



# **ANÁLISIS DEL PRECIO DE LA ELECTRICIDAD EN EL MERCADO ELÉCTRICO ESPAÑOL: SU RELACIÓN CON LAS TECNOLOGÍAS DE GENERACIÓN Y EMISIONES**

Danilo Atilio Pich Ponti

Trabajo entregado para la obtención del grado de **MÁSTER EN  
“TECNOLOGÍA AMBIENTAL”**

Modalidad: Investigación

Diciembre / 2022

Director:

José Enrique García Ramos



**D. José Enrique García Ramos**, DNI: 75443776-G, Catedrático de Universidad, Departamento de Ciencias Integradas, área de Física Aplicada, Universidad de Huelva,

**INFORMA:**

Que el trabajo titulado “**Análisis del precio de la electricidad en el mercado eléctrico español: su relación con las tecnologías de generación y emisiones**” presentado por **D. Danilo Atilio Pich Ponti**, con **N.I.E.: Y9194232L**, ha sido realizado bajo mi dirección, y autorizo su presentación y defensa como **Trabajo Fin de Máster** (Modalidad: Trabajo de Investigación), para el Máster Universitario en Tecnología Ambiental de la Universidad de Huelva.

En Huelva, a 09 de diciembre de 2022

Fdo.: José Enrique García Ramos



## RESUMEN

En el actual contexto de crisis medioambiental, el uso intensivo de combustibles fósiles con la finalidad principal de la generación de energía (a nivel mundial, aproximadamente el 67% de la misma es generada a través de este tipo de combustibles), ha llevado a un aumento sostenido de los niveles de emisión de los distintos gases de efecto invernadero (GEI), particularmente CO<sub>2</sub>, el cual es uno de los grandes responsables del cambio climático.

España no es ajena a esta situación, ya que su *mix* energético (para la obtención de energía eléctrica) tiene una dependencia directa a los combustibles fósiles, en lo que a sus tecnologías de generación respecta. Debido a que la electricidad cobra cada vez más relevancia como fuente de energía final en el país, lo propio hace su precio, como así también las tecnologías que intervienen en su generación (principalmente las de fuente renovables, en vistas de la política de cero emisiones propuesta por la Unión Europea para el año 2050).

Este trabajo se plantea estudiar la interrelación existente entre el precio final de la energía eléctrica en España (en el mercado *spot*) y las distintas tecnologías de generación (en cuanto a su volumen de producción y potencia instalada de las mismas), como así también el precio de las principales *commodities* que afectan a estas tecnologías, junto con otras variables (como la meteorología). Con esto, se pretende establecer el peso que tienen dichas variables (y principalmente las tecnologías emisoras de GEI) en el precio final de la electricidad para posteriormente realizar un pronóstico considerando posibles escenarios (como el desmantelamiento del parque nuclear) a los fines de entender cómo impactaría esto en el precio final de la energía eléctrica.

**Palabras clave:** precio de la electricidad, precio del gas natural, derechos de emisión, tecnologías de generación, emisiones, ARIMA, regresión Lasso.

## **ABSTRACT**

In the current context of environmental crisis, the intensive use of fossil fuels for the main purpose of power generation (worldwide, approximately 67% of power is generated through this type of fuel), has led to a sustained increase in the emission levels of different greenhouse gases (GHG), particularly CO<sub>2</sub>, which is one of the main drivers of climate change.

Spain is no outside of this situation, as its energy mix (for electricity generation) is directly dependent on fossil fuels, as far as its generation technology is concerned. As electricity is becoming more and more relevant as a final energy source in the country, the same happens with its price, as well as with its generation technologies (mainly the renewable ones, taking into consideration the zero emissions policy proposed by the European Union for the year 2050).

The aim of this work is to study the interrelation among the final price of electricity in Spain (in the spot market) and the different generation technologies (in terms of their production volume and installed capacity), as well as the price of the main commodities affecting these technologies, along with other factors (such as meteorology). Therefore, this study tries to find the contribution of those variables (and mainly that of GHG-emitting technologies) in the electricity final price, to further forecast that price considering possible scenarios (such as the decommissioning of the nuclear plants), in order to understand how this would impact on the final price of electricity.

**Key words:** electricity price, natural gas price, emission rights, generation technologies, emissions, ARIMA, Lasso regression.

## **AGRADECIMIENTOS**

Con la presentación de este trabajo culmina un camino intenso, y quisiera agradecer a todas las personas que han formado parte, de una u otra forma, de este proceso.

En primer lugar, agradecer a mi familia que me ha acompañado en cada momento, dándome su apoyo incondicional a pesar de la distancia. Sin ellos y su confianza en mí, esto no sería posible. Carla, Diane y Pablo, gracias por estar siempre, los amo.

A mis amigos, los de siempre y los que he hecho aquí, quienes han sido testigos y partícipes de este proceso, y siempre han sabido tener la palabra adecuada para acompañarme y alentarme a seguir hacia adelante.

A mis profesores, que con vocación han transmitido sus conocimientos y experiencias, incentivando a sus alumnos a mejorar como estudiantes y como profesionales.

A mi director de TFM, José Enrique García Ramos, quien me ha guiado a lo largo de este trabajo, y ha colaborado brindando todo su tiempo y dedicación para que logre llegar al objetivo. Estoy muy agradecido por toda su paciencia y orientación para llevar adelante este trabajo.

Por último, pero no menos importante, a la Universidad de Huelva por abrirme sus puertas, por darme la chance de continuar con mi formación como profesional y brindarme educación de calidad, poniendo a disposición todas las herramientas necesarias.

**Danilo**

# ÍNDICE

RESUMEN.....	5
ABSTRACT .....	6
AGRADECIMIENTOS .....	7
ÍNDICE .....	8
ÍNDICE DE FIGURAS .....	9
ÍNDICE DE TABLAS.....	10
ÍNDICE DE ABREVIATURAS.....	10
CAPÍTULO 1. INTRODUCCIÓN.....	12
1.1. Introducción: el mercado eléctrico español y su regulación.....	12
1.1.1 El mercado eléctrico español .....	12
1.1.2 Organismos regulatorios .....	12
1.2. El mercado binacional: MIBEL .....	12
1.3. El operador del mercado eléctrico y sus funciones.....	13
1.4. Organización del mercado <i>spot</i> y la formación de precios.....	15
1.4.1 Organización del mercado <i>spot</i> .....	15
1.4.2 Mecanismo de formación de precios .....	15
1.5. Influencia de los distintos parámetros en la conformación del precio .....	16
1.6. Objetivos .....	18
CAPÍTULO 2. MATERIALES Y MÉTODOS.....	20
2.1 Herramienta de análisis .....	20
2.2 Datos analizados .....	20
2.2.1 Obtención de datos .....	20
2.2.2 Completamiento de datos faltantes .....	22
2.2.3 Conjunto final de datos.....	22
2.3 Métodos de análisis .....	22
2.3.1 Determinación de las variables principales .....	23
2.3.2 Aplicación de un modelo de previsión para las variables principales - ARIMA .....	26
2.3.3 Previsión del precio futuro a partir de las variables proyectadas.....	28
2.3.4 Definición de escenarios hipotéticos de generación y previsión a partir de estos.....	28
2.3.5 Cálculo de emisiones del parque de generación de los distintos escenarios .....	28
CAPÍTULO 3. RESULTADOS Y DISCUSIÓN .....	29
3.1 Determinación de la correlación entre las series de datos.....	29
3.1.1 Modelo SEM.....	29

3.1.2 Regresión lineal tipo Lasso .....	33
3.2 Aplicación de un método de previsión.....	38
3.2.1 Previsión de variables para el escenario base – método ARIMA .....	38
3.3 Elaboración de escenarios futuros.....	56
3.4 Previsión del precio final de la electricidad – Regresión lineal tipo Lasso y comparación con escenarios futuros .....	57
3.5 Cálculo de emisiones .....	63
CAPÍTULO 4. CONCLUSIONES .....	65
REFERENCIAS .....	67

## ÍNDICE DE FIGURAS

Figura 1: <i>Ratio</i> de interconexiones del sistema eléctrico de España al año 2020. Fuente: ENTSO-E ( <i>European Network of Transmission System Operators</i> , asociación europea para la cooperación de los gestores de redes de transporte)... 13	
Figura 2: Ejemplo de casación horaria en el mercado <i>spot</i> , para una hora concreta del día.....	16
Figura 3: diagrama de trayectorias del modelo SEM n°1. ....	30
Figura 4: diagrama de trayectorias del modelo SEM n°2. ....	32
Figura 5: Representación del ajuste por regresión lineal.....	35
Figura 6: Representación del segundo ajuste por regresión lineal de tipo Lasso.....	38
Figura 7: Gráfico de correlación de valores estimados (modelo) versus valores reales.....	38
Figura 8: Gráfico de estimación de generación por tecnología de carbón vs. valores reales, y previsión a seis meses.....	41
Figura 9: Gráfico de estimación de generación por tecnología nuclear vs. valores reales, y previsión a seis meses.....	42
Figura 10: Gráfico de estimación de generación por tecnología hidroeléctrica vs. valores reales, y previsión a seis meses. ....	43
Figura 11: Gráfico de estimación de generación por tecnología de ciclo combinado vs. valores reales, y previsión a seis meses.....	44
Figura 12: Gráfico de estimación de generación por tecnología eólica vs. valores reales, y previsión a seis meses.....	45
Figura 13: Gráfico de estimación de generación por tecnología solar térmica vs. valores reales, y previsión a seis meses. ....	46
Figura 14: Gráfico de estimación de generación por tecnología solar fotovoltaica vs. valores reales, y previsión a seis meses. ....	47
Figura 15: Gráfico de estimación de generación por tecnología de cogeneración vs. valores reales, y previsión a seis meses. ....	48
Figura 16: Gráfico de estimación de importación de energía eléctrica vs. valores reales, y previsión a seis meses.....	49
Figura 17: Gráfico de estimación de potencia instalada de tecnología nuclear vs. valores reales, y previsión a seis meses. ....	50
Figura 18: Gráfico de estimación de potencia instalada de tecnología de carbón vs. valores reales, y previsión a seis meses. ....	51

Figura 19: Gráfico de estimación de potencia instalada de tecnología de ciclo combinado vs. valores reales, y previsión a seis meses. ....	52
Figura 20: Gráfico de estimación de potencia instalada de tecnología eólica vs. valores reales, y previsión a seis meses. ....	53
Figura 21: Gráfico de estimación de potencia instalada de tecnología solar fotovoltaica vs. valores reales, y previsión a seis meses.....	54
Figura 22: Gráfico de estimación de potencia instalada de tecnología de cogeneración vs. valores reales, y previsión a seis meses. ....	55
Figura 23: Gráfico de estimación de precio del gas natural vs. valores reales, y previsión a seis meses. ....	56
Figura 24: Previsión del precio de la electricidad en el "Escenario base". ....	58
Figura 25: Previsión del precio de la electricidad en el "Escenario 1". ....	59
Figura 26: Previsión del precio de la electricidad en el "Escenario 2". ....	60
Figura 27: Previsión del precio de la electricidad en el "Escenario 3". ....	61
Figura 28: Comparación del precio de la electricidad en los distintos escenarios con su valor real. ....	62
Figura 29: Comparación de emisiones reales y calculadas en los escenarios. ....	64

## ÍNDICE DE TABLAS

Tabla 1: Coeficientes resultantes de la aplicación del método SEM para el modelo nº1. ....	30
Tabla 2: Coeficientes resultantes de la aplicación del método SEM para el modelo nº2. ....	32
Tabla 3: coeficientes obtenidos por el método de regresión lineal Lasso.....	34
Tabla 4: coeficientes obtenidos por el método de regresión lineal Lasso en segunda instancia. ....	36
Tabla 5: coeficientes del método ARIMA obtenidos por AutoARIMA y los finalmente utilizados.....	39
Tabla 6: Comparativa del precio de la electricidad en los distintos escenarios, con respecto al valor real. ....	62
Tabla 7: factores de emisión publicados por REE. ....	63

## ÍNDICE DE ABREVIATURAS

€: Euros

€/MWh: Euros por megavatio hora

AEMET: Agencia Estatal de Meteorología

ARIMA: *AutoRegressive Integrated Moving Average*

ATR: Acceso de Terceros a la Red

CNMC: Comisión Nacional de los Mercados y la Competencia

CO<sub>2</sub>: Dióxido de carbono

ENTSO-E: *European Network of Transmission System Operators*

EUA: *European Union Allowances*

GHG: *greenhouse gases*

GEI: Gases de efecto invernadero

Lasso: *Least Absolute Shrinkage and Selection Operator*

MIBEL: Mercado Ibérico de la Electricidad

MIBGAS: Mercado Ibérico del Gas

MinCoTur: Ministerio de Turismo, Industria y Comercio de España

MW: Megavatio

MWh: Megavatio hora

NEMO: *Nominated Electricity Market Operator*

OMI: Operador del Mercado Ibérico

OMIE: Operador del Mercado Ibérico de la Electricidad

OMIP: Operador del Mercado Ibérico – Polo Portugués

PNIEC: Plan Nacional Integrado de Energía y Clima

REE: Red Eléctrica de España

SEM: *Structural Equation Modelling*

tCO<sub>2</sub> eq.: Valores unitarios de toneladas equivalentes de dióxido de carbono por unidad de energía generada

# **CAPÍTULO 1. INTRODUCCIÓN**

## **1.1. Introducción: el mercado eléctrico español y su regulación**

### **1.1.1 El mercado eléctrico español**

El mercado eléctrico se define como la totalidad de las transacciones de compra-venta de energía eléctrica, llevadas a cabo por los sujetos de mercado en el mercado diario, intradiario, futuro y los demás servicios derivados de éstos. Este es, en resumen, un mercado mayorista donde acuden los productores por un lado a ofertar, y por otro lado las comercializadoras (y eventualmente también, grandes consumidores que compran directamente) a demandar.

Con la sanción de la Ley 54/1997, se pone en marcha el 1 de enero de 1998 el mercado eléctrico español. Esta ley liberalizó el mercado eléctrico, dividiéndolo en generación, transmisión, distribución y comercialización. Al día de hoy, está vigente la ley 24/2013, la cual mantiene la citada estructura.

### **1.1.2 Organismos regulatorios**

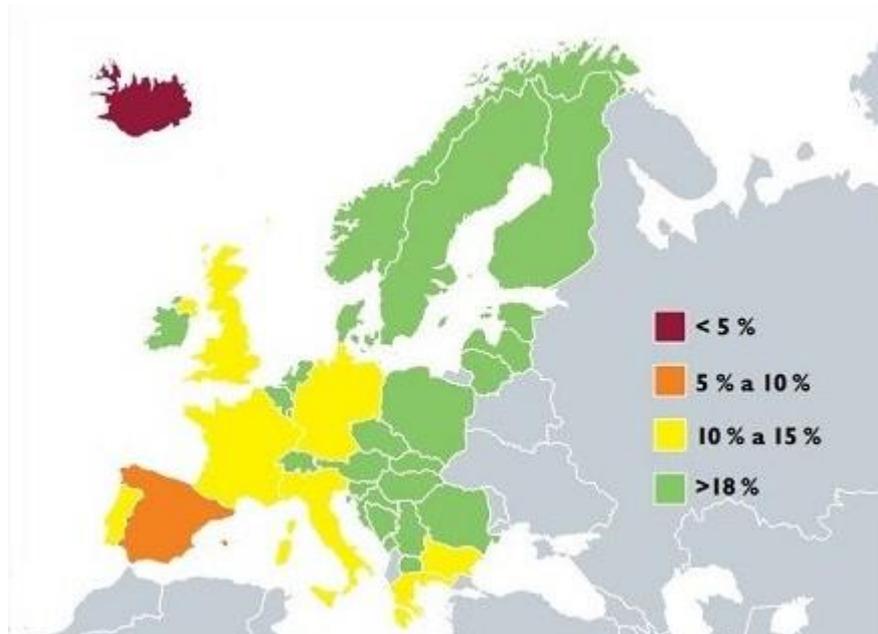
Este mercado eléctrico está regulado por dos organismos distintos con objetivos diferentes pero complementarios. Por un lado, el Ministerio de Turismo, Industria y Comercio (MinCoTur), el cual está a cargo de legislar y regular administrativamente el mercado y todas sus entidades participantes, estableciendo anualmente las tarifas integrales, las tarifas de accesos de terceros a la red (ATR), como así también los premios e incentivos para regímenes especiales o los tipos de impuestos especiales a aplicar.

Por otro lado, se encuentra la Comisión Nacional de los Mercados y la Competencia (CNMC), antes conocida como Comisión Nacional de Energía (la cual se encargaba de regular solamente el mercado eléctrico), el cual es un organismo público de regulación (independiente del gobierno) y sometido a control parlamentario. Esta entidad se encarga de asegurar la competencia efectiva en los mercados energéticos (sector eléctrico y de hidrocarburos), velando por su objetividad y transparencia para el beneficio de todas las partes que operan en el sistema, incluyendo a los consumidores.

## **1.2. El mercado binacional: MIBEL**

El caso de la península ibérica (España y Portugal) constituye una particularidad en el contexto europeo, en lo que refiere al sector eléctrico, ya que se encuentra

relativamente desconectada de Europa central y otros frentes en términos de producción y distribución de electricidad, como se puede apreciar en la siguiente imagen (Figura 1) que ilustra, para diferentes países europeos, su *ratio* de interconexión al año 2020.



**Figura 1:** *Ratio* de interconexiones del sistema eléctrico de España al año 2020. Fuente: ENTSO-E (*European Network of Transmission System Operators*, asociación europea para la cooperación de los gestores de redes de transporte).

Con el fin de mejorar esta situación, los gobiernos de ambos países buscaron una forma de robustecer sus mercados eléctricos y lograr reforzarlos a través de una política integradora. De este modo surgió el Mercado Ibérico de la Electricidad (MIBEL), como una iniciativa conjunta de los gobiernos de España y Portugal con el fin de constituir un mercado binacional de electricidad. El MIBEL fue concebido con el acuerdo Internacional de Santiago en 2004, y quedó finalmente instituido con la firma de un nuevo acuerdo en 2008.

Este mercado le permite a cualquier consumidor en el espacio de la península ibérica comprar energía eléctrica en un régimen de libre competencia, de cualquier productor u operador en España o Portugal.

Dentro del MIBEL, los paquetes negociados son extremadamente variados, desde ventas de energía cerradas con meses de anticipación a operaciones para un momento determinado, cerradas solo unas horas antes.

### **1.3. El operador del mercado eléctrico y sus funciones.**

Con la firma del mencionado acuerdo de 2008 con el que se instituye el MIBEL, se le da también estructura legal al Operador del Mercado Ibérico (OMI). Esta entidad está conformada en mitades iguales por propiedad española y portuguesa.

Del mencionado OMI se desprenden dos entidades particularmente relevantes:

- Operador del Mercado Ibérico de la Electricidad (OMIE): esta institución es el Operador Designado del Mercado Eléctrico (NEMO por las siglas en inglés de *Nominated Electricity Market Operator*) el cual es el encargado de gestionar el mercado diario e intradiario de la electricidad. A este conjunto de mercados se le denomina también el “*pool* eléctrico”, y al precio que surge del mismo, “precio *spot*”. Sus funciones principales son:
  - Permitir el libre acceso al mercado de los diferentes agentes implicados: a mayor competencia, mayores beneficios para los consumidores finales
  - Crear un sistema energético transparente e independiente: el objetivo es que el consumidor final o cualquier agente pueda acceder a los datos e información que desee, relacionados con los precios de la electricidad.
  - Comunicación constante a los agentes del mercado: diariamente, toda persona física o jurídica recibirá una actualización de los precios del día siguiente, así como todo lo relacionado con los pagos y cobros correspondientes.
- Operador del Mercado Ibérico – Polo Portugués (OMIP): esta entidad se encarga de gestionar el mercado a futuro, es decir aquellos donde la entrega de energía se negocia con plazos mayores a 24 horas. Aquí los contratos se negocian a plazo por semanas, meses, trimestres e incluso años. En este caso, compradores y vendedores pueden intervenir en diferentes mercados futuros:
  - Contratos bilaterales: comprador y vendedor intercambian energía por precio, que puede acordarse de dos maneras.
    - Entrega física de energía, en la que el vendedor pone la energía a disposición del comprador y éste pagará el precio fijado.
    - Liquidación financiera: cuando no hay entrega física de energía, pero las partes acuerdan pagarse mutuamente la diferencia entre el precio del mercado diario en el momento de la liquidación del contrato y el precio acordado en el contrato.
  - Mercado de futuros organizado: los contratos estandarizados se negocian en función del importe nominal de las operaciones.
  - Mercado no organizado: son mercados no estandarizados, donde los agentes del mercado cierran las operaciones a través de terceros intermediarios (*brokers*).

El presente trabajo hace foco en este precio *spot*, debido a que el mismo refleja la realidad inmediata del mercado eléctrico, a diferencia de los mercados futuros o a plazo, que tienen una incertidumbre asociada a su naturaleza especulativa.

Como se mencionó anteriormente, OMIE y OMIP son solo dos de las cuatro entidades que pertenecen a OMI, pero a los fines de este análisis, OMIE es la que presenta mayor interés, ya que de esta surge el precio *spot* que será analizado en el presente trabajo.

