD del año j , DD_i^j representa los HDD que la provincia i tuvo en el año j , P_i la población de la provincia i y P_T la población total española.

$$DD^j = \sum_p^i DD_i^j \cdot \frac{P_i}{P_T} \quad (25)$$

4.1.2.1.3. Resultados del cálculo de HDD en España

Los resultados para el cálculo de HDD en España pueden observarse en el Gráfico 11. Se observa que los valores absolutos obtenidos mediante el método II son mayores que los del método I. Por otro lado, en el Gráfico 12 pueden observarse los resultados para los HDD de ambos métodos desagregados según la población. Ambas series de datos son bastante parecidas, sin embargo, tras utilizar la ecuación $DD^j = \sum_p^i DD_i^j \cdot \frac{P_i}{P_T}$ (25) sobre los resultados del Gráfico 11, las distancias entre ambas series se acortan o estrechan para distintos años.

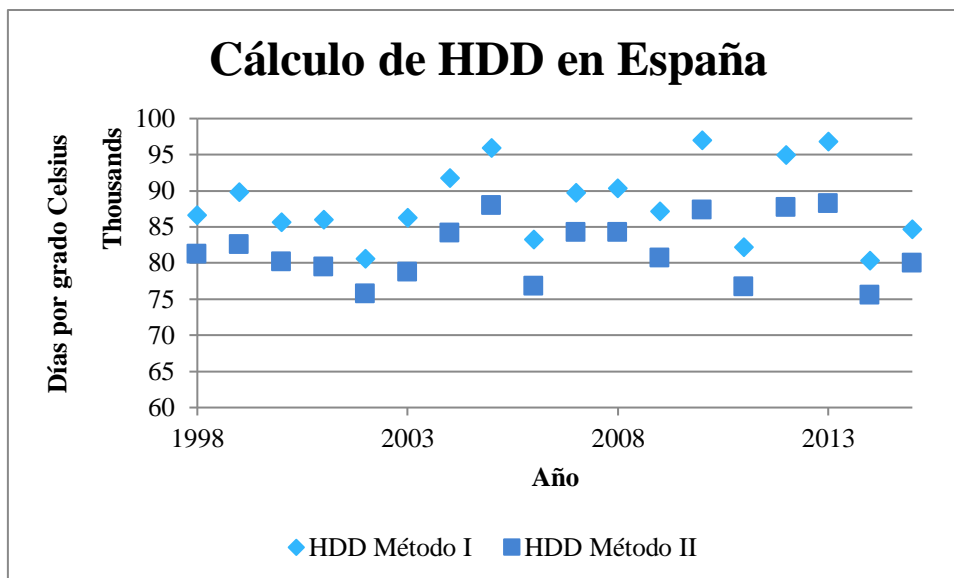


Gráfico 11 : Resultados para el cálculo de HDD en España. Elaboración propia a parit de datos de AEMET.

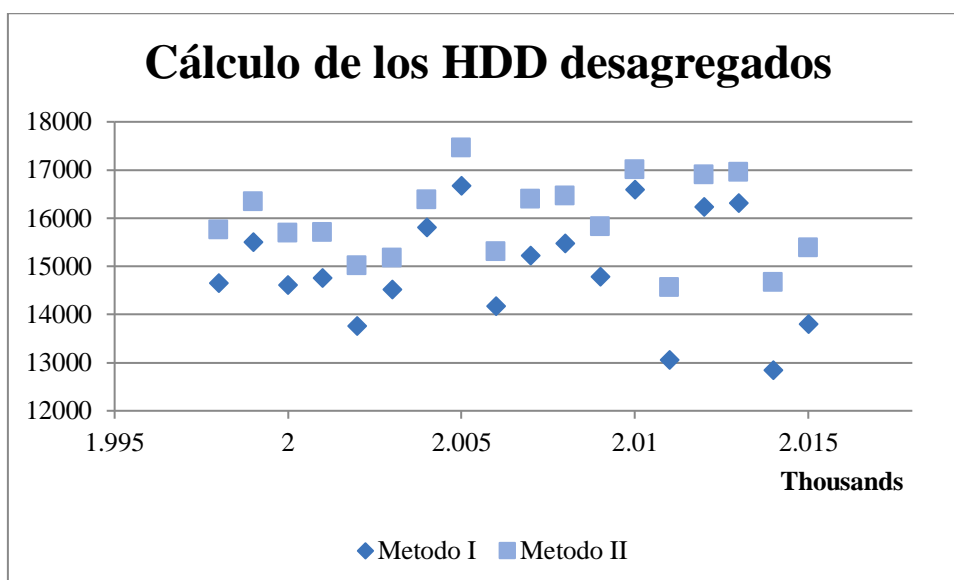


Gráfico 12: HDD en España desagregados según la población. Elaboración propia a partir de datos del INE y de AEMET.

4.1.2.1.4. Resultados del cálculo de CDD en España.

En el Gráfico 14 pueden observarse los resultados del cálculo de los CDD desagregados por la población. En este caso, los valores absolutos siguen siendo mayores para el método II. Respecto a su comparación con los datos de HDD, puede afirmarse que:

- Las diferencias entre ambos métodos son menos acusadas que en el caso de los HDD.
- Los valores absolutos de los HDD presentes en el Gráfico 12 son mucho mayores que los de los CDD del Gráfico 14.

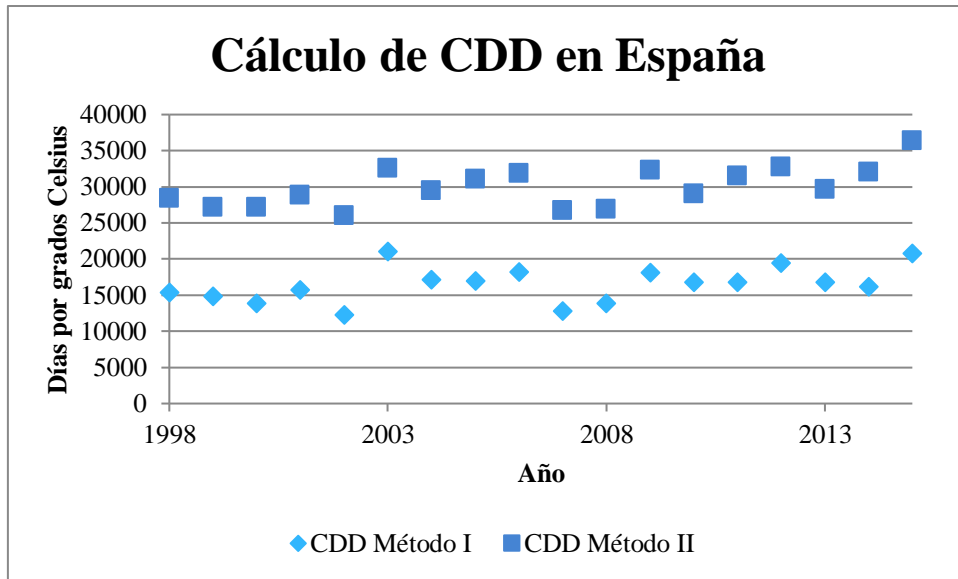


Gráfico 13 : Cálculo de los CDD en España. Elaboración propia a partir de datos de AEMET.

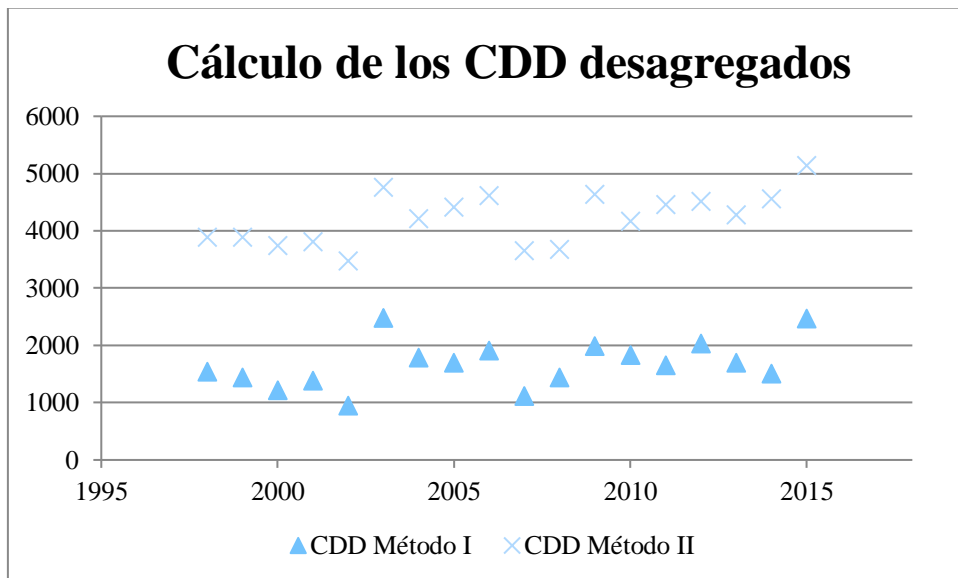


Gráfico 14 : Resultados del cálculo de los CDD desagregados. Fuente: Elaboración propia a partir de datos del AEMET y del INE.

4.1.2.1.5. Otros datos de HDD

Otros datos sobre los HDD y CDD en España, calculados siguiendo la ecuación $HDD = T_{base}^{inv} - T_{media}$ (19) y ponderándolos según la población de cada estación meteorológica, fueron extraídos de la base de datos de ODYSSEE-MURE [ODYS17] y se presentan en el Gráfico 15.

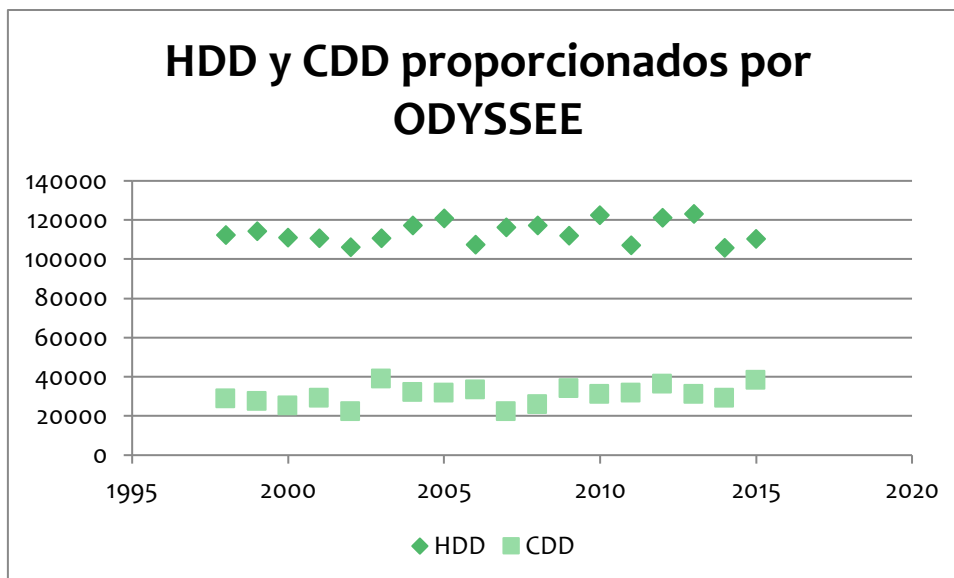


Gráfico
Estos

4.1.2.1.6. Comparación de las series de datos

Puesto que los resultados numéricos difieren bastante para ambas series de datos, es preciso analizar la variabilidad normalizada (ver ecuación $Z_i = \frac{X_i - \bar{X}}{\sigma(X)}$ (27)) de éstas para poder establecer comparaciones. Estos resultados pueden observarse en el Gráfico 16.

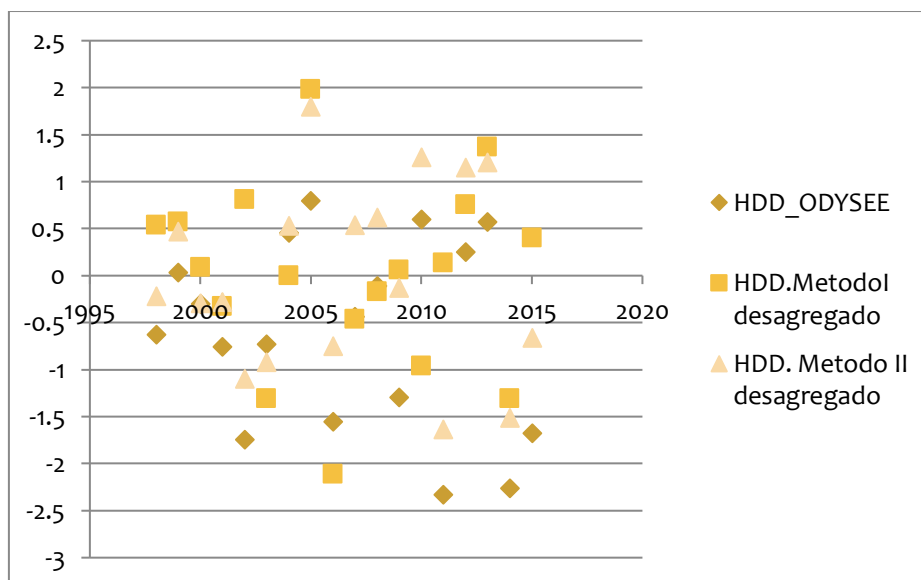


Gráfico 16 : Variabilidad normalizada de los distintos datos de HDD desagregados. Elaboración propia a partir de datos de ODYSSEE, AEMET e INE.

Se puede deducir que todas las series presentan una variabilidad parecida, pero los datos no presentan la misma evolución respecto a los años. Respecto a las series de datos de ODYSSEE y del método I, estos últimos parecen tener una menor dispersión. En

definitiva, ninguna de las series es resultado de un factor de escala con otra, por lo que las tres se utilizarán para la realización de los modelos lineales.

En el Gráfico 17 se presenta lo mismo pero para los datos de CDD. Puede deducirse que los datos de CDD están más próximos que los de HDD.

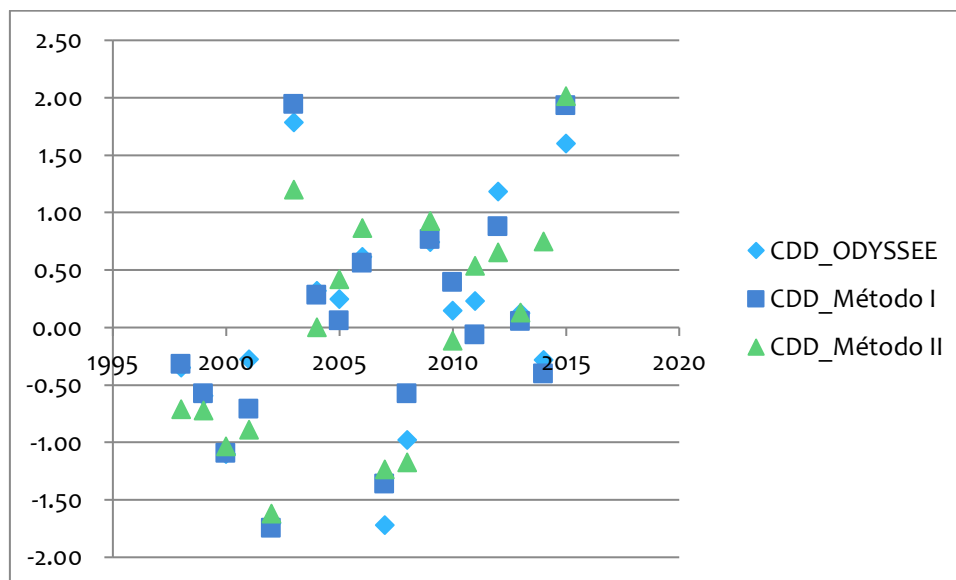


Gráfico 17: Variabilidad normalizada de los distintos datos de CDD.

Como se ve en el apartado Taltal, los datos que se esto son los de la ecuación esta, que són los de la gráfica. Por ello, todas son variables relevantes.

4.1.2.2. Cuantificación de la innovación tecnológica

Para expresar numéricamente el factor tecnológico se ha utilizado el indicador ODEX [ODYS17]. Éste mide el ahorro energético que se da en un año debido a las mejoras en las eficiencias de los aparatos, y su fórmula de cálculo se expresa a continuación [ODYSsf]:

$$ODEX = \frac{E}{E+ES} \cdot 100\% \quad (26)$$

En la ecuación anterior, E representa la energía consumida en calefacción en un determinado año, y ES el ahorro energético producido debido a la tecnología.

Para el cálculo de tal índice, ODYSSEE estima el ahorro energético (ES) cada año respecto a un año base (en este caso, el año 2000). Para cada periodo, se multiplica la variación del consumo medio de un determinado uso final o maquinaria por el número de aparatos existentes. Por último, para evitar excesivas fluctuaciones, el ODEX se calcula como resultado de una media móvil de tres años. El objetivo es evitar la influencia del comportamiento, ciclos laborales y errores estadísticos entre otras variaciones [ODYSsf].

Este índice se calcula en los ámbitos del transporte, del consumo residencial, de la industria y finalmente global. Ya que sólo se poseen datos del ODEX a partir del año

2000 (año base), para los años anteriores (1999 y 1998) se ha considerado también la unidad.

Para el cálculo del índice relativo a calefacción del sector residencial, ODYSSEE considera el consumo por m² en condiciones normales [ODYSsf]. En el caso del sector servicios, también existen varios índices [ODYSsf]:

- Eficiencia bruta: es el ODEX que se computa directamente de las variaciones del consumo de energía real.
- Ahorro negativo de energía: este término corresponde al hecho de que el consumo energético aumente en lugar de reducirse. Se considera que su origen no puede ser la eficiencia técnica sino otros factores externos.
- Eficiencia técnica: es el resultado de sustraer de la eficiencia bruta el ahorro negativo.

Para los modelos se han utilizado el indicador de eficiencia técnica, que es el que resume los avances de la tecnología real y las mejoras del consumo.

Respecto al consumo de aire acondicionado, no se disponen de datos de ODEX pertinentes referidos a España.

4.1.2.3. Cuantificación de población y riqueza

Esta sección engloba diferentes datos relativos a la población y el nivel de riqueza en España, tales como la población (fuentes: INE [INE_sf] (2) y la base de datos “Datos Macro” [DATOs]), el PIB a precios constantes (fuente: Banco Mundial [BANCsf]) y la renta per cápita a precios constantes (elaborada a partir de los datos anteriores).

4.1.2.4. Resumen de las variables explicativas

En definitiva, las variables con que se pretende explicar el consumo energético de uso térmico de los sectores residencial y servicios en España pueden clasificarse en:

Nombre corto	Tipo	Descripción	Unidad
HDD M.I.	Climática	HDD calculados según el método I.	Días por grado celsius
HDD M.I.D.	Climática	HDD calculados según el método I y desagregados según la población.	Días por grado Celsius
HDD M.II.D.	Climática	HDD calculados según el método II.	Días por grado Celsius
HDD M.II.D.	Climática	HDD calculados según el método II y desagregados según la población.	Días por grado Celsius
HDD ODYSSEE	Climática	HDD proporcionados por ODYSSEE.	Días por grado Celsius
CDD M.I.	Climática	CDD calculados según el método I.	Días por grado celsius
CDD M.I.D.	Climática	CDD calculados según el método I y desagregados según la población.	Días por grado Celsius
CDD M.II.	Climática	CDD calculados según el método II.	Días por grado

			Celsius
CDD M.II.D.	Climática	CDD calculados según el método II y desagregados según la población.	Días por grado Celsius
CDD ODYSSEE	Climática	CDD proporcionados por ODYSSEE.	Días por grado celsius
T. med.	Climática	Temperatura media.	Grado Celsius
T. med. D.	Climática	Temperatura media desagregada por la población.	Grado Celsius
T. med. ver.	Climática	Temperatura media en verano.	Grado Celsius
T. med. ver. D.	Climática	Temperatura media en verano desagregada según la población.	Grado Celsius
T. med. inv.	Climática	Temperatura media en invierno.	Grado Celsius
T. med. inv. D.	Climática	Temperatura media en invierno desagregada según la población.	Grado Celsius
PIB	Población, desarrollo y riqueza	Producto Interior Bruto español a precios constantes.	Millones de euros constantes (2002)
Población	Población	Población	Millones de habitantes
Renta per cápita	Riqueza	Renta per cápita española a precios constantes.	Miles de euros constantes (2002)
ODEX R.	Tecnológica	Índice de eficiencia energética en consumo de calefacción del sector residencial	-
ODEX S.	Tecnológica	Índice de eficiencia energética del sector servicios	-

Tabla 8: Resumen de variables explicativas a utilizar.

