

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE ESTUDIOS ESTADÍSTICOS

Máster en Minería de Datos e Inteligencia de Negocio



**Estudios de la correlación sectorial del riesgo en la crisis  
financiera de 2008.**

Héctor Jorge Castillejo Ruiz

Tutor: Lorenzo Escot Mangas

Madrid, noviembre de 2017

# Índice

Introducción .....	1
Contexto del estudio.....	2
Objetivos .....	7
Metodología empleada.....	7
Base teórica.....	8
Planificación .....	12
Pasos.....	13
Herramientas.....	14
Desarrollo del trabajo y principales resultados.....	15
Selección de datos.....	15
Catálogo de las series.....	17
Preparación de los datos.....	20
Carga.....	20
Limpieza.....	21
Estimación de modelos preliminares.....	23
Creación de índices sintéticos.....	23
Modelos univariantes (GARCH/apARCH).....	26
Modelos multivariantes (DCC) .....	29
Estimación de modelos predictivos finales .....	30
Estimación para el sector del petróleo y la energía.....	32
Estimación para el sector de materiales básicos, industria y construcción.....	34
Estimación para el sector de bienes de consumo.....	36
Estimación para el sector de servicios de consumo.....	38
Estimación para el sector de servicios financieros e inmobiliarias.....	39
Estimación para el sector de tecnología y telecomunicaciones.....	41
Estimación para las series sectoriales sintéticas.....	43
Conclusiones .....	44
Futuro estudio y aplicaciones.....	45
Bibliografía .....	47
Referencias.....	48
Anexos.....	49
Código R .....	74

## Índice de gráficos y tablas

Gráfico 1 – Serie de los precios a cierre de la cotización del IBEX 35, entre 2007 y 2008, con eventos registrados.....	5
Gráfico 2 – Rendimientos del IBEX 35, entre 2007 y 2008, con eventos registrados.....	6
Gráfico 3 – Rendimientos al cuadrado del IBEX 35, entre 2007 y 2008.....	6
Gráfico 4 – Planificación del proyecto (Diagrama de Gantt).....	12
Tabla 1 – Ranking de Contratación de Renta Variable en el SIBE en 2008. 50 primeras empresas..	16
Tabla 1A – Ranking de Contratación de Renta Variable en el SIBE en 2008. Empresas de la 51 a la 100.....	16
Tabla 2 – Conjunto de valores agrupados en el sector 1 de petróleo y energía.....	18
Tabla 3 – Conjunto de valores agrupados en el sector 2 de materiales básicos, industria y construcción.....	19
Tabla 4 – Conjunto de valores agrupados en el sector 4 de bienes de consumo.....	19
Tabla 5 – Conjunto de valores agrupados en el sector 4 de servicios de consumo.....	19
Tabla 6 – Conjunto de valores agrupados en el sector 5 de servicios financieros e inmobiliaria.....	19
Tabla 7 – Conjunto de valores agrupados en el sector 6 de tecnología y telecomunicaciones.....	19
Gráfico 5 – Serie de cotizaciones a cierre de Sabadell.....	22
Gráfico 6 – Serie de cotizaciones a cierre de Grifols.....	22
Gráfico 7 – Serie de cotizaciones a cierre de Bankinter.....	22
Gráfico 8 – cotizaciones estimadas para las series sintéticas de los sectores.....	24
Gráfico 8A – superposición de las cotizaciones estimadas para las series sintéticas de los sectores..	25
Gráfico 8B – Rendimientos diarios de los sectores.....	25
Gráfico 9 - Número de series que mejor se ajusta a cada modelo, dentro de los mejores modelos seleccionados.....	27
Gráficos 9A,9B,9C,9D,9E y 9F – Estimaciones de la varianza realizadas por los mejores modelos GARCH recogidos en la Tabla AT2.....	28
Tabla 8 – Estadísticos de modelos de prueba DCC.....	30
Tabla 9 - Ejemplo de matriz Sectorial.....	31
Gráfico 10 – Matriz media geométrica de las matrices de correlación estimadas.....	33
Gráfico 11 – Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas.....	33
Gráfico 12 – Matriz media geométrica de las matrices de correlación estimadas para el sector de materiales básicos, industria y construcción.....	34

Gráfico 13 - Clúster de 4 grupos de la matriz, de la media geométrica, de las matrices de correlación estimadas para el sector de materiales básicos, industria y construcción.....	35
Gráfico 14 – Matriz media geométrica de las matrices de correlación estimadas para el sector de bienes de consumo.....	36
Gráfico 15 - Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas para el sector de bienes de consumo.....	37
Gráfico 16 – Matriz media geométrica de las matrices de correlación estimadas para el sector de bienes de consumo.....	38
Gráfico 17 - Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas para el sector de servicios de consumo.....	39
Gráfico 18 – Matriz media geométrica de las matrices de correlación estimadas para el sector de servicios financieros e inmobiliarias.....	40
Gráfico 19 - Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas para el sector de servicios financieros e inmobiliarias.....	41
Gráficos A1 a A20 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.....	54
Gráficos A21 a A40 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.....	55
Gráficos A41 a 60 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.....	56
Gráficos A61 a 80 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.....	57
Gráficos A81 a 100 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.....	58
Gráficos A101 a 120 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.....	59
Gráficos A121 a 140 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.....	60
Gráficos A141 a 160 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.....	61
Gráficos A161 a 180 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.....	62
Gráficos A181 a 200 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.....	63

Gráficos A201 a 210 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	64
Gráficos A211 a 220 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	65
Gráficos A221 a 230 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	66
Gráficos A231 a 240 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	67
Gráficos A241 a 250 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	68
Gráficos A251 a 260 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	69
Gráficos A261 a 270 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	70
Gráficos A271 a 280 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	71
Gráficos A281 a 290 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	72
Gráficos A291 a 300 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.....	73

# Introducción

Este estudio estadístico se realiza como Trabajo de Fin de Máster, del Máster en Minería de datos e Inteligencia de Negocios, enfocado en el campo del riesgo financiero.

Como caso concreto de estudio se plantea la evaluación de la correlación en el riesgo de precios, medido a través de la varianza, de diferentes valores del Mercado Continuo Español, durante la conocida coloquialmente como “crisis del 2008”.

Englobado dentro del riesgo financiero, el riesgo de mercado es la volatilidad que se da en la serie de rendimientos para una serie de cotizaciones. Un estimador insesgado, pero no eficiente, de dicha serie de rendimientos es la serie de rendimientos al cuadrado, que es indicador del riesgo y como se ve en el estudio lejos de ser constante está autocorrelacionado. Esta característica de los rendimientos al cuadrado hace posible el uso de modelos GARCH univariantes para la estimación de la varianza.

Para la consecución del proyecto se ha seleccionando un conjunto de datos suficientemente amplio, compuesto de series temporales, depurando los mismos de manera cuidadosa, para conseguir una calidad razonable y, posteriormente, aplicar sobre éstos algunos modelos de la familia de los Autorregresivos de Heterocedasticidad Condicional, ARCH (Robert F. Engle, 1982).

Estos modelos, que permiten representar la volatilidad, tienen gran variedad de aplicaciones que aprovechan las diferentes características de las series. En este caso, de entre todos ellos, se seleccionan dos tipos, los apARCH y GARCH, cuyo comportamiento se prueba sobre el conjunto muestral aquí expuesto, permitiendo así la selección de los que mejor se comportan. Posteriormente, aprovechando los modelos previamente elegidos, se estiman modelos GARCH multinomiales mediante aproximaciones DCC, para estimar las covarianzas en los movimientos de todas las series temporales financieras analizadas.

La estrategia empírica se desarrolla en dos bloques principales, con sendos niveles de agrupación lógica de las series. En un primer bloque se agrupan las series por sectores y se

realiza una estimación de GARCH multivariante, para todos los valores de un mismo sector, que permite analizar las covarianzas en el riesgo dentro de un mismo sector o intra sector. En un segundo bloque de la estrategia, se resumen los sectores en índices sintéticos generados expresamente para cada uno de ellos, tomando como base todas las series que lo componen. Después de crear las nuevas series sectoriales, se estima otro GARCH multinomial, para todas ellas, para analizar a su vez las covarianzas entre sectores.

## ***Contexto del estudio***

Históricamente la sociedad ha sufrido grandes crisis de distintos tipos, algunas de ellas intrínsecamente financieras, entre las que se encuentra la que servirá de trasfondo en el planteamiento aquí expuesto, que a posteriori tuvo un impacto sobre la economía real [1].

Los orígenes de las crisis son diversos y suelen tener como base la caída del precio de algún subyacente, factor o sector. Por ejemplo, una de las primeras crisis que quedaron documentadas fue la ocurrida en Holanda en el siglo XVII, cuyo principal subyacente fueron los tulipanes. Posteriormente y de manera recurrente el sistema ha tenido otras muchas crisis, en algunos casos de mayor impacto como pudo ser la de la Gran Depresión en los años 20. Provocada por la caída de precios en el sector agrícola, que derivó en un crack bursátil conocido como “crack del 29” de grandes dimensiones y efectos que perduraron durante muchos años, llegando incluso a desestabilizar el equilibrio geopolítico mundial.

“Los historiadores están de acuerdo en afirmar que la Gran Depresión, que se inició en 1929 y se prolongó durante los años treinta, es una de las épocas fundamentales de la historia de los Estados Unidos. Pero no sólo afectó a los EE.UU. [...] se hizo sentir en prácticamente todas las economías de mercado del mundo [...]. Los historiadores también se ponen de acuerdo en afirmar que fue una de las razones principales para el desarrollo del nazismo que más tarde condujo a la II Guerra Mundial.”

Paul Krugman y Robin Wells. Macroeconomía: Introducción, 2006.

Si bien no es el objetivo de este documento el estudio de todas las crisis acaecidas durante la historia universal, queda clara la importancia de las mismas y su eventual estudio

como herramienta para tratar de localizar su origen, desarrollo y la evaluación del impacto que tienen en la sociedad. Dicho estudio lleva en algunos casos a tomar medidas institucionales como por ejemplo la implantación de los Acuerdos de Basilea III para la regulación bancaria [2] en respuesta a la crisis de 2008 que, como ya se ha indicado, será foco principal aquí.

Se analizan algunos aspectos de la crisis de 2008 ya que, para el estudio del riesgo, es un periodo de cierta anormalidad, en el que se dieron volatilidades extremas por los acontecimientos ocurridos, lo que la hace una época interesante desde el punto de vista de la varianza, siendo identificada además entre las crisis de mayor impacto. A día de hoy es la crisis acaecida más reciente, dándose en una época en la que las tecnologías de la información estaban plenamente implantadas, lo que permite mayor facilidad y fiabilidad a la hora de obtener datos, sobre todo para un conjunto tan amplio de valores.

Bajo el foco de dicha crisis es difícil determinar unos límites temporales con exactitud. Las crisis suelen ir acompañadas de hechos o noticias, que provocan reacciones en los mercados que están asociados con alteraciones en la varianza a su vez asociada al riesgo, pero al darse muchos y con un impacto diverso es difícil distinguir los más relevantes. Es complejo por tanto localizar cuál es el desencadenante exacto, que fijaría el inicio por ejemplo, y según se buscan referencias o estudios se definen diferentes plazos dependiendo del punto de vista del autor/a.

Durante 2007 se detectaron eventos de cierta importancia, como el comportamiento extraño de grandes bancos en los Estados Unidos de América, por ejemplo Citigroup, Bear Sterns y American Home Mortgage. La intervención de algunos de los principales actores financieros institucionales, a nivel mundial, como la Reserva Federal de los Estados Unidos de América, el Banco Central Europeo o el Banco de Japón desde agosto del mismo año en adelante, que decidieron tomar diversas medidas como la inyección de fondos. El cenit de todos los hechos ocurridos se podría localizar en la famosa declaración de quiebra, de Lehman Brothers, el día 15 de septiembre de 2008. Por tanto se fijan los años de 2007 y 2008 como plazo para nuestros datos, que con alta probabilidad se encontrarán en el periodo en el que se desarrolló dicha crisis.

En España, que es el país en el que cotizaban los valores seleccionados en ese momento, se dan dos factores adicionales que pudieron afectar: el efecto de una burbuja inmobiliaria y unas elecciones generales en marzo de 2008. Se tienen ambos en cuenta, el primero de ellos a la hora de analizar los resultados y el segundo a la hora de seleccionar los datos, tomándolos a partir de marzo de ese año.

De manera ilustrativa podemos ver, en el Gráfico 1, la serie de cotización del IBEX-35, que aglutina los 35 valores con mayor capitalización incluidos en nuestro conjunto muestral. En el Gráfico 2, los rendimientos para la misma serie, por último, en el Gráfico 3 los rendimientos al cuadrado que, como se comentó al principio, es un indicador de riesgo y se aprecia que no es constante. En los gráficos se puede ver una alteración en el comportamiento de la serie, muy evidente, con alteraciones más pronunciadas en 2008. Por ejemplo, en el primero de los tres se observa una tendencia algo “alcista”, con sesgo ascendente, hasta finales de 2007. Desde ese momento hay una inversión en la tendencia, que pasa a ser claramente bajista, descendente, hasta casi finales de ese mismo año. Se marcan, en los tres gráficos, ciertos hechos o eventos para identificar visualmente el comportamiento de las tres series en el momento que se dieron. La lista incluida es la siguiente:

#### 2007

1. 18 de abril: Citigroup declara pérdidas de 5.000 millones de dólares.
2. 17 de septiembre: la Reserva Federal de EE.UU. recorta medio punto los tipos de interés.
3. 6 de diciembre: el Presidente de los EE.UU., George W. Bush, anuncia un plan para aliviar la crisis hipotecaria.

#### 2008

4. 21 de enero: el IBEX tiene la mayor caída de su historia, los principales índices mundiales sufren caídas también.
5. 22 de mayo: el banco suizo UBS lanza una operación para cubrir parte de sus millonarias pérdidas (35.000 millones de dólares).
6. 4-7 de julio: el día 4 el Banco Central Europeo sube los tipos al 4,2%, tres días después las bolsas europeas se desploman por las dudas del mercado subprime.

7. 15 de septiembre: Lehman Brothers se declara en quiebra.
8. 10 de octubre: el IBEX sufre la mayor caída de su historia.
9. 12 de octubre: la Unión Europea cierra un acuerdo para garantizar los préstamos entre bancos.
10. 13 de octubre: el IBEX tiene la mayor subida de su historia.

Las caídas en el IBEX, Gráfico 1, se reflejan también en los rendimientos negativos del Gráfico 2, en el que se aprecia que los diferentes eventos representados en los gráficos se traducen en grandes oscilaciones de la tasa de rendimiento, esto es en grandes volatilidades o varianzas, que pueden medirse inicialmente por el cuadrado del los rendimientos, Gráfico 3. En este gráfico de rendimientos diarios del IBEX, Gráfico 2, puede observarse que la volatilidad de los rendimientos (véase también el gráfico3), es decir el riesgo de los rendimientos del ibex, lejos de ser estacionario o constante presenta comportamientos agrupados, es decir periodos de relativa estabilidad sin grandes oscilaciones, y periodos con una alta volatilidad o alto riesgo. Este comportamiento del riesgo es un indicio suficiente para explorar modelos GARCH que traten de capturar dicha recurrencia en la volatilidad de los mercados.

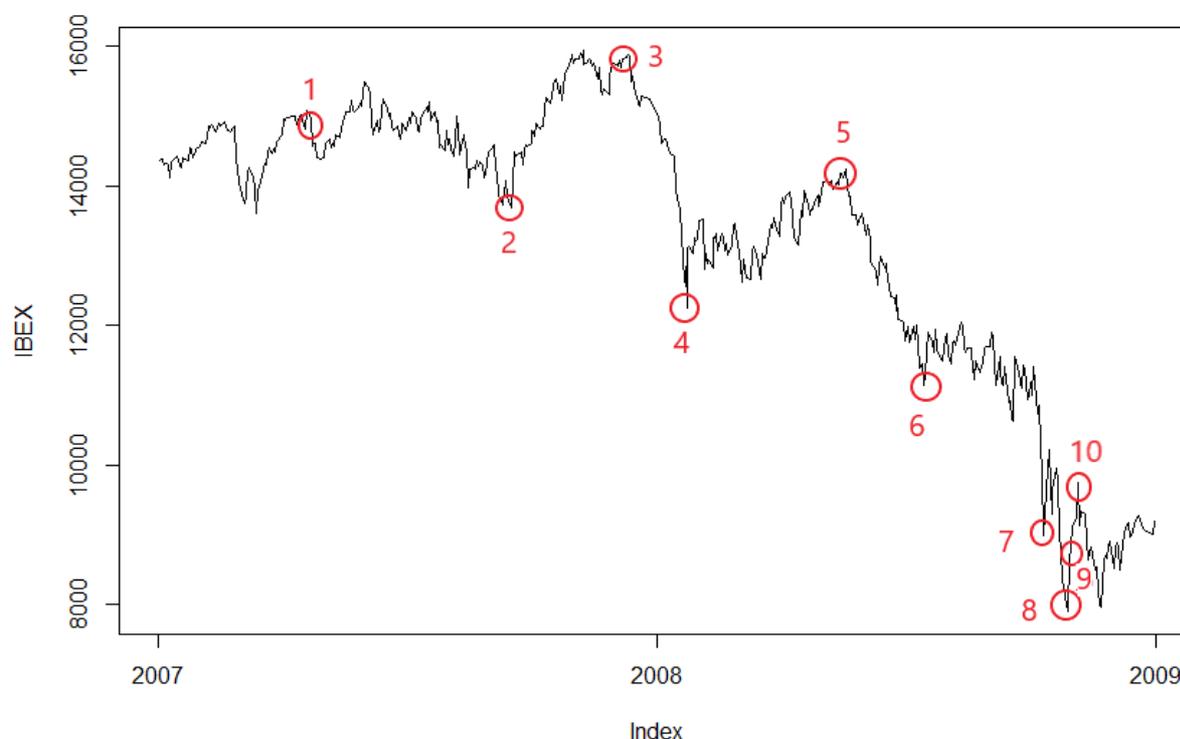


Gráfico 1 – Serie de los precios a cierre de la cotización del IBEX 35, entre 2007 y 2008, con eventos registrados.

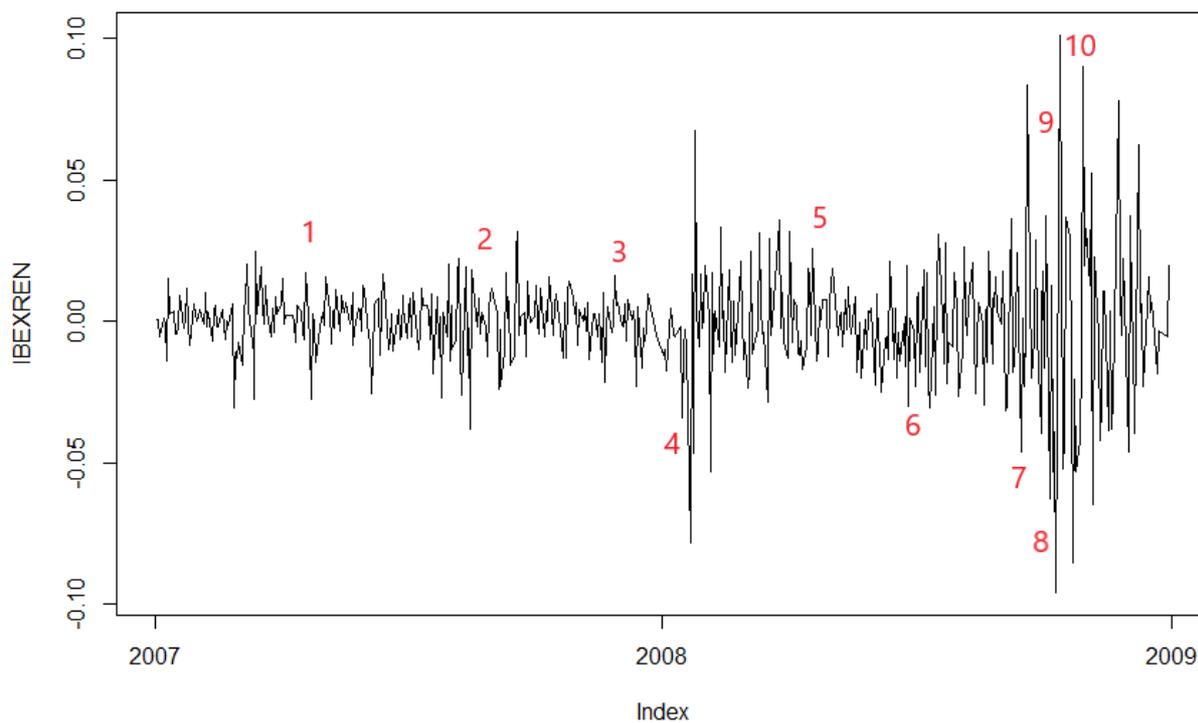


Gráfico 2 – Rendimientos del IBEX 35, entre 2007 y 2008, con eventos registrados.

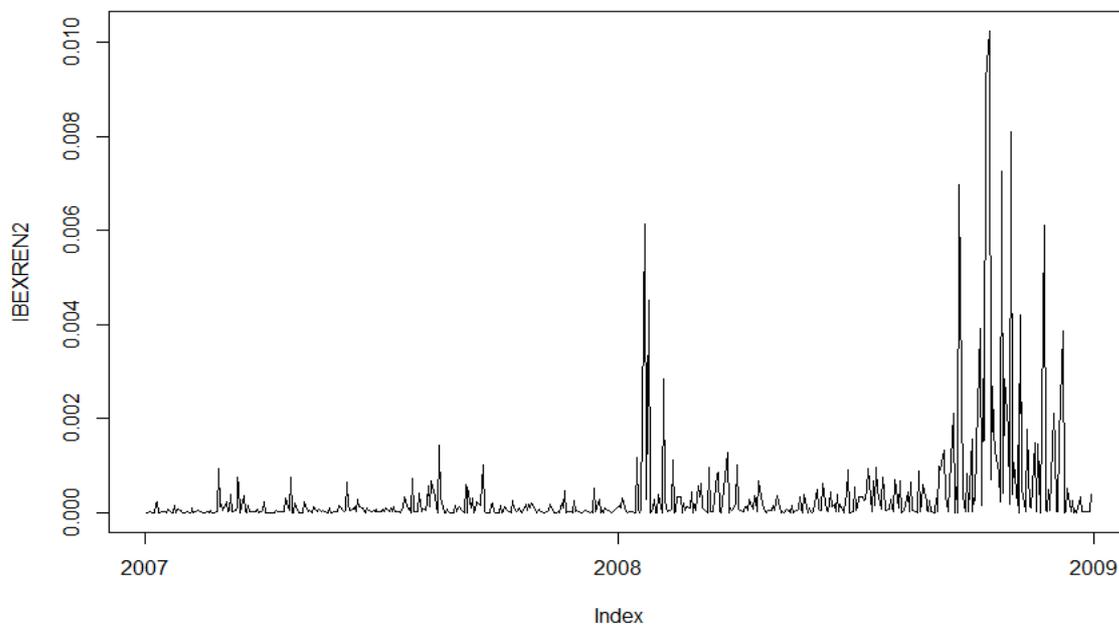


Gráfico 3 – Rendimientos al cuadrado del IBEX 35, entre 2007 y 2008.

## **Objetivos**

Organizando las series antes citadas, en sectores, se proponen dos objetivos: el principal tiene como fin el estudio de la correlación en el riesgo de precios medido a través de su varianza, para el conjunto de los valores que conforman cada uno de los sectores, obteniendo así una correlación intrasectorial, dentro de un mismo sector, del riesgo de precios.

De manera secundaria se tiene como objetivo, que complementa al anterior, el evaluar si existe algún tipo de correlación, del riesgo de precios medido también a través de sus varianzas, de los sectores representados por índices sintéticos obteniendo una correlación intersectorial.

Para ello se obtiene la matriz de varianzas y covarianzas del riesgo de precios, medido por los modelos, para extraer su correlación. En el primer objetivo la correlación es interna de cada los sector y en la segunda externa, entre los diferentes los sectores, todo ello en el periodo seleccionado de 2008.

Por último se evalúa la agrupación en clústers, de la evolución en el tiempo de correlaciones de los riesgos, para analizar en qué medida el riesgo sufrió un efecto contagio a través de los diferentes valores y sectores económicos o si se trató de fenómenos aislados e independientes.

## **Metodología empleada**

Para el desarrollo se han planteado diversas fases, que se desglosan en pasos y tareas concretos y que tienen como base teórica la que se expone en este apartado.

## **Base teórica**

Para el estudio se utilizan modelos GARCH univariantes, que son generalización de los procesos ARCH, junto con modelos apARCH, realizando una comparativa de ambos. Éstos permiten modelizar las varianzas de las series, en este caso financieras. Adicionalmente para calcular las correlaciones de las estimaciones de la varianza, se utilizan modelos DCC, que utilizan para su aplicación los modelos GARCH multivariantes que son una generalización de los univariantes antes citados. Para la selección de los mejores modelos se utiliza el criterio de información de Akaike y por último se aplican clústers jerárquicos de tipo ward a los resultados obtenidos de los modelos DCC. Todos ellos quedan descritos brevemente en las siguientes líneas.

Se definen los procesos conocidos como ARCH presentándolos el autor de la siguiente manera:

“Los modelos econométricos tradicionales asumen una predicción de varianza de un solo periodo constante. Para generalizar esta asunción implausible, una nueva clase de procesos estocásticos llamados procesos autorregresivos de heterocedasticidad condicional (ARCH) son presentados en este *paper*.”

Robert F. Engle 1982.

Según se indica en el documento de definición de los modelos ARCH, hasta ese momento los intentos de medir la varianza asumían que ésta era constante en el tiempo. La nueva definición de procesos aplicados a series temporales es algo más flexible, en este caso se documentó con un caso práctico basado en la serie de inflación del Reino Unido entre los años 1958 y 1977, con periodos de alta volatilidad, tanto caídas como subidas bruscas.

El proceso ARCH(q), o ARCH de orden q, se puede expresar de la siguiente manera (Ruey S. Tsay, 2006):

$$a_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_m a_{t-m}^2$$

Donde  $a_t$  es la serie de rendimientos de un valor y  $\varepsilon_t$  es un proceso con media 0 y varianza 1. La varianza, momento de grado 2, depende por un lado de un  $\alpha$  cero que tiene que ser mayor estrictamente que 0 y por otro lado de los rendimientos al cuadrado de los m pasos anteriores dependiendo cada uno de ellos de otra componente  $\alpha$ .

Posteriormente se trata de generalizar los procesos anteriormente citados y define entonces los procesos autorregresivos generalizados de heterocedasticidad condicional, GARCH (Bollerslev, 1986). La búsqueda de nuevos modelos se plantea ya que “aunque el modelo ARCH es sencillo, en ocasiones requiere de muchos parámetros para describir la volatilidad de un proceso de rendimientos de un valor” (Ruey S. Tsay, 2006).

La ecuación que representa un modelo GARCH(p,q) se puede expresar de la siguiente manera (Ruey S. Tsay, 2006):

$$a_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$$

La ecuación de la serie coincide con el modelo ARCH salvo la parte que define la varianza, que en este caso depende de una componente  $\beta$  y las varianzas de los pasos anteriores.

Como evolución de estos modelos aparece más adelante el modelo apARCH (Ding, Granger y Engle, 1993) que incluyen un nuevo parámetro  $\gamma$  que permite que los efectos del pasado sean asimétricos, esto es que no sean iguales cuando son positivos que cuando sean negativos. Cubren momentos de diverso grado a través de la parametrización de su  $\delta$ , cuya expresión puede ser la siguiente (Christian Kleiber y Achi Zeileis, 2008):

$$\sigma_t^\delta = w + \sum_{i=1}^p \alpha_i (|a_{t-i}| - \gamma_i a_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta$$

Para este modelo se pueden parametrizar los anteriores ARCH y GARCH parametrizando para el primero con  $\delta=2$ ,  $\gamma=0$  y  $\beta=0$  y el segundo  $\delta=2$  y  $\gamma=0$ .

Una vez estimada la varianza de las series, a través de modelos GARCH, se detecta una nueva necesidad de estimar modelos multinomiales. Utilizando un número finito de series temporales, dentro del sector financiero, se plantean nuevos modelos multivariantes que permiten medir la volatilidad de manera conjunta. “Las volatilidades mutivariantes tienen muchas aplicaciones financieras importantes. Ellas juegan un rol importante en la selección de carteras y la distribución de activos, y se pueden utilizar para computar el valor en riesgo (VER) de una posición financiera compuesta de múltiples activos” (Ruey S. Tsay, 2006).