## **1.4. Organización del mercado *spot* y la formación de precios**

### **1.4.1 Organización del mercado *spot***

El mercado eléctrico mayorista *spot*, bajo la órbita de OMIE, está estructurado en un mercado diario y un mercado intradiario, el cual a su vez se divide en mercado de subastas intradiario e intradiario continuo.

El mercado diario (también llamado de acoplamiento único diario) tiene por objeto la realización de transacciones de electricidad mediante una subasta, donde se da la presentación de ofertas para la venta y compra de electricidad por parte de los agentes del mercado para las veinticuatro horas del día siguiente.

Los mercados intradiarios son una herramienta para que los agentes del mercado puedan ajustar su programa de mercado diario en función de las necesidades en tiempo real, mediante la presentación de ofertas de compra y venta de energía, con un funcionamiento muy similar a la del mercado diario. Esto es lo que permite asegurar el balance continuo entre generación y demanda que requiere la física de la electricidad.

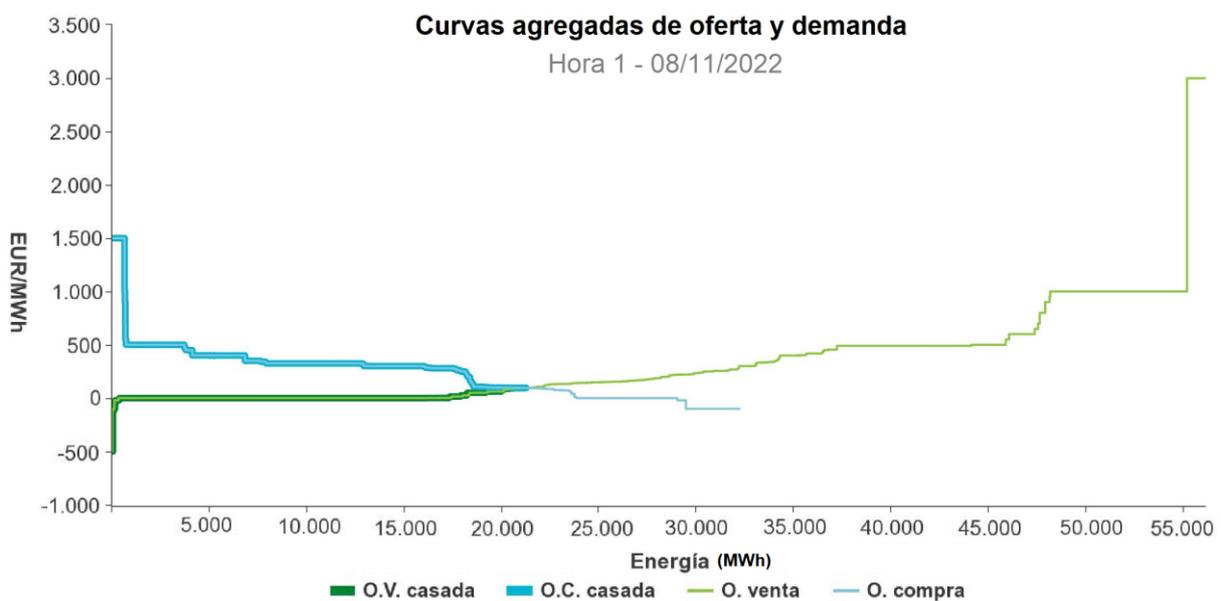
### **1.4.2 Mecanismo de formación de precios**

La mecánica de formación del precio final de la electricidad es compleja y se lleva a cabo por medio de un software con un algoritmo, que se encarga del acoplamiento de la oferta con la demanda (también conocida como casación).

De forma simplificada, la formación del precio se puede entender como el punto de equilibrio resultante del cruce entre dos curvas: la de demanda (perfilada por el operador del sistema eléctrico Red Eléctrica de España o REE) y la de oferta (resultante de la oferta de los distintos productores), para cada hora del programa.

El operador del mercado (OMIE), recibe todas las ofertas de venta y de compra presentadas en el *pool* eléctrico. Estas ofertas provienen de agentes representantes de centrales productoras (para la venta) y de comercializadores, encargados de suministrar a los clientes finales, como así también de grandes consumidores que compran directamente en el *pool* (para la compra).

Para cada una de las sesiones horarias, OMIE ordena las ofertas recibidas de menor a mayor precio para la venta y de mayor a menor precio para la compra, como bien se aprecia en la figura que se adjunta a continuación (Figura 2). La casación del mercado diario es de tipo marginal y el precio para una hora de un día se produce en el cruce de ambas curvas. Ese precio será asignado a todas las ofertas de venta que hubieran quedado por debajo del cruce, y a todas las ofertas de compra que estuvieran por encima.



**Figura 2:** Ejemplo de casación horaria en el mercado *spot*, para una hora concreta del día.

## 1.5. Influencia de los distintos parámetros en la conformación del precio

De lo anteriormente expuesto se aprecia que el proceso de fijación de precio es una operación compleja donde se balancea tanto la oferta como la demanda para cada hora en el sistema eléctrico. Adicionalmente, por la naturaleza de la determinación del precio (oferta-demanda) sucede que en los períodos de mayor demanda el precio sube, ya que es cuando entran en funcionamiento las tecnologías de generación más costosas, y al ser un sistema de tipo marginal, son estas las que marcarán el precio. De este modo, al tener el mercado eléctrico un marcado carácter volátil, es importante tener en cuenta que los precios ofertados por los productores tienen aparejados una serie de costos inherentes a la tecnología de producción y a su naturaleza, como así también existen externalidades ajenas a los productores, los cuales pueden influir en

el precio final. Casado y Larrea (2017) hacen una clasificación de estos factores (los listados en el punto 1,2,5 y 6), a los que se le agregan dos más en el presente análisis, de modo que a los fines de este estudio se puede clasificar a estos factores en:

1. Precio de las materias primas o *commodities*: cada una de las tecnologías de generación requiere de una materia prima distinta en función de su naturaleza. Aun así, a grandes rasgos se las puede clasificar en dos grandes grupos: aquellas que requieren un *input* energético (como combustible nuclear o gas natural) y aquellas que no lo requieren como tal, como las renovables (solar térmica y fotovoltaica, eólica, hidráulica etc.).  
En el caso específico de España y en virtud de las tecnologías de generación presentes, es el precio del gas natural el que tiene mayor relevancia y por tanto será estudiado en este trabajo. Esto se debe a que, de la potencia instalada de centrales térmicas, la mayoría corresponde a las que funcionan a base de este tipo de combustible (centrales de turbina de gas de ciclo combinado y cogeneración).
2. Derechos de emisión: un segundo factor a considerar son los llamados derechos de emisión o EUA (por sus siglas en inglés de *European Union Allowances*), los cuales son un costo adicional para aquellas productoras que utilizan tecnologías emisoras de GEI. Estos EUA (bonos a pagar como derecho para emitir cierta cantidad de CO<sub>2</sub>) se negocian en mercados de similar naturaleza a los existentes en MIBEL (*spot*, a plazo, etc.), y cumplen una doble función de penalizar a los productores que emiten en grandes cantidades de GEI a la vez que resultan un incentivo a reducir dichas emisiones. Los precios de los EUA han aumentado de manera sostenida desde su implementación, y es un costo que potencialmente se puede trasladar al precio final, por lo que su precio *spot* se ha incorporado como otro factor a analizar.
3. Estructura de generación: dentro del proceso de formación del precio, resulta de vital importancia considerar qué componente del *mix* de generación está produciendo en un determinado momento, como así también el volumen que está produciendo, ya que la última de estas tecnologías en entrar en el mercado es la que marcará el precio *spot*. Esto puede variar no solo día a día, sino de manera horaria, como en el caso de la solar fotovoltaica (la cual solo puede producir en un determinado rango horario donde haya radiación solar suficiente), la eólica (con la cual pasa algo similar, solo entra en producción cuando la velocidad del viento esté dentro de un umbral determinado), o también las térmicas (las cuales pueden ofrecer su producción en un determinado momento por estrategia de comercialización, o pueden estar imposibilitadas de producir por algún desperfecto técnico).
4. Potencia instalada: adicionalmente al factor del volumen aportado por cada tecnología de generación en cada hora, también resulta relevante incorporar al análisis la conformación del parque de generación (como valor de potencia instalada para cada tecnología) a nivel nacional, como un indicador del peso

que tiene cada tecnología en la matriz energética. Esto brinda una referencia del potencial de producción que tiene cada tecnología, pudiendo o no alcanzarse ese valor máximo por razones de necesidad, técnicas, indisponibilidades, etc., sin embargo, este valor será el límite máximo (en potencia) que podrá aportar cada tecnología.

5. Factores meteorológicos: el factor meteorológico puede ser relevante en un análisis de este tipo, ya que condiciona de diferente manera a las tecnologías de generación que pueden depender de los mismos (como la solar o eólica), y a su vez presentan una cierta dificultad en su previsión, lo que hace que los productores de dichas tecnologías dependientes de estos factores vean reducida su capacidad de pronosticar con precisión sus volúmenes de generación.
6. Otros factores: económicos, regulatorios, geopolíticos, etc. Estos factores (que pueden jugar un papel importante en el precio final de la electricidad) se hallan incluidos de manera implícita en este análisis, debido a que (por el período temporal elegido) las variables analizadas se encuentran bajo el efecto del conflicto bélico entre Rusia y Ucrania. Este suceso ha tenido como consecuencia más notable (a nivel global) un aumento sustancial en los precios del gas natural, como así también una disminución en su distribución desde Rusia.

## **1.6. Objetivos**

El objetivo principal de este Trabajo de Fin de Máster es, en una primera instancia, identificar si existen correlaciones entre los distintos factores de conformación de precio final de la electricidad analizados y determinar cuáles de estos tienen relevancia en el valor final del precio para luego, aplicando un método de previsión, poder realizar un pronóstico de los precios en base a dichos factores. Paralelamente, planteando distintos escenarios (generados a partir de la variación de los valores de distintas variables) se busca conocer como impactarán estos en el precio futuro.

A tal fin, se plantean entonces una serie de objetivos específicos, a saber:

1. Analizar la información disponible al respecto de las variables a analizar, y recabar la que sea relevante para el mayor período temporal disponible.
2. Someter los datos a un proceso de análisis que permita identificar las correlaciones entre las distintas series de datos, para así encontrar las variables más relevantes y descartar las que tengan menor relevancia, para con esto conformar un modelo del precio final de la electricidad en el mercado *spot*.

3. Aplicar un método de previsión para series temporales univariadas a cada una de las variables identificadas como relevantes, para luego incorporar los valores obtenidos en el modelo de precio desarrollado y así obtener una previsión del valor del precio de la electricidad, comparándolo luego con el precio real para el mismo período.
4. Establecer distintos escenarios hipotéticos a futuro (donde se varíe el valor de distintas variables), aplicar nuevamente el modelo para obtener el precio final de la electricidad para dichos escenarios y finalmente, comparar estos valores con el precio real para el mismo período.
5. Realizar el cálculo de las emisiones en ambos casos (puntos 3 y 4, tanto para los valores proyectados como para los que resulten de los escenarios hipotéticos) y luego compararlas con las emisiones reales para el mismo período.

## CAPÍTULO 2. MATERIALES Y MÉTODOS

### 2.1 Herramienta de análisis

Para el presente análisis, se utilizó el lenguaje informático de alto nivel de programación Python a través de Jupyter Lab, un entorno interactivo para ejecución de código en navegadores web, muy útil para el análisis exploratorio de datos. Este conjunto (Python y Jupyter Lab) conforma el entorno de análisis.

Adicionalmente, para la recopilación de todos los datos de entrada como así también para la elaboración de los escenarios hipotéticos, se utilizó un software de procesamiento de hojas de cálculo (Microsoft Excel).

A través de estas herramientas, se aplicaron todos los métodos (matemáticos, estadísticos y econométricos) que más abajo se detallan.

### 2.2 Datos analizados

Las series de datos analizadas corresponden a los valores de precio horario del mercado *spot* español de electricidad, además de las correspondientes a los factores mencionadas en el apartado 1.5 “Influencia de los distintos parámetros en la conformación del precio”. Cabe destacar que, a excepción de un caso, la totalidad de estos datos fueron obtenidos de portales públicos y abiertos de información de las correspondientes entidades oficiales, lo que asegura la transparencia y certeza en la veracidad de los mismos.

#### 2.2.1 Obtención de datos

Primeramente, se seleccionó el período de estudio común para todas las variables. El mismo va desde el 01/01/2017 al 31/05/2022. El límite temporal inferior está condicionado por la información disponible correspondiente al histórico horario de generación por tecnología, ya que previo al 01/01/2017 no se dispone de datos con alta granularidad (horaria). Por otro lado, el límite temporal superior está dado por los valores del precio de la electricidad en el mercado eléctrico *spot* español, ya que posteriormente a junio del 2022 los precios horarios se ven afectados por la aplicación de la llamada “Singularidad Ibérica”, un mecanismo de desacople del precio de la electricidad respecto a la del gas natural. Es por esto que, para evitar introducir un efecto distorsivo derivado de este mecanismo en los modelos de previsión usados, se ha optado por tomar valores hasta el 31 de mayo de 2022.

A continuación, se detallan las series de datos obtenidas como así también otros detalles de las mismas:

1. Precios horarios de la electricidad base en el mercado eléctrico *spot* español
  - Fuente: Operador del Mercado Ibérico de la Electricidad (OMIE).
  - Unidad: €/MWh.

- Granularidad: horaria.
  - Enlace: <https://www.omie.es/es/file-access-list>
2. Histórico horario de generación por tecnología: las tecnologías de generación consideradas fueron todas las que han estado en actividad en el período considerado y que han ofertado su producción en el mercado *spot* de OMIE. Estas son: hidráulica, nuclear, carbón, turbinas de gas de ciclo combinado, eólica, solar fotovoltaica, solar térmica, cogeneración, residuos y minihidráulica.
    - Fuente: Operador del Mercado Ibérico de la Electricidad (OMIE).
    - Unidad: MWh.
    - Granularidad: horaria.
    - Enlace: <https://www.omie.es/es/market-results/daily/daily-market/hourly-power-technologies>
  3. Histórico de importación de energía eléctrica: en este apartado se consideraron las importaciones totales hacia España, es decir, desde Portugal en el marco del MIBEL, y desde otros países fuera de este mercado.
    - Fuente: Operador del Mercado Ibérico de la Electricidad (OMIE).
    - Unidad: MWh.
    - Granularidad: horaria.
    - Enlace: <https://www.omie.es/es/market-results/daily/daily-market/hourly-power-technologies?scope=daily&date=2022-01-01>
  4. Potencia instalada del parque productor español: en este caso se consideraron las potencias correspondientes a las tecnologías de generación implicadas en este estudio y previamente detalladas.
    - Fuente: Red Eléctrica de España (REE).
    - Unidad: MW.
    - Granularidad: mensual (extrapolada a horaria).
    - Enlace: <https://www.ree.es/es/datos/generacion/potencia-instalada>
  5. Precio del gas natural en el mercado *spot* español.
    - Fuente: Mercado Ibérico del Gas (MIBGAS, entidad de análoga al MIBEL dedicada a la gestión del gas natural).
    - Unidad: €/MWh.
    - Granularidad: diaria (extrapolada a horaria).
    - Enlace: <https://www.mibgas.es/es/file-access>
  6. Precios de los EUA (*European Union Allowances*)
    - Fuente: Energy Market Price (compañía privada que se encarga de recopilar precios de diversos mercados).
    - Unidad: €/ton.
    - Granularidad: diaria (extrapolada a horaria).
    - Enlace: <https://www.energymarketprice.com/home/>
  7. Temperatura media mensual de España y anomalía respecto a la media histórica: a los fines de extender el análisis a todo el territorio español.
    - Fuente: Agencia Estatal de Meteorología (AEMET).

- Unidad (para temperatura media mensual y anomalía): °C.
  - Granularidad: mensual (extrapolada a horaria).
8. Precipitación media mensual de España y desviación porcentual respecto a la media histórica.
- Fuente: Agencia Estatal de Meteorología (AEMET).
  - Unidad:
    - i. Precipitación media mensual: mm.
    - ii. Desviación porcentual: adimensional (porcentaje).
  - Granularidad: mensual (extrapolada a horaria).
  - Enlace: [https://www.aemet.es/es/serviciosclimaticos/vigilancia\\_clima/resumenes](https://www.aemet.es/es/serviciosclimaticos/vigilancia_clima/resumenes)

### 2.2.2 Completamiento de datos faltantes

Dentro de las series de datos analizadas, se encontró que solo 2 de ellas presentaban datos faltantes, pero en ambos casos estos valores representaban una mínima proporción del total de datos de la serie. Estos valores faltantes estaban presentes en:

- Precios horarios de la electricidad base en el mercado *spot* español: en esta serie de datos se han detectados datos faltantes correspondientes a 24hs, por lo que se obtuvo dicha información del portal Energy Market Prices (fuente para la serie de datos n°6).
- Precios de los EUA (*European Union Allowances*): en esta serie se detectó el faltante de aproximadamente un 8% de los datos (dispersos a lo largo de toda la serie), por lo que se procedió al completamiento de los datos faltantes a través de la media aritmética de los valores superiores e inferiores inmediatamente próximos.

### 2.2.3 Conjunto final de datos

Luego de obtenidos los datos y completados los mismos, se conformó un archivo de extensión .xlsx aunando todos los datos en una serie temporal horaria multivariable, donde cada serie de datos (variable) corresponde a una columna. En total, el mismo consta de 47.424 registros de 24 variables distintas, correspondiendo tres de ellas a precios, ocho a volúmenes de producción de energía eléctrica, una a volúmenes de importación de energía eléctrica, ocho a potencia instalada de distintas tecnologías de generación y cuatro a factores meteorológicos. Este archivo es el que ha servido como entrada al entorno de análisis y se ha utilizado como base para todos los análisis subsiguientes.

## 2.3 Métodos de análisis

Con el conjunto de datos ya conformado, a los fines de poder analizar los datos en la secuencia planteada a continuación (determinación de las variables principales,

conformación del modelo, proyección de las variables y previsión del precio final), se utilizaron diversos métodos que fueran acordes en cada caso.

### 2.3.1 Determinación de las variables principales

Luego de obtenidas las series de datos, organizadas las mismas y completados los datos faltantes, se procedió a realizar un primer análisis a los fines de buscar las correlaciones directas e indirectas que pudieran existir entre las distintas variables y el precio final *spot* de la electricidad, a través de un modelo (aplicando el método de *Structural Equation Modelling* o SEM, por sus siglas en inglés) que permitiera evaluar dichas relaciones (Suhr, 2006), para determinar cuáles de las variables estudiadas presentaban mayor grado de relación con el precio final de la electricidad.

En una instancia posterior, se utilizó un segundo modelo para llegar al mismo objetivo de eliminar las variables con menor relevancia (Müller y Guido, 2007), como así también determinar la expresión que relacionara el precio con las variables que lo conforman (Regresión lineal de tipo Lasso).

#### 2.3.1.1 Structural Equation Modelling (SEM)

El modelado estructural de ecuaciones o SEM es una metodología estadística multivariable que se utiliza para estimar relaciones complejas entre variables observadas y latentes. Una variable observada es aquella que se mide directamente, mientras que una variable latente es aquella que no se mide exactamente, o cuyo valor puede ser inferido a través de un modelo matemático desde otras variables que sí son directamente medidas. El modelo resultante se define como una declaración estadística sobre las relaciones entre las variables, y se suele representar a través de un diagrama de trayectorias.

El modelo SEM se conforma por dos partes:

- la parte estructural: vincula las variables latentes entre sí mediante un sistema de ecuaciones lineales;
- la parte de medición: especifica las influencias lineales de las variables latentes en relación a las variables observadas.

Matemáticamente se lo puede plantear de la siguiente manera: si  $\eta$  es un vector de las variables latentes, y  $y$  un vector de variables observadas, las dos partes del modelo son las siguientes:

$$\begin{cases} \eta = B\eta + \varepsilon \\ y = \Lambda y + \delta \end{cases} \quad (1)$$

donde  $B$  y  $\Lambda$  son matrices con parámetros lineales, y  $\varepsilon$  y  $\delta$  son términos de error independientes.

El resultado del modelo es un coeficiente de estimación basado en las relaciones planteadas entre las distintas variables, el cual tiene un valor comparativo entre ellas, es decir, a mayor coeficiente, más peso tiene una variable frente a otra con respecto

al valor final del precio de la electricidad. Esto brinda una gran flexibilidad ya que se pueden plantear distintas hipótesis para luego estudiarlas a través de sus coeficientes, sin embargo, posee una limitación: el método SEM no es un método explorativo; es decir que no puede encontrar nuevas relaciones más allá de las planteadas, por lo que depende plenamente de las hipótesis planteadas.

Adicionalmente, también se obtiene el valor- $p$  de las variables analizadas (más comúnmente llamado *p-value*, por su nombre en inglés), el cual se define como el nivel de significancia mínimo no arbitrario, valor que debe ser menor de 0.05 para considerar a una variable significativa. En el contexto del análisis por método SEM, esto implica que si el valor de una variable latente independiente es menor a 0.05, se puede concluir que dicha variable afecta a la variable latente dependiente (precio final de la electricidad).