4.1.2.5. Formato de las variables

Puesto que las magnitudes y los valores de las variables explicativas son muy dispares, para cuantificar el efecto de la variación de una de ellas se ha decidido trabajar con valores homogéneos. Con ello se evita que ciertos coeficientes parezcan despreciables frente a otros pese a que las variables tengan un gran poder explicativo, ya que diferencias en las magnitudes provoca necesariamente diferencias en los coeficientes.

Para el cálculo de estos valores se ha seguido el procedimiento descrito en la ecuación

$$Z_i = \frac{x_i - \bar{X}}{\sigma(X)} \quad (27).$$

$$Z_i = \frac{x_i - \bar{X}}{\sigma(X)} \quad (27)$$

Donde Z_i es el resultado de la homogenización del valor i de la variable X , \bar{X} es la media de dicha variable y $\sigma(X)$ es su desviación típica.

Trabajar con valores homogéneos provoca que los coeficientes de correlación sean proporcionales a la variación de los datos de las variables explicativas, y que dicho

coeficiente no dependa de las diferencias entre las magnitudes de la variable dependiente y las explicativas.

4.2. Consumo de calefacción del sector residencial

En este apartado se exploran las distintas posibilidades para la elaboración de un modelo de regresión lineal del consumo de calefacción del sector residencial con las diferentes variables explicativas, y finalmente se escoge el modelo más apropiado.

4.2.1. Modelos de regresión lineal de calefacción del sector residencial

En un primer momento se realizó un modelo de regresión lineal (modelo A) teniendo en cuenta los HDD desagregados calculados con el método 1, la población, la renta per cápita a precios constantes y el índice tecnológico ODEX. Se eligieron estas variables por que incluyen los aspectos climatológicos, demográficos, nivel de riqueza y eficiencia en las tecnologías de calefacción.

Los resultados pueden observarse en la siguiente tabla:

Poder explicativo del modelo				
F-test	13.5	P-valor (F-test)	0.000149	
R²	0.806	R² ajustado	0.746	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.386		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	±0.0911	72.8290	2.2956e-18
HDD M.I.D.	0.1001	±0.0993	1.0086	0.3316
Población	0.6540	±0.2453	2.6655	0.0194
Renta per cápita	0.3923	±0.1436	2.7331	0.0171
ODEX	0.5302	±0.2152	2.4644	0.0284

Tabla 9: Resultados del modelo A.

Respecto al modelo anterior pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.75.
- Las variables Población, Renta per cápita y ODEX tienen un poder explicativo a un nivel de significación del 5%.
- La variable climática en este caso no parece aportar información al modelo, ya que su p-valor es superior al 10% de nivel de significación.
- Todos los coeficientes son positivos, luego un incremento en cualquiera de las variables provoca un incremento en el consumo de calefacción en el sector residencial.

Puesto que eliminar la variable climática del modelo carece de sentido, de todas las observaciones puede deducirse que es necesario sustituir la variable climática de este modelo por otra cuyo p-valor asociado resulte más bajo. Los resultados se muestran en el apartado 4.2.1.1.

Otra posibilidad es deducir que el poder explicativo variables empleadas se solapa, por lo que se trataron de sustituir o eliminar las otras variables para obtener un modelo con mayor poder explicativo (apartado 4.2.1.2).

4.2.1.1. Modelo A.1.

El modelo A.1 está basado en el modelo A cambiando la variable referente al clima (HDD desagregados según el método I) por las demás variables disponibles. El objetivo es obtener un p-valor asociado a dicha variable climática que sea más bajo que en el caso A, manteniendo bajos el resto de p-valores.

4.2.1.1.1. Modelo A.1.1.

En este modelo se sustituyó la variable climática del modelo A por los HDD no desagregados calculados con el método I. Los resultados se agrupan en la siguiente tabla:

Poder explicativo del modelo				
F-test	15.9	P-valor (F-test)	6.28e-05	
R ²	0.83	R ² ajustado	0.778	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.361		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	±0.0850	77.981	9.457e-19
HDD M.I.	0.1723	±0.0984	1.7524	0.1032
Población	0.4673	±0.2549	1.8338	0.0896
Renta per cápita	0.4143	±0.1339	3.0925	0.0086
ODEX	0.3653	±0.2216	1.6481	0.1232

Tabla 10: Resultados del modelo A.1.1.

Puede observarse que:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A y sus valores son más altos, siendo el coeficiente ajustado menor que 0.78.
- Todos los coeficientes estimados siguen siendo positivos.

- El p-valor asociado a la variable climática es menor que en el modelo A, sin embargo no puede aceptarse a un nivel de significación del 10%. Esto mismo ocurre con la variable ODEX.
- Los p-valores del resto de variables pueden aceptarse al nivel de significación del 10%.

La principal conclusión de las observaciones es que el modelo A.1.1. no puede considerarse válido.

4.2.1.1.2. Modelo A.1.2.

En este caso se utilizaron los datos de HDD proporcionados por ODYSSEE. Los resultados pueden observarse en la

Poder explicativo del modelo				
F-test	16.5	P-valor (F-test)	5.17e-05	
R ²	0.835	R ² ajustado	0.785	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.355		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.7374	±0.1008	66.8470	6.9747e-18
HDD ODYSSEE	0.1706	±0.0903	1.8898	0.0813
Población	0.5724	±0.2305	2.4823	0.0275
Renta per cápita	0.3998	±0.1299	3.0782	0.0088
ODEX	0.4209	±0.2059	2.0436	0.0618

Tabla 11.

Poder explicativo del modelo				
F-test	16.5	P-valor (F-test)	5.17e-05	
R ²	0.835	R ² ajustado	0.785	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.355		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.7374	±0.1008	66.8470	6.9747e-18
HDD ODYSSEE	0.1706	±0.0903	1.8898	0.0813
Población	0.5724	±0.2305	2.4823	0.0275

Renta per cápita	0.3998	± 0.1299	3.0782	0.0088
ODEX	0.4209	± 0.2059	2.0436	0.0618

Tabla 11: Resultados del modelo A.1.2.

Puede anotarse que:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A y sus valores son más altos, siendo el coeficiente ajustado menor que 0.79.
- Todos los coeficientes estimados siguen siendo positivos.
- Los p-valores asociados a cada una de las variables pueden aceptarse a un nivel de significación del 10%.

De todo ello puede afirmarse que el modelo A.1.2. es válido a un nivel de significación del 10%. Por este motivo, se adjuntan en la Figura 1 las propiedades sus residuos.

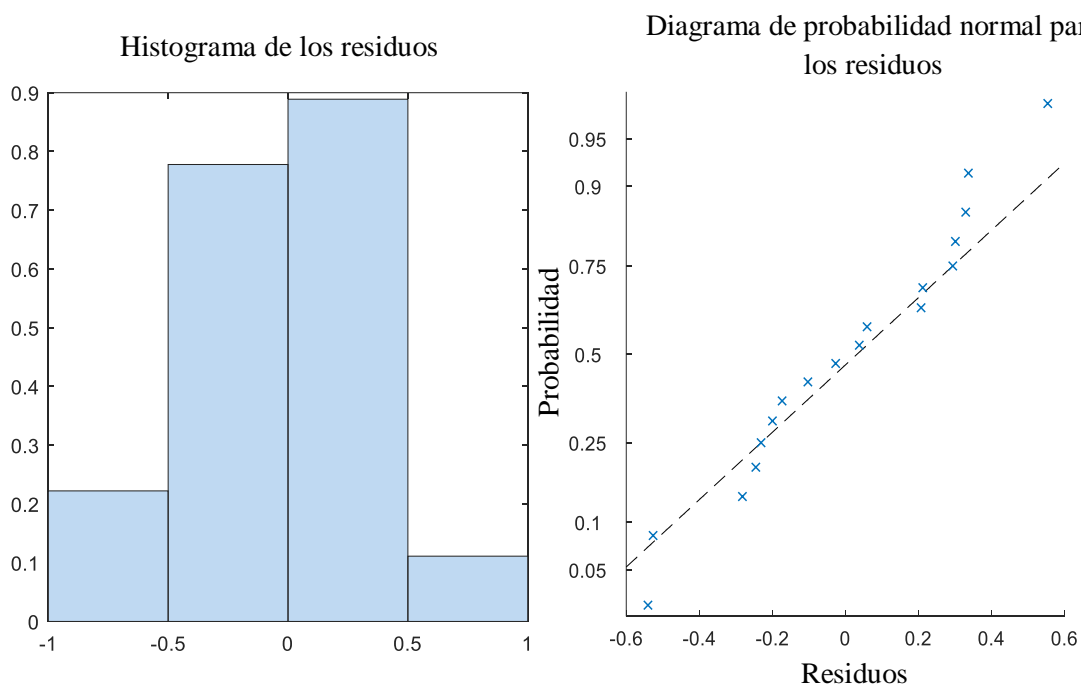


Figura 1 : Propiedades de los residuos del modelo A.1.2.

Los residuos de este modelo se distribuyen de manera casi simétrica y su distribución se ajusta a una distribución normal de media nula y varianza constante. En definitiva, el modelo A.1.2. es válido para su explotación.

4.2.1.1.3. Modelo A.1.3.

En este modelo se incluyeron los HDD calculados según el método II. Los resultados fueron los siguientes:

Poder explicativo del modelo				
F-test	15.7	P-valor (F-test)	6.77e-05	
R ²	0.828	R ² ajustado	0.776	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.363		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	±0.0855	77.5180	1.0217e-18
HDD M.II.	0.1881	±0.1108	1.6972	0.1134
Población	0.7596	±0.2372	3.2020	0.0069
Renta per cápita	0.2371	±0.1481	1.6009	0.1334
ODEX	0.6729	±0.2192	3.0703	0.0089

Tabla 12: Resultados del modelo A.1.3.

Se pueden realizar las siguientes afirmaciones a cerca del modelo A.1.3:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A y sus valores son más altos, siendo el coeficiente ajustado menor que 0.78.
- Todos los coeficientes estimados siguen siendo positivos.
- El p-valor asociado a la variable climática es menor que en el modelo A, sin embargo no puede aceptarse a un nivel de significación del 10%. Esto mismo ocurre con la variable Renta per cápita.
- Los p-valores del resto de variables pueden aceptarse al nivel de significación del 10%.

La principal conclusión de las observaciones es que el modelo A.1.3 no puede considerarse válido.

4.2.1.1.4. Modelo A.1.4.

En este caso, se utilizaron los HDD calculados según el método II desagregados según la población. Los resultados fueron los siguientes:

Poder explicativo del modelo				
F-test	19.8	P-valor (F-test)	1.96e-05	
R ²	0.859	R ² ajustado	0.815	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.329		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	±0.0776	85.4640	2.8799e-19
HDD M.II.D.	0.2120	±0.0844	2.5106	0.0261
Población	0.4937	±0.2196	2.2481	0.0426
Renta per cápita	0.4052	±0.1200	3.3762	0.0050
ODEX	0.3831	±0.1923	1.9918	0.0678

Tabla 13: Resultados del modelo A.1.4.

Las siguientes conclusiones pueden obtenerse a cerca de los resultados:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A y sus valores son más altos, siendo el coeficiente ajustado menor que 0.82.
- Todos los coeficientes estimados siguen siendo positivos.

- Los p-valores asociados a cada una de las variables pueden aceptarse a un nivel de significación del 10%.

Puesto que el modelo tienen un gran poder explicativo y todos los coeficientes son menores que el nivel de significación del 10%, este modelo se considera válido. Por este motivo se añaden las propiedades de sus residuos en la Figura 2.

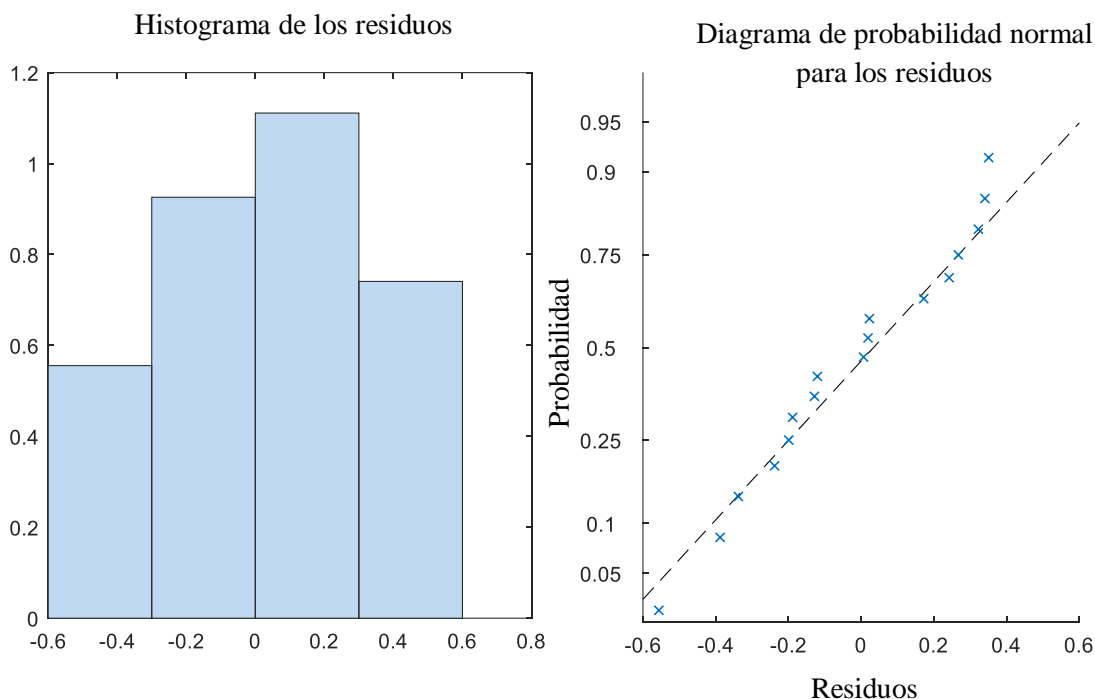


Figura 2: Propiedades de los residuos del modelo A.1.4.

Los residuos se distribuyen de manera simétrica respecto a la media, y se ajustan a una distribución normal de media cero y varianza constante. Puede concluirse que el modelo A.1.4 es un modelo válido para su explotación.

4.2.1.1.5. Otros.

También se probó sustituir la variable de HDD por la temperatura media, la temperatura media desagregada, la temperatura media en invierno y la temperatura media en invierno desagregada, sin embargo se obtuvieron en todos los casos p-valores de los t.test individuales superiores al 10%.

4.2.1.2. Modelo A.2.

Para la realización de este modelo la hipótesis de base es que el poder explicativo de las variables del modelo A se solapaban, por lo que se debía variar las variables implicadas en el modelo e incluso eliminar una de ellas. Se utilizan los HDD desagregados según el método I como variable climática, ya que fue para esta serie de datos para la que se obtuvieron los mejores resultados en el modelo A.1.

4.2.1.2.1. Modelo A.2.1.

En un primer momento se decidió eliminar la variable del ODEX para observar si su poder explicativo se solapaba y se podía llegar a un modelo con menos variables y p-valores individuales más bajos. Los resultados pueden observarse en la Tabla 14.

Poder explicativo del modelo				
F-test	20.7	P-valor (F-test)	2.08e-05	
R ²	0.816	R ² ajustado	0.776	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.362		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	±0.0854	77.6330	7.5299e-20
HDD M.II.D.	0.2628	±0.0886	209665	0.0102
Población	0.0962	±0.1009	0.9536	0.3565
Renta per cápita	0.5600	±0.1007	5.5637	6.9831e-05

Tabla 14: Resultados del modelo A.2.1.

Pueden realizarse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A y sus valores son más altos, siendo el coeficiente de correlación ajustado menor que 0.78.
- Todos los coeficientes estimados son siendo positivos.
- El p-valor de la variable Población no puede aceptarse a un nivel de significación del 10%. El resto de variables son significativas al 1%.
- Respecto al modelo B, el p-valor asociado a la población aumenta desproporcionadamente al eliminar el ODEX.

En conclusión, el modelo A.2.1 no puede considerarse válido.

4.2.1.2.2. Modelo A.2.2.

Para este caso se decidió conservar el ODEX, pero variar la renta per cápita por el producto interior bruto.

Poder explicativo del modelo				
F-test	20.5	P-valor (F-test)	1.61e-05	
R ²	0.863	R ² ajustado	0.821	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.324		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	±0.0764	86.791	2.3580e-19
HDD M.II.D.	0.2151	±0.0832	2.5841	0.0227
PIB	0.7412	±0.2122	3.4932	0.0040
Población	0.0200	±0.3267	0.0611	0.9522
ODEX	0.3493	±0.1941	1.799	0.0953

Tabla 15: Resultados del modelo A.2.2.

De la tabla anterior puede afirmarse lo siguiente:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A y sus valores son más altos, siendo el valor del coeficiente de correlación ajustado menor que 0.83.
- Todos los coeficientes estimados son siendo positivos.
- El p-valor de la variable Población no puede aceptarse a un nivel de significación del 10%.
- El resto de variables son significativas para los niveles habituales.
- Respecto al modelo B, el p-valor asociado a la población aumenta desproporcionadamente al eliminar la renta per cápita.

Por lo tanto, el modelo A.2.2 no puede considerarse apto para su explotación.

4.2.1.2.3. Modelo A.2.3.

En este modelo también se elimina el ODEX, y se sustituyen renta per cápita y población por el PIB. Este cambio fue motivado por el incremento del p-valor de la población en los modelos A.2.1 y A.2.2.

Los resultados se muestran en la siguiente tabla:

Poder explicativo del modelo				
F-test	19.8	P-valor (F-test)	6.25e-05	
R ²	0.725	R ² ajustado	0.688	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.428		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	± 0.1008	65.7760	7.01e-20
HDD M.II.D.	0.2562	± 0.1046	2.4498	0.0271
PIB	0.5682	± 0.1046	5.4324	6.93e-05

Tabla 16 : Resultados del modelo A.2.3.

De la tabla anterior puede afirmarse lo siguiente:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A, sin embargo sus valores son más bajos, siendo inferiores a 0.69.
- Todos los coeficientes estimados son siendo positivos.
- Todos los p-valores pueden aceptarse a un nivel de significación del 1%

De todas las observaciones anteriores puede determinarse que el modelo anterior es apropiado, por lo que se adjuntan gráficas del comportamiento de sus residuos.

