



Universidad Miguel Hernández

Facultad de Ciencias Sociales y Jurídicas de Elche
Grado en Administración y Dirección de Empresas



Trabajo Fin de Grado

Curso académico 2017/2018

Modelos econométricos para la predicción a corto plazo
de cotizaciones del Ibex-35

ALUMNO: Ginés García Soler

TUTOR: Agustín Pérez Martín

ÍNDICE

1. RESUMEN.....	1
2. INTRODUCCIÓN	2
3. FASE 1: SELECCIÓN DE LOS TÍTULOS.....	4
3.1. FACTORES DE DECISIÓN	5
3.2. SELECCIÓN DE EMPRESAS	7
3.2.1. EXPLICACIÓN DE LAS EMPRESAS SELECCIONADAS	7
4. FASE 2: IDENTIFICACIÓN DE MODELOS Y LONGITUD DE LAS SERIES TEMPORALES A UTILIZAR	12
4.1. MODELOS ESTADÍSTICOS.....	13
4.1.1. MÉTODO MEDIANTE ESQUEMA ADITIVO	14
4.1.3. MÉTODOS HOLT WINTERS MULTIPLICATIVO Y ADITIVO	15
4.1.4. MODELOS <i>ARMA</i>	16
4.1.5. MODELOS <i>ARIMA</i>	17
4.1.6. MODELOS <i>SARIMA</i>	17
4.1.7. SELECCIÓN DEL MODELO ÚTIL	18
4.1.8. SELECCIÓN DE MODELOS PARA LAS EMPRESAS SELECCIONADAS	24
4.2. LONGITUD DE LAS SERIES TEMPORALES	55
4.2.1. IDENTIFICACIÓN DE LA LONGITUD DE LAS SERIES TEMPORALES	56
4.2.1.1. MODELOS RESULTANTES CON 6 MESES DE LONGITUD DE SERIE	57
5. FASE 3: PROCESO DE PREDICCIÓN.....	58
5.2. INTERPRETACIÓN DE LAS TABLAS DE RESULTADOS	71
5.3. RESUMEN GRÁFICO DE LOS RESULTADOS	72
6. FASE 4: COMPROBACIÓN Y ANÁLISIS DE ERRORES	74
6.1. VERTIENTE DE ANÁLISIS DE ERRORES EN DATOS EXACTOS	75
6.2. VERTIENTE DE ANÁLISIS DE TENDENCIA PRONOSTICADA	77
7. INTERPRETACIÓN DE LOS RESULTADOS OBTENIDOS	78
8. LIMITACIONES DEL ESTUDIO.....	80
9. CONCLUSIONES RESULTANTES DEL TRABAJO.....	82

ANEXO I. Pruebas prácticas..... 84

BIBLIOGRAFÍA..... 88



1. RESUMEN

En este documento se llevará a cabo a cabo las pruebas empíricas necesarias para evaluar la efectividad de modelos estadísticos a la hora de predecir las cotizaciones futuras a corto plazo de acciones cotizadas en bolsa, más concretamente en nuestro mercado organizado, el IBEX-35.

El objetivo principal de este estudio es conocer si los métodos estadísticos que se procederán a desarrollar pueden servir para la predicción a corto plazo de precios de acciones de diferentes empresas cotizadas.

En primer lugar, se procederá a la selección de empresas. Para ello se escogerán distintas empresas de diferentes sectores económicos. La selección también tendrá en cuenta la liquidez de las compañías y la variabilidad de cada título con respecto al mercado (beta), eligiendo títulos con más movilidad que el mercado y títulos con menos movilidad que el mercado.

En segundo lugar, y ya con las empresas seleccionadas, se realizarán análisis estadísticos con el software informático SPSS Statistics, con el fin de identificar dos variables clave para el estudio. La primera variable a identificar será de naturaleza cualitativa, ya que se determinará qué modelo se llevará a cabo para cada título, pudiendo resultar que el estudio se realice con el mismo modelo para todos los títulos o, por el contrario, que cada acción se plantee con un método distinto. Esta solución se reflejará en el ajuste que tenga cada modelo con su serie temporal y en la calidad de los resultados. La segunda variable a identificar será de tipo cuantitativa, debido a que se determinará la cantidad de datos que utilizaremos en cada serie económica para realizar sus respectivos modelos. En la Fase 2 se comentarán con más detalle de cada una de estas dos variables y de su importancia para la resolución del estudio.

En tercer lugar, con las empresas seleccionadas y los modelos y cantidad de datos identificados, se procederá a la realización de los cálculos estadísticos que utilizaremos

para las predicciones finales. Estos cálculos se llevarán a cabo mediante el software informático R, con R Studio y R Commander.

En cuarto lugar, con las predicciones realizadas anteriormente guardadas, se llevarán a cabo las comprobaciones de dichas predicciones. Los días siguientes se constatará si efectivamente las predicciones se ajustan a la realidad, comparándolas con los datos reales de las cotizaciones actualizadas en dicho día. En esta fase determinaremos los aciertos y errores de nuestros modelos.

Por último, ya con los datos del estudio acabados, se dará una interpretación de las predicciones y de los aciertos y errores cometidos, intentando discernir la efectividad de dichos modelos estadísticos en nuestro intento de predecir cotizaciones futuras a corto plazo.

Palabras clave: modelos econométricos, series temporales, Ibex-35, predicciones, suavizados exponenciales, *ARMA*, *ARIMA*.

2. INTRODUCCIÓN

El estudio que se lleva a cabo está estrechamente relacionado con el Análisis Técnico que utilizan los analistas bursátiles para dar predicciones sobre precios de acciones, fluctuaciones de divisas o movimientos de los mercados. Al igual que el Análisis Técnico, nos basaremos en datos pasados para la elaboración de los modelos, esto significa que existirán limitaciones a la hora de predecir movimientos puesto que estos sistemas se verán condicionados por esos aspectos contextuales de los datos seleccionados, pudiendo no resultar exactos en el futuro cercano o no ajustarse a la tendencia ya creada por estos datos pasados.

El análisis técnico moderno se asienta en las bases elaboradas por el analista económico Charles Dow (1851-1902), en su estudio titulado Teoría Dow¹. Desde entonces se han utilizado diferentes métodos basados en lectura de gráficos y modelos estadísticos para

¹Véase: [https://pi.bnpparibas.es/cursos/99/an%C3%A1lisis-t%C3%A9cnico-i-5-teor%C3%ADa-de-dow-\(y-de-por-qu%C3%A9-nunca-utiliz%C3%B3-el-an%C3%A1lisis-t%C3%A9cnico\)](https://pi.bnpparibas.es/cursos/99/an%C3%A1lisis-t%C3%A9cnico-i-5-teor%C3%ADa-de-dow-(y-de-por-qu%C3%A9-nunca-utiliz%C3%B3-el-an%C3%A1lisis-t%C3%A9cnico))

intentar predecir los precios futuros de acciones cotizadas en bolsa, que es nuestro objetivo. Ha habido multitud de casos en los que, por una parte, el modelo creado ha sido exitoso y posee capacidad de predicción, y otros, por el contrario, en los que los modelos creados no sirven para explicar los comportamientos futuros de los valores subyacentes. Nos detendremos en el caso más significativo, a mi modo de ver. Este caso es el de Martin A. Armstrong, un analista económico que desarrolló sus propios modelos para predecir los movimientos de los mercados financieros. Su trabajo tuvo un gran éxito y sus modelos fueron perseguidos por instituciones públicas y privadas. Entre los pronósticos acertados de Martin Armstrong se encuentran un repunte en los precios de los productos básicos en 1977 y la crisis financiera rusa de 1998.

Nuestros modelos, a diferencia del ejemplo citado de Martin Armstrong, se basan únicamente en datos numéricos, y las predicciones resultantes son a corto plazo, en concreto diarias. La elección de predicciones diarias es resultado del tiempo de preparación del presente estudio, ya que no sería posible hacer predicciones a medio o largo plazo debido a que no se tendría tiempo de observar el comportamiento de los modelos seleccionados y contrastar sus predicciones con los datos reales que se fueran sucediendo. Este carácter diario del análisis nos fuerza a realizar las fases del estudio en dos partes diferenciadas. Por una parte, se desarrollarán todos los cálculos y análisis necesarios para obtener las predicciones deseadas, y por otro, se realizarán las comprobaciones oportunas con los datos reales que vayan apareciendo diariamente para ver si los modelos se han ajustado a la realidad, posteriormente daremos una interpretación de cada caso. Esta segunda parte se llevará a cabo necesariamente al finalizar la primera parte, puesto que está estrictamente basada en ella, y si no se tienen las predicciones antes no es posible realizar nada más.

Para la realización de los modelos, se obtendrán los datos necesarios de las cotizaciones de las acciones seleccionadas mediante la página web “*Yahoo Finance*”. Existen multitud de sitios web donde es posible encontrar estos datos, se utilizará ésta por ser de gran confianza y actualizarse diariamente.

La composición del documento consta de nueve apartados. En los apartados 3 y 4 se desarrollarán las fases 1 y 2 del trabajo. En estas primeras fases se llevarán a cabo los procesos pre-estudio, los cuales son la selección de las acciones a pronosticar, y la

evaluación de los modelos estadísticos y las longitudes de las series temporales de datos posibles. Mediante estas fases se obtendrán qué acciones se pronosticarán, qué modelos se utilizarán, y qué longitud temporal se analizará en cada caso. En el apartado 5 se desarrollará la fase 3, en la cual se elaborarán todos los cálculos necesarios para obtener las predicciones deseadas. Esta fase, junto con la elección del modelo estadístico, son el cuerpo del estudio y las más importantes sin duda puesto que, si no se hacen correctamente, todo el trabajo estará mal ejecutado. Posteriormente, en el apartado 6, se comprobarán los resultados obtenidos, así como los errores derivados de los cálculos anteriores, esto será la fase 4. Esta parte será la que indique el resultado final de nuestro estudio, y nos dirá si el objetivo principal se cumple o no.

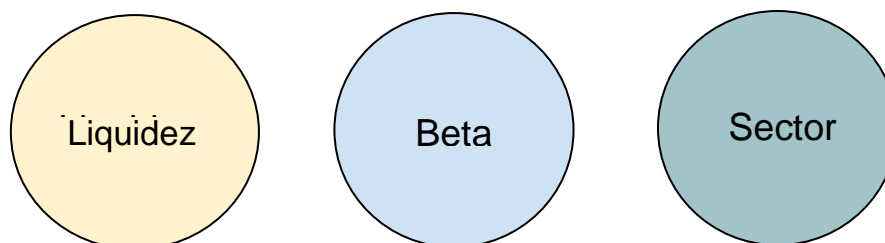
Por último, en el apartado 7, interpretaremos los resultados obtenidos en el estudio completo. Será necesario contextualizar las predicciones e interpretarlas según el entorno de cada empresa. Además, se explicarán las posibles desviaciones de los resultados con respecto a la realidad desde una perspectiva empresarial y no estadística, para así poder comprender el porqué de estas posibles desviaciones dejando a un lado los cálculos matemáticos y visualizando el entorno de cada compañía en su estado actual.

3. FASE 1: SELECCIÓN DE LOS TÍTULOS

Los modelos estadísticos que se van a utilizar en este estudio podrían aplicarse a todo tipo de series temporales de valores y a multitud de campos de actuación diferentes sin la necesidad de pertenecer al sector financiero. En este caso sí que será estrictamente financiero ya que nos dedicaremos a la predicción de las cotizaciones de determinadas acciones en el mercado organizado de la bolsa de valores española.

Para escoger las empresas a pronosticar, nos hemos basado en tres factores: el primero es la liquidez que poseen, el segundo las betas de variabilidad con respecto al mercado, y el tercero el sector económico al que pertenecen.

Figura 1. Factores de decisión



Fuente: elaboración propia

Como todas las empresas están englobadas en el índice IBEX-35, estos factores no son determinantes, ya que este índice está creado con empresas que poseen liquidez, empresas con distintas betas, y empresas de casi todos los sectores económicos. Debido a esto, lo que se quiere conseguir con la selección no es escoger a las mejores empresas, sino contemplar acciones de distintas naturalezas y características para que nuestro estudio contemple diversas posibilidades de actuación.

3.1. FACTORES DE DECISIÓN

Para la selección de las empresas se tendrán en cuenta tres factores. El primero está relacionado con la liquidez de las empresas, el segundo con la variabilidad con respecto al mercado de cada una de ellas, y el tercero con el sector al que pertenezcan las empresas escogidas.

El primer factor no resulta determinante en cuanto a las predicciones y al estudio en sí, pero sí que sería importante si este trabajo se llevara a la práctica por profesionales del sector, ya que los modelos, al ser de carácter diario, necesitan de empresas con alto grado de liquidez para hacer frente a las variaciones diarias del mercado que son recogidas en los modelos estadísticos. Si una empresa no tuviera la liquidez suficiente sería inútil usar las predicciones puesto que no podríamos recibir los beneficios económicos de éstas. Debido a este factor, las empresas que se seleccionarán pertenecerán al IBEX-35, ya que se sabe que si una empresa está recogida en este índice

quiere decir que ha cumplido una serie de requisitos de liquidez, así como otros requisitos en los que no nos detendremos por no aportar fundamento a nuestro estudio.

El segundo factor a tener en cuenta es la variabilidad de los títulos con respecto al mercado, lo que se conoce como *Beta* de una acción. Este factor indicará si los precios de las acciones se mueven más que el mercado o menos que el mercado. Un valor menor que 1 significará que se mueve menos que el mercado, por el contrario, si su valor es mayor que 1 será porque se mueve más que el mercado (siendo el mercado el IBEX-35). Puesto que el estudio tiene como objetivo demostrar la efectividad de los modelos estadísticos para la predicción de los precios de los títulos seleccionados, escogeremos acciones de los dos tipos, unas con una beta inferior a 1 y otras con una beta superior a 1. De este modo se podrá observar si la beta condiciona a las predicciones o si por el contrario no influye en los resultados.

Los datos de las betas de cada empresa, los obtendremos de las páginas web www.investing.com y www.finance.yahoo.com. Este valor varía según la página web donde lo encuentres, por lo tanto, se tendrá que identificar en ambas webs para contrastar de qué valor se trata.

Por último, el tercer factor corresponderá con el sector económico al que pertenezcan las empresas seleccionadas. Para poder estudiar mejor la eficacia de los sistemas de predicción estadísticos, se pondrán a prueba los modelos con tres sectores económicos diferentes. Los sectores escogidos son: el sector financiero, el sector de energía y combustible y el sector de la industria y la construcción. Se elegirán dos empresas representantes de cada uno de los sectores mencionados.

3.2. SELECCIÓN DE EMPRESAS

Teniendo en cuentas los factores explicados en el punto anterior, las empresas escogidas para el estudio son:

Figura 2. Empresas seleccionadas



Fuente: elaboración propia

Para que los resultados sean presentados de una manera más visual, se utilizará un código de colores para cada una de las empresas. Se empleará el azul para BBVA, el rojo para Santander, el amarillo para ArcelorMittal, el verde para Iberdrola, el lila para Siemens Gamesa, y el naranja para Repsol.

3.2.1. EXPLICACIÓN DE LAS EMPRESAS SELECCIONADAS

A continuación, se detallarán las empresas seleccionadas teniendo en cuentas los factores de decisión marcados para el estudio y el contexto de la propia entidad.

BBVA

Banco Bilbao Vizcaya Argentaria es una entidad bancaria global que fue fundada en 1857 y que ofrece servicios financieros en más de 30 países, con más de 70 millones de clientes. Posee un posicionamiento de liderazgo en el mercado español, es la mayor entidad bancaria en México, y además cuenta con franquicias líderes en América del

Sur y en “Sunbelt”, en Estados Unidos. Cuenta a su vez con una presencia relevante en Turquía.

Las actividades económicas del grupo están enfocadas a mercados de gran crecimiento, y dentro de su filosofía está inculcada la tecnología como una ventaja competitiva clave. El equilibrio entre zonas de actuación emergentes y desarrolladas permite una diversificación adecuada de las fuentes de ingresos y proporciona un alto grado de reiteración en sus resultados.

Dentro del sector financiero, escogemos el banco BBVA por su alta relevancia en nuestro mercado y por su fuerte posicionamiento en la cuota de mercado relativa del negocio bancario en España. Además, cumple con el requisito de liquidez.

En cuanto al factor de la variabilidad con respecto al mercado, en el sector financiero de nuestro estudio, tanto BBVA como Santander poseen una beta mayor a 1, lo que significa que se mueven más que el mercado. Esto es así debido a que todas las empresas de este sector que pertenecen al IBEX-35 poseen una beta superior a 1, de tal modo que no ha sido posible escoger una empresa con beta menor que 1 para las demostraciones de nuestros modelos.



El Banco Santander es la segunda compañía bancaria española que se considerará para el estudio. Tiene la sede en la ciudad de Santander, en Cantabria, donde desempeñaron sus primeras actividades en 1857. En cuanto a su negocio, la banca minorista le aporta cerca del 70% de su beneficio (2014). A finales de 2017, poseía activos por valor de 1.444.305 millones de euros, siendo así la mayor entidad española por volumen de activos, gracias en gran parte a la beneficiosa compra del Banco Popular, a mediados de 2017. Cuenta también con cerca de 14.000 oficinas, y más de 130 millones de clientes repartidos de manera global.

Acompañando a BBVA en el sector financiero, se ha escogido al Banco Santander por su posición líder en el marco bancario español y por su fuerte posicionamiento tanto nacional como internacional, cotizando en el DOW JONES y en el EURO STOXX 50, además de en el IBEX-35, lo que demuestra un alto grado de fortaleza y liquidez.

Como hemos comentado anteriormente, la beta de Santander también es mayor a 1, al igual que la de BBVA, lo que significa que posee una variabilidad superior al mercado. En este caso no se podrá observar si influye la variable beta al no tener otra empresa de este sector con una beta inferior a 1.



ArcelorMittal es una empresa presente en 60 países y con una plantilla de cerca de los 200.000 empleados. Es el mayor productor siderúrgico a nivel mundial, y cuenta con instalaciones industriales en 18 países y es capaz de producir 113 millones de toneladas de acero líquido al año.

Dentro del sector de la industria y la construcción, se ha escogido ArcelorMittal, una compañía líder en el sector nacional e internacional y con presencia en multitud de países. Además de cotizar en el IBEX-35, también lo hace en la bolsa de Nueva York, Ámsterdam, París y Luxemburgo, lo que asegura el grado de liquidez necesaria para el estudio.

En cuanto a la beta de sus acciones, posee un valor superior a 1, lo que significa que se mueve diariamente más que nuestro mercado de referencia.



“Siemens Gamesa Renewable Energy nació en abril de 2017 con la fusión de Gamesa Corporación Tecnológica y Siemens *Wind Power*. La historia de Gamesa está marcada por el espíritu de innovación y una exitosa expansión hacia nuevos mercados. Lo que empezó como un pequeño taller mecánico en el norte de España acabó convirtiéndose

rápidamente en una compañía global centrada en la gestión de instalaciones industriales, en el sector de la automoción y en el desarrollo de nuevas tecnologías.” (<https://www.siemensgamesa.com/es-es/sobre-nosotros/historia-de-la-empresa>)

La segunda compañía escogida para el sector de industria y construcción es Siemens Gamesa. Las características principales de esta organización son la innovación y la flexibilidad en su mercado. En este caso la compañía no goza del mismo grado de liderazgo internacional en su mercado que ArcelorMittal, no obstante, es una empresa cotizada en el IBEX-35 y por ello cumple nuestro requisito de liquidez.

Con respecto a la beta de Siemens Gamesa, su valor es inferior a 1, a diferencia de ArcelorMittal, lo que significa que su variabilidad es menor que la del mercado de referencia. Esto nos puede dar una idea de si una beta inferior a 1 nos facilita el seguimiento del precio de sus acciones o si, por el contrario, nos entorpece la capacidad de movimiento de nuestros modelos.



Iberdrola es un grupo empresarial español que se dedica a la producción, distribución y comercialización de energía. Tiene su sede en la ciudad de Bilbao, en el País Vasco, y es una de las grandes empresas eléctricas a nivel internacional. En España, es la segunda compañía en cuanto a producción eléctrica, pero la primera en cuanto a capitalización bursátil y cuarta del mundo. Además, apuesta por la energía eólica y está a la cabeza en este sector de manera global.

Al igual que las empresas anteriormente detalladas, se ha escogido a Iberdrola en el sector de energía y combustible por su posicionamiento líder en su mercado y su relevancia a nivel nacional e internacional. La compañía cotiza en el IBEX-35, lo que nos asegura un nivel de liquidez óptimo para el estudio.

En cuanto a la variabilidad con respecto al mercado o beta, posee un valor inferior a 1, al igual que Siemens Gamesa, lo que representa que la movilidad del precio de sus

acciones es inferior a la movilidad general del IBEX-35. Con estas dos compañías con betas inferiores a 1, se intentará dar interpretación a su menor movimiento, si se viera reflejado en los resultados de los modelos estadísticos.



Por último, Repsol será la segunda compañía del sector energético del estudio. Es una multinacional del sector de la energía y la petroquímica, con sede social en Madrid. Fue fundada en 1987 y en sus principios estuvo formada por varias compañías, las cuales pertenecían previamente al Instituto Nacional de Hidrocarburos (INH).

Sus actividades económicas constaban de producción, exploración, refinado de petróleo y gas, y transporte. A su vez, realiza actividades de producción, distribución y comercialización de productos derivados del petróleo, petroquímicos y gas natural.