En este estudio se ha utilizado un modelo SEM para analizar el modelo conceptual que especifica las relaciones causales entre el precio de la energía eléctrica y las demás variables relevantes. La razón por la cual se ha elegido este modelo es el hecho de que permite relacionar distintas variables de forma visual (a través del diagrama de trayectorias) y establecer distintas relaciones más allá de una relación directa o lineal entre las mismas. Con los resultados obtenidos se busca determinar la relación entre las distintas variables y su influencia en el precio final de la electricidad.

#### **2.3.1.2 Regresión lineal – *Machine Learning* y método Lasso**

Con una finalidad análoga a la anterior, se aplicó un método de regresión lineal para poder identificar las variables más relevantes y descartar aquellas que tuvieran menor relevancia, a través de una técnica de *Machine Learning*.

##### *Machine Learning* o aprendizaje automático

El *Machine Learning* (cuya traducción al español sería “aprendizaje automático”) es una rama de la inteligencia artificial que permite que, a través de algoritmos, un programa tenga la habilidad de identificar patrones en datos masivos (“*big data*”) y pueda realizar predicciones, sin ser expresamente programado para esto.

Dentro del *machine learning*, se encuentran dos ramas muy diferenciadas en función a su naturaleza y forma de funcionamiento: el aprendizaje automático supervisado y no supervisado. En el primero, sólo se conocen los datos de entrada y no se le brinda datos de salida al algoritmo, por lo que estos métodos suelen ser más difíciles de entender y evaluar. En el caso del aprendizaje supervisado, el usuario le proporciona al algoritmo los pares de entradas y salidas deseadas, y el algoritmo encuentra la manera de producir la salida deseada en función de la entrada, siendo capaz de crear una salida para una entrada que nunca antes ha visto, sin la ayuda de un humano.

Este último, el aprendizaje automático supervisado, se denomina así porque primero se debe construir un modelo a partir de estos pares de entrada/salida (que constituyen el denominado conjunto de entrenamiento) y a partir de allí sí hacer predicciones precisas para datos nuevos y nunca vistos. Estos datos “no vistos”, con los que se puede probar el modelo, son una parte de los datos de entrada que ha sido reservada específicamente a tal fin (a la cual se le llama conjunto de prueba), es decir, que el modelo no se ha entrenado con ellos por lo que sirven para aplicar dicho modelo y analizar si el resultado que devuelve es adecuado o no.

El aprendizaje supervisado suele requerir un esfuerzo humano para construir el conjunto de entrenamiento, pero después automatiza (y a menudo acelera) una tarea que de otro modo sería laboriosa o inviable. Este tipo de aprendizaje automático es uno de los más exitosamente aplicados y difundidos, y se utiliza para resolver básicamente dos tipos de problemas: de clasificación, donde el objetivo es predecir una etiqueta de clase (es decir, obtener un resultado a partir de la elección desde una lista predefinida de posibilidades), y de regresión, donde el objetivo es predecir un número continuo.

En consecuencia, para el presente análisis se utilizarán herramientas de regresión del aprendizaje automático supervisado; ya que se le brinda al algoritmo datos de entrada y de salida (valores de todos los factores que componen el precio final de la electricidad y valores de dicho precio, respectivamente), se conforma un conjunto de datos de entrenamiento, y el algoritmo luego genera automáticamente la salida (un modelo para determinar una serie de números continuos: precios finales de la electricidad).

### Modelos regresivos de aprendizaje automático: el modelo lineal

Dentro de los modelos de regresión del aprendizaje automático supervisado hay varios algoritmos desarrollados; en este análisis se ha utilizado un modelo lineal.

Los modelos lineales son modelos ampliamente estudiados y muy utilizados, no solo actualmente sino desde hace ya varias décadas, debido a su simplicidad y fácil interpretación. Este tipo de modelos en su esencia, hacen una predicción utilizando una función lineal de las características de entrada.

Para la regresión, la fórmula general de predicción de un modelo lineal es la siguiente:

$$\hat{y} = w[0] * x[0] + w[1] * x[1] + \dots + w[p] * x[p] + b \quad (2)$$

En donde  $x[0]$  a  $x[p]$  denota las características (en este ejemplo, el número de características es  $p$ ) de un solo punto de datos,  $w$  y  $b$  son parámetros del modelo que son aprendidos por el mismo, e  $\hat{y}$  es propiamente la predicción que hace el modelo. Existen muchos modelos diferentes de regresión lineal, los cuales se diferencian básicamente en cómo el modelo aprende los parámetros  $w$  y  $b$  desde la serie de datos, y cómo se puede controlar la complejidad del modelo. De todos estos modelos

disponibles, se ha utilizado el método de los mínimos cuadrados con regresión tipo Lasso. Esto se debe a que este tipo de modelo es útil cuando se tiene una gran cantidad de variables y se espera que solamente alguna de ellas sea representativa.

El modelo Lasso (por sus siglas en inglés *Least Absolute Shrinkage and Selection Operator*, que se puede traducir como “Operador de mínima contracción absoluta y selección”) es un método de análisis de regresión que combina un modelo de regresión con un procedimiento de contracción de algunos parámetros hacia cero y selección de variables, imponiendo una restricción o una penalización sobre los coeficientes de regresión.

Este método restringe los coeficientes ( $w$ ) para que sean lo más pequeños posible de tal manera que ciertos coeficientes sean exactamente cero, lo que implica que algunas variables son directamente ignoradas por el modelo. Esta característica (conocida como regularización) se controla a través de un coeficiente denominado *alpha*, el cual determina que tanto serán disminuidos los coeficientes.

Esto se traduce entonces en una forma de selección automática de variables, ya que hace al modelo más fácil de interpretar y puede revelar las variables más importantes del mismo. Esta característica de regularización ha resultado ser clave a la hora de la elegir este modelo en específico para el análisis, ya que permite seleccionar solo las variables más relevantes. De este modo, aplicando este método no solamente se obtiene un modelo sencillo de interpretar (ya que solo seleccionará una parte de todas las variables de entrada, lo que simplifica el modelo a nivel interpretativo) sino que también se aligera la carga computacional requerida a lo largo de todo el análisis.

El modelo de Lasso se puede describir como:

$$RSS_{Lasso} = \sum_{i=1}^n (y_i - f(x_i))^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (3)$$

Una vez aplicado este modelo, donde los coeficientes  $\beta_j$  corresponden a los coeficientes del ajuste lineal, se eliminaron las variables cuyos coeficientes de regresión fueran cero, y se obtuvo la expresión para el cálculo del precio final de la electricidad a partir de las variables seleccionadas y sus correspondientes coeficientes.

### **2.3.2 Aplicación de un modelo de previsión para las variables principales - ARIMA**

Una vez determinadas las variables más relevantes dentro de las analizadas, y conformado el modelo para la determinación del precio a partir de las variables de entrada seleccionadas, se procedió a realizar una proyección a futuro de los valores de cada una de ellas utilizando el modelo ARIMA.

## El modelo ARIMA

Actualmente existe una gran variedad de métodos de previsión para series temporales univariadas, sin embargo, se ha decantado por el método ARIMA ya que este posee un alto grado de ajuste, es apropiado para trabajar con grandes series temporales y puede manejar series temporales con estacionalidad (gracias a su componente de integración).

El método ARIMA (acrónimo del inglés *AutoRegressive Integrated Moving Average*, o modelo autoregresivo integrado de promedio móvil) es un modelo estadístico para series temporales ampliamente utilizado en estadística y econometría, que se vale de los tres elementos mencionados (autoregresión, integración y media móvil) para describir un valor como una función lineal de datos anteriores y errores debidos al azar.

El modelo ARIMA es un modelo desarrollado en la década de 1960, sistematizado por George Box y Gwilym Jenkins, por lo que también se lo suele conocer como el “modelo Box-Jenkins” (Box y Jenkins, 1970).

Este modelo se puede describir como:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \dots + \phi_q \varepsilon_{t-q} \quad (4)$$

Según la anterior expresión se puede entender al valor de la precisión como la sumatoria de una constante más una combinación lineal de valores de *lags* (rezagos) de  $Y$  más una combinación lineal de errores de previsión rezagados.

Este modelo tiene la característica de poder realizar previsiones sobre series de datos tanto estacionales como no estacionales. A estas últimas, les aplica un proceso de diferenciación para llegar a la estacionariedad. Además, puede incluir un componente de tendencia, y uno cíclico o estacional. Todas estas características se introducen a través de una serie de coeficientes llamados “p”, “d” y “q” (a los cuales el modelo es muy sensible), y que son requeridos para poder realizar la previsión a futuro, por lo que este tipo de modelos se suele denotar como  $ARIMA_{(p,d,q)}$ . Estos términos corresponden a:

- p: número de términos autoregresivos, también conocido como “lag order”;
- d: el número de veces que se diferencian las observaciones; también conocido como grado de diferenciación; y
- q: el tamaño de la ventana de la media móvil; también conocido como orden de la media móvil.

El proceso de determinación de estos parámetros es complejo ya que los mismos condicionarán la forma que tome el modelo, sin embargo, existen aplicaciones ya implementadas dentro de los mismos modelos ARIMA que permiten realizar un

análisis de la serie y llegar a determinar automáticamente el valor de estos parámetros; y esto es lo que se ha utilizado en el presente estudio. Las únicas excepciones se han dado donde, luego del análisis de la previsión hecha por el modelo, se han corregido manualmente estos valores para obtener un mejor ajuste del mismo.

### **2.3.3 Previsión del precio futuro a partir de las variables proyectadas**

Una vez realizada la previsión de los valores futuros de las variables que componen el precio final de la electricidad (según el modelo definido), se procedió a utilizar nuevamente el modelo de composición del precio final de la electricidad (obtenido a través del método Lasso de regresión lineal), pero esta vez los valores de entrada utilizados son los valores de las variables seleccionadas ya proyectados a futuro.

### **2.3.4 Definición de escenarios hipotéticos de generación y previsión a partir de estos**

Tomando en consideración las variables analizadas, se establecieron tres escenarios *ad hoc* de generación en virtud de las condiciones actuales del parque de generación (tecnologías disponibles y previsión de crecimiento, como en el caso de la fotovoltaica o la eólica), las condiciones previstas a futuro (disminución del aporte de ciertas tecnologías en el parque productor, como la nuclear) y las posibles variaciones que pudieran surgir (como, por ejemplo, una normalización en el precio del gas natural). Con estos escenarios plasmados en las variables de entrada, se procedió a utilizar nuevamente el modelo de composición del precio final de la electricidad, para obtener dicho valor para cada uno de los escenarios planteados.

### **2.3.5 Cálculo de emisiones del parque de generación de los distintos escenarios**

Finalmente, considerando las previsiones realizadas de los valores de las variables que componen el precio final de la electricidad, para cada uno de los cuatro escenarios desarrollados, se realizó el cálculo de las emisiones del parque de generación en el período proyectado de seis meses.

Para esto se utilizaron los factores de emisión (valores unitarios de toneladas equivalentes de dióxido de carbono -tCO<sub>2</sub> eq.- por unidad de energía generada) para las distintas tecnologías emisoras involucradas (turbinas de gas de ciclo combinado, centrales de carbón y cogeneración), publicados por el operador del sistema eléctrico español (el grupo Red Eléctrica España, o REE).

La obtención del valor final de las emisiones resulta simplemente de hacer el producto entre el volumen de energía eléctrica producido por su correspondiente factor de emisión. Estos valores son finalmente comparados con los valores de emisión reales para el período analizado.

## **CAPÍTULO 3. RESULTADOS Y DISCUSIÓN**

En este apartado se detallan los resultados obtenidos del análisis realizado a través de los métodos previamente explicados. A los fines de estructurar este apartado, los resultados y sus correspondientes discusiones son presentados en distintas instancias, siguiendo el orden establecido en los objetivos inicialmente planteados.

### **3.1 Determinación de la correlación entre las series de datos**

A los fines de poder determinar la correlación entre la serie de datos y elaborar finalmente un modelo que relacione a las distintas variables con el precio final de la electricidad, se aplicó primeramente un modelo SEM (cuyos resultados no fueron lo suficientemente contundentes) y en segunda instancia, un modelo de regresión lineal tipo Lasso.

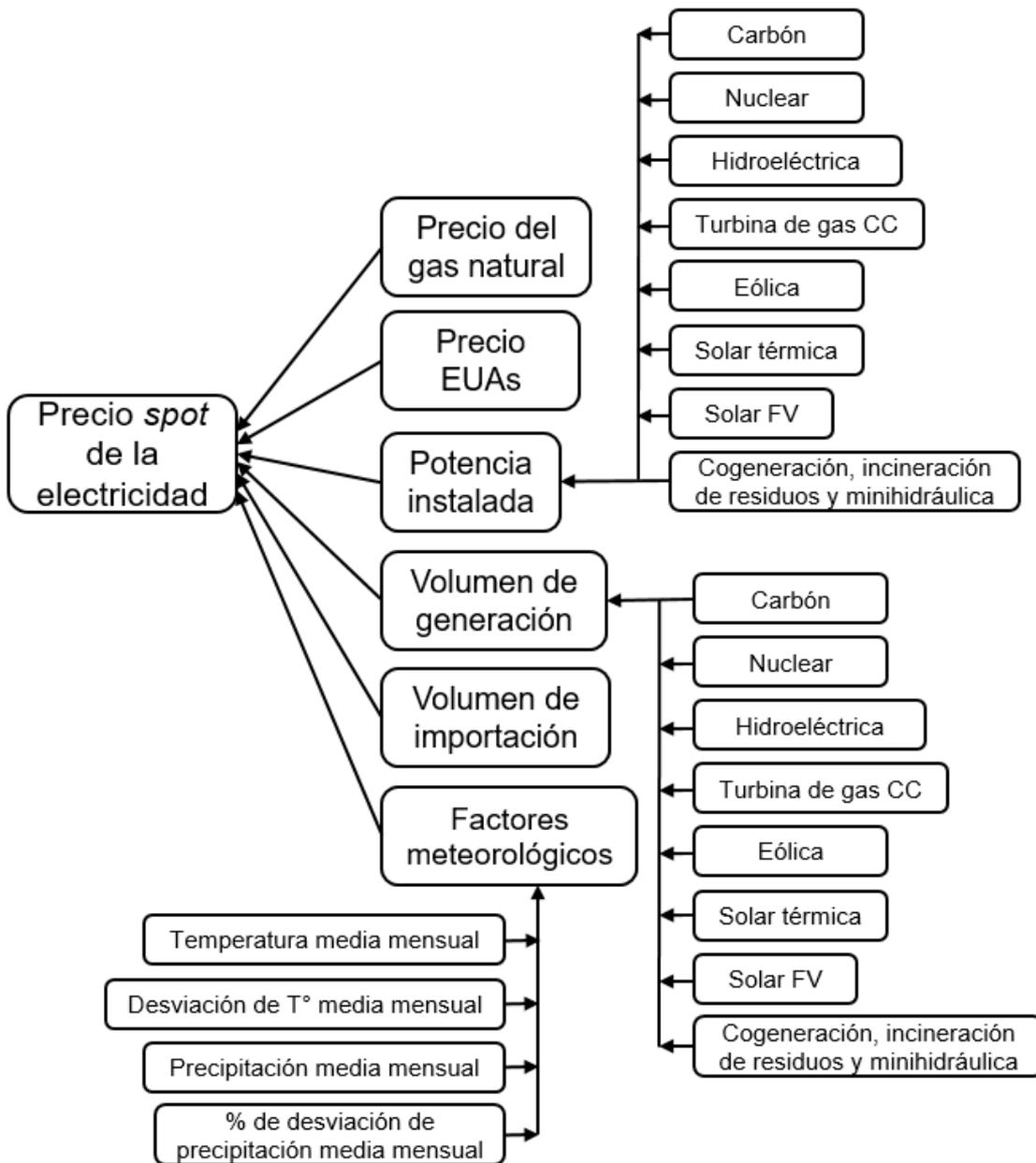
Finalmente, los resultados de este último fueron los utilizados para conformar el modelo final.

#### **3.1.1 Modelo SEM**

Este método fue utilizado (como se mencionó en el apartado 2.3.1.1) ya que posibilita plantear una interrelación entre las distintas variables y analizarlas visualmente a través del diagrama de trayectorias, facilitando su interpretación.

A los efectos de aplicar este modelo, se catalogó como variables observadas a aquellas que no estaban compuestas por otras variables en sí mismas (como el precio del gas natural, el precio de los EUA, o los valores de importación de energía eléctrica), y como variables latentes a aquellas que se conformaban por más variables dentro de las analizadas (como el caso de la potencia del parque generador, que es la sumatoria de las potencias individuales de cada una de las tecnologías de generación, o los factores meteorológicos, que se ha definido por los valores medios y las desviaciones de la temperatura y la precipitación).

El modelo propuesto en un primer lugar, denominado Modelo 1 (Figura 3) interrelaciona todas las variables de manera directa con el precio, y es el que se presenta a continuación:



**Figura 3:** diagrama de trayectorias del modelo SEM n°1.

Los resultados obtenidos de la aplicación del método SEM con el modelo n°1 fueron los siguientes:

**Tabla 1:** Coeficientes resultantes de la aplicación del método SEM para el modelo n°1.

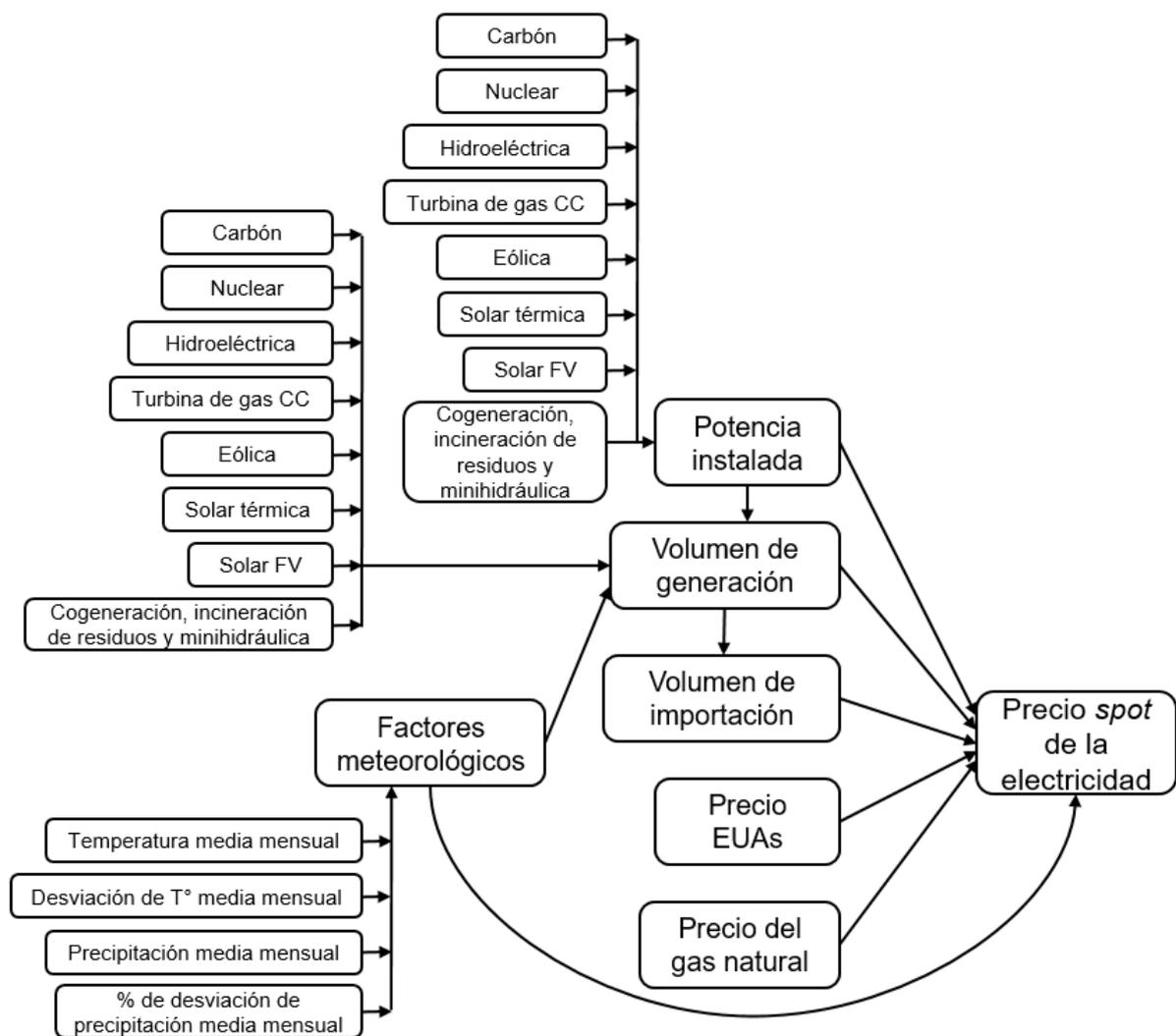
Variable	Estimación	p-value
Precio de gas natural	2.095	0.0000
Precio de EUAs	0.066	0.0000
Volumen de generacion	1.637	0.8257
Potencia del parque generador	0.132	0.0000
Temperatura media mensual	0.185	0.0000
Anomalía de temp. media mensual	0.080	0.3685
Precipitación media mensual	-0.048	0.0000
Desv. porc. de precipitación media mensual	-1.067	0.0000

A la hora de interpretar los resultados (Tabla 1), se aprecia que la variable que se determina como más relevante para la conformación del precio (en la columna “Estimación”) es el precio del gas natural en España, seguida por el volumen de generación de las distintas tecnologías y desviación porcentual de la precipitación media mensual. A su vez, se observa que los valores del p-valor para la generación resultan mayores a 0.05, por lo que se los considera no significativos, situación que resulta inicialmente contraintuitiva.