Uno de los primeros modelos planteados es el modelo diagonal VEC (Bollerslev, Engle y Wooldridge, 1988), donde generalizando el uso del modelo de medias móviles con ponderación exponencial (EWMA) se expresa de la siguiente manera (Ruey S. Tsay, 2006):

$$\Sigma_t = A_0 + \sum_{i=1}^m A_i \odot (a_{t-i} a'_{t-i}) + \sum_{j=1}^s B_j \odot \Sigma_{t-j}$$

Donde  $p$  y  $q$  son enteros no negativos  $A_i$  y  $B_j$  son matrices simétricas y  $\odot$  denota el producto Hadamard, que es el producto elemento a elemento. Se denota como VEC(m,s) o DVEC(m,s) y es la generalización de un GARCH univariante para que sea multinomial.  $\Sigma_t$  es la matriz de varianzas y covarianzas condicionales. Existen diferentes enfoques de dicha matriz dependiendo del modelo que se utilice para la estimación de la misma.

En un planteamiento multivariante, se define un nuevo modelo que es combinación no lineal de los procesos GARCH univariantes conocido como modelo de Correlaciones Condicionales Constantes, CCC (Bollerslev, 1990). Basándose en las estimaciones de las varianzas, de un cierto conjunto de series expresadas por un conjunto de modelos GARCH, se plantea la matriz de varianzas y covarianzas condicionales  $\Sigma_t$ . Ésta se descompone, en dos matrices diagonales y una matriz de correlación  $\rho$ , usando los coeficientes de correlación condicional y las varianzas de  $a_t$  de la siguiente manera:

$$\Sigma_t \equiv [\sigma_{ij,t}] = D_t \rho D_t$$

Donde  $\rho$  es la matriz de correlaciones condicionales de  $a_t$ , que en este caso se asume como constante, y  $D_t$  es una matriz diagonal  $k \times k$  que contiene las desviaciones condicionales estándar de los elementos de  $a_t$ .

En una ampliación del CCC se plantea el modelo de Correlaciones Condicionales Dinámicas, DCC (Engle, 2002), que plantea un único matiz que consiste en asumir la matriz de relaciones  $\rho$  como variable en el tiempo. Se define con la siguiente ecuación (Ruey S. Tsay, 2006):

$$\Sigma_t \equiv [\sigma_{ij,t}] = D_t \rho_t D_t$$

El DCC se aplica en dos fases: una primera de ajuste de modelos, que incluye la estimación de modelos GARCH o similares, que mejor se ajustan a las series. Una segunda fase estima la matriz de varianzas y covarianzas haciendo uso de los modelos seleccionados, y ésta al poder descomponerse, como se indica anteriormente, permite extraer la matriz  $\rho_t$  de correlación que es la que se toma como referencia en este estudio.

Para la selección de los mejores modelos se utiliza el criterio de información de Akaike, AIC (Hirotugu Akaike, 1983), que tiene en cuenta el número de parámetros y la función de verosimilitud y que refleja un mejor modelo cuanto más pequeño sea éste.

Una vez estimadas las matrices de correlaciones entre las diferentes series de rendimientos con el modelo DCC, se aplica un clúster jerárquico de tipo ward (Joe H. Ward, Jr., 1963) que utiliza un criterio de varianza mínima para estimar los niveles de clúster, aplicando para ello el algoritmo de Lance-Williams [3]. En este caso para poder calcular el clúster se realiza la siguiente operativa: se estima  $\rho_t$  a través del modelo DCC, es decir una matriz de correlación para cada instante  $t$ . A continuación, al tener diversas matrices resultantes, para poder hacer el clúster es necesario resumirlas, por lo que se calcula la media geométrica del conjunto de matrices de correlación, como se explica más adelante.

## Planificación

El estudio se planifica en diferentes fases, con el objetivo de seguir un proceso completo y coherente. Las fases, cuyo cronograma se puede ver en el Gráfico 4, comienzan por una primera, de estudio y documentación, en la que se determinan los datos y modelos a utilizar, se categorizan los valores por sectores y se recaba el máximo de información posible para el correcto desarrollo del estudio.

Una segunda fase, en la que se obtienen y prepararán los datos que son utilizados posteriormente con los modelos. A continuación en la tercera fase se llevan a cabo estimaciones iniciales, para evaluar el rendimiento de los diferentes modelos, con el objetivo de detectar los de mejor comportamiento. Se realizan comparativas de modelos GARCH y apARCH.

En una cuarta fase se realizan las estimaciones finales, con los mejores modelos recogidos en la fase anterior, realizando primero un estudio interno de los sectores para evaluar la correlación de las medidas de las varianzas de los valores, mediante los modelos DCC de cada uno de los sectores. En un segundo paso se realizará ese mismo estudio sobre series sintéticas, que se generan previamente, de los sectores.

Los resultados quedarán agrupados de manera lógica, explicados y plasmados en este informe, que será sintetizado en la presentación, preparada en una quinta fase, que se utilizará como defensa del proyecto.

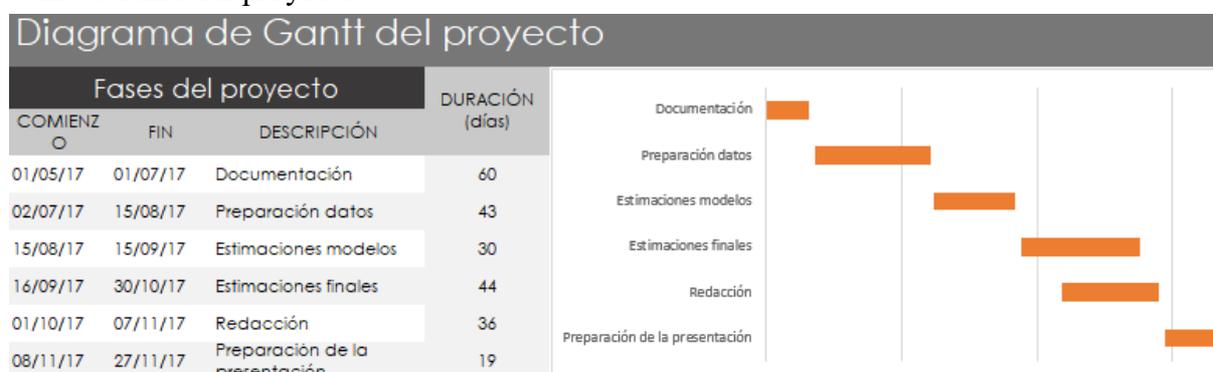


Gráfico 4 – Planificación del proyecto (Diagrama de Gantt).

## **Pasos**

A continuación se detallan los pasos en los que se plantea el desarrollo, tenidos en cuenta en la planificación antes descrita.

- Estudio y documentación: se hace acopio de detalles, de mercados y bolsas españoles, para la selección de las series, su categorización y la recopilación de datos asociados a las mismas, como por ejemplo hechos relevantes. Se localizan los diferentes orígenes de los datos, en los que se encuentren disponibles las series seleccionadas.
- Preparación: se realiza una LT (carga y transformación).
  - Carga: se realiza la carga de datos desde sistemas origen.
  - Transformación: se realiza una primera transformación, mediante una tanda de limpieza para obtener series con calidad suficiente para trabajar, por ejemplo eliminando anomalías y *missings*.
- Estimaciones previas de modelos:
  - Creación de series sintéticas: se analizan las series, realizando una segunda tanda de transformaciones, necesarias para su manejo como, por ejemplo, se agrupan las series en sectores y se generan series sintéticas, que sirvan de indicadores, para cada uno de los sectores identificados.
  - Estimaciones: a continuación se realiza una prueba de modelos GARH y apARCH, que permite crear una tabla comparativa para elegir así los modelos que mejor comportamiento tienen con las series. Además se realiza una prueba de rendimiento con modelos DCC para ver su comportamiento general.
- Estimaciones finales y obtención de resultados: se realizan estimaciones finales de los mejores modelos identificados en las fases anteriores, analizando los resultados.
- Valoración y conclusiones: se valoran los resultados obtenidos y se detallan las conclusiones globales del proyecto.

## ***Herramientas***

Para la realización del tratamiento de los datos, realización de pruebas y estudio de modelos se elige el software R, incluyendo el propio lenguaje de programación y el entorno Rstudio para el manejo de los datos y funciones. Como parte de este software se incluye un conjunto de librerías de entre las que se utilizan las siguientes en este proyecto:

Tseries, paquete que permite realizar diversas operaciones sobre series temporales por ejemplo se usa para la carga de datos, desde Yahoo, a través de la función `get.hist.quotes`.

Corrplot, orientado al análisis de matrices de correlación, se usa en el estudio tanto visualizaciones de éstas como para aplicar clústers jerárquicos.

Zoo permite crear y manejar estructuras orientadas al uso de series temporales de tipo zoo, que entre otras ventajas tiene la de organizar en listas ordenadas los datos, siguiendo en este caso un orden diario, por lo que se puede hacer referencias a días concretos o extraer rangos de fechas o submuestras, lo que facilita su análisis.

Readr permite la carga de archivos csv y el formateo parcial de los datos que se cargan.

Quantmod es una librería de modelado cuantitativo financiero, se utiliza para aplicar operaciones sobre todos los elementos de una serie financiera, en este caso la media geométrica, por medio de la función `apply.daily`.

Fgarch y rugarch son librerías para aplicación de modelos Garch, con sus diversas variantes. fgarch se utiliza en las fases iniciales de pruebas de modelos, posteriormente se usa rugarch ya que las estimaciones se pueden reaprovechar para aplicar directamente el modelo DCC.

Hclust, permite realizar diversos tipos de clústers jerárquicos, entre otros el tipo ward que se utiliza en este estudio.

Para la redacción del informe y generación de tablas, se utiliza software ofimático dentro de la suite Libre Office, a su vez se utiliza dentro de la suite Office 365 de Microsoft

tanto Power Point para la presentación por motivos de compatibilidad y Excel para la generación del diagrama de Gantt.

Por último se utiliza el software NotePad++, para manipular los archivos csv y la limpieza de textos o valores cuando se pasan de unos formatos a otros, por ejemplo los datos que se reciben por la consola de RStudio.

## **Desarrollo del trabajo y principales resultados**

### ***Selección de datos***

El conjunto de datos utilizado se compone de las series temporales, que recogen los datos de cotización a cierre de mercado, de los cien valores más negociados durante el año 2008. El listado incluye la denominación social junto con el montante total efectivo en millones de euros se puede ver en la Tabla 1 y la 1A , para cada valor, y ha sido obtenido del Sistema de Interconexión Bursátil Español (SIBE) para el año 2008 [4].

Es complejo encontrar un origen de datos homogéneo, ya que cada sistema, empresa o entidad, que registra las cotizaciones, tiende a utilizar técnicas y criterios diferentes para su almacenamiento. Ésto lleva a tener series incompletas, con diversas anomalías y con datos determinados de maneras dispares, lo que permite encontrar diferentes datos para una misma serie en un mismo periodo. Por ejemplo algunas series se determinan teniendo en cuenta los hechos relevantes que se dan en el mercado, como por ejemplo ampliaciones de capital, splits, contrasplits y todo tipo de operaciones que pueden afectar en el valor final de cotización y que en ocasiones pero no siempre, son tenidos en cuenta, incrementando o decrementando su valor.

Para evitar en la medida de lo posible las anomalías en los datos, generadas por los problemas descritos, se determina el uso de diferentes orígenes, en este caso: Yahoo, Invertia y en algunos casos Google Finance, para que en caso de existir una carencia o inconsistencia se puedan contrastar desde distintos sitios.

VALOR	EFFECTIVO (MILL)	VALOR	EFFECTIVO (MILL)
1 BANCO SANTANDER, S.A.	277596,38	26 IBERIA, LINEAS AEREAS DE ESPAÑA, S.A.	6056,73
2 TELEFONICA ,S.A.	239510,84	27 BANKINTER,S.A.	5639,88
3 BANCO BILBAO VIZCAYA ARGENTARIA, S.A.	171660,05	28 CRITERIA CAIXACORP, S.A.	5515,6
4 IBERDROLA, S.A.	99339,02	29 INDRA SISTEMAS, S.A., SERIE A	5272,07
5 REPSOL YPF,S.A.	76363,71	30 GESTEVISION TELECNCO,S.A.	4908,84
6 UNION FENOSA,S.A.	34851,07	31 TECNICAS REUNIDAS, S.A.	3466,58
7 BANCO POPULAR ESPAÑOL, S.A.	30349,98	32 INMOBILIARIA COLONIAL, S.A.	3280,82
8 INDUSTRIA DE DISEÑO TEXTIL,S.A."INDITEX"	26516,26	33 SACYR VALLEHERMOSO, S.A.	3053,07
9 GAMESA CORPORACION TECNOLOGICA, S.A.	15638,26	34 SOLARIA ENERGIA Y MEDIO AMBIENTE, S.A.	2656,95
10 ACS,ACTIV. DE CONST.Y SERVICIOS S.A	14538,04	35 PROMOTORA DE INFORMACIONES,S.A.	2339,8
11 RED ELECTRICA CORPORACION, S.A.	12768,05	36 SDAD.GRAL AGUAS DE BARCELONA, S.A.	2197,02
12 ACCIONA,S.A.	12147,07	37 BANCO ESPAÑOL DE CREDITO,S.A.	2195,49
13 GAS NATURAL SDG, S.A.	11833,25	38 ABENGOA, S.A.	2102,36
14 GRUPO FERROVIAL, S.A.	11248,6	39 OBRASCON HUARTE LAIN, S.A.	2097,75
15 IBERDROLA RENOVABLES, S.A.	10655,81	40 SEDA DE BARCELONA, S.A. (LA)	1813,38
16 ABERTIS INFRAESTRUCTURAS, S.A.	10593,36	41 METROVACESA S.A.	1484,92
17 ENAGAS, S.A.	9719,42	42 ZELTIA, S.A.	1470,89
18 BANCO DE SABADELL, S.A.	7649,11	43 ANTENA 3 DE TELEVISION, S.A.	1443,43
19 ACERINOX, S.A.	7060	44 SOL MELIA,S.A.	1197,73
20 MAPFRE, S.A.	6978,89	45 TUBACEX,S.A.	1193
21 ENDESA, SOCIEDAD ANONIMA	6962,42	46 LABORATORIOS ALMIRALL, S.A.	1185,95
22 FOMENTO DE CONSTR. Y CONTRATAS S.A.	6727,59	47 CORPORACION FINANCIERA ALBA, S.A.	1182,54
23 BOLSAS Y MERCADOS ESPAÑOLES	6486,9	48 BANCO PASTOR, S.A.	1162,53
24 GRIFOLS, S.A.	6430,89	49 AVANZIT, S.A.	1158,34
25 CINTRA CONC.DE INF.DE TRANSPORTE,S.A.	6173,15	50 EBRO PULEVA, S.A.	1107,67

Tabla 1 – Ranking de Contratación de Renta Variable en el SIBE en 2008. 50 primeras empresas.

VALOR	EFFECTIVO (MILL)	VALOR	EFFECTIVO (MILL)
51 ZARDOYA OTIS, S.A.	1089,65	76 CAMPOFRIO ALIMENTACION,S.A.	207,26
52 CIA.ESPAÑOLA DE PETROLEOS, S.A.	1072,17	77 CODERE, S.A.	203,63
53 SOS CUETARA, S.A.	1031,02	78 AMPER, S.A.	196,91
54 BANCO DE VALENCIA, S.A.	976,75	79 BARON DE LEY, S.A.	191,57
55 PROSEGUR S.A., CIA. DE SEGURIDAD	952,07	80 PARQUESOL INMOBILIARIA Y PROJ., S.A.	187,39
56 NH HOTELES, S.A.	832,88	81 ITINERE INFRAESTRUCTURAS, S.A.	182,63
57 GRUPO CATALANA DE OCCIDENTE S.A.	782,83	82 CORPORACION DERMOESTETICA, S.A.	176,4
58 AFIRMA GRUPO INMOBILIARIO, S.A.	668,45	83 LABORATORIOS FARMACEUTICOS ROVI, S.A.	173,29
59 CONST.Y AUXILIAR DE FERROCARRILES S.A.	652,95	84 SNIACE	152,73
60 FAES FARMA,S.A.	580,28	85 URBAS GUADAHERMOSA, S.A.	146,08
61 VISCOFAN, S.A.	493,65	86 DURO FELGUERA, S.A.	139,13
62 GRUPO EMPRESARIAL ENCE,S.A.	474,19	87 CLINICA BAVIERA, S.A.	115,23
63 VOCENTO, S.A.	443,27	88 VIDRALA S.A.	111,54
64 GRAL. DE ALQUILER DE MAQUINARIA, S.A.	373,2	89 PESCANOVA, S.A.	110,73
65 REYAL URBIS, S.A.	360,17	90 MONTEBALITO, S.A.	103,61
66 PAPELES Y CARTONES DE EUROPA,S.A.	358,52	91 FLUIDRA, S.A.	103,41
<b>CAJA DE AHORROS DEL MEDITERRANEO</b>	<b>354,7</b>	92 NATRACEUTICAL,S.A.	102,68
67 ERCROS S.A.	320,99	<b>DINAMIA CAPITAL PRIVADO,S.A.</b>	<b>89</b>
68 RENTA CORPORACION REAL ESTATE, S.A.	268,6	93 CIE AUTOMOTIVE, S.A.	86,45
69 CEMENTOS PORTLAND VALDERRIVAS, S.A.	265,41	94 BANCO GUIPUZCOANO, S.A.	85,26
70 VUELING AIRLINES, S.A.	256,13	95 RENTA 4 SERVICIOS DE INVERSION, S.A.	82,07
71 TUBOS REUNIDOS,S.A.	253,25	96 AZKOYEN S.A.	77,27
72 SERVICE POINT SOLUTIONS,S.A.	223,76	97 VERTICE 360 GRADOS, S.A.	75,69
73 FERSA ENERGIAS RENOVABLES, S.A.	212,31	98 TECNOCOM,TELECO. Y ENERGIA,S.	69,39
74 MECALUX, S.A.	209,19	99 URALITA, S.A.	67,27
75 REALIA BUSINESS, S.A.	207,84	100 ESPAÑOLA DEL ZINC, S.A.	64,29

Tabla 1A – Ranking de Contratación de Renta Variable en el SIBE en 2008. Empresas de la 51 a la 100.

Fuente: Bolsas y Mercados Españoles (BME) – Estadística de los Mercados.

## **Catálogo de las series**

Previo a la carga de los datos, se lleva a cabo un pequeño estudio, para recabar información relativa a las series, que sirva para el manejo y organización de los datos. Se recuperan los códigos bursátiles de los valores, en inglés *tickers*, para simplificar la identificación de las series. Se buscan detalles relevantes asociados a éstas, como pueden ser las fusiones, los cambios de denominación, suspensiones de cotización, absorciones o adquisiciones.

A la hora de realizar la búsqueda, empresas que tenían una denominación en el plazo indicado pasan a tener una nueva, tras cambios corporativos, asociando sus datos a la nueva denominación o viceversa, lo que requiere cierta investigación. Se da por ejemplo en el caso de Iberia que actualmente pertenece al grupo IAG, tras su fusión con British Airways. Se realiza un pequeño recuento de estos hechos que se incluyen en la siguiente lista de hechos relevantes.

- Iberia pasa a ser International Airlines Group en 2010.
- Cintra Concesión de Transportes deja de cotizar en diciembre de 2009.
- Unión Fenosa y Gas Natural S.A. Fusionan en Gas Natural Fenosa.
- Critería CaixaCorp S.A. Pasa a denominarse CaixaBank.
- Banco Español de Crédito (Banesto) es absorbido por el Banco de Santander en el año 2013.
- Zeltia pasa a denominarse PharmaMar en el año 2015.
- Avanzit pasa a denominarse Ezentis.
- Cepsa deja de cotizar en 2011 tras una Opa.
- Sos Cuétara S.A. Pasa a denominarse Deoleo.
- Construcciones Reyal e Inmobiliaria Urbis se fusionan en junio de 2007 pasando a ser Reyal Urbis.
- Uralita pasa a denominarse COEMAC.

Se implantó el 1 de enero de 2005 la llamada Clasificación Sectorial y Subsectorial unificada [5], la cual categoriza todas las empresas, que cotizan en las diferentes bolsas españolas, entre las cuales se encuentran las seleccionadas en este documento. Esta clasificación es la que vertebra la agrupación en el estudio, aplicando dos niveles de orden. Un primer nivel de clasificación, compuesto por sectores, y un segundo nivel que divide estos primeros en subsectores. Se distinguen seis grupos principales de sectores:

- Petróleo y energía.
- Materiales básicos, industria y construcción.
- Bienes de consumo.
- Servicios de consumo.
- Servicios financieros e inmobiliarias.
- Tecnología y Telecomunicaciones.

Estos seis sectores, pasan a ser los conjuntos en los que se agrupan las series, para ello se accede también a la página del BME donde aparecen inventariadas las empresas siguiendo esta clasificación, las que no aparecen por no existir a día de hoy se clasifican buscando empresas similares de las que sí se tienen categoría. Se recopilan los subsectores también para cada una de las series.

Los conjuntos generados son los que aparecen en las Tabla 2 a la Tabla 7, donde se detallan los *tickers* de los valores, para cada grupo sectorial, indicando entre paréntesis el subsector, donde, por ejemplo, 1.1 indica el sector 1 subsector 1.

Sector 1	Sector del petróleo y energía				Número de Valores 14			
	IBE(1.2)	REP(1.1)	UNF(1.2)	REE(1.2)	GAS(1.2)	IBR(1.4)	ENG(1.2)	ELE(1.2)
	SLR(1.4)	AGS(1.3)	CEP (1.1)	ENC(1.4)	FRS(1.4)	MTB(1.4)		

*Tabla 2 – Conjunto de valores agrupados en el sector 1 de petróleo y energía.*

Sector 2	Materiales básicos, industria y construcción				Número de Valores 24			
	GAM(2.2)	ACS(2.3)	ANA(2.3)	FER(2.3)	ACX(2.1)	FCC(2.3)	TRE(2.6)	SCYR(2.3)
	ABG(2.6)	OHL(2.3)	SED(2.5)	TUB(2.1)	ZOT(2.2)	CAF(2.2)	GALQ(2.6)	ERC(2..5)
	CPL(2.4)	TRG(2.1)	MDF(2.6)	FDR(2.6)	CIE(2.1)	AZK(2.2)	CCMC(2.3)	ESZ(2.1)

*Tabla 3 – Conjunto de valores agrupados en el sector 2 de materiales básicos, industria y construcción.*

Sector 3	Bienes de consumo				Número de valores 17			
	ITX(3.2)	GRF(3.5)	ZEL(3.5)	ALM(3.5)	EBRO(3.1)	OLE(3.1)	FAE(3.5)	VIS(3.1)
	PAC(3.3)	MLX(3.6)	CPF(3.1)	BD(3.1)	ROVI(3.5)	SNC(3.2)	VID(3.6)	PVA(3.1)
	NAT(3.5)							

*Tabla 4 – Conjunto de valores agrupados en el sector 4 de bienes de consumo.*

Sector 4	Servicios de consumo				Número de valores 17			
	ABE(4.5)	CIN(4.5)	IBLA(4.4)	TL5(4.3)	PRS(4.3)	A3M(4.3)	MEL(4.1)	PSG(4.6)
	NH(4.1)	VOC(4.3)	VLG(4.4)	SPS(4.2)	CDR(4.1)	ITI(4.5)	DERM(4.6)	CBAV(4.6)
	VER(4.3)							

*Tabla 5 – Conjunto de valores agrupados en el sector 4 de servicios de consumo.*

Sector 5	Servicios financieros e inmobiliarias				Número de valores 23			
	SAN(5.1)	BBVA(5.1)	POP(5.1)	SAB(5.1)	MAP(5.2)	BME(5.6)	BKT(5.1)	CABK(5.1)
	COL(5.5)	BTO(5.1)	MVC(5.5)	ALB(5.3)	PAS(5.1)	BVA(5.1)	GCO(5.2)	QBT(5.5)
	REY(5.5)	REN(5.5)	RLIA(5.5)	PSL(5.5)	UBS(5.5)	GUI(5.1)	R4(5.6)	

*Tabla 6 – Conjunto de valores agrupados en el sector 5 de servicios financieros e inmobiliaria.*

Sector 6	Tecnología y Telecomunicaciones				Número de valores 5
	TEF (6.1)	IDR(6.2)	EZE(6.1)	AMP(6.2)	TEC(6.2)

*Tabla 7 – Conjunto de valores agrupados en el sector 6 de tecnología y telecomunicaciones.*

## **Preparación de los datos**

### **Carga**

La primera dificultad se encuentra a la hora de obtener los datos, como se ha descrito anteriormente, al no tener un único punto de obtención. Los datos se obtienen en este orden:

1. Se intentan descargar desde el portal de Yahoo Finance: siempre que se encuentre el valor que se busca. En una primera revisión se localizan 75 valores en Yahoo, aunque posteriormente por la calidad de los mismos o la indisponibilidad se obtienen de otras fuentes.
2. Se buscan en el portal financiero de Invertia: la práctica totalidad de los valores que no se localizaban en Yahoo sí aparecen en este portal. Algunas series que no tenían la calidad suficiente se obtienen también de aquí.
3. Adicionalmente se toma como referencia el portal de Google Finance para comprobar cotizaciones de días puntuales, por ejemplo los que tienen anomalías, o contrastar las series obtenidas.

El formato de carga es distinto dependiendo de la fuente, lo que se corrige para tener un formato de datos único con el que trabajar. Las series disponibles en Yahoo se descargan a través de la función *get.hist.quotes*, las que se descargan desde el portal de Invertia se reciben en formato excel, como tablas, con las siguientes características:

Los campos de fecha tienen dos formatos distintos: 30/12/08 ó 30-dic-2008.

La cotización a cierre de los valores tiene los decimales separados por comas.

Se encuentran caracteres especiales en las columnas.

Se realizan pruebas de carga directa, desde los archivos excel, al programa de Rstudio a formato zoo, pero siendo estas fallidas se determina usar un formato intermedio, en este caso csv.

## Limpieza

Tras cargar los datos se dispone de 100 series temporales con características diferentes. La mayor parte de ellas tienen 510 observaciones, comenzando el día 2 de enero de 2007 y terminando el 30 de diciembre del 2008, sin embargo algunas de ellas tienen menor tamaño ya que comienzan en fechas posteriores. La serie que tiene el menor tamaño es la de ESZ con 226 observaciones que comienza el 2 de febrero de 2008, lo que nos limitará el periodo de las estimaciones para aplicar el modelo DCC.

Para comenzar a trabajar se visualizan las series de los valores de cotización diaria, a cierre de mercado para cada uno de los valores, reflejadas en los gráficos A1 a A100 del anexo. Analizándolas se observan varios tipos de anomalías entre las que se encuentran las de cambio de nivel, como por ejemplo en Banco Sabadell, que se observa en el Gráfico 5. En ese caso atiende a un split por el cual cada acción pasó a dividirse en cuatro acciones nuevas, aunque al realizar la transformación se aplicó mal como se puede apreciar. En otros casos se debe probablemente a problemas en el almacenamiento o generación de los datos ya que no se detectan hechos relevantes como el indicado para SAB.

Por otro lado se detectan *missings* en muchas de ellas, que corresponden a días en los que no abrieron los mercados y se incluyen los días con valores no disponibles (NA). Adicionalmente se detectan *outliers* como en el caso de BKT, Gráfico 6, o series con bastantes datos anómalos como en el caso de GRF Gráfico 7. Todas estas anomalías se corrigen, ya que por un lado los modelos que representan la varianza son sensibles a outliers [6] y además es necesario que las series tengan todos los días completos, con su valor correcto para poder estimarlos. Se realizan varias modificaciones para corregirlas, por un lado en series como la de SAB se reconstruyen los valores de la serie tomando los datos hasta el salto y aplicando los porcentajes de variación diarios al último día previo al split, para el resto de casos se intenta recuperar los valores faltantes, se reemplazan las series por algunas con mejor calidad o como última opción, en casos muy contados, se eliminan los días.