Acompañando a Iberdrola dentro de nuestro sector de energía y combustible, se ha seleccionado a Repsol. Con esta compañía se ha querido hacer un hueco al mercado de los hidrocarburos. Como las empresas anteriormente mencionadas, Repsol goza de un fuerte posicionamiento y de un alto reconocimiento de marca tanto en España como en Latinoamérica.

A diferencia de su compañero Iberdrola en nuestro sector de energía y combustible, Repsol posee una beta superior a 1, lo que significa que tiene una variabilidad del precio de sus acciones superior al mercado de referencia. Esto ayudará a comparar los dos modelos y discernir si existe relación entre la beta y el ajuste de los modelos estadísticos creados.

4. FASE 2: IDENTIFICACIÓN DE MODELOS Y LONGITUD DE LAS SERIES TEMPORALES A UTILIZAR

Una vez realizado el proceso de selección de empresas, se identificará qué modelo o modelos se van a utilizar para realizar las predicciones. Cada título escogido puede tener unas características diferentes en la serie temporal de su cotización, por ello, se deberán realizar pruebas de cada serie para obtener qué modelo se ajusta mejor a los datos reales.

Para establecer qué modelo se va a seguir en cada acción, se realizará una comparación de modelos en el software SPSS Statistics, de manera que se observe cada modelo y se puedan descartar los que menos se ajusten a los datos reales. Este ajuste a los datos reales, o efectividad del modelo, vendrá dado por datos estadísticos como el p-valor o el estadístico de los contrastes. Este paso se explicará ampliamente en el siguiente subapartado.

Por otro lado, se deberá establecer la cantidad de datos que se utilizarán para realizar el análisis de las predicciones. Esta cantidad vendrá determinada por dos aspectos:

- En primer lugar, el estudio es a corto plazo, por lo que no tendría sentido coger una serie temporal de 10 años, por ejemplo. Al ser de carácter diario nuestras predicciones, se deberán escoger la cantidad de datos que más se ajuste a la última etapa de variaciones de precios que han experimentado nuestras acciones. Estas etapas de variaciones se identificarán a partir de un análisis de la tendencia y estacionalidad de cada serie temporal, resultando el tiempo que se deberá considerar en cada modelo para conseguir el mayor ajuste posible a la realidad.
- En segundo lugar, se deberá prestar atención a los hechos que hayan tenido lugar en los últimos meses y que hayan podido afectar al funcionamiento de cada empresa, o que hayan podido alterar la variabilidad de los precios de sus acciones. Dentro del entorno de cada compañía seleccionada e incluso dentro de cada compañía, han podido producirse cambios sustanciales que afecte a las variaciones de las series temporales de los precios de sus acciones. Esto es algo

que debemos tener en cuenta y saber visualizarlo, ya que estos cambios pueden alterar las series temporales rompiendo sus tendencias o estacionalidad, lo que provocaría que se tuviera que modificar la cantidad de datos a utilizar, reduciéndolos a la etapa desde que se han producido las variaciones hasta la fecha del inicio de los cálculos.

Para esta selección de los modelos y cantidad de datos, se usará el software informático SPSS Statistics. Este programa ayudará a identificar fácilmente qué modelo utilizar en cada caso y qué longitud de serie temporal deberemos emplear. Igual que en el caso de las fuentes de datos, existen multitud de programas informáticos útiles para realizar los cálculos necesarios para las predicciones, en nuestro estudio utilizaremos SPSS Statistics por su intuitiva interfaz y su fácil manejo. Como ya se ha señalado, posteriormente se utilizará el programa R, con RStudio y RCommander, para calcular las predicciones. El motivo de esta separación de tareas por programas es la manejabilidad de cada uno. Mientras que en SPSS Statistics deberíamos hacer cada paso a manualmente, así como los cambios entre empresas, con el programa R Studio podremos realizar con más agilidad los cálculos de las diferentes empresas de manera casi automática, lo que nos ahorrará tiempo. Además, se podrá comprobar los resultados de la selección de modelos y datos en dos programas distintos, lo que dará más seguridad en los resultados finales del trabajo.

4.1. MODELOS ESTADÍSTICOS

En este punto de la Fase 2, nos centraremos en la elección del modelo estadístico que más se ajuste a los datos reales para así favorecer a que las predicciones realizadas sean lo más fieles posibles a los datos futuros. Para llevar a cabo estas predicciones existen multitud de modelos que pueden ser susceptibles de utilizar. se escogerán unos tipos de modelos estadísticos útiles para realizar predicciones, y se probará cada uno de los modelos con las series temporales de las acciones seleccionadas. El programa SPSS Statistics dará qué modelo se ajusta más a los datos y procederemos posteriormente a usar ese modelo en el programa R para realizar las predicciones.

Debido al carácter diario de las predicciones, se deberán utilizar modelos estadísticos que se ajusten a predicciones a corto plazo. Existen diversos métodos que podríamos utilizar para calcular las predicciones, como el aditivo, multiplicativo, Holt Winters Aditivo, Holt Winters Multiplicativo, *ARMA*, *ARIMA* o *SARIMA*. Estos son algunos de los modelos que se han contemplado para la realización de este estudio. A continuación, se pasará a detallarlos para decidir con qué modelo nos quedamos para las predicciones en general. Posteriormente, y como se verá a continuación, habrán diversas combinaciones dentro del modelo que se seleccione, y en ese momento es cuando se deberán hacer comprobaciones para conocer qué variación del método es más adecuada.

En primer lugar, se describirá brevemente el funcionamiento de los posibles modelos con los que se podrían realizar las predicciones, exponiendo las fórmulas y las características que los componen.

En segundo lugar, pasaremos a decantarnos por alguno de los modelos detallados a través de un proceso de decisión que se basará en la utilidad de cada modelo y en el ajuste que tendría en nuestro estudio.

4.1.1. MÉTODO MEDIANTE ESQUEMA ADITIVO

Supone que las observaciones se generan como suma de las cuatro componentes principales de una serie temporal, es decir:

$$Y_t = t_t + c_t + e_t + r_t$$

Siendo las cuatro componentes: tendencia secular (t), variación estacional (e), variación cíclica (c), y variación irregular, residual, aleatoria o accidental (r).

En este caso cada componente se expresa en el mismo tipo de unidad que las observaciones. La variación residual, en este modelo, es independiente de las demás componentes, es decir la magnitud de dichos residuos no depende del valor que tome cualquier otra componente de la serie, (análogamente la variación estacional y la cíclica son independientes de las demás componentes).

4.1.2. MÉTODO MEDIANTE ESQUEMA MULTIPLICATIVO

Supone que las observaciones se generan como producto de las cuatro componentes, es decir:

$$Y_t = t_t \cdot c_t \cdot e_t \cdot r_t$$

En este modelo (multiplicativo puro) la tendencia secular se expresa en el mismo tipo de unidad que las observaciones, y el resto de las componentes en tanto por uno. Aquí no se cumple la hipótesis de independencia del esquema aditivo.

Otro tipo de modelo multiplicativo que si la cumple llamado modelo multiplicativo mixto es el siguiente:

$$Y_t = t_t \cdot c_t \cdot e_t + r_t$$

Existen otros modelos que combinan esquemas aditivos y multiplicativos, tratando de resolver las carencias o inconvenientes de los modelos más sencillos. Señalaremos que generalmente el modelo multiplicativo es el que mejor se adapta a la descripción de variables económicas.

4.1.3. MÉTODOS HOLT WINTERS MULTIPLICATIVO Y ADITIVO

Este método está diseñado para series que sigan una tendencia lineal y además tengan factor estacional.

Esquema Multiplicativo

Se desarrollará para el esquema mixto siguiente:

$$Y_t = T_t \cdot E_t + e_t \quad \text{siendo } T_t = b_0 + b_1 \cdot t$$

Con lo cual: $Y_t = (b_0 + b_1 t) \cdot E_t + e_t$

En el modelo Holt-Winters se producen tres ecuaciones de alisado para estimar

b_0 , b_1 y E_t . El coeficiente de alisado puede ser distinto.

El nivel o componente permanente (b_0) se estima mediante la variable de alisado S_t , la pendiente de la recta (b_1) con b_{1t} y E_t con C_t .

Esquema Aditivo

Se desarrollará para el esquema siguiente:

$$Y_t = (b_0 + b_1 t) + E_t + e_t$$

Se utilizan tres ecuaciones de alisado como antes para estimar b_0 , b_1 y E_t .

4.1.4. MODELOS ARMA

“En estadística, los modelos autorregresivos de media móvil (en inglés *AutoRegressive Moving Average models*, abreviados *ARMA*), también llamados *Modelos Box-Jenkins*, se aplican a series temporales de datos. Dada una serie temporal de datos X_t , el modelo *ARMA* es una herramienta para entender y, aún más, para predecir futuros valores de la serie. El modelo está formado por dos partes, una parte autorregresiva (AR) y otra de media móvil (MA). El modelo se conoce con el nombre de modelo *ARMA* (p,q), donde p es el orden de la parte autorregresiva y q es el orden de la parte de media móvil.” (https://es.wikipedia.org/wiki/Modelo_autorregresivo_de_media_m%C3%B3vil)

- Un proceso AR (1) está dado por:

$$X_t = c + \phi X_{t-1} + \epsilon_t$$

Donde ϵ_t es un proceso de ruido blanco con media cero y varianza σ^2 . El proceso es de covarianza estacionaria si $|\phi| < 1$. Si $|\phi| = 1$, entonces X_t tiene una raíz unitaria.

- La notación MA(q) se refiere a un modelo de media móvil de orden q :

$$X_t = \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}$$

Donde $\theta_1, \dots, \theta_q$ son los parámetros del modelo y $\epsilon_t, \epsilon_{t-1}, \dots$ son, de nuevo, los términos de error.

- La notación *ARMA* (p, q) se refiere a un modelo con p términos autorregresivos y q términos de media móvil. Este modelo combina los modelos AR e MA:

$$X_t = \epsilon_t + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i}$$

4.1.5. MODELOS *ARIMA*

“En estadística y econometría, en particular en series temporales, un modelo autorregresivo integrado de promedio móvil o *ARIMA* (acrónimo del inglés *autoregressive integrated moving average*) es un modelo estadístico que utiliza variaciones y regresiones de datos estadísticos con el fin de encontrar patrones para una predicción hacia el futuro. Se trata de un modelo dinámico de series temporales, es decir, las estimaciones futuras vienen explicadas por los datos del pasado y no por variables independientes.”

(https://es.wikipedia.org/wiki/Modelo_autorregresivo_integrado_de_media_m%C3%B3vil)

El método *ARIMA* se utiliza cuando la serie estudiada no es estacionaria. Cuando esto se produce, se debe diferenciar la serie hasta convertirla en estacionaria, por lo tanto, el parámetro añadido I hace referencia al número de veces que se ha diferenciado la serie para convertirla en estacionaria.

Esto se refleja en modelos *ARIMA* (p, d, q), donde p y q son los valores determinados por el método *ARMA* (p, q), y el valor de d son las veces que se ha diferenciado la serie.

4.1.6. MODELOS *SARIMA*

“En estadística, a menudo, las series temporales poseen un componente estacional que se repite en todas las observaciones. Para hacer frente a la estacionalidad, los procesos *ARIMA* han sido generalizados, estableciendo los modelos *SARIMA* (*Seasonal Autoregressive Integrated Moving Average Model*).” (Gekko, 2017)

“Escribimos $X_t \sim ARIMA(p, d, q) \times (P, D, Q)$. La idea es que los modelos *SARIMA* son modelos *ARIMA* (p, d, q) cuyos residuos a_t son *ARIMA* (P, D, Q). Con *ARIMA* (P, D, Q)

proponemos modelos *ARIMA* cuyos operadores se definen en B_s y en potencias sucesivas.” (Gekko, 2017)

Como bien publica esta página web económica, los modelos *SARIMA* son modelos desarrollados a partir de *ARIMA* y, a su vez, de *ARMA*. Se utilizan cuando la serie es estacional, es decir, que posee patrones que se repiten en determinadas fechas.

4.1.7. SELECCIÓN DEL MODELO ÚTIL

Una vez explicados cada uno de los modelos que podríamos utilizar para nuestro estudio, pasaremos a seleccionar el modelo adecuado para realizar las predicciones. Cada modelo que hemos expuesto tiene su utilidad en función del número de datos o del horizonte temporal de las predicciones.

Según las características descritas en cada modelo estadístico, se podría pensar que el método que más se ajustará a los requisitos será el *ARIMA* (p,d,q) . Las razones son sencillas de sintetizar, pues para nuestras complejas series de datos los demás modelos mostrados no serían capaces de realizar las predicciones que necesitamos o, al menos, no como nosotros queremos, ya que se ajustarían menos a la realidad por su poca capacidad de adaptación a diferentes series. Por otra parte, el método *SARIMA* (p,d,q) (P,D,Q) nos podría ser útil si nuestras series tuvieran algún tipo de factor estacional, esto lo deberemos comprobar antes de realizar los cálculos.

Método *ARIMA* y sus procedimientos

Con la publicación de G. P. E. Box y G. M. Jenkins “*Times Series Analysis: Forecasting and Control*” en la década de los 70’s, se generaron nuevas herramientas para el cálculo de predicciones de series temporales, cuya técnica se llamó metodología Box-Jenkins, también conocida como metodología *ARIMA*.

El objetivo establecido de la metodología Box-Jenkins fue identificar y estimar modelos estadísticos que podían ser interpretados como generador de información de la muestra. En este sentido, si los modelos son destinados a la predicción, debe suponerse que las características de las series son constantes en el tiempo. Por lo tanto, las predicciones se

realizan sobre la base de que los modelos son estacionarios o estables en media y varianza. A continuación, se pasará a exponer el procedimiento a seguir para la elaboración de un modelo *ARIMA* con fines predictivos.

A continuación, expondremos el proceso por el cual llevaremos a cabo la identificación de los modelos y de la longitud de serie. Cabe resaltar que este procedimiento será el seguido tanto si se tratan de pruebas para modelos de la familia *ARIMA* como si se trata de pruebas para cualquier otro modelo contemplado, por ejemplo, los suavizados exponenciales, que son las alternativas que utilizaremos en nuestras pruebas.

Paso 0: Completar las series temporales de las cotizaciones

Este paso lo denominamos “Paso 0” debido a que no es una fase del proceso como los demás, sino que se trata de la preparación de la base de datos previa a los cálculos del verdadero proceso.

Al descargar los datos de la web *Yahoo Finance*, nos damos cuenta de que en las fechas en las que es fiesta no aparecen cotizaciones, por lo tanto lo primero que debemos hacer es completar la serie y rellenar las cotizaciones de los días de fiesta. En el caso del IBEX-35, que es el mercado en el cual hemos seleccionado las empresas, las fiestas son las correspondientes a las de España. Por lo tanto, los días que deberemos rellenar serán: el día 1 de mayo, Día del Trabajo, y los días señalados por la Navidad, el 25 y 26 de diciembre, así como el primer día del año festivo, el 1 de enero.

Considerando las opciones que tenemos para completar las series, no consideramos que la media entre los dos días en los que se encuentra el día de fiesta sea el método más adecuado, ya que no sabemos si se ajustaría a la realidad o no, por lo tanto, completaremos la serie utilizando la cotización del día anterior al día de fiesta. De esta manera imaginaremos las cotizaciones como si hubieran repetido las del día anterior, y así no caemos en hipótesis.

Paso 1: introducir los datos y convertir a serie temporal

Ya con las series completas, se introducen los datos en el software SPSS Statistics con formato CSV, un formato en el que los datos son presentados de manera seguida y delimitados por comas. Es un formato que se usa con frecuencia para ser abiertos en programas estadísticos.

Una vez se tengan los datos ejecutados en el programa, pasaremos a convertir la serie de cotizaciones en una serie temporal, para así poder utilizar las herramientas que nos proporciona el programa para trabajar las series temporales y seleccionar los modelos para las predicciones.

Paso 2: Gráfico de secuencia

En los datos con los que se está trabajando, se presentan varias variables: fecha, apertura, máximo, mínimo, cierre, cierre ajustado y volumen. Para desarrollar los modelos nosotros utilizaremos el cierre ajustado, ya que lo que buscamos predecir es el valor con el que el título cierra día a día.

En primer lugar, se le pedirá al programa que muestre el gráfico de secuencia de la variable “Cierre ajustado”. Este gráfico ayuda a identificar si existe estacionalidad en la serie, dato que se necesita saber para llevar a cabo un modelo estadístico u otro.

Observando el gráfico de secuencia nos daremos cuenta de si existe o no una estacionalidad determinada, lo que significaría que los datos se comportan de manera cíclica o que siguen un patrón identificable.

Paso 3: Estudiar la variabilidad

Para estudiar la variabilidad se utilizará la función “Explorar” dentro de los estadísticos descriptivos. Le pedimos que muestre la estimación de la potencia a la que se debería transformar la serie, en caso de que fuera necesario, para que fuera de varianza constante, que es lo que perseguimos.

Aparece el gráfico que muestra la variabilidad en la serie, y bajo aparece la prueba de Levene. La prueba de Levene, se define como: “prueba estadística inferencial utilizada para evaluar la igualdad de las varianzas para una variable calculada para dos o más grupos. Algunos procedimientos estadísticos comunes asumen que las varianzas de las poblaciones de las que se extraen diferentes muestras son iguales. La prueba de Levene evalúa este supuesto.” (Wikipedia.org, 2018), proporcionará una estimación de potencia a la cual se debería elevar la serie para convertirla en constante en varianza, si fuera necesario.

El programa que usamos para conocer la potencia de transformación nos proporcionará un valor exacto al que debemos elevar la serie, no obstante, se utilizan una serie de redondeos para hacer más sencilla la transformación y la “des-transformación” posterior. Estos valores redondeados se recogen en la siguiente tabla:

Tabla 1. Transformaciones

λ	$T\lambda$
-1	$1/X_t$
-0,5	$1/\sqrt{X_t}$
0	$\ln X_t$
0,5	$\sqrt{X_t}$
1	no transformar

Fuente: elaboración propia

λ : hace referencia a la estimación de potencia que nos ofrece el programa estadístico, de manera que se debería ajustar al valor que más se acercara.

$T\lambda$: hace referencia a la transformación que se debería realizar en la serie original.

Paso 4: Transformar en serie de media constante

Otro aspecto que se deberá conseguir, además de la variabilidad constante, es que la serie posea una media constante. Para saber si la serie cumple este requisito nos fijaremos en las autocorrelaciones simples y parciales.

Dentro de la pestaña de “Predicciones” encontramos la función “Autocorrelaciones” que tiene el programa para realizar este paso. Le pedimos las autocorrelaciones de la variable “Cierre Ajustado”, mediante aproximación de Barlett y que nos muestre 16 retardos.

La autocorrelación simple debe converger a 0 rápidamente para que nuestra serie cumpla el requisito de la media, si esto no sucede, se deberá diferenciar la serie hasta que esto se produzca.

El proceso de diferenciación de la serie consiste en crear una variable en la cual el valor de la variable X_i sea igual a la resta de los valores en la variable “Cierre Ajustado” ($Y_i - Y_{i-1}$).

Una vez dibujados los gráficos de las autocorrelaciones de la serie diferenciada, se utilizarán los retardos de los gráficos (barras azules verticales) para identificar el modelo *ARIMA* que más se ajuste.

Paso 5: Identificar los parámetros de los modelos, y crearlos

Con las autocorrelaciones que ha mostrado el programa, se pueden interpretar los retardos para identificar los parámetros del modelo *ARIMA* (p,d,q). Si el primer retardo de las autocorrelaciones se saliera de alguno de los dos márgenes de confianza (inferior o superior), diríamos que los parámetros (p) y (q) tienen un valor igual al número de barras que se salen, en la autocorrelación simple por parte de (q), y en la autocorrelación parcial por parte de (p). El parámetro (d) corresponde al número de veces que hemos diferenciado la serie. Para asegurarnos de que vamos a utilizar el modelo adecuado, procedemos a diferenciar de nuevo para conocer si las autocorrelaciones nos dan parámetros que se ajusten más al modelo.

Una vez estén estimados los parámetros, se deberán contemplarlos también de manera sobreajustada, es decir, incrementando en 1 unidad alguno de los parámetros (p) o (q), pero sin alterar el parámetro que corresponde a la diferencia. Aumentaremos el parámetro (p) o el (q), pero nunca los dos a la vez. Una vez tengamos todas las posibilidades, pasaremos a comprobar la viabilidad de cada uno de ellos.

Una vez se hayan identificado las alternativas de parámetros *ARIMA* que se pueden utilizar, se pasa a comprobar cada una de las posibles opciones para conocer qué parámetros se ajustan mejor. Para ello se utilizará la función “Crear modelos tradicionales” dentro de la pestaña de “Predicciones” en el programa informático. Se introducen los parámetros que hemos considerado y le pedimos que nos estime los parámetros para saber si los que hemos puesto tienen significación, o por el contrario, no la tienen y debemos quitarlos. Este procedimiento lo hacemos con todas las alternativas posibles que hemos sacado anteriormente, para comprobar qué parámetros son mejores para la predicción.

Como hemos comentado, la función “Estimación de los parámetros” dirá si nuestros parámetros tienen significación o no, esto lo interpretaremos observando el valor que aparezca en la columna de “Significación” en cada uno de los parámetros, menos en el de la diferencia. Si dicho valor es inferior a 0,05, podremos afirmar que nuestro parámetro es importante y por lo tanto podremos continuar con nuestro modelo.