En el caso de los valores negativos de Estimación, estos son lo que se conoce como “caso Heywood”, y se pueden deber a una muestra demasiado pequeña como para estimar adecuadamente los parámetros, a datos que contienen valores atípicos o a un modelo especificado que no es apropiado para los datos.

En virtud de la presencia de estos valores negativos, y descartándose que se puedan deber a las dos primera razones mencionadas (ya que la muestra posee un tamaño considerable y los valores se encuentran en rangos adecuados), se determinó que se debía generar un segundo modelo SEM para poder contrastar los resultados obtenidos.

El nuevo modelo propuesto, denominado Modelo 2 (Figura 4), en lugar de interrelacionar directamente a las variables, plantea a las variables precio de EUAs y precio del gas natural como variables observadas, y a las tecnologías y potencias de generación, factores meteorológicos y volumen de importación, como variables latentes. Esta segunda clasificación no responde al hecho de que las mencionadas variables latentes no sean directamente medibles, sino a un planteo conceptual a los fines de determinar la influencia indirecta de estas variables en el valor final del precio. El modelo propuesto es el que se plantea a continuación:



**Figura 4:** diagrama de trayectorias del modelo SEM n°2.

Los resultados obtenidos de la aplicación del método SEM con el modelo n°2 fueron los que se presentan a continuación:

**Tabla 2:** Coeficientes resultantes de la aplicación del método SEM para el modelo n°2.

Variable	Estimación	p-value
Precio de gas natural	2.092	0.000
Precio de EUAs	0.124	0.000
Potencia del parque generador	0.069	0.000
Temperatura media mensual	0.452	0.000
Anomalía de temp. media mensual	-0.369	0.081
Precipitación media mensual	-0.036	0.004
Dev. porc. de precipitación media mensual	-2.453	0.000

De los resultados obtenidos de la aplicación del segundo modelo (Tabla 2), se aprecia que las variables que presentan mayor relevancia son la desviación porcentual de la precipitación media mensual, el precio del gas natural y la temperatura media mensual. En este caso, todos los p-valor (menos el de la anomalía de temperatura media mensual) son menores a 0.05 por lo que dichas variables resultan significativas.

Análogamente a lo sucedido con el modelo n°1, se encuentra que hay varios valores negativos del parámetro Estimación (concretamente el 43% de ellos), por lo que se determina que el modelo establecido no es concluyente al respecto de la selección de variables.

En virtud de que los resultados obtenidos a través de la aplicación de ambos modelos SEM no permitieron establecer con claridad las variables más importantes en relación con el precio final de la electricidad, es que se plantea utilizar un segundo método de selección de variables (que a su vez permitirá también obtener un modelo del precio final de la electricidad), cuya aplicación se plasma a continuación.

### **3.1.2 Regresión lineal tipo Lasso**

Con la finalidad de determinar cuáles son las variables más relevantes a la hora de definir un modelo para el cálculo del precio final de la electricidad y luego con estas poder obtener propiamente el modelo para la determinación de dicho precio, se aplicó el método de regresión lineal tipo Lasso.

Este método de regresión lineal de aprendizaje automático supervisado permite (a través de su coeficiente de regularización *alpha*) forzar la regularización de las variables de modo que el coeficiente de las mismas sea lo menor posible.

La regularización implica básicamente evitar el sobreentrenamiento del modelo (también conocido como *overfitting* en inglés), es decir, que el modelo se apegue demasiado a los datos de entrenamiento ya conocidos de tal manera tal que, cuando se enfrente a nuevos datos, no pueda generar un resultado preciso. En consecuencia, el modelo solo puede dar resultados altamente precisos para los datos conocidos, pero muy imprecisos para datos nuevos.

Para evitar esto, se deben utilizar valores extremos del coeficiente de regularización *alpha* (siendo *alpha*=1 el valor por defecto), por lo que el valor utilizado fue *alpha*=10. De este modo, se separan los valores en un conjunto de datos de entrenamiento (un 75% del total de los datos disponibles) y un conjunto de datos de prueba (el 25% restante). Aplicando el método (con el valor de *alpha* establecido), se obtienen entonces los coeficientes de cada una de las variables (Tabla 3).

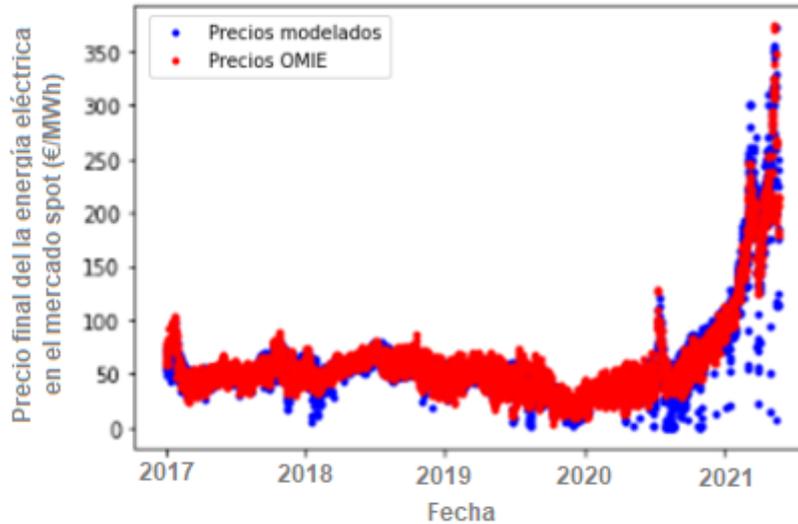
Cabe mencionar que, a la hora de aplicar este método, se realizó el ajuste con los valores de las variables promediados mensualmente, ya que es a estos valores a los que se aplicará posteriormente el método de previsión ARIMA. Se ampliará al respecto en la sección correspondiente a métodos de previsión.

**Tabla 3:** coeficientes obtenidos por el método de regresión lineal Lasso.

	<b>Coeficiente</b>	<b>Unidad</b>
Generación por tecnología de carbon	1.19E-03	€/MWh <sup>2</sup>
Generación por tecnología nuclear	-1.59E-03	€/MWh <sup>2</sup>
Generación por tecnología hidroeléctrica	-2.15E-03	€/MWh <sup>2</sup>
Generación por tecnología de ciclo combinado	8.62E-04	€/MWh <sup>2</sup>
Generación por tecnología eólica	-1.12E-03	€/MWh <sup>2</sup>
Generación por tecnología solar térmica	-3.26E-03	€/MWh <sup>2</sup>
Generación por tecnología solar fotovoltaica	2.38E-04	€/MWh <sup>2</sup>
Generación por tecnología de cogeneración y otros	2.09E-03	€/MWh <sup>2</sup>
Importacion de energía eléctrica	2.06E-03	€/MWh <sup>2</sup>
Potencia de tecnología hidráulica	0.00E+00	€/MWh MW
Potencia de tecnología nuclear	-9.24E-03	€/MWh MW
Potencia de tecnología de carbón	2.51E-04	€/MWh MW
Potencia de tecnología de ciclo combinado	-1.46E-02	€/MWh MW
Potencia de tecnología eólica	-2.14E-02	€/MWh MW
Potencia de tecnología solar fotovoltaica	1.40E-02	€/MWh MW
Potencia de tecnología solar térmica	0.00E+00	€/MWh MW
Potencia de tecnología de cogeneración y otros	-6.74E-03	€/MWh MW
Precio de EUAs	0.00E+00	ton/MWh
Precio de gas natural	1.76E+00	-
Temperatura media mensual	0.00E+00	€/MWh °C
Anomalía de temp. media mensual	0.00E+00	€/MWh °C
Precipitación media mensual	0.00E+00	€/MWh mm
Desv. porc. de precipitación media mensual	0.00E+00	€/MWh

De los coeficientes obtenidos en esta primera instancia (Tabla 3), se observan que se han eliminado 7 variables a través de la regresión, ya que su coeficiente es igual a 0. Esto es producto de la mencionada regularización del método Lasso, de modo tal que ahora el modelo ha sido reducido, de las 23 variables iniciales (excluyendo propiamente al precio *spot* de la electricidad), a 16 variables finales.

Si se representan los valores de precio obtenidos a través del modelo contrastados con el precio real (Figura 5), se aprecia el alto nivel de ajuste del modelo.



**Figura 5:** Representación del ajuste por regresión lineal.

Se debe realizar la salvedad de que si bien el ajuste se ha realizado con los valores de precio promediados mensualmente como ya se ha mencionado, para una mejor apreciación de la gráfica se han representado los valores horarios en la gráfica (Figura 5).

Esta representación gráfica es útil para tomar dimensión del grado de ajuste del modelo, sin embargo, se requiere de valores analíticos para poder juzgar propiamente la capacidad del modelo de ajustarse a los valores de precio.

En tal sentido, para evaluar analíticamente el ajuste se pueden utilizar diversos parámetros estadísticos. Si bien determinar analíticamente la bondad de un ajuste puede resultar complejo, se han seleccionado indicadores que fueran representativos y pudieran aportar información clara sobre el ajuste. Los indicadores utilizados en este caso fueron:

- el valor de varianza explicada o  $r^2$ , que representa la proporción de la varianza de una variable dependiente que se explica por una o varias variables independientes. Conceptualmente se define como:

$$r^2 = \frac{\text{Varianza total explicada por el modelo}}{\text{Varianza total}} \quad (5)$$

- MSE (acrónimo inglés de *Mean Squared Error*, o error medio cuadrático) que consiste en el promedio del cuadrado de la diferencia entre los valores reales y los valores estimados por la regresión, cuya raíz cuadrada indica la diferencia promedio entre los valores reales y los estimados. Esta se define como:

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n} \quad (6)$$

donde  $y_i$  es el valor real,  $\hat{y}_i$  el valor estimado y  $n$  el número de observaciones.

- MAPE (acrónimo inglés de *Mean Absolute Percentage Error*, o error medio absoluto porcentual), un indicador que mide la exactitud de un método para la

construcción ajustada de valores de series temporales a través del tamaño del error (absoluto) en términos porcentuales. Se define como:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (7)$$

Los valores obtenidos de estos parámetros fueron los siguientes:

- $r^2$  de prueba = 0,990;
- $r^2$  de entrenamiento = 0,995;
- MSE= 40,156;
- MAPE= 0,083.

En el caso del valor de  $r^2$  (tanto para el conjunto de datos de prueba como el de entrenamiento) se aprecia que el mismo es muy cercano a 1, indicando que el modelo explica correctamente el valor del precio de la electricidad.

Es importante destacar que el valor de este parámetro en el caso del conjunto de datos de entrenamiento es muy próximo al de los datos de prueba, lo que es un buen indicador de que no se ha producido *overfitting* o sobreentrenamiento.

Por otra parte, el MSE presenta un valor de 40,15, el cual es un valor esperable en orden de magnitud teniendo en cuenta las diferencias que se observan entre los datos reales y los predichos. Por último, el MAPE toma un valor más bajo.

A los fines de confirmar el ajuste hallado y determinar finalmente cuáles eran las variables a descartar, se procedió a realizar nuevamente la regresión lineal, pero esta vez sobre las variables que no habían sido descartadas en una primera instancia. En esta segunda instancia se selecciona nuevamente un coeficiente de regularización *alpha* igual a 10.

Para esta segunda instancia, los coeficientes finales obtenidos fueron los siguientes:

**Tabla 4:** coeficientes obtenidos por el método de regresión lineal Lasso en segunda instancia.

	Coeficiente	Unidad
Generación por tecnología de carbon	1.23E-03	€/MWh2
Generación por tecnología nuclear	-1.65E-03	€/MWh2
Generación por tecnología hidroeléctrica	-2.19E-03	€/MWh2
Generación por tecnología de ciclo combinado	8.21E-04	€/MWh2
Generación por tecnología eólica	-1.10E-03	€/MWh2
Generación por tecnología solar térmica	-3.11E-03	€/MWh2
Generación por tecnología solar fotovoltaica	5.97E-05	€/MWh2
Generación por tecnología de cogeneración y otros	2.11E-03	€/MWh2
Importacion de energía eléctrica	2.02E-03	€/MWh2
Potencia de tecnología nuclear	-9.69E-03	€/MWh MW
Potencia de tecnología de carbón	2.89E-04	€/MWh MW
Potencia de tecnología de ciclo combinado	-1.53E-02	€/MWh MW

Potencia de tecnología eólica	-2.31E-02	€/MWh MW
Potencia de tecnología solar fotovoltaica	1.50E-02	€/MWh MW
Potencia de tecnología de cogeneración y otros	-7.06E-03	€/MWh MW
Precio de gas natural	1.75E+00	-

Los valores obtenidos (Tabla 4) muestran que, bajo el análisis realizado, el valor final del precio de la electricidad depende principalmente del precio del gas natural, seguido por las potencias instaladas de la tecnología eólica, de centrales térmicas de ciclo combinado y de solar fotovoltaica.

En cuanto a los resultados expuestos, se puede observar que ninguna de las variables seleccionadas ha sido eliminada por el método Lasso, lo que indica que dichas variables son representativas para la conformación del modelo de precio.

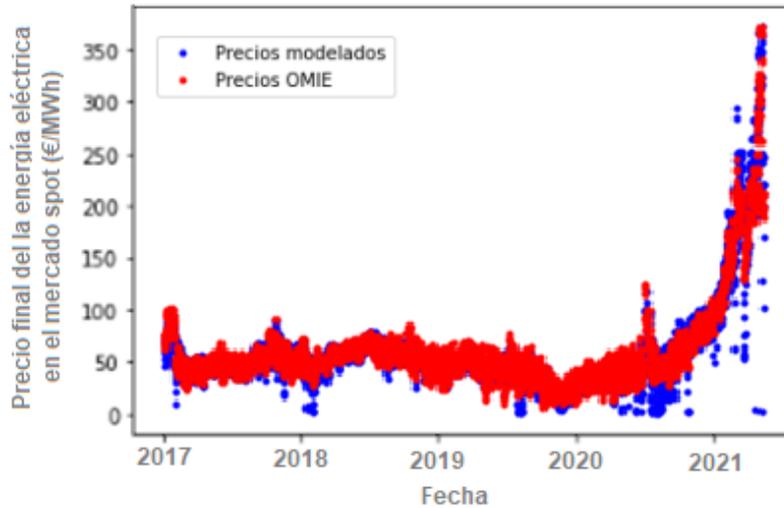
A su vez, los valores de los parámetros estadísticos para esta segunda aplicación del método son los que se presentan a continuación:

- $r^2$  de prueba = 0,989;
- $r^2$  de entrenamiento = 0,995;
- MSE= 41,985;
- MAPE= 0,087.

Nuevamente se aprecia que ambos  $r^2$  tienen valores cercanos a 1, y se da nuevamente el hecho de que el  $r^2$  del conjunto de datos de entrenamiento es muy similar al de prueba, lo que indica la ausencia de sobreentrenamiento del modelo. Por otra parte, el valor final del MSE de 41,985 es un resultado esperable (similar al obtenido en primera instancia) considerando la diferencia entre datos reales y los predichos, y finalmente el MAPE presenta nuevamente un valor más pequeño.

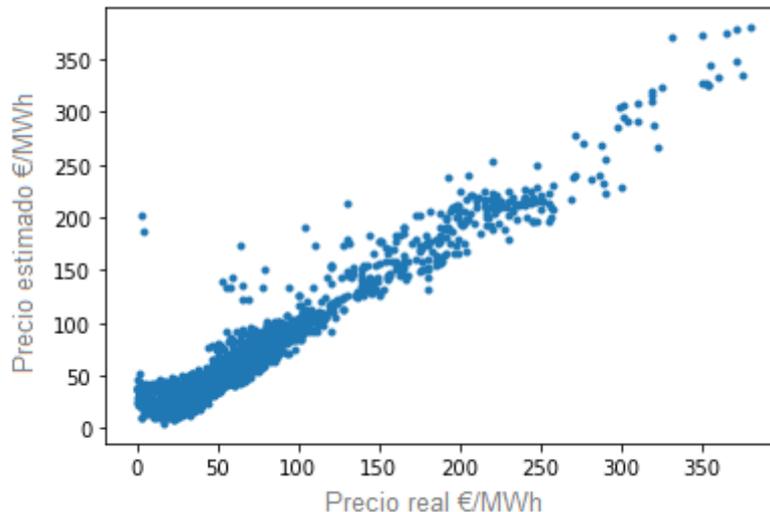
Con estos valores se aprecia que, a pesar de que el modelo se ha simplificado (al haber eliminado varias variables) el ajuste conserva las propiedades iniciales, y es con estos valores que se determina el modelo final con el que se conformará el precio *spot* de la electricidad.

Representando este nuevo ajuste de forma visual (Figura 6), se obtiene una gráfica muy similar a la obtenida en primera instancia, lo que reafirma el concepto de que el ajuste es muy similar al anterior, aunque el modelo haya sido simplificado.



**Figura 6:** Representación del segundo ajuste por regresión lineal de tipo Lasso.

Finalmente, si se representan gráficamente los valores estimados por el modelo frente a los valores de precio reales (Figura 7), se puede apreciar el alto grado de linealidad, y por ende de correlación, que estos valores poseen. Cabe mencionar que el caso de los valores de precios reales muy bajos (algunos cercanos a cero) generan en la regresión una desviación importante debido a la naturaleza del cálculo de esta, lo que puede explicar los valores fuera de rango que se aprecian.



**Figura 7:** Gráfico de correlación de valores estimados (modelo) versus valores reales.

## 3.2 Aplicación de un método de previsión

### 3.2.1 Previsión de variables para el escenario base – método ARIMA

Una vez determinadas las variables principales y realizado el ajuste que lleva al modelo definitivo (que se utilizará para la determinación del precio), se procedió a realizar la previsión de los valores de cada una de las variables. Con esto, se obtendrán valores a futuro de cada una de ellas, que servirán de entrada al modelo

ya conformado con el método de regresión lineal tipo Lasso. En este estudio, se ha realizado una previsión con un horizonte de 6 meses. A este escenario se le denominará Escenario Base, para diferenciarlo de los tres escenarios hipotéticos que se plantearán a continuación.

Para realizar las previsiones se implementó un modelo ARIMA, el cual se aplica a series temporales univariadas, razón por la cual se aplicó a todas las variables una por una.

A la hora de aplicar el modelo, se ha debido reescalar la muestra (de datos horarios a datos mensuales) por dos razones: por un lado, esto disminuye la carga de cálculo requerida para poner en práctica el modelo (el cual de otra forma requiere no solo un gran poder computacional, sino que también consume una gran cantidad de tiempo), y, por otro lado, porque esta metodología es más adecuada para previsiones con períodos mensuales.

Una vez reescalados los datos, se procedió a utilizar una funcionalidad (denominada AutoARIMA, la cual viene incorporada en el entorno de trabajo) que permite analizar la muestra (la totalidad de la serie) y obtener de forma automática los coeficientes  $p$ ,  $d$  y  $q$ . Cabe mencionar que la determinación de estos coeficientes es de suma importancia para la correcta aplicación del método, ya que estos condicionarán el ajuste correspondiente. De este modo, se evita introducir errores al estimar los coeficientes de manera manual. De igual manera, llegado el momento de aplicar propiamente el método ARIMA, se comprobó la idoneidad de los valores obtenidos de los coeficientes, y se corrigió manualmente en caso de que los mismos fueran inconsistentes con la variable a proyectar (por ejemplo, si los coeficientes llevaran a una proyección con valores negativos de potencia instalada o volumen generado).