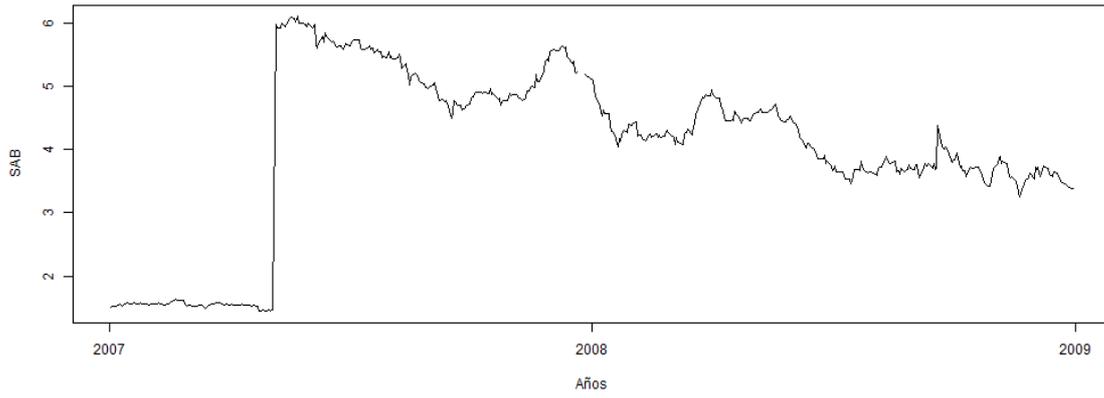


Gráfico 5 – Serie de cotizaciones a cierre de Sabadell.

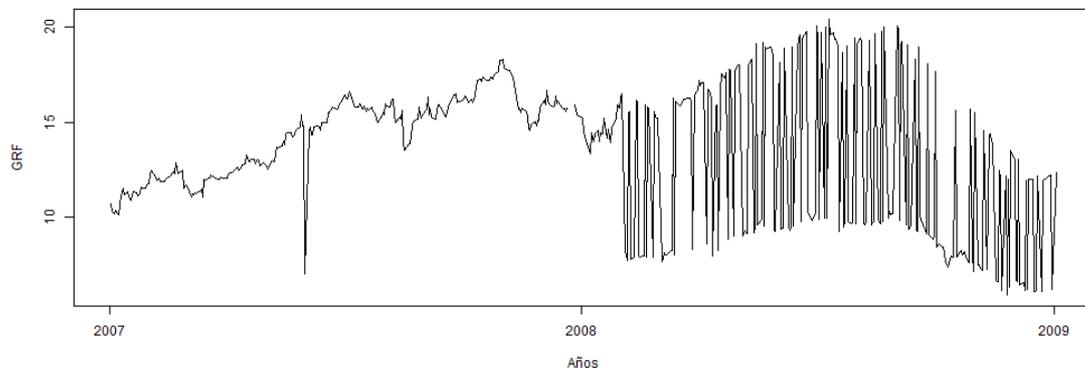


Gráfico 6 – Serie de cotizaciones a cierre de Grifols.

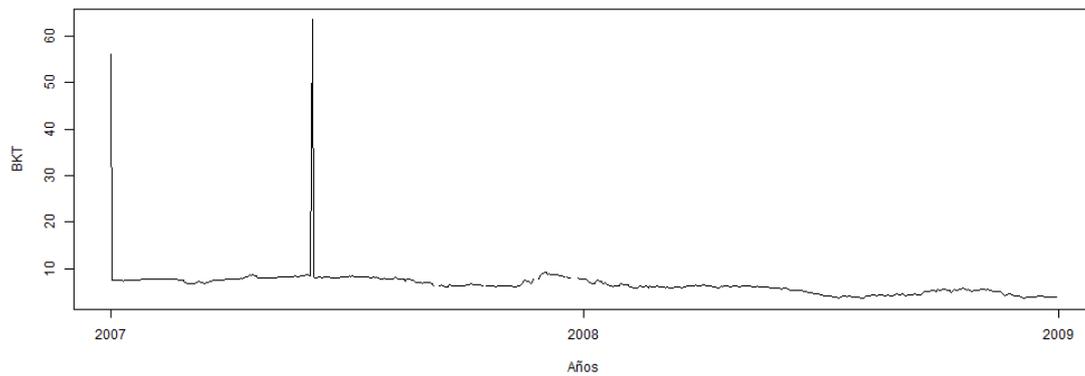


Gráfico 7 – Serie de cotizaciones a cierre de Bankinter.

## ***Estimación de modelos preliminares***

Una vez se tienen datos, con calidad suficiente, se definen las series de rendimientos diarios, entendiendo el rendimiento diario como la diferencia del logaritmo neperiano de los precios, a cierre de cotización, para cada uno de los días que completan cada una de las series. Por tanto el rendimiento en un día  $t$ , se calcula como la diferencia del logaritmo neperiano del precio de cierre, para ese día  $t$ , menos el logaritmo neperiano del precio de cierre para el día anterior  $t-1$ . Son estas series y no las de los datos originales las que se utilizan en los modelos apARCH y GARCH.

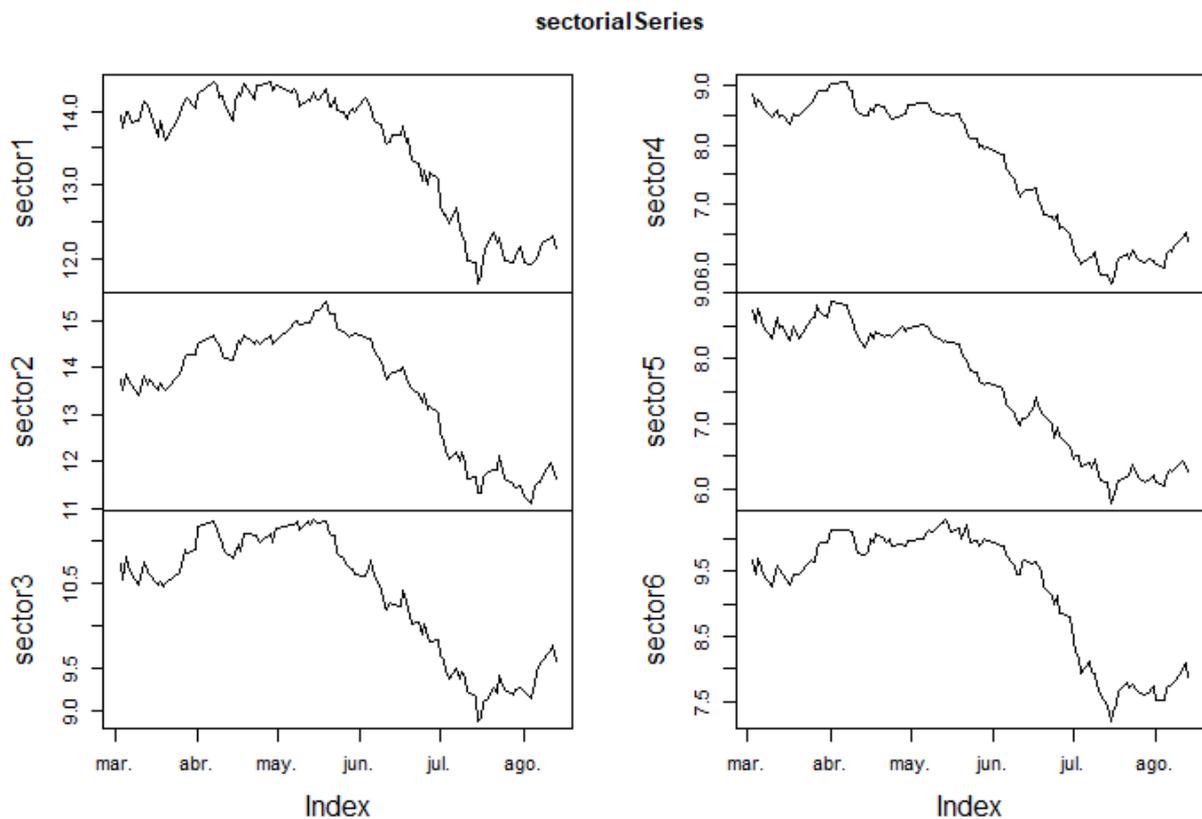
## **Creación de índices sintéticos**

El análisis intrasectorial se realiza de manera directa, simplemente agrupando las diferentes series en sectores. Para el resto de estimaciones, sin embargo, no existe un índice representativo para cada uno de los sectores, existen algunos índices de futuros y derivados que cubren algunos de los valores, pero no existe ninguno que cubra exactamente los que se usan aquí y que permita realizar una estimación intersectorial. Se generan, entonces, para cubrir esta carencia, seis índices sintéticos, usando para ello las series de cotizaciones a cierre de cada uno de los valores que los componen y estimando la media geométrica diaria de todos. Por tanto se tendrá un sólo valor que será media geométrica de los valores de todas las series para cada uno de los días. El periodo que se toma para generarlos es el determinado en la estimación de los modelos predictivos finales, que más adelante se explica, que abarca del 3 de marzo al 13 de agosto de 2008 y que cumple con el mínimo de 100 observaciones que se requieren para aplicar el modelo DCC, como también se explica en la estimación de modelos preeliminares.

Tomando el ejemplo del sector 6, de tecnología y telecomunicaciones por ser el menor tamaño, se tiene la media geométrica de las cotizaciones a cierre de TEF, IDR, EZE AMP, TEC, teniendo por tanto un único valor para el primero de los días, se realiza la misma

operación para cada uno de los días de la serie, consiguiendo así una serie única que deriva de las cinco que componen el sector.

Las gráficas de los datos estimados, de cada serie sintética, para los sectores son las que se pueden ver en el los Gráficos 8 y 8A. Al observar todas ellas se aprecia una cierta tendencia en cada una de las series, que hasta mediados de mayo es de un cierto sesgo alcista y a partir de entonces pasa a ser claramente bajista hasta mediados de julio, ya se puede apreciar una evolución acompasada en la que el comportamiento parece tener cierta similitud en cuanto a la tendencia.



*Gráfico 8 – cotizaciones estimadas para las series sintéticas de los sectores.*

### Series sintéticas de los sectores

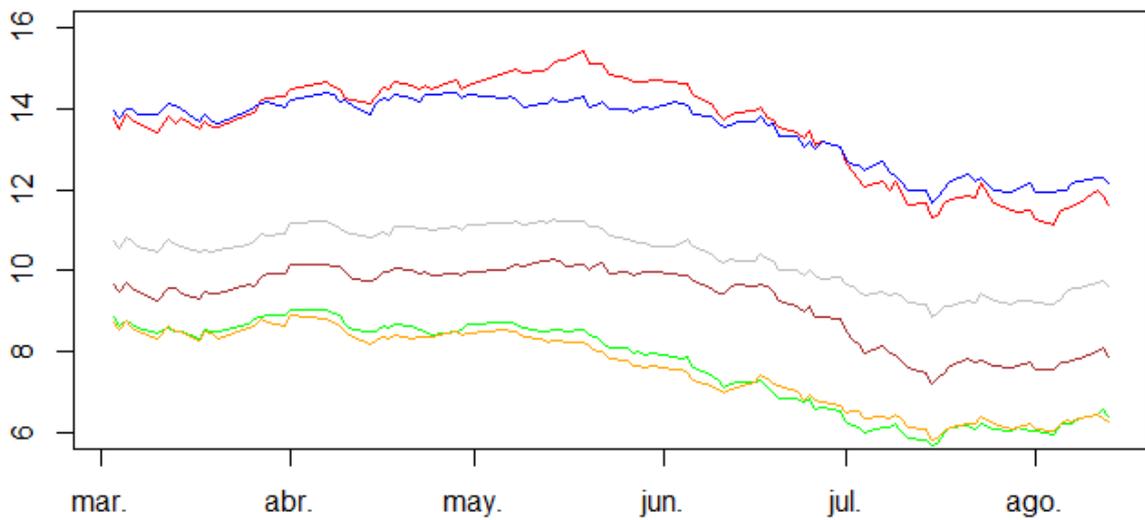


Gráfico 8A – superposición de las cotizaciones estimadas para las series sintéticas de los sectores.

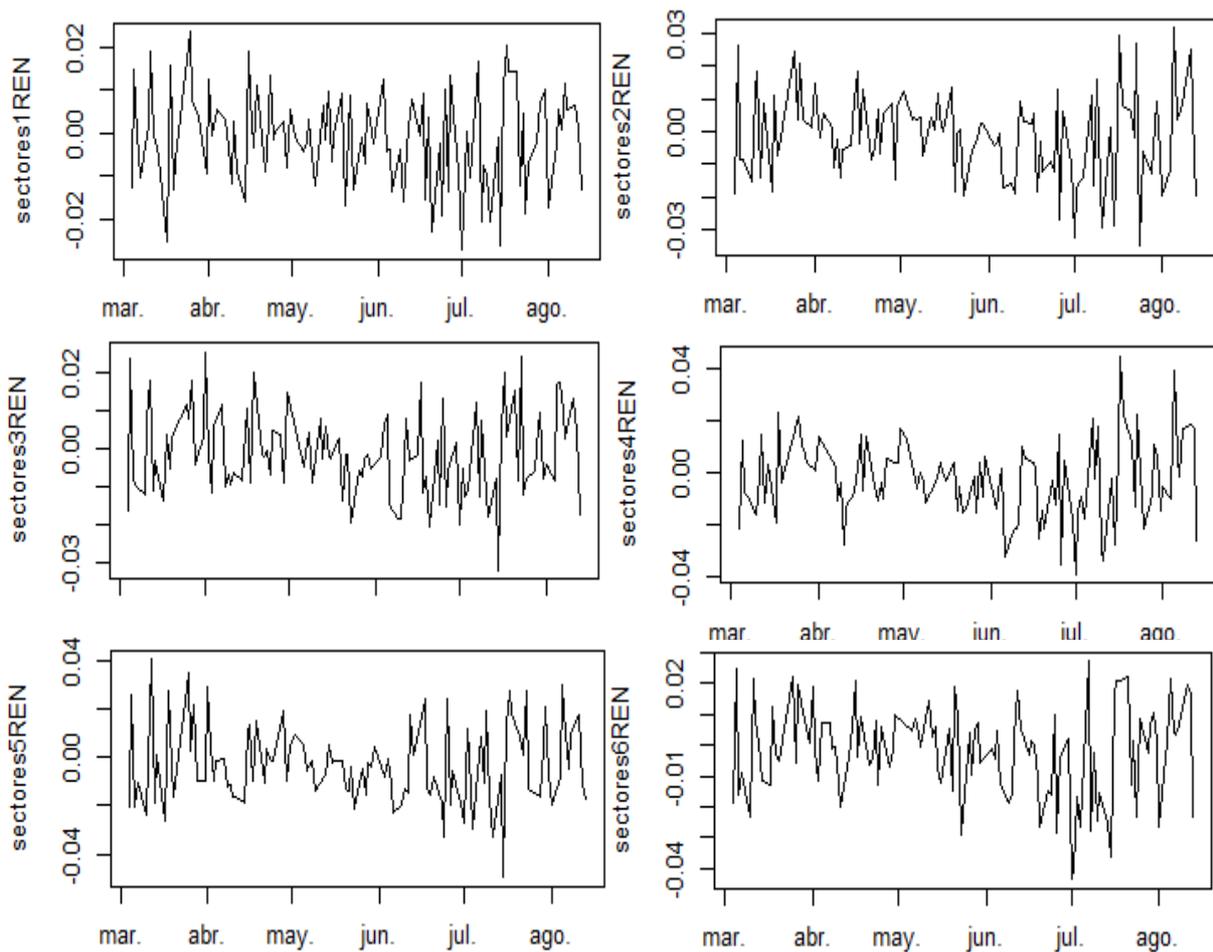


Gráfico 8B – Rendimientos diarios de los sectores.

## Modelos univariantes (GARCH/apARCH)

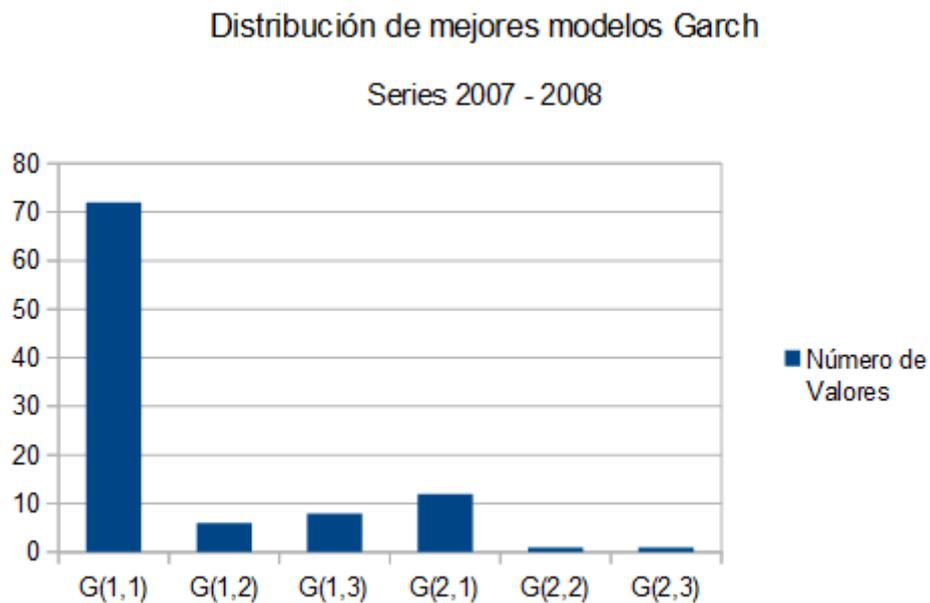
Se realiza una primera estimación de modelos, valorando los mismos a través del criterio de información de Akaike, que determinará cuál de ellos tiene mejor comportamiento para una serie dada. Para dichas estimaciones se comparan tanto modelos GARCH como apARCH, estos segundos con la peculiaridad de forzar que tengan momento de grado dos, o lo que es lo mismo se fija un  $\delta$  igual a dos, para que reflejen la varianza y ver así si la asimetría tiene un peso importante.

Se prueban ambos modelos asumiendo por un lado que  $\varepsilon_t$  tiene distribución normal y por otro que tiene distribución t-student, otro lado se utilizan valores de las componentes p y q en un rango de uno a tres. Se prueban por tanto  $2*2*9=36$  modelos por valor, teniendo en total para todas las series 3.600 modelos, de los que se seleccionan únicamente 400 y se utilizan los 100 con menor AIC para los cálculos multivariantes.

Para realizar dichos cálculos se preparan varios métodos en R que ejecutan secuencialmente los modelos indicados, sobre cada serie almacenada, incrementando las componentes en cada iteración y guardando el modelo calculado en cada pasada si el AIC resultante es menor que el que el mejor modelo estimado hasta el momento. Los métodos se pueden encontrar en el anexo del informe. Se tiene así, al final de la ejecución, una lista con los mejores modelos para cada serie que se pueden ver en la Tabla AT1, en los anexos, donde aparecen remarcados en verde los mejores y en naranja los que se han utilizado si el mejor modelo no se puede aplicar por tener mal comportamiento al estimar el modelo DCC.

El resultado, que se observa en el Gráfico 9, muestra cuántos valores han mostrado como mejor a cada uno de los modelos. De las nueve combinaciones posibles, de p y q, sólo seis de ellas se encuentran las que generan los mejores modelos. Entre los modelos con mejor comportamiento en un mayor número de series son, por mayor número de valores, el modelo GARCH(1,1), GARCH(2,1), GARCH(1,3) y GARCH(1,2), en menor medida GARCH(2,2) y GARCH(2,3). La mayor parte de los modelos se representan mejor con modelos que tiene valores bajos de p y q, lo que implica que sus ecuaciones serán mucho menos complejas,

además los modelos apARCH obtienen resultados con un AIC siempre algo mayor que los GARCH, aunque en ocasiones la diferencia es mínima, como se observa en la Tabla AT1.

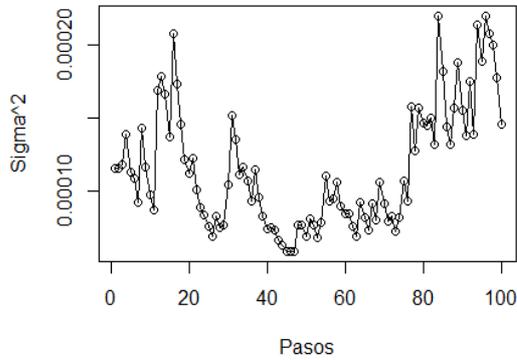


*Gráfico 9 - Número de series que mejor se ajusta a cada modelo, dentro de los mejores modelos seleccionados.*

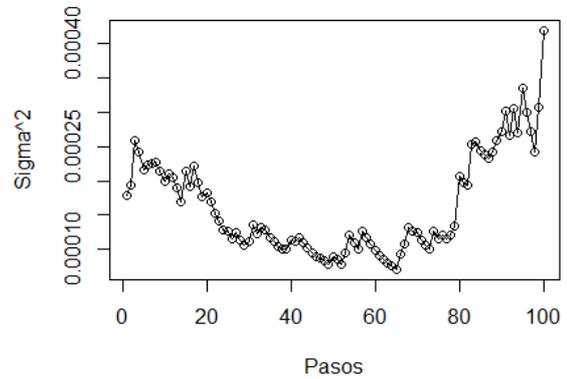
Se realiza la misma prueba de rendimiento para las series sintéticas, cuyo resultado se aprecia en la Tabla AT2. Además se pueden ver las estimaciones de la varianza, que se realizan recogiendo el parámetro sigma que estiman los mejores modelos para dichas series y que se pueden ver en los Gráficos 9A a 9F.

Gráficos 9A,9B,9C,9D,9E y 9F – Estimaciones de la varianza realizadas por los mejores modelos GARCH recogidos en la Tabla AT2.

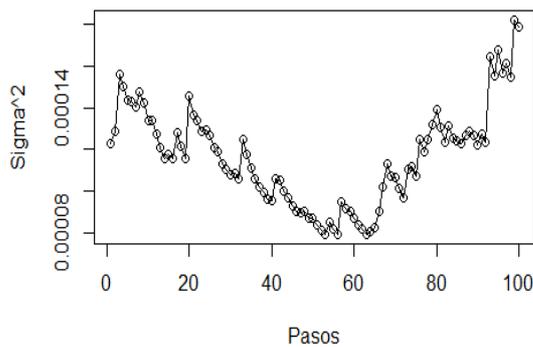
**Estimación de la varianza - Sector 1**



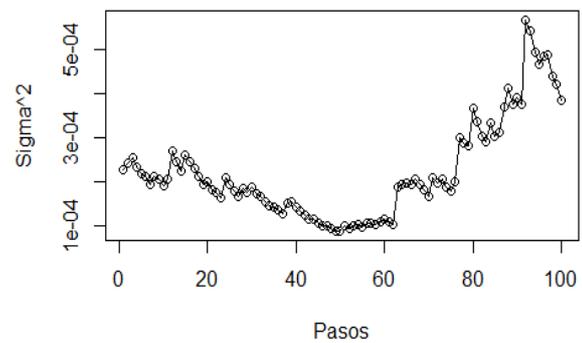
**Estimación de la varianza - Sector 2**



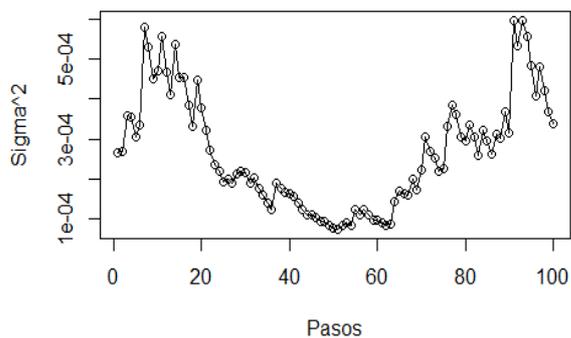
**Estimación de la varianza - Sector 3**



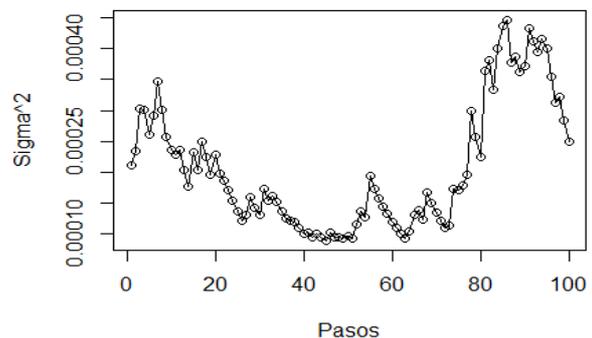
**Estimación de la varianza - Sector 4**



**Estimación de la varianza - Sector 5**



**Estimación de la varianza - Sector 6**



## **Modelos multivariantes (DCC)**

Los modelos GARCH y APARCH univariantes, estiman consistentemente las varianzas condicionales de las series de rendimientos. Para estimar consistentemente las matrices de correlación de los rendimientos de un conjunto de series pueden utilizarse los modelos GARCH multinomiales, y en particular los modelos DCC, que utilizan las mejores estimaciones univariantes del apartado anterior para obtener la estimación de las matrices de correlación.

Para la estimación se usan dos niveles de agrupación de los datos. Por un lado un nivel sectorial, internamente como se ha descrito antes y por otro un segundo nivel más alto, entre los sectores a través de series sintéticas. Se tienen como datos de entrada de los modelos, las diferentes varianzas representadas por los modelos GARCH para cada uno de los valores, en el caso de la estimación interna de cada sector, y para cada una de las series sectoriales en el caso de la estimación intersectorial.

Como prueba de comportamiento se preparan unas primeras pruebas de rendimiento del modelo DCC, en primer lugar se identifica un número mínimo de 100 observaciones por serie, para poder lanzar los modelos se definen entonces de manera aleatoria conjuntos, de series de valores, de 25, 50 y 55 de ellas, con 105 observaciones cada una. La ejecución es correcta pero se detecta un límite de datos cercano a las 55 series, a partir del cual el programa no es capaz de realizar las estimaciones, por ser las matrices de los cálculos intermedios excesivamente grandes.

En un primer análisis, de los resultados de las pruebas, se detectan ciertos problemas para manejar el conjunto de datos de salida. Hay que tener en cuenta que, para estimar la correlación, el modelo DCC calcula una matriz de correlación diaria, en el caso de estas pruebas se tienen 105 observaciones correspondientes a 105 días, por lo que se obtiene ese número de matrices con una magnitud de hasta 55 por 55 elementos, lo que como se ha indicado previamente hace difícil el manejo de los resultados y su interpretación a priori. En un primero momento, en el que todavía no se tenía las agrupaciones por sectores, por ejemplo,

se decide limitar los conjuntos de series a un número máximo de 25, lo que permite visualizar y tratar los resultados de manera óptima, por encima de ese número la labor se vuelve bastante tediosa. Los conjuntos finales, que componen los sectores, tienen todos ellos menos de 25 series de valores, por lo que esta problemática no se da en el resto del estudio.

Se pueden apreciar algunos estadísticos en la Tabla 8, que muestran, por ejemplo, cómo el AIC es menor cuanto mayor es el número de series y/o parámetros, pero en contraposición la log-verosimilitud es mayor.

N. Series	AIC	Log-verosimilitud	Número de parámetros
25	-140.96	7802.49	402
50	-176.21	10682.31	1427
55	-200.30	12222.58	1707

*Tabla 8 – Estadísticos de modelos de prueba DCC.*

En estas pruebas se han detectado por tanto dos factores a tener en cuenta: el límite de series sobre las que se pueden trabajar, 55 en este caso, y el mínimo de observaciones por serie, de 100.

### **Estimación de modelos predictivos finales**

Una de las principales dificultades, encontradas al realizar el proyecto, radica en el manejo no sólo de las series sino también de los resultados. El tener 100 series en paralelo con un número amplio de observaciones, que llegan hasta las 510, plantea algunas complejidades a la hora de estimar los modelos.

El número de observaciones no es homogéneo, por ejemplo, algunas series comienzan más tarde que otras, es el caso de Española del Zinc, que comienza a cotizar en febrero de 2008, muchas otras comienzan a cotizar a mediados de 2007, al menos se detectan nueve de ellas que empiezan a cotizar a partir de mediados de ese año.

Teniendo en cuenta las limitaciones antes citadas, se crea una ventana inicial de 100 observaciones, como referencia, comenzando el día 1 de marzo de 2008. Se toma cada serie y se separa un subconjunto muestral, para cada una de ellas, tomando como primera observación la correspondiente a esa primera fecha e incluyendo las 99 observaciones posteriores.

Al unificar las series por sectores se crean matrices, una por cada sector como se indicó antes, que contienen los diferentes subconjuntos muestrales de cada uno de ellos. Cada serie pasa a ser una columna en la nueva estructura, que a modo de ejemplo se ve en la Tabla 9.