Existirán casos en los que los parámetros no aparezcan como significativos, o que la tendencia estimada de los valores pronosticados sea nula (plana), en estos casos se deberá probar los modelos que nos ofrece el programa SPSS Statistics, que son los suavizados exponenciales de tendencia, concretamente el Simple, el de Holt, el de Brown y el Amortiguado. Estos modelos no serán útiles para predecir el valor exactos de nuestras cotizaciones, pero sí que nos servirán para predecir la tendencia que seguirá la serie en el futuro, que es un aspecto muy relevante.

Paso 6: Estudiar los residuos

Por último, una vez seleccionados los modelos y parámetros que se van a utilizar, se deben analizar sus residuos para comprobar si son ruido blanco o no. Los residuos son el resultado de la diferencia entre el valor pronosticado y el real. La característica de ruido blanco hace referencia a cuando los residuos no poseen información del modelo y por lo tanto son descartables, si los residuos nos dieran que no son ruido blanco, explicarían algo del modelo y por lo tanto no podríamos considerarlos residuos y el modelo no valdría.

Para comprobar este requisito se llevarán a cabo tres pruebas, en el caso de modelos de la familia *ARIMA* serán cuatro:

- Prueba de media 0: identificar si la media de los residuos es 0, si no lo es, no sería un modelo adecuado.
- Prueba de variabilidad: se comprueba que la variabilidad de la serie de los residuos puede considerarse constante. Utilizaremos para ello dos márgenes, uno inferior y otro superior, en el gráfico de secuencia para observar si la serie se sale o no, lo correcto es que no se salga. Los valores de los márgenes inferior y superior los calcularemos multiplicando por (-3) y por (3) respectivamente sus desviaciones típicas.
- Prueba de incorrelación: se debe confirmar que los datos de las series de los residuos están incorrelados, si no fuera así significaría que poseen información relevante para el modelo y no podríamos aprobar su uso.
- Por último, y solo en el caso de los modelos de la familia *ARIMA*, se deberá comprobar que la serie de los residuos sigue una distribución Normal, para que pueda considerarse ruido blanco *Gaussiano*.

4.1.8. SELECCIÓN DE MODELOS PARA LAS EMPRESAS SELECCIONADAS

Una vez descritos los pasos que vamos a seguir, comenzamos con las pruebas para nuestras seis empresas. Para cada una de ellas utilizaremos los mismos pasos, lo que no quiere decir que nos aparezcan los mismos resultados, de hecho, no sería lo normal.

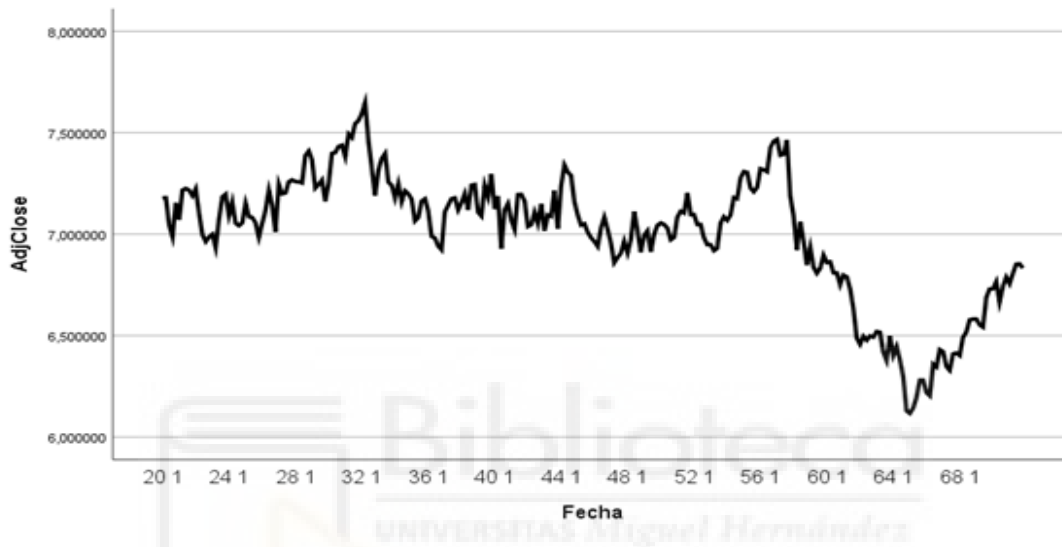
Los pasos 0 y 1 los omitiremos, ya que son dos pasos preparatorios y en todos los casos se realizaría de la misma manera. Las pruebas que se van a realizar a continuación están estudiadas sobre bases de datos de un año y de 6 meses, todas ellas con fin el día 14 de mayo de 2018.

BBVA

Paso 2

Dibujamos el gráfico de secuencia para comprobar si existen factores estacionales en la serie.

Figura 3. Gráfico de secuencia



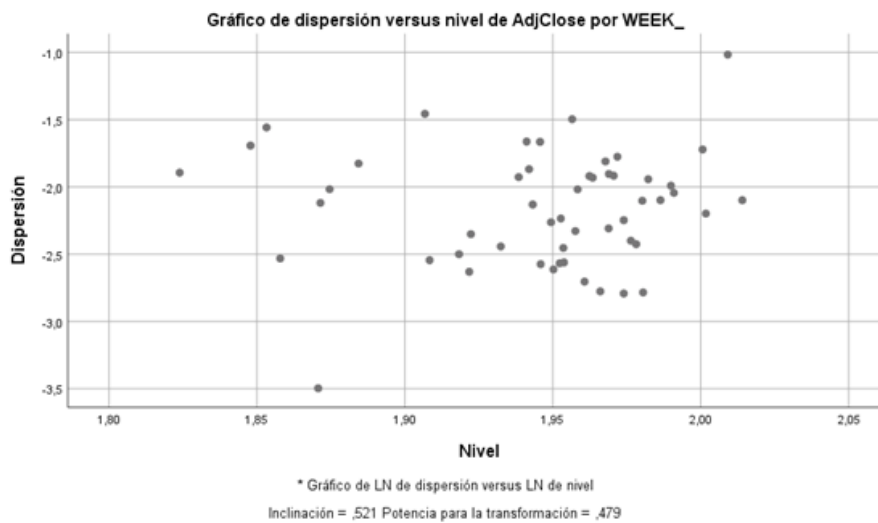
Fuente: elaboración propia mediante SPSS Statistics

Observamos que la serie no contiene movimientos estacionales, por lo tanto, seguimos con el estudio mediante modelos no estacionales.

Paso 3

A continuación, estudiamos la variabilidad de la serie para constatar que se trata de una serie con variabilidad constante. Para ello usamos la función “Explorar” y le pedimos que nos muestre el gráfico de dispersión y la estimación de la potencia a la que deberíamos elevar la serie si no fuera estable en variabilidad.

Figura 4. Gráfico de dispersión



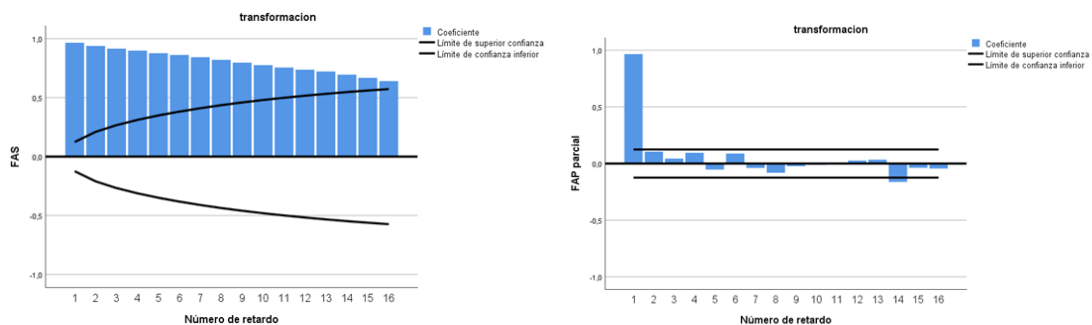
Fuente: elaboración propia mediante SPSS Statistics

El programa nos propone una estimación de potencia para elevar la serie, por lo tanto, procedemos a transformarla para convertirla en estable en variabilidad.

Paso 4

Continuamos con los gráficos de las autocorrelaciones simple y parcial para conocer si la media de la serie es constante.

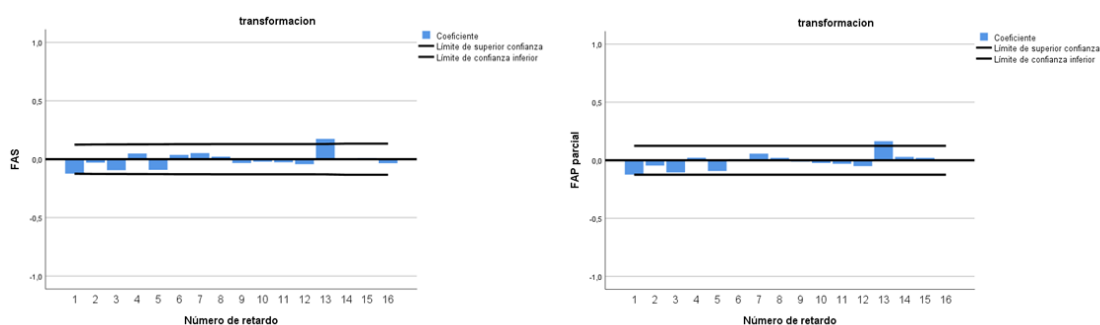
Figura 5. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Observamos que los retardos de la autocorrelación simple no convergen a cero rápidamente, por lo tanto, realizamos la diferenciación de la serie para convertirla en media constante, y dibujamos otra vez las autocorrelaciones para comprobarlo.

Figura 6. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Ahora ya tenemos una serie de media constante, ya que los retardos convergen a cero. Como podemos observar, ningún retardo se sale de los intervalos de confianza, debido a esto, podemos estar ante un $ARIMA(0,1,0)$, no obstante, probamos con los parámetros sobreponderados: $(1,1,0)$ y $(0,1,1)$. Probaremos también con una diferenciación más para contemplar más posibilidades, o sea, $ARIMA(1,2,0)$, ya que con una diferenciación más se nos sale un retardo para el parámetro (p).

Paso 5

Procedemos ahora a las pruebas de los modelos que hemos obtenido. Además de los modelos que nos han resultado, probaremos también el grupo los suavizados exponenciales que nos ofrece el programa, para así contemplar más posibilidades.

Tabla 2. Resultados de los modelos

Modelo	RECM	R ²	Tendencia favorable
$ARIMA(0,1,0)$	0,014	0,933	No
$ARIMA(1,1,0)$	0,014	0,934	No
$ARIMA(0,1,1)$	0,014	0,934	No
$ARIMA(1,2,0)$	0,018	0,894	Si
S. Simple	0,014	0,934	No
S. de Holt	0,014	0,934	Si
S. de Brown	0,015	0,922	Si
S. Amortiguado	0,014	0,934	No

Fuente: elaboración propia

Modelo: modelo a analizar.

Tendencia favorable: contrastar si la tendencia de la predicción se ajusta a la tendencia original de la serie.

RMSE: Raíz del Error Cuadrático Medio, es mejor cuanto menor sea, ya que indica menor error en las predicciones.

R²: Raíz Cuadrática, mejor cuanto mayor es, indica la cantidad en porcentaje que los datos explican el modelo.

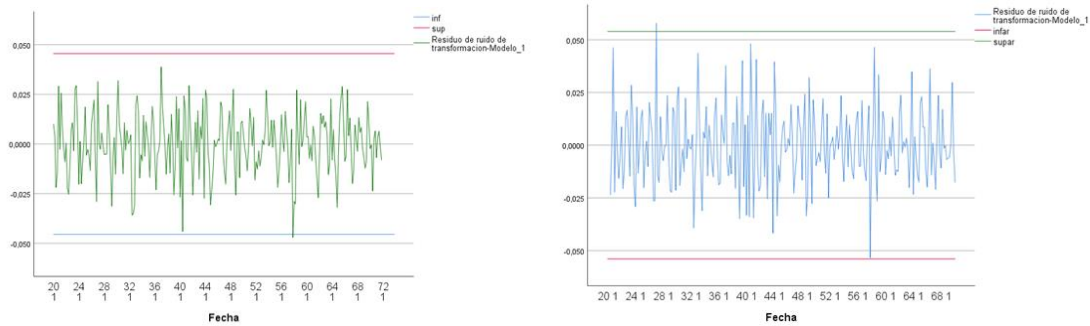
Paso 6

Como podemos observar en la tabla anterior, los únicos modelos que cumplen el requisito de tendencia favorable son el suavizado exponencial de Brown, el de Holt y el *ARIMA* (1,2,0), por lo tanto, procedemos a analizar sus residuos para comprobar si se tratan de ruido blanco, y en el caso del *ARIMA*, ruido blanco gaussiano.

En primer lugar, estudiamos si la media de los residuos es cero. Para ello utilizamos la “Prueba T para una muestra”. Efectivamente, el programa nos resuelve que las significaciones de los tres residuos son mayores que 0,05, por lo tanto, consideramos que la media sí es igual a cero en los dos residuos de los dos modelos.

En segundo lugar, dibujamos el gráfico de secuencia juntos con los límites inferior y superior para estudiar la variabilidad de la serie de los residuos. Los límites están calculados multiplicando por (-3) y (3) la desviación típica de los residuos.

Figura 7. Gráficos de secuencia



Fuente: elaboración propia mediante SPSS Statistics

Podemos apreciar que la línea de los residuos está dentro de los límites calculados a partir de la desviación típica de la serie en los tres casos, por lo tanto, podemos concluir en que sí presentan variabilidad constante.

En tercer lugar, y sólo para el modelo *ARIMA*, comprobamos si sigue una distribución normal los residuos. Para ello utilizamos la prueba K-S y los gráficos P-P. En efecto, las dos pruebas nos confirman que los residuos sí que siguen una distribución normal, por lo tanto podemos continuar con este modelo también.

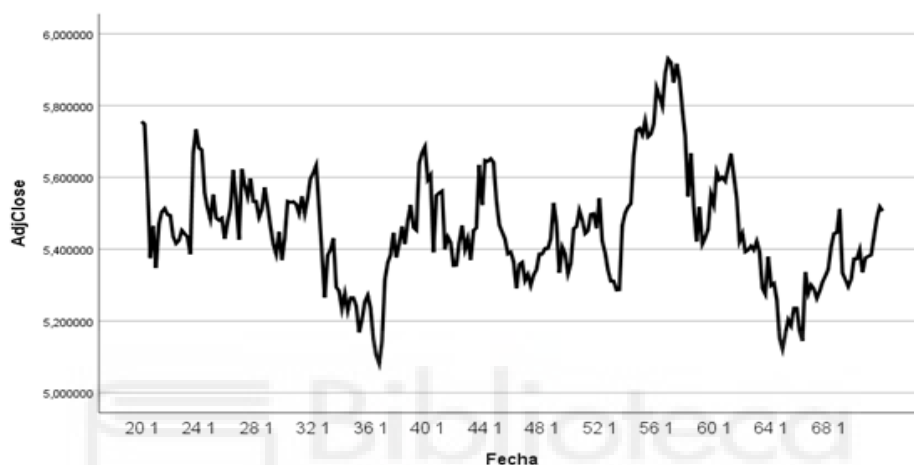
Por último, comprobamos si los datos de los residuos están incorrelados, para ello dibujaremos las autocorrelaciones de los residuos y nos fijaremos en los retardos y en los estadísticos Box-Ljung.

El programa nos muestra que los residuos del modelo de suavizado de Holt sí que están incorrelados, por el contrario, los residuos del suavizado de Brown y el *ARIMA*, no lo están. Por lo tanto, el modelo que utilizaremos para realizar las predicciones será el suavizado exponencial de Holt.

Paso 2

En primer lugar, como hemos venido haciendo, dibujamos el gráfico de secuencia para observar si existe algún factor estacional.

Figura 8. Gráfico de secuencia



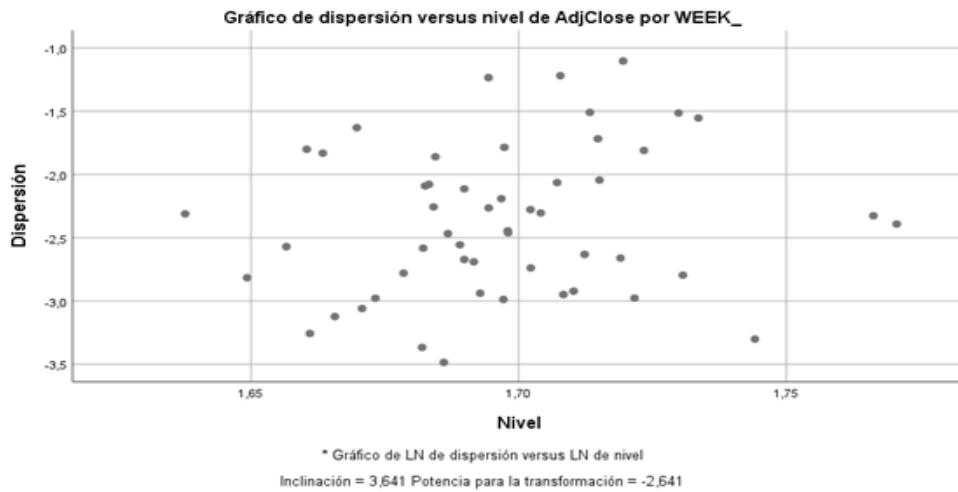
Fuente: elaboración propia mediante SPSS Statistics

Se puede observar que la serie no posee a simple vista ningún movimiento estacional.

Paso 3

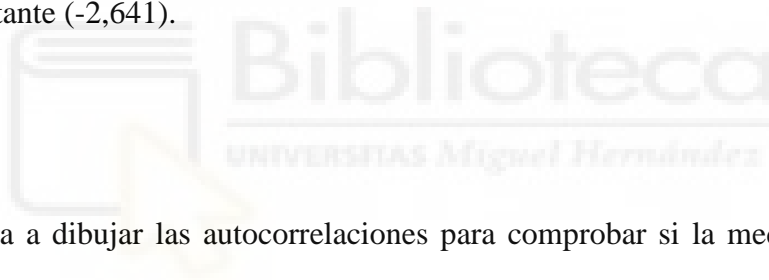
Continuamos con el estudio de la variabilidad. Utilizamos la función “Explorar” para que nos muestre el gráfico de dispersión y la estimación de la potencia a la que deberíamos transformar la variable si lo requiriera.

Figura 9. Gráfico de dispersión



Fuente: elaboración propia mediante SPSS Statistics

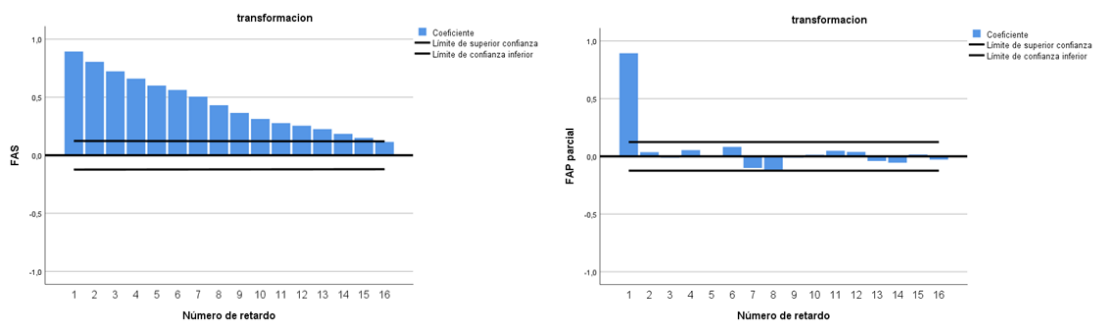
SPSS nos muestra su opción de estimación de potencia para convertir nuestra serie a varianza constante (-2,641).



Paso 4

Pasamos ahora a dibujar las autocorrelaciones para comprobar si la media de nuestra serie es constante.

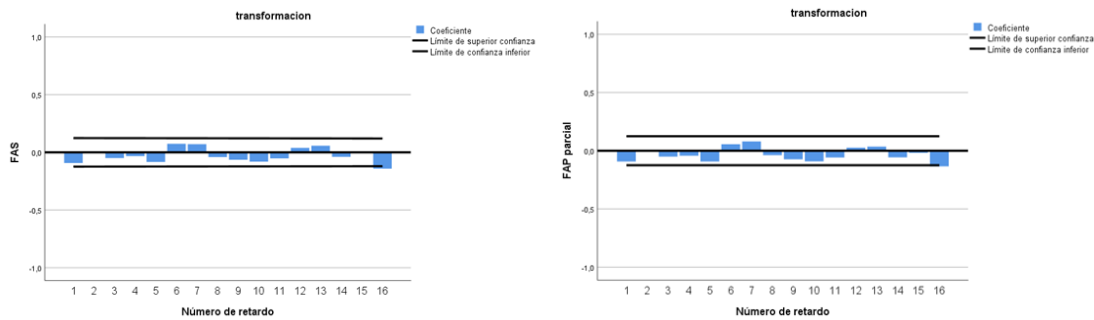
Figura 10. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Como nos ha estado pasando hasta ahora, la autocorrelación simple no cumple con la condición de decrecer rápidamente a cero, por lo tanto, realizamos la diferenciación de la serie, y volvemos a dibujar sus autocorrelaciones.