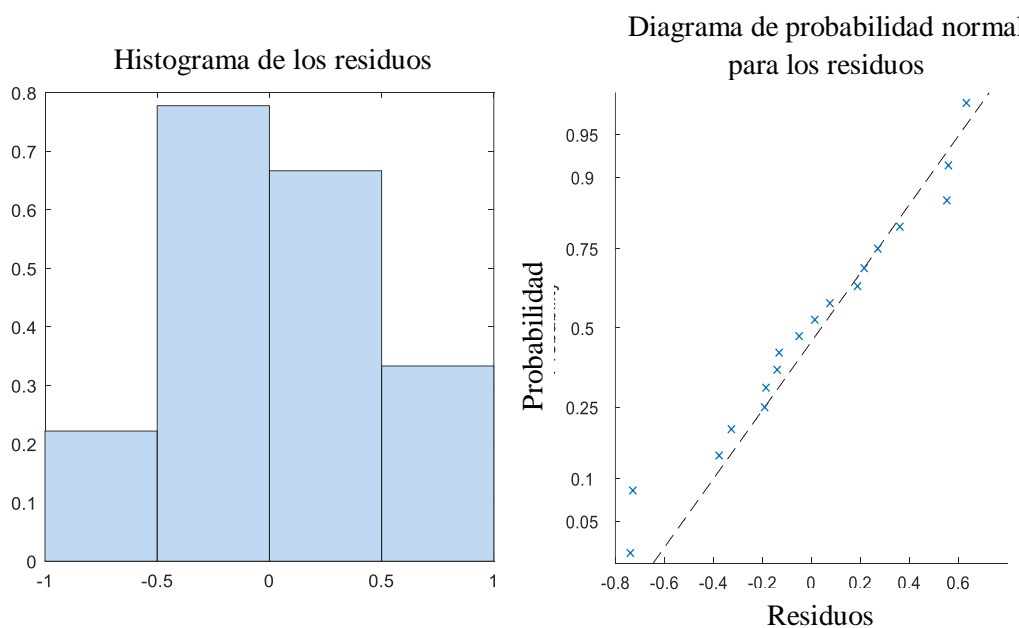


Figura 3: Distribución de los residuos del modelo A.2.3.

Los residuos se distribuyen de manera simétrica respecto a la media, y se ajustan a una distribución normal de media cero y varianza constante. Puede concluirse que el modelo A.2.3 es un modelo válido para su explotación.

4.2.1.2.4. Modelo A.2.4.

Este modelo está basado en el anterior (modelo A.2.3) pero incluyendo el ODEX. Los resultados pueden observarse en la Tabla 17:

Poder explicativo del modelo				
F-test	29.4	P-valor (F-test)	2.66e-06	
R ²	0.863	R ² ajustado	0.834	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.312		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.6314	0.0736	90.0680	9.44e-21
HDD M.II.D.	0.2164	0.0771	2.8066	0.0140
PIB	0.7528	0.0908	8.2896	9.03e-07
ODEX	0.3389	0.0902	3.7584	0.0021

Tabla 17 : Resultados del modelo A.2.4.

Del modelo anterior pueden extraerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene poder explicativo para todos los niveles de significación habituales.
- Los coeficientes de correlación y correlación ajustada difieren en menor medida que en el modelo A y sus valores son más altos, siendo inferiores a 0.84.
- Todos los coeficientes estimados son siendo positivos.
- Todos los p-valores pueden aceptarse a un nivel de significación del 5%.

Por lo tanto, el modelo anterior se puede considerar apto, por lo que se adjuntan los gráficos de distribución de sus residuos.

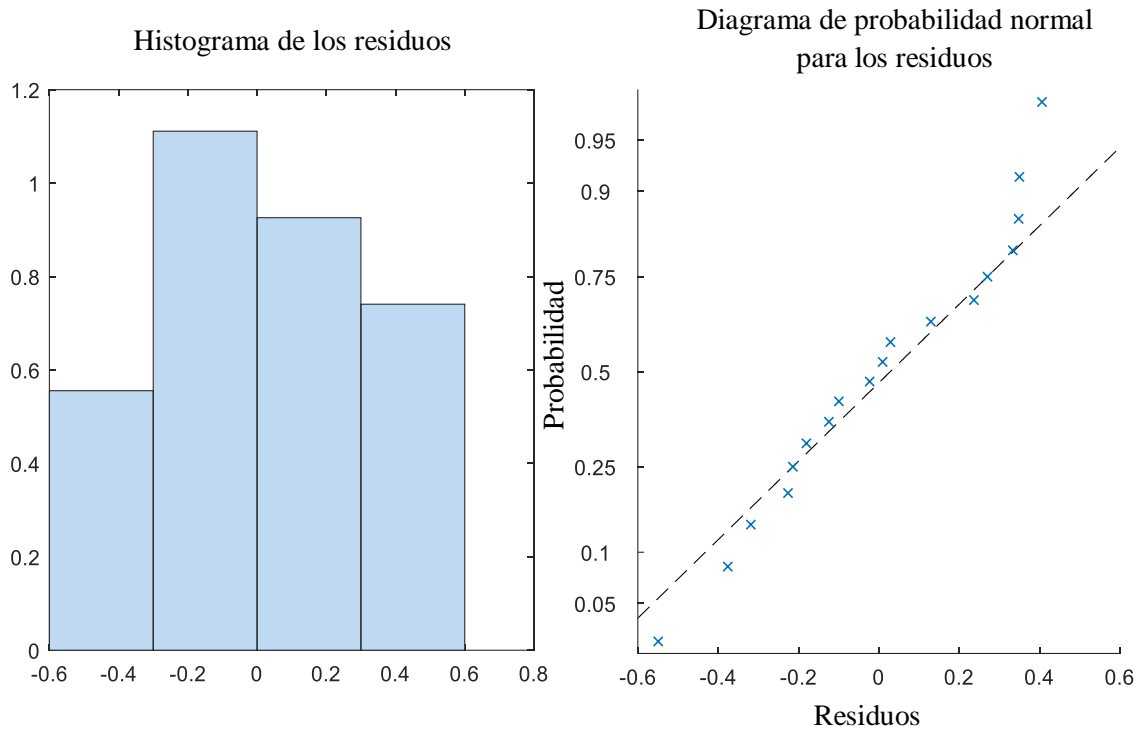


Figura 4 : Distribución de los residuos del modelo A.2.4.

De la Figura 4 puede concluirse los residuos se distribuyen de manera simétrica respecto a la media y se ajustan a una distribución normal de media cero y varianza constante. Por lo tanto, el modelo A.2.4 es un modelo válido para su explotación.

4.2.2. Resultados del modelo de regresión para el consumo de calefacción en el sector residencial.

De los modelos cuyo objetivo era explicar el comportamiento del consumo de calefacción en el sector residencial, sólo cuatro se consideran aptos el modelo A.1.2, el A.1.4, A.2.3 y el A.2.4. Tanto el A.1.2 como el A.1.4 utilizan las variables de población, renta per cápita y el índice tecnológico ODEX, pero mientras el primero utiliza los HDD proporcionados por ODYSSEE, el segundo utiliza los HDD calculados de manera desagregada siguiendo el método II. El tercero utiliza sólo los HDD desagregados calculados con el método II y el PIB, y el cuarto añade la variable de ODEX.

Los dos primeros pueden considerarse válidos a un nivel de significación del 10% y sus residuos siguen una distribución normal de media cero y varianza constante. Así mismo, los resultados de los coeficientes de correlación y correlación ajustada son parecidos, así como el error cuadrático de los residuos. Por último, el valor de los coeficientes para cada variable es prácticamente el mismo en los dos modelos (con la excepción de la variable población, cuyo coeficiente es de 0.7 en el modelo A.1.2 y de 0.5 en el modelo A.1.4). Por ese motivo, ambos resultados podrían utilizarse indistintamente para explicar el consumo de calefacción en el sector residencial.

El tercer modelo puede ser aceptado a un nivel de significación del 5%, aunque la correlación entre las variables explicativas y el consumo en calefacción del sector residencial es menor que en el caso anterior. Por último, el cuarto modelo puede aceptarse al 5% de significación y los niveles de correlación son comparables a los dos primeros.

Por lo tanto, el más apropiado es el A.2.4, ya que puede ser aceptado a un nivel de significación del 5% y posee altos coeficientes de correlación.

En la Tabla 18 puede observarse el peso relativo de cada variable en este modelo:

Modelo	Clima	PIB	Población	Renta per cápita	ODEX
A.2.4.	16.54%	57.55%	-	-	25.91%

Tabla 18 : Peso de las variables en el modelo A.2.4.

El peso en el modelo de cada variable (PM) ha sido calculado utilizando la siguiente ecuación:

$$PM(\%) = \frac{\hat{\beta}_i}{\sum \hat{\beta}_j} \cdot 100 \% \quad (28)$$

Puede deducirse que, tanto la variación en la riqueza como en la población (unidas en el PIB) son los factores más importantes en la explicación del consumo de calefacción en el sector residencial. El segundo factor más influyente es la variación en las mejoras tecnológicas, dejando la componente climática con el menor peso.

Tras haber escogido el modelo, puede calcularse la componente corregida del consumo de calefacción del sector residencial teniendo en cuenta el clima utilizando la ecuación $Y^* = Y - \hat{\beta}_1 \cdot X_1$ (9). En el siguiente gráfico se compara el consumo real con la componente corregida según el modelo A.2.4.

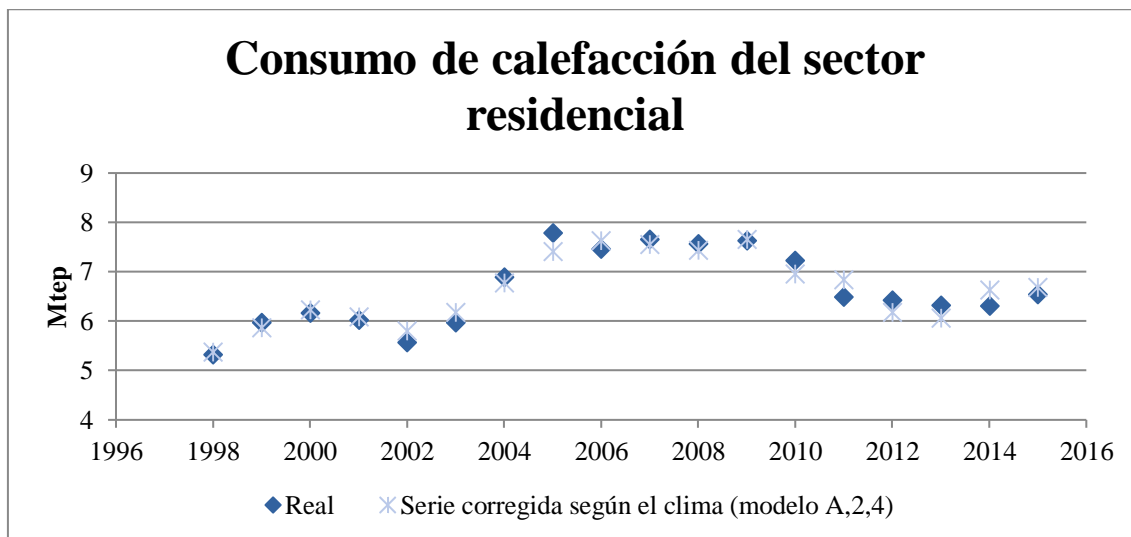


Gráfico 18: Comparación del consumo de calefacción del sector residencial real y la serie corregida con el modelo A.2.4.

Los datos de ambas series son bastante parecidos, presentando diferencias en torno al 5% respecto a la serie original en los años más desfavorables.

4.3. Consumo de calefacción del sector servicios

Este apartado exploran los posibles modelos de regresión múltiple para el consumo de calefacción del sector servicios y presenta, de entre todos ellos, el más apropiado para su representación.

4.3.1. Modelos de regresión del consumo del calefacción del sector servicios

Basado en los resultados anteriores, las variables que constituyen el modelo inicial (modelo B) para el caso del sector servicios son los HDD desagregados calculados con el método II, el PIB y el ODEX técnico del sector servicios. Se han elegido estas variables porque dan cuenta del clima, el número de personas, la actividad económica y la eficiencia en los equipos.

Poder explicativo del modelo				
F-test	16.5	P-valor (F-test)	7.24e-05	
R ²	0.779	R ² ajustado	0.732	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.308		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4454	0.0727	33.6410	8.56e-15
HDD M.II.D.	-0.1978	0.0755	-2.6204	0.0202
PIB	-0.1320	0.1477	-0.8937	0.3866
ODEX	-0.6104	0.1471	-4.1503	0.0009

Tabla 19 : Resultados del modelo B.

Pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.74.
- El p-valor del poder explicativo de las variables Población, Renta per cápita y ODEX es muy elevado, y el de los HDD es aceptable a un nivel de significación del 5%.
- Los coeficientes son negativos, lo que indica una relación inversa entre estas variables y el consumo de calefacción del sector servicios.

Puede deducirse que el modelo es explicativo, pero el poder de las variables se solapa y se distorsiona, por lo que aparecen p-valores altos en la mayoría de variables y coeficientes negativos (que carecen de sentido).

Como posibles caminos, se ha escogido sustituir la variable climática por otra (modelo B.1), eliminar una de las variables para evitar solapamientos (modelo B.2) y combinaciones de ambas (modelo B.3).

4.3.1.1.1. Modelo B.1.

En este apartado se sustituye la variable climática presente en el modelo B por otra cuya participación en el modelo no resulte en coeficientes negativos y que contribuya a reducir los p-valores del resto de variables.

4.3.1.1.2. Modelo B.1.1.

En esta sección se estudia el efecto de sustituir la variable climática del modelo B por todas las variantes de variables de HDD que se poseen. Los resultados más importantes se presentan en la siguiente tabla:

Variable considerada	Variabes con alto p-valor (>10%)	Variabes con coeficientes negativos
HDD M.II.	Todas	PIB y ODEX
HDD M.I.D.	HDD y PIB	PIB y ODEX
HDD M.I.	PIB	Todas
HDD ODYSSEE	PIB	Todas

Tabla 20: Resumen de los resultados más importantes del modelo B.1.1.

Por lo tanto, ninguna de estas variantes puede considerarse válida.

4.3.1.1.3. Modelo B.1.2.

En este caso, se estudia el efecto de sustituir la variable climática HDD por otras, tales como la temperatura media anual, la temperatura media anual durante el invierno, y esta última ponderada por la población.

Variable considerada	Variabes con alto p-valor (>10%)	Variabes con coeficientes negativos
T. med.	PIB	PIB y ODEX
T. med. D.	PIB	PIB y ODEX
T. med. Inv.	PIB	PIB y ODEX
T. med. Inv. D.	PIB	PIB y ODEX

Tabla 21 : Resumen de los resultados más importantes del modelo B.1.2.

Se deduce que ninguna de las posibilidades resultó ser apropiada.

4.3.1.2. Modelo B.2.

En este apartado se sustituyen o eliminan las variables no climáticas del modelo B con el objetivo de obtener coeficientes positivos y p-valores más bajos para todas las variables que compongan el modelo. Las variables que se consideran son: PIB, población y ODEX.

4.3.1.2.1. Modelo B.2.1.

En este modelo se sustituye la variable del PIB por la de la población, para dar cuenta del gran incremento experimentado en el consumo del sector servicios (ver Gráfico 8).

Poder explicativo del modelo				
F-test	16.1	P-valor (F-test)	8.02e-05	
R ²	0.776	R ² ajustado	0.728	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.311		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4454	0.0732	33.3920	9.49e-15
HDD M.II.D.	-0.2246	0.0799	-2.8092	0.0139
Población	0.5848	0.7675	0.7619	0.4588
ODEX	0.0826	0.7649	0.1079	0.9156

Tabla 22 : Resultados del modelo B.2.1.

Pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.73.
- El p-valor del poder explicativo de las variables Población y ODEX es muy elevado, y el de los HDD es aceptable a un nivel de significación del 5%.
- El coeficiente de la variable climática es negativo, lo cual carece completamente de sentido, ya que indica una relación inversa entre estas variables y el consumo de calefacción del sector servicios.

Es evidente que este modelo no puede explotarse.

4.3.1.2.2. Modelo B.2.2.

Este modelo está basado en el anterior, del cual se ha eliminado la variable del ODEX, ya que posee el p-valor más alto. Los resultados se presentan en la siguiente tabla:

Poder explicativo del modelo				
F-test	25.9	P-valor (F-test)	1.36e-05	
R ²	0.776	R ² ajustado	0.746	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.300		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4454	0.0708	34.5490	1.02e-15
HDD M.II.D.	-0.2218	0.0734	-3.0244	0.0085
Población	0.5023	0.0734	6.8482	5.52e-06

Tabla 23 : Resultados del modelo B.2.2.

En este caso, es posible afirmar que:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.75.
- Los p-valores pueden aceptarse a cualquier nivel de significación.
- El coeficiente de la variable climática es negativo, lo cual carece completamente de sentido, ya que indica una relación inversa entre estas variables y el consumo de calefacción del sector servicios.

Por lo tanto, este modelo no puede considerarse válido.

4.3.1.2.3. Modelo B.2.3.

Este modelo está basado en el B.2, del cual se ha decidido eliminar el PIB. Los resultados se presentan en la Tabla 25:

Poder explicativo del modelo				
F-test	28.9	P-valor (F-test)	7.23e-06	
R ²	0.794	R ² ajustado	0.766	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.288		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.3033	0.0808	28.5100	1.76e-14
HDD M.II.D.	-0.2288	0.0706	-3.2427	0.0055
ODEX	-0.4470	0.0706	-6.3345	1.04e-05

Tabla 24 : Resultados del modelo B.2.3.

Respecto al modelo anterior pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.73.
- Los p-valores pueden aceptarse a cualquier nivel de significación.
- Ambos coeficientes son negativos, lo cual carece completamente de sentido, ya que indica una relación inversa entre estas variables y el consumo de calefacción del sector servicios.

En definitiva, este modelo no puede considerarse apto para su explotación.

4.3.1.2.4. Modelo B.2.4.

Por último, este modelo está basado en el modelo B.2, del cual se ha eliminado el ODEX. Los resultados son los siguientes:

Poder explicativo del modelo				
F-test	7.72	P-valor (F-test)	0.0049	
R ²	0.507	R ² ajustado	0.442	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.445		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4454	0.1049	23.3160	3.38e-13
HDD M.II.D.	-0.2122	0.1088	-1.9508	0.0700
PIB	0.3952	0.1088	3.6320	0.0024

Tabla 25 : Resultados del modelo B.2.4.

Del modelo B.2.4 puede afirmarse que:

- El p-valor del F-test indica que tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.45, lo cual es un valor bajo.
- El efecto del PIB puede ser aceptado a cualquier nivel de significación y el de los HDD es aceptable a un nivel de significación del 10%.
- El coeficiente de la variable climática es negativo, lo cual carece completamente de sentido, ya que indica una relación inversa entre estas variables y el consumo de calefacción del sector servicios.

Luego este modelo no es apropiado para su explotación.

4.3.1.3. Modelo B.3.

Para la realización de los nuevos modelos, se ha decidido combinar las variantes del Modelo B.2. con las del Modelo B.1. En los modelos B.3.1 y B.3.2 se sustituye la variable climática presente en el modelo B.2.2; en el B.3.3 se sustituye la variable climática del B.2.4.

4.3.1.3.1. Modelo B.3.1.

En este modelo se sustituyó la variable climática del modelo B.2.2 por la temperatura media. Los resultados pueden observarse en la siguiente tabla:

Poder explicativo del modelo				
F-test	24.4	P-valor (F-test)	1.94e-05	
R ²	0.765	R ² ajustado	0.733	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.308		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4453	0.0725	33.7340	1.46e-15
T. med.	0.2216	0.0783	2.8320	0.0126
Población	0.4089	0.0782	5.2255	0.0001

Tabla 26 : Resultados del modelo B.3.1.