	SAN	BBVA	...	PAS
01/01/07	1	6		9
03/01/07	2	7		8
04/01/07	3	8		6

*Tabla 9 - Ejemplo de matriz Sectorial.*

En el momento de generar la matriz se detecta que al casar las series entre ellas, osea al generar las nuevas columnas y combinarlas dentro de las matrices, algunas submuestras no tienen todos los valores, en parte debido al proceso de limpieza. Es necesario que todas las series encajen perfectamente coincidiendo todos los días para poder estimar las correlaciones, lo que por las diferentes transformaciones o simplemente por la no existencia de algún dato no se da. Se modifica la ventana, ajustando los valores inválidos, paulatinamente escogiendo finalmente el periodo comprendido entre el 3 de marzo de 2008 y el 13 de agosto de 2008.

La librería rmGARCH permite lanzar el modelo DCC, con sus dos fases, siendo posible definir para la primera fase el modelo GARCH(p,q) que se quiere utilizar, que permite configurar para cada serie su modelo concreto. Siguiendo la Tabla AT1 de comparativa GARCH-apARCH, antes detallada, se selecciona el mejor modelo para cada valor. En caso de que el modelo seleccionado para un valor no tenga buen comportamiento, a la hora de estimar el modelo DCC, se selecciona el segundo mejor modelo y se destacado en la misma tabla con color naranja como se ha indicado.

Para poder mostrar un resultado concreto de los modelos, que incluya el valor de la varianza a lo largo del periodo, se estima una única matriz de correlaciones. Para que dicha matriz contenga la información completa del periodo, se usa para cada una de las posiciones de la matriz, el valor de la media geométrica, resultado de aplicarla sobre las matrices de correlación para cada día estimado, de cada uno de los modelos de las diferentes series. Dicho valor se consigue aplicando esta media a cada uno de los datos correspondientes a cada serie y en cada una de las n matrices de correlación estimadas por el modelo DCC. Por tanto sea M la matriz que contiene las medias geométricas y  $M_{1,1}$  su primer elemento, sean M1 a MN las n matrices de correlación diaria de los modelos aplicados a las series, entonces:

$$M_{1,1} = \sqrt[n]{M1_{1,1} * M2_{1,1} * \dots * MN_{1,1}}$$

## Estimación para el sector del petróleo y la energía

Se ejecuta el modelo DCC para el sector 1, el resultado, en el Gráfico 10, refleja una alta correlación entre algunos de los valores, en concreto los que tienen perfil energético como Gas Natural (GAS), Red Eléctrica (REE), Enagás (ENG), Endesa (ELE), Iberdrola Renovables (IBR), Iberdrola (IBE) y Unión Fenosa (UNF) que coinciden en este caso con Repsol (REP), esta última petrolera. Se observa, en contraposición, una correlación débil de Montebalito (MTB), Solaria (SLR) y Fersa (FRS) con el resto, e inexistente o muy próxima a cero, en el caso de Cepsa (CEP) y Aguas de Barcelona (AGS).

Al aplicar un clúster ward, de cuatro niveles, se obtienen los grupos de valores que se aprecian en la matriz del Gráfico 11. Comparando los subsectores reales del sector, cuatro en total, con los clústers, se distingue que el primero de ellos contiene dos valores, que pertenecen al subsector 1.4 Solaria (SLR) y Fersa (FRS); el segundo valores de los sectores 1.1 Cepsa (CEP) y 1.3 Aguas de Barcelona (AGS); el tercero un conjunto perteneciente en su totalidad al subsector 1.2 Gas Natural (GAS), Red Eléctrica (REE) y Enagás (ENG) y el cuarto y último del sector 1.1 Repsol (REP), el 1.2 Iberdrola (IBE), Unión Fenosa (UNF) y Endesa (ELE) y el 1.4 Iberdrola Renovables (IBR), Grupo Empresarial Ence (ENC) y Montebalito (MTB). Por tanto, en el caso del sector de petróleo y energía, la agrupación por

clústers derivada de la matriz de correlaciones no presenta demasiadas similitudes con los subsectores reales.

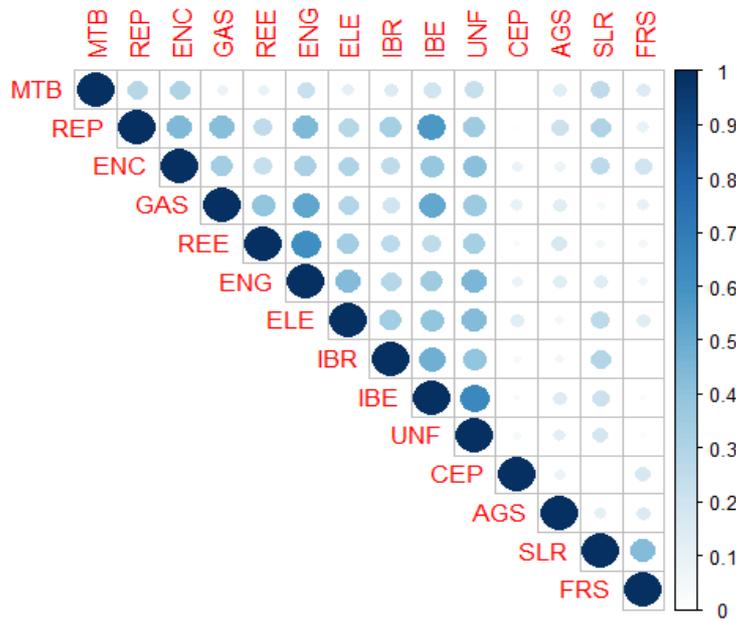


Gráfico 10 – Matriz media geométrica de las matrices de correlación estimadas.

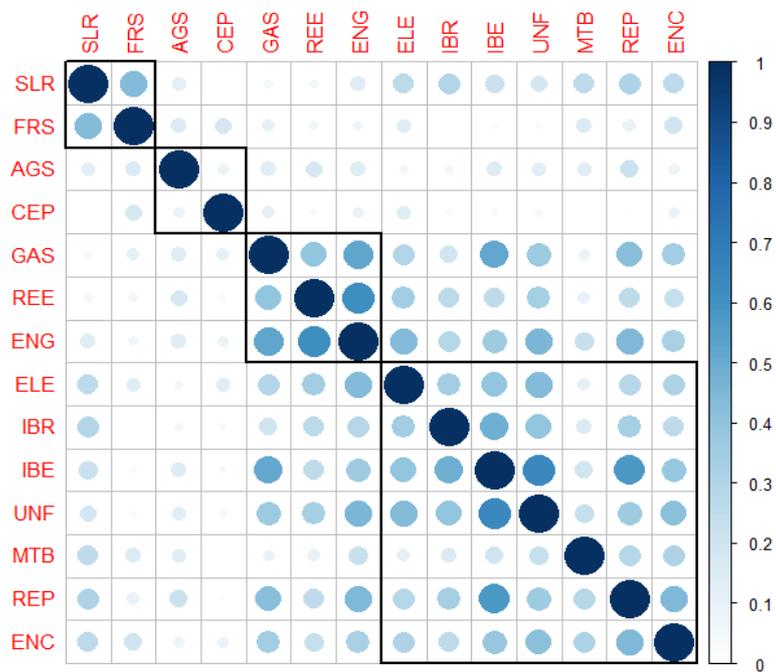


Gráfico 11 – Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas.

## Estimación para el sector de materiales básicos, industria y construcción

En el caso del segundo sector, se realiza el mismo proceso y se observa como resultado la matriz del Gráfico 12. Se aprecia una alta correlación generalizada, a excepción de Uralita (CCMC), Española del Zinc (ESZ), Ercros (ERC), CIE Automotive (CIE) y Azkoyen (AZK), que bien tienen una correlación baja, con el resto de valores o incluso no tienen correlación, por ejemplo CCMC con Construcción y Auxiliar de Ferrocarriles (CAF), AZK, Gamesa (GAM) y Técnicas Reunidas (TRE).

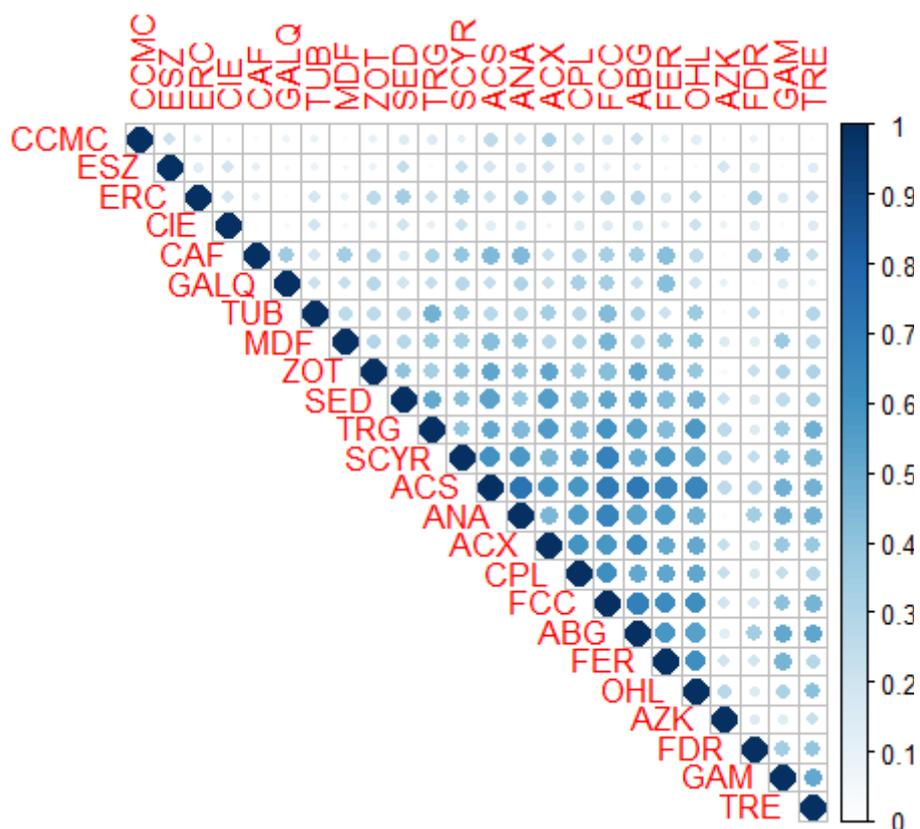


Gráfico 12 – Matriz media geométrica de las matrices de correlación estimadas para el sector de materiales básicos, industria y construcción.

El Gráfico 13 no nos da información destacable a nivel individual con respecto a sus subsectores, los valores se agrupan en clústers que en general no coinciden con los

subsectores reales. Si destaca a nivel global el clúster número 3, que se ve en el centro de la matriz, ya que éste reúne a 13 de los 25 valores que conforman el sector y éstos tienen una alta correlación. No a nivel de cada uno de los subsectores, pero sí a nivel del conjunto del sector, se puede determinar una alta correlación en un gran número de valores, siendo algo menor en el resto, como ya se había indicado antes.

Del clúster central destaca, también, que en él se encuentran todos los valores, menos uno, pertenecientes al subsector 2.3, de la construcción Actividades de Construcciones y Servicios (ACS), Acciona (ANA), Ferrovial (FER), Fomentos de Construcciones y Contratas (FCC), Sacyr (SCYR) y Obrascon Huarte Lain (OHL), y el único de los valores de dicho subsector que no aparece Uralita (CCMC) no se dedicaba puramente a la construcción, sino a la fabricación de materiales de construcción.

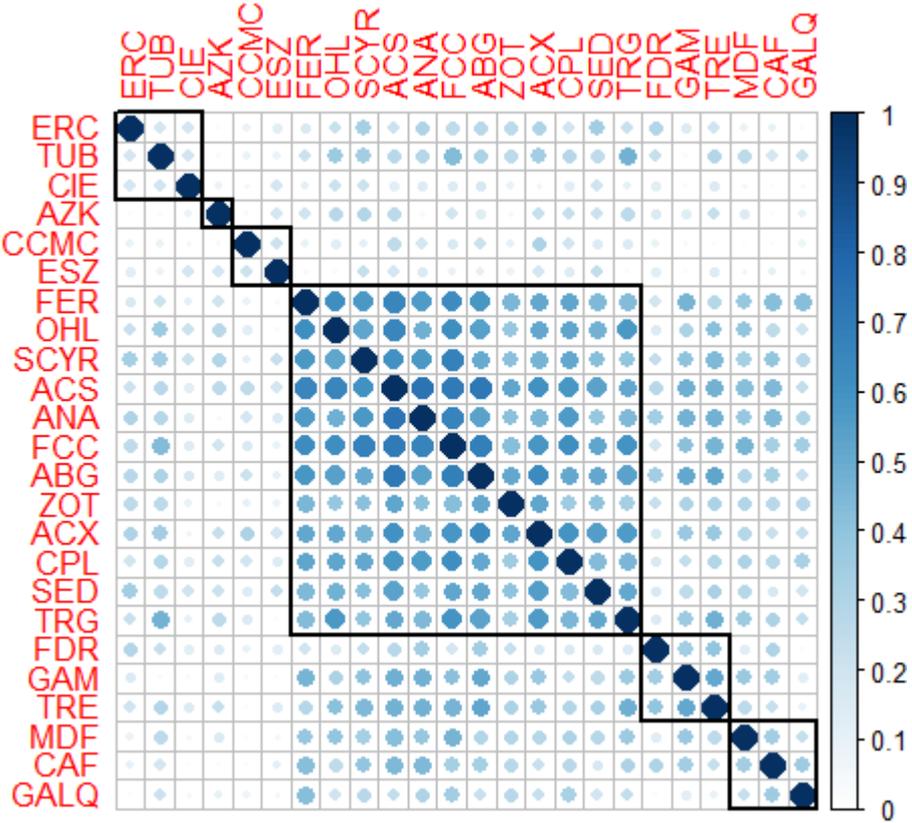


Gráfico 13 - Clúster de 4 grupos de la matriz, de la media geométrica, de las matrices de correlación estimadas para el sector de materiales básicos, industria y construcción.

Se aprecia por tanto una coincidencia clara en ese grupo de valores, en cuanto a la correlación de la varianza. De los otros cinco valores que lo componen, casi todos tienen

cierta relación con la construcción, por ejemplo, Tubos Reunidos (TRG) se dedica a la ingeniería civil, Zardoya Otis (ZOT) a la fabricación e instalación de ascensores, Acerinox (ACX) a la fabricación de acero inoxidable, que en parte distribuye al sector de la construcción y Cementos Portland (CPL) es fabricante de cemento. La Seda de Barcelona (SED) no tiene una relación en cuanto al perfil de empresa, por dedicarse a la fabricación de materiales químicos.

### Estimación para el sector de bienes de consumo

En el sector de bienes de consumo no se aprecia ninguna alta correlación y la correlación existente es bastante baja, existiendo bastantes valores incorrelados como se puede observar en el Gráfico 14. Entre los valores que mayor correlación tienen se encuentran

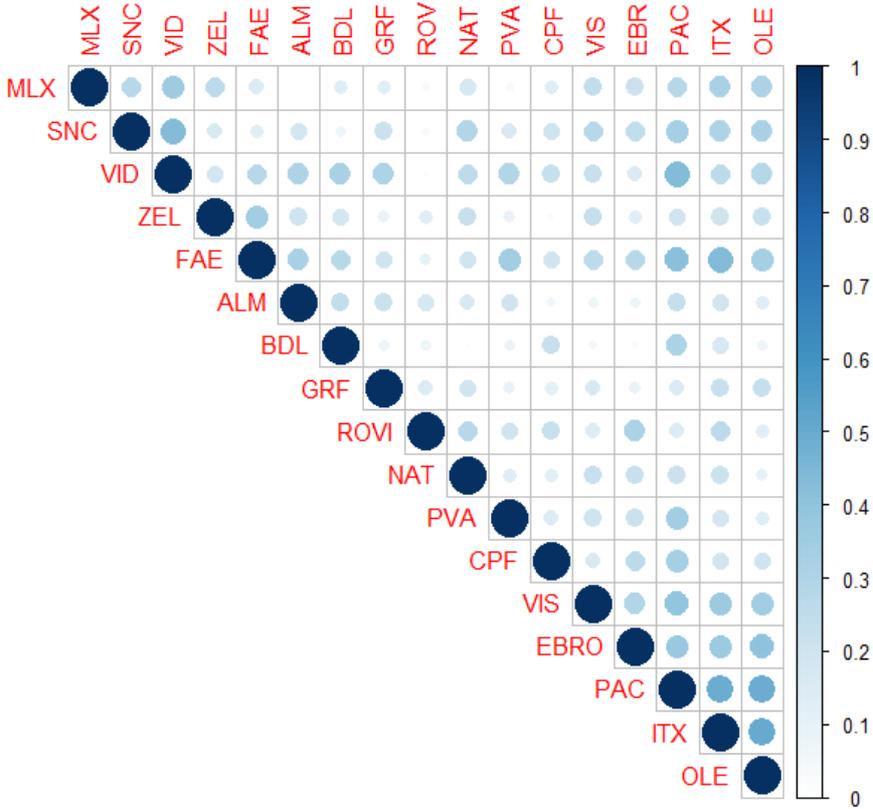


Gráfico 14 – Matriz media geométrica de las matrices de correlación estimadas para el sector de bienes de consumo.

Con respecto a la generación del clúster jerárquico, que se aprecia en el Gráfico 15, el subsector 3.1 Ebro Foods (EBRO), Deoleo (OLE), Viscofan (VIS), Campo Frío (CPF), Barón de Ley (BDL) y Puleva (PVA) se divide en tres clústers el clúster número 1 que comparten con el sector 3.2 Inditex (ITX) y el 3.3 Papeles y Cartones de Europa (PAC), el clúster número 3 donde sólo se encuentra un valor, BDL, por último el 5 compuesto únicamente por valores de ese subsector.

El clúster número 2 se compone de todos los valores del subsector 3.6 Mecalux (MLX), Vidrala (VID) uno del 3.5 Grifols (GRF) y otro del 3.2 Sniace (SNC). El número 3 lo componen valores del subsector 3.5 Zeltia (ZEL), Faes Farma (FAE), Laboratorios Almirall (ALM) combinado con el ya citado del 3.1. Por último los dos últimos valores restantes del subsector 3.5 tienen su propio clúster.

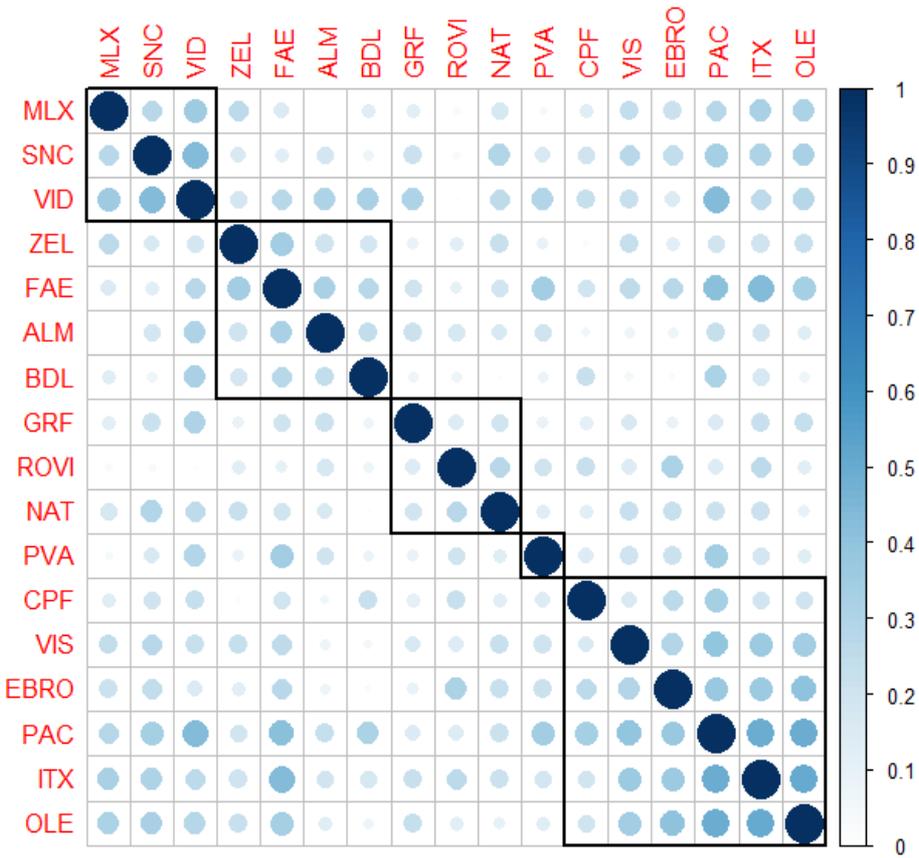


Gráfico 15 - Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas para el sector de bienes de consumo.

## Estimación para el sector de servicios de consumo

En el Gráfico 16 se incluye la matriz para el sector de servicios de consumo, en este caso hay algo más de correlación que en el caso de bienes de consumo, pero no tanta como se dio en el sector que incluía la construcción. Destaca un caso que prácticamente no tiene correlación con el resto de valores y pertenece al subsector 4.1 Codere (CDR) y correlación media entre las aerolíneas del subsector 4.4 Iberia Líneas Aéreas (IBLA) y Vueling (VLG).

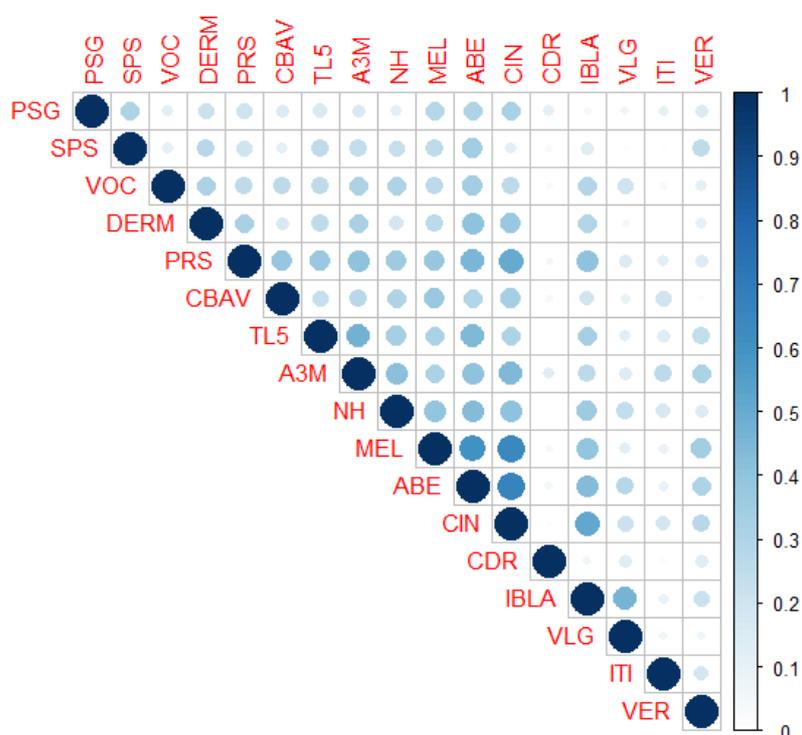


Gráfico 16 – Matriz media geométrica de las matrices de correlación estimadas para el sector de bienes de consumo.

Al aplicar el clúster de seis grupos, por tener ese número de subsectores, el valor que no tiene correlación tiene su propio clúster, las aerolíneas que comentábamos antes también tienen uno propio, el resto se reparten en clústers de dos valores cada uno y por último aparece un conjunto mayor que contiene tres valores del subsector 4.3 Telecinco (TL5), Atresmedia (A3M) y Prisa (PRS), los componentes restantes del subsector 4.1 Hoteles Meliá

(MEL) y NH Hoteles (NH), dos de los tres valores del 4.5 Abertis Infraestructuras (ABE) y Cintra (CIN) y uno del 4.6 Clínica Baviera (CBVA).

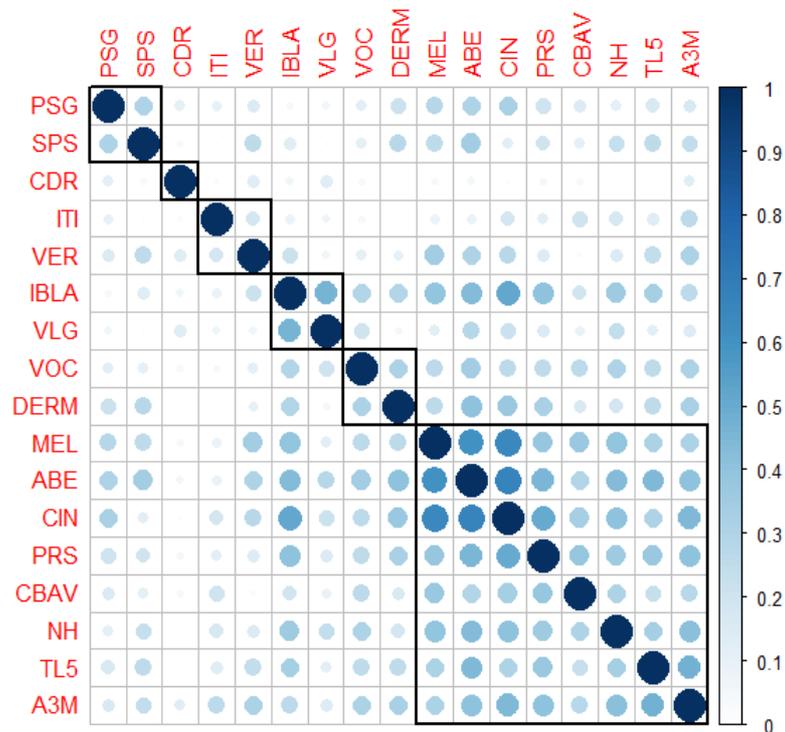


Gráfico 17 - Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas para el sector de servicios de consumo.

## Estimación para el sector de servicios financieros e inmobiliarias

En el caso del sector financiero e inmobiliaria el Gráfico 18 refleja una alta correlación muy clara entre los bancos, de hecho se podría dividir la matriz en dos, incluyendo por un lado las columnas de la matriz a la izquierda de la columna de Metrovacesa (MVC) y por otro las restantes, teniendo en ese caso agrupados todos los bancos salvo Banco Guipuzcoano (GUI) en la parte izquierda y en la parte derecha todas las inmobiliarias. Revisando la correlación de algunos de los bancos por ejemplo Santander (SAN) y Banco Bilbao Vizcaya (BBVA) tienen una correlación de 0.92, que es la mayor de todas seguida de la de BBVA y Popular (POP) de 0.845. En la parte superior la correlación es muy baja o nula.

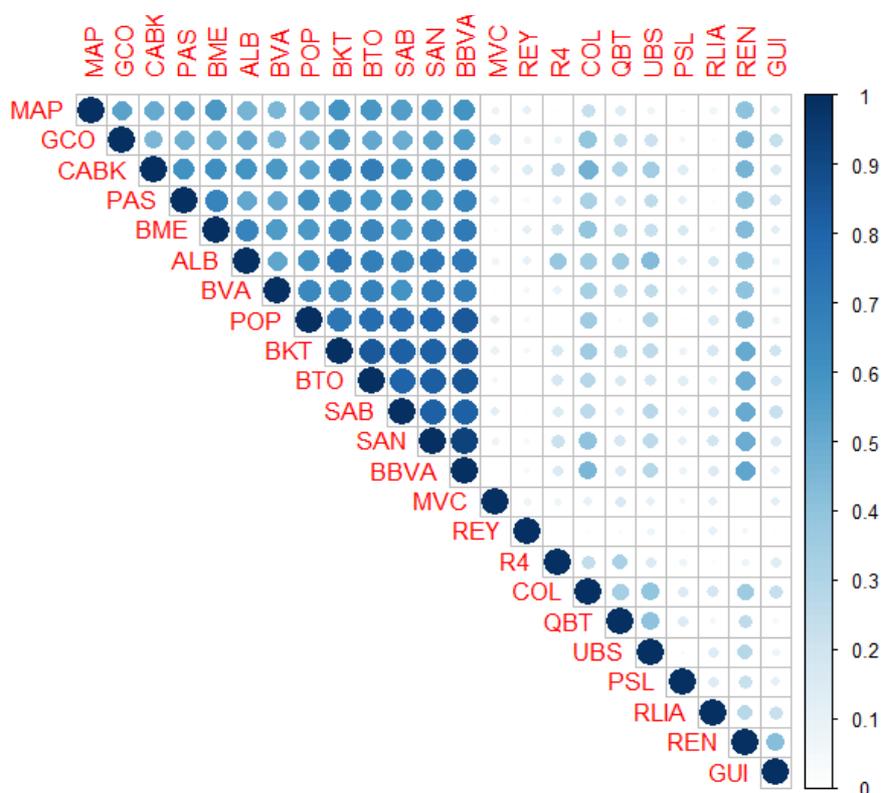


Gráfico 18 – Matriz media geométrica de las matrices de correlación estimadas para el sector de servicios financieros e inmobiliarias.