Figura 11. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Ahora los retardos ya nos quedan entre nuestros intervalos de confianza, por lo tanto, ya tenemos asegurado que la serie sea de media constante. Como podemos observar, no se sale ningún retardo de los límites de confianza, por lo tanto, el modelo *ARIMA* que podría ajustarse sería (0,1,0).

Paso 5

A causa de asegurarnos de los parámetros, probaremos con el *ARIMA* (0,1,0), (0,1,1) y (1,1,0), además del *ARMA* (0,1), (1,1), (2,1) y (1,0). Los *ARMA* los probaremos para comprobar si la serie sin diferenciar puede ajustarse a los modelos, ya que como hemos comentado antes, los gráficos de las autocorrelaciones nos salían dudosos. A parte de los de la familia *ARIMA*, probaremos también con los suavizados, ya que podrían ajustarse a las predicciones.

Después de probar los casos posibles, los resultados nos muestran que podríamos quedarnos con: el suavizado exponencial de Brown, el *ARMA* (1,1), el *ARMA* (2,1) y el *ARMA* (1,0). Estas son las características que presentan cada uno de ellos:

Tabla 3. Resultados de los modelos

Modelo	RECM	R ²	Tendencia favorable
<i>ARMA</i> (1,0)	0,00	0,799	Si
S. de Brown	0,00	0,76	Si

Fuente: elaboración propia

Modelo: modelo a analizar.

Tendencia favorable: contrastar si la tendencia de la predicción se ajusta a la tendencia original de la serie.

RMSE: Raíz del Error Cuadrático Medio, es mejor cuanto menor sea, ya que indica menor error en las predicciones.

R²: Raíz Cuadrática, mejor cuanto mayor es, indica la cantidad en porcentaje que los datos explican el modelo.

Paso 6

Una vez que tenemos los modelos posibles que podemos utilizar para las predicciones, vamos a estudiar sus residuos para constatar que se tratan de ruido blanco, aquellos modelos que no cumplan esta condición, deberemos eliminarlos de los posibles.

En primer lugar, comprobamos que los residuos siguen una media cero. Para ello utilizamos la función de “Prueba T para una muestra”. Como resultado de esta prueba, confirmamos que los residuos de los cuatro modelos posibles tienen media cero.

En segundo lugar, estudiamos la variabilidad de los residuos. Para ello hacemos lo mismo que hemos venido haciendo, representamos el gráfico de secuencia de los residuos junto con un límite inferior y superior, los cuales constan de la desviación típica de cada modelo multiplicada por (-3) y (3).

En efecto, todas las series de los residuos están dentro de los márgenes de la variabilidad aceptada, por lo tanto, confirmamos la condición de la variabilidad para todos.

En tercer lugar, y únicamente para los procesos *ARMA*, comprobaremos que los residuos siguen una distribución normal, para así confirmar si se tratan de ruido blanco Gaussiano, que es el requisito que deben cumplir los residuos de la familia *ARIMA*.

Para ello utilizamos dos pruebas, la función “Prueba K-S” y “Gráficos P-P”. Aunque las pruebas K-S no han sido positivas, los gráficos P-P nos confirman que sí que siguen una distribución normal, por lo tanto, seguimos con el proceso.

Por último, vamos a conocer si los datos de los residuos están incorrelados, para ello utilizaremos la misma función que antes para las autocorrelaciones simples y parciales. Para que los residuos cumplan el requisito, los retardos no deben salirse de los intervalos de confianza, y todos los estadísticos de Box-Ljung deben ser mayores que 0,05.

Observando las autocorrelaciones de los residuos de los posibles modelos que hemos seleccionado, podemos apreciar que el modelo *ARMA* sí que cumplen la condición de que sus residuos sean incorrelados, sin embargo, el último modelo, el correspondiente al suavizado exponencial de Brown, no cumple este requisito, y por lo tanto debemos eliminarlo de los posibles modelos que podríamos utilizar.

Como conclusión de este apartado, podemos decir que nos quedamos finalmente con el modelo *ARMA* (1,0) para la realización de las previsiones que haremos posteriormente.



Paso 2

Realizamos el gráfico de secuencia.

Figura 12. Gráfico de secuencia



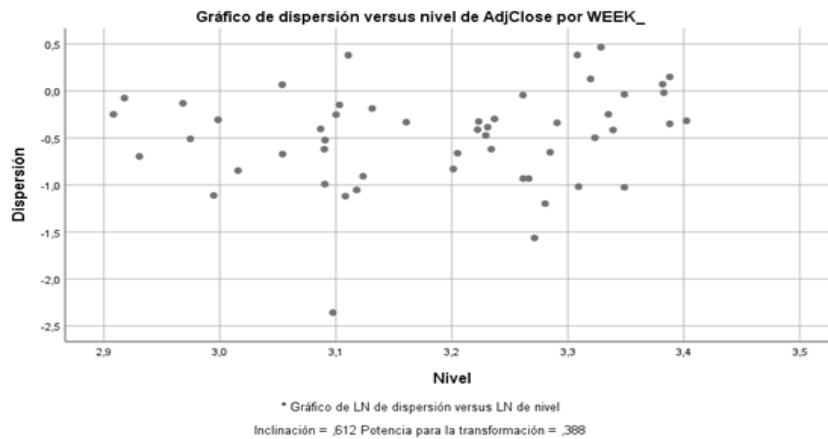
Fuente: elaboración propia mediante SPSS Statistics

Comprobamos que a simple vista no posee ningún factor estacional.

Paso 3

Variabilidad mediante la función “Explorar”, y dibujamos el gráfico de dispersión para obtener la estimación de la potencia a la que debería ser elevada la serie si tuviéramos que transformarla.

Figura 13. Gráfico de dispersión



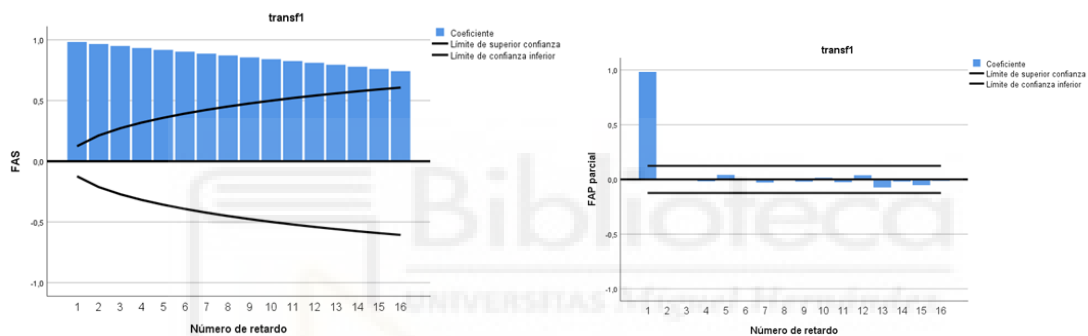
Fuente: elaboración propia mediante SPSS Statistics

Observamos que nos propone una estimación de potencia de 0,388. Existen muchas posturas sobre si elevar la serie al valor exacto, o si redondearlo a la unidad o a la mitad de la unidad, tanto en negativo como en positivo. En nuestro caso, utilizaremos el valor exacto que nos proporciona el programa. Utilizamos la función “Crear variable” para crear esta serie transformada, y así poder trabajar con ella en los pasos siguientes.

Paso 4

Procedemos a realizar las autocorrelaciones simples y parciales de la serie transformada anteriormente.

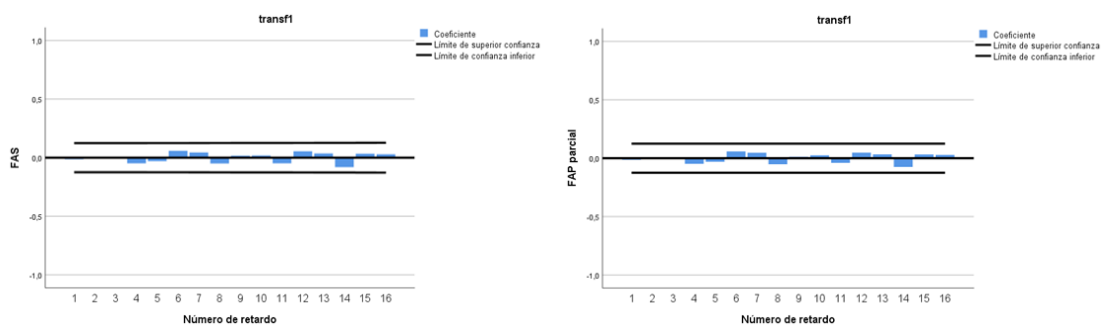
Figura 14. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Como podemos observar, la autocorrelación simple de la serie transformada no tiende a cero rápidamente, por lo tanto, consideramos que no es de media constante, y procedemos a diferenciarla para intentar que así sea.

Figura 15. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

En estas nuevas autocorrelaciones vemos que los retardos ya tienden a cero, por lo tanto, en el caso de que escogiéramos el método *ARIMA*, nuestros parámetros serían (p,d,q) (0,1,0).

Paso 5

Comenzamos a probar los parámetros posibles en la función “Crear modelos tradicionales”, y le pedimos que nos muestre la significación de nuestros parámetros.

Figura 16. Parámetros modelo ARIMA

				Estimación	SE	t	Sig.
transf1-Modelo_1	transf1	Ninguna transformación	Constante	,004	,003	1,346	,179
			Diferencia	1			

Fuente: elaboración propia mediante SPSS Statistics

Nos resuelve que la constante del modelo no es significativa, puesto que la significación es mayor que 0,05. Debido a esto y a que no tenemos ningún valor en ninguno de los otros dos parámetros, podemos considerar una alternativa al modelo *ARIMA*.

El método que más puede ajustarse a nuestros datos y propósitos es el método del suavizado exponencial. En nuestro software SPSS Statistics, nos da la opción de realizar diversos suavizados como el simple, el de Holt, el de Brown y el amortiguado. Tendremos que realizar una prueba con cada alternativa para comprobar qué método es mejor para nuestra serie.

Tabla 4. Resultados de los modelos

Modelo	RECM	R²	Tendencia favorable
<i>ARIMA</i> (0,1,0)	0,045	0,981	Si
S. Simple	0,045	0,981	No
S. de Holt	0,045	0,981	Si
S. de Brown	0,05	0,978	Si
S. Amortiguado	0,045	0,981	Si

Fuente: elaboración propia

Modelo: modelo a analizar.

Tendencia favorable: contrastar si la tendencia de la predicción se ajusta a la tendencia original de la serie.

RMSE: Raíz del Error Cuadrático Medio, es mejor cuanto menor sea, ya que indica menor error en las predicciones.

R²: Raíz Cuadrática, mejor cuanto mayor es, indica la cantidad en porcentaje que los datos explican el modelo.

Como apreciamos en la tabla, a simple vista casi todos los modelos presentan las mismas características, la única alternativa que podríamos descartar por su falta de ajuste a la tendencia original serie la del suavizado exponencial simple.

Paso 6

Para poder discernir qué modelo es el más apropiado, y teniendo en cuenta que las características de los 4 seleccionados son muy parecidas, analizaremos los residuos de cada uno para comprobar que son ruido blanco y que no contienen información necesaria del modelo.

En primer lugar, comprobamos que la media de los residuos es igual a cero, para ello utilizamos la función “Prueba T para una muestra” para cada uno de los residuos de los modelos posibles.

Figura 17. Resultado Prueba T para una muestra

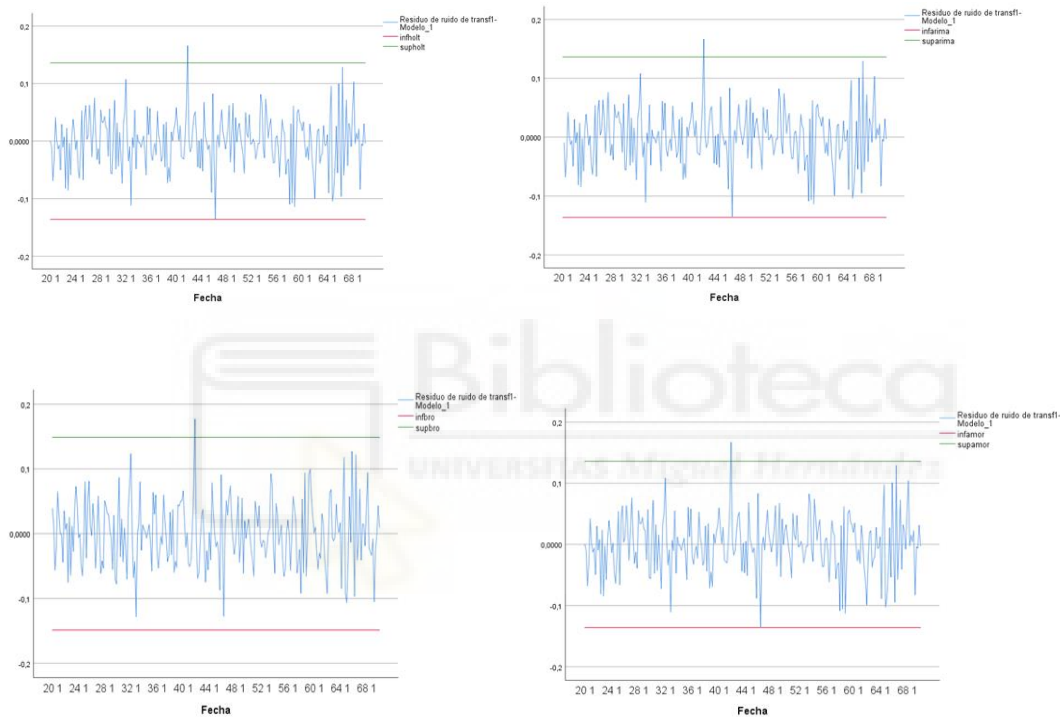
Prueba para una muestra						
Valor de prueba = 0						
	t	gl	Sig. (bilateral)	Diferencia de medias	95% de intervalo de confianza de la diferencia	
					Inferior	Superior
Residuo de ruido de transf1-Modelo_1	,000	249	1,000	,00000	-,0057	,0057
Residuo de ruido de transf1-Modelo_1	-,339	250	,735	-,00097	-,0066	,0047
Residuo de ruido de transf1-Modelo_1	,112	250	,911	,00035	-,0058	,0065
Residuo de ruido de transf1-Modelo_1	,074	250	,941	,00021	-,0054	,0058

Fuente: elaboración propia mediante SPSS Statistics

Vemos que las significaciones de los residuos de todos los modelos son mayores que 0,05, por lo tanto, podemos afirmar que los residuos siguen una media igual a cero.

En segundo lugar, vamos a dibujar el gráfico de los residuos para observar si poseen varianza constante. Para ello lo dibujaremos entre las líneas del límite inferior y límite superior, que son resultado de multiplicar por (-3) y por (3) la desviación típica de cada uno de los residuos.

Figura 18. Gráficos de secuencia de los residuos



Fuente: elaboración propia mediante SPSS Statistics

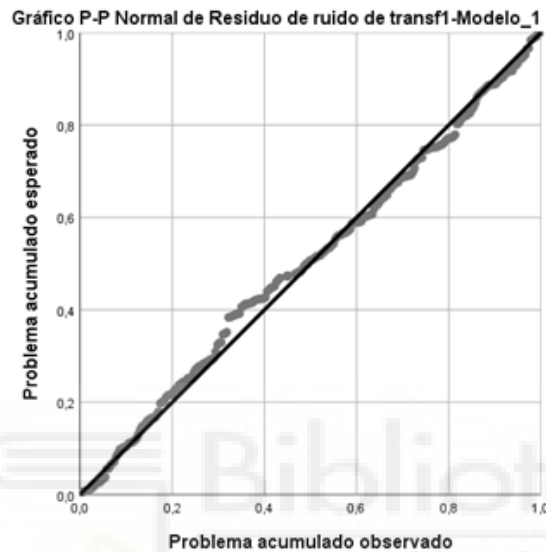
Como nos muestran los gráficos, todos los residuos poseen varianza constante, y por lo tanto son homocedásticos.

En tercer lugar, estudiamos la normalidad de la serie, pero únicamente del modelo ARIMA (0,1,0), debido a que los suavizados no requieren que la serie de los residuos sigan una distribución normal.

Usaremos dos maneras de estudiar si siguen una distribución normal: con una prueba Kolmogorov-Smirnov (K-S) para una muestra, o con el gráfico P-P de los residuos.

- a) K-S de una muestra: no nos da un resultado significativo para poder concluir si sigue distribución normal o no, por lo tanto, dibujaremos el gráfico P-P.
- b) Gráfico P-P: comprobamos que los residuos están en torno a la tendencia de la normal, por lo tanto, podemos concluir en que sí que sigue una distribución normal.

Figura 19. Gráfico P-P d un residuo



Fuente: elaboración propia mediante SPSS Statistics

Por último, examinaremos si estos residuos están incorrelados, es decir, no presentan dependencia uno de otro. Para ello utilizaremos la misma función que para hacer las autocorrelaciones que hicimos al principio del proceso, pero ahora introduciendo los datos de los residuos de cada modelo. Para comprobar si son incorrelados o no, nos fijaremos en el estadístico de Box-Ljung, si todos son mayores a 0,05, significaría que sí que son incorrelados.

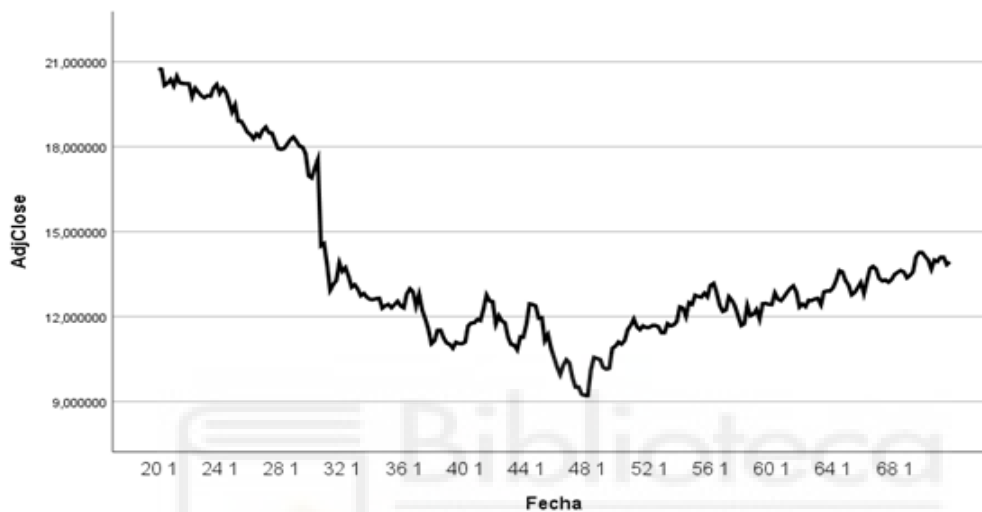
Como resultado tenemos que todos los modelos presentan residuos incorrelados, a excepción de un modelo, el modelo de suavizado exponencial de Brown. Por lo tanto, decidimos descartar este modelo de los posibles, para quedarnos finalmente con tres alternativas factibles para realizar nuestras predicciones:

- ARIMA (0,1,0)
- Suavizado exponencial de Holt
- Suavizado exponencial amortizado

Paso 2

Dibujamos el gráfico de secuencia para observar si contiene algún factor estacional.

Figura 20. Gráfico de secuencia



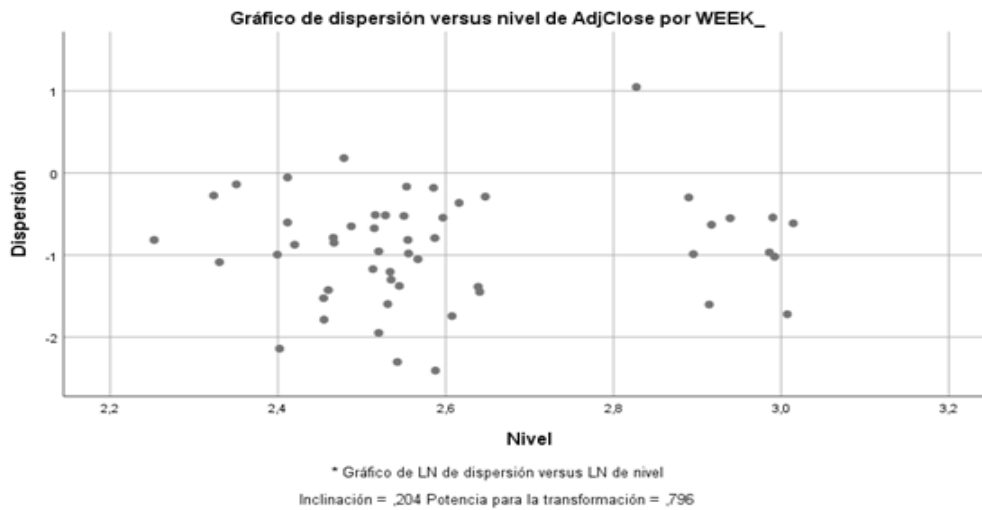
Fuente: elaboración propia mediante SPSS Statistics

Comprobamos que a simple vista no muestra comportamiento estacional ni se observan patrones de movimientos, por lo tanto, podemos proseguir con los modelos no estacionales que consideramos.

Paso 3

Estudiamos la variabilidad con la función “Explorar” y le pedimos que nos muestre una estimación de potencia para elevar la variable en caso de que fuera necesario la transformación para convertirla en estable en variabilidad.