Los valores obtenidos por AutoARIMA y los finalmente utilizados se recopilan (Tabla 5) a continuación:

**Tabla 5:** coeficientes del método ARIMA obtenidos por AutoARIMA y los finalmente utilizados.

Variable	Coef. AutoARIMA (p, d, q)	Coef. aplicado (p, d, q)
Generación por tecnología de carbón	(0,1,0)	(2,1,0)
Generación por tecnología nuclear	(0,0,1)	(2,0,1)
Generación por tecnología hidroeléctrica	(1,1,1)	(2,1,0)
Generación por tecnología de ciclo combinado	(0,1,0)	(2,0,0)
Generación por tecnología eólica	(1,0,0)	(1,0,0)
Generación por tecnología solar térmica	(1,1,0)	(4,1,0)
Generación por tecnología solar fotovoltaica	(1,0,0)	(2,0,0)
Generación por tecnología de cogeneración y otros	(0,1,1)	(0,1,1)
Importación de energía eléctrica	(1,0,0)	(0,1,0)
Potencia de tecnología nuclear	(0,1,0)	(0,1,0)
Potencia de tecnología de carbón	(2,2,1)	(2,2,1)
Potencia de tecnología de ciclo combinado	(0,1,0)	(0,1,0)
Potencia de tecnología eólica	(1,1,3)	(1,1,3)
Potencia de tecnología solar fotovoltaica	(1,2,1)	(1,2,1)

Potencia de tecnología de cogeneración y otros	(0,1,0)	(0,1,0)
Precio del gas natural	(2,2,0)	(2,2,0)

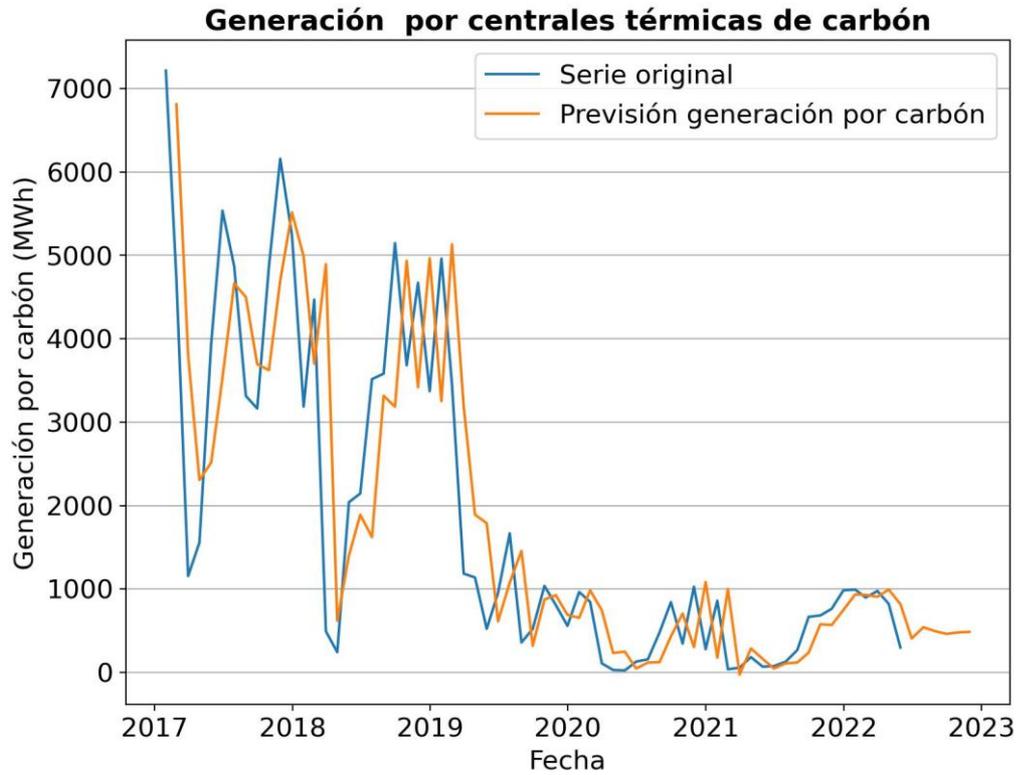
Por último, ya determinados los coeficientes, se implementó propiamente el método ARIMA para cada una de las variables, y se obtuvo una previsión de precio a 6 meses. A continuación, se analizan dichas variables a lo largo del período temporal estudiado, y se grafica superpuesto el valor de las previsiones hechas (a través del método ARIMA) para el mismo período, extendido 6 meses más. Para una mejor interpretación, este análisis se categoriza dividiendo a las variables según su naturaleza.

Cabe mencionar que, en algunas de las previsiones, a pesar de un buen ajuste en la curva de previsión se observa un retraso o *delay* en los valores. Este fenómeno es inherente al modelo ARIMA, es una concesión o *trade-off* por la estabilidad y la capacidad de respuesta del modelo, que resulta ser clave para los modelos a largo plazo.

### 3.2.1.1 Previsión de generación e importación

Los volúmenes de generación de las distintas tecnologías analizadas resultan muy dispares a lo largo del tiempo, ya sea por cuestiones operativas (como las centrales de ciclo combinado, que usualmente entran al mercado como la última tecnología para cubrir los picos de demanda), como así también por cuestiones intrínsecas a su principio de funcionamiento (como las tecnologías que dependen de recursos renovables ligados a las condiciones meteorológicas). Esto hace que su modelado y previsión resulten en una mayor complejidad, lo que se traduce en el hecho de que el modelo en algunos casos presenta un muy buen ajuste a la curva de generación, y en otros casos no es capaz de captar dichas variaciones. Particularmente en las tecnologías de generación solar (térmica y fotovoltaica), se aprecia que el modelo presenta un muy buen ajuste ya que interpreta correctamente la estacionalidad que las caracteriza.

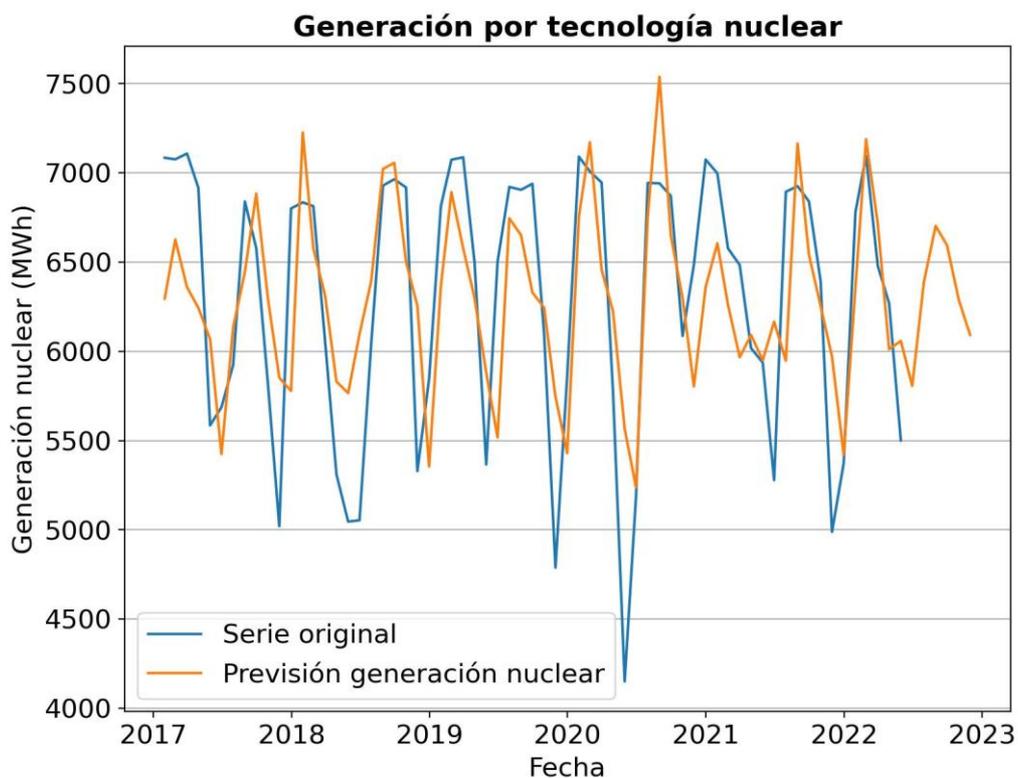
- Generación por tecnología de carbón



**Figura 8:** Gráfico de estimación de generación por tecnología de carbón vs. valores reales, y previsión a seis meses.

En este caso, se observa que la predicción hecha por el modelo (Figura 8) se ajusta en buena medida a los valores reales, haciendo una previsión casi constante de la producción en base a carbón para los meses proyectados.

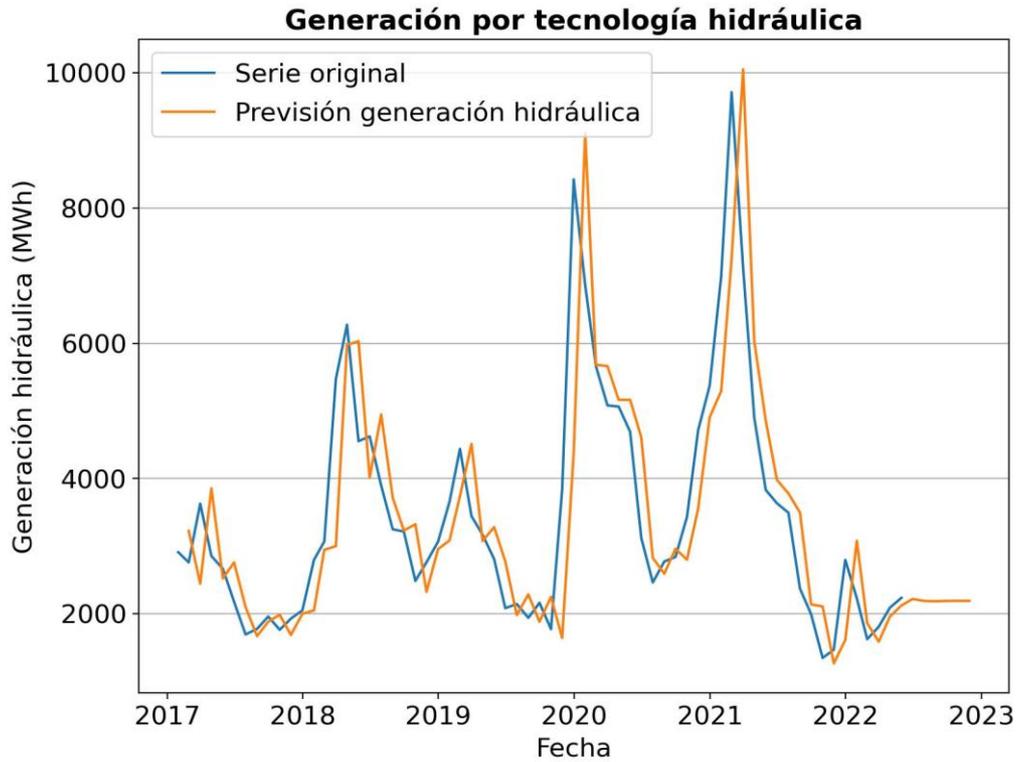
- Generación por tecnología nuclear



**Figura 9:** Gráfico de estimación de generación por tecnología nuclear vs. valores reales, y previsión a seis meses.

En el caso de la generación nuclear (Figura 9), el modelo no es lo suficientemente capaz de captar la amplitud de las oscilaciones propias de la curva de generación de esta tecnología. Es precisamente esta gran variabilidad la que puede ser la causa de esto.

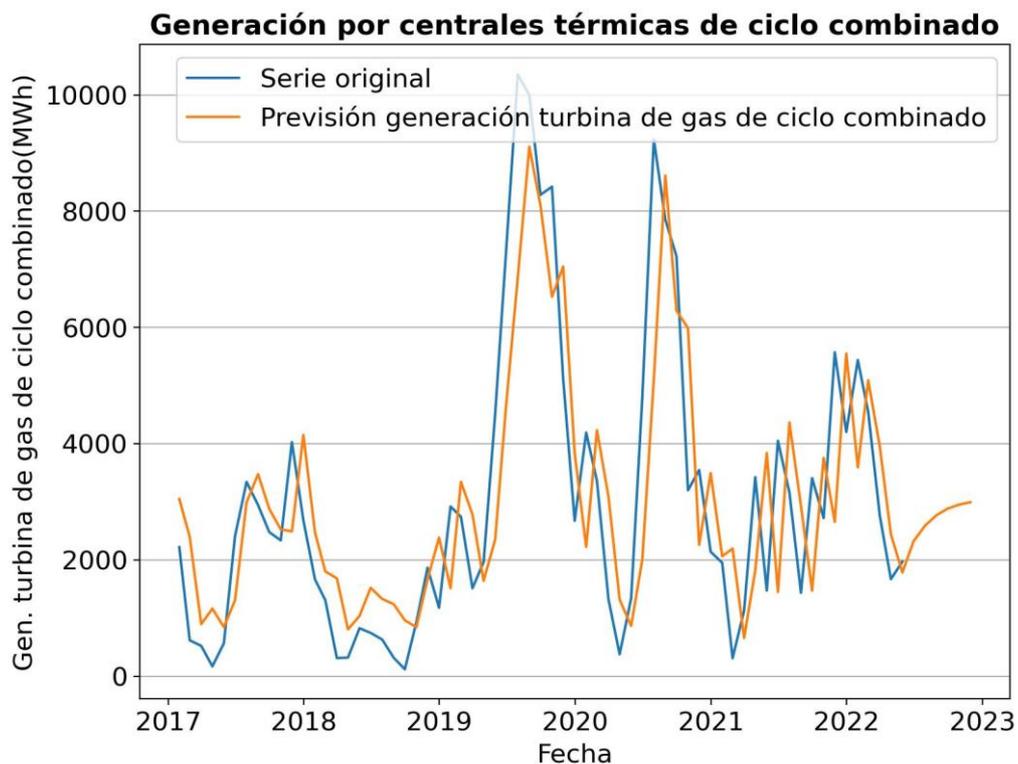
- Generación por tecnología hidroeléctrica



**Figura 10:** Gráfico de estimación de generación por tecnología hidroeléctrica vs. valores reales, y previsión a seis meses.

En el caso de la generación hidroeléctrica (Figura 10), el modelo es capaz de captar la estacionalidad y variaciones en la misma, y hace una proyección casi constante de la producción en el período estudiado.

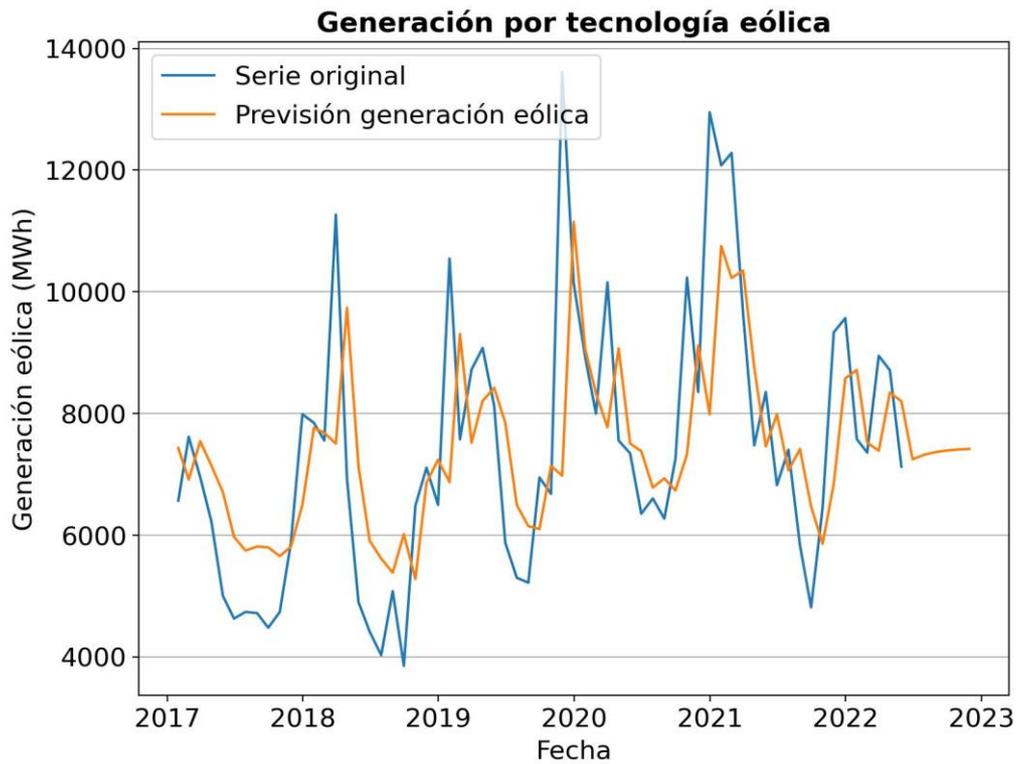
- Generación por tecnología de ciclo combinado



**Figura 11:** Gráfico de estimación de generación por tecnología de ciclo combinado vs. valores reales, y previsión a seis meses.

Para las centrales térmicas de ciclo combinado (Figura 11), se observa que el modelo también se ajusta en buena medida a la curva de producción, realizando una proyección al alza para el horizonte de 6 meses.

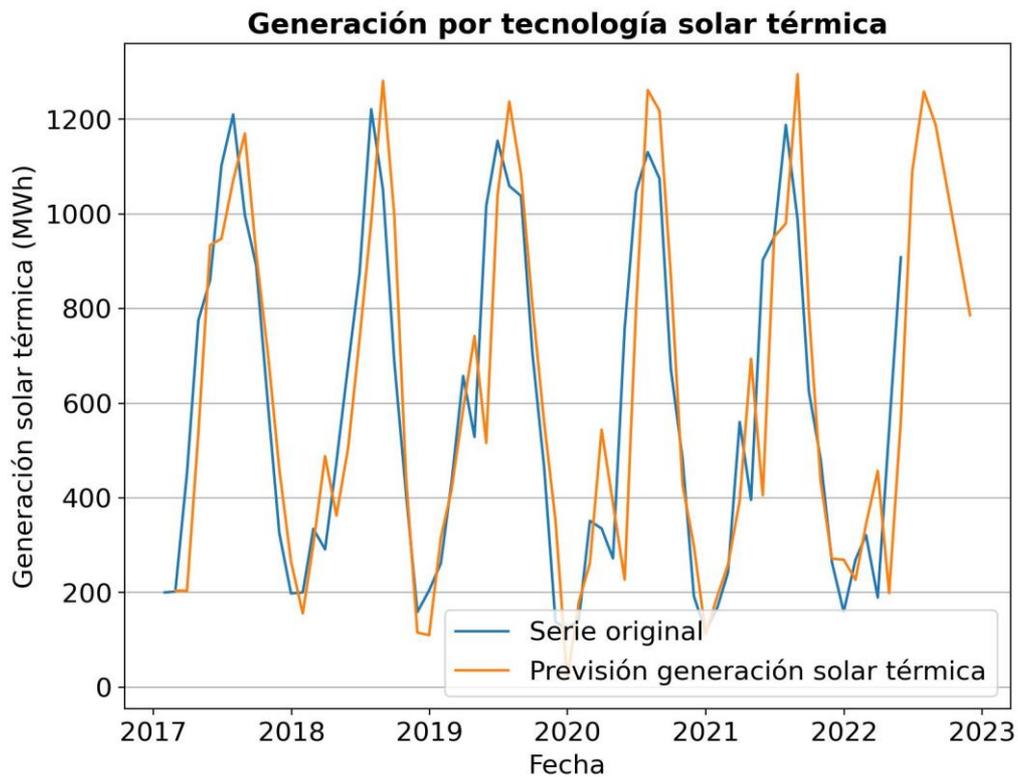
- Generación por tecnología eólica



**Figura 12:** Gráfico de estimación de generación por tecnología eólica vs. valores reales, y previsión a seis meses.

En lo que respecta a la generación eólica (Figura 12), se aprecia que el modelo asimila correctamente las variaciones estacionales de la misma, aunque no logra captar la amplitud de las oscilaciones en la curva de producción, proyectando valores de generación que resultan casi constantes para el período proyectado de 6 meses.

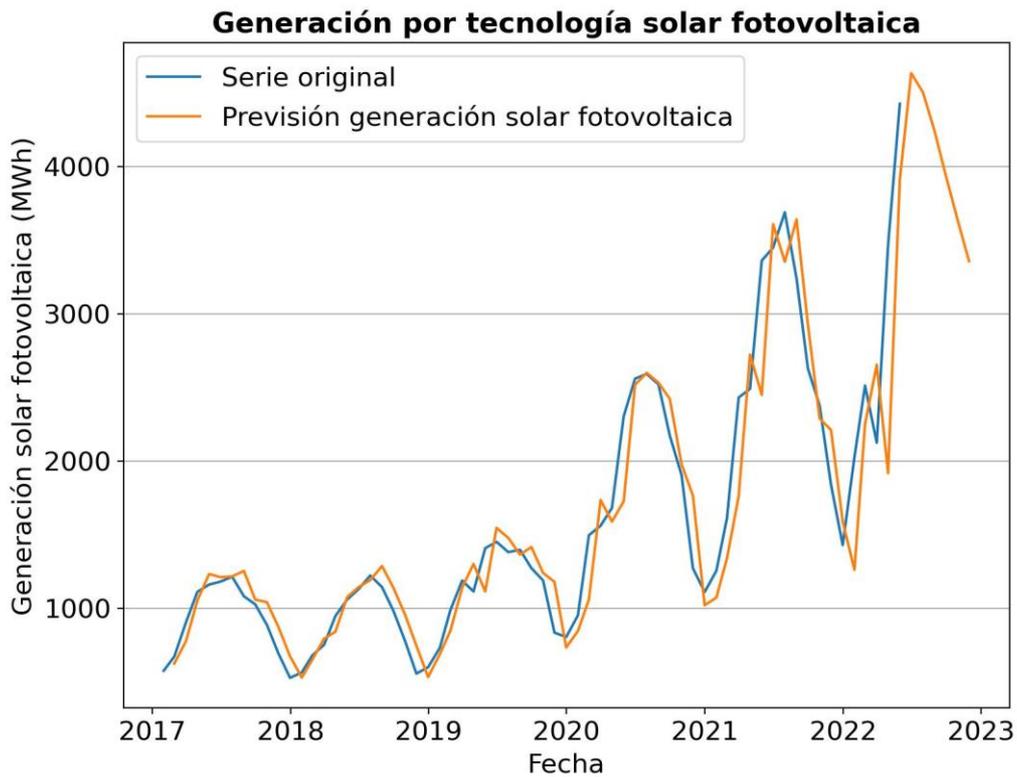
- Generación por tecnología solar térmica



**Figura 13:** Gráfico de estimación de generación por tecnología solar térmica vs. valores reales, y previsión a seis meses.