En cuanto a los clústers en el Gráfico 19 aparecen 5 clústers, entre los que destaca sin duda el primero de ellos que es el conglomerado de todos los bancos del subsector 5.1 SAN, BBVA, POP, Sabadell (SAB), Bankinter (BKT), CaixaCorp (CABK), Banesto (BTO), Banco Pastor (PAS) y Banco de Valencia (BVA) junto con aseguradoras del subsector 5.2 Mapfre (MAP), Gaesco (GCO) y el único valor del 5.3 Corporación Financiera Alba (ALB). En un solo clúster aparecen tres subsectores completos salvo un valor, GUI, con una correlación muy alta.

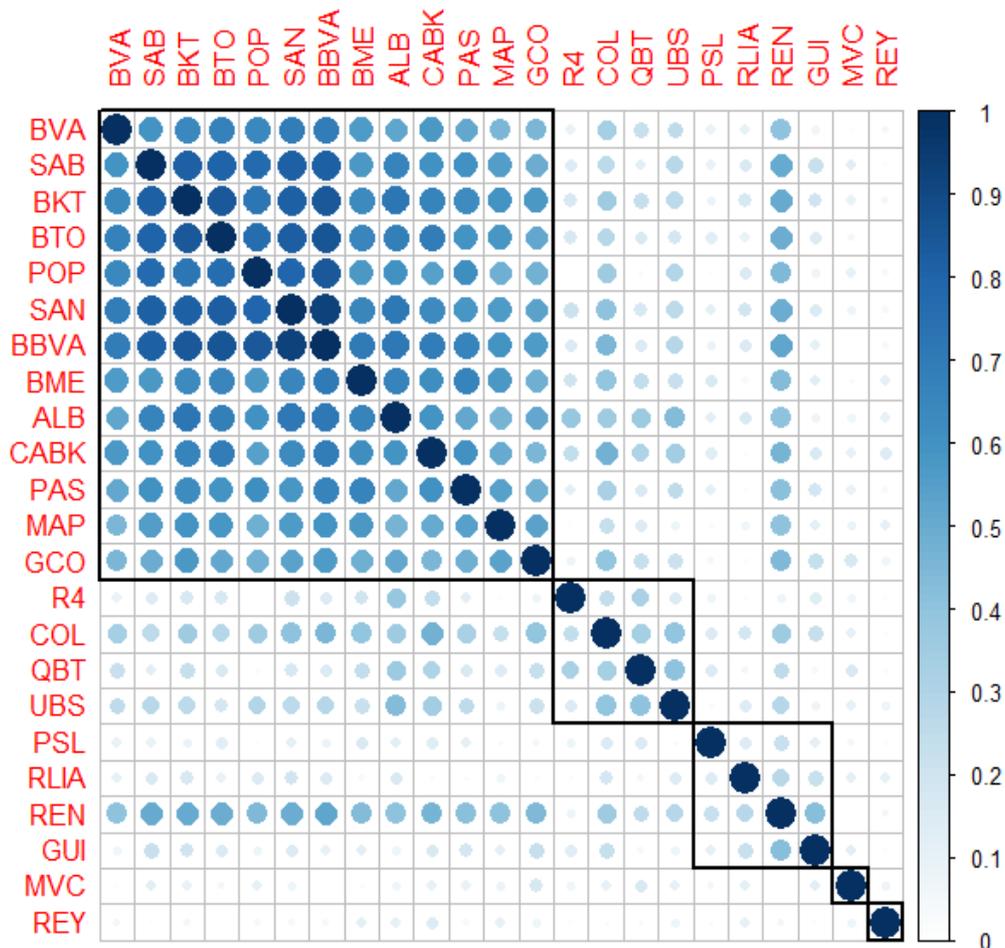


Gráfico 19 - Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas para el sector de servicios financieros e inmobiliarias.

## Estimación para el sector de tecnología y telecomunicaciones

En el último de los sectores, de menor tamaño, visible en el Gráfico 20, la correlación es baja y media en la mayoría de los casos, salvo entre Amper (AMP) y Ezentis (EZE), que es media-alta de 0.58. El clúster de dos grupos, en el Gráfico 21, separa a Tecnocom (TEC) del resto de los valores. Se detecta una correlación bastante aproximada para todos los valores entre sí aunque no es de excesiva intensidad.

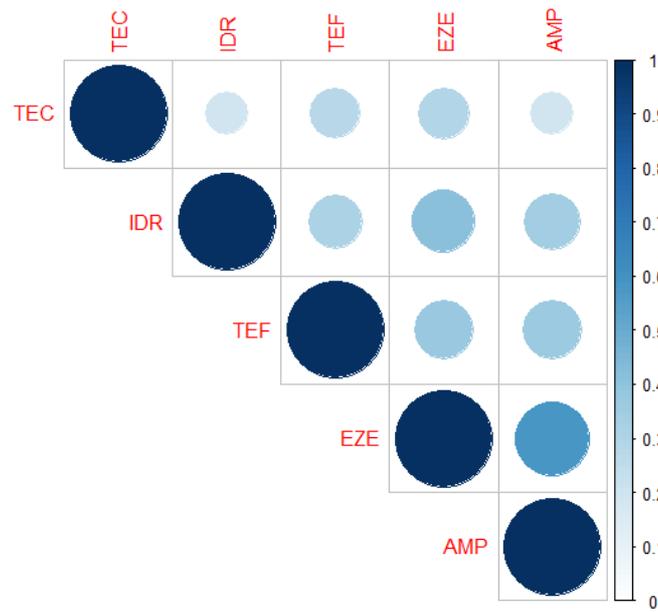


Gráfico 20 – Matriz media geométrica de las matrices de correlación estimadas para el sector de tecnología y telecomunicaciones.

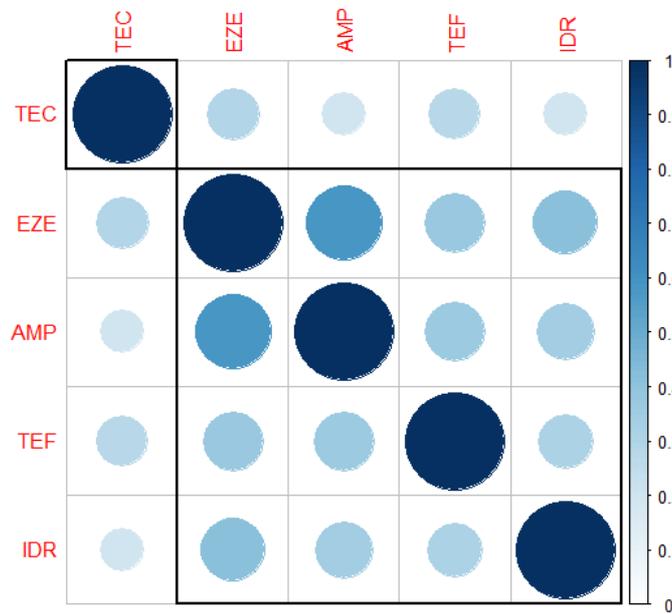


Gráfico 21 - Clúster de 4 grupos de la matriz media geométrica de las matrices de correlación estimadas para el sector de tecnología y telecomunicaciones.

## Estimación para las series sectoriales sintéticas

Al estimar las matrices de correlación sin clúster, Gráfico 22, y con él, Gráfico 23, se observan los siguientes resultados: la correlación es relativamente alta, con un valor mínimo de 0.64 entre los sectores 1 con el 4 y 5. El sector 2 está muy correlado con 3, 4, 5 y 6 con valores entre 0.809 y 0.83.

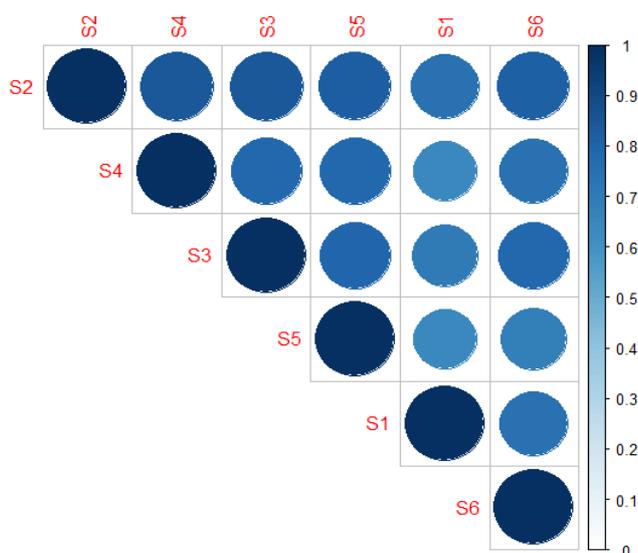


Gráfico 22 – Matriz media geométrica de las matrices de correlación estimadas para el conjunto de series sintéticas.

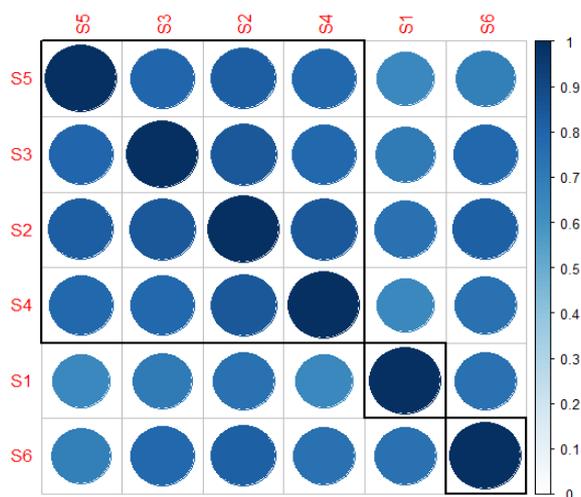


Gráfico 23 - Clúster de tres grupos de la matriz media geométrica de las matrices de correlación estimadas para el conjunto de series sintéticas.

## Conclusiones

Enlazando con el inicio del proyecto, las crisis en general, volvemos al foco principal, la crisis de 2008. En el desarrollo del estudio se ha podido analizar el riesgo, asociado a a dicha crisis, a través de la medición de la varianza, por medio de modelos GARCH y apARCH, cuya correlación se ha desgranado mediante modelos DCC.

De la experiencia adquirida se puede afirmar que, de los modelos probados, los que mejor comportamiento tienen para este conjunto muestral, en gran medida, son los GARCH(1,1).

En cuanto a los estudios de la correlación, de los que derivan los resultados asociados a los dos objetivos principales, se obtiene información relevante, tanto para el intrasectorial como el intersectorial. A nivel intrasectorial los valores muestran un comportamiento variado, de la correlación, que permite la creación de tres grupos representativos.

El primero de los grupos lo conforman el sector de materiales básicos, industria y construcción, junto con el de servicios financieros e inmobiliarias. En este conjunto se aprecia un gran número de valores con alta correlación, claramente detectable, entre valores que tienen el mismo perfil empresarial, como pueden ser el conjunto de los bancos, las constructoras, algunas empresas de materiales o inmobiliarias, todos ellos se agrupan junto con grupos de valores que no tienen ninguna o muy poca correlación. Por tanto los sectores que pertenecen a este primer grupo se pueden denominar como “polarizados”, donde subconjuntos de cada uno de estos sectores tienen o bien una fuerte correlación entre ellos o ninguna o muy baja correlación.

Un segundo grupo lo conforman los sectores de servicios de consumo y bienes de consumo y del petróleo y la energía. La correlación interna de estos sectores es algo más débil que la apreciada en el primer grupo. A diferencia del primero que tiene una cierta “polarización”, en este caso el comportamiento es muy variado, no se distinguen grupos de

valores que tengan alta correlación, pero sí algunos que tienen una correlación media, como se aprecia en el clúster más grande mostrado en el Gráfico 17.

El tercer y último grupo es el de tecnología y telecomunicaciones, que muestra una correlación media-alta generalizada.

Respecto a las series sintéticas de los sectores, hay una correlación bastante alta entre todos ellos, lo cual refuerza la idea mostrada en el momento de generación de los índices. Al igual que las series mostraban una cierta similitud en cuanto a su desarrollo en el tiempo y la tendencia tenía también similitudes, tras analizar los resultados queda patente el hecho de que a todo esto se suma correlación en las varianzas de los índices.

En conclusión se detectan correlaciones en las mediciones de las varianzas, que se asocian al riesgo de mercado, en mayor o menor medida, como se ha documentado en el estudio, con diferente intensidad sobre todo a nivel intrasectorial.

## **Futuro estudio y aplicaciones**

Como evolución del estudio que se ha desarrollado aquí se plantean diferentes opciones de ampliación o aplicación de los resultados. Una vez detectada cierta correlación, en el riesgo de mercado, posteriormente se pueden generar varias líneas de estudio y aplicaciones, algunas de ellas se describen en las siguientes líneas.

En cuanto a la ampliación del estudio se propone ampliar el conjunto de modelos autorregresivos que modelizan las series de varianza, por si se detectase algún modelo con características que definiese de manera más precisa los datos aquí modelizados. Hay que tener en cuenta que aunque los modelos no avanzan con excesiva rapidez, si surgen nuevas alternativas o variantes que aprovechan características concretas de las series.

Del mismo modo se plantea el uso de alternativas a los modelos DCC, que pudieran mostrar un mejor rendimiento, e incluso la realización de un estudio de la correlación a nivel

global de las 100 series, lo que daría información adicional sobre la correlación a nivel más atómico entre todas las empresas estudiadas, de ser posible a nivel de rendimiento.

A raíz de los resultados obtenidos se plantean algunas aplicaciones, por ejemplo, la aplicación del mismo estudio en diferentes etapas históricas, como pueden ser otras de las crisis mencionadas al comienzo del trabajo. En ese caso se tendrán que tener en cuenta las limitaciones ya descritas con respecto a los datos.

Otra de las aplicaciones más interesantes puede ser la de realizar un estudio en tiempo real o casi real, usando para ello técnicas de minería de datos como las usadas también en este proyecto, pero combinadas con tecnología Big Data, en la que se pudiera ampliar el estudio de manera cuantitativa. Se podrían en ese caso recibir, diariamente, muestras de datos que fueran realizando un estudio incremental, siendo en este caso importante el rendimiento de los modelos utilizados. En este caso sería más plausible la división en varias líneas de trabajos por sectores.

Una de las aplicaciones inmediatas es la de la estimación del Valor en Riesgo (VER) de carteras, ya que el cálculo del VER requiere de la estimación de las matrices de varianzas y covarianzas diarias. La importancia de la estimación consistente del VeR es fundamental para las entidades financieras, nótese en este sentido que es uno de los requisitos a los que están sujetas las entidades financiera europeas por los acuerdos de Basilea II.

Realizando una extrapolación trasversal, se podría contrastar el resultado actual, con hechos o factores de la economía real. Así, en el periodo indicado, se podrían localizar efectos que la correlación del riesgo aquí expuesta pudiera haber tenido sobre la economía, lo que podría derivar en informes que pudieran utilizarse por instituciones públicas o privadas para aplicar medidas de contención de los efectos.

Como se puede ver es amplio el campo de acción por tener el estudio inicial un amplio foco, afectando a diversos sectores y datos.

## Bibliografía

Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation. *Econometrica*, Vol. 50, No. 4 , 987-1007.

Engle, R. F. (2002). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *Journal of Business & Economic Statistics* Vol. 20, No. 3, 339-350.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticit. *Journal of Econometrics*. Vol. 31, No. 3, 307–327.

Bollerslev, T. (2008). Glossary to ARCH (GARCH). *Research Paper, 2008-49*.  
[ftp://ftp.econ.au.dk/creates/rp/08/rp08\\_49.pdf](ftp://ftp.econ.au.dk/creates/rp/08/rp08_49.pdf)

Luc Bauwens, Sébastien Laurent y Jeroen V. K. Rombouts (2006). Multivariate Garch Models a survey. *Journal of Applied Econometrics*. Vol. 21, 79-109.

Tsay, Ruey S. (2006) *Analysis of Financial Time Series* (2 ed.). *Wiley-Blackwell*.

Ghalanos, A. (2016), Manual del paquete RMGARCH “Multivariate GARCH Models”.  
Disponible en: <https://cran.r-project.org/web/packages/rmgarch/rmgarch.pdf>

Ghalanos, A. (2015), The rmgarch models: Background and properties, V 1.3-0. Disponible en: [https://cran.r-project.org/web/packages/rmgarch/vignettes/The\\_rmgarch\\_models.pdf](https://cran.r-project.org/web/packages/rmgarch/vignettes/The_rmgarch_models.pdf)

## Referencias

- [1] Lafont, I. (2008). “El contagio de la crisis financiera a la economía real derrumba las Bolsas”. El País, 16 de octubre de 2008. Disponible en:  
[https://elpais.com/diario/2008/10/16/economia/1224108003\\_850215.html](https://elpais.com/diario/2008/10/16/economia/1224108003_850215.html)
- [2] González-Páramo, J. M. (2010). “Basilea III y el nuevo marco macroprudencial”. Banco Central Europeo. Disponible en:  
[https://www.ecb.europa.eu/press/key/date/2010/html/sp101119\\_1.es.html](https://www.ecb.europa.eu/press/key/date/2010/html/sp101119_1.es.html)
- [3] Legendre, P y Murtagh, F. (2014). “Ward’s Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward’s Criterion?”. *Journal of Classification*, Vol. 31, No 3, 274-295. Disponible en: <https://link.springer.com/article/10.1007/s00357-014-9161-z>
- [4] ”Ranking contratación Bursátil del año 2008”. Bolsas y Mercados Españoles. Disponible en: [http://www.bolsasymercados.es/docs/infmercado/2008/esp/IM\\_visualizarPagina.aspx?codigo=II\\_1\\_D#esp2](http://www.bolsasymercados.es/docs/infmercado/2008/esp/IM_visualizarPagina.aspx?codigo=II_1_D#esp2)
- [5] Bolsas y Mercados Españoles (2010). “BME - CLASIFICACIÓN SECTORIAL BURSÁTIL, Descripción general y especificaciones por subsector”. Marzo 2010. Disponible en: <http://www.bolsamadrid.es/esp/Empresas/ClasificacionSect.aspx>
- [6] Carnero, M.A., Peña, D. Y Ruiz, E. (2007). “Effects of Outliers on the Identification and Estimation of GARCH Models” *Journal of Time Series Analysis*, Vol. 28, No. 4, 471-497. Disponible en:  
[http://rady.ucsd.edu/faculty/directory/valkanov/pub/classes/mfe/docs/HansenLunde\\_Garch.pdf](http://rady.ucsd.edu/faculty/directory/valkanov/pub/classes/mfe/docs/HansenLunde_Garch.pdf)

## Anexos

Tabla AT1 – Tabla de comparativa de Gach-apARCH con los resultados de selección de los mejores modelos para cada serie.

Valor	apARCH						GARCH					
	Normal			T-Student			Normal			T-Student		
	p	q	AIC	p	q	AIC	p	q	AIC	p	q	AIC
SAN	3	3	-4,651	1	2	-4,718	2	1	-5,057	2	1	-5,071
TEF	2	1	-5,360	3	1	-5,180	3	1	-5,399	1	1	-5,458
BBVA	3	3	-4,757	1	2	-4,812	2	1	-5,106	2	1	-5,122
IBE	2	1	-4,738	1	3	-4,616	1	1	-4,826	1	1	-4,902
REP	2	2	-4,798	1	2	-4,853	2	1	-5,058	1	1	-5,130
UNF	1	2	-5,395	1	3	-5,483	1	1	-5,618	1	1	-5,732
POP	3	3	-4,686	1	3	-4,786	1	1	-4,994	1	1	-5,050
ITX	1	3	-4,692	1	3	-4,561	1	1	-4,803	1	1	-4,847
GAM	2	2	-4,161	1	1	-4,046	3	3	-4,370	2	1	-4,399
ACS	1	3	-4,897	1	3	-4,720	1	1	-4,969	1	1	-5,008
REE	2	1	-5,299	3	1	-5,224	2	2	-5,300	1	1	-5,398
ANA	2	2	-4,066	2	2	-3,940	2	1	-4,244	2	1	-4,270
GAS	2	2	-4,784	1	3	-4,620	3	1	-4,960	2	1	-4,972
FER	2	2	-4,333	1	3	-4,309	1	1	-4,664	1	1	-4,684
IBR	1	1	-3,215	1	3	-3,756	1	1	-4,124*	1	1	-4,124
ABE	3	2	-4,922	1	2	-4,984	2	1	-5,196	1	1	-5,270
ENG	2	1	-5,330	3	1	-5,155	1	1	-5,343	1	1	-5,370
SAB	1	2	-5,120	1	3	-5,285	2	2	-5,313	1	3	-5,493
ACX	1	3	-4,608	1	3	-4,553	1	1	-4,813	1	1	-4,848
MAP	2	2	-4,654	1	2	-4,483	1	1	-4,772	1	1	-4,790
ELE	1	2	-5,542	3	2	-3,950	2	2	-5,713	1	1	-6,130
FCC	1	2	-4,388	1	3	-4,419	1	2	-4,738	1	1	-4,753
BME	1	3	-4,436	1	3	-4,283	1	1	-4,577	1	1	-4,599
GRF	2	1	-4,644	1	3	-4,498	1	3	-4,701	1	1	-4,761

CIN	3	2	-4,374	1	3	-4,286	1	1	-4,656	1	1	-4,674
IBLA	2	1	-4,106	1	3	-4,026	1	1	-4,227	1	1	-4,302
BKT	1	2	-4,408	1	2	-4,276	1	1	-4,590	1	1	-4,61
CABK	3	3	-4,009	1	1	-4,439	1	1	-4,611	1	1	-4,714
IDR	1	2	-5,300	3	1	-5,096	1	1	-5,335	1	1	-5,355
TL5	1	3	-4,438	1	3	-4,668	1	3	-4,852	1	1	-4,934
TRE	2	2	-4,160	1	2	-3,887	2	1	-4,309	2	1	-4,305
COL	1	1	-3,445	-	-	-	1	1	-3,685	1	1	-3,869
SCYR	1	1	-3,61	1	2	-3,628	1	2	-3,833	1	1	-3,891
SLR	1	3	-2,95	1	3	-3,319	1	1	-3,427	1	1	-3,541
PRS	1	1	-4,14	1	2	-4,231	2	3	-4,511	1	1	-4,554
AGS	1	3	-5,06	1	1	-5,692	1	3	-5,616	1	2	-5,916
BTO	1	2	-4,785	1	3	-4,842	1	1	-5,136	1	1	-5,142
ABG	1	3	-4,042	1	3	-4,011	1	1	-4,376	1	1	-4,378
OHL	1	3	-4,038	1	3	-4,022	1	1	-4,313	1	1	-4,339
SED	1	1	-3,789	1	3	-4,160	1	1	-4,308	1	1	-4,387
MVC	1	1	-5,319	-	-	-	2	3	-6,018	1	1	-6,433
ZEL	1	3	-4,044	1	3	-4,047	1	2	-4,164	1	1	-4,283
A3M	3	3	-4,009	1	3	-4,607	1	1	-4,839	1	1	-4,892
MEL	1	2	-4,069	1	2	-4,266	2	1	-4,483	2	1	-4,539
TUB	1	2	-4,301	1	2	-4,364	1	1	-4,598	1	1	-4,657
ALM	1	3	-1,949	1	2	-4,506	1	2	-4,387	1	1	-4,743
ALB	1	2	-4,552	1	2	-4,554	1	1	-4,853	1	1	-4,887
PAS	1	2	-4,393	1	3	-4,798	1	2	-4,897	1	1	-4,996
EZE	1	3	-3,123	1	3	-3,541	1	3	-3,613	1	1	-3,750
EBRO	1	3	-4,976	1	3	-4,996	3	1	-5,248	1	1	-5,306
ZOT	1	3	-4,061	1	2	-2,841	2	3	-4,251	1	1	-4,538
CEP	1	3	-7,363	1	3	-7,320	1	2	-7,433	1	2	-7,543
OLE	1	2	-4,543	1	3	-4,375	1	1	-4,722	1	1	-4,728
BVA	2	1	-4,912	1	2	-4,791	2	3	-5,014	1	1	-5,078
PSG	2	1	-4,908	3	1	-4,748	1	1	-4,935	1	1	-4,992
NH	1	1	-4,066	1	3	-4,475	3	1	-4,623	1	1	-4,701
GCO	1	2	-4,543	1	3	-4,375	1	1	-4,722	1	1	-4,728
QBT	1	3	-1,383	1	1	-2,022	3	3	-2,629	1	3	-3,305
CAF	1	1	-4,601	1	2	-4,774	1	3	-4,880	1	3	-5,027
FAE	3	1	-4,200	1	3	-4,113	1	3	-4,355	1	3	-4,407
VIS	1	2	-5,125	3	1	-4,931	1	2	-5,132	1	2	-5,154
ENC	1	3	-3,934	1	3	-4,206	3	1	-4,447	1	1	-4,536
VOC	3	1	-4,156	1	2	-4,805	1	3	-4,877	1	1	-5,037
GALQ	1	2	-3,919	1	3	-4,171	1	1	-4,375	1	1	-4,416
REY	3	3	-5,749	3	3	-5,355	2	2	-6,087	1	1	-6,546

PAC	3	2	-4,462	1	3	-4,527	1	1	-4,749	1	1	-4,791
ERC	2	1	-6,328	2	2	-6,395	1	1	-6,356	2	1	-6,547
REN	3	1	-3,891	1	2	-3,743	1	1	-4,023	1	1	-4,077
CPL	3	1	-4,203	1	3	-4,559	1	1	-4,636	2	1	-4,787
VLG	1	2	-2,659	2	1	-2,921	3	1	-3,569	1	1	-3,894
TRG	1	2	-4,316	1	3	-4,344	1	1	-4,634	1	1	-4,659
SPS	3	3	-3,948	1	2	-4,312	2	1	-4,419	2	1	-4,551
FRS	3	1	-3,638	1	2	-3,827	1	2	-3,777	1	1	-4,020
MLX	1	3	-4,122	1	3	-4,238	2	1	-4,427	1	1	-4,499
RLIA	3	1	-4,755	1	1	-4,877	1	1	-5,071	1	1	-5,158
CPF	1	3	-4,853	1	2	-5,077	1	3	-5,167	1	1	-5,335
CDR	2	3	-3,462	2	1	-3,459	1	2	-4,021	1	1	-4,379
AMP	3	2	-4,623	1	3	-4,835	3	1	-4,994	1	1	-5,077
BDL	2	1	-5,222	1	2	-4,617	2	1	-5,293	2	1	-5,476
PSL	3	3	-3,920	1	2	-4,921	2	3	-5,055	1	1	-5,167
ITI	3	2	-4,164	3	3	-4,260	3	3	-4,790	2	2	-5,080
DERM	1	3	-3,905	1	3	-4,157	1	2	-4,193	1	1	-4,358
ROVI	2	2	-4,280	1	2	-4,443	3	1	-4,601	1	1	-4,686
SNC	1	1	-3,779	2	1	-3,623	2	3	-4,170	1	1	-4,334
UBS	2	2	-3,016	1	1	-2,925	1	3	-3,743	1	3	-3,923
MDF	1	3	-4,692	1	3	-4,811	1	3	-5,012	1	2	-5,064
CBAV	2	1	-3,985	1	1	-4,340	1	3	-4,463	1	1	-4,574
VID	1	2	-4,802	1	2	-4,788	1	3	-4,906	1	2	-5,020
PVA	1	3	-5,128	1	3	-5,168	1	3	-5,185	2	3	-5,354
MTB	2	3	-3,418	3	2	-3,236	2	3	-4,003	1	3	-4,241
FDR	1	1	-4,096	1	3	-4,201	1	3	-4,406	1	1	-4,468
NAT	2	1	-4,441	1	3	-5,152	1	2	-5,290	1	1	-5,390
CIE	1	3	-4,440	1	3	-4,670	1	3	-4,760	1	3	-4,884
GUI	1	3	-5,029	1	3	-5,341	1	2	-5,363	1	1	-5,497
R4	1	1	-5,162	1	2	-5,061	2	2	-6,158	2	1	-6,425
AZK	1	3	-4,404	1	3	-4,719	1	1	-4,892	1	1	-4,973
VER	3	1	-3,463	1	3	-3,564	3	3	-3,701	1	1	-3,771
TEC	1	3	-4,202	1	3	-4,454	2	1	-4,450	1	2	-4,667
CCMC	1	1	-5,008	3	2	-3,575	1	1	-5,004	1	3	-5,430
ESZ	1	2	-3,066	1	3	-3,432	1	1	-3,565	1	1	-3,690