Figura 21. Gráfico de dispersión



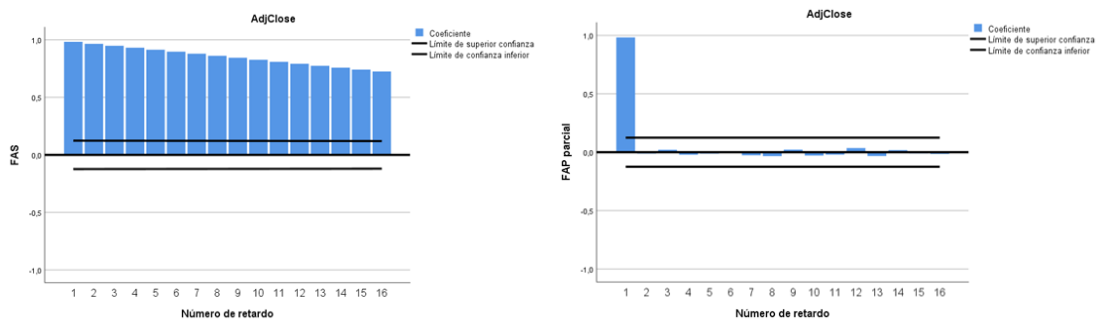
Fuente: elaboración propia mediante SPSS Statistics

Nos propone una estimación de potencia de 0,796, por lo tanto, siguiendo con nuestro método del redondeo, lo redondeamos a la unidad, y por lo tanto no es necesario transformar la variable.

Paso 4

Ahora procedemos a dibujar las autocorrelaciones para comprobar su estabilidad en media. Utilizaremos para ello, como hemos venido haciendo, las autocorrelaciones simple y parcial

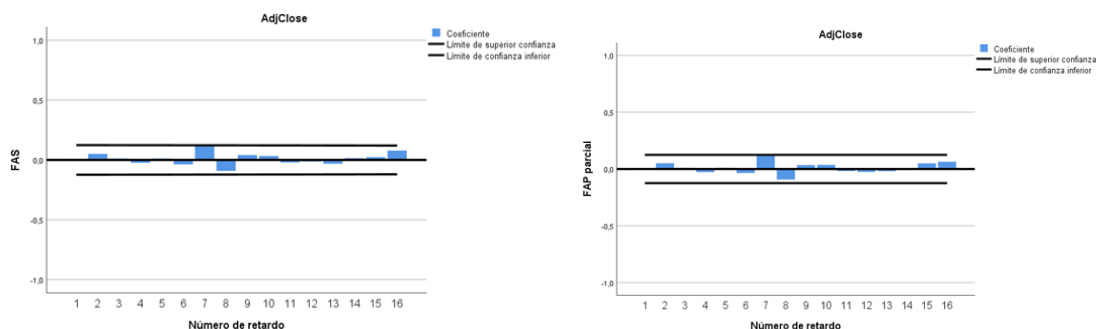
Figura 22. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Observamos que la autocorrelación simple no tiende a cero, por lo tanto, realizamos la diferenciación de la serie y dibujamos sus autocorrelaciones.

Figura 23. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Con esta transformación ya podemos observar que sí que posee una media constante, además, podemos apreciar que ningún retardo se sale de los intervalos de confianza, por lo tanto, estaríamos ante un *ARIMA* (0,1,0).

Paso 5

Nos aparece que la constante del modelo no es significativa, aunque no por demasiado. Por lo tanto, debido a esto y a que no presenta valores para ninguno de los otros dos parámetros, podemos considerar otros modelos. Los más apropiados para nuestro estudio, como hemos comentado anteriormente, son los suavizados.

Tabla 5. Resultados de los modelos

Modelo	RECM	R ²	Tendencia favorable
<i>ARIMA</i> (0,1,0)	0,341	0,987	Si
S. Simple	0,342	0,987	No
S. de Holt	0,342	0,987	Si
S. de Brown	0,368	0,985	No
S. amortizado	0,341	0,987	Si

Fuente: elaboración propia

Modelo: modelo a analizar.

Tendencia favorable: contrastar si la tendencia de la predicción se ajusta a la tendencia original de la serie.

RMSE: Raíz del Error Cuadrático Medio, es mejor cuanto menor sea, ya que indica menor error en las predicciones.

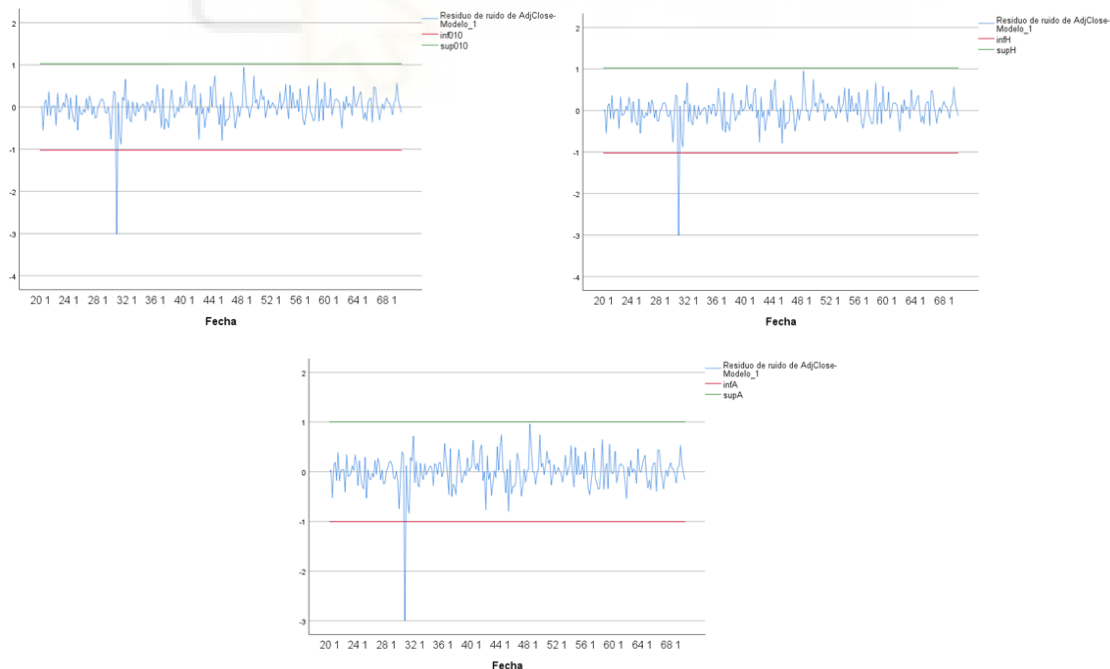
R²: Raíz Cuadrática, mejor cuanto mayor es, indica la cantidad en porcentaje que los datos explican el modelo.

Paso 6

Procedemos a analizar los residuos para comprobar que se tratan de ruido blanco. Como ya sabemos, si no se tratara de ruido blanco, los residuos contendrían información acerca del modelo, y por lo tanto no podríamos seguir con el modelo analizado.

En primer lugar, comprobamos que las medias de los residuos de los tres modelos posibles son cero. Y en efecto, las significaciones del modelo “Prueba T para una muestra”, nos confirman que poseen media cero. Posteriormente, en segundo lugar, estudiamos la variabilidad de la serie de los residuos para comprobar que es constante, para ello utilizamos los márgenes inferior y superior para observar si se mueve dentro de dichos márgenes.

Figura 24. Gráficos de secuencia de los residuos



Fuente: elaboración propia mediante SPSS Statistics

Como se puede observar, aunque existe algún valor atípico en la serie, podemos afirmar que la serie posee una variabilidad constante.

En tercer lugar, y únicamente para modelos *ARIMA*, procedemos al análisis de la serie para comprobar que sigue una distribución Normal. Para ello utilizamos las pruebas de K-S y los gráficos P-P. Efectivamente, el programa nos resuelve que los residuos superan las dos pruebas.

Por último, comprobar si los datos de los residuos están incorrelados. Confirmamos que así es, los tres modelos presentan residuos incorrelados, por lo tanto, podemos concluir que tenemos tres modelos posibles para las predicciones:

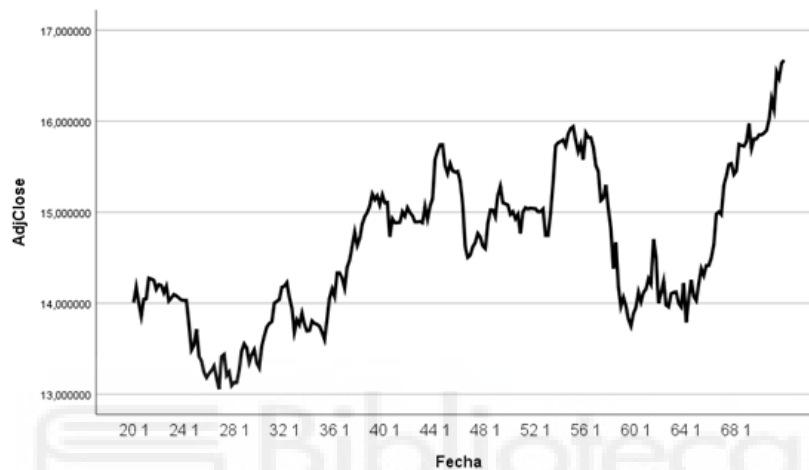
- *ARIMA* (0,1,0)
- Suavizado de Holt
- Suavizado Amortiguado



Paso 2

Como anteriormente, empezamos con el gráfico de secuencia para comprobar si posee la serie factores estacionales.

Figura 25. Gráfico de secuencia



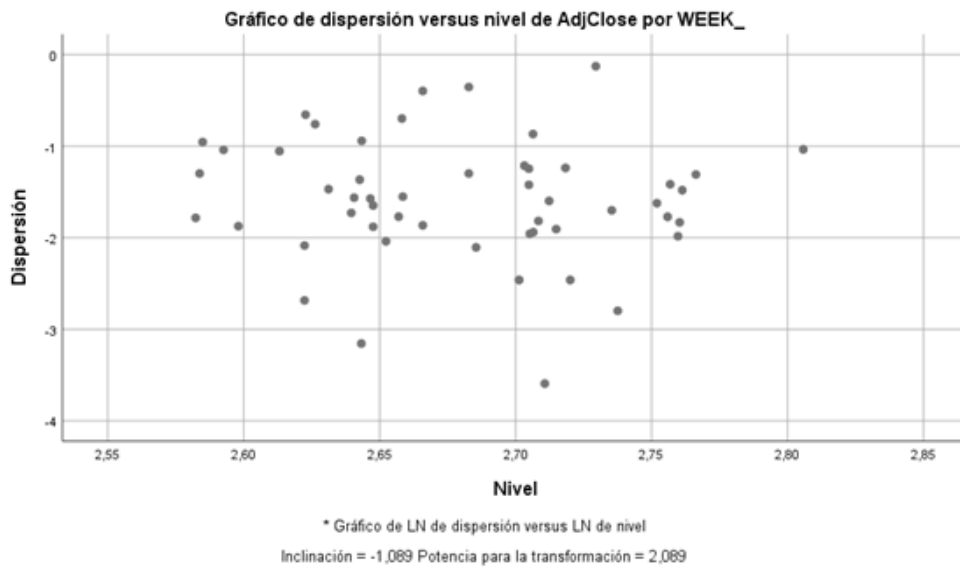
Fuente: elaboración propia mediante SPSS Statistics

se puede observar que a simple vista la serie no posee ningún movimiento estacional, por lo tanto, continuamos con el proceso.

Paso 3

Como solemos hacer, continuamos con el estudio de la variabilidad mediante la función “Explorar”, y le pedimos que nos muestre el gráfico de dispersión con la estimación de la potencia a la que deberíamos elevar la serie para transformarla si fuera necesario.

Figura 26. Gráfico de dispersión



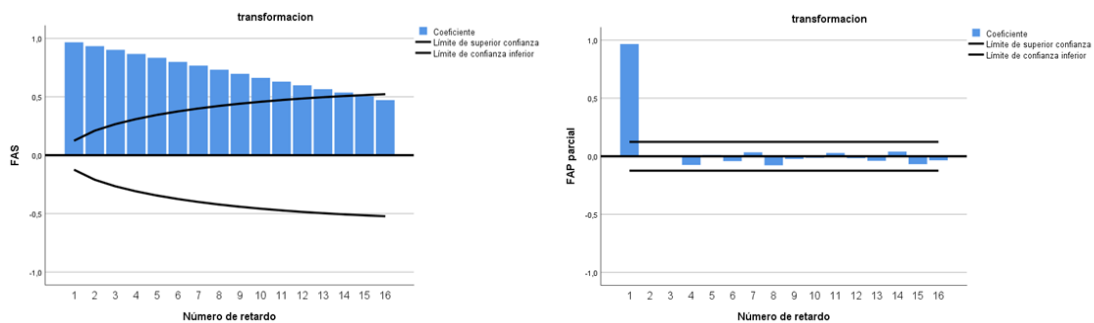
Fuente: elaboración propia mediante SPSS Statistics

Nos da la estimación de la potencia a la que debemos elevar la serie para transformarla en varianza constante, por lo tanto, procedemos a ello y la transformamos.

Paso 4

Continuamos dibujando las autocorrelaciones simples y parciales para comprobar que la serie transformada de varianza constante también posee una media constante.

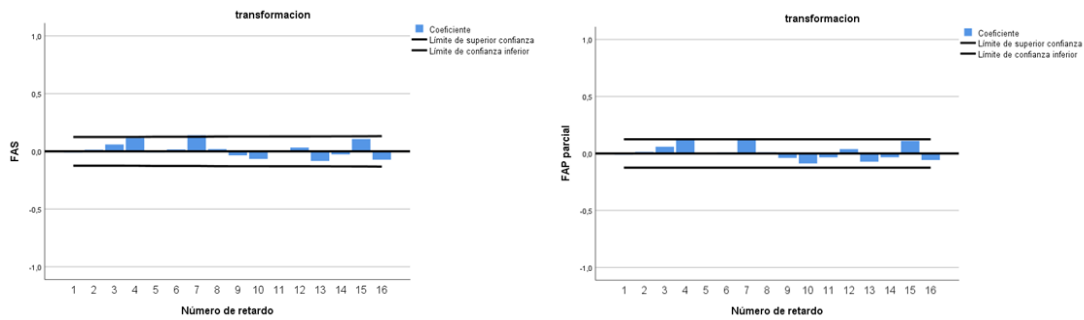
Figura 27. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

La autocorrelación simple no tiende a cero rápidamente, por lo que procedemos a realizar la diferenciación de la serie, y dibujamos nuevamente las autocorrelaciones para conocer si esta vez sí que es de media constante.

Figura 28. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Ahora podemos ver que los retardos sí convergen a cero, por lo tanto, se interpreta que la serie ya posee una media constante. Como ninguno de los retardos se sale de los intervalos de confianza, podemos probar con el *ARIMA* (0,1,0). No obstante, probaremos también con los valores sobreponderados (1,1,0) y (0,1,1).

Paso 5

Es hora de probar los modelos que nos ha proporcionado el programa, a parte de ellos, probaremos también, como solemos hacer, con los suavizados exponenciales.

Una vez probados los modelos posibles, estas son las características de cada uno de ellos:

Tabla 6. Resultados de los modelos

Modelo	RECM	R ²	Tendencia favorable
<i>ARIMA</i> (0,1,0)	6,06	0,957	Si
S. de Holt	6,061	0,957	Si
S. de Brown	6,416	0,952	Si

Fuente: elaboración propia

Modelo: modelo a analizar.

Tendencia favorable: contrastar si la tendencia de la predicción se ajusta a la tendencia original de la serie.

RMSE: Raíz del Error Cuadrático Medio, es mejor cuanto menor sea, ya que indica menor error en las predicciones.

R²: Raíz Cuadrática, mejor cuanto mayor es, indica la cantidad en porcentaje que los datos explican el modelo.

Paso 6

Una vez que se tienen los posibles modelos, vamos a estudiar sus residuos para confirmar que son ruido blanco y que no poseen información acerca de la serie.

En primer lugar, se comprueba que los residuos tienen media cero mediante la función “Prueba T para una muestra”.

En efecto, todos los residuos de los modelos poseen media cero, ya que sus significaciones son mayores que 0,05, por lo que se puede continuar con todos ellos.

En segundo lugar, estudiamos la variabilidad realizando el gráfico de secuencia de las series de los residuos junto con sus intervalos inferiores y superiores, los cuales se calculan multiplicando por (-3) y (3) la desviación típica de cada serie de residuos.

Dibujamos los gráficos de secuencia y los resultados nos informan de que los tres modelos superan esta prueba de variabilidad, ya que las series de los residuos están dentro de los márgenes inferiores y superiores.

En tercer lugar, y únicamente para el *ARIMA*, realizaremos la prueba de la normalidad. Los modelos de la familia *ARIMA*, como ya hemos comentado, necesitan que se cumpla el requisito de que sus residuos sigan una distribución normal, por lo tanto, comprobamos esto mediante la función “Prueba K-S” y los gráficos P-P.

La prueba K-S no es superada por los residuos, sin embargo, los gráficos P-P nos muestran que sí que siguen una distribución normal los residuos del modelo *ARIMA* seleccionado.

Por último, comprobamos si los datos de las series de los residuos están incorrelados, condición indispensable para poder seguir adelante con los modelos. Para ello utilizaremos las autocorrelaciones y los estadísticos Box-Ljung.

Los resultados que nos proporciona el programa nos hacen darnos cuenta de que los datos de los residuos del tercer modelo seleccionado, suavizado exponencial de Brown, no están incorrelados. Por lo tanto, debemos eliminar este modelo de los posibles modelos para realizar las predicciones.

Podemos concluir en que podemos seleccionar para las predicciones dos de los tres modelos posibles que hemos estudiado, éstos son:

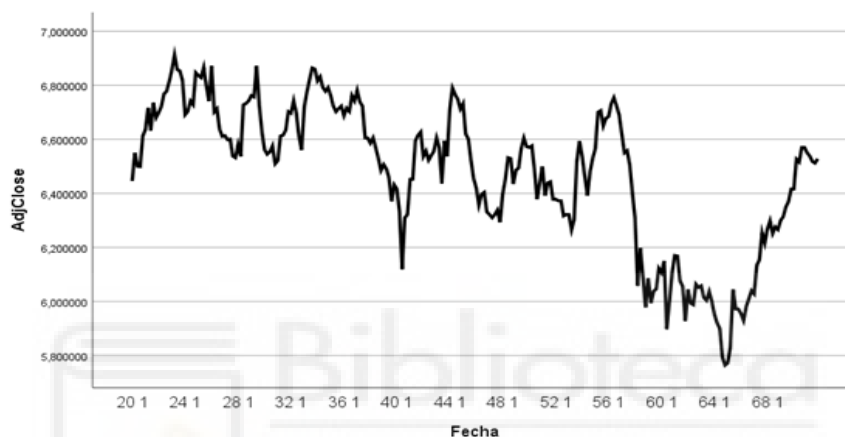
- *ARIMA* (0,1,0)
- Suavizado exponencial de Holt



Paso 2

Como solemos hacer, dibujamos el gráfico de secuencia para discernir si existen factores estacionales en la serie.

Figura 29. Gráfico de secuencia



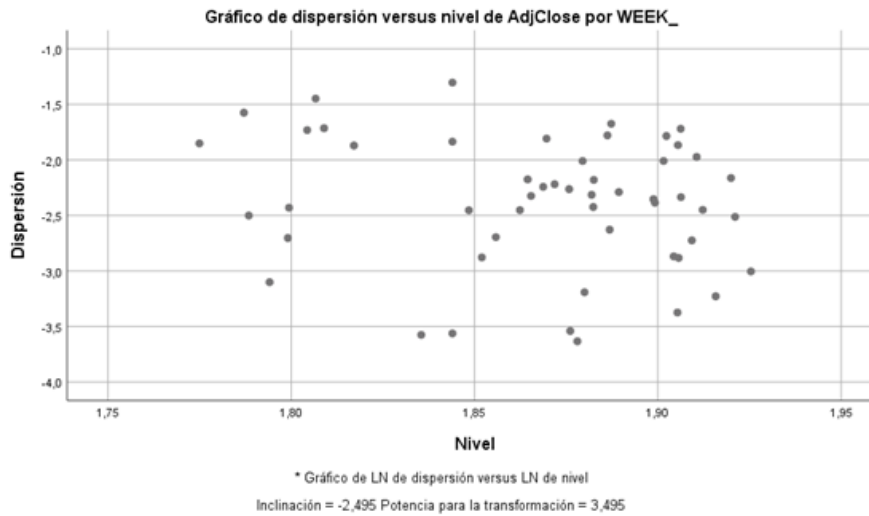
Fuente: elaboración propia mediante SPSS Statistics

Observamos que la serie no posee movimientos estacionales, por lo tanto, proseguimos con los modelos no estacionales.

Paso 3

Seguidamente, se analiza la variabilidad de la serie para comprobar que es constante, Utilizaremos, como hemos estado haciendo, la función “Explorar” y le pediremos que nos dibuje el gráfico de dispersión y que nos proporcione una estimación de potencia para elevar la serie en el caso de que no fuera constante en variabilidad.