En lo que hace a la tecnología solar térmica (Figura 13), se encuentra la particularidad de que la misma casi no ha variado su potencia instalada en el período estudiado, por lo que los valores de generación dependen casi exclusivamente de los factores meteorológicos, mostrando una gran estacionalidad. En consecuencia, los valores son claramente estacionales, y el modelo interpreta correctamente esto, previendo que se alcance un pico de producción en el mes de julio (característico de este mes, el de mayor radiación solar en el año), para luego ir disminuyendo en los meses posteriores.

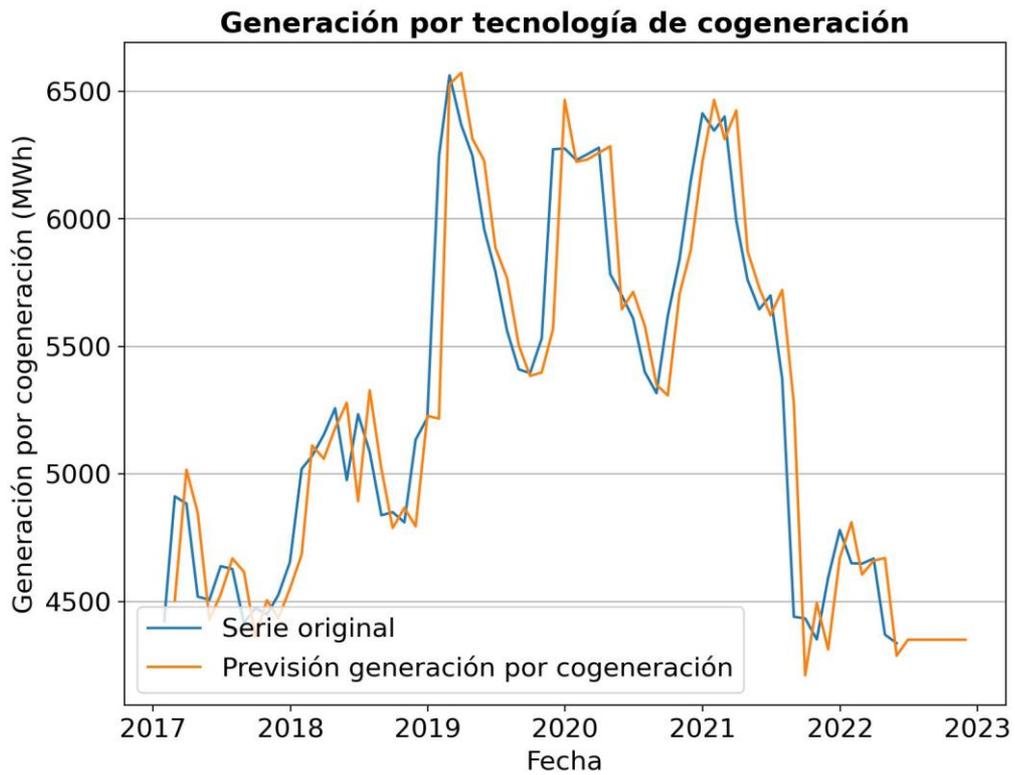
- Generación por tecnología solar fotovoltaica



**Figura 14:** Gráfico de estimación de generación por tecnología solar fotovoltaica vs. valores reales, y previsión a seis meses.

Respecto a la generación solar fotovoltaica (Figura 14), se aprecia un patrón estacional como en el caso anterior, pero con el agregado del aumento sostenido de la potencia instalada a lo largo del período analizado (ver Figura 21 sobre potencia instalada de tecnología solar fotovoltaica). En consecuencia, este tipo de generación presenta picos (coincidentes con los meses de mayor radiación solar) cada vez más elevados, los cuales son correctamente interpretados por el modelo que. En este caso, el modelo prevé (como en el anterior caso de la generación solar térmica) valores de generación con un pico de producción en el mes de julio (característico de este mes de mayor radiación solar en el año para España), para luego ir disminuyendo en meses subsiguientes, acorde a la época del año.

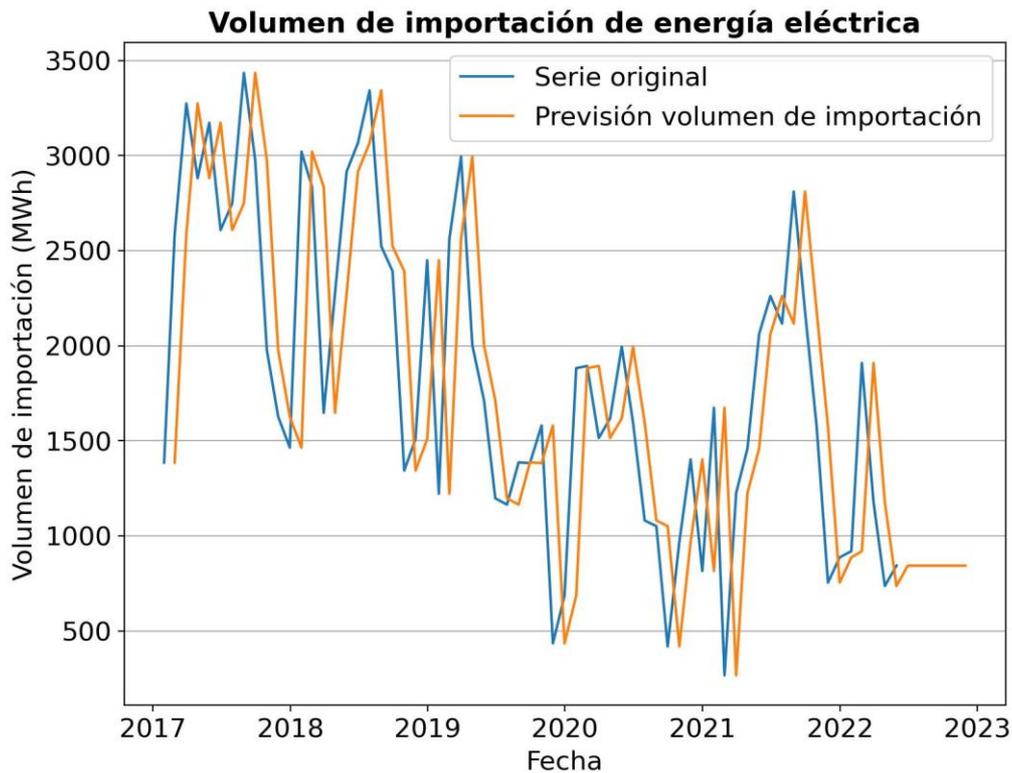
- Generación por tecnología de cogeneración y otros



**Figura 15:** Gráfico de estimación de generación por tecnología de cogeneración vs. valores reales, y previsión a seis meses.

Para el caso de la cogeneración (Figura 15), se puede apreciar que el modelo se ajusta de gran manera a la curva de producción, sin embargo, la proyección para los meses siguientes presenta un valor cuasi constante.

- Importación de energía eléctrica



**Figura 16:** Gráfico de estimación de importación de energía eléctrica vs. valores reales, y previsión a seis meses.

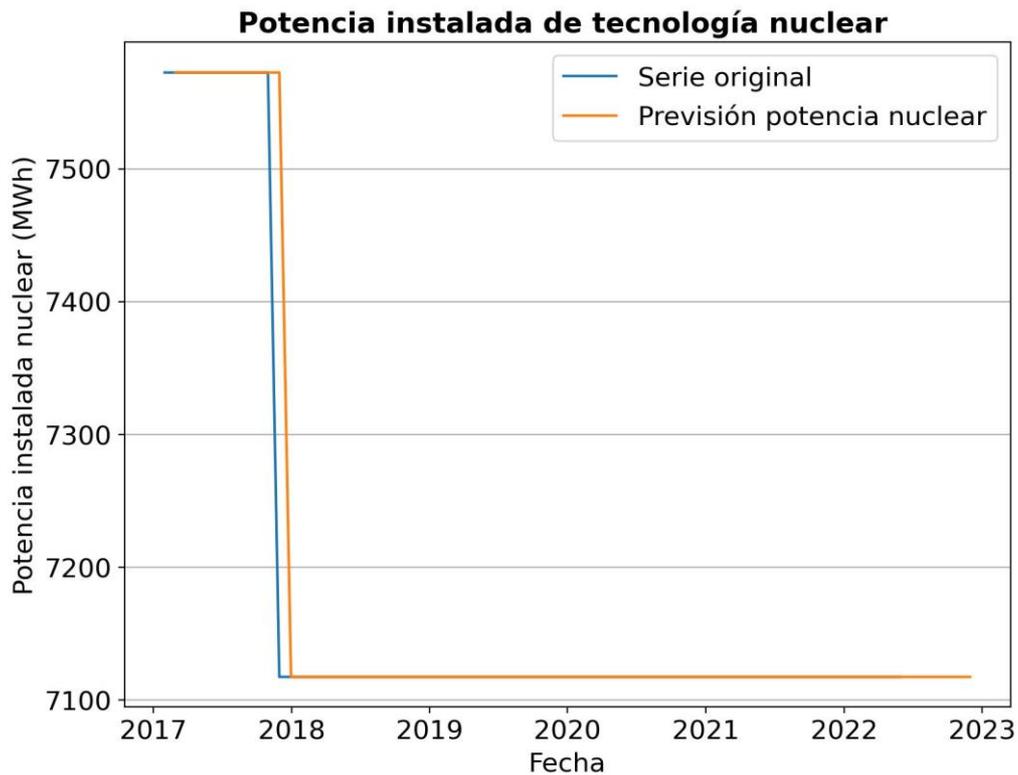
Analizando el caso de la importación de energía eléctrica (Figura 16), se puede ver que se produce algo similar al caso anterior, donde el modelo estima adecuadamente los valores a lo largo del período estudiado, pero a la hora de realizar la proyección, estos presentan un valor casi constante.

### 3.2.1.2 Previsión de potencia instalada

En el caso de la potencia de las distintas tecnologías involucradas en este análisis, se observa, por un lado, tecnologías cuya potencia instalada apenas ha variado durante el último lustro (centrales de ciclo combinado o nucleares), y, por otro lado, tecnologías que han visto variar su potencia instalada de forma notoria (centrales térmicas de carbón, eólica y solar fotovoltaica).

Para las primeras, se aprecia que el modelado es más sencillo y la previsión resulta en valores constantes, y para las segundas, el modelo igualmente presenta un muy buen ajuste, pero la previsión varía en función de la tendencia de cada una de estas tecnologías.

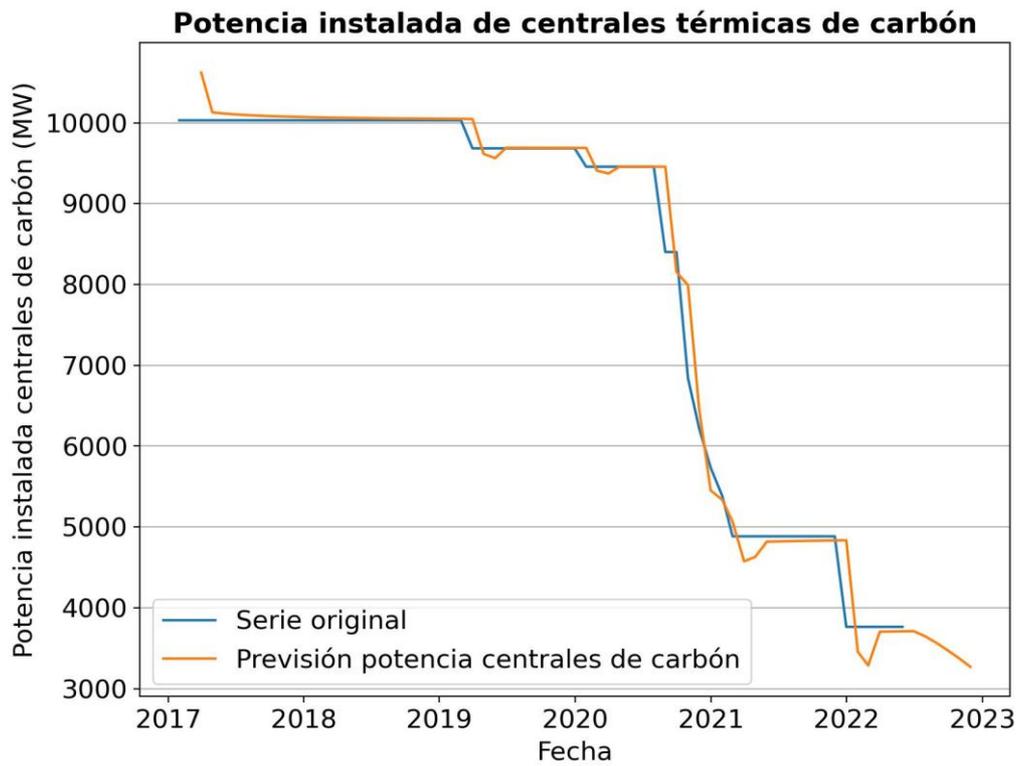
- Potencia de tecnología nuclear



**Figura 17:** Gráfico de estimación de potencia instalada de tecnología nuclear vs. valores reales, y previsión a seis meses.

En lo que respecta a potencia instalada nuclear (Figura 17), el modelo estima con gran precisión los valores de generación y realiza una previsión constante de la misma, cuya razón radica en el hecho de que la potencia instalada de esta tecnología ha sufrido una sola variación en el período analizado, por lo que el cálculo se simplifica de gran manera.

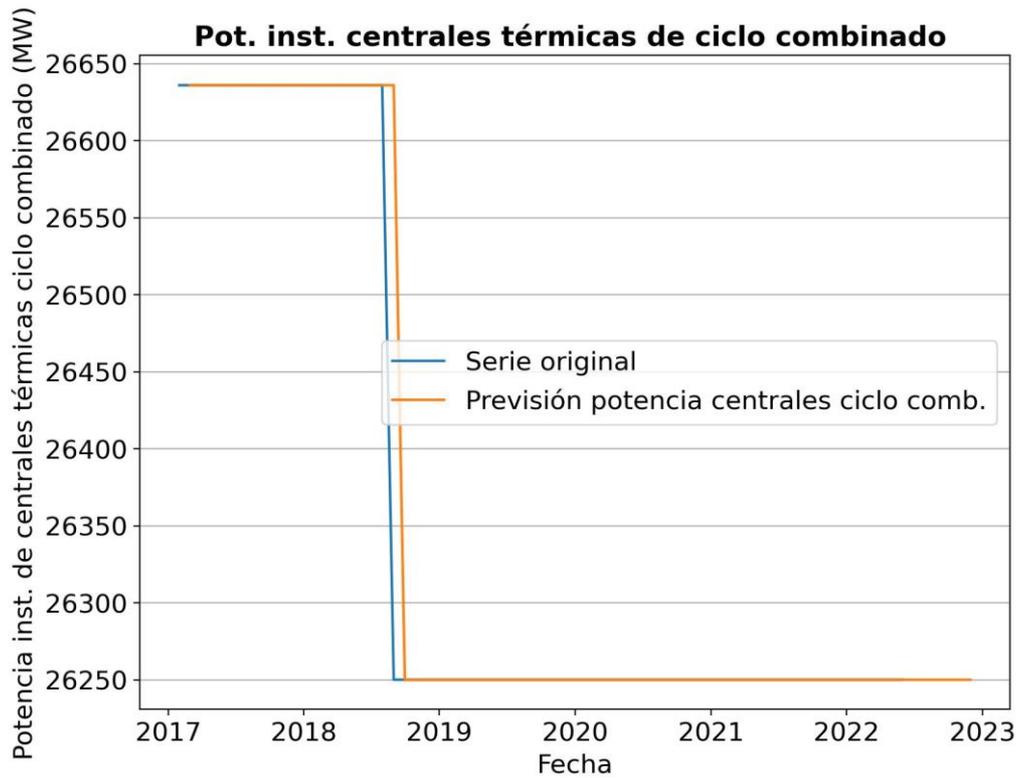
- Potencia de tecnología de carbón



**Figura 18:** Gráfico de estimación de potencia instalada de tecnología de carbón vs. valores reales, y previsión a seis meses.

Para el caso de la potencia instalada de centrales térmicas de carbón (Figura 18) se observa, como en otros casos, que el modelo se ajusta en gran medida a la variación de esta, y a su vez la previsión que realiza resulta coherente con la tendencia general.

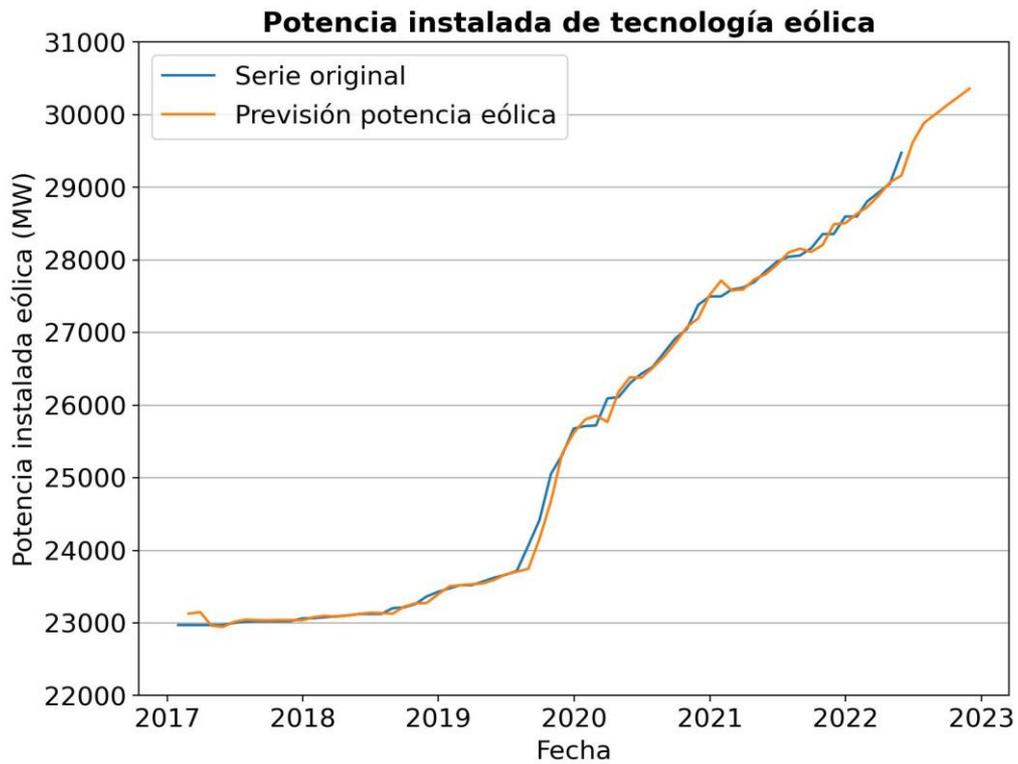
- Potencia de tecnología de ciclo combinado



**Figura 19:** Gráfico de estimación de potencia instalada de tecnología de ciclo combinado vs. valores reales, y previsión a seis meses.

El caso de la potencia instalada de centrales térmicas de ciclo combinado (Figura 19), se observa una gran similitud con el caso de la potencia instalada nuclear. El modelo estima con gran precisión los valores de generación y previsión resulta en valores constantes, ya que este valor apenas ha variado en el período analizado, por lo que el modelado resulta ser más simple.