Tabla AT2 – Tabla de comparativa de Gach-apARCH con los resultados de selección de los mejores modelos para las series sintéticas.

	apARCH						GARCH					
	Normal			T-Student			Normal			T-Student		
Valor	p	q	AIC	p	q	AIC	p	q	AIC	p	q	AIC
Sector 1	1	2	-5,449	1	1	-5,682	2	1	-6,191	3	1	-6,156
Sector 2	3	1	-5,304	1	2	-5,276	1	1	-5,797	1	1	-5,777
Sector 3	2	1	-5,462	1	2	-5,617	1	1	-6,105	2	3	-6,051
Sector 4	1	2	-4,451	1	1	-5,063	1	1	-5,542	1	3	-5,502
Sector 5	3	1	-5,213	1	1	-4,936	1	1	-5,414	1	1	-5,395
Sector 6	2	1	-5,451	1	2	-5,123	1	1	-5,641	1	1	-5,621

Listado de sectores y subsectores:

## 1. Petróleo y Energía

1.1 Petróleo

1.2 Electricidad y Gas

1.3 Agua y Otros

1.4 Energías Renovables

## 2. Mat. Básicos, Industria y Construcción

2.1 Mineral, Metales y Transformación

2.2 Frabricación y Montaje Bienes de Equipo

2.3 Construcción

2.4 Materiales de Construcción

2.5 Industria Química

2.6 Ingeniería y Otros

## 3. Bienes de Consumo

3.1 Alimentación y Bebidas

3.2 Textil, Vestido y Calzado

3.3 Papel y Artes Gráficas

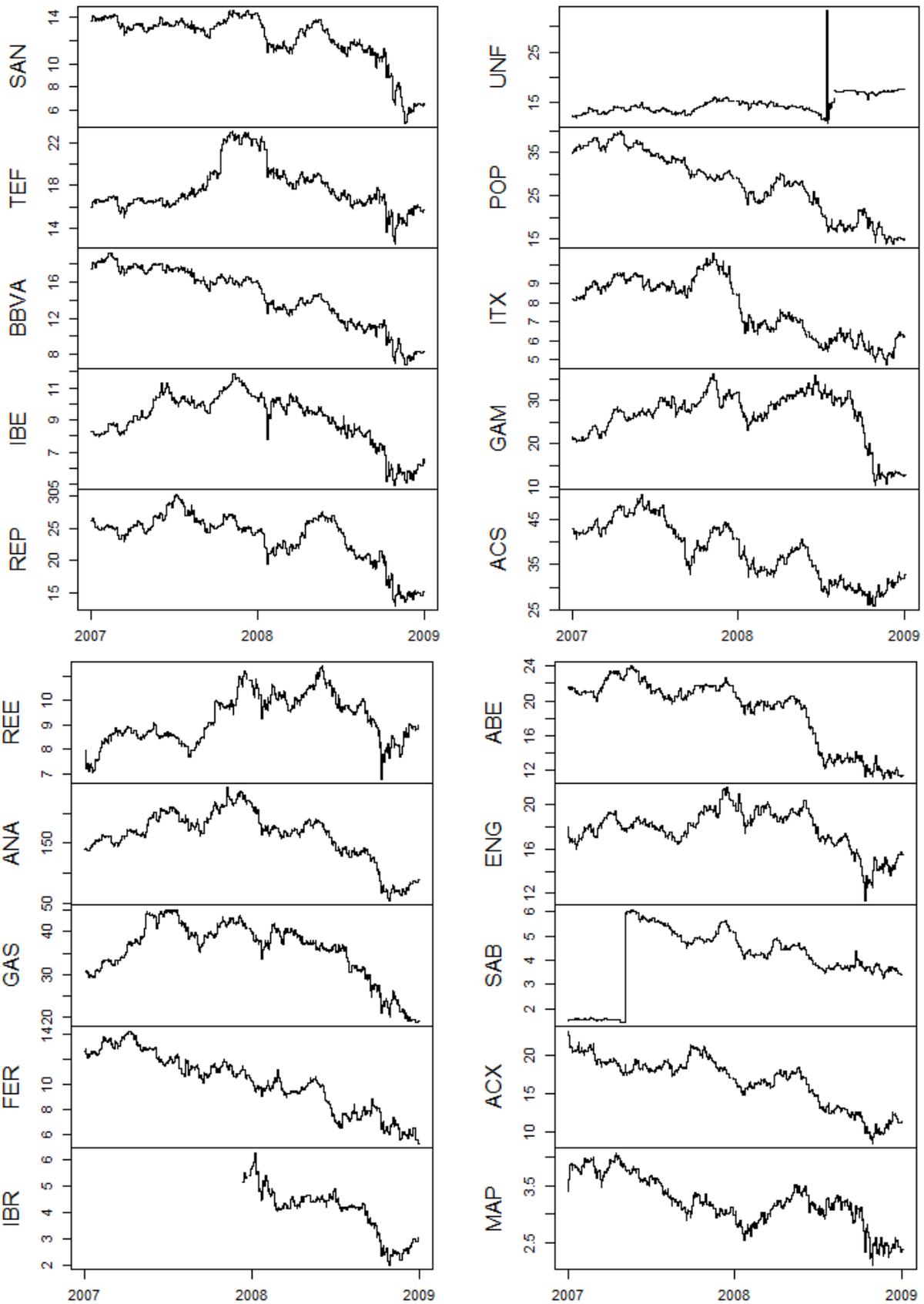
3.5 Productos Farmacéuticos y Biotecnología

3.6 Otros Bienes de Consumo

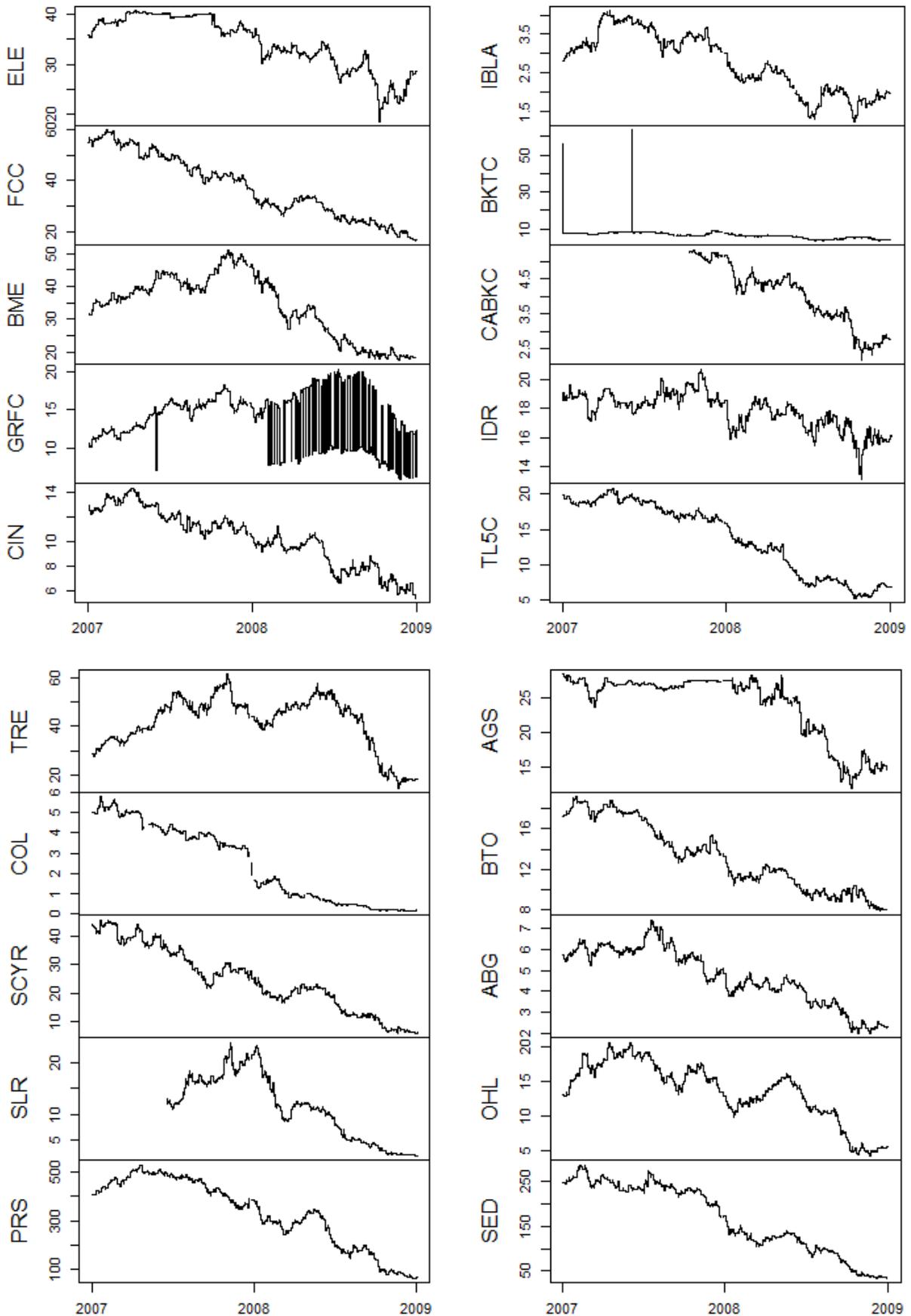
## 4. Servicios de Consumo

- 4.1 Ocio, Turismo y Hostelería
- 4.2 Comercio
- 4.3 Medios de Comunicación y Publicidad
- 4.4 Transporte y Distribución
- 4.5 Autopistas y Aparcamientos
- 4.6 Otros Servicios
- 5. Servicios Financieros e Inmobiliarios
  - 5.1 Bancos y Cajas de Ahorro
  - 5.2 Seguros
  - 5.3 Cartera y Holding
  - 5.4 SICAV
  - 5.5 Inmobiliarias y Otros
  - 5.6 Servicios de Inversión
- 6. Tecnología y Telecomunicaciones
  - 6.1 Telecomunicaciones y Otros
  - 6.2 Electrónica y Software

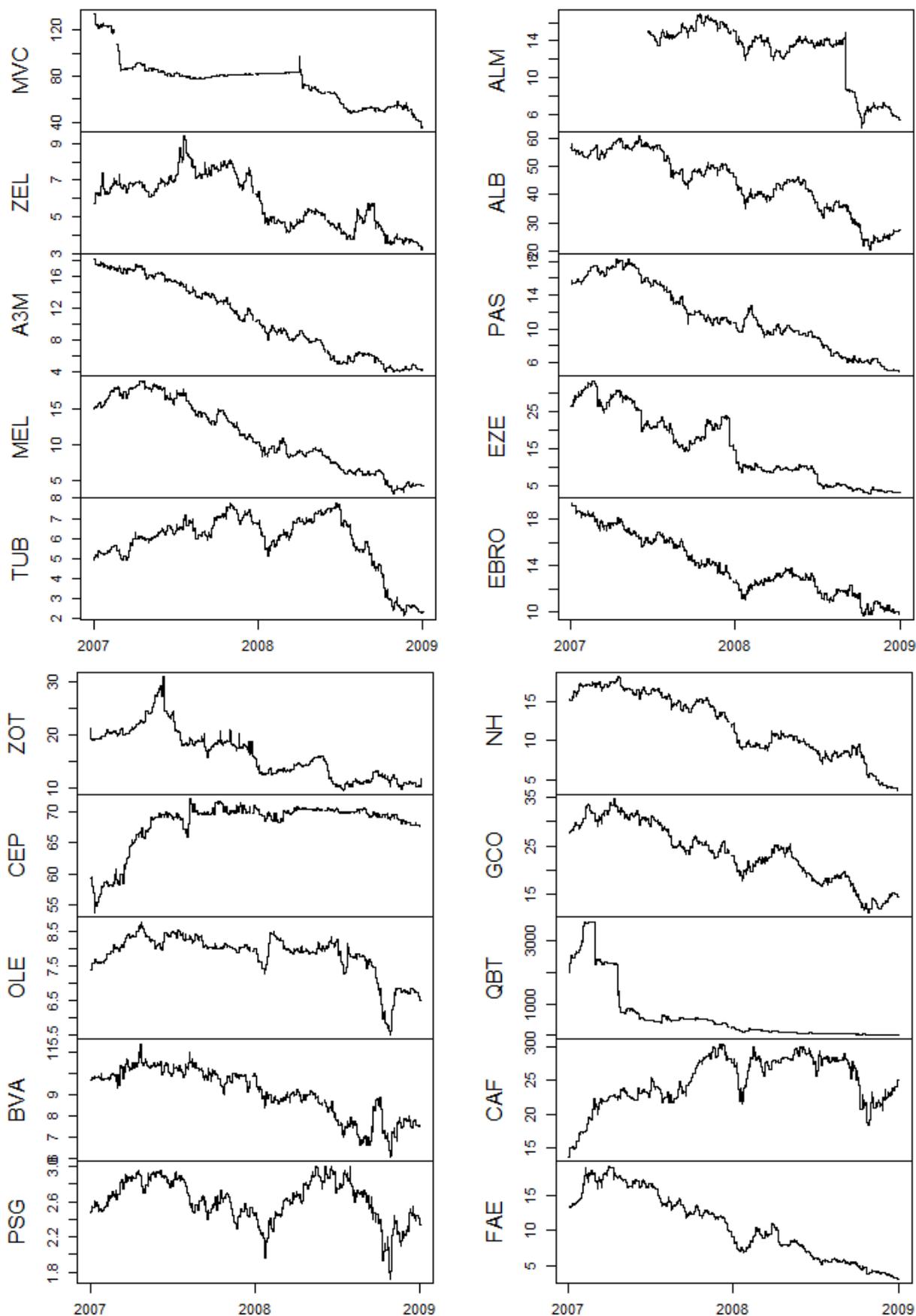
Gráficos A1 a A20 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.



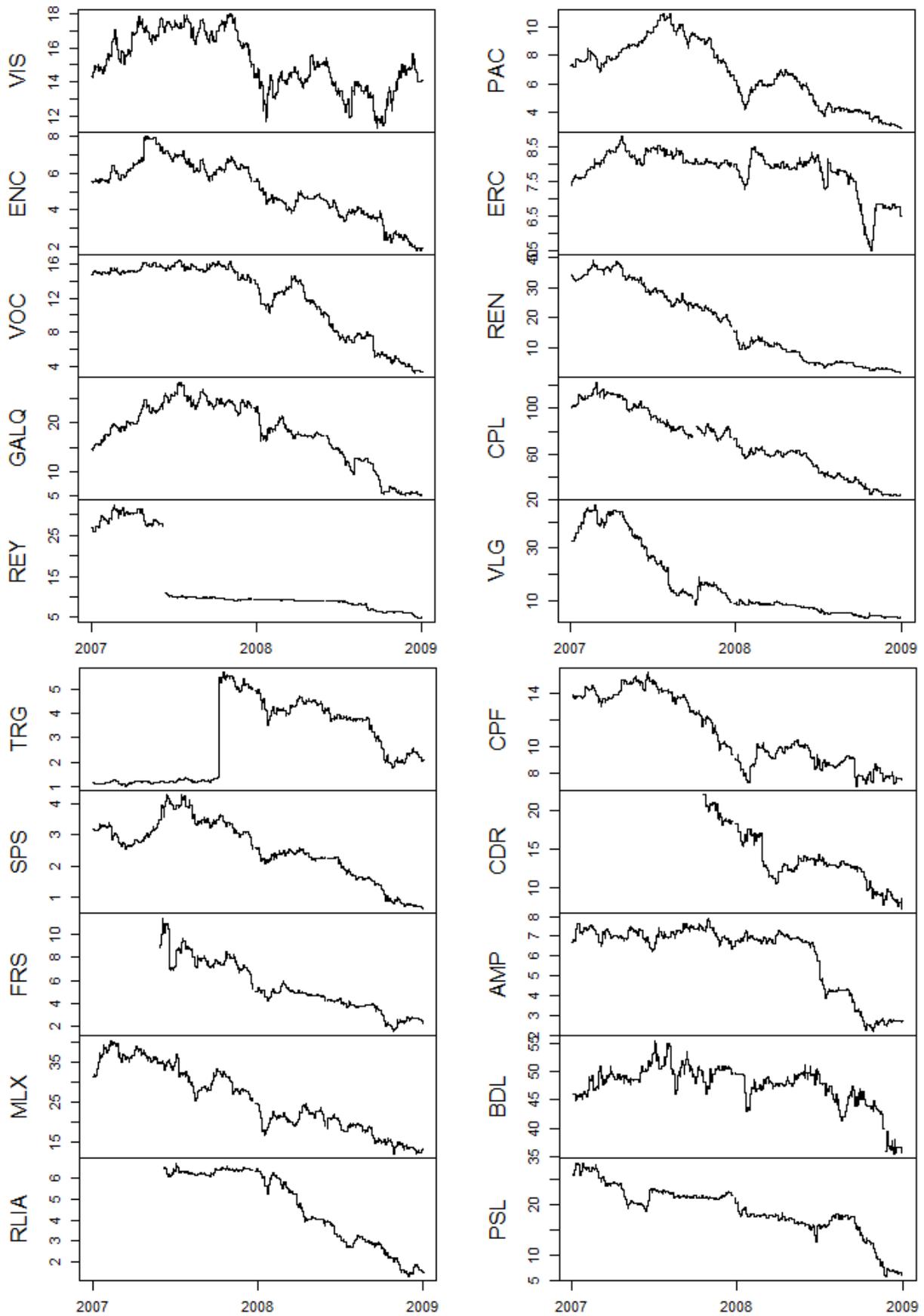
Gráficos A21 a A40 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.



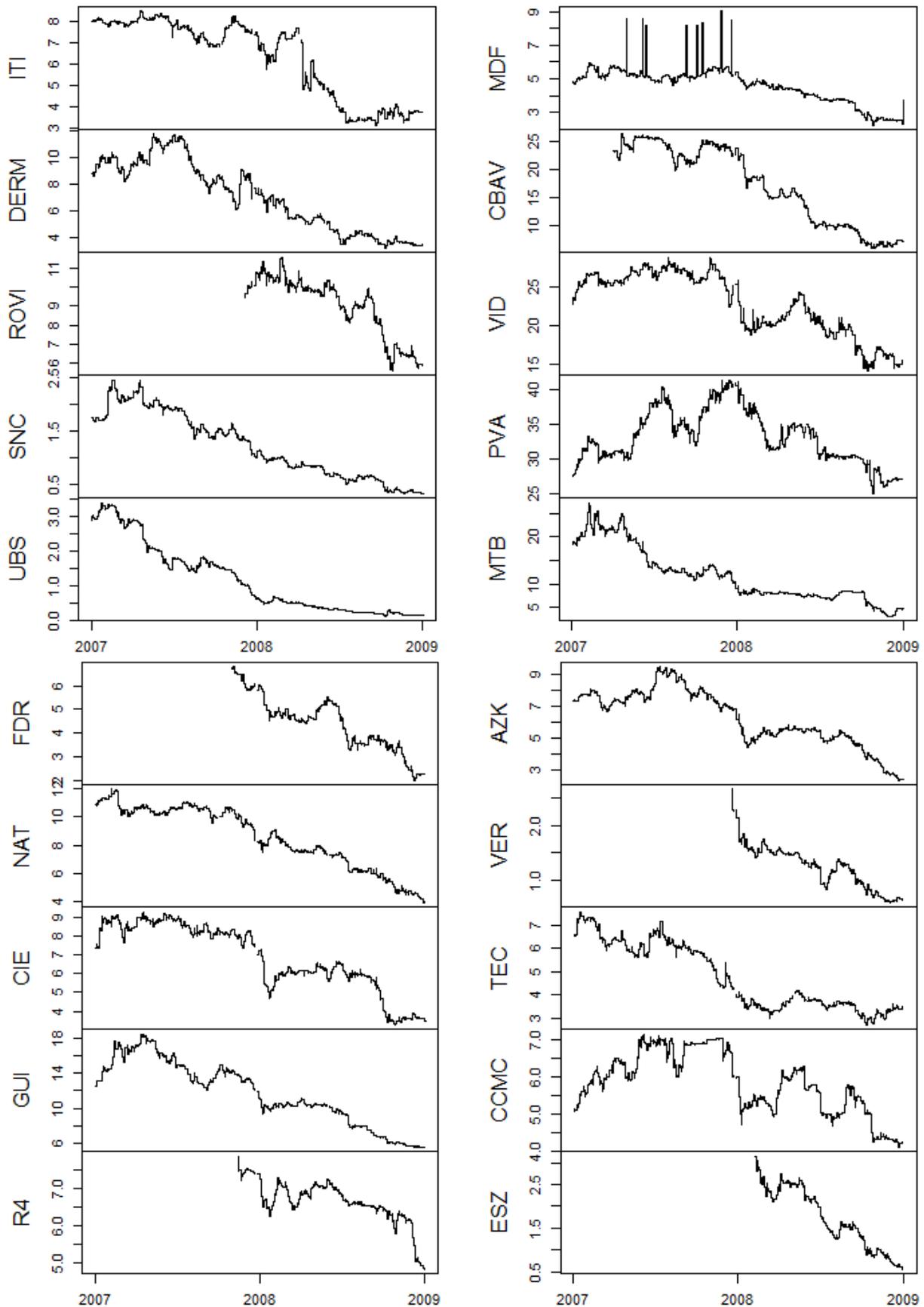
Gráficos A41 a 60 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.



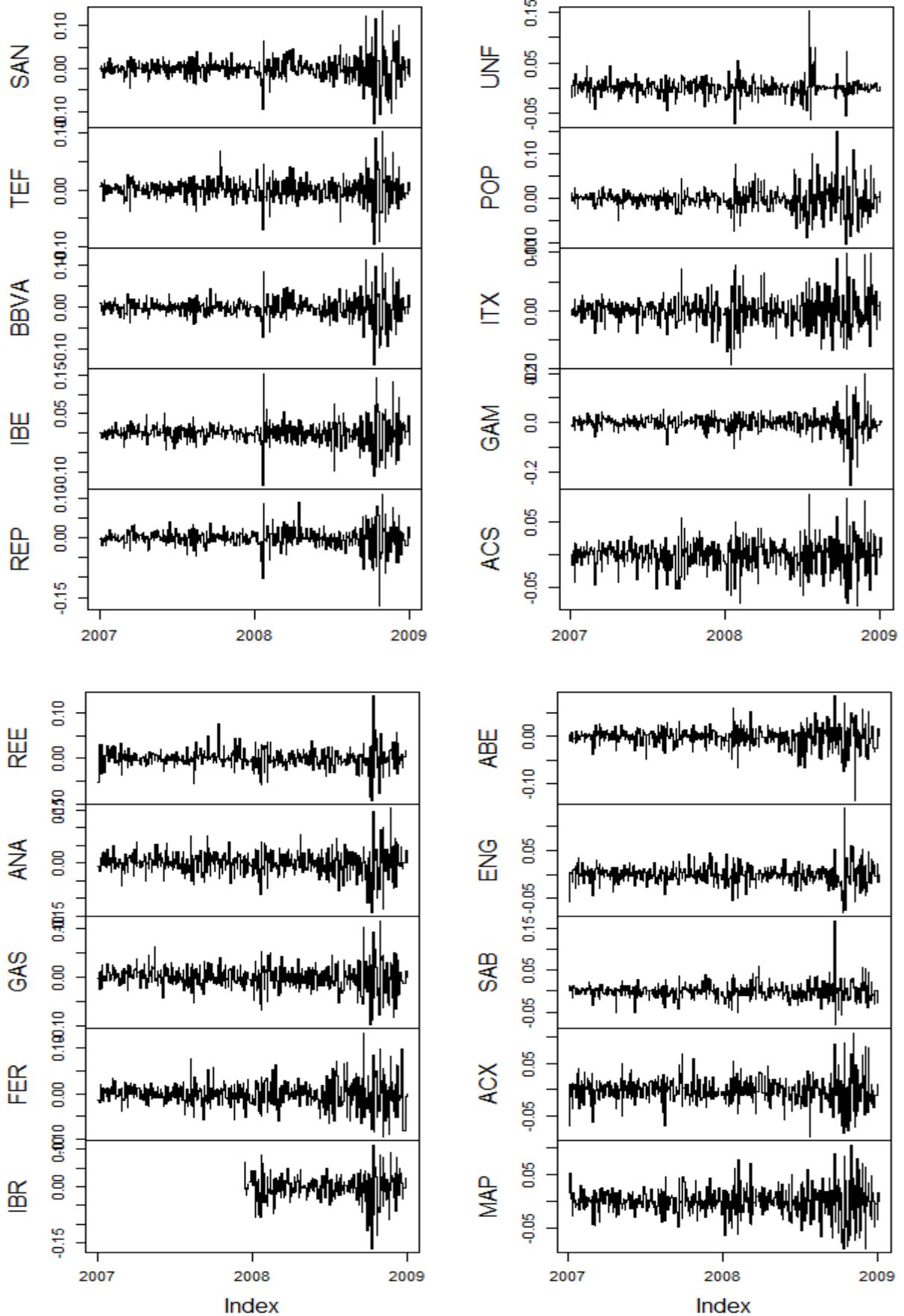
Gráficos A61 a 80 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.



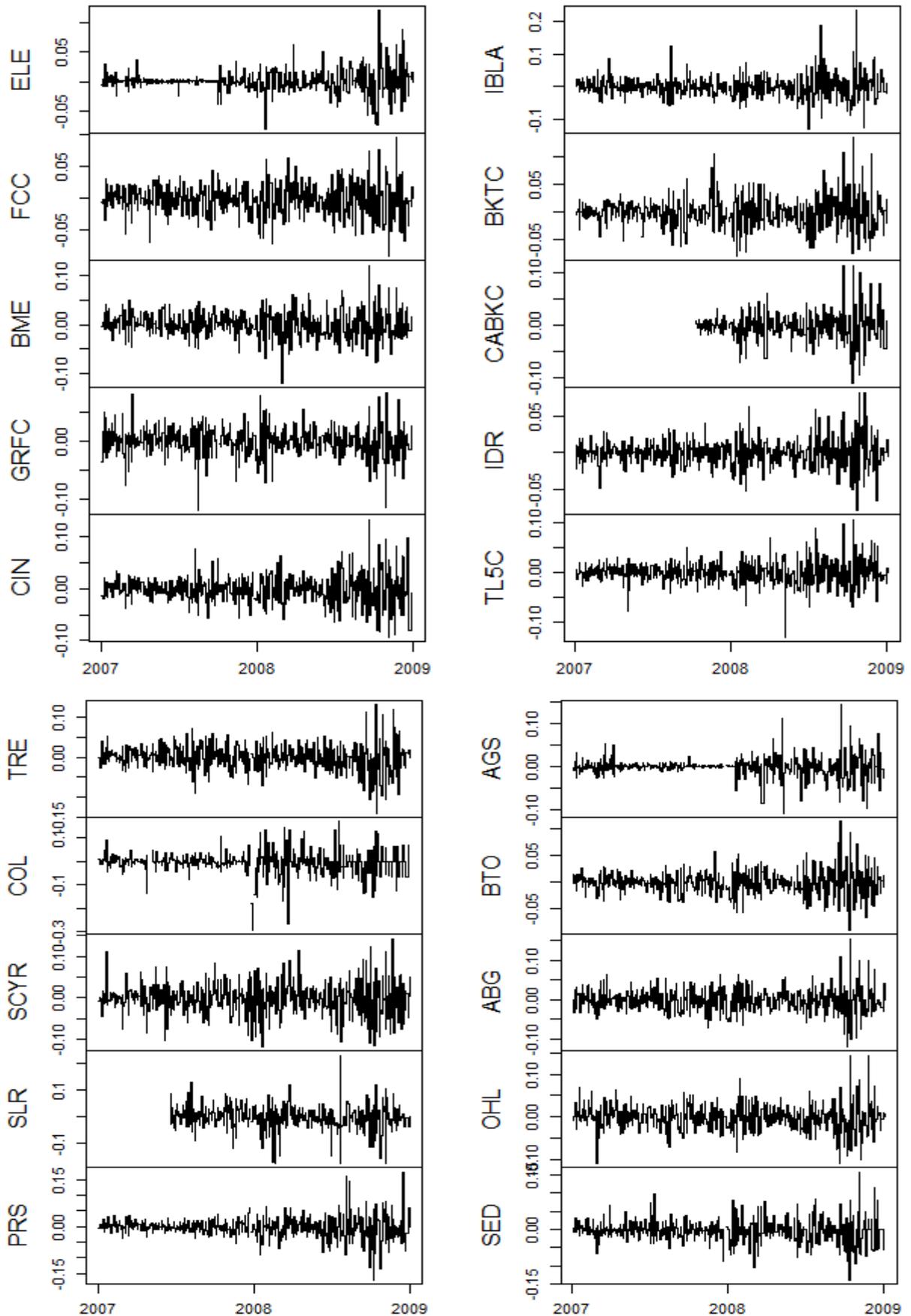
Gráficos A81 a 100 – Representación gráfica de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores, con anomalías.