Figura 30. Gráfico de dispersión



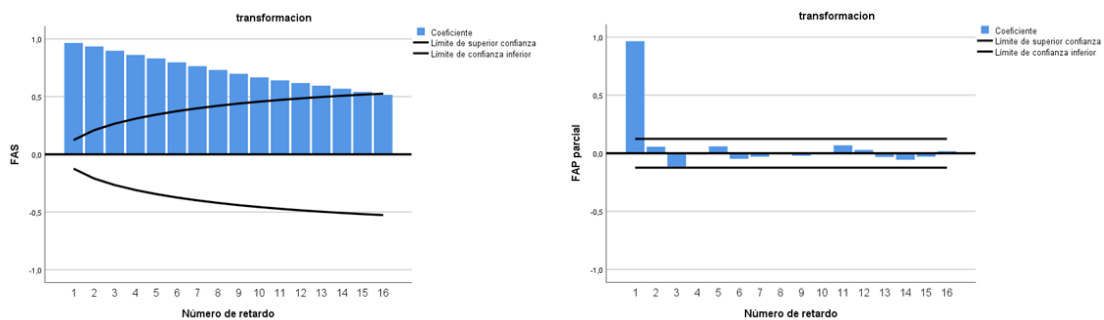
Fuente: elaboración propia mediante SPSS Statistics

Como se ve en la imagen, el programa propone una estimación de potencia de 3,495. Por lo tanto, se ha de crear una variable que estará formada por la serie estudiada elevada a dicha estimación de potencia.

Paso 4

Comenzamos ahora con los gráficos de las autocorrelaciones. Dibujamos las simples y las parciales y comprobamos si la serie posee media constante.

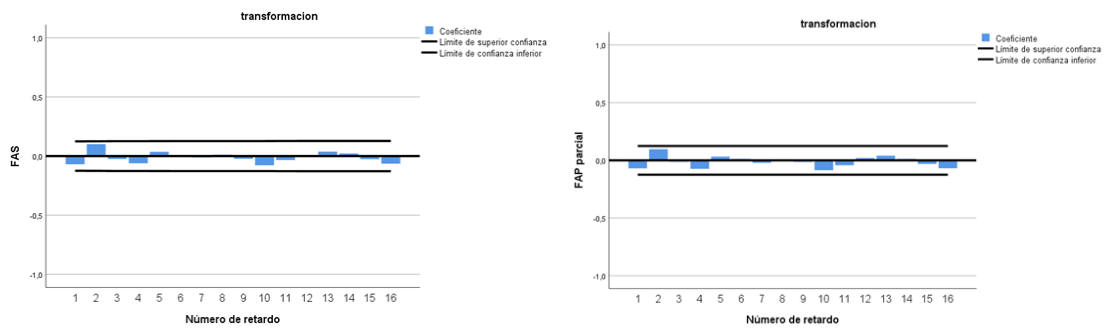
Figura 31. Autocorrelaciones, simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Los retardos de la autocorrelación simple no convergen rápidamente a cero, por lo tanto, procedemos a diferenciar la serie y dibujar las autocorrelaciones de nuevo.

Figura 32. Autocorrelaciones. simple izquierda y parcial derecha



Fuente: elaboración propia mediante SPSS Statistics

Ahora los retardos ya se comportan como una serie de media constante, además, como podemos observar, ningún retardo se sale de los intervalos de confianza, por lo tanto, podríamos estar ante un modelo *ARIMA* (0,1,0). No obstante, como solemos hacer, probaremos con los parámetros sobreponderados: (1,1,0) y (0,1,1).

Paso 5

Una vez estudiadas las variabilidades y la media de la serie, empezamos con las pruebas de los modelos. Para ello, probaremos con los modelos *ARIMA* que hemos obtenido, y con los suavizados que disponemos en SPSS.

Tabla 7. Resultados de los modelos

Modelo	RECM	R ²	Tendencia favorable
<i>ARIMA</i> (0,1,0)	25,63	0,93	No
<i>ARIMA</i> (1,1,0)	25,62	0,931	No
<i>ARIMA</i> (0,1,1)	25,63	0,931	No
S. Simple	25,53	0,93	No
S. de Holt	25,604	0,931	Si
S. de Brown	25,578	0,919	Si
S. Amortiguado	25,56	0,931	No

Fuente: elaboración propia

Modelo: modelo a analizar.

Tendencia favorable: contrastar si la tendencia de la predicción se ajusta a la tendencia original de la serie.

RMSE: Raíz del Error Cuadrático Medio, es mejor cuanto menor sea, ya que indica menor error en las predicciones.

R²: Raíz Cuadrática, mejor cuanto mayor es, indica la cantidad en porcentaje que los datos explican el modelo.

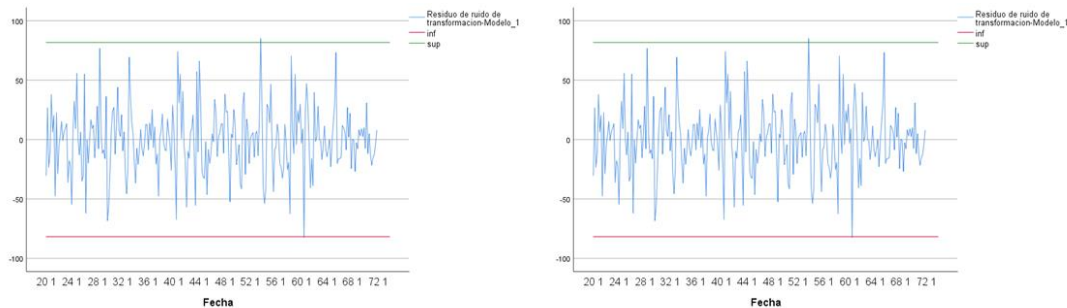
Paso 6

En el análisis de esta serie, podemos observar que sólo dos modelos cumplen el requisito de tendencia favorable, que hace referencia a la dirección de la predicción y si se ajusta o no a la dirección real de la serie. Por lo tanto, sólo estudiaremos los residuos de estos dos modelos.

En primer lugar, comprobamos que la media de la serie de los residuos es cero. Para ello utilizamos “Prueba T para una muestra” para comparación de medias. El programa nos resuelve que las significaciones son mayores que 0,05, por lo tanto, aceptamos que la media de los residuos es cero.

En segundo lugar, estudiamos la variabilidad mediante el gráfico de secuencia. Para ello dibujamos la serie de los residuos junto con los límites inferior y superior, los cuales están calculados multiplicando por (-3) y (3) la desviación típica de los residuos.

Figura 33. Gráficos de secuencia de los residuos



Fuente: elaboración propia mediante SPSS Statistics

Se puede apreciar que los residuos están dentro de los límites marcados por la desviación típica, por lo tanto, cumplen el requisito de la variabilidad constante.

Por último, analizamos si los datos de los residuos están incorrelados, es decir, que no contienen información de la serie de datos original que sea necesaria. Para ello utilizamos las autocorrelaciones de los residuos y nos fijamos en los estadísticos Box-Ljung y en si los retardos están todos entre los intervalos de confianza.

En efecto, SPSS nos resuelve que todos los estadísticos Box-Ljung de los dos residuos de los dos modelos, son mayores de 0,05, además, ningún retardo de los gráficos de autocorrelaciones se sale de los intervalos de confianza. Por lo tanto, podemos afirmar que los residuos están incorrelados. Para concluir, en Iberdrola probaremos con estos dos modelos:

- Suavizado exponencial de Holt
- Suavizado exponencial de Brown

En la siguiente tabla mostramos el resumen los modelos obtenidos para cada una de nuestras empresas con un año de cantidad de datos seleccionados para el estudio:

Tabla 8. Resumen de los modelos obtenidos a 1 año

BBVA	Suavizado de Holt		
SANTANDER			<i>ARMA (1,0)</i>
IBERDROLA	Suavizado de Holt	Suavizado de Brown	
REPSOL	Suavizado de Holt		<i>ARIMA (0,1,0)</i>
ARCELORMITTAL	Suavizado de Holt		<i>ARIMA (0,1,0)</i>
SIEMENS GAMESA	Suavizado de Holt		<i>ARIMA (0,1,0)</i>

Fuente: elaboración propia

4.2. LONGITUD DE LAS SERIES TEMPORALES

La longitud de las series temporales hace referencia a la cantidad de datos pasados que introduciremos en nuestros modelos para obtener las predicciones. Como hemos

comentado anteriormente, este factor es importante debido al carácter diario de nuestro estudio. Al ser diarias las predicciones, no podemos basarnos en una serie temporal demasiado larga, ya que podría contener perturbaciones o variaciones que no fueran útiles para el desarrollo del modelo estadístico. Estas perturbaciones pueden ser por ejemplo: crisis financieras, reestructuraciones empresariales, cambios en entornos competitivos, etc. Estos sucesos podrían ser útiles para obtener predicciones a medio o largo plazo, ya que afectan a la vida de cada empresa y a su evolución general, no obstante, no resultarían útiles para nuestro estudio puesto que estos sucesos no afectan a las variaciones diarias de los precios de las acciones.

Las variaciones diarias en los precios de las acciones escogidas, que es lo que a nosotros nos interesa, no son resultado de grandes cambios coyunturales o crisis financieras que afecten a sus entornos, sino que más bien son resultado del propio juego de oferta y demanda que sufre cada uno de los títulos diariamente. Este juego oferta-demanda está sujeto a especuladores e inversores de todo tipo, tanto personas físicas como personas jurídicas o entes públicos. Por ello deberemos ajustar la longitud de cada serie temporal a la última etapa de variaciones que hayan sufrido los precios de nuestras acciones, descartando los movimientos anteriores que no influyen en el desarrollo diario de los títulos.

Para intentar comprender qué cantidad de datos es mejor utilizar para calcular las predicciones, realizaremos el mismo procedimiento que hemos hecho en el apartado anterior para identificar los modelos posibles en cada empresa, pero cambiando el horizonte temporal, esta vez las pruebas se realizarán con 6 meses de datos, en vez de un año.

4.2.1. IDENTIFICACIÓN DE LA LONGITUD DE LAS SERIES TEMPORALES

En este apartado se llevarán a cabo las pruebas estadísticas para selección de modelos, pero basándonos en 6 meses de datos, siendo la fecha de fin la misma que en las bases de datos de un año. Como hemos comentado, las pruebas y los cálculos son exactamente los mismos, pero cambiando la variable de la longitud de la serie.

Puesto que los pasos a seguir son los mismos que en el apartado anterior, no nos detendremos en cada uno de ellos como hemos hecho anteriormente. Se propondrán directamente los modelos que han resultado con el cambio de la longitud a 6 meses. Cabe destacar que, aunque no se hayan reflejado los pasos y las imágenes, todos los cálculos están desarrollados en el programa SPSS Statistics.

4.2.1.1. MODELOS RESULTANTES CON 6 MESES DE LONGITUD DE SERIE

A continuación, se exponen los modelos estadísticos resultantes de las pruebas realizadas a partir de diferente base temporal, esta vez con 6 meses de bases de datos.

Tabla 9. Resumen de modelos obtenidos a 6 meses

BBVA	Suavizado de Holt		
SANTANDER	Suavizado de Holt	Suavizado de Brown	ARMA (1,0)
IBERDROLA	Suavizado de Holt		ARMA (1,0)
REPSOL	Suavizado de Holt	Suavizado de Brown	
SIEMENSGAMESA	Suavizado de Holt		
ARCELORMITTAL	Suavizado de Holt	Suavizado de Brown	ARIMA (0,1,0)

Fuente: elaboración propia

Como se puede observar, algunas empresas ajustan los mismos modelos tanto si la base de datos utilizada es de un año como si es de 6 meses, no obstante, la mayoría de las empresas cambian los modelos, o incrementa o disminuye el número de modelos posibles. Esto se debe en parte a que las tendencias de las series cambian significativamente según la longitud de base de datos que utilizemos para analizarlas.

De momento, no se sabe si las predicciones que calcularemos se ajustarán mejor a los modelos en un año o en 6 meses. Los modelos basados en suavizados de tendencia exponenciales nos van a ser útiles para predecir la posible tendencia que tendrá la serie en el futuro, en cambio, los modelos de la familia *ARIMA* puede que nos ajusten mejor los valores exactos en cada día de cotizaciones. Cabe destacar que no obtendremos en

ningún caso el dato exacto de la cotización, los modelos utilizados nos servirán, pero hasta cierto punto.

5. FASE 3: PROCESO DE PREDICCIÓN

En este apartado se comenzará con el proceso de predicción para conocer finalmente si nuestros modelos aciertan en sus cálculos o si, por el contrario, no se ajustan a los datos reales de las cotizaciones de nuestras empresas. Este proceso se llevará a cabo en el software informático R Studio.

Al realizar la selección de modelo y las predicciones en programas distintos, se podrá contrastar también si los modelos resultantes de SPSS Statistics son realmente correctos o si no funcionan en R Studio. Lo normal en nuestros casos es que los modelos identificados sean útiles en este segundo programa tanto como en el primero, ya que se basan en los mismos parámetros y en los mismos cálculos. Si por el contrario se observara que los modelos no son significativos en R Studio, sería motivo de reevaluación de dichos modelos y replanteamiento de los mismos.

5.1. PREDICCIONES EN RSTUDIO DE LAS SERIES TEMPORALES

A continuación, se procederá a realizar las predicciones en el software RStudio a partir de los modelos identificados anteriormente en el SPSS. En las siguientes tablas se resumirán los datos obtenidos referentes a los cálculos de las predicciones, posteriormente, procederemos a su explicación.

Guía para las tablas de resultados

Valor real: cotizaciones reales en las fechas seleccionadas (X).

Valor pronosticado: valor que hemos obtenido mediante R Studio (Y).

Error: resultado de $(Y_t - X_t)$

Error cuadrático: el error anterior elevado al cuadrado (E^2).

Tendencia real: tendencia de la serie real (S) en los datos seleccionados, el valor 1 simboliza una subida, y el valor 0 una bajada.

Tendencia pronosticada: tendencia (T) que hemos predicho mediante R Studio.

Diferencia abs: diferencia entre la tendencia real y pronosticada ($S_t - T_t$).

Prob. fallo u.: probabilidad de fallar en la tendencia (Pf), medida en forma unitaria.

Prob. acierto (%): probabilidad de acierto (1 - Pf), medida en forma porcentual.

RECM: Raíz del Error Cuadrático Medio, con fórmula ($\sqrt{\Sigma E^2/\text{número de casos}}$)

BBVA

Suavizado de Holt

Longitud de serie: 1 año

Tabla 10. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	6,833						
15/05/18	6,729	6,837	0,108	0,012	0	1	1
16/05/18	6,62	6,838	0,218	0,047	0	1	1
17/05/18	6,707	6,839	0,132	0,017	1	1	0
18/05/18	6,58	6,840	0,260	0,067	0	1	1
21/05/18	6,508	6,841	0,333	0,111	0	1	1
Suma				0,254		Prob. fallo u.	0,8
				RECM	0,2255	Prob. acierto(%)	20

Fuente: elaboración propia

Suavizado de Holt

Longitud de serie: 6 meses

Tabla 11. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs.
14/05/18	6,833						
15/05/18	6,729	6,848	0,119	0,014	0	1	1
16/05/18	6,62	6,863	0,243	0,059	0	1	1
17/05/18	6,707	6,878	0,171	0,029	1	1	0
18/05/18	6,58	6,894	0,314	0,098	0	1	1
21/05/18	6,508	6,909	0,401	0,160	0	1	1
Suma				0,362		Prob. fallo u.	0,8
				RECM	0,2689	Prob. acierto(%)	20

Fuente: elaboración propia



ARMA (1,0)

Longitud de serie: 1 año

Tabla 12. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	5,537						
15/05/18	5,526	5,502	-0,024	0,001	0	0	0
16/05/18	5,46	5,498	0,038	0,001	0	0	0
17/05/18	5,363	5,494	0,131	0,017	0	0	0
18/05/18	5,206	5,491	0,285	0,081	0	0	0
21/05/18	5,169	5,488	0,319	0,102	0	0	0
Suma				0,202		Prob. fallo u.	0
				RECM	0,2010	Prob. acierto (%)	100

Fuente: elaboración propia

ARMA (1,0)

Longitud de serie: 6 meses

Tabla 13. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	5,537						
15/05/18	5,526	5,502	-0,024	0,001	0	0	0
16/05/18	5,46	5,497	0,037	0,001	0	0	0
17/05/18	5,363	5,494	0,131	0,017	0	0	0
18/05/18	5,206	5,490	0,284	0,081	0	0	0
21/05/18	5,169	5,487	0,318	0,101	0	0	0
Suma				0,201		Prob. fallo u.	0
				RECM	0,2003	Prob. acierto (%)	100

Fuente: elaboración propia



Suavizado de Holt

Longitud de serie: 6 meses

Tabla 14. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	5,537						
15/05/18	5,526	5,521	-0,005	0,000	0	0	0
16/05/18	5,46	5,533	0,073	0,005	0	1	1
17/05/18	5,363	5,546	0,183	0,033	0	1	1
18/05/18	5,206	5,558	0,352	0,124	0	1	1
21/05/18	5,169	5,571	0,402	0,162	0	1	1
Suma				0,324		Prob. fallo u.	0,8
				RECM	0,2547	Prob. acierto(%)	20

Fuente: elaboración propia

Suavizado de Brown

Longitud de serie: 6 meses

Tabla 15. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	5,537						
15/05/18	5,526	5,507	-0,019	0,000	0	-	-
16/05/18	5,46	5,507	0,047	0,002	0	-	-
17/05/18	5,363	5,507	0,144	0,021	0	-	-
18/05/18	5,206	5,507	0,301	0,090	0	-	-
21/05/18	5,169	5,507	0,338	0,114	0	-	-
Suma				0,228		Prob. fallo u.	-
				RECM	0,2134	Prob. acierto(%)	-

Fuente: elaboración propia



Suavizado de Holt

Longitud de serie: 1 año

Tabla 16. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	6,53						
15/05/18	6,55	6,539	-0,011	0,000	1	1	0
16/05/18	6,51	6,551	0,041	0,002	0	1	1
17/05/18	6,56	6,562	0,002	0,000	1	1	0
18/05/18	6,6	6,574	-0,026	0,001	1	1	0
21/05/18	6,54	6,585	0,045	0,002	0	1	1
Suma				0,005		Prob. fallo u.	0,4
				RECM	0,0301	Prob. Acierto (%)	60

Fuente: elaboración propia

Suavizado de Brown

Longitud de serie: 1 año

Tabla 17. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	6,53						
15/05/18	6,55	6,527	-0,023	0,001	1	-	-
16/05/18	6,51	6,527	0,017	0,000	0	-	-
17/05/18	6,56	6,527	-0,033	0,001	1	-	-
18/05/18	6,6	6,527	-0,073	0,005	1	-	-
21/05/18	6,54	6,527	-0,013	0,000	0	-	-
Suma				0,007		Prob. fallo u.	-
				RECM	0,0384	Prob. Acierto (%)	-

Fuente: elaboración propia

Suavizado de Holt

Longitud de serie: 6 meses

Tabla 18. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	6,53						
15/05/18	6,55	6,539	-0,011	0,000	1	1	0
16/05/18	6,51	6,549	0,039	0,002	0	1	1
17/05/18	6,56	6,560	0,000	0,000	1	1	0
18/05/18	6,6	6,570	-0,030	0,001	1	1	0
21/05/18	6,54	6,580	0,040	0,002	0	1	1
Suma				0,004		Prob. fallo u.	0,4
				RECM	0,0289	Prob. acierto (%)	60

Fuente: elaboración propia

ARMA (1,0)

Longitud de serie: 6 meses

Tabla 19. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	6,53						
15/05/18	6,55	6,521	-0,029	0,001	1	0	1
16/05/18	6,51	6,514	0,004	0,000	0	0	0
17/05/18	6,56	6,507	-0,053	0,003	1	0	1
18/05/18	6,6	6,501	-0,099	0,010	1	0	1
21/05/18	6,54	6,494	-0,046	0,002	0	0	0
Suma				0,016		Prob. fallo u.	0,6
				RECM	0,0559	Prob. Acierto (%)	40

Fuente: elaboración propia



ARIMA (0,1,1)

Longitud de serie: 6 meses

Tabla 20. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	6,53						
15/05/18	6,55	6,521	-0,029	0,001	1	0	1
16/05/18	6,51	6,514	0,004	0,000	0	0	0
17/05/18	6,56	6,507	-0,053	0,003	1	0	1
18/05/18	6,6	6,501	-0,099	0,010	1	0	1
21/05/18	6,54	6,494	-0,046	0,002	0	0	0
Suma				0,016		Prob. fallo u.	0,6
				RECM	0,0559	Prob. acierto (%)	40

Fuente: elaboración propia

Suavizado de Holt

Longitud de serie: 1 año

Tabla 21. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	16,67						
15/05/18	16,74	16,743	0,003	0,000	1	1	0
16/05/18	16,75	16,814	0,064	0,004	1	1	0
17/05/18	17,09	16,884	-0,206	0,042	1	1	0
18/05/18	17,07	16,954	-0,116	0,013	0	1	1
21/05/18	17,14	17,023	-0,117	0,014	1	1	0
Suma				0,074		Prob. fallo u.	0,2
				RECM	0,1214	Prob. acierto (%)	80

Fuente: elaboración propia

ARIMA (0,1,0)

Longitud de serie: 1 año

Tabla 22. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	16,67						
15/05/18	16,74	16,670	-0,070	0,005	1	-	-
16/05/18	16,75	16,670	-0,080	0,006	1	-	-
17/05/18	17,09	16,670	-0,420	0,176	1	-	-
18/05/18	17,07	16,670	-0,400	0,160	0	-	-
21/05/18	17,14	16,670	-0,470	0,221	1	-	-
Suma				0,569		Prob. fallo u.	-
				RECM	0,3372	Prob. acierto (%)	-