De este modelo puede deducirse que:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.74.
- El modelo es significativo al 5%.
- Todos los coeficientes son positivos.

Por lo tanto, el modelo anterior puede considerarse apto para su explotación. Por ese motivo, en la Figura 5 se muestra la distribución de sus residuos.

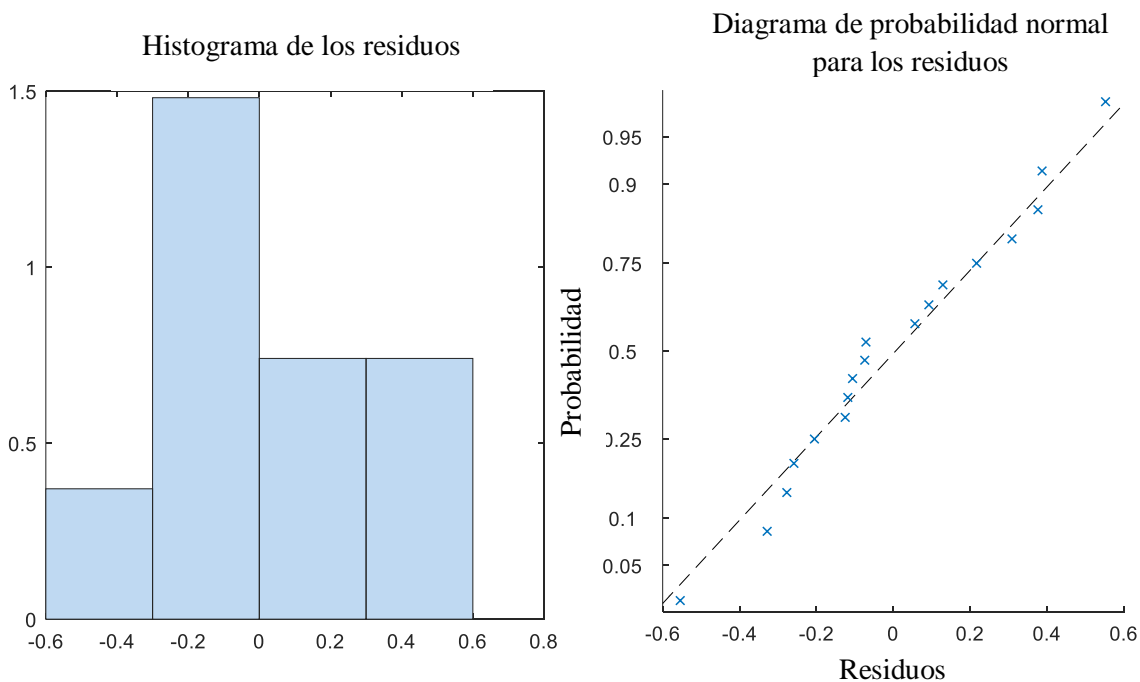


Figura 5 : Distribución de los residuos del modelo B.3.1.

Se observa que los residuos se distribuyen de manera aproximadamente simétrica y que se ajustan a una distribución normal de media cero y varianza constante. Por lo tanto, este modelo resulta apto.

4.3.1.3.2. Modelo B.3.2.

En este modelo, se sustituyeron los HDD desagregados calculados según el método II del modelo B.2.2 por la temperatura media de los meses de invierno. Los resultados se presentan en la siguiente tabla:

Poder explicativo del modelo				
F-test	27.3	P-valor (F-test)	9.94e-16	
R ²	0.785	R ² ajustado	0.756	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.294		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4453	0.0693	35.2730	7.56e-16
T. med. Inv.	0.2315	0.0726	3.1888	0.0061
Población	0.4329	0.0726	5.9638	2.60e-05

Tabla 27 : Resultados del modelo B.3.2.

Del modelo B.3.2 se puede sintetizar que:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.76.
- El modelo es significativo al 1%.
- Todos los coeficientes son positivos.

Por lo tanto, el modelo anterior puede considerarse apto para su explotación, por lo que se adjunta la distribución de sus residuos en la figura a continuación:

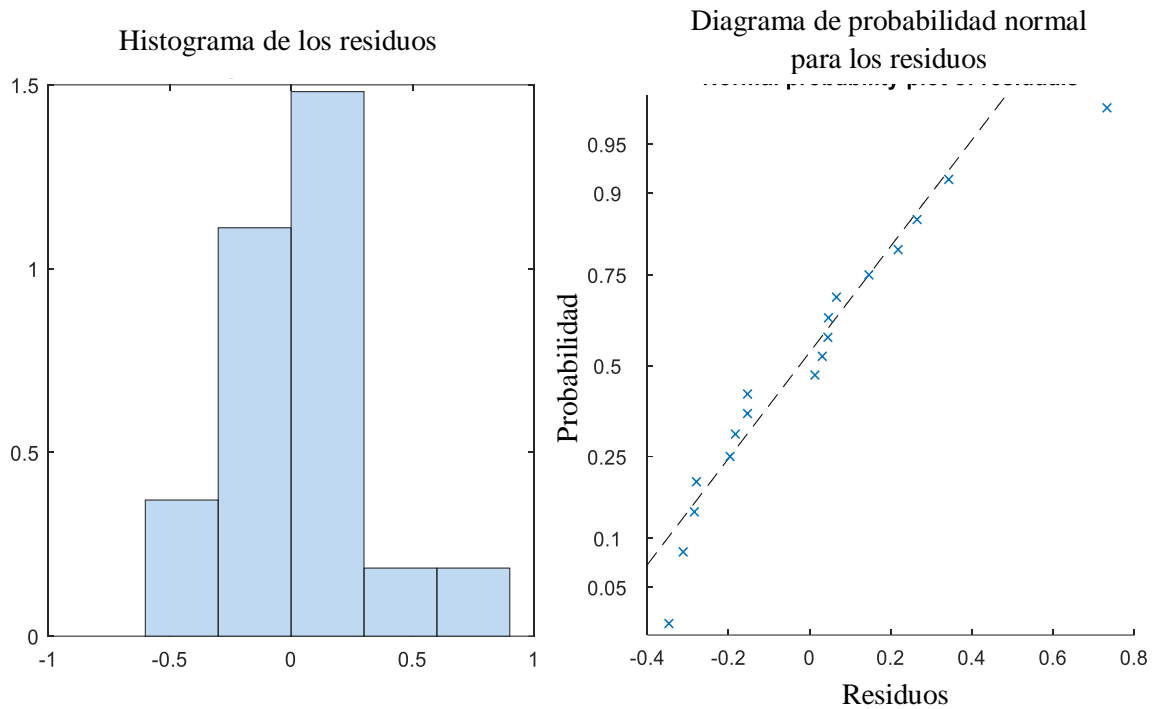


Figura 6: Distribución de los residuos del modelo B.3.2.

La distribución de los residuos es aproximadamente simétrica (el histograma no es simétrico por la acción de un único residuo).

4.3.1.3.3. Modelo B.3.3.

En este modelo, se sustituyeron los HDD desagregados calculados según el método II del modelo B.2.2 por la temperatura media de los meses de invierno. Los resultados se presentan en la siguiente tabla:

Poder explicativo del modelo				
F-test	19.4	P-valor (F-test)	6.85e-05	
R ²	0.722	R ² ajustado	0.684	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.335		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4453	0.0789	31.012	5.019e-15
T. med. Inv. D.	0.185	0.0876	2.1109	0.0520
Población	0.406	0.0876	4.6324	0.0003

Tabla 28 : Resultados del modelo B.3.3.

Del modelo B.3.3 se puede sintetizar que:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.69.
- El modelo es significativo al 1%.
- Todos los coeficientes son positivos.

Por lo tanto, el modelo anterior puede considerarse apto para su explotación, por lo que se adjunta la distribución de sus residuos en la figura a continuación:

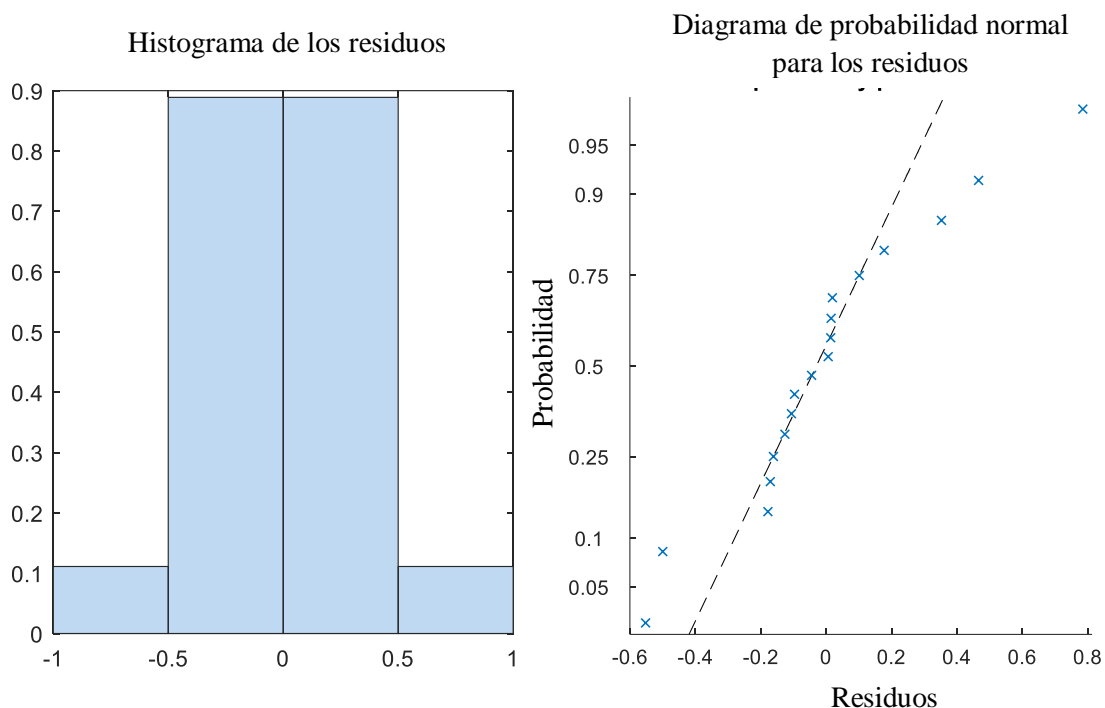


Figura 7: Distribución de los residuos del modelo B.3.3.

La distribución de los residuos es perfectamente simétrica, pero función de probabilidad normal se pierde fuertemente en ambas colas. En conclusión, este modelo no puede aceptarse como válido.

4.3.1.3.4. Otras variantes del modelo B.2.1.

En este apartado se recogen los resultados del resto de variantes del modelo B.2.1 al sustituir su variable climática por otra.

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
HDD M.II	HDD	-
HDD M.I.D.	HDD	-
HDD M.I.	-	HDD
HDD ODYSSEE	-	HDD
T. med. D.	Temperatura media desagregada	-

Tabla 29 : Resultados más importantes del resto de variantes del modelo B.2.2.

Respecto a los resultados del modelo B.2.1 con la temperatura media en invierno calculada de manera desagregada, el modelo puede aceptarse a un nivel de significación del 5%, pero el coeficiente de correlación es más bajo que en los casos B.3.1 y B.3.2.

4.3.1.3.5. Variantes del modelo B.2.3.

La Tabla 30 muestra los resultados más importantes de las variantes propuestas para el modelo B.2.2. La conclusión principal es que ninguna de ellas resultó constituir un modelo válido.

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
HDD M.II.	HDD	ODEX
HDD M.I.D.	HDD	ODEX
HDD M.I.	-	HDD y ODEX
HDD ODYSSEE	-	HDD y ODEX
T. med	-	ODEX
T. med. D.	T. med. D.	ODEX
T. med. Inv.	-	ODEX
T. med. Inv. D.	-	ODEX

Tabla 30 : Resultados más importantes de las variantes del modelo B.2.2.

4.3.1.3.6. Modelo B.3.3.

En este caso, se sustituyeron los HDD desagregados calculados según el método II del modelo B.2.3 por la temperatura media. Los resultados pueden hallarse en la siguiente tabla:

Poder explicativo del modelo				
F-test	10.4	P-valor (F-test)	0.0014	
R ²	0.582	R ² ajustado	0.526	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.410		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4453	0.0966	25.3090	1.01e-13
T. med.	0.2737	0.1023	2.6754	0.0173
PIB	0.3037	0.1023	2.9688	0.0096

Tabla 31 : Resultados del modelo B.3.4.

Respecto al modelo anterior pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.53, por lo que se considera que la correlación es baja.
- El modelo es significativo al 1%.
- Todos los coeficientes son positivos.

Se

4.3.1.3.7. Otras variantes del modelo B.2.4.

La

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
-----------------------------	-------------------------------	--------------------------------

HDD M.II.	HDD	-
HDD M.I.D.	HDD	-
HDD M.I.	-	HDD
HDD ODYSSEE	-	HDD
T. med. D.	T. med. D.	-

Tabla 32 : Resultados más importantes del resto de variantes del modelo B.2.3.

Los resultados de las variantes con las temperaturas medias en invierno agregada y desagregada presentaron bajos coeficientes de correlación, como en el caso B.3.3.

4.3.2. Resultados de los modelos de consumo de calefacción del sector servicios

En el caso del sector servicios, se obtuvo un gran número de casos con coeficientes de correlación negativos y p-valores altos. Además, las variables climáticas basadas en los HDD han resultado ser inútiles. Los únicos modelos válidos para este conjunto de datos fueron los modelos B.3.1 y B.3.2. El primero trata de explicar el consumo de calefacción del sector servicios mediante las variables de temperatura media y población y el segundo la temperatura media en invierno y la población. Por lo tanto, las variables explicativas de ambos tienen un origen muy similar.

El

E

l

P

e

s

Modelo	Temperatura media en invierno	Población
B.3.2	34.84%	65.16%

Tabla 33: Peso de las variables del modelo B.3.2.

Por lo tanto, para el sector servicios sólo hay dos variables influyentes: la población y la temperatura media en invierno.

En el siguiente gráfico se compara el consumo real del sector servicios con la serie corregida según el clima calculada a partir del modelo B.3.2.

l

a

s

v

a

r

i

a

b

l

e

s

e

n

e

s

t

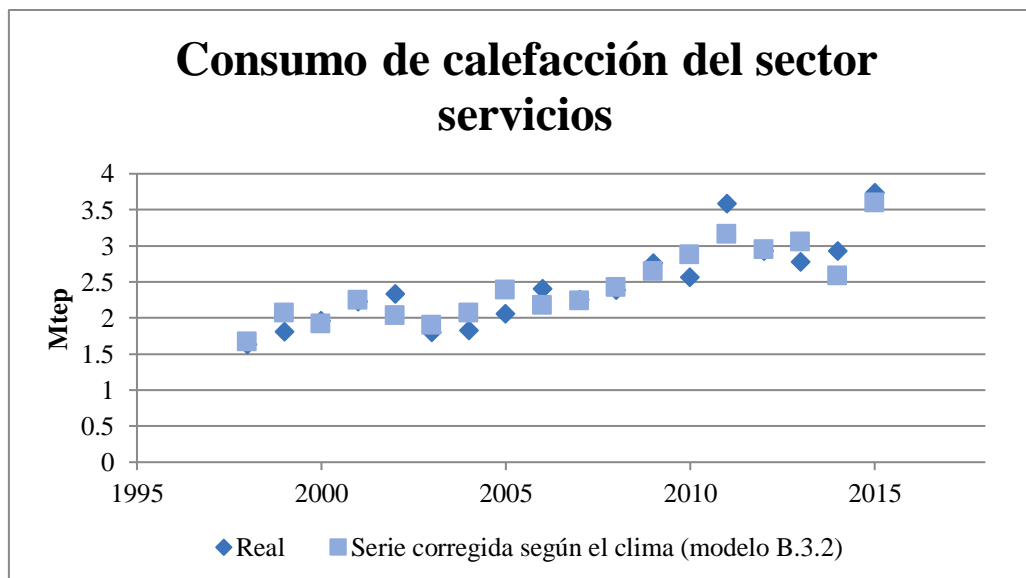


Gráfico 19 : Comparación del consumo de calefacción del sector servicios real y la serie corregida según el clima siguiendo el modelo B.3.2.

La serie corregida se calculó siguiendo el mismo procedimiento descrito en la ecuación **Error! Reference source not found.**, adaptando sus elementos al consumo del sector servicios y al modelo B.3.2. Los datos difieren más que en el caso del sector residencial, llegando a presentar diferencias en torno al 12% para ciertos años.

4.4. Consumo de aire acondicionado del sector residencial

En este apartado se exploran las distintas posibilidades para la elaboración de un modelo de regresión lineal del consumo de calefacción del sector residencial con las diferentes variables explicativas, y finalmente se escoge el modelo más apropiado.

4.4.1. Modelos del consumo de aire acondicionado del sector residencial

Para este conjunto de datos no se dispone de ninguna variable que dé cuenta de las mejoras tecnológicas. Por este motivo, las variables que constituyen el modelo inicial (modelo C) para el caso del consumo de aire acondicionado del sector residencial son los CDD desagregados calculados con el método II y el PIB. Los resultados pueden observarse en la siguiente tabla:

Poder explicativo del modelo				
F-test	64.8	P-valor (F-test)	4.17e-08	
R²	0.896	R² ajustado	0.882	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.0176		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	0.1033	0.0041	24.9610	1.24e-13
CDD M.II.D.	0.0058	0.0046	1.2543	0.2289
PIB	0.0459	0.0046	9.9138	5.59e-8

Tabla 34: Resultados del modelo C.

Respecto al modelo anterior puede observarse que:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.890.
- El PIB tiene un poder explicativo significativo al 1%, sin embargo, los CDD no pueden aceptarse para ningún nivel de significación en el modelo.
- Todos los coeficientes son positivos.

Se concluye que este modelo no puede ser considerado válido. Por este motivo, puesto que el p-valor de la variable climática es el único superior al nivel de significación escogido, se decidió en un primer momento sustituir esta variable por otra también de origen climático (modelo C.1). En segundo lugar, sustituir el PIB por población y renta per cápita, ya que en el sector residencial se ha de tener en cuenta tanto población como

nivel de riqueza individual (modelo C.2). Finalmente, se decidió probar a eliminar las variables climáticas, ya que su participación en los modelos parecía no tener ninguna contribución.

4.4.1.1. Modelo C.1.

En este apartado se trató de sustituir la variable climática por otra de la que se dispone.

4.4.1.1.1. Modelo C.1.1.

En este modelo se sustituyeron los CDD desagregados del método II por los del método I y los proporcionados por ODYSSEE. Ambas alternativas presentaban altos coeficientes de correlación y un p-valor del F-test aceptable a todos los niveles de significación. Sin embargo, el p-valor de las variables climáticas resultó ser demasiado elevado para poder aceptarse a un nivel de significación del 10%, tal y como se resume en la siguiente tabla.

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
CDD M.I.D.	CDD M.I.D.	-
CDD ODYSSEE	CDD ODYSSEE	-
CDD M.I.	CDD M.I.	-
CDD M.II.	CDD M.II.	-
T. med. ver.	T. med. ver	T. med. ver.

Tabla 35 : Principales resultados del modelo C.1.1.

4.4.1.1.2. Modelo C.1.2.

En este caso se decidió sustituir la variable climática del modelo C por la temperatura media en los meses estivales desagregada según la población. Los resultados de este modelo se recogen en la Tabla 36.

Poder explicativo del modelo				
F-test	73.5	P-valor (F-test)	1.78e-08	
R²	0.907	R² ajustado	0.895	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.0166		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	0.1033	0.0039	26.418	5.41e-13
T. med. ver. D.	0.0088	0.0046	1.8881	0.0785
PIB	0.0437	0.0047	9.3581	1.18e-07

Tabla 36 : Resultados del modelo C.1.2.