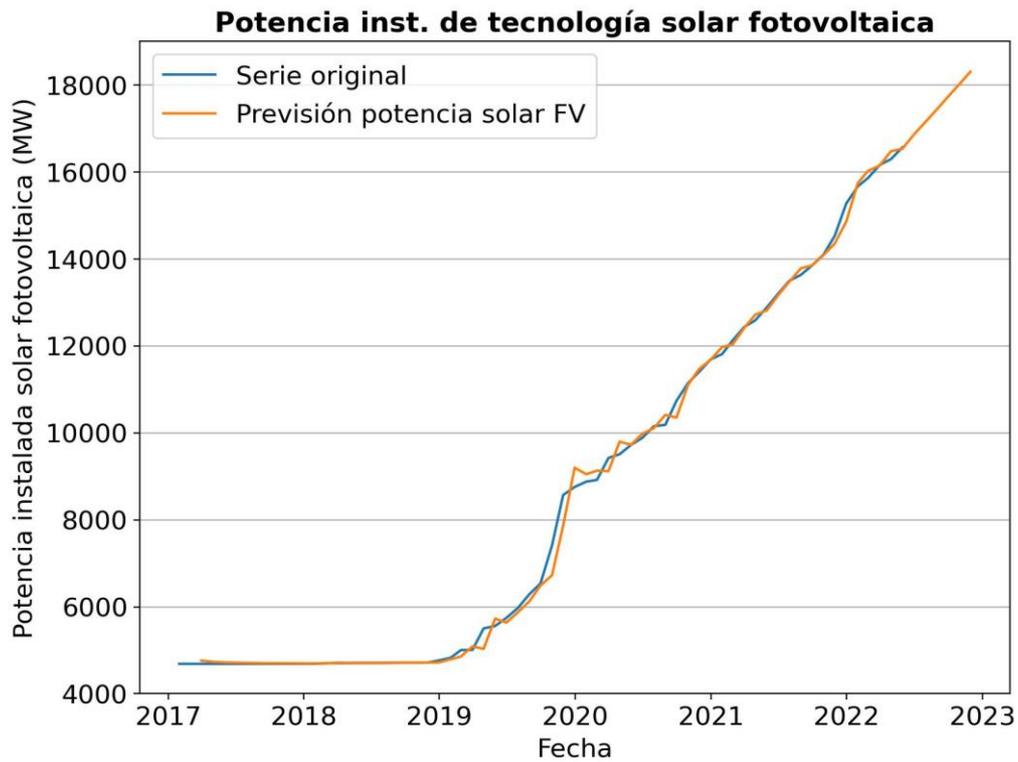
- Potencia de tecnología eólica



**Figura 20:** Gráfico de estimación de potencia instalada de tecnología eólica vs. valores reales, y previsión a seis meses.

La potencia instalada de tecnología eólica (Figura 20) ha aumentado de manera considerable durante el último lustro en el país, situación que se ve reflejada en la gráfica. El modelo estima con gran precisión esta evolución y realiza a su vez una previsión al alza de este valor.

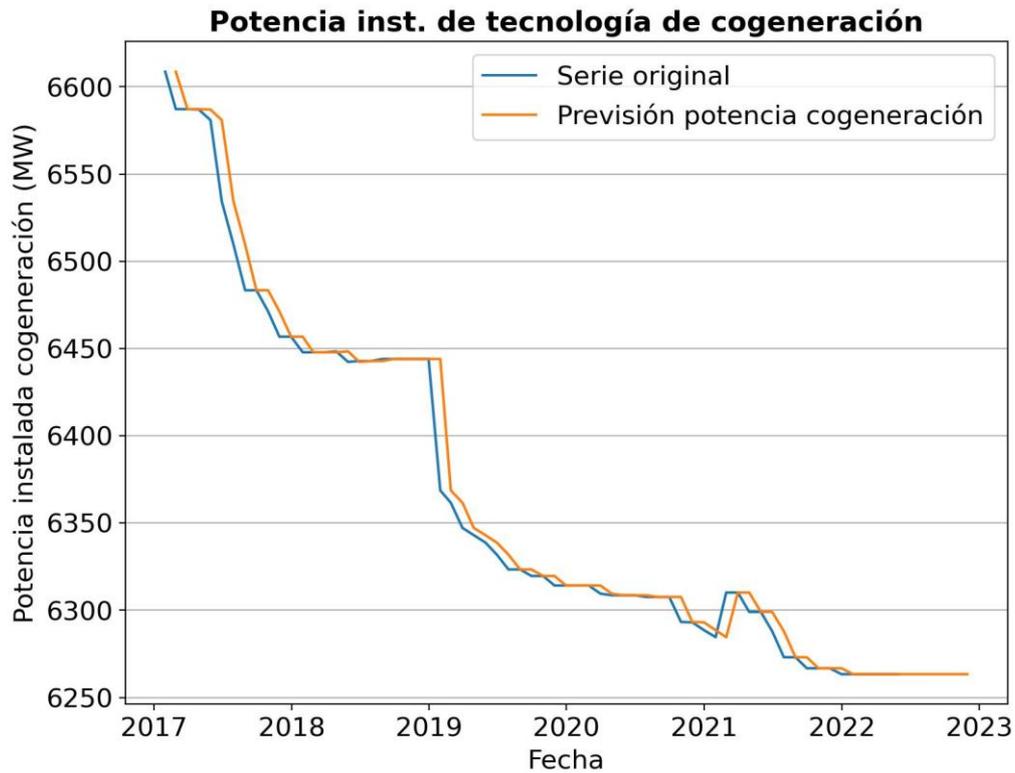
- Potencia de tecnología solar fotovoltaica



**Figura 21:** Gráfico de estimación de potencia instalada de tecnología solar fotovoltaica vs. valores reales, y previsión a seis meses.

En lo que respecta a la potencia instalada de tecnología solar fotovoltaica (Figura 21), se aprecia una similitud evidente (a diferente escala) con el caso de la tecnología eólica, habiendo registrado un crecimiento considerable en el período analizado. Análogamente a lo sucedido con el caso de la tecnología eólica, el modelo presenta un muy buen ajuste respecto a esta evolución, y realiza un pronóstico al alza para estos valores.

- Potencia de tecnología de cogeneración



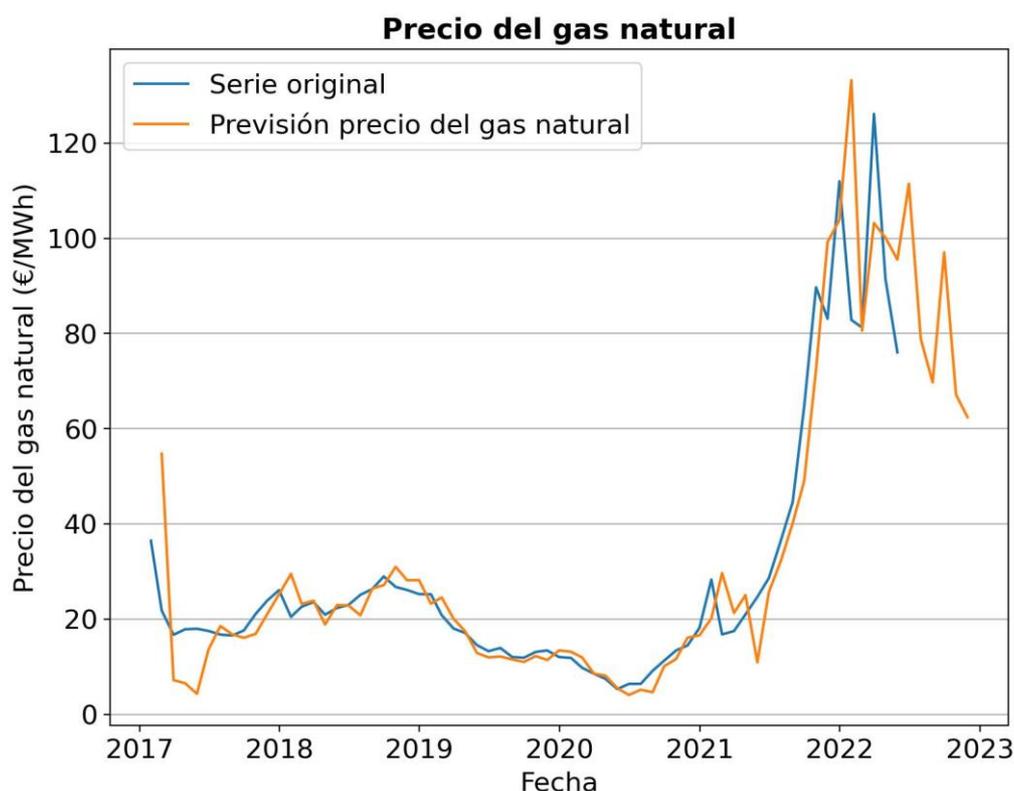
**Figura 22:** Gráfico de estimación de potencia instalada de tecnología de cogeneración vs. valores reales, y previsión a seis meses.

La potencia instalada de tecnología de cogeneración (Figura 22), se aprecia una similitud evidente con otros casos modelados previamente donde el modelo presenta un muy buen ajuste respecto a esta evolución, pero a la hora de realizar la proyección, estos presentan un valor casi constante.

### 3.2.1.3 Previsión del precio del gas natural

El caso del gas natural, si bien se trata de un valor de precio (no de volumen ni potencia de generación, como los antes analizados), presenta igualmente un buen ajuste del modelo.

- Precio de gas natural



**Figura 23:** Gráfico de estimación de precio del gas natural vs. valores reales, y previsión a seis meses.

En lo que respecta al precio del gas natural (Figura 23), puede verse que ha habido una sustancial variación en la cotización de esta *commodity* durante el período estudiado. El modelo presenta un muy buen ajuste a dicha variación, previendo a su vez valores fluctuantes para el período proyectado, con una tendencia general a la baja.

Una vez realizadas las previsiones para cada una de las variables, se conforma con las mismas un conjunto de datos que servirá de entrada para el modelo generado en el apartado 3.1.2 (a través de la regresión lineal de tipo Lasso), con el cual se realizará la conformación del precio final de la electricidad en el que se denominará “Escenario base”, es decir, el supuesto caso en que las variables sigan su evolución esperada (lo que usualmente se conoce como un escenario “*business-as-usual*”).

### 3.3 Elaboración de escenarios futuros

En esta instancia, se procedió a elaborar 3 escenarios hipotéticos a futuro (6 meses) variando el aporte de distintas variables. Con esto se buscó estudiar las posibles implicaciones de diversos factores que condicionan al parque generador y demás variables. En consecuencia, los escenarios planteados fueron los siguientes:

1. Reducción de la potencia nuclear instalada y aumento de las renovables: el Plan Nacional Integrado de Energía y Clima (conocido por sus siglas PNIEC,

un plan del gobierno de España elaborado para dar respuesta a los desafíos en estas áreas de cara a los años venideros) prevé una disminución de la potencia nuclear instalada del 57% para el 2030, por lo que se adopta esta hipótesis para el primer escenario. De forma análoga, también se disminuye el volumen de generación nuclear. Por otro lado, se plantea una compensación de dicho déficit con un incremento en la potencia y generación solar fotovoltaica y eólica, ya que la primera ha crecido en una tasa del 30% en el período 2020-2021 y la segunda un 7% en el mismo período, por lo que se aplican estas tasas en la proyección. Los restantes valores de potencia instalada se mantienen constantes, y para los valores de generación (que no fueran los mencionados) se adoptó la media del período observado.

2. Disminución del precio de gas natural: en este escenario se plantea una disminución del 50% del precio de esta *commodity* (hasta llegar al valor promedio del período analizado), debido a los diversos factores coyunturales que indican que precio (luego de aumentar sustancialmente) se estabilizará en valores históricos. Entre estos se incluye: la aparición de nuevos proveedores en el mercado europeo, el llenado de las reservas de gas a valores de casi el 90% en la mayoría del continente, la reacción de los mercados a las medidas propuestas por la Comisión Europea, entre otros. Para los valores de potencia, los mismos se han supuesto constantes, y para los valores de generación, se adoptó la media del período observado.
3. Disminución de un 25% de la generación hidroeléctrica en el período considerado (con respecto al valor promedio), normalización del precio del gas natural y crecimiento de las energías renovables. Para este escenario se supuso una disminución de la producción hidroeléctrica del 25%, en un contexto de escasez hídrica (como el que se vive actualmente y el que plantean varios pronósticos para los próximos años en la península ibérica), compensado con un aumento de la producción de las renovables de tipo eólica y solar fotovoltaica (como lo propuesto en el escenario 1), sumados a una posible normalización del precio del gas (análoga a lo propuesto en el escenario 2). Nuevamente, los valores de potencia (que no sean eólica ni fotovoltaica) se han supuesto constantes, y para los valores de generación (excepto las afectadas por esta hipótesis), se adoptó la media del período observado.

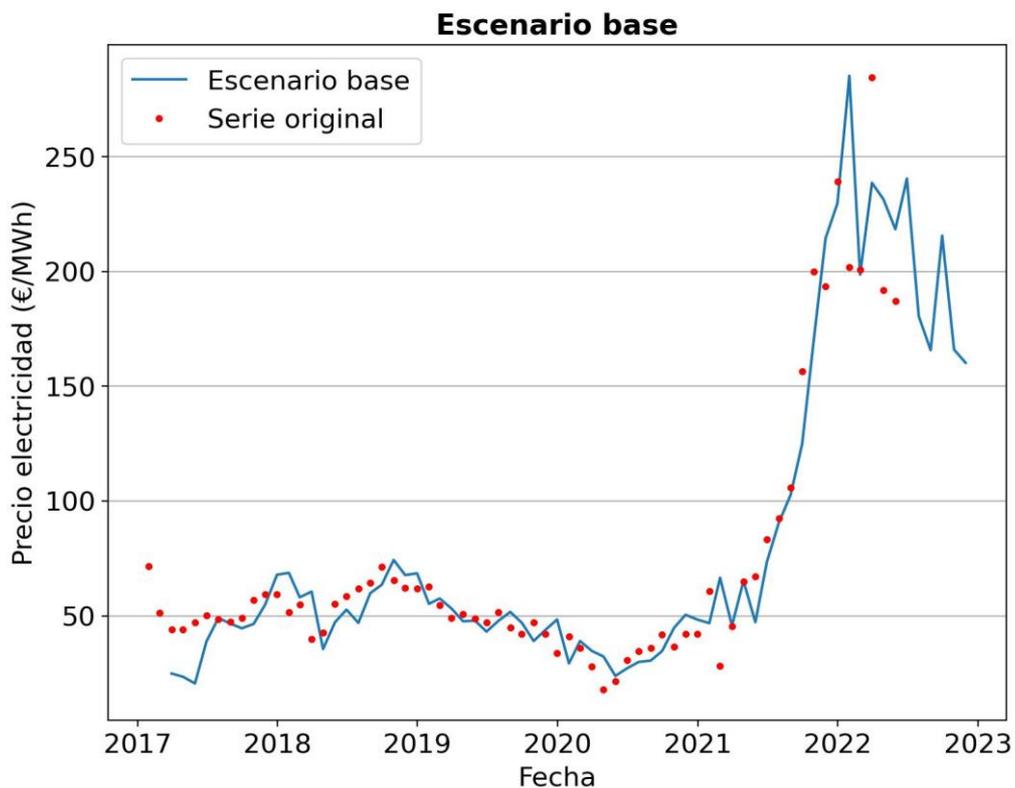
### **3.4 Previsión del precio final de la electricidad – Regresión lineal tipo Lasso y comparación con escenarios futuros**

Una vez realizada la previsión del escenario base (a través del modelo ARIMA), y ya planteados los escenarios hipotéticos a futuro, se procede a utilizar dichos valores como entradas para el modelo de regresión lineal definido en el punto 3.1.2, a los fines

de realizar la conformación del precio final de la energía en el mercado *spot* de MIBEL en un horizonte de 6 meses.

Para esto, primeramente, se realizó la modelación del precio en base a las previsiones realizadas para cada una de las variables, en el llamado "Escenario base". El resultado de las previsiones del modelo ARIMA para cada una de las variables es recogido en un nuevo conjunto de datos que tendrá entonces el valor de las 16 variables para todo el período estudiado, más el período proyectado de 6 meses. Este conjunto de datos es introducido en el modelo del precio elaborado anteriormente, y se obtienen los siguientes resultados:

- Escenario base



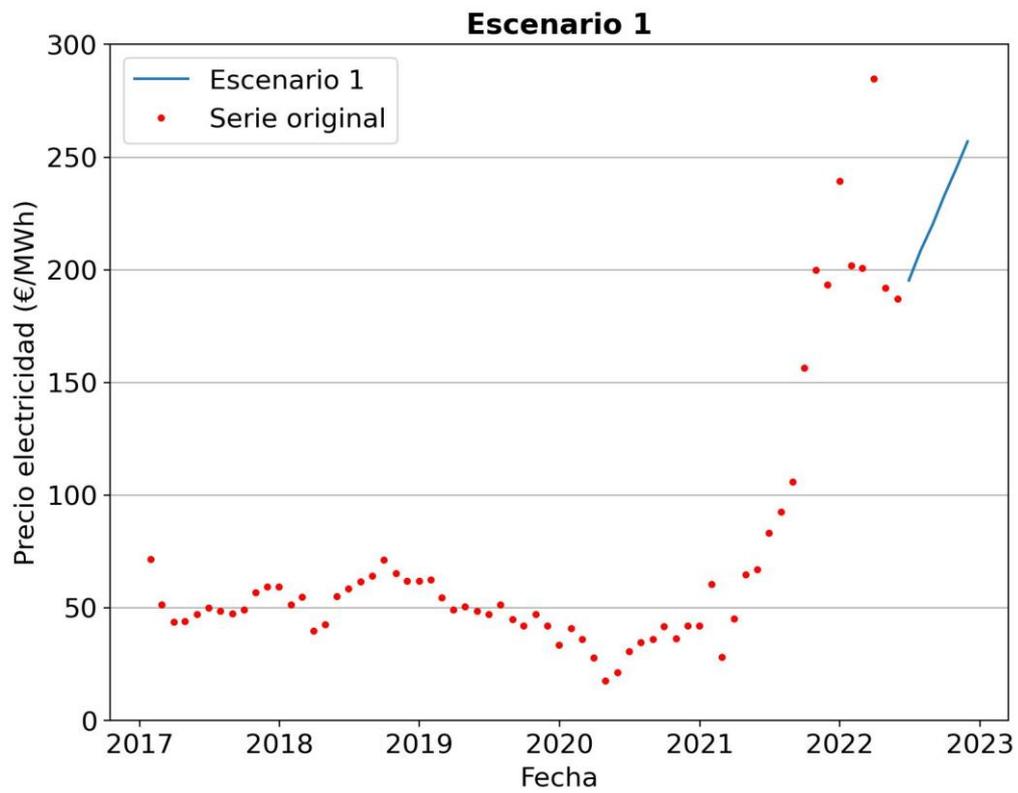
**Figura 24:** Previsión del precio de la electricidad en el "Escenario base".

Graficando la evolución de precios bajo el escenario base (Figura 24) se observa que en un caso de "*business-as-usual*", el precio del MWh alcanzaría valores que superarían los 250 € para luego fluctuar, pero ya siguiendo una tendencia descendente.

En una segunda instancia se realiza este mismo procedimiento, pero esta vez utilizando como valores de entrada para el modelo de precio los valores planteados en los escenarios hipotéticos.

Los resultados obtenidos fueron los siguientes:

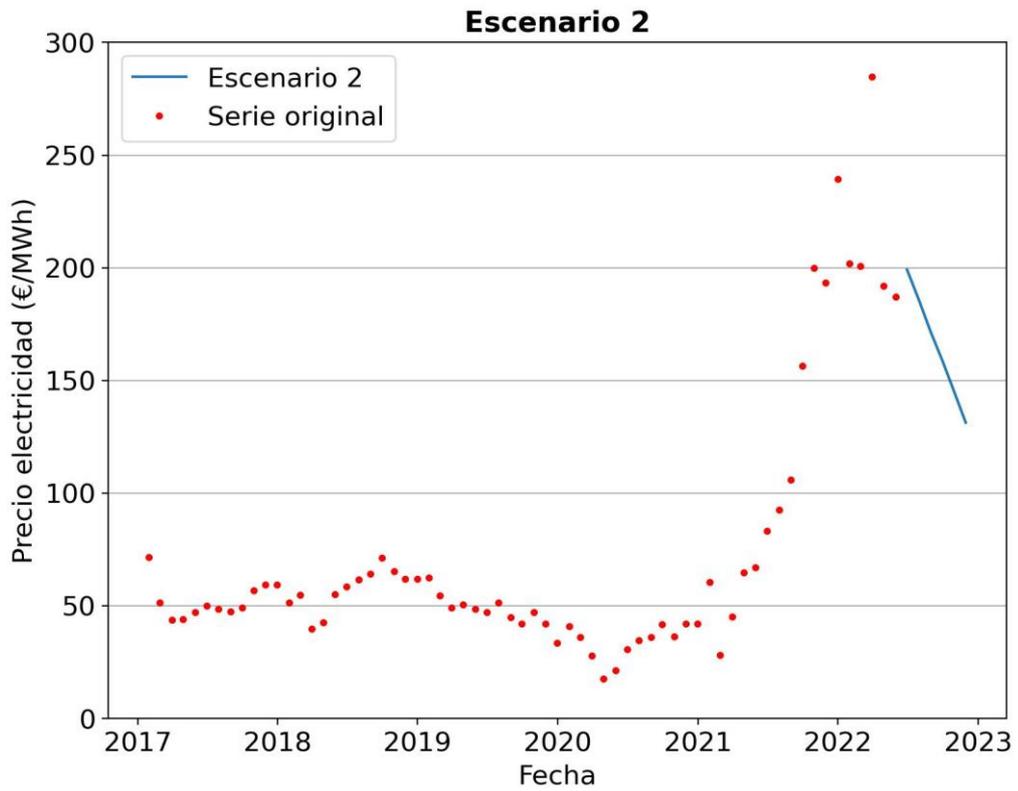
- Escenario 1:



**Figura 25:** Previsión del precio de la electricidad en el "Escenario 1".

En el caso del escenario 1 (Figura 25) se observa que, disminuyendo la producción nuclear y compensado con producción renovable, el valor de precio del MWh presenta una clara tendencia al alza.