Fuente: elaboración propia

Suavizado de Holt

Longitud de serie: 6 meses

Tabla 23. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	16,67						
15/05/18	16,74	16,769	0,029	0,001	1	1	0
16/05/18	16,75	16,858	0,108	0,012	1	1	0
17/05/18	17,09	16,947	-0,143	0,020	1	1	0
18/05/18	17,07	17,036	-0,034	0,001	0	1	1
21/05/18	17,14	17,124	-0,016	0,000	1	1	0
Suma				0,034		Prob. fallo u.	0,2
				RECM	0,0830	Prob. acierto (%)	80

Fuente: elaboración propia

Suavizado de Brown

Longitud de serie: 6 meses

Tabla 24. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	16,67						
15/05/18	16,74	16,668	-0,072	0,005	1	-	-
16/05/18	16,75	16,668	-0,082	0,007	1	-	-
17/05/18	17,09	16,668	-0,422	0,178	1	-	-
18/05/18	17,07	16,668	-0,402	0,161	0	-	-
21/05/18	17,14	16,668	-0,472	0,222	1	-	-
Suma				0,573		Prob. fallo u.	-
				RECM	0,3386	Prob. acierto (%)	-

Fuente: elaboración propia

Suavizado de Holt

Longitud de serie: 1 año

Tabla 25. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	13,94						
15/05/18	14,03	13,926	-0,104	0,011	1	0	1
16/05/18	13,79	13,913	0,123	0,015	0	0	0
17/05/18	14,2	13,899	-0,301	0,091	1	0	1
18/05/18	14,25	13,885	-0,365	0,133	1	0	1
21/05/18	14,345	13,871	-0,474	0,224	1	0	1
Suma				0,474		Prob. fallo u.	0,8
				RECM	0,3079	Prob. acierto (%)	20

Fuente: elaboración propia

ARIMA (0,1,0)

Longitud de serie: 1 año

Tabla 26. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	13,94						
15/05/18	14,03	14,040	0,010	0,000	1	1	0
16/05/18	13,79	14,140	0,350	0,123	0	1	1
17/05/18	14,2	14,241	0,041	0,002	1	1	0
18/05/18	14,25	14,341	0,091	0,008	1	1	0
21/05/18	14,345	14,442	0,097	0,009	1	1	0
Suma				0,142		Prob. fallo u.	0,2
				RECM	0,1688	Prob. acierto(%)	80

Fuente: elaboración propia

Suavizado de Holt

Longitud de serie: 6 meses

Tabla 27. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	13,94						
15/05/18	14,03	13,958	-0,072	0,005	1	1	0
16/05/18	13,79	13,977	0,187	0,035	0	1	1
17/05/18	14,2	13,995	-0,205	0,042	1	1	0
18/05/18	14,25	14,013	-0,237	0,056	1	1	0
21/05/18	14,345	14,031	-0,314	0,099	1	1	0
Suma				0,237		Prob. fallo u.	0,2
				RECM	0,2177	Prob. acierto (%)	80

Fuente: elaboración propia



Suavizado de Holt

Longitud de serie: 1 año

Tabla 28. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	30,386						
15/05/18	30,112	30,439	0,327	0,107	0	1	1
16/05/18	30,436	30,493	0,057	0,003	1	1	0
17/05/18	30,6	30,548	-0,052	0,003	1	1	0
18/05/18	30,2	30,603	0,403	0,162	0	1	1
21/05/18	29,57	30,658	1,088	1,184	0	1	1
Suma				1,459		Prob. fallo u.	0,6
				RECM	0,5402	Prob. acierto (%)	40

Fuente: elaboración propia

ARIMA (0,1,0)

Longitud de serie: 1 año

Tabla 29. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	30,386						
15/05/18	30,112	30,601	0,489	0,239	0	1	1
16/05/18	30,436	30,818	0,382	0,146	1	1	0
17/05/18	30,6	31,035	0,435	0,189	1	1	0
18/05/18	30,2	31,253	1,053	1,109	0	1	1
21/05/18	29,57	31,472	1,902	3,618	0	1	1
Suma				5,300		Prob. fallo u.	0,6
				RECM	1,0296	Prob. acierto (%)	40

Fuente: elaboración propia



Suavizado de Holt

Longitud de serie: 6 meses

Tabla 30. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	30,386						
15/05/18	30,112	30,314	0,202	0,041	0	0	0
16/05/18	30,436	30,534	0,098	0,010	1	1	0
17/05/18	30,6	30,762	0,162	0,026	1	1	0
18/05/18	30,2	30,997	0,797	0,634	0	1	1
21/05/18	29,57	31,239	1,669	2,784	0	1	1
Suma				3,495		Prob. fallo u.	0,4
				RECM	0,8361	Prob. acierto (%)	60

Fuente: elaboración propia

ARIMA (0,1,0)

Longitud de serie: 6 meses

Tabla 31. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	30,386						
15/05/18	30,112	30,607	0,495	0,245	0	1	1
16/05/18	30,436	30,834	0,398	0,159	1	1	0
17/05/18	30,6	31,069	0,469	0,220	1	1	0
18/05/18	30,2	31,311	1,111	1,234	0	1	1
21/05/18	29,57	31,561	1,991	3,964	0	1	1
Suma				5,821		Prob. fallo u.	0,6
				RECM	1,0790	Prob. acierto(%)	40

Fuente: elaboración propia



Suavizado de Brown

Longitud de serie: 6 meses

Tabla 32. Resultados

Fecha	Valor real	Valor pronosticado	Error	Error cuadrático	Tendencia real	Tendencia pron.	Diferencia abs
14/05/18	30,386						
15/05/18	30,112	30,350	0,238	0,057	0	-	-
16/05/18	30,436	30,350	-0,086	0,007	1	-	-
17/05/18	30,6	30,350	-0,250	0,062	1	-	-
18/05/18	30,2	30,350	0,150	0,023	0	-	-
21/05/18	29,57	30,350	0,780	0,609	0	-	-
Suma				0,758		Prob. fallo u.	-
				RECM	0,3894	Prob. acierto (%)	-

Fuente: elaboración propia

5.2. INTERPRETACIÓN DE LAS TABLAS DE RESULTADOS

Para una mayor comprensión de las tablas de resultados que hemos obtenido en R Studio, se procederá a interpretar y analizar algunos aspectos destacables de los datos resultantes.

Las tablas del punto anterior hacen referencia a los cálculos que hemos tenido que llevar a cabo en el proceso de predicción y de elección del mejor resultado. Este proceso de predicción consta de dos partes diferenciadas y dependientes, en cierta medida. La primera parte del proceso se basa en obtener los valores exactos pronosticados a partir de los modelos escogidos, es decir, calcular la cotización exacta que el título tendría en los cinco días posteriores. La segunda parte estaría enfocada a interpretar las tendencias de las series y comparar la real con la pronosticada.

En la primera parte del proceso, la que calcula el valor exacto de la cotización futura, la interpretación es objetiva y clara, ya que basta con comparar el dato real con el pronosticado para discernir si hemos acertado o no. En este sentido, nuestros modelos no se ajustan a los datos exactos reales en ninguno de los casos, es decir, el precio real y el pronosticado no son iguales ni en ninguno de los modelos ni en ninguna de las empresas. Desde el principio del trabajo, incluso desde antes de plantear el estudio, sabíamos con cierta seguridad que esto no iba a ocurrir, ya que es prácticamente imposible predecir el valor exacto de un título corporativo, al menos con los modelos tradicionales que hemos tratado en este estudio. Estas fluctuaciones de precios que queremos predecir están sujetas a multitud de variables condicionales, a especuladores y a simples cambios de los entornos competitivos, por lo que hacen tarea casi imposible acertar con el valor exacto de la cotización futura.

No obstante, desde el otro enfoque de los resultados, el que hace referencia a las tendencias descritas por las series, los datos no son tan desalentadores. En este caso tratamos de predecir la tendencia que seguirá la serie de la cotización en los cinco días siguientes, y posteriormente obtener una probabilidad de acierto de la tendencia predicha. Esta interpretación de resultados es más subjetiva que la anterior, ya que los datos obtenidos, pueden no ajustarse a los datos intradía, pero sí que pueden ajustarse a la tendencia semanal de la serie sin tanto problema. De las 24 tablas que forman

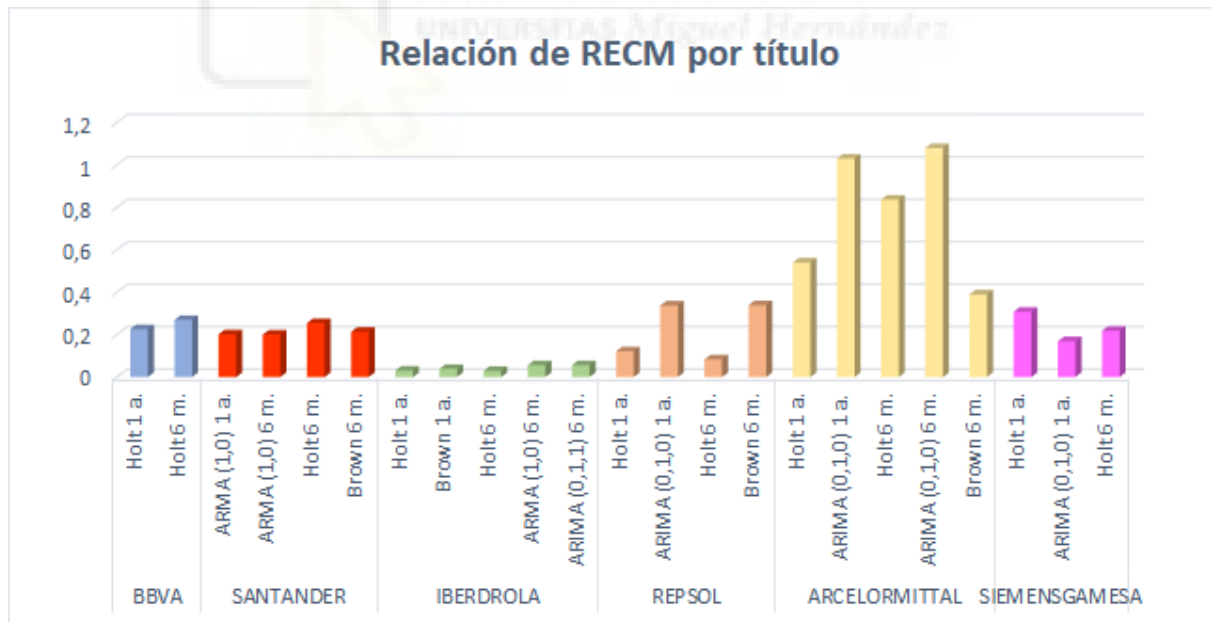
nuestros resultados, en 9 de ellas coinciden la tendencia pronosticada y la real, son menos de la mitad de los casos, pero tenemos que tener en cuenta que desde el enfoque del dato exacto, como hemos dicho, no se acierta en ningún caso.

Otro enfoque dependiente del primero, el de dato exacto, es la comparación de las raíces de los errores cuadráticos medios (RECM). Este valor lo que nos indica es la distancia entre el dato real y el pronosticado, cuanto mayor sea, mayor será la desviación de nuestras predicciones.

5.3. RESUMEN GRÁFICO DE LOS RESULTADOS

Con el objetivo de visualizar mejor el contenido de las tablas de resultados anteriores, se han creado dos gráficos de barras donde se reflejan la relación de las raíces de los errores cuadráticos medios (*gráfico 1*), y las probabilidades de acierto para cada modelo (*gráfico 2*).

Gráfico 1. Relación de RECM por título



Fuente: elaboración propia, datos extraídos de RStudio.

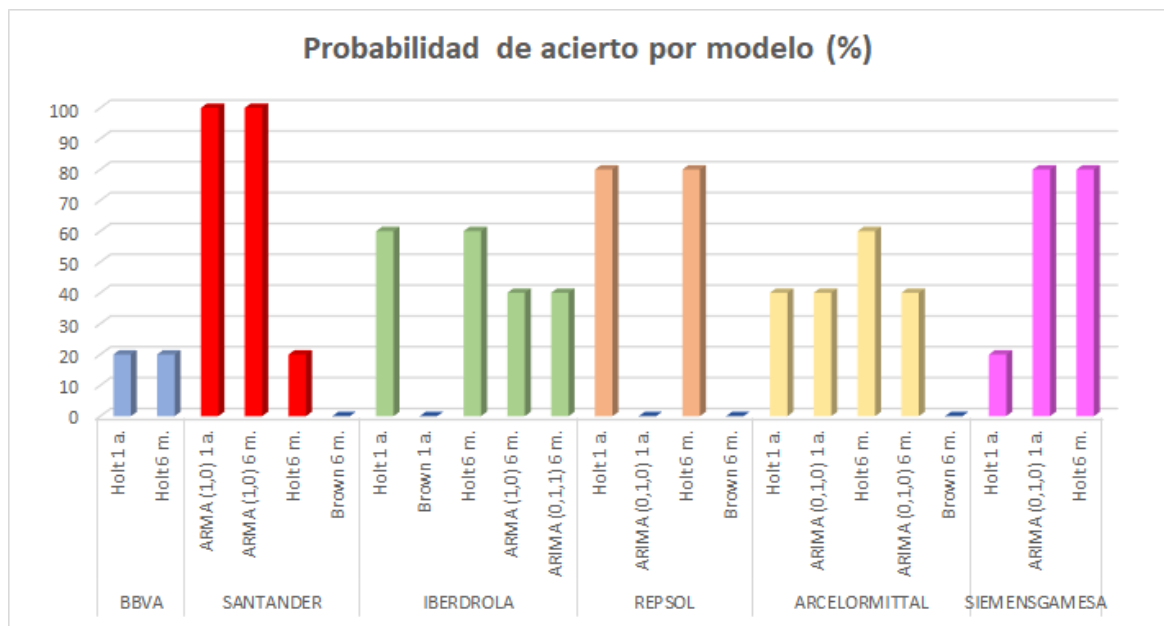
En este gráfico, la variable protagonista es el RECM. Como hemos comentado anteriormente, las raíces de los errores cuadráticos medios de los títulos hacen referencia a la desviación de sus pronósticos con respecto a los datos reales. Es una

medida utilizada en estadística, en el campo de análisis de series económicas, para comparar dos o más modelos mediante sus errores de predicción. De esta manera, un modelo con un RECM menor, implicará que la predicción realizada en dicho modelo se ajusta más a los datos reales.

Una vez contextualizada la variable, se puede observar que la empresa ArcelorMittal posee los RECM mayores, es decir, que sus predicciones se ajustan menos a la realidad que las de las otras empresas. Esto no quiere decir que no acierte en la tendencia, simplemente los datos exactos pronosticados son más lejanos a los reales que en las demás. En el caso de que pronosticara bien la tendencia, sí que sería el RECM una variable clave a la hora de eliminar modelos, no obstante, se debe tener en cuenta que el valor del RECM está influido por el valor en sí de la cotización, es decir, un título con una cotización más elevada puede tener fluctuaciones más pronunciadas en términos absolutos que una cotización más baja. Por ejemplo, si un título cotiza a 100 euros la acción, sus fluctuaciones serán más pronunciadas en términos absolutos que las de un título cotizado a 1 euro la acción, una subida del 2% no es la misma en los dos casos.

A continuación, se detallará el gráfico de las probabilidades.

Gráfico 2. Probabilidad de acierto por modelo (%)



Fuente: elaboración propia, datos extraídos de RStudio.

Por otra parte, tenemos este gráfico donde se muestran las probabilidades de acierto de cada uno de los modelos en cada una de las empresas. Esta tasa de acierto se ha obtenido ponderando con 1=subida y 0=bajada los movimientos de las cotizaciones dentro de los cinco días pronosticados. Por ejemplo, si la tendencia pronosticada acierta en cuatro días de cinco posibles, el porcentaje de acierto de este modelo en la tendencia intrasemanal sería del 80%.

Como podemos observar, las tasas de acierto más altas las encontramos en dos de los modelos desarrollados para el Banco Santander, con un 100% de acierto. Le siguen de cerca dos modelos correspondientes a Repsol y dos de SiemensGamesa. Se puede apreciar también en el gráfico que algunos de los modelos no presentan puntuación, esto se debe a que las predicciones realizadas en dicho modelo eran iguales para los cinco días, es decir, tendencia nula, y por supuesto, errónea.

Esta interpretación de la tendencia pronosticada es un eje fundamental del estudio, ya que, de las dos vertientes de resultados (datos exactos y tendencias), es la única que acierta sensiblemente con los datos reales. En el caso anterior de los datos exactos pronosticados, vimos que sería casi imposible acertar con los datos reales, debido a ello, la predicción de la tendencia cobra importancia, al ser la predicción más fiable de nuestro estudio. Esto quiere decir que, en el caso de una operación bursátil de compra o de venta, podríamos sacar rentabilidad ejecutando posiciones basándonos en el pronóstico de la tendencia de nuestras series. Cabe destacar que, aunque una cotización real fluctúe de manera intradía o de un día para otro, lo importante de este enfoque es conseguir pronosticar con éxito la tendencia a lo largo de los cinco días, es decir, si el valor a día cinco es mayor o menor que el del día uno.

6. FASE 4: COMPROBACIÓN Y ANÁLISIS DE ERRORES

Una vez concluido el proceso de predicción, vamos a realizar un análisis detallado de los errores que han cometido los modelos en sus resultados. Como hemos procedido anteriormente, consideraremos los errores en dos vertientes diferentes y dependientes, la

predicción de datos exactos y la predicción de la tendencia seguida. Es necesario hacer esta distinción debido a que no se tratan de errores de la misma naturaleza y, aunque dependen el uno del otro, no tienen por qué comportarse de la misma manera.

A la hora de examinar los distintos errores, debemos tener en cuenta que algunos de nuestros modelos utilizados son útiles básicamente para predecir tendencias, no datos exactos, como es el caso de los suavizados exponencial es, tanto el de Holt (suavizado doble) como el de Brown (suavizado triple). Esto quiere decir que a estos modelos no se les exige que predigan con exactitud los datos requeridos, sino que pronostiquen la tendencia que seguirá dicha serie en el horizonte temporal que marquemos. En cambio, los modelos utilizados de la familia *ARIMA*, sí que son susceptibles de exigir precisión, o por lo menos, mayor ajuste.

Seguidamente, pasamos al análisis de las dos vertientes de errores, donde intentaremos discernir el porqué de los errores y dar una posible solución al respecto. Además, introduciremos el criterio de la *beta* de los títulos relacionándolo, si se puede, con los resultados obtenidos.

6.1. VERTIENTE DE ANÁLISIS DE ERRORES EN DATOS EXACTOS

En primer lugar, enfocaremos nuestra atención en las desviaciones numéricas de nuestros pronósticos con respecto a los datos reales. En esta vertiente, consideramos los datos exactos que nos han proporcionado los modelos y los comparamos con los datos que se han ido sucediendo en los días posteriores, de esta manera conocemos si los modelos utilizados han propuesto valores ajustados a la realidad, o si por el contrario han predicho de manera lineal y poco ajustada.

Si analizamos los errores de los modelos utilizados para el Banco BBVA, podemos observar que van aumentando día a día, lo que significa que los datos pronosticados se van alejando cada vez más de los datos reales. Esto se produce cuando el modelo considera una tendencia lineal en el sentido opuesto a la tendencia real. En la otra cara de la moneda del sector financiero de nuestro estudio, encontramos al Banco Santander. Sus errores se comportan de igual manera que los de BBVA, sin embargo, presentan una diferencia clara y muy destacable. Esta diferencia es que los errores de los modelos

de Santander, aunque tienden a aumentar día a día como los de BBVA, no lo hacen de la misma manera. Estos errores se separan cada vez más de los datos reales, pero la tendencia lineal que siguen es en el mismo sentido que la tendencia real de la serie, por lo tanto, aunque se vayan separando pronóstico y realidad, lo hacen en la misma dirección, por lo que los convierte en modelos más fiables que los anteriores de BBVA.

Dentro del sector de la energía y los combustibles, encontramos a Iberdrola. Sus errores se intercalan entre positivos y negativos, lo que significa que los datos pronosticados se cruzan con los reales, es decir, que pretenden un mayor ajuste. Las distancias entre los datos predichos y los reales de Iberdrola son las más bajas de todas las empresas. A pesar de ello, los valores se van alejando cada vez más de los reales, comportándose de forma similar que los modelos referidos a BBVA. Por lo tanto, no debemos interpretar como buenos los resultados obtenidos sin tener en cuenta la dirección que siguen los errores. Por otro lado, tenemos los modelos desarrollados para la empresa Repsol, del mismo sector. Esta vez, aunque tenga una relación de errores mayor que su compañero de sector, la dirección de estos errores sí que es la adecuada a la serie real. Si consideráramos la fiabilidad de los modelos únicamente por la variable RECM, podríamos indicar que son mejores los modelos de Iberdrola, sin embargo, hemos aprendido durante el estudio que un menor RECM no quiere decir selección segura, hemos de comprobar la dirección de los errores para interpretar ampliamente los modelos y conseguir una conclusión fiable.

Por último, en el barco de la industria y la construcción, tenemos a ArcelorMittal y a SiemensGamesa. Los modelos desarrollados para la primera presentan los errores más altos de todos los modelos realizados, esto se debe en parte a la cotización significativamente más elevada que la del resto de títulos. El comportamiento de estos errores es similar a los correspondientes a BBVA e Iberdrola, esto es, los errores se desvían cada vez más de la trayectoria real de la serie. Por el contrario, en el caso de SiemensGamesa, aunque los errores también se distancian cada vez más de la serie real, la dirección de éstos es la óptima para la predicción, es decir, la misma que la de la serie real, salvo en un modelo, en el que los errores se desvían en la dirección contraria a los reales.