Respecto al modelo anterior pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.890, ligeramente superior al modelo C.
- El PIB tiene un poder explicativo significativo al 1%, y la temperatura media puede aceptarse a un nivel de significación del 10 %.
- Todos los coeficientes son positivos.

Por lo tanto, este modelo puede considerarse válido para su explotación. Por este motivo, se adjuntan las gráficas de la distribución de sus residuos.

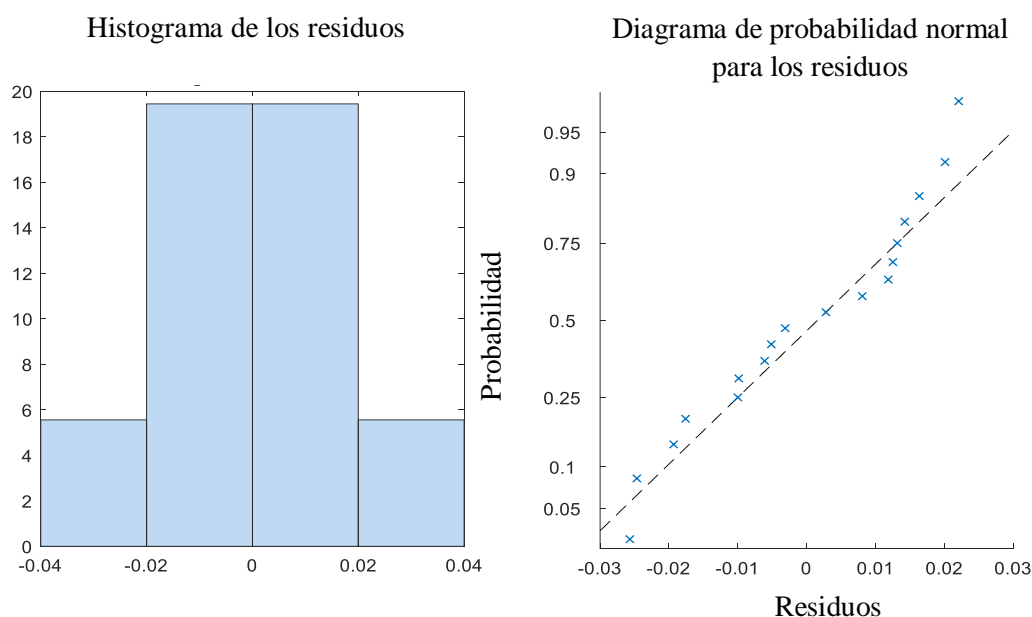


Figura 8: Distribución de residuos del modelo C.1.2.

Los residuos del modelo C.1.2 se distribuyen de manera simétrica y su distribución se ajusta a la de una distribución normal de media cero y varianza constante. En definitiva, el modelo C.1.2 puede considerarse válido.

4.4.1.2. Modelo C.2.

En este caso, se decidió sustituir el PIB por población y renta per cápita a la espera de mejores resultados en la correlación del modelo. Por otro lado, al separar riqueza y población la influencia de estas variables podría ser distinta sobre la variable térmica. Además

4.4.1.2.1. Modelo C.2.1.

En este caso, se ha mantenido la variable climática del modelo C y se ha sustituido el PIB por población y renta per cápita.

Poder explicativo del modelo				
F-test	351	P-valor (F-test)	2.1e-13	
R²	0.987	R² ajustado	0.984	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.0065		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	0.1033	0.0015	67.798	4.99e-19
CDD M.II.D.	-0.0016	0.0018	-0.8497	0.4098
Población	0.0454	0.0021	21.7450	3.45e-12
Renta per cápita	0.0110	0.0018	6.0588	2.9427

Tabla 37 : Resultados del modelo C.2.1.

Respecto al modelo anterior pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.990, por lo que la correlación entre los datos es extremadamente alta.
- Mientras que los CDD no pueden aceptarse para ningún nivel de significación en el modelo, el resto de variables tiene poder explicativo significativo al 1%.
- El coeficiente de los CDD es negativo, lo cual indica una relación inversa entre el consumo de aire acondicionado del sector residencial y los CDD carente de sentido.

Por lo tanto, este modelo no puede considerarse válido.

4.4.1.2.2. Otras variantes del modelo C.2.1.

Ante las mejoras de correlación en el modelo C.2.1. respecto al modelo C, se decidió también probar con el resto de variables climáticas disponibles. Pese a que el coeficiente de correlación ostentaba valores cercanos a 0.98 en todos los casos, las variables climáticas pasaron a tener un alto p-valor e incluso coeficientes negativos.

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
CDD M.I.D.	CDD M.I.D.	CDD M.I.D.
CDD ODYSSEE	CDD ODYSSEE	CDD ODYSSEE
CDD M.I.	CDD M.I.	-
CDD M.II.	CDD M.II.	-
T. med. ver.	T. med. ver	T. med. ver.
T. med. ver. D.	T. med. ver. D.	T. med. ver. D.

Tabla 38 : Principales fallos de las variantes del modelo C.2.1.

4.4.1.3. Modelo C.3.

Por último, ante los malos resultados de la variable climática para todos los casos, se decidió probar modelos sin dichas variables.

4.4.1.3.1. Modelo C.3.1.

Este modelo sólo recoge el efecto de la población y la renta per cápita en el consumo de aire acondicionado del sector residencial. Los resultados pueden observarse en la siguiente tabla:

Poder explicativo del modelo				
F-test	536	P-valor (F-test)	1.12e-14	
R²	0.986	R² ajustado	0.984	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.0064		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	0.1033	0.0015	68.4350	3.88e-20
Población	0.0445	0.0018	25.0460	1.18e-13
Renta per cápita	0.0111	0.0018	6.2676	1.51e-05

Tabla 39 : Resultados del modelo C.3.1.

De la tabla anterior pueden extraerse las siguientes conclusiones

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.990, por lo que la correlación entre los datos es extremadamente alta.
- Todas las variables tienen poder explicativo significativo al 1% y sus coeficientes son positivos.

Por lo tanto, el modelo resulta válido. En la Figura 9 se muestra la distribución de los residuos del modelo C.3.1.

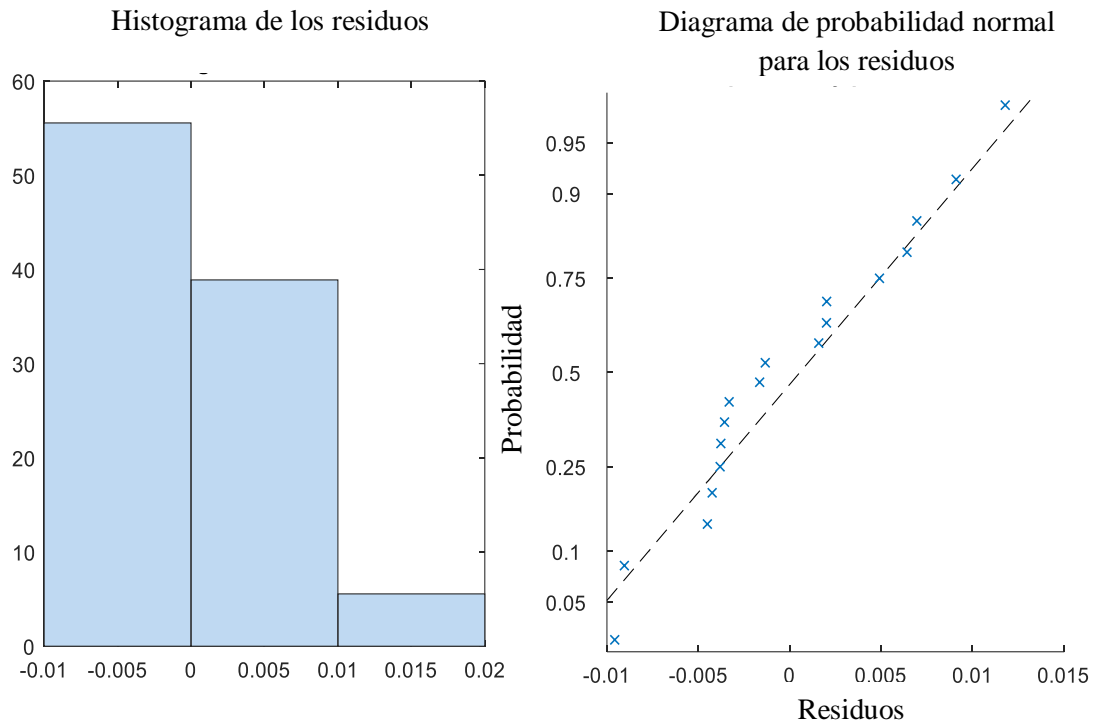


Figura 9 : Distribución de los residuos del modelo C.3.1.

Como la distribución de los modelos es prácticamente simétrica y se ajusta aproximadamente a una distribución normal, el modelo puede considerarse apto.

4.4.1.3.2. Modelo C.3.2.

En este último modelo sólo se recoge el efecto del PIB. Los resultados se muestran en la tabla a continuación:

Poder explicativo del modelo				
F-test	124	P-valor (F-test)	6.18e-09	
R²	0.885	R² ajustado	0.878	
Observaciones	18	Grados de libertad	16	
Error medio al cuadrado		0.0179		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	0.1033	0.0042	24.5250	4.04e-14
PIB	0.0482	0.0043	11.1160	6.18e-09

Tabla 40 : Resultados del modelo C.3.2.

De la tabla anterior pueden extraerse las siguientes conclusiones

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.88.
- Todas las variables tienen poder explicativo significativo al 1% y sus coeficientes son positivos.

Puesto que los parámetros y test de hipótesis dan pie a aceptar el modelo, en la Figura 10 se adjunta la distribución de sus residuos.

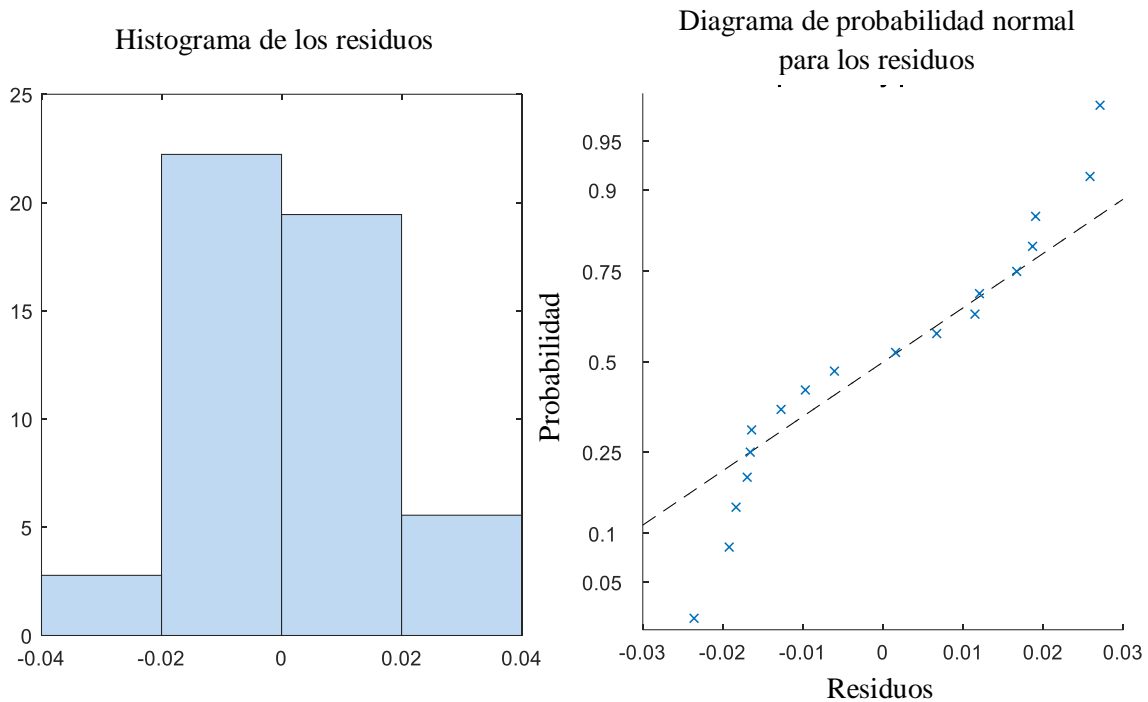


Figura 10 : Distribución de residuos del modelo C.3.2.

Puede observarse que, pese a que los residuos se distribuyen de forma cuasi simétrica, su distribución se aleja de la probabilidad normal en ambas colas. Por este motivo, el modelo no puede considerarse válido.

4.4.2. Resultados de los modelos de aire acondicionado del sector residencial

Considerando los resultados obtenidos sobre el consumo de aire acondicionado en el sector residencial, los únicos modelos aptos son el C.1.2, el C.3.1 y el C.3.2. El primero de ellos es significativo al 10% y un coeficiente de correlación ajustado de 0.89 mientras que los otros dos son significativos al 1% y poseen un coeficiente de correlación de 0.98 y 0.88 respectivamente.

Se puede concluir que el modelo que mejor explica este consumo es el C.3.1, el cual tiene como variables explicativas la población y la renta per cápita. La ausencia de variable climática en este modelo puede deberse a que el mercado de aire acondicionado está aún en fase de desarrollo, por lo que un incremento en el consumo del sector residencial en el aire acondicionado puede deberse a un incremento de las viviendas en que se utiliza, lo cual se explica por el incremento de la población y la riqueza.

En la siguiente tabla puede observarse el peso de las diferentes variables en el modelo:

Modelo	Población	Renta per cápita
C.3.1	80.00%	20.00%

Tabla 41 : Peso de las variables en el modelo C.3.1.

En conclusión, actualmente el consumo de aire acondicionado en España en el sector residencial no depende de las condiciones térmicas.

4.5. Consumo de aire acondicionado del sector servicios

En este apartado se exploran las distintas posibilidades para la elaboración de un modelo de regresión lineal del consumo de calefacción del sector residencial con las diferentes variables explicativas, y finalmente se escoge el modelo más apropiado.

4.5.1. Modelos del consumo de aire acondicionado del sector servicios

Para definir el consumo del aire acondicionado en el sector servicios se dispone de variables que dan cuenta del clima, de la población, de la riqueza y de las mejoras tecnológicas. Por lo tanto, en un principio se decidió construir un modelo que las incluyese a todas (Modelo D). Los resultados pueden observarse en la Tabla 42 : Resultados del modelo D. Tabla 42.

Poder explicativo del modelo				
F-test	40	P-valor (F-test)	4.05e-07	
R²	0.896	R² ajustado	0.873	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.114		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	1.8999	0.0269	70.7050	2.77e-19
CDD M.II.D.	0.0288	0.0324	0.8886	0.3892
PIB	0.3146	0.0547	5.7557	4.97e-05
ODEX	0.0279	0.0586	0.4753	0.6419

Tabla 42 : Resultados del modelo D.

Del modelo anterior pueden deducirse las siguientes conclusiones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.88.
- Tanto el ODEX como los CDD no pueden aceptarse para ningún nivel de significación habitual. Sin embargo, el PIB puede aceptarse a todos los niveles de significación.
- Todos los coeficientes son positivos.

Debido al bajo p-valor de las variables CDD y ODEX este modelo no puede considerarse válido.

4.2.2.1. Variaciones del modelo D sin tener en cuenta la variable tecnológica.

En este caso se decidió descartar la influencia de la variable tecnológica ODEX en el modelo. Pese a que los coeficientes de correlación de todos los modelos eran parecidos al caso base, en todos los casos la variable climática resultó tener un p-valor demasiado alto, tal y como se resume en la Tabla 43.

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
CDD M.II.D.	CDD M.II.D.	-
CDD M.I.D.	CDD M.I.D.	-
CDD ODYSSEE	CDD ODYSSEE	-
T. med. ver. D.	T. med. ver. D.	-
T. med. ver.	T. med. ver.	-

Tabla 43: Principales problemas de las variantes del modelo D en las cuales se eliminó el ODEX.

En definitiva, un modelo lineal que sólo incluya el PIB y una de las variables climáticas anteriores no puede considerarse apto.

4.2.2.2. Modelo D.1.

Este modelo relaciona el consumo de aire acondicionado del sector servicios única y exclusivamente con el PIB, ante los resultados obtenidos en los apartados anteriores. Los resultados se presentan en la siguiente tabla:

Poder explicativo del modelo				
F-test	50.4	P-valor (F-test)	9.46e-20	
R²	0.915	R² ajustado	0.897	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.103		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	1.8999	0.0242	78.5020	6.44e-20
PIB	0.1367	0.0367	3.7150	0.0023

Tabla 44 : Resultados del modelo D.1.

De la tabla anterior pueden extraerse las siguientes conclusiones

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.90, valor ligeramente superior al modelo D.
- El PIB tienen poder explicativo significativo al 1% y sus coeficientes son positivos.

Por lo tanto, el modelo resulta explicativo y se incluye la distribución de sus residuos en la Figura 11. Los residuos no se distribuyen de manera simétrica, sin embargo, su distribución se aproxima a una normal de media cero y varianza constante, por lo que el modelo puede considerarse apto.

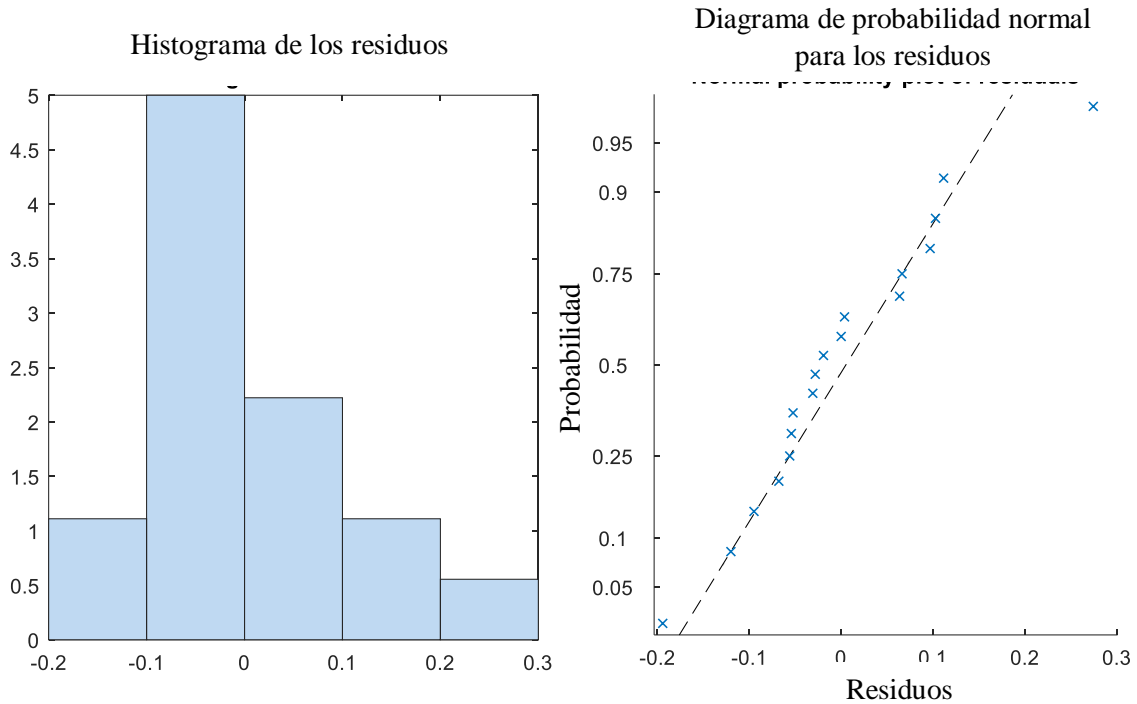


Figura 11: Distribución de residuos del modelo D.1.