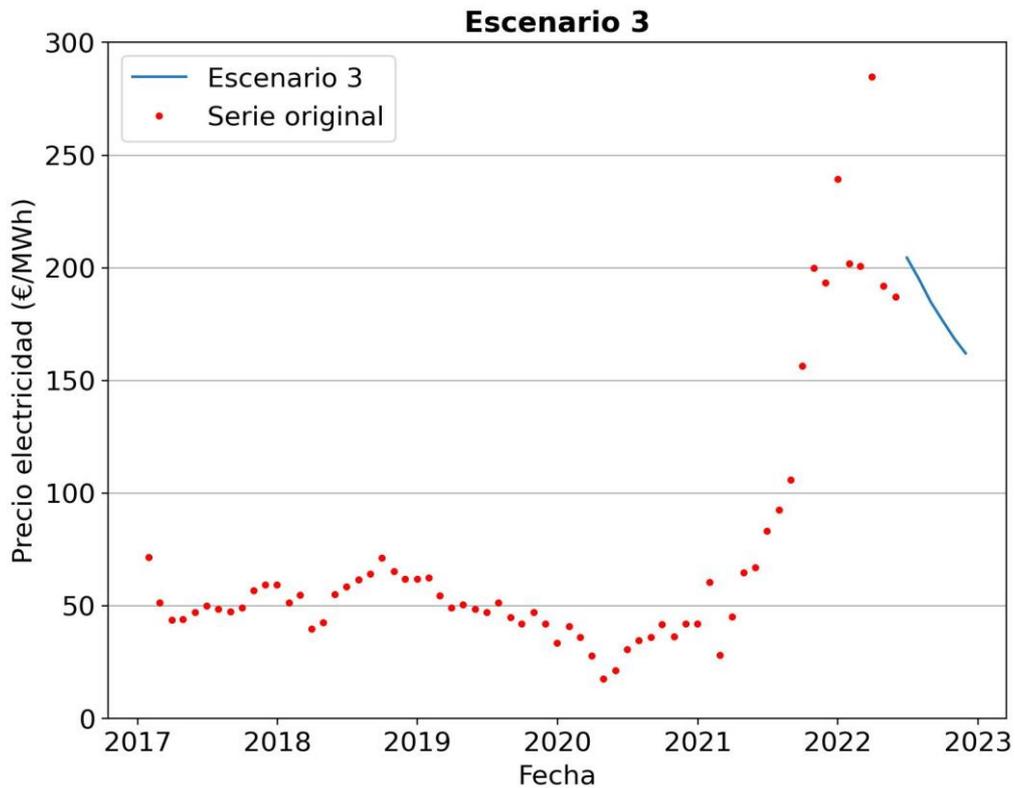
- Escenario 2:



**Figura 26:** Previsión del precio de la electricidad en el "Escenario 2".

En este segundo escenario hipotético (Figura 26), ligado a una normalización del precio del gas natural, se puede apreciar la influencia de esta *commodity* en precio final de la electricidad, ya que el precio del MWh experimentaría en este caso una disminución que va de la mano con la del gas natural.

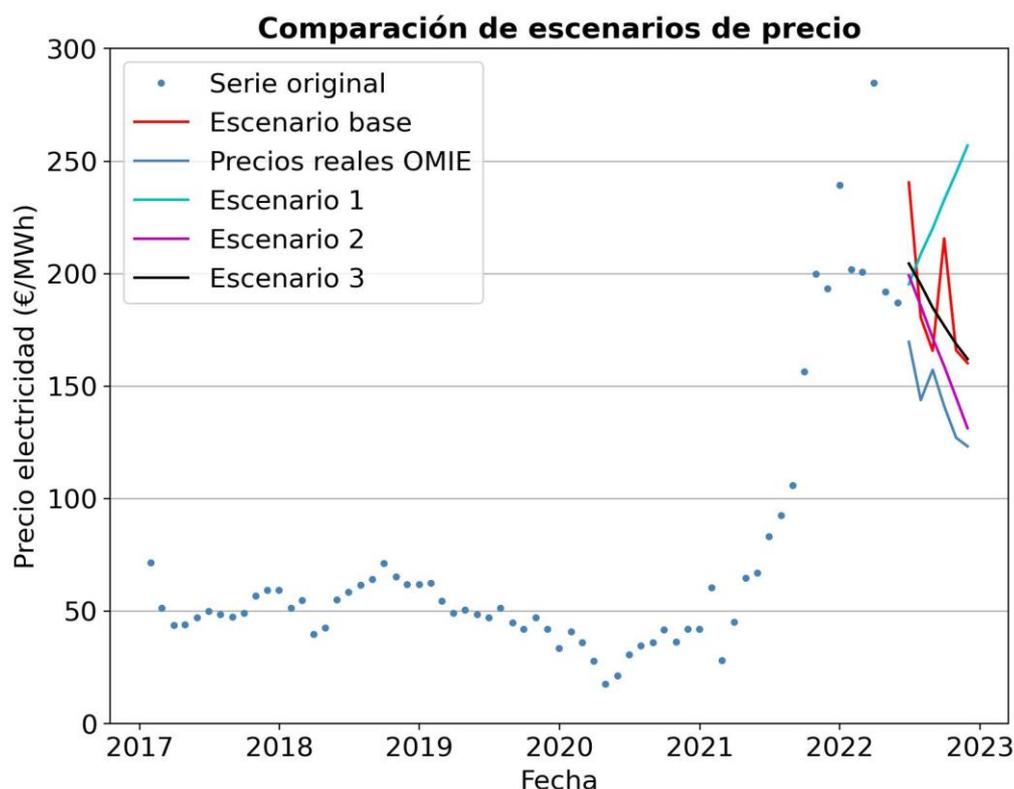
- Escenario 3:



**Figura 27:** Previsión del precio de la electricidad en el "Escenario 3".

En este último escenario hipotético (Figura 27), donde se plantea una disminución de la producción hidroeléctrica, compensada por un aumento de las renovables (análogo al del escenario 1) y una normalización del precio del gas natural (como la del escenario 2), se aprecia nuevamente una disminución del precio, sin embargo, aquí puede deberse tanto a la reducción del precio del gas natural como a otras razones (como el aumento de la potencia y generación renovable). De igual manera, considerando el peso que tiene el gas natural en la formación del precio (según lo analizado en el punto 3.1.2), se puede explicar entonces la disminución del precio final de la electricidad por la disminución del precio de la mencionada *commodity*.

- Comparación de escenarios



**Figura 28:** Comparación del precio de la electricidad en los distintos escenarios con su valor real.

**Tabla 6:** Comparativa del precio de la electricidad en los distintos escenarios, con respecto al valor real.

Período	Escenario base	Escenario 1	Escenario 2	Escenario 3	Precios reales
Junio 2022	240.53 €	195.33 €	199.20 €	204.51 €	169.72 €
Julio 2022	180.51 €	208.75 €	185.73 €	195.20 €	143.82 €
Agosto 2022	165.77 €	220.12 €	171.47 €	184.94 €	157.32 €
Septiembre 2022	215.64 €	232.79 €	158.95 €	176.88 €	141.19 €
Octubre 2022	166.00 €	244.76 €	145.03 €	168.80 €	127.15 €
Noviembre 2022	160.26 €	256.95 €	131.37 €	162.11 €	123.28 €

Contrastando entre sí los distintos escenarios analizados (Figura 28 y Tabla 6), observamos:

- el “Escenario base”, si bien logra prever el comportamiento de la curva de precio, lo hace con un cierto error. Esta diferencia, sin embargo, puede deberse a la Singularidad Ibérica (tope al precio del gas implementado en junio del 2022 como medida del gobierno para paliar el aumento de precios), lo que ha implicado fácticamente una disminución en el precio de la electricidad en el mercado *spot*;
- el “Escenario 1”, al presentar una curva ascendente de precios, diverge de lo ocurrido realmente, y es el que mayor diferencia presenta con los valores reales;
- el “Escenario 2”, si bien no representa la curva de precios reales, es aun así el que toma los valores más similares a dicho precio. Esto refuerza otra vez lo

mencionado al respecto del peso que tiene el precio del gas natural en la conformación del precio final de la electricidad;

- el “Escenario 3”, por último, presenta una curva de precios que queda entre medio de la planteada por los escenarios “1” y “2”, sin embargo, su forma es muy similar a la del escenario 2, lo cual se puede explicar en el hecho de que ambos tienen como factor común, la disminución del precio del gas natural.

De la comparación presentada, se puede apreciar que la previsión realizada a través del modelo ARIMA logra interpretar el comportamiento de la curva de precios, pero con el costo de tener un error asociado. Por otro lado, solamente los escenarios en donde se ha planteado una disminución del precio del gas (situación que se corresponde con lo efectivamente ocurrido) son aquellos que han mostrado correspondencia con los valores reales de precio de la electricidad, lo que refleja la influencia de este valor en la conformación del precio en el *pool* eléctrico.

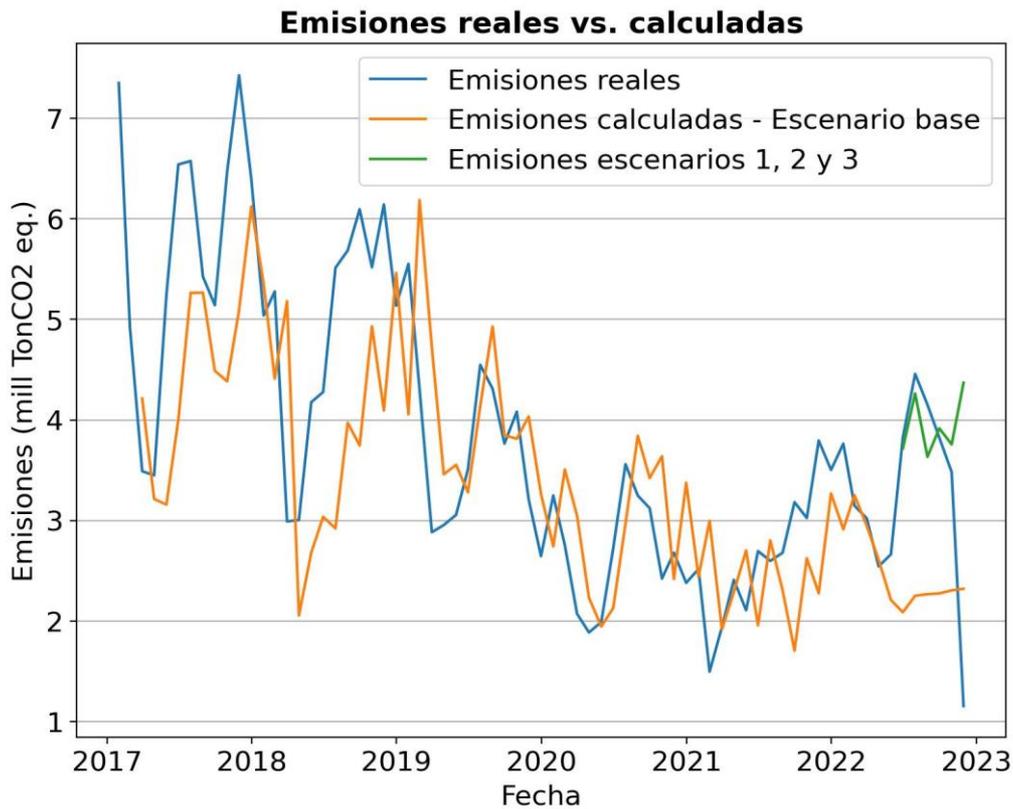
### 3.5 Cálculo de emisiones

Una vez realizadas las previsiones en función de los cuatro escenarios, y utilizando los factores de emisión publicados por Red Eléctrica de España, se procedió al cálculo de las emisiones generadas por las tecnologías emisoras que forman parte del parque de producción de España para todo el período temporal analizado (en el caso del “Escenario base”) y para el período proyectado de seis meses (para los 4 escenarios). Los valores de factores de emisión utilizados son los que se presentan en la siguiente tabla:

**Tabla 7:** factores de emisión publicados por REE.

<b>Tecnología de generación</b>	<b>Factor de emisión (tCO<sub>2</sub>eq./MWh)</b>
Central térmica de carbón	0,95
Central térmica de ciclo combinado (gas natural)	0,37
Cogeneración	0,38

Aplicando los factores correspondientes (Tabla 7) se obtienen entonces los valores de emisión (Figura 29).



**Figura 29:** Comparación de emisiones reales y calculadas en los escenarios.

De la gráfica se puede observar:

1. los valores de emisión estimados por el modelo ARIMA prevén una buena aproximación al comportamiento de la curva de emisiones reales, pero con un error asociado, en donde para el período proyectado hace una previsión por debajo de los valores reales;
2. los valores de emisiones de los 3 escenarios hipotéticos planteados son iguales, debido a que en dichos escenarios no se ha variado el aporte de las tecnologías emisoras de CO<sub>2</sub>;
3. estos valores se asemejan en magnitud a los valores de las emisiones reales (en mayor medida que los del “Escenario base”), lo que puede explicarse por el hecho de que estos valores resultan del promedio de la generación de las distintas tecnologías (para el período analizado), multiplicado por los factores de emisión;
4. los valores de emisiones (en línea con lo observado con respecto al aumento sostenido de potencia instalada y generación renovable) muestran una clara tendencia a la baja en el último lustro, con la excepción del presente año 2022 donde se observa un aumento de estos; y
5. los valores reales de las emisiones resultan complejos de prever debido a la multitud de factores que entran en juego en los mismos.

## CAPÍTULO 4. CONCLUSIONES

El objetivo central de este Trabajo Fin de Máster ha sido estudiar, analizar y determinar la contribución de distintas variables a la conformación del precio de final de la electricidad en el mercado eléctrico español, para poder realizar una previsión de dicho valor a futuro (en base a previsiones de cada uno de los componentes individuales, por un lado, y de escenarios hipotéticos por otro lado), y con esto poder también determinar la implicancia en las emisiones del parque generador del país.

A continuación, se enlistan las conclusiones más importantes de este trabajo:

1. El precio final de la electricidad en el mercado *spot* de España ha crecido considerablemente en el último lustro, pasando de un precio promedio menor a 53 €/MWh en el año 2017 a un valor promedio que ronda los 213 €/MWh en lo que va del año 2022.
2. Dado el sistema marginalista de composición del precio final de la electricidad, este está fuertemente influenciado por el precio del gas natural. Un mecanismo de formación del precio que logre el desacople de estos precios del gas natural con respecto a los de la electricidad (como la llamada “Singularidad Ibérica”, mecanismo implementado en España) puede ser efectivo a la hora de generar dicho desacople.
3. Realizar una previsión del precio final de la electricidad resulta muy complejo debido a la gran volatilidad del mismo, por su propia naturaleza y la gran sensibilidad que posee a múltiples variables externas (incluyendo situaciones coyunturales tales como conflictos bélicos o cambios regulatorios).
4. El modelo ARIMA resulta útil para la previsión de diversas variables que componen el precio final de la electricidad, no solo porque permite incorporar elementos como la estacionalidad (característica de varias de las componentes del precio final) sino porque es capaz de modelar escenarios altamente variables, como el que se viene dando desde finales de 2021.
5. Realizar una previsión de las emisiones del parque generador del país resulta una tarea compleja, debido a la gran cantidad de factores que influyen; no solamente considerando las tecnologías emisoras y sus factores de emisión, sino las distintas condicionantes que pueden hacer que una cierta tecnología comience a generar en un determinado momento y, por ende, a emitir GEI. Esto es característico por ejemplo de las centrales térmicas de ciclo combinado, las cuales suelen comenzar a generar cuando ya las restantes tecnologías han ofertado todo su volumen en el mercado.
6. Las emisiones del parque generador en España muestran una tendencia decreciente en el último lustro, debido a la creciente potencia instalada de generación de energías renovables. Este valor de emisiones ha de disminuir

en tanto las energías renovables continúen en aumento, como lo vienen haciendo de manera sostenida.

## REFERENCIAS

Arnedillo O. (2007). *Modelos de mercado eléctrico. Paradigma competitivo y alternativas de diseño*. Economía industrial, ISSN 0422-2784, Nº 364, 2007. pp. 39-54

Box G. (2015). *Time Series Analysis: Forecasting and Control*. WILEY.

Box G., Jenkins G. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Casado Galdeano U., Larrea Basterra M. (2017). *Determinantes del precio de la electricidad en España*. Estadística Española, Volumen 59, número 194 / 2017, pp. 119-149.

Comisión Europea, Dirección General de Acción por el Clima. (2019). *Going climate-neutral by 2050: a strategic long-term vision for a prosperous, modern, competitive and climate-neutral EU economy*, Oficina de publicaciones de la Unión Europea. <https://data.europa.eu/doi/10.2834/02074>

Chien T., Hu J. (2008). *Renewable energy: An efficient mechanism to improve GDP*. Energy Policy 36 (2008) 3045– 3052.

Díaz Mendoza A., Larrea Basterra, M., Álvarez Pelegry E., Mosácula Atienza, C. (2015). *De la liberalización (Ley 54/1997) a la reforma (Ley 24/2013) del sector eléctrico español*. Cuadernos Orkestra 2015/10 ISSN 2340-7638

Energy Market Prices (s.f.). Precios de los derechos de emisión EUA. Recuperado de: <https://www.energymarketprice.com/home/>

García M. (7 de febrero de 2019). *Time Series Forecasting (1): Initial analysis*. Towards Data Science. <https://towardsdatascience.com/time-series-forecasting-1-initial-analysis-87e017501e98>

Grupo de Trabajo I, contribución para el Sexto Informe de Evaluación del IPCC (2021). *Climate Change 2021: The Physical Science Basis*. Intergovernmental Panel on Climate Change.

International Energy Agency. Energy Statistics Data Browser. (s.f.). Recuperado de: <https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser?country=WORLD&fuel=Energy%20supply&indicator=ElecGenByFuel>

Korstanje J. (27 de junio de 2021). *Structural Equation Modeling*. Towards Data Science. <https://towardsdatascience.com/structural-equation-modeling-dca298798f4d>

Maklin C. (25 de mayo de 2019). *ARIMA Model Python Example — Time Series Forecasting*. Towards Data Science. <https://towardsdatascience.com/machine-learning-part-19-time-series-and-autoregressive-integrated-moving-average-model-arima-c1005347b0d7>

McKinsey Global Institute (2022). The net-zero transition. What it could cost, what it could bring. McKinsey Global Institute.

Mercado Ibérico del Gas (s.f.). Resultados anuales. Recuperado de: <https://www.mibgas.es/es/file-access>

Meshcheryakov G., Igolkina A. (2019). *Semopy: A Python package for Structural Equation Modeling*. *Structural Equation Modeling: A Multidisciplinary Journal*, 27:6, pp. 952-963.

Morata, A. (2014). *Guía de escenarios regionalizados de cambio climático sobre España a partir de los resultados del IPCC-AR4*. Agencia Estatal de Meteorología de España

Mouakher A., Inoubli W., Ounoughi C., Ko A. (2022). EXPECT: EXplainable Prediction Model for Energy ConsumPTion. *Mathematics* 2022, 10, 248. <https://doi.org/10.3390/math10020248>

Müller A., Guido S. (2017). *Introduction to Machine Learning with Python – A guide for data scientists*. O'Reilly Media

Operador del Mercado Ibérico de la Electricidad. (s.f.) Energía horaria por tecnologías. Recuperado de: <https://www.omie.es/es/market-results/daily/daily-market/hourly-power-technologies>

Operador del Mercado Ibérico de la Electricidad. (s.f.) Precios horarios del mercado diario en España. Recuperado de: <https://www.omie.es/es/file-access-list>

Red Eléctrica de España (2020). *Emisiones de CO<sub>2</sub> asociadas a la generación de electricidad en España*. Red Eléctrica de España, pp. 3-4

Red Eléctrica de España (2022). *Informe del sistema eléctrico español: sistema eléctrico 2021*. Red Eléctrica de España.

Red Eléctrica de España (s.f.) Serie histórica de potencia instalada. Sistema eléctrico nacional. Recuperado de: <https://www.ree.es/es/datos/generacion/potencia-instalada>

Rubio S. (23 de septiembre de 2021). *El mercado eléctrico y el sistema marginalista de fijación de precios*. El País. <https://elpais.com/economia/2021-09-23/el-mercado-electrico-y-el-sistema-marginalista-de-fijacion-de-precios.html>

*Statistical Learning – Lasso*. (s.f.). DataCadamia. [https://datacadamia.com/data\\_mining/lasso](https://datacadamia.com/data_mining/lasso)

Suhr D. (2006). *The Basics of Structural Equation Modeling*.

Tibshirani R. (1996). *Regression Shrinkage and Selection via the lasso*. *Journal of the Royal Statistical Society. Series B (methodological)*. Wiley. 58 (1): 267–88.

Verma Y. (24 de mayo de 2022). *Quick way to find p, d and q values for ARIMA*. Analytics India Mag. <https://analyticsindiamag.com/quick-way-to-find-p-d-and-q-values-for-arima/>

*What Is ARIMA Modeling?* (s.f.). Master's in Data Science. <https://www.mastersindatascience.org/learning/statistics-data-science/what-is-arima-modeling/>

Ministerio para la Transición Ecológica y el Reto Demográfico. (2020). *Plan Nacional Integrado de Energía y Clima (PNIEC) 2021-2030*. Ministerio para la Transición Ecológica y el Reto Demográfico.