En cuanto a las *betas* de los títulos, podemos afirmar que no tienen influencia en los pronósticos de las series, puesto que el comportamiento de los errores ha sido totalmente independiente de la variabilidad de las empresas con respecto al mercado. Esto lo podemos apreciar en que los errores de las empresas con *betas* superiores a 1, como Repsol y ArcelorMittal, se han comportado de manera totalmente diferente uno de otro, lo mismo sucede con las empresas con *betas* inferiores a 1, Iberdrola y SiemensGamesa, comportándose de manera opuesta en la tendencia de sus errores y, por lo tanto, en su diagnóstico.

6.2. VERTIENTE DE ANÁLISIS DE TENDENCIA PRONOSTICADA

Una vez comentada la vertiente de los errores de datos exactos, y explicada la situación de cada empresa con respecto a éstos, vamos a cambiar de plano para fijarnos esta vez en el pronóstico de la tendencia de cada serie. En el punto anterior se ha hecho una breve referencia a esta vertiente. Cuando decíamos que los errores, aunque se iban separando, iban en la misma dirección que la serie real, nos referíamos precisamente a esta predicción de la tendencia a la que ahora prestamos atención. Esta predicción se ha convertido durante el estudio en la más fiable de las dos, ya que es la única que acierta con los resultados pronosticados y, por lo tanto, la manera más fiable de interpretar los datos obtenidos.

Observando el *gráfico 2* y las tablas de resultados, podemos darnos cuenta de que las empresas con una probabilidad de acierto de más del 70%, es decir, con 4/5 o 5/5 días acertados, son las que sus modelos han dado con la predicción correcta de la tendencia seguida de la serie. Estas empresas y sus modelos son: Santander, con *ARMA* (1,0) a un año y *ARMA* (1,0) a 6 meses; Repsol, con el suavizado de Holt a un año y el suavizado de Brown a 6 meses; y SiemensGamesa, con *ARIMA* (0,1,0) a un año y suavizado de Holt a 6 meses.

Esta vertiente de análisis de errores nos ha dado mejores noticias que la anterior, por lo tanto, a la hora de llevar a la práctica este estudio, consideraríamos únicamente estas empresas con estos modelos concretos para un asesoramiento financiero en materia de inversión. Los modelos con RECM más bajos que no tuvieran aprobada esta vertiente,

aún con menos errores, no serían adecuados para nuestras recomendaciones, ya que no estarían siguiendo la misma dirección que el mercado.

7. INTERPRETACIÓN DE LOS RESULTADOS OBTENIDOS

A partir de este momento, el proceso de predicción queda finalizado. En el apartado anterior hemos dado una pequeña interpretación de los errores que han cometido nuestros modelos, así como de la importancia de la buena predicción de la tendencia seguida. En esta parte del estudio nos dedicaremos a relacionar los resultados obtenidos con la situación actual en los mercados financieros y, en concreto, con la situación de las empresas que hemos estudiado.

El mercado financiero español se ha mostrado recuperado de la crisis de 2008, y las empresas españolas han visto cómo sus beneficios han vuelto a situarse a niveles de antes de la crisis. Esto se produce en medio de una tormenta política en nuestro país, lo que altera en cierta medida los movimientos naturales de las empresas autóctonas. Sin embargo, las compañías hacen el caso justo a la política y se centran en sus rentabilidades y eficiencia.

En el caso de las empresas del sector financiero escogidas para este estudio, las dos compañías se tratan de bancos, con su parte comercial y de inversión. Este sector bancario es el que más ha visto revolcarse sus cifras de negocio, al estar sus beneficios muy ligados a tipos de intereses procedentes de Bruselas y a la disposición del público por la inversión, factor este último que se ha visto perjudicado por la pasada crisis. El producto principal de estos bancos son los préstamos de todo tipo, productos que se ven afectados por estos tipos de interés que ponen en desde Europa. Al disminuir estos tipos de interés a niveles de en torno a 0, los márgenes de beneficio de los bancos a la hora de realizar préstamos han caído notablemente, a la vez que aumentaban sus niveles de morosidad y activos tóxicos. Actualmente ya respiran aliviados y sus beneficios están volviendo a niveles pre-crisis, esto lo ha hecho posible una inminente subida de tipos de interés y a un saneamiento de activos tóxicos. Además, al haberse reducido sus

márgenes de beneficios, han optado por mayor rotación de clientes, aumentando su número de clientes mediante fusiones, como la de Santander y Popular, o campañas de fidelización.

Relacionando lo expuesto con nuestras predicciones, vemos que mientras BBVA se ha iniciado en una senda bajista, Santander es estable en su tendencia alcista, lo que ha hecho más fácil la labor de predicción por parte de los modelos estadísticos, no ocurriendo lo mismo con los modelos desarrollados para BBVA.

Por otro lado, el sector de la energía y combustible, representado por Repsol e Iberdrola, está en auge desde el final de la 2010. Iberdrola por su parte ha remontado su cotización desde 2012 hasta situarse a niveles de antes de la crisis. Esto ha sido posible gracias a la cantidad de nuevos contratos conseguidos por la compañía, especialmente en latinoamérica, donde se le considera una compañía de calidad y muy competitiva. Su compañero en nuestro estudio, Repsol, no ha sufrido la misma suerte. Su tendencia es negativa desde el inicio de la crisis hasta ahora, aún con repuntes positivos la compañía no consigue crear una tendencia alcista estable durante años, pese a que ha recuperado niveles pre-crisis también.

En cuanto a las predicciones de estos títulos, observamos que los modelos desarrollados para Iberdrola, pese a tener menor sus errores no aciertan en la tendencia de la serie real, por lo que no es óptima la predicción. Sin embargo, los modelos referidos a Repsol sí que han sabido leer la tendencia negativa que posee el título, producido quizás por su última etapa bajista.

Por último, si nos referimos al campo de la industria y la construcción, tenemos a ArcelorMittal y a SiemensGamesa, los cuales han seguido una trayectoria parecida después de la crisis, pero con distinto final. Ambas compañías sufrieron fuertes caídas en 2008 con el comienzo de la crisis financiera, en el caso de ArcelorMittal se dejó más de dos tercios de su cotización y SiemensGamesa paso de cotizaciones en torno a 25 euros la acción a cifras en torno a 5 euros la acción. Esta situación cambiaría totalmente a partir de 2012 para SiemensGamesa, que recuperó cifras de entorno a los 20 euros por acción hasta hoy en día. No ocurre lo mismo con ArcelorMittal, la cual no ha vuelto a sus niveles de cotización pre-crisis y en su tendencia no se muestran indicios de una

posible rotura al alza. Pese a ser la principal compañía siderúrgica mundial, no ha conseguido remontar sus beneficios y se ha visto estancada en los últimos tres años, a diferencia de su compañera de sector.

En lo referente a nuestros modelos, en el caso de ArcelorMittal, sí que pronosticaban esta senda alcista que podría haberse producido, sin embargo, como hemos comentado, la compañía sigue estancada y mientras que nuestros modelos proponían crecimiento, la cotización no ha hecho lo propio y ha caído durante la semana que hemos estudiado. Por otra parte, Siemens Gamesa sí que se ha encaminado hacia la trayectoria creciente en su cotización. A pesar de haber sufrido una caída considerable a mediados de 2017, sus mínimos actuales están por encima de los máximos post-crisis, lo que es señal de una subida en el futuro.

8. LIMITACIONES DEL ESTUDIO

Desde el comienzo del Análisis Técnico, se ha reprochado que esta disciplina utiliza únicamente datos pasados para elaborar sus cálculos y sus estudios, por lo que movimientos como las roturas de tendencia o los cambios en los entornos producían alteraciones que eran imposibles de predecir. Sin embargo, con el éxito de modelos estadísticos y otros métodos de análisis técnico, como el de Martin Armstrong mencionado en la introducción, su uso lejos de decrecer aumenta. Es importante indicar que el éxito de un modelo estadístico está muy ligado al objeto subyacente que predice, es decir, no es lo mismo predecir una tendencia cíclica del paro, por ejemplo, que las fluctuaciones de divisas. Algunas series son más susceptibles de predicción que otras, y esa característica es lo primero que se debe averiguar a la hora de decidir el campo de estudio a analizar.

Las roturas de tendencia o los cambios en los entornos competitivos, por suerte o por desgracia, son muy frecuentes y afectan a los modelos que usamos para desarrollar las predicciones. Nosotros hemos utilizado bases de datos de un año y de 6 meses, considerando que una mayor amplitud de la serie no sería adecuada y desvirtuaría los

resultados. Al no trabajar con series con componentes cíclicos, estas roturas o cambios son imposibles de predecir, puesto que, en visión de los modelos, son totalmente aleatorias. Si pusiéramos el caso de una persona que quisiera invertir en valores como los de nuestro estudio, esta persona sí que sería capaz de intentar adelantarse a los acontecimientos y aprovechar las fluctuaciones, ya que puede disponer de información cualitativa de cada empresa y conocer los acontecimientos que les afecten. Por el contrario, un modelo estadístico solo se basa en la serie que se le viene dada, sin considerar ningún aspecto de la vida cotidiana de la empresa.

Si nos referimos al presente estudio, los objetos que hemos desarrollado para su predicción son empresas del IBEX-35. Las acciones corporativas son objetos especiales a la hora de realizar sus predicciones. La razón por la que son especiales es que su comportamiento puede pasar de estable a dinámico en cuestión de horas en un mismo día. Con esta perspectiva podemos interpretar mejor los resultados que hemos obtenido en nuestros modelos, ya que esto es el motivo de la mayoría de errores cometidos en las predicciones.

Los modelos estadísticos que hemos utilizado han predicho las tendencias con el supuesto de que las cotizaciones seguían una tendencia estable en su última etapa de variabilidad, es decir, que mantendrían la dirección en los días venideros. Esto, como hemos comprobado, no ha sido así. Las cotizaciones han girado la tendencia en las tres empresas que poseen menos de un 70% de probabilidad de acierto, debido a esto, las predicciones de tendencia no han sido correctas, ya que los modelos seguían su propia tendencia estable mientras que las cotizaciones reales se daban la vuelta.

En resumen, las limitaciones de nuestros modelos son esos movimientos normales y constantes que sufren las series económicas de las cotizaciones de acciones corporativas. En el caso de que estos movimientos fueran cíclicos, bastaría con escoger una serie de tiempo lo suficientemente amplia como para contener en ella esos factores que se repiten en el tiempo, si este fuera el caso los modelos estadísticos habrían predicho a la perfección los movimientos futuros. Esta afirmación es posible gracias al estudio de más series temporales en otros trabajos realizados anteriormente. Como esto no ha ocurrido, no hemos logrado el nivel de eficacia que pretendíamos con los modelos seleccionados.

9. CONCLUSIONES RESULTANTES DEL TRABAJO

Una vez considerados todos los aspectos relacionados con nuestro pequeño experimento, podemos sacar en claro algunas conclusiones. Como bien explicamos al principio del documento, nuestro objetivo principal era conocer si los modelos estadísticos planteados serían capaces de predecir con cierta fiabilidad el comportamiento futuro a corto plazo de acciones cotizadas en el IBEX-35. Destacamos que el objetivo es comprobar si los modelos sirven o no sirven para las predicciones, no acertar en la totalidad de nuestros resultados. Desde un primer momento, en la fase de preparación del estudio, sabíamos que los movimientos de las cotizaciones de las empresas no iban a seguir tendencias claras, y que sus movimientos estarían ligados a acontecimientos y resultados internos de cada compañía. No obstante, nos planteábamos la posibilidad de que a corto plazo sí que fueran capaces de predecir con cierta facilidad la tendencia seguida de las series. El resultado que hemos obtenido no ha sido del todo desalentador. Si bien es cierto que la tasa de acierto en cuanto a datos exactos ha sido cero, no ha ocurrido lo mismo con las predicciones de las tendencias, llegando incluso a predecir con un 100% de acierto la tendencia seguida en dos de los casos desarrollados.

Actualmente, y ha venido siendo así durante casi toda la historia financiera moderna, estos movimientos en los precios de acciones a los que hacemos referencia también son provocados por especuladores bursátiles que llevan a cabo su trabajo comprando acciones baratas para venderlas caras, lo que altera el juego de oferta y demanda soportado por las cotizaciones sin ninguna razón subyacente al propio negocio, ya que “una operación de inversión es aquella que, después de realizar un análisis exhaustivo, promete la seguridad del principal y un adecuado rendimiento. Las operaciones que no satisfacen estos requisitos son especulativas.” (Graham, El inversor inteligente, 2017) Esto quiere decir que, si intentáramos predecir los precios de las acciones únicamente basándonos en datos financieros internos de las compañías, seguramente no llegaríamos a los mismos niveles a los que llega el mercado en realidad, puesto que muchos de los movimientos y roturas de tendencias son provocados por meros rumores, pánico financiero o información sin fundamentos.

A pesar de todas estas causas externas que afectan a nuestras series económicas, nuestros modelos no hacen distinción de las subidas o bajas que se producen, únicamente las considera por igual para elaborar un pronóstico que siga la tendencia normal de la serie estudiada. Esto nos puede llevar a considerar que, si conseguimos una serie con suficientes registros de actividad de cambios de tendencia concatenados, podríamos llegar a predecir casi todas las fluctuaciones venideras. Bien, pues esto no es así. Aún teniendo una serie que contenga toda la información que podríamos necesitar, los entornos competitivos son dinámicos y evolucionan afectando de manera distinta a cada empresa. Las empresas que saben adaptarse a los entornos futuros son las supervivientes del tejido empresarial. Con esto queremos expresar que, aún teniendo esa serie sobreinformada, el nivel del precio de las acciones empresariales puede hundirse o elevarse sin que haya ocurrido previamente en el pasado, lo que sería imposible de predecir para los modelos estadísticos utilizados.

Como conclusión, podemos decir que la predicción de cotizaciones bursátiles del IBEX-35 utilizando modelos estadísticos no podría considerarse fiable a la hora de asesoramiento financiero en materia de inversión, ni tampoco para especulación. Podemos acogernos a los resultados de las predicciones de las tendencias, los cuales no han sido del todo desfavorables, pero estaríamos contradiciendo las interpretaciones y conclusiones que hemos descrito. Aunque hayamos predicho correctamente algunas de las tendencias seguidas, sabemos por lo estudiado que, si estas cotizaciones hubieran sufrido un cambio de tendencia, por el motivo que fuera, nuestros modelos no habrían sido capaces de predecirla, y habrían seguido con sus tendencias preestablecidas. Esto quiere decir, que el resultado de nuestro estudio es negativo por parte de predicciones de los precios de acciones del IBEX-35 mediante modelos estadísticos.

ANEXO I. Pruebas prácticas

En este anexo propuesto llevaremos a cabo dos experimentos con distintas operaciones y procedimientos. En primer lugar, se realizará una hipotética operativa de compraventa de tres valores seleccionados previamente, escogiendo los títulos que han presentado mejores probabilidades de acierto en la tendencia. En segundo lugar, se compararán los distintos títulos en base a la rentabilidad obtenida en cada uno de ellos en el primer día de los pronosticados, se escogerá la acción que más rentabilidad presente y se realizarán pruebas día a día para conocer si el beneficio diario podría superar al semanal.

Experimento 1

Ya finalizado nuestro trabajo, vamos a considerar un supuesto ficticio. En este caso práctico supondremos que hemos decidido comprar acciones de las empresas que nos han dado un porcentaje de acierto en la tendencia mayor a 70%, es decir, las que han acertado en su resultado. Para el desarrollo de este caso no tendremos en cuenta los gastos adicionales de las operaciones de compraventa de acciones que se producen en la vida real, ni la relación del binomio rentabilidad-riesgo, consideraremos únicamente los beneficios o pérdidas obtenidos mediante las operaciones relacionadas a nuestro estudio.

- Partimos de una base de efectivo de 120.000 euros, los cuales se destinan íntegramente a la compra de acciones sugeridas por nuestros modelos.
- La compra se realizaría a partes iguales entre las empresas que nos han dado más de un 70% de probabilidad de acierto en la tendencia, estas empresas serían: Santander, Repsol y Siemens Gamesa (40.000 euros para cada título).

A continuación, mostraremos el precio al inicio de los cinco días y el precio final, resultando como nuestro beneficio la diferencia entre ambos.

Santander → **Precio inicial** = 5,537 | **Precio final** = 5,169

Repsol → **Precio inicial** = 16,67 / **Precio final** = 17,14

Siemens Gamesa → **Precio inicial** =13,94 / **Precio final** = 14,345

Como podemos observar, en el caso de las acciones de Santander, tendríamos que realizar una operación de venta, en vez de compra, ya que el precio sigue una tendencia a la baja.

Obtención del resultado

- Santander $\rightarrow |(40.000*5,537) - (40.000*5,169)| = 14.720$
- Repsol $\rightarrow |(40.000*16,67) - (40.000*17,14)| = 18.800$
- Siemens Gamesa $\rightarrow (40.000*13,94) - (40.000*14,345) = 16.200$

Beneficio = $14.720 + 18.800 + 16.200 = 49.720$ euros

Capital final obtenido = $120.000 + 49.720 = 169.720$ euros

Rentabilidad de la operación (i) $\rightarrow 120.000 (1+i) = 169.720$

$i = 0,4143$

Rentabilidad de la operación 41,43%

Esta rentabilidad obtenida únicamente hace referencia a la rentabilidad simple que resultaría al realizar dichas operaciones. Es una rentabilidad elevada que lleva consigo un riesgo también elevado ya que, como se ha explicado en las limitaciones del estudio, los cambios o roturas de tendencia no estarían pronosticados y por lo tanto estaríamos expuesto a que la cotización no se comporte como se espera, lo cual no sería difícil en un entorno bursátil normal.

Experimento 2

En este caso, se va a proceder a comparar entre las distintas rentabilidades obtenidas en una supuesta operación de compra de acciones en el primer día de pronósticos, y escogeremos la acción que posea la mayor de ellas. Se utilizarán los valores pronosticados resultantes de los modelos estadísticos que hayan dado mayores probabilidades de acierto, y en el caso de que sean iguales estas probabilidades, se escogerá el modelo con menor RECM.

Tabla 33. Comparación entre empresas

Empresa	Valor Real*	Valor pronosticado	Beneficio por acción* €	Beneficio relativo %
Santander	5,537	5,502	-0,035	-0,00640
BBVA	6,833	6,837	0,004	0,00055
Iberdrola	6,53	6,539	0,009	0,00131
Repsol	16,67	16,769	0,099	0,00596
ArcelorMittal	30,386	30,314	-0,072	-0,00238
SiemensGam.	13,94	14,040	0,100	0,00718

Fuente: elaboración propia

*Valor Real: cotización de la acción en el día 0 de los valores reales.

*Beneficio por Acción: cantidad en unidades monetarias de la diferencia de cotizaciones reales y pronosticadas.

Como se puede observar en la Tabla 33, si nos guiáramos por el beneficio relativo por acción, escogeríamos los títulos de SiemensGamesa por poseer mayor rentabilidad en el primer día pronosticado. A continuación, se procederá a calcular su rentabilidad según el día en el que se vendan las acciones compradas desde el día 0 (14/05/18), para un capital invertido de 10.000 euros, y se comparará el beneficio obtenido en el caso de que se retirara el capital invertido pasado un día, o por el contrario se siguiera hasta el final del pronóstico (cinco días).

- Compra de acciones con 10.000 euros, a un precio de 13,94 euros por acción.

$$\text{Número de acciones compradas} = 10.000 / 13,94 = 717,36$$

Se establece la compra de 717 acciones de la empresa SiemensGamesa.

En la siguiente tabla (Tabla 34) se muestra el beneficio dependiendo del día en el que se vendan las acciones. Los valores de la columna “Beneficio por día” hacen referencia al beneficio monetario que se obtendrían si las acciones compradas el día 14/05/18 se vendieran en cada uno de los sucesivos días posteriores, es decir, la diferencia entre la cotización de cada día pronosticado multiplicado por 717 acciones, y la cotización del día 0 multiplicado por 717 acciones.

Tabla 34. Beneficio según día de venta

Fecha	Valor real*	Valor pronosticado	Beneficio por día
14/05/18	13,94		
15/05/18		14,040	71,81
16/05/18		14,140	143,71
17/05/18		14,241	215,73
18/05/18		14,341	287,85
21/05/18		14,442	360,07

Fuente: elaboración propia

*Valor Real: cotización de la acción en el día 0 de los valores reales.

Como se puede comprobar, el beneficio aumenta día a día, lo que significa que la operación óptima consistiría en no retirar el capital invertido hasta el último día pronosticado, para así conseguir el máximo beneficio posible (360,07 euros). Por lo tanto, podemos concluir que el beneficio diario no podría superar al semanal. Este caso se produce de esta manera debido a que las cotizaciones pronosticadas van aumentando de manera lineal en cada uno de los días, por lo que en ningún caso el beneficio de un día podría ser superior al del día siguiente.

BIBLIOGRAFÍA

Gekko. (2017, junio 24). *Blablanegocios.com*. From blablanegocios.com:
<http://blablanegocios.com/sarima-que-es/>

Wikipedia.org. (2018, marzo 19). www.wikipedia.org/Prueba_de_Levene. From
www.wikipedia.org: https://es.wikipedia.org/wiki/Prueba_de_Levene

Graham, B. (2017). *El inversor inteligente*. New York: DEUSTO.