4.2.2.3. Modelo D.2.

Para la concepción de este modelo se utilizaron las variables población y renta per cápita para dar cuenta del efecto del PIB por separado. Así mismo, se eliminó la variable tecnológica ya que su p-valor era el más elevado en el modelo D. En este apartado se muestran los resultados para la temperatura media desagregada en verano, y los resultados para el resto de casos pueden verse en el siguiente apartado.

Poder explicativo del modelo				
F-test	50.4	P-valor (F-test)	9.46e-20	
R²	0.915	R² ajustado	0.897	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.103		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	1.8999	0.0242	78.5020	6.44e-20
T. med. ver. D.	0.0590	0.0331	1.7834	0.0962
Población	0.1367	0.0367	3.7150	0.0023
Renta per cápita	0.1798	0.0287	6.2631	2.08e-05

Tabla 45 : Resultados del modelo D.2.

De la tabla anterior pueden extraerse las siguientes conclusiones

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.90, valor ligeramente superior al modelo D.

- La población y la renta per cápita tienen poder explicativo significativo al 1% y sus coeficientes son positivos.
- La temperatura media desagregada en los meses de verano puede ser aceptada a un nivel de significación del 10% y su coeficiente es positivo.

Por lo tanto, el modelo anterior podría considerarse válido a un nivel de significación del 10%. En la Figura 12 puede observarse la distribución de los residuos de dicho modelo. Puede concluirse que, pese a que los residuos no se distribuyen de manera simétrica, su distribución de probabilidad sí se aproxima a una normal, por lo que el modelo puede considerarse válido.

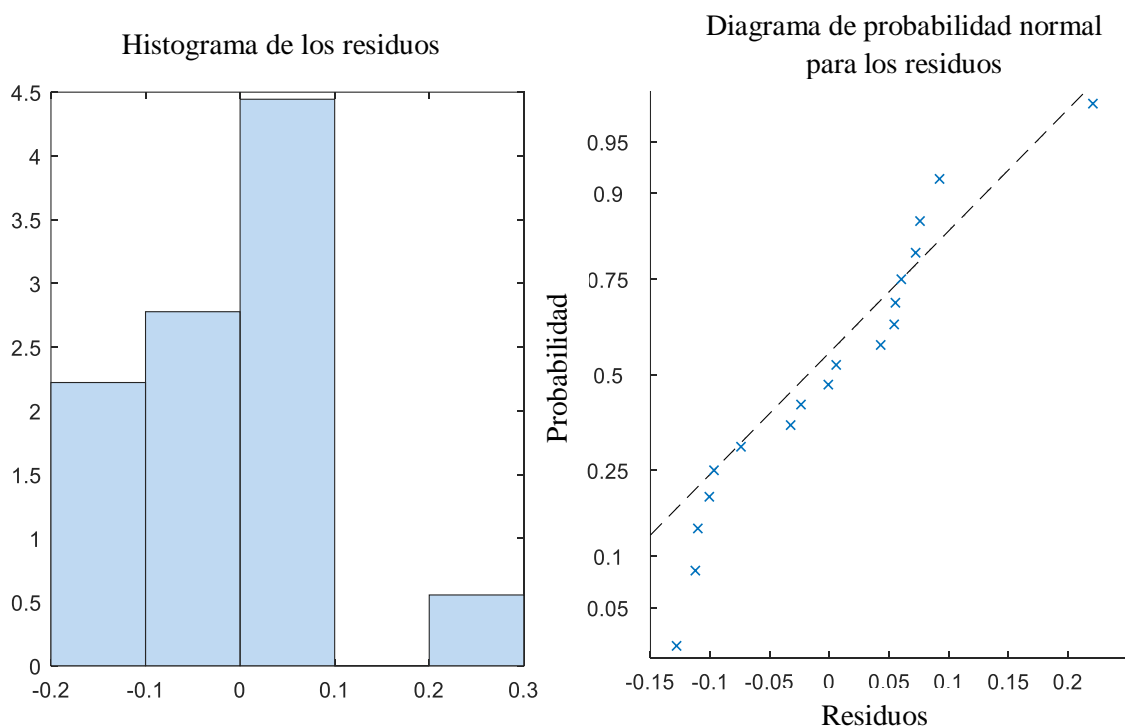


Figura 12: Distribución de los residuos del modelo D.2.

4.2.2.4. Otras variaciones del modelo D.2.

En este apartado se recogen los inconvenientes encontrados al sustituir los CDD desagregados método II por otra de las variables climáticas disponibles. Ninguna combinación resultó en un modelo satisfactorio, tal y como muestra la Tabla 46.

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
CDD M.II.D.	CDD M.II.D.	-
CDD M.I.D.	CDD M.I.D.	-
CDD ODYSSEE	CDD ODYSSEE	-
CDD. M.I.	CDD M.I.	CDD. M.I.
CDD M.II.	CDD M.II	CDD M.II.
T. med. ver.	T. med. ver.	T. med. ver.

Tabla 46: Principales inconvenientes de intercambiar la variable climática del modelo D.2.

4.2.2.4.1. Modelo D.2.1.

Ante los resultados obtenidos en los apartados 4.2.2.3 (*Modelo D.2.*) y 4.2.2.4. (*Otras variaciones del modelo D.2.*), se decidió eliminar así mismo la variable climática del modelo D.2. Los resultados pueden observarse en la siguiente tabla:

Poder explicativo del modelo				
F-test	64.7	P-valor (F-test)	4.23e-08	
R²	0.896	R² ajustado	0.882	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.110		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	1.8999	0.0259	73.3510	1.37e-20
Población	0.1782	0.0305	5.8460	3.21e-05
Renta per cápita	0.1734	0.0305	5.6874	4.31e-05

Tabla 47 : Resultados del modelo D.2.1.

Pueden realizarse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.90, valor ligeramente superior al modelo D.
- Tanto la población como la renta per cápita tienen poder explicativo significativo al 1% y sus coeficientes son positivos.

Por lo tanto, en principio el modelo D.2.1. puede considerarse apto para su explotación. En la Figura 13 se representan los residuos de este modelo. Puede deducirse que, pese a que no se distribuyen de manera simétrica, su distribución de probabilidad se aproxima a una distribución normal, por lo que el modelo puede aceptarse.

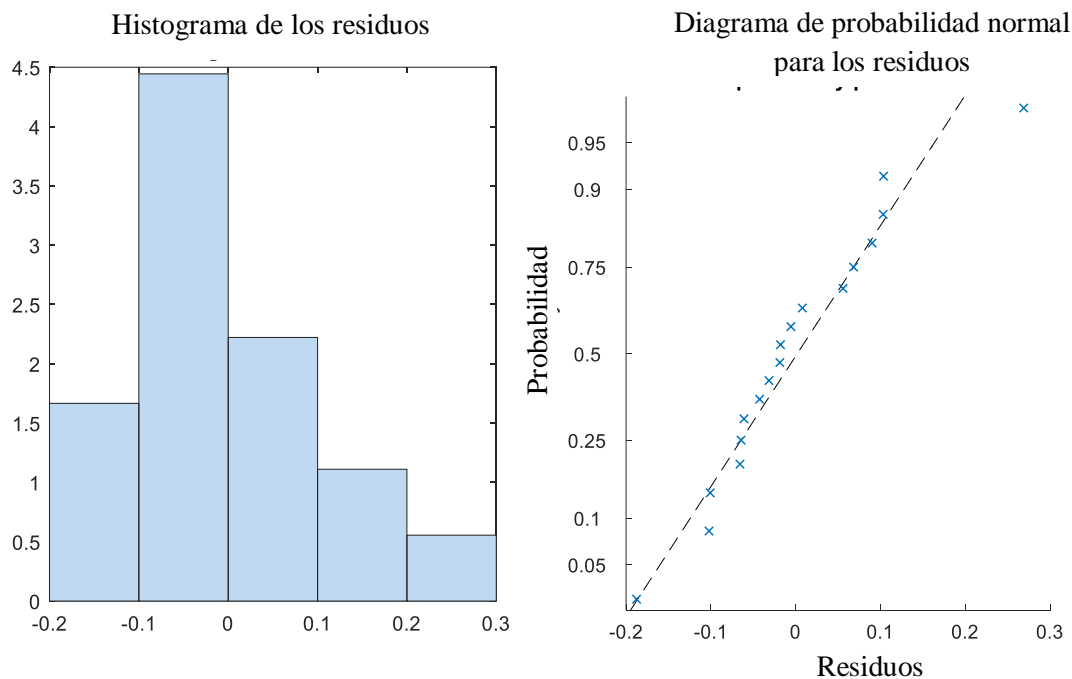


Figura 13 : Representación de los residuos del modelo D.2.1.

4.2.2.5. Modelos sin variable climática pero con variable tecnológica.

En este apartado se comprueba el efecto negativo para los modelos D.1 y D.2.1 de añadir la variable de mejoras tecnológicas. Los principales problemas pueden hallarse en la tabla a continuación.

Variable considerada	P-valor alto (>10%)	Coefficientes negativos
Renta per cápita y población	ODEX	-
PIB	ODEX	

Tabla 48 : Problemas de añadir la variable ODEX a los modelos D.1 y D.2.1.

4.5.2. Resultados de los modelos de aire acondicionado del sector servicios

Los modelos que mejor describen el consumo de aire acondicionado del sector servicios son el D.1, el D.2 y el D.2.1. El primero es significativo al 1% y tiene un coeficiente de correlación ajustado del 0.89; el segundo es significativo al 10% y tiene un coeficiente de correlación de 0.89 y el tercero es significativo al 1% con un coeficiente de correlación ajustado de 0.88. Es evidente que el modelo que representa más fielmente es bien el D.1 o el D.2.1.

El modelo D.1 explica el consumo de aire acondicionado del sector servicios sólo a partir del PIB mientras que el D.2.1 lo hace a partir de la población y de la renta per cápita. En ambos, los residuos se ajustan muy bien a la distribución normal y la diferencia en la correlación es prácticamente inexistente. Además, ambos tienen en cuenta los mismos factores: población y riqueza. Se ha decidido que el modelo más apropiado es el D.1, ya que el PIB es mejor indicador de actividad económica. Por otro lado, ese modelo tiene más grados de libertad y un coeficiente ligeramente más elevado.

En definitiva, las variaciones en consumo de aire acondicionado por el sector servicios son explicadas en su totalidad por las variaciones en el PIB, el cual mide la actividad

económica, y con ella, el crecimiento en número de oficinas que pueden utilizar tales dispositivos.

No existe en este caso correlación con el clima. Esto puede deberse, al igual que en el sector residencial, a que el aire acondicionado constituye un mercado en crecimiento, y por lo tanto un incremento en el consumo puede deberse a la adquisición de nuevos equipos por diferentes negocios.

4.6. Datos corregidos según las condiciones climáticas de ODYSSEE

4.6.1. Modelos del consumo de calefacción corregido del sector residencial

En este apartado se ha tratado de explicar un conjunto de datos facilitados por ODYSSEE que describen el consumo total en calefacción del sector residencial en España corregido según la situación climática. La

Poder explicativo del modelo				
F-test	13.4	P-valor (F-test)	0.0002	
R²	0.742	R² ajustado	0.687	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.460		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	± 0.1084	63.7100	1.1894e-18
Población	0.4649	± 0.2916	1.5943	0.1332
Renta per cápita	0.4721	± 0.1652	2.8573	0.0127
ODEX	0.2552	± 0.2559	0.9976	0.3354

Tabla 49 muestra los resultados para las otras tres variables principales (modelo E).

Poder explicativo del modelo				
F-test	13.4	P-valor (F-test)	0.0002	
R²	0.742	R² ajustado	0.687	
Observaciones	18	Grados de libertad	13	
Error medio al cuadrado		0.460		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	± 0.1084	63.7100	1.1894e-18
Población	0.4649	± 0.2916	1.5943	0.1332
Renta per cápita	0.4721	± 0.1652	2.8573	0.0127
ODEX	0.2552	± 0.2559	0.9976	0.3354

Tabla 49 : Resultados del modelo E.

De la tabla anterior puede decirse lo siguiente:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.7.
- La única variable con poder explicativo aceptable para los niveles de significación convencionales es la Renta per cápita.
- Todos los coeficientes son positivos, luego un incremento en cualquiera de las variables provoca un incremento en el consumo de calefacción en el sector residencial.

De todo lo anterior puede deducirse que el modelo E no es apropiado. Para obtener un modelo con p-valores más bajos asociados a todas las variables, en un principio se trató de eliminar la variable ODEX, ya que presenta el p-valor más alto.

4.6.1.1. Modelo E.1.

La

Poder explicativo del modelo				
F-test	19.6	P-valor (F-test)	6.48e-05	
R²	0.724	R² ajustado	0.687	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.460		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	±0.1083	63.7200	1.1269e-19
Población	0.5769	±0.1275	4.5237	0.0004
Renta per cápita	0.2033	±0.1275	1.5944	0.1317

Tabla 50 contiene los resultados tras haber eliminado la variable ODEX del modelo E.

Poder explicativo del modelo				
F-test	19.6	P-valor (F-test)	6.48e-05	
R²	0.724	R² ajustado	0.687	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.460		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	±0.1083	63.7200	1.1269e-19
Población	0.5769	±0.1275	4.5237	0.0004
Renta per cápita	0.2033	±0.1275	1.5944	0.1317

Tabla 50 : Resultados del modelo E.1.

Puede afirmarse que:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.

- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.55.
- La variable de la Renta per Cápita no puede aceptarse para ningún nivel de significación convencional. Por el contrario, la variable de población puede aceptarse para todos los niveles de significación convencionales.
- Todos los coeficientes son positivos, luego un incremento en cualquiera de las variables provoca un incremento en el consumo de calefacción en el sector residencial.

Por lo tanto, el modelo no es válido. Se decidió dar un paso más y eliminar la variable de renta per cápita.

4.6.1.1.1. Modelo E.1.1.

La

Poder explicativo del modelo				
F-test	8.49	P-valor (F-test)	0.0102	
R²	0.347	R² ajustado	0.306	
Observaciones	18	Grados de libertad	16	
Error medio al cuadrado		0.684		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	±0.1612	42.8000	6.2289e-18
Población	0.4834	±0.1660	2.9130	0.0102

Tabla 51 muestra los resultados del modelo E.1 tras eliminar la variable de la renta per cápita (modelo E.1.1):

Poder explicativo del modelo				
F-test	8.49	P-valor (F-test)	0.0102	
R²	0.347	R² ajustado	0.306	
Observaciones	18	Grados de libertad	16	
Error medio al cuadrado		0.684		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	±0.1612	42.8000	6.2289e-18
Población	0.4834	±0.1660	2.9130	0.0102

Tabla 51 : Resultados del modelo E.1.1.

Las siguientes afirmaciones pueden extraerse de la tabla anterior:

- El p-valor del F-test indica que el modelo es explicativo para un nivel de significación del 5%.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.35.
- Todos los coeficientes son positivos, luego un incremento en cualquiera de las variables provoca un incremento en el consumo de calefacción en el sector residencial.
- Todas las variables tienen poder explicativo a un nivel de significación del 5%.

Pese a que los p-valores son apropiados, el modelo no se considera apto debido a los bajos valores de los coeficientes de correlación y correlación ajustada y el alto valor del error medio.

4.6.1.2. Modelo E.2.

La muestra los resultados del modelo B tras eliminar la variable de la renta per cápita y población y sustituirlo por el PIB, que engloba ambas (modelo E.2.):

Poder explicativo del modelo				
F-test	33.5	P-valor (F-test)	2.79e-20	
R²	0.677	R² ajustado	0.656	
Observaciones	18	Grados de libertad	16	
Error medio al cuadrado		0.481		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	± 0.1135	60.8380	2.3181e-20
PIB	0.6755	± 0.1168	5.7856	2.7861e-05

Tabla 52 : Resultados del modelo E.2.

Las siguientes afirmaciones pueden extraerse de la tabla anterior:

- El p-valor del F-test indica que el modelo es explicativo para un nivel de significación del 1%.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.66.
- El coeficiente del PIB es positivo, luego un incremento en cualquiera de las variables provoca un incremento en el consumo de calefacción en el sector residencial.
- Todas las variables tienen poder explicativo a un nivel de significación del 1%.

De todo lo anterior, puede deducirse que el modelo es apto.

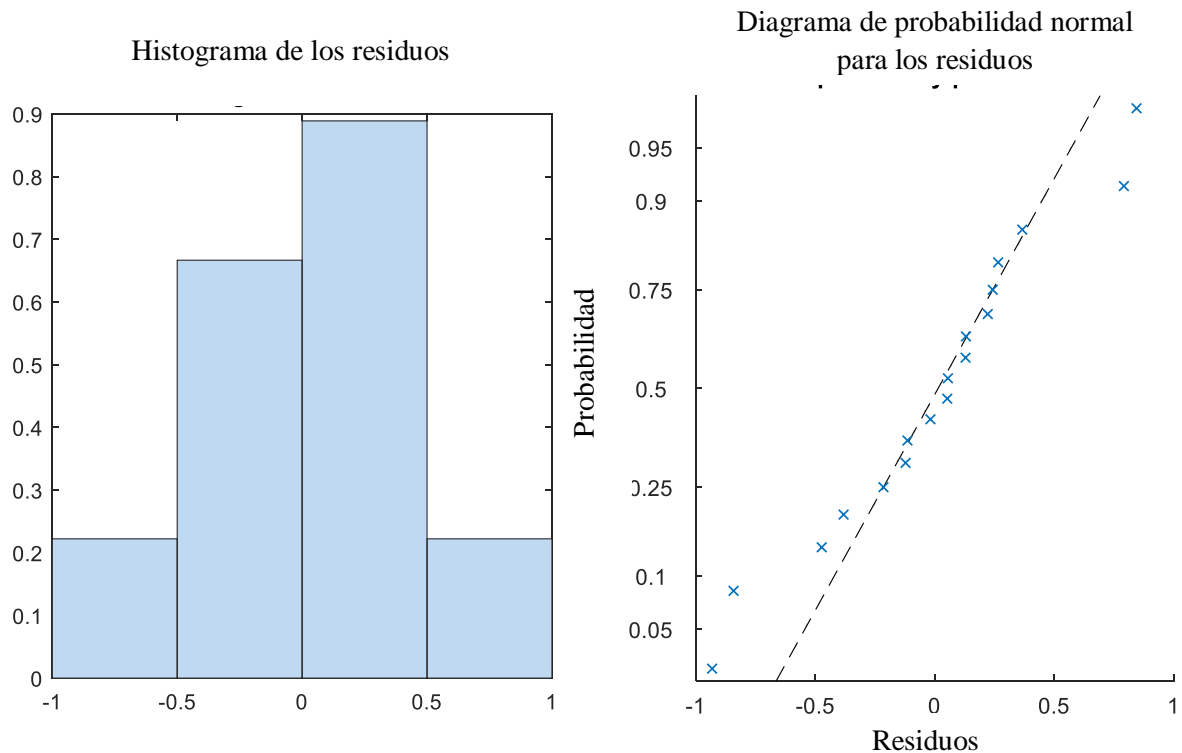


Figura 14 : Distribución de los residuos del modelo E.2.

Análisis de algunos indicadores de sostenibilidad en España

Los residuos se distribuyen de manera aproximadamente simétrica respecto a la media, y se ajustan a una distribución normal de media cero y varianza constante. Puede concluirse que el modelo E.2 es un modelo válido para su explotación.