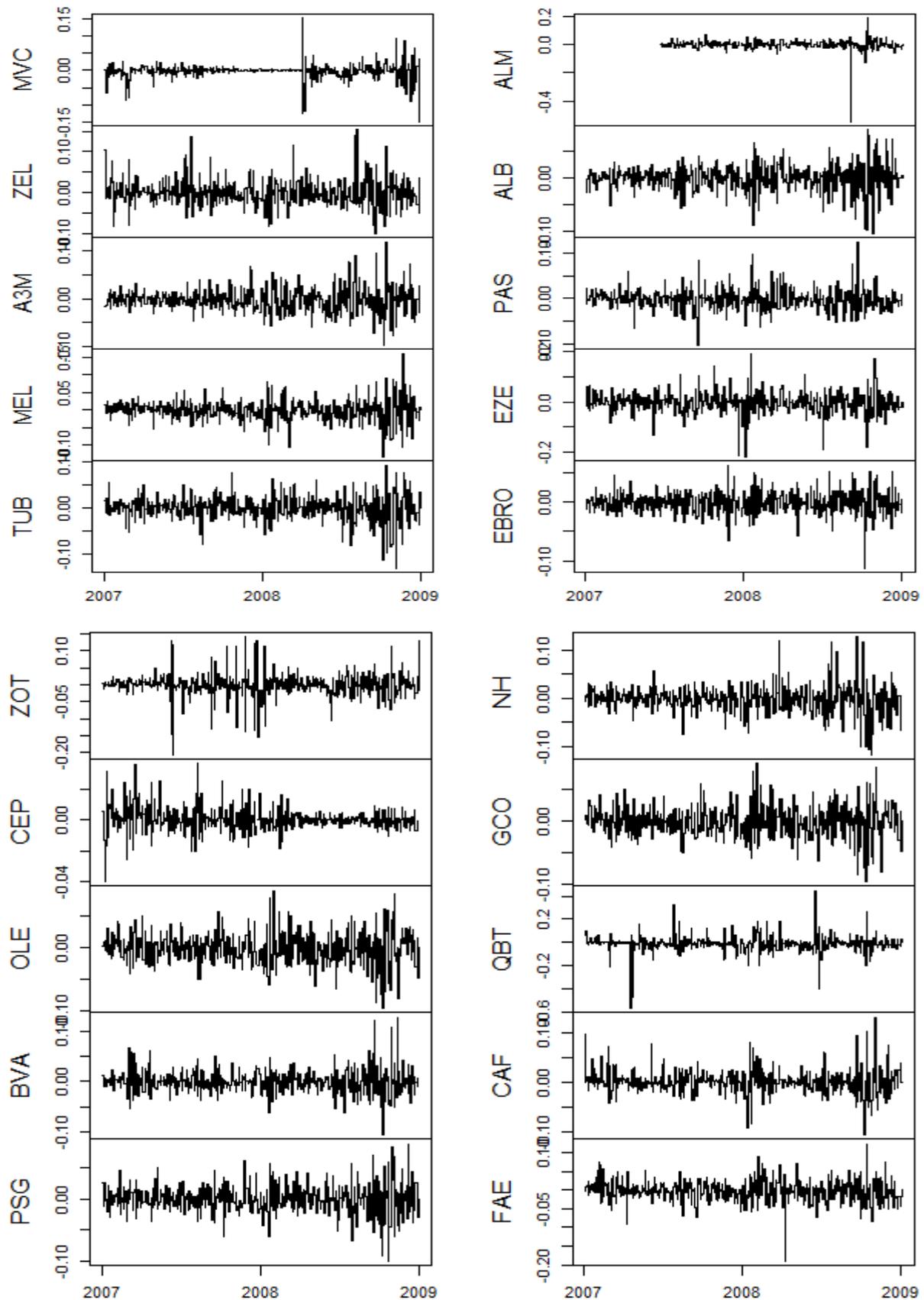
Gráficos A101 a 120 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.



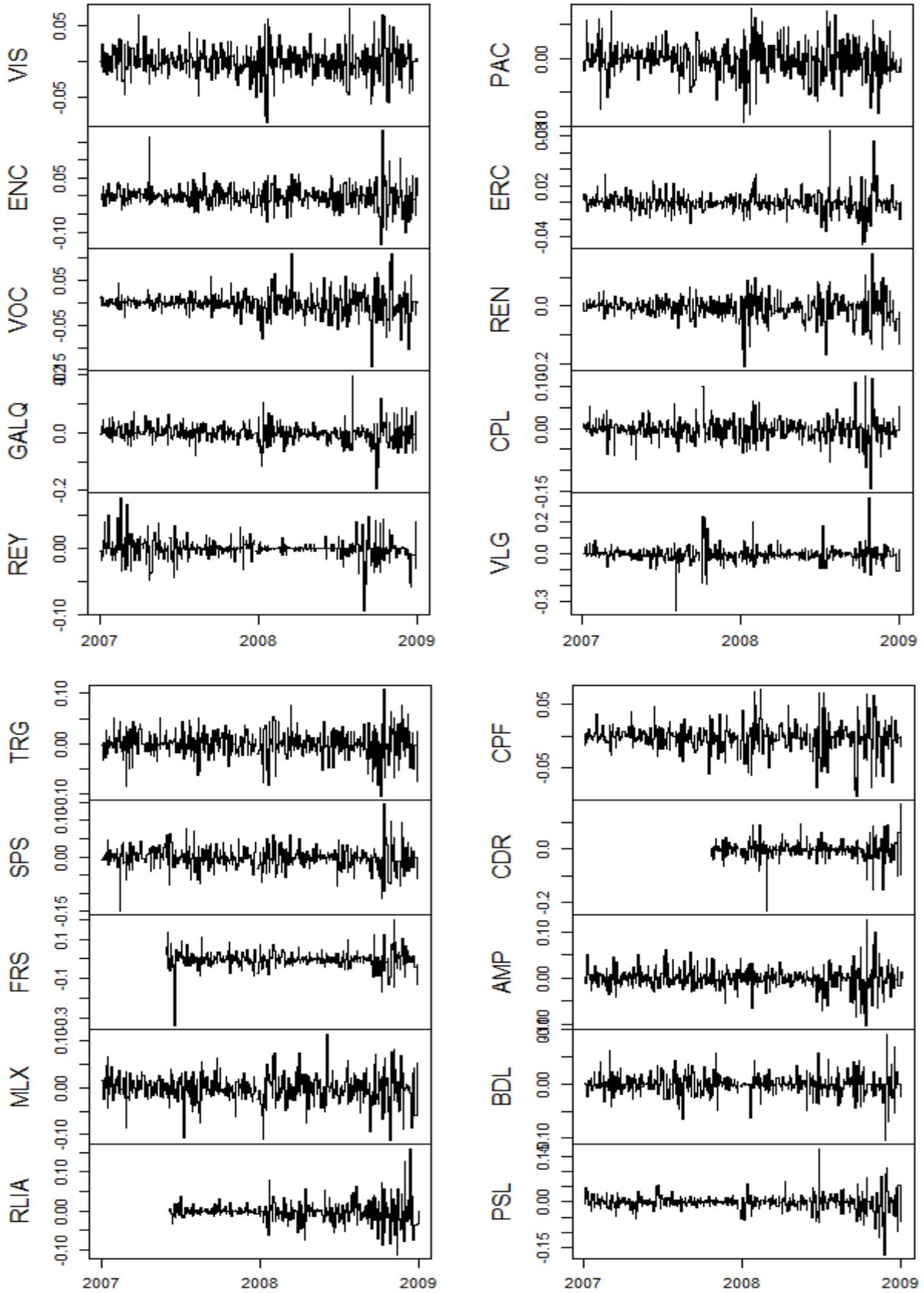
Gráficos A121 a 140 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.



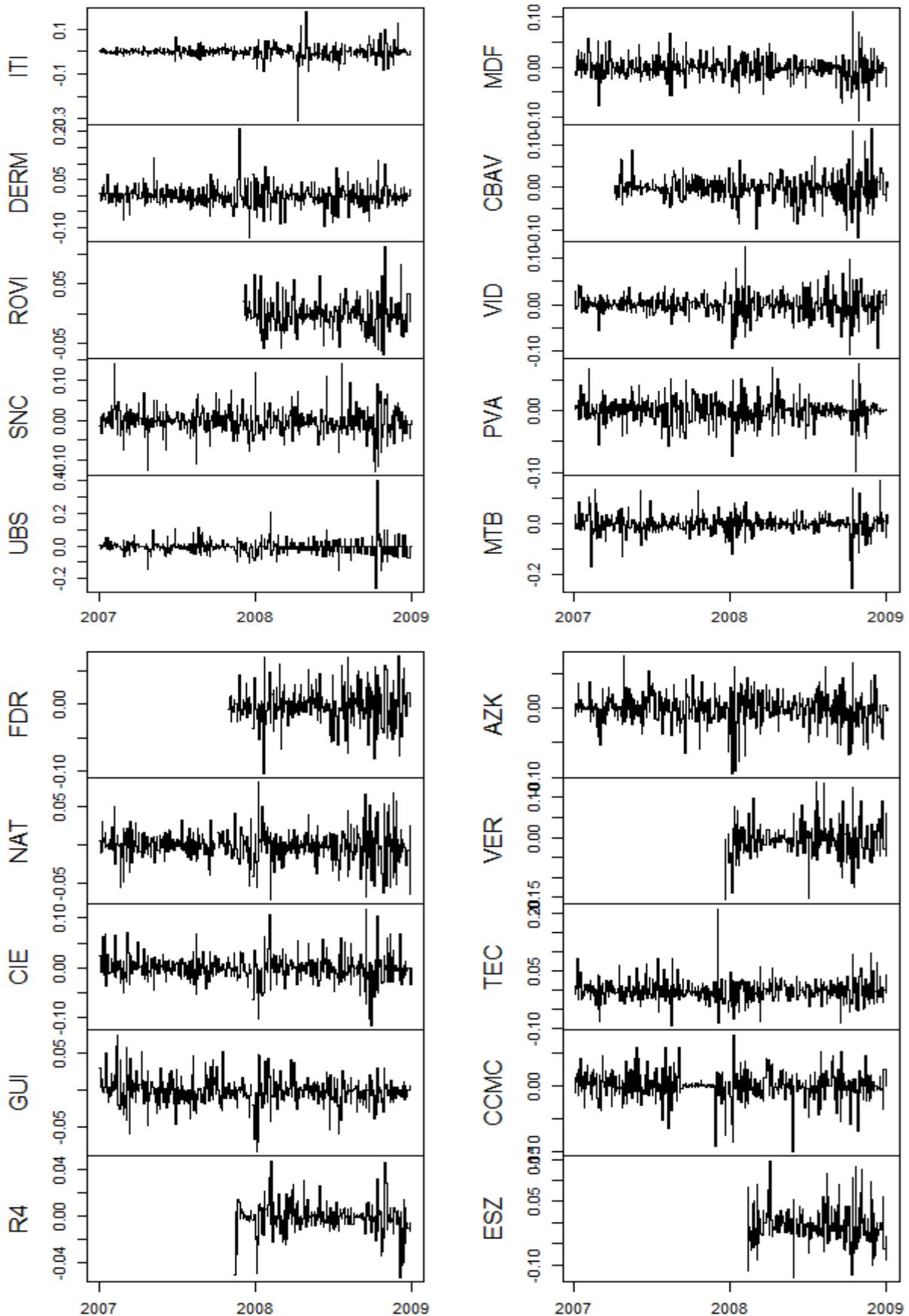
Gráficos A141 a 160 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.



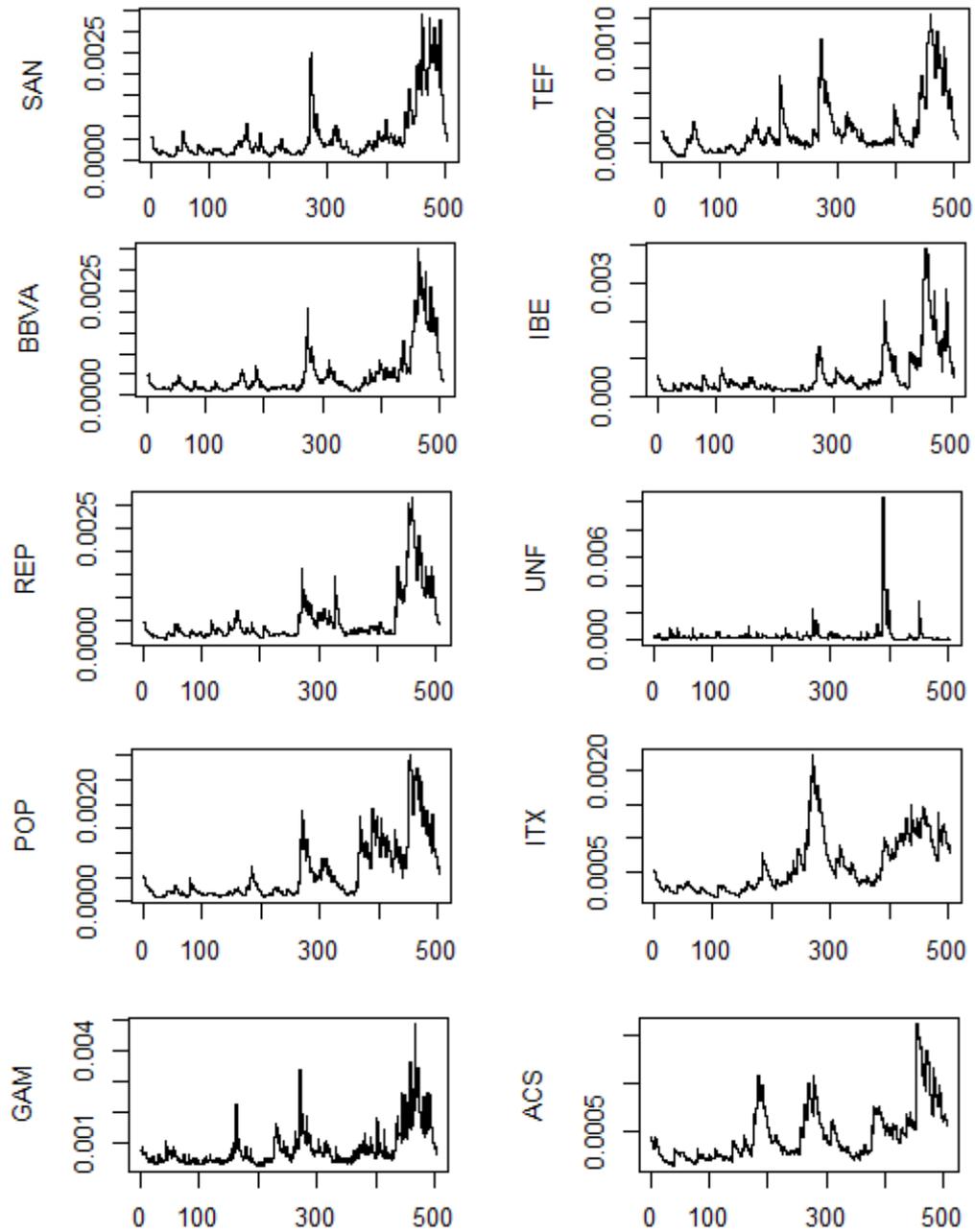
Gráficos A161 a 180 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.



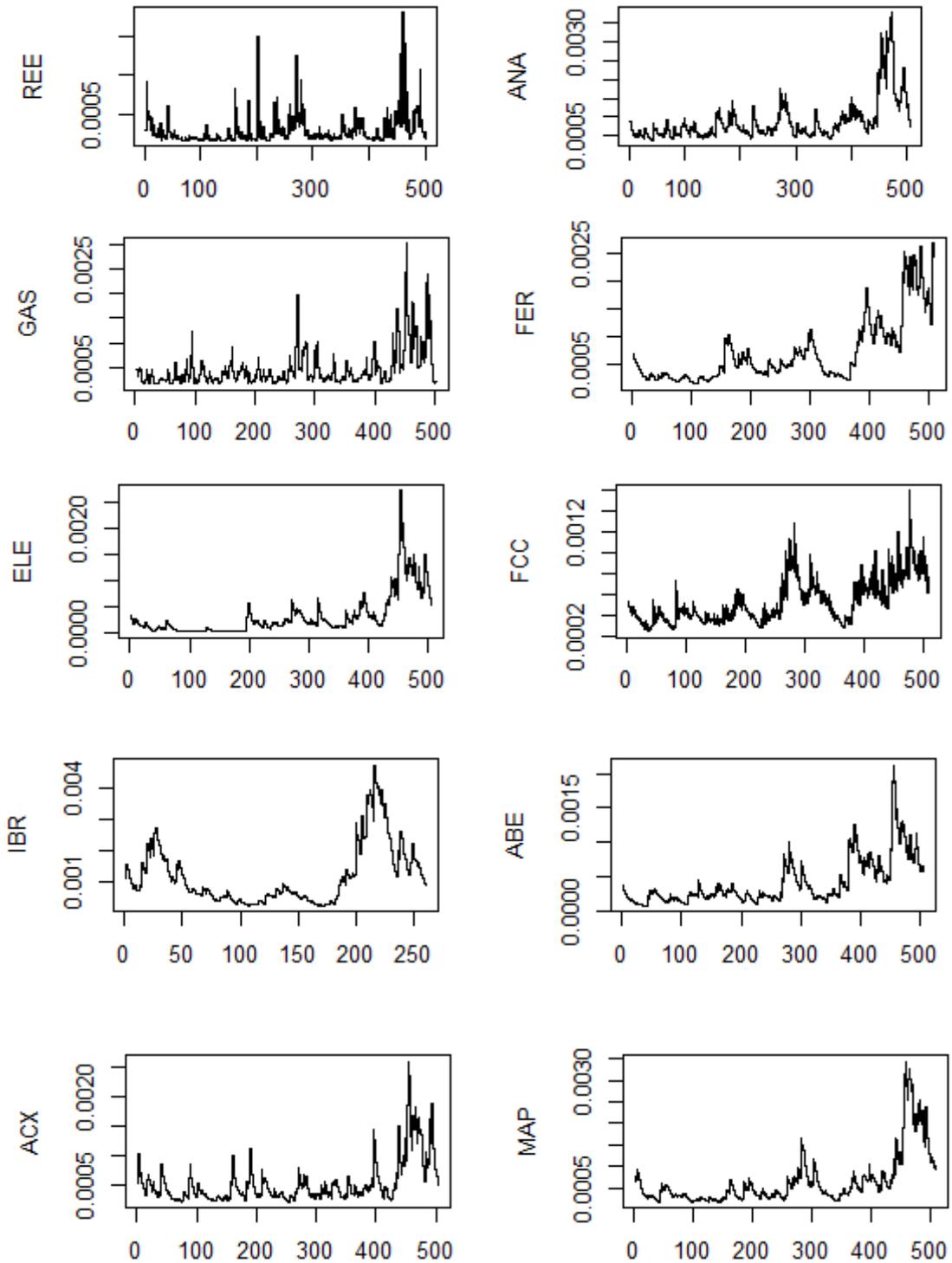
Gráficos A181 a 200 – Representación gráfica de los rendimientos de las series de cotizaciones diarias, generadas a partir del precio a cierre de la cotización de los valores.



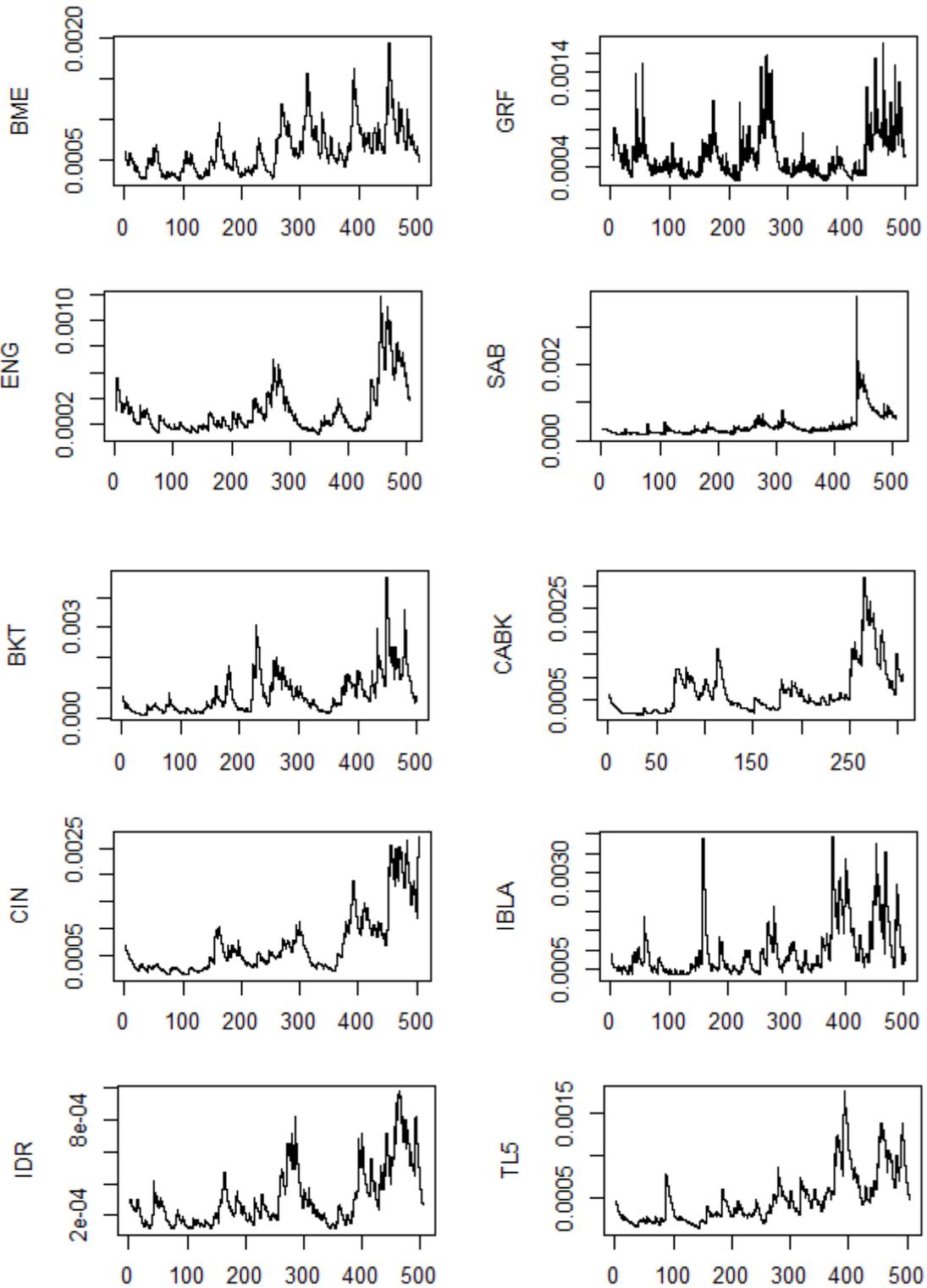
Gráficos A201 a 210 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



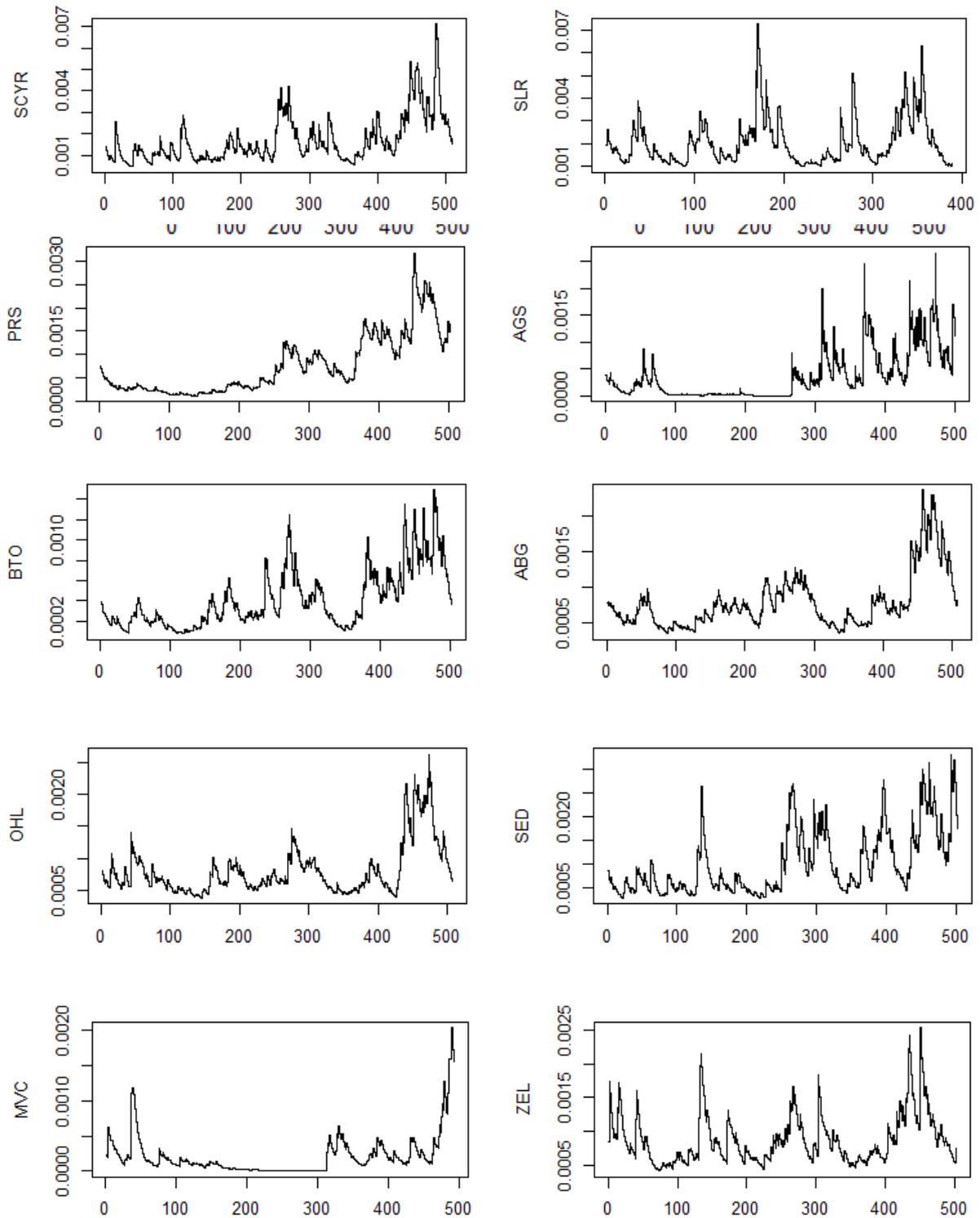
Gráficos A211 a 220 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