4.6.1.3. Modelo E.3.

Este modelo está basado en el anterior (modelo E.2) pero incluyendo la variable del ODEX. Los resultados se presentan en la siguiente tabla:

Poder explicativo del modelo				
F-test	22.4	P-valor (F-test)	3.09e-05	
R²	0.750	R² ajustado	0.716	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.438		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	6.9031	0.1031	66.9410	5.39e-20
PIB	0.8141	0.1251	6.5070	9.90e-06
ODEX	0.2616	0.1251	2.0908	0.0540

Tabla 53 : Resultados del modelo E.3.

Las siguientes afirmaciones pueden extraerse de la tabla anterior:

- El p-valor del F-test indica que el modelo es explicativo para un nivel de significación del 1%.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.72.
- El coeficiente del PIB es positivo, luego un incremento en cualquiera de las variables provoca un incremento en el consumo de calefacción en el sector residencial.
- Todas las variables tienen poder explicativo a un nivel de significación del 10%.

De todo lo anterior, puede deducirse que el modelo es apto.

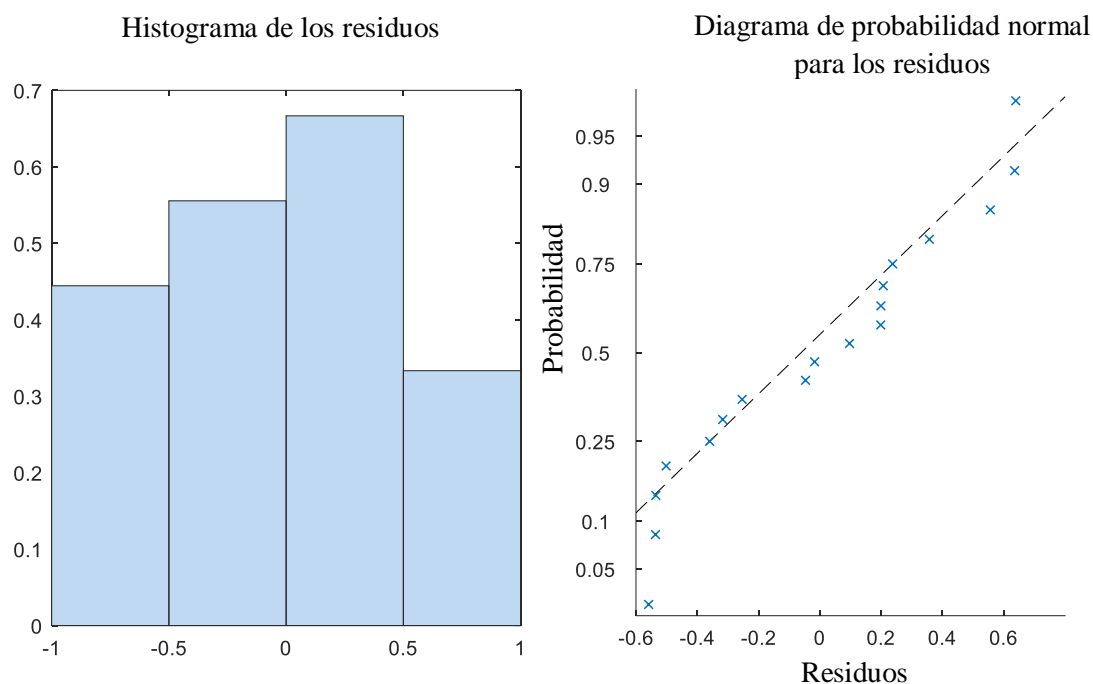


Figura 15: Distribución de los residuos del modelo E.3.

Los residuos se distribuyen de manera aproximadamente simétrica respecto a la media, y se ajustan a una distribución normal de media cero y varianza constante. Puede concluirse que el modelo E.2 es un modelo válido para su explotación.

4.6.2. Resultados del consumo de calefacción corregido del sector residencial

Los únicos modelos explicativos de este conjunto de datos son el E.2 y el E.3. El primero sólo tiene el PIB como variable explicativa y puede ser aceptado con un nivel de significación del 5%, sin embargo, presenta una correlación de 0.66. El segundo tiene como variables explicativas el PIB y el ODEX, sin embargo sólo puede ser aceptado a un nivel de significación del 10%, pero la correlación entre sus variables es más alta (0.72). Según los criterios enunciados anteriormente, el modelo más apropiado sería el E.2.

Estos modelos tienen las mismas variables explicativas que dos de los desarrollados en este proyecto (a excepción de la variable climática). Se trata del modelo A.2.3, explicado en el apartado 4.2.1.2.3 tiene como variables explicativas la variable climática y el PIB y el A.2.4, desarrollado en el apartado 4.2.1.2.4 tiene, además de las anteriores, el ODEX. Sin embargo, las variables del modelo más apropiado de la serie corregida por ODYSSEE no coinciden con las del modelo más apropiado de la serie de calefacción. Esto puede ser porque los resultados de los HDD de ODYSSEE no coinciden con los de los modelos A.2.3 y A.2.4, que son los calculados según el método II y desagregados por la población.

Puede deducirse que la ecuación $H = Hn \cdot \frac{HDD}{HDDn}$ (18) puede ayudar a determinar la predicción del consumo de calefacción del sector residencial si el año hubiera sido normal, ya que los resultados de los modelos de

regresión de la serie corregida con dicha ecuación es coherente con varios modelos aptos desarrollados en este documento.

4.6.3. Modelos del consumo de calefacción corregido del sector servicios

En este apartado se trata de explicar el consumo del sector servicios corregido según la componente climática. En la Tabla 54 se muestran los resultados para las variables población, renta per cápita y ODEX (modelo F).

Poder explicativo del modelo				
F-test	10.7	P-valor (F-test)	0.0006	
R ²	0.697	R ² ajustado	0.632	
Observaciones	18	Grados de libertad	14	
Error medio al cuadrado		0.364		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4320	0.0859	28.3120	9.271e-14
Población	0.2309	0.8542	0.2703	0.7909
Renta per cápita	-0.1068	0.1011	-1.0556	0.3091
ODEX	-0.3141	0.8520	-0.3687	0.7179

Tabla 54: Resultados del modelo F.

Respecto al modelo anterior pueden hacerse las siguientes observaciones:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo el valor del coeficiente ajustado inferior a 0.64.
- El p-valor del poder explicativo de todas las variables es demasiado elevado, no pudiendo aceptarse ninguna para los niveles de significación habituales.
- Los coeficientes del ODEX y de la Renta per cápita son negativos, lo que indica una relación inversa entre estas variables y el consumo de calefacción del sector servicios.

Puede deducirse que el modelo es explicativo, pero el poder de las variables se solapa y se distorsiona, por lo que aparecen p- y coeficientes negativos (que carecen de sentido) en algunas.

Como posibles caminos, se ha escogido sustituir las variables de población y renta per cápita por el PIB (modelo F.1) y eliminar algunas de las variables para evitar solapamientos (modelo F.2).

4.6.3.1. Modelo F.1

Este modelo busca comprobar la relación entre PIB, ODEX y el consumo del sector servicios corregido por el clima.

Poder explicativo del modelo				
F-test	17.0	P-valor (F-test)	0.0001	
R ²	0.694	R ² ajustado	0.653	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.354		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4320	0.0834	29.1720	1.256e-14
PIB	-0.1756	0.1685	-1.0423	0.3138
ODEX	-0.6435	0.1685	-3.8192	0.0017

Tabla 55: Resultados del modelo F.1.

De los resultados contenidos en la Tabla 55 se puede concluir lo siguiente:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo ligeramente superiores que en el modelo F.
- El p-valor del poder explicativo del PIB es demasiado elevado, no pudiendo aceptarse ninguna para los niveles de significación habituales.
- Ambos coeficientes son negativos.

Por lo tanto, este modelo no puede aceptarse como válido.

4.6.3.2. Modelo F.2

En este caso los modelos buscan una mayor adecuación mediante la eliminación de una de las variables. La siguiente tabla muestra los resultados para la población y el ODEX.

Poder explicativo del modelo				
F-test	17.0	P-valor (F-test)	0.0001	
R ²	0.694	R ² ajustado	0.653	
Observaciones	18	Grados de libertad	15	
Error medio al cuadrado		0.354		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4320	0.0834	29.1720	1.256e-14
Población	-0.1756	0.1685	-1.0423	0.3138
ODEX	-0.6435	0.1685	-3.8192	0.0017

Tabla 56: Resultados del modelo F.2.

De la tabla anterior puede concluirse lo siguiente:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.

- Los coeficientes de correlación y correlación ajustada difieren, siendo ligeramente superiores que en el modelo F.
- El p-valor del poder explicativo de la población es demasiado elevado, no pudiendo aceptarse ninguna para los niveles de significación habituales.
- Ambos coeficientes son negativos.

Por lo tanto, este modelo no puede aceptarse como válido.

4.6.3.3. Modelo F.2.1.

El modelo F.2.1 estudia la correlación entre la población y el consumo de calefacción del sector servicios corregido por las condiciones climáticas. Los resultados se muestran en la siguiente tabla:

Poder explicativo del modelo				
F-test	32.4	P-valor (F-test)	3.3e-05	
R ²	0.669	R ² ajustado	0.649	
Observaciones	18	Grados de libertad	16	
Error medio al cuadrado		0.356		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4320	0.0839	28.9870	2.94e-15
Población	0.4915	0.0863	5.6926	3.33e-05

Tabla 57: Resultados del modelo F.2.1.

Sobre este modelo puede concluirse lo siguiente:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada difieren, siendo ligeramente superiores a los del modelo F.
- Los p-valores individuales pueden aceptarse a cualquier nivel de significación.
- La correlación es positiva.

Por lo tanto, este modelo puede considerarse apto para su explotación, por lo que se adjunta así mismo la distribución de sus residuos en la Figura 16:

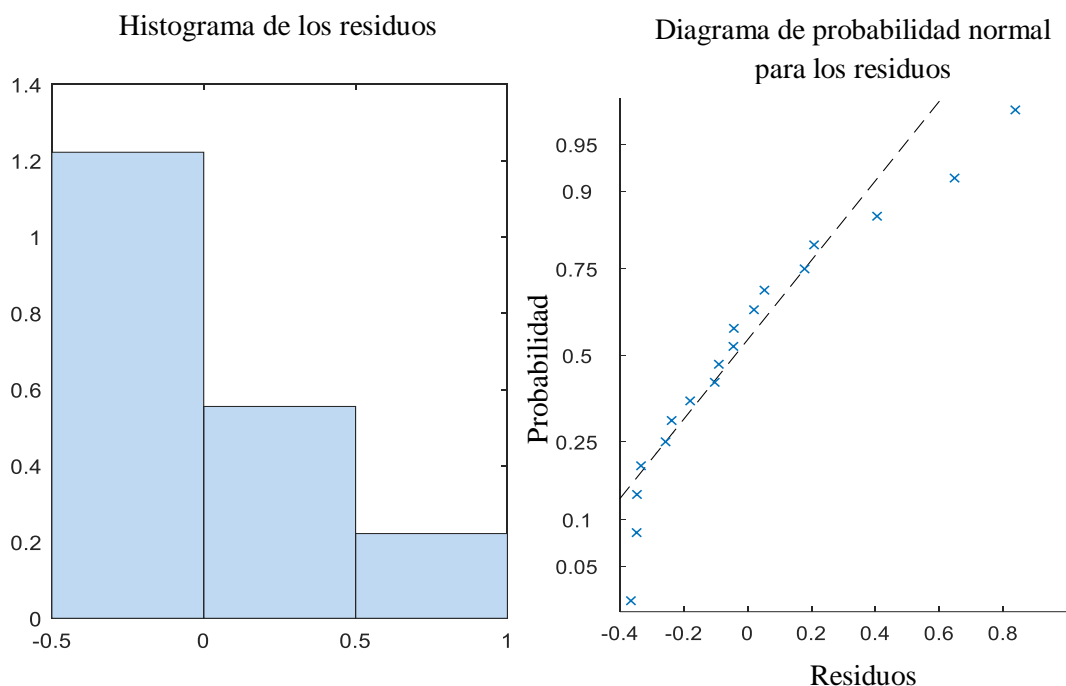


Figura 16 : Distribución de los residuos del modelo F.2.1.

Los residuos en este caso no puede considerarse que se distribuyan de manera simétrica, y difieren de una distribución normal significativamente en ambas colas. Luego este modelo tampoco debería considerarse válido.

4.6.3.4. Modelo F.2.2.

El modelo F.2.2 estudia la correlación entre la población y el consumo de calefacción del sector servicios corregido por las condiciones climáticas. Los resultados se muestran en la Tabla 58:

Poder explicativo del modelo				
F-test	10.4	P-valor (F-test)	0.0051	
R ²	0.397	R ² ajustado	0.359	
Observaciones	18	Grados de libertad	16	
Error medio al cuadrado		0.481		
Contribución de cada variable				
Variable	Estimación	Intervalo de confianza	t-test	p-valor (t-test)
β_0	2.4320	0.1134	21.4530	3.23e-13
PIB	0.3782	0.1166	3.2423	0.0051

Tabla 58 : Resultados del modelo F.2.2.

De la tabla anterior puede concluirse lo siguiente:

- El p-valor del F-test indica que el modelo tiene un gran poder explicativo.
- Los coeficientes de correlación y correlación ajustada son inferiores a 0.5.
- Los p-valores individuales de las variables pueden ser aceptados a cualquier nivel de significación.

- La correlación es positiva.

Debido a la baja correlación, este modelo no puede aceptarse como válido.

4.6.4. Resultados del consumo de calefacción corregido del sector servicios

En el caso del sector servicios no pudo encontrarse un modelo apropiado para explicar el consumo de calefacción corregido por la componente climática. Esto puede deberse a que dichos datos se obtienen con la ecuación $H = Hn \cdot \frac{HDD}{HDDn}$ (18), la cual incluye cálculos utilizando HDD, los cuales no aportan información sobre el consumo de calefacción del sector servicios.

4.7. Conclusiones sobre la demanda térmica en España

4.7.1. Conclusiones sobre el consumo de calefacción

En este apartado se detallan diversas razones por las que se considera posible que no exista una relación entre el consumo de calefacción del sector servicios y la temperatura exterior.

En primer lugar, resultó que ninguna de las variables climáticas más verosímiles para cuantificar el efecto de la temperatura exterior en el consumo de calefacción fue relevante para el sector servicios. Se considera que los HDD son la variable más verosímil porque incluye una temperatura de confort interna y la diferencia de la temperatura exterior con esta. De hecho, en el cálculo de la temperatura media en los meses fríos, estos se definen como los meses cuyos HDD calculados según el método I son distintos de cero.

En segundo lugar, algunos de los modelos de regresión realizados sobre la serie corregida de ODYSSEE del sector residencial (el E.2 y el E.3) resultaron ser coherentes con dos de los modelos aptos explicados en este proyecto (el A.2.3 y el A.2.4). Sin embargo, no pudo extraerse ningún modelo válido de la serie corregida de ODYSSEE del sector servicios.

En el Gráfico 20 se compara la tendencia para cada año del consumo real en calefacción en comparación al consumo corregido para ambos sectores. Los picos en sentido positivo representan puntos en los que el consumo real supera el corregido y el contrario para los picos en sentido negativo.

Comparación de la diferencia entre las series reales y las corregidas

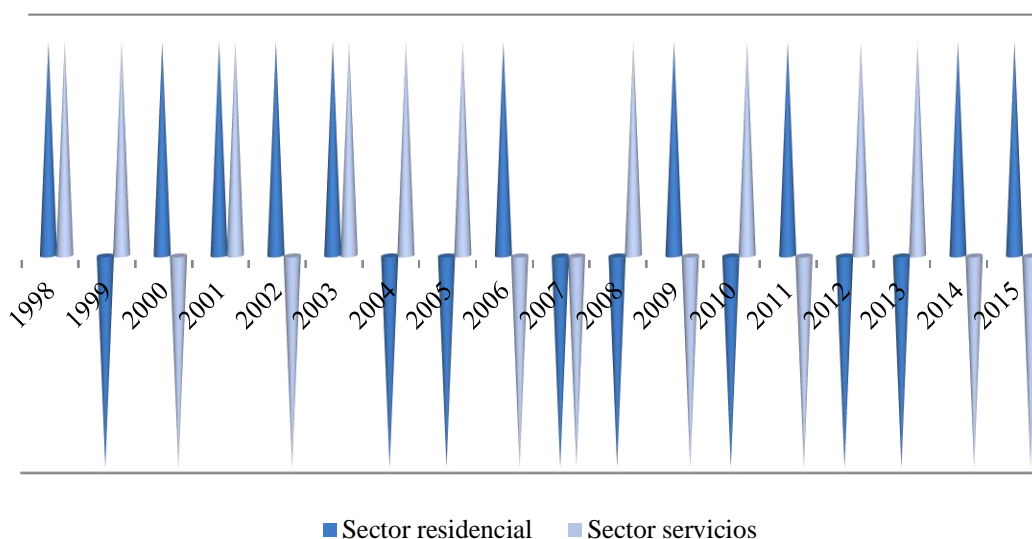


Gráfico 20: Comparación de la diferencia entre componente real y corregida para ambos sectores.

Puede observarse que el sentido de ambas diferencias raramente coincide para los sectores residencial y servicios. Es decir, en la mayoría de los años la componente corregida para ambos sectores da respuestas opuestas al consumo real. Este hecho puede significar que los modelos son incoherentes, ya que uno predice que ese mismo año el consumo de un sector debe aumentar mientras el del otro sector debe disminuir. Si bien es cierto que los dos sectores son distintos, si los modelos son correctos y las condiciones climáticas las mismas el comportamiento no debería variar de manera opuesta durante un año normal.

Entonces, podría deducirse que el consumo de calefacción del sector servicios no está relacionado con la temperatura exterior. Si este fuera el caso, entonces España estaría perdiendo oportunidades de eficiencia energética.

4.7.2. Conclusiones sobre el uso de aire acondicionado

Respecto al consumo de aire acondicionado, la falta de modelos explicativos tanto del sector residencial como del sector servicios puede ser consecuencia de que el mercado de aire acondicionado no está aún saturado, y por lo tanto, el incremento en consumo de aire acondicionado de un año puede deberse a la adquisición de equipos de aire acondicionado por ambos sectores ese año, pero no a un incremento de la temperatura exterior respecto al año anterior.

5. Resultados

En este proyecto se ha determinado tanto el producible hidráulico normal en el total del territorio español como el consumo de los sectores residencial y servicios durante un año corriente en España.

Se ha determinado que el producible hidráulico sigue una distribución log-normal $LN(10,1198; 0,3041)$, y que por lo tanto la cantidad de producible hidráulico en España durante un año típico corresponde con la esperanza de tal distribución, es decir, 26.004 GWh. Este resultado se obtuvo mediante inferencia de parámetros sobre una muestra del producible hidráulico anual español, seguido de contrastes de hipótesis de bondad de ajuste tales como el test de Kolmogorov-Smirnov o el estadístico Anderson-Darling. Sin embargo, el valor obtenido depende de la muestra, la cual de haber sido mayor habría sido más robusto, al haberse podido aplicar además el test Chi-cuadrado de bondad de ajuste.

Para el consumo de calefacción del sector residencial se ha establecido un modelo de regresión lineal que depende de variables climáticas, de riqueza, demográficas y tecnológicas. Este modelo fue aceptado a un 5% de significación y posee un coeficiente de correlación de 0,86 y uno de correlación ajustada de 0,83. Las variables en cuestión fueron los HDD M.II.D., el ODEX residencial y el PIB a precios constantes.

Se trata de un modelo robusto no sólo por los resultados de los diferentes contrastes de hipótesis aplicados sobre él, sino por la coherencia de las variables escogidas y por la comparación de este modelo con otras series corregidas según el efecto del clima. Los datos del consumo del sector residencial para un año típico pueden determinarse sustrayendo la componente climática de dicho modelo.

En el caso del consumo del sector servicios también se ha determinado un modelo de regresión lineal. Éste fue aceptado a un 1% de significación y poseía un coeficiente de correlación de 0,78 y de correlación ajustado de 0,73. El modelo incluye solamente una variable climática, la temperatura media en invierno; y el PIB.

Sin embargo, existen motivos para creer que la demanda térmica del sector servicios no posee una relación directa con la temperatura exterior, ya que las variables más verosímiles para cuantificar el efecto de ésta no resultaron útiles para este sector, tampoco se encontraron modelos de otras series corregidas compatibles y por último la cantidad corregida en este proyecto no sigue la misma tendencia respecto a los datos originales que la del sector residencial. Si esto fuese correcto, el sector servicios podría incrementar su eficiencia energética con un uso de los sistemas de calefacción más responsable.

Respecto al aire acondicionado, no se pudieron obtener modelos de regresión lineal concluyentes que incluyesen una variable climática. El motivo más verosímil puede ser que el mercado de aire acondicionado aún se encuentra en expansión, por lo que familias y empresas continúan adquiriendo equipos cada año, lo que provoca el incremento del consumo.

6. Limitaciones y futuras consideraciones

En el análisis de las variables aleatorias que representan la meteorología, se ha considerado que se trata de variables independientes con una distribución desconocida inmutable. Si se valora esta afirmación, se están ignorando los posibles efectos del cambio climático en la forma de esta distribución.

El cambio climático altera el nivel de precipitaciones y temperatura exterior hacia valores más extremos según la propensión de las regiones. Por ejemplo, en España el cambio climático provoca más precipitaciones de carácter torrencial en el litoral mediterráneo mientras acentúa la sequía en zonas del interior peninsular [SEVI14]. Por otro lado, para realizar predicciones meteorológicas futuras, la Asociación Española de Meteorología (AEMET) trabaja en colaboración con entidades nacionales e internacionales para desarrollar modelos a nivel regional que cuantifiquen los futuros cambios en el clima con un determinado grado de incertidumbre [AEMEsf].

Por ese motivo, la distribución de densidad de probabilidad del producible hidráulico podrá sufrir cambios en el futuro debido a este factor. Además, el resultado obtenido de los contrastes de bondad de ajuste no fue obtenido rechazando la hipótesis nula, por lo que cualquier otra distribución sobre la cual se realicen los test de bondad de ajuste y no pueda rechazarse la hipótesis nula también podría constituir una distribución de densidad de probabilidad válido.

Existen otras fuentes de energía renovables cuyo uso en un año típico puede ser estudiado, tales como la energía eólica y la solar. El principal problema de estudio de la energía eólica puede ser que, según el modelo del generador, las velocidades idóneas del viento cambiarán de unas granjas eólicas a otras. Además, sólo interesa la velocidad del viento en dichos puntos de España. Así mismo, para la energía solar sería preciso estudiar la radiación media así como las horas de cielo despejado. Sin embargo, el estudio de estas variables también podría estar sujeto al cambio climático.

Finalmente, podría estudiarse también el consumo de uso térmico del sector industrial, que es el que más energía consume.

7. Conclusiones

Del proyecto pueden extraerse las siguientes conclusiones:

- El producible hidráulico español durante un año representativo es de 26.004 GWh.
- El modelo de regresión lineal múltiple del uso de calefacción del sector residencial se considera un modelo robusto y coherente, tanto por las pruebas de inferencia paramétrica como por la coherencia de sus variables explicativas y con otra serie de datos corregida.
- Según ese modelo, la demanda de calefacción del sector residencial se explica por el PIB a precios constantes, el ODEX residencial y los HDD desagregados calculados según el método II. Los datos corregidos según las condiciones climáticas pueden calcularse según la ecuación $Y^* = Y - 0,2164 \cdot X_1^r$ (29), donde Y^* corresponde a la serie corregida en mtep, Y a la serie original en mtep y X_1^r a los HDD desagregados calculados según el método II normalizados:

$$Y^* = Y - 0,2164 \cdot X_1^r \quad (29)$$

- El modelo de regresión lineal múltiple de la demanda térmica del sector servicios se considera menos robusto debido a las variables que emplea y la falta de coherencia con la serie corregida del sector residencial. Con ello se pretende plantear una posible falta de correlación entre el uso de calefacción del sector residencial y la temperatura exterior.
- De acuerdo al modelo, la demanda de calefacción del sector servicios se explica por la temperatura media en invierno y el PIB a precios constantes. La serie corregida según las condiciones climáticas puede hallarse de acuerdo a la ecuación $Y^* = Y - 0,2315 \cdot X_1^s$ (30). En ella, Y^* es la serie corregida en mtep, Y la original también en mtep y X_1^s hace referencia a la temperatura media en invierno normalizada.

$$Y^* = Y - 0,2315 \cdot X_1^s \quad (30)$$

- No se ha obtenido ningún modelo satisfactorio que relacione la temperatura exterior con el consumo de aire acondicionado tanto del sector residencial como del sector servicios. Este hecho puede ser consecuencia de que el mercado de aire acondicionado aún no se encuentra saturado.

8. Bibliografía

- [AEME18] Asociación Española de Meteorología (AEMET), 2018.
Obtenido por cortesía de AEMET.
- [AEMEsf] Asociación Española de Meteorología (AEMET), “Generación de escenarios regionalizados del cambio climático”, s.f.
Retirado de http://www.aemet.es/es/idi/clima/escenarios_CC en julio de 2018.
- [BANCSf] Banco Mundial, “PIB (UMN a precios constantes)”, s.f.
Retirados datos entre 1998 y 2017 en enero de 2018 de <https://datos.bancomundial.org/indicador/NY.GDP.MKTP.KN?end=2016&locations=ES&start=2002>
- [CARO13] Caro R., J. Reneses, “Estadística. No soy mala... me han dibujado así”. Thomson Reuters, Pamplona 2013.
- [CBPE17] (1) Cátedra BP de Energía y Sostenibilidad . “Observatorio de la Energía y Sostenibilidad en España. Edición 2016”, 2017.
- [CBPE17] (2) Cátedra BP de Energía y Sostenibilidad, “Anejo: tabla de datos”, 2017.
- [CBPEsf] (1) Cátedra BP de Energía y Sostenibilidad, “Presentación de la Cátedra BP de Energía y Sostenibilidad”, s.f.
Recuperado en octubre de 2017 de <https://www.comillas.edu/es/catedra-bp-de-energia-y-sostenibilidad/presentacion>
- [CBPEsf] (2) Cátedra BP de Energía y Sostenibilidad, “Objetivos de la Cátedra BP de Energía y Sostenibilidad”, s.f.
Recuperado en octubre de 2017 de <https://www.comillas.edu/es/catedra-bp-de-energia-y-sostenibilidad/objetivos>
- [CBPEsf] (3) Cátedra BP de Energía y Sostenibilidad, “Observatorio”, s.f.
Recuperado en octubre de 2017 de <https://www.comillas.edu/es/catedra-bp-de-energia-y-sostenibilidad/observatorio>
- [DATOfs] Datos Macro, “España, población”, s.f.
Retirados datos entre 1998 y 2001 en enero de 2018 de

<https://www.datosmacro.com/demografia/poblacion/espana>

- [DUTTsf] Dutton J., “EGEE 102: Energy Conservation and Environmental Protection: Heating Degree Days”, s.f.
Retirado en julio de 2018 de <https://www.education.psu.edu/egee102/node/2055>
- [EUCO14] European Union Council (EUCO), “Conclusions October 2014 about the 2030 climate and energy framework”, 2014.
- [HURCsf] Hospital Universitario Ramón y Cajal, “Contrastes de hipótesis”, s.f.
Recuperado en marzo de 2018 de http://www.hrc.es/bioest/Introduccion_ch.html
- [INE_18] Instituto Nacional de Estadística (INE), “Cifras oficiales de población resultante de la revisión del Padrón municipal a 1 de enero: Resumen por provincias”, 2018.
Retirados datos entre 1998 y 2017 en marzo de 2018 de <http://www.ine.es/jaxiT3/Datos.htm?t=2852>
- [INE_18] (2) Instituto Nacional de Estadística (INE), “Población residente por fecha, sexo y edad”, s.f.
Retirados datos entre 2002 y 2017 en enero de 2018 de <http://www.ine.es/jaxiT3/Datos.htm?t=9663>
- [INVEsf] Investopedia, “Definition of Heating Degree Day – HDD”, s.f.
Retirado en julio de 2018 de <https://www.investopedia.com/terms/h/heatingdegreeday.asp>
- [KISB10] Kisbye P, “Test de Kolmogorov-Smirnov”, 2010.
- [LINA18] Linares P., Comentario sobre el método de cálculo de los HDD y los CDD, 2018.
- [MATHsf] MathWorks, “Soporte. Documentación: fitdist”, s.f.
Recuperado en julio de 2018 de <https://es.mathworks.com/help/stats/fitdist.html>
- [MINIsf] Soporte Minitab, “El estadístico de Anderson-Darling”, s.f.
Recuperado en enero de 2018 de <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/normality/the-anderson-darling-statistic/>

- [MONGsf] Monge J., P Á, “Estadística no paramétrica: prueba Chi-cuadrado”, s.f.
- [ODYS17] ODYSSEE, “ODYSSEE database – july 2017”, 2017.
Recuperado en marzo de 2018 de la base de datos ODYSSEE del proyecto ODYSSEE-MURE en <http://www.odyssee-mure.eu/>
- [ODYSsf] ODYSSEE-MURE, “Definition of ODEX indicators in ODYSSEE database”, s.f.
- [ODYSsf] (2) ODYSSEE-MURE, “Q&A. Result: normal climate”, s.f.
Retirado en marzo de 2017 de <http://www.odyssee-mure.eu/faq/result/3/>
- [PCUE12] Parlamento y Consejo de la Unión Europea, “Directiva 2012/27/UE del parlamento europeo y del consejo”, 2012.
- [REE_17] Red Eléctrica de España (REE), “Series estadísticas nacionales: Producible hidráulico y reservas hidroeléctricas”, 2017.
- [REE_sf] Red Eléctrica de España (REE), “Sala de prensa: Glosario”, s.f.
Recuperado en julio de 2018 de <http://www.ree.es/es/sala-de-prensa/glosario>.
- [SEVI14] Sevillano E., “Un clima cada vez más extremo”, El País, 2014.

9. Anexos

9.1. Anexo 1: Lista de indicadores empleados en los Observatorios.

Los indicadores utilizados en la elaboración de los informes se enumeran y definen a continuación [CBPE17] (1):

- Indicadores sobre la energía primaria.
 - *Consumo total de energía primaria.* Este indicador muestra los registros de energía primaria total consumida, comparando los valores para España, la UE-15, los restantes estados del OCDE y finalmente el resto del mundo.
 - *Consumo de energía primaria per cápita.* Este indicador tiene en cuenta el anterior y los cambios población mundial. En él aparece la situación de España frente a la de la UE-15, la OCDE, y la media mundial.
 - *Intensidad energética primaria.* Se calcula mediante el cociente entre la energía primaria utilizada y el PIB y representa la energía primaria necesaria para producir una unidad monetaria. En el Observatorio se comparan las mismas situaciones que para el indicador anterior.
 - *Energía primaria por PIB: Intensidad energética.* De manera análoga al anterior, muestra el valor medio de energía en su forma final de consumo necesario para producir una unidad de riqueza. En este caso sólo se representa la situación española frente a la UE-15.
 - *Consumo de energía primaria en España.* Este indicador pone de manifiesto el total de energía primaria consumido en España y su repartición entre las distintas fuentes de energía.
 - *Composición de energías renovables en la energía primaria.* Este último es un gráfico circular que representa, en porcentaje, las fuentes de energía renovables que han contribuido a la generación de energía renovable en España en el último año.
- Indicadores sobre emisiones.
 - *Emisiones de CO₂ per cápita.* Expone, en toneladas de CO₂ por habitante, la media mundial comparada a España, a la UE-15 y a la OCDE sin incluir los países de la UE-15.
 - *Emisiones de Gases de Efecto Invernadero (GEI).* Exhibe el total mundial anual de gases de efecto invernadero emitidos cada año a la atmósfera y repartido entre UE-15, la OCDE y el resto de países.
 - *Emisiones de CO₂ por PIB: Intensidad de emisiones.* Representan, de manera análoga a la intensidad energética, las emisiones necesarias para producir una unidad de riqueza. Se compara el total para España, la UE-15, la OCDE y la media mundial.
- Indicadores sobre precios:
 - *Índice de precios “Total Energy” real de la IEA.* Es un índice ponderado de los precios finales a los consumidores, obtenido a partir de Eurostat.

- *Precios de la electricidad.* Este indicador compara la evolución del precio de un kilowatio hora sin impuestos para España y la UE tanto en el sector doméstico como en el industrial.
- *Precios del gas natural.* Exponen los cambios en el precio del gas natural de manera análoga al indicador anterior.
- *Precios de los carburantes.* De manera análoga a los dos anteriores, exhiben los precios de diésel y gasolina sin plomo 95.
- *Precios de los recursos energéticos.* Indica, para los diferentes años, la evolución del precio de los siguientes combustibles: un barril de petróleo crudo Brent, una tonelada métrica de carbón Northwest, un millón de BTUs de gas natural de la UE y finalmente un kilogramo de uranio U308.
- *Precio medio ponderado del CO₂ en Europa.* Exhibe la evolución del precio de emitir una tonelada de CO₂ dentro del EU Emissions Trading System.
- Indicadores sobre la movilidad de mercancías o viajeros en España.
 - *Movilidad interior de mercancías en España.* Este indicador expresa las toneladas por kilómetro transportadas como mercancías en España.
 - *Movilidad interior de viajeros en España.* Representa los pasajeros por kilómetro que utilizan un determinado medio de transporte cada año.
- Diagramas de Sankey.
 - *Fuentes de energía, usos y transformaciones.* En este diagrama se representan las distintas fuentes de energía y las transformaciones que sufren hasta llegar a su uso final por sectores.
 - *Origen de las emisiones de CO₂.* En él se representan las distintas fuentes de energía emisoras de CO₂, nacionales e importadas, y su evolución hacia pérdidas o emisiones de los distintos sectores.
 - *Flujos económicos.* Este diagrama muestra importaciones de energía primaria y su transformación en usos finales, pérdidas, impuestos o exportaciones.
 - *Externalidades.* En este caso se muestran algunas fuentes de energía y su contribución neta a cada sector más las externalidades que generan, tales como emisiones de otros gases contaminantes.
 - *Balance exergético.* En él se distinguen las fuentes de energía nacionales e importadas y su transformación en pérdidas o trabajo útil para los distintos sectores.

9.2. Anexo 2: Fuentes de información utilizadas en la elaboración de indicadores.

Los indicadores son construidos a partir de datos obtenidos de diversas fuentes. Estas últimas se clasifican y enumeran a continuación [CBPE17] (2):

- 1) En el contexto mundial destacan las siguientes fuentes:
 - Del Banco Mundial se obtiene:

- a. La población mundial así como la de los países de la OCDE.
 - b. El Producto Interior Bruto (PIB) total mundial, de los países de la OCDE, de la UE-15 y finalmente de España en dólares corrientes. También provee cifras del PIB a precio constante para los siguientes casos: UE-27, UE-15 y España.
 - Eurostat ofrece las referencias enumerados a continuación:
 - a. La población total de la UE-27, UE-15 y España.
 - b. El PIB medido en euros constantes desde 2010 para la UE-27, la UE-15 y España.
 - c. El consumo de energía primaria tanto de la UE-27 como de la UE-15.
 - La Agencia Internacional de la Energía procura los siguientes elementos:
 - a. El consumo total de energía primaria mundial y de los países de la OCDE.
 - b. Las emisiones de CO₂ por uso de energía en Mt de CO₂ equivalente mundiales, de los países de la OCDE, la UE-27 y la UE-15.
 - c. Los precios finales de la energía al consumidor tanto en el total de la energía, productos petrolíferos, electricidad y gas natural distinguiendo para todos los casos entre sector residencial e industrial. Todos ellos se recogen para el total de países de la OCDE, la OCDE europea y España.
 - BP Statistical review proporciona las referencias enumeradas debajo:
 - a. El precio de los distintos recursos energéticos.
 - b. Las reservas de los siguientes recursos energéticos: petróleo, carbón, gas natural.
 - La Agencia Internacional de la Energía Atómica proporciona los datos sobre las reservas de Uranio.
 - De las Naciones Unidas se obtiene el número de personas sin acceso a formas de energía modernas.
 - Blue Next suministra el precio medio ponderado del CO₂ en Europa.
 - La Universidad de California ofrece la concentración de dióxido de carbono mundial en partes por millón.
 - Finalmente, los siguientes conjuntos de datos son de elaboración propia:
 - a. El consumo total de energía primaria en España, así como sus emisiones por producción de energía.
 - b. El consumo total de energía primaria per cápita, así como por PIB.
- 2) En el contexto nacional, las fuentes empleadas se enumeran a continuación:
- El Instituto Nacional de Estadística proporciona tanto datos poblacionales como referentes al PIB español total y per cápita.
 - Del Ministerio de Fomento se obtienen las referencias relativas a la construcción de infraestructuras, así como del movimiento de personas y mercancías en los diversos medios de transporte.

- De la Dirección General de Movilidad y Transporte de la Comisión Europea se obtienen cifras con relación a la movilidad de personas y mercancías en la UE-27.
 - De la Agencia Estatal Europea del Medio Ambiente se adquiere todo lo relativo a las emisiones de los distintos gases contaminantes por sectores.
 - De Eurostat se reciben los números relativos a los precios de la electricidad, el gas natural y los combustibles empleados en el transporte en carretera tanto para el caso español como para la UE-27.
 - La Agencia Tributaria aporta referencias en el coste total de la energía importada.
 - Finalmente, son de elaboración propia los datos relativos al nivel de dependencia del extranjero.
- 3) Por último, diversas fuentes contribuyen a la elaboración de los diferentes diagramas de Sankey. La gran mayoría del conjunto de datos de elaboración propia a partir de otros, sin embargo, otras fuentes también son de gran utilidad en la elaboración de este conjunto de indicadores.

En los diagramas de fuentes, transformaciones y usos de energía destacan:

- La Corporación de Reservas Estratégicas de Productos Petrolíferos determina los autoconsumos en refinería de gas natural, las exportaciones de productos petrolíferos y las importaciones de petróleo y productos derivados.
- Del Ministerio de Industria, Energía y Turismo se adquieren los datos sobre los autoconsumos totales así como de la energía total exportada. Así mismo, facilita la mayor parte de las cifras correspondiente a la energía primaria utilizada para la producción de electricidad y cogeneración, así como las relevantes a la importación de carbón.
- El Instituto para la Diversificación y el Ahorro de Energía elabora así mismo ciertos registros sobre el consumo total de energía primaria.
- Eurostat provee ciertas notas sobre el consumo de combustibles para el uso final, los autoconsumos de petróleo y gas natural en refinerías así como la energía eléctrica para consumo final en los principales sectores.

En los diagramas de origen de las emisiones de CO₂ en el sector energético español destaca únicamente el Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. Éste proporciona datos de consumo de petróleo, carbón y gas natural en el sector de transporte, el industrial y usos diversos, así como el total de emisiones de dióxido de carbono imputables a pérdidas y a la producción de energía primaria.