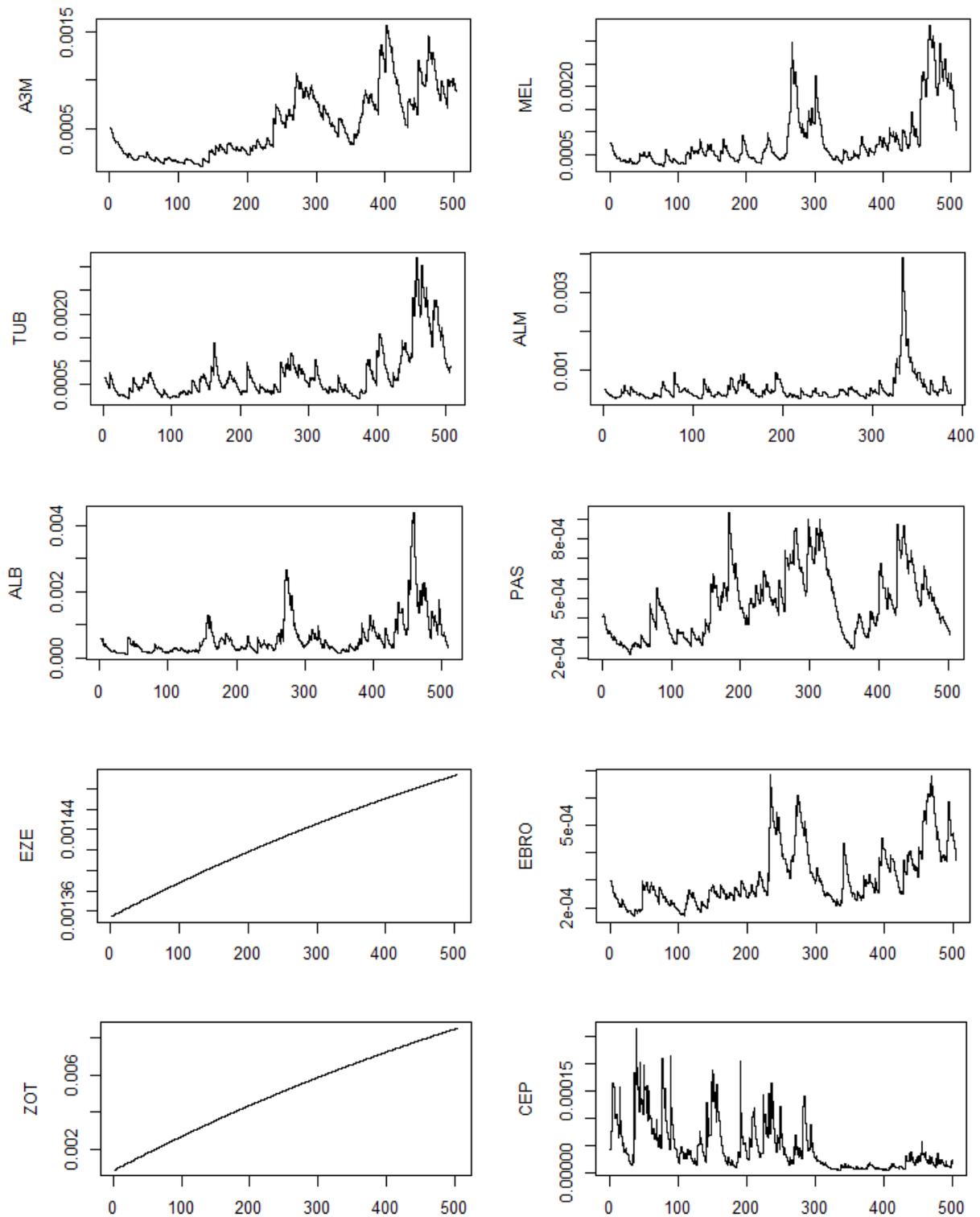
Gráficos A221 a 230 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



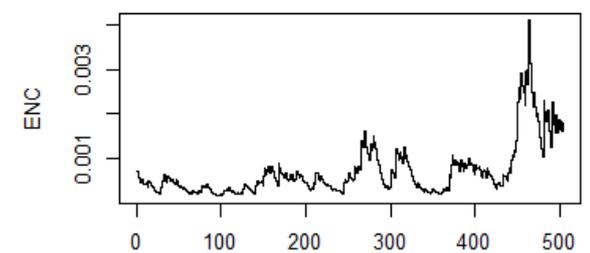
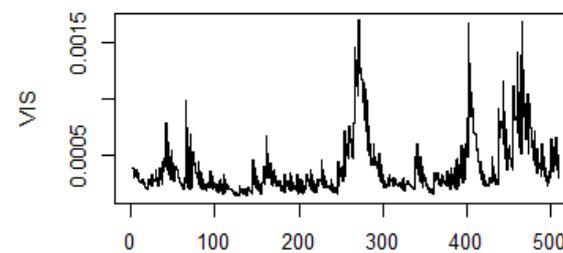
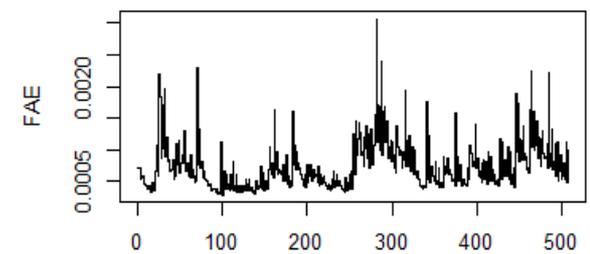
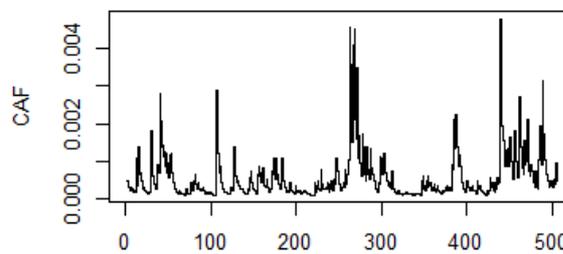
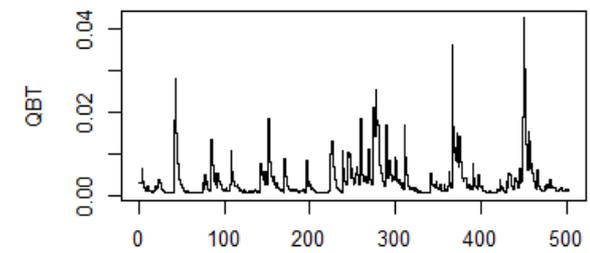
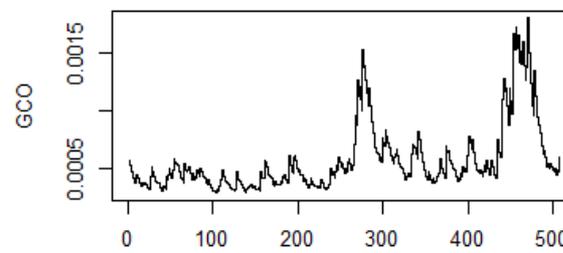
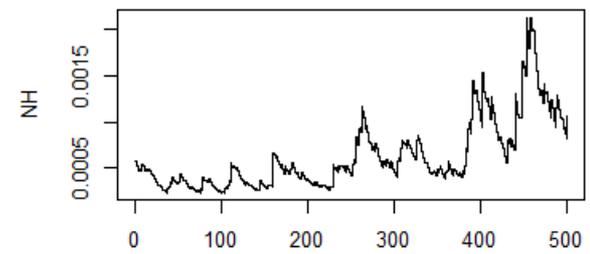
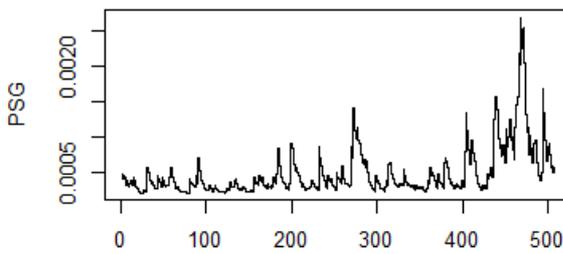
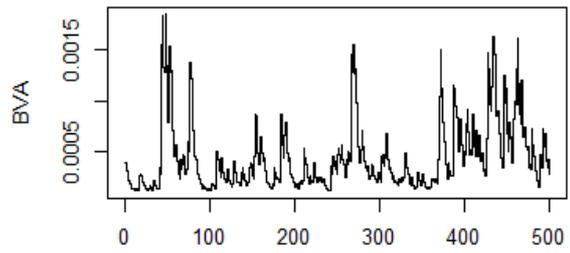
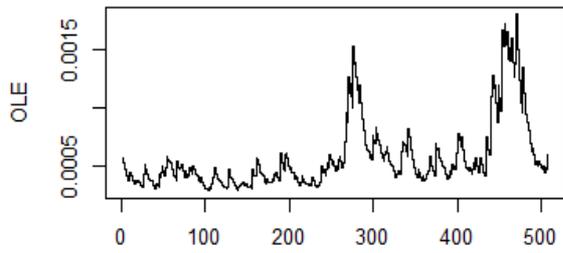
Gráficos A231 a 240 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



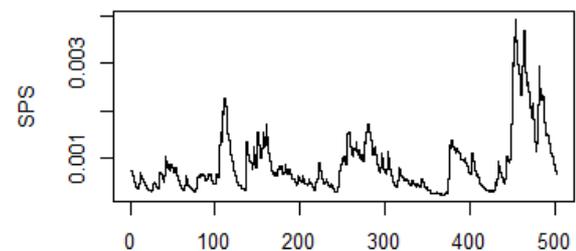
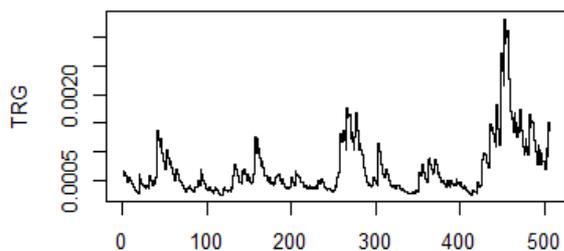
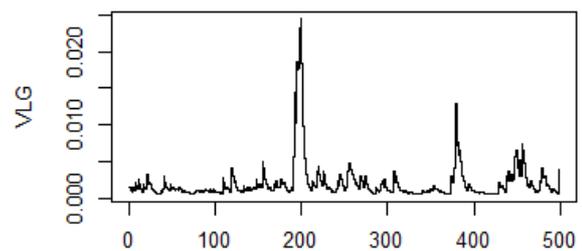
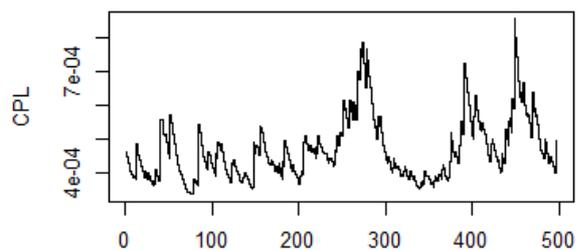
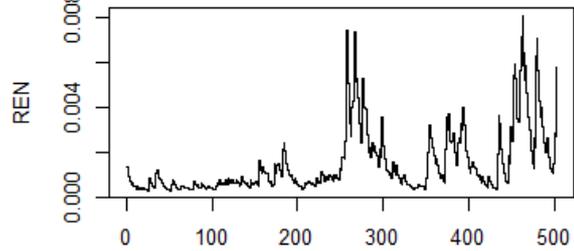
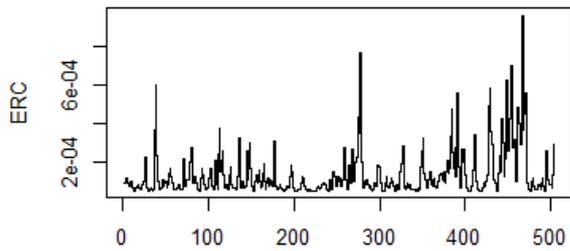
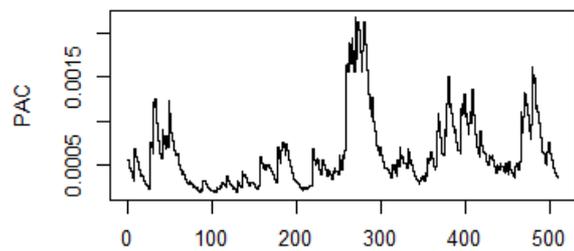
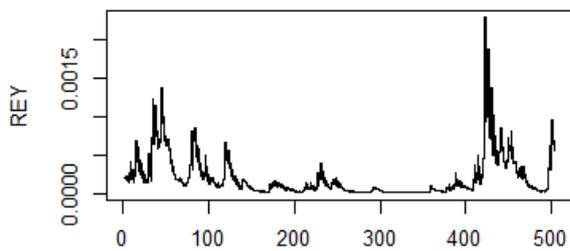
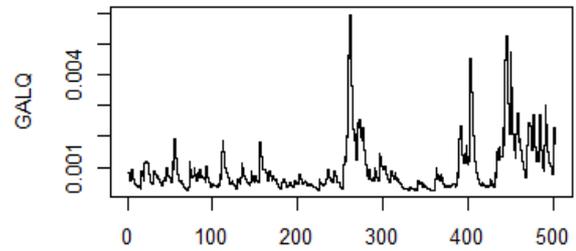
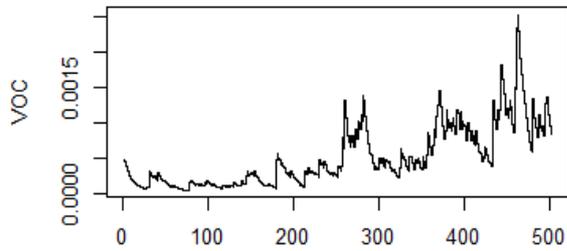
Gráficos A241 a 250 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



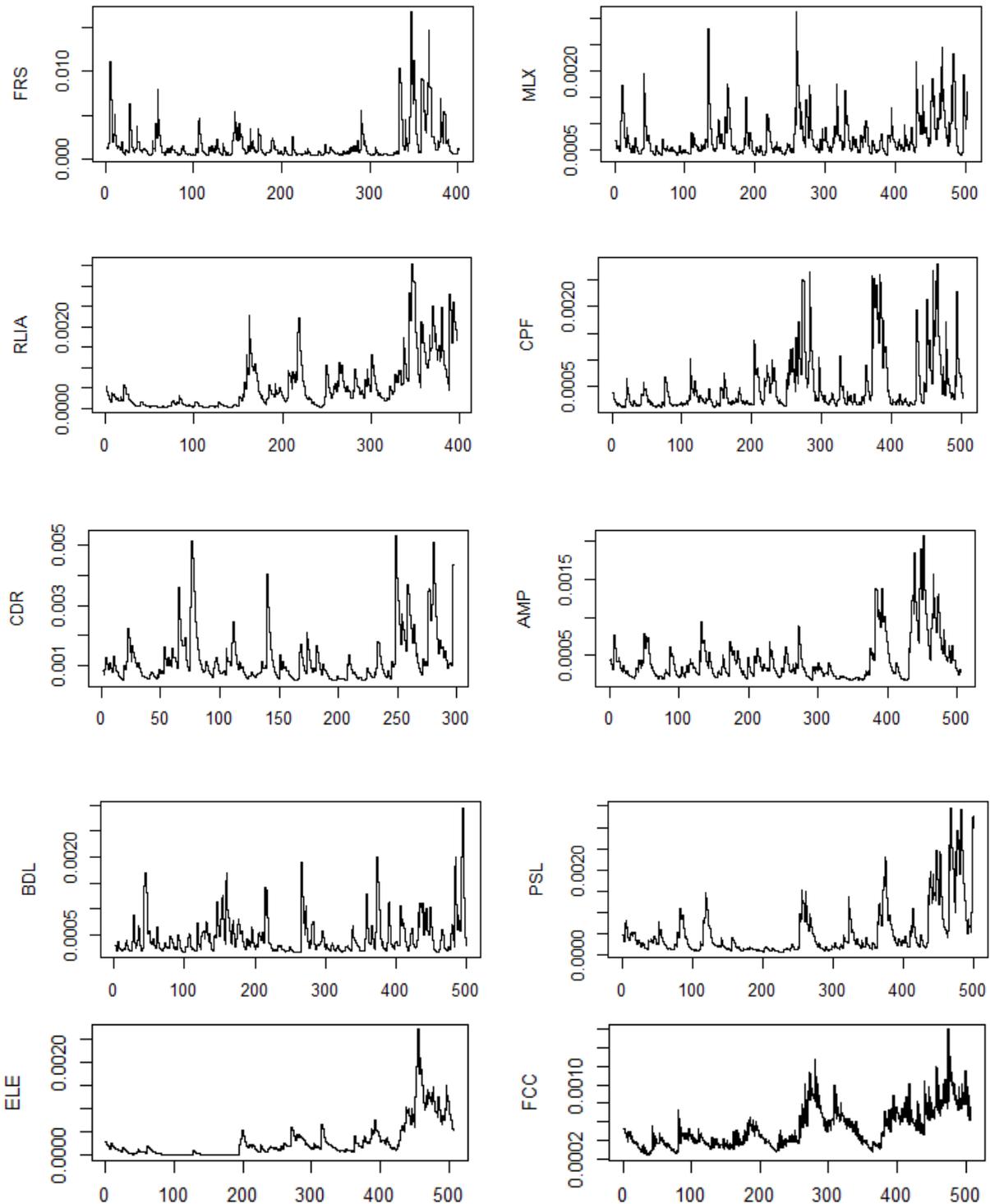
Gráficos A251 a 260 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



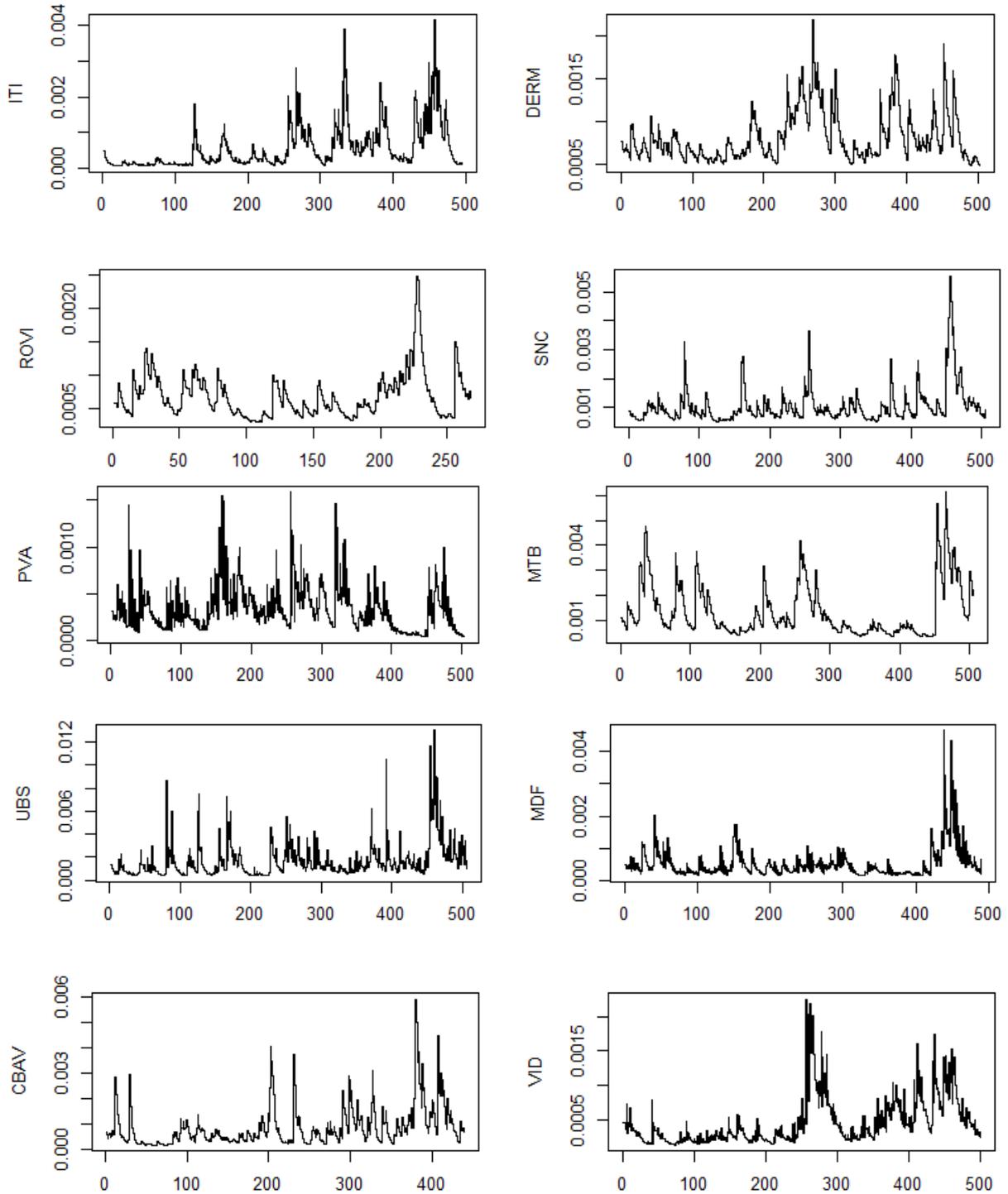
Gráficos A261 a 270 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



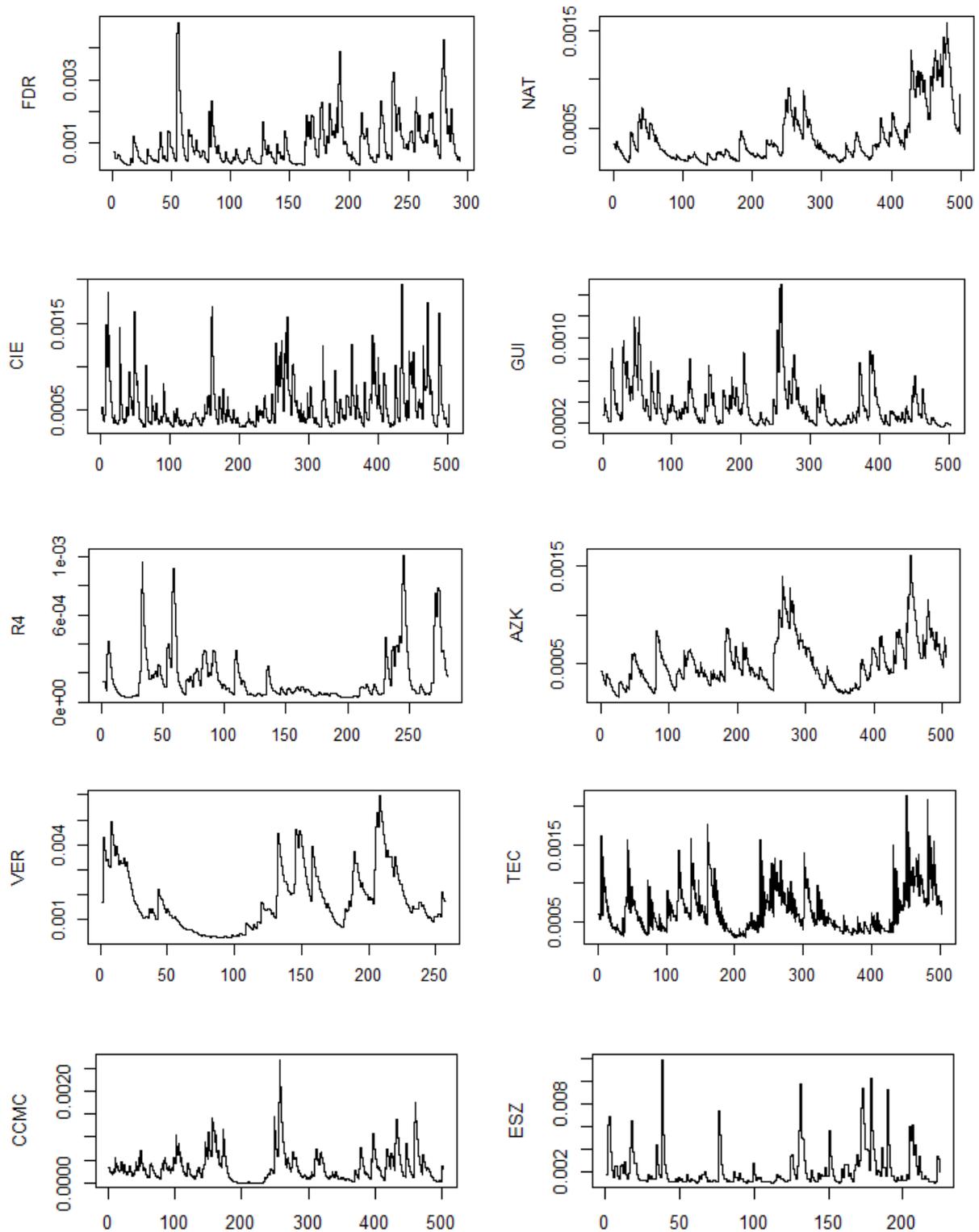
Gráficos A271 a 280 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



Gráficos A281 a 290 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



Gráficos A291 a 300 – Representación gráfica de la varianza estimada por los modelos GARCH con mejor comportamiento.



## Código R

El código utilizado es bastante amplio, ya que se han ido realizando muchas pruebas. Algunas de las líneas están incluidas en este apartado.

### Selección de los mejores modelos.

Funciones que se utilizan para la selección del mejor modelo:

```
calculaMejorModelo = function(modelo,distribucion,delta){
  ptm <- proc.time()
  mejoresAparch<-data.frame(matrix(ncol=5,nrow = length(tickers)),
                              stringsAsFactors=FALSE)
  names(mejoresAparch)<-c("Valor","p", "q","AIC","MSE")
  print(delta)
  for(i in 1:length(tickers)){
    mejoresAparch[i,1]<-tickers[i]
    if(as.integer(delta)==2)
    {
      auxGarch <-lanzaModeloConDelta(eval(as.name(paste(tickers[i],"REN",sep="")))),
                                    modelo,distribucion)
    } else{
      auxGarch <-lanzaModelo(eval(as.name(paste(tickers[i],"REN",sep="")))),
                            modelo,distribucion)
    }
    mejoresAparch[i,2]<-auxGarch[1]
    mejoresAparch[i,3]<-auxGarch[2]
    mejoresAparch[i,4]<-auxGarch[3]
    mejoresAparch[i,5]<-auxGarch[4]

    # print(paste(tickers[i],auxGarch[1],":",auxGarch[2],":",round(auxGarch[3],digits =
6),":",auxGarch[4],":",modelo, sep=""))
  }
  print(proc.time() - ptm)
  return(mejoresAparch)
```

```
}
```

```
lanzaModeloConDelta = function(rendimientos,modelo,distribucion){  
  miGarch<-c(0,0,999999,0)  
  ctrl = list(delta = 2)  
  for (i in 1:3) {  
    for (j in 1:3) {  
      spec = ugarchspec(mean.model= list(armaOrder= c(0,0)),  
                        variance.model= list(garchOrder= c(i,j), model = modelo),  
                        distribution.model= distribucion )  
  
      fit = tryCatch( ugarchfit(spec, rendimientos, solver = "solnp", solver.control = ctrl),  
                    error=function( err ) FALSE, warning=function( warn ) FALSE )  
  
      if( !is.logical( fit ) ){  
        if (as.double(Infocriterio(fit)[1])<as.double(miGarch[[3]])) {  
          mse<-0 #el mse se utilizaba en la primera versión de prueba antes de usar el AIC,  
                #posteriormente se deja de usar y se reemplaza por 0 para aligerar los cálculos  
          miGarch<-c(as.integer(i),as.integer(j),Infocriterio(fit)[1],mse)  
  
        }  
      }  
    }  
  }  
  return(miGarch)  
}
```

```
lanzaModelo = function(rendimientos,modelo,distribucion){  
  miGarch<-c(0,0,999999,0)  
  for (i in 1:3) {  
    for (j in 1:3) {  
      spec = ugarchspec(mean.model= list(armaOrder= c(0,0)),  
                        variance.model= list(garchOrder= c(i,j), model = modelo),  
                        distribution.model= distribucion )
```

```

fit = tryCatch( ugarchfit(spec, rendimientos), error=function( err ) FALSE,
warning=function( warn ) FALSE )
if( !is.logical( fit ) ){
  if (as.double(infocriterio(fit)[1])<as.double(miGarch[[3]])) {
    mse<-0 #el mse se utilizaba en la primera versión de prueba antes de usar el AIC,
          #posteriormente se deja de usar y se reemplaza por 0 para aligerar los cálculos
    miGarch<-c(as.integer(i),as.integer(j),infocriterio(fit)[1],mse)
  }
}
}
}
return(miGarch)
}

```

```

mejoresAparch20072008norm <-calculaMejorModelo('apARCH','norm',2)
mejoresAparch20072008std <-calculaMejorModelo('apARCH','std',2)
mejoresGarch20072008norm <-calculaMejorModelo('sGARCH','norm',0)
mejoresGarch20072008std <-calculaMejorModelo('sGARCH','std',0)

```

### **Preparación de las series en el periodo de marzo a agosto de 2008.**

Se recuperan los rendimientos de las series para cada valor y se almacenan en la variable con nombre ticker del valor más el sufijo ES.

```

for(i in 1:length(tickers)){
  name<-gsub(".MC$", "", tickers[i])
  subconjunto<-window(eval(as.name(paste(name,'REN',sep=""))),
                      start = as.Date("2008-03-01"),
                      end = as.Date("2008-08-13"))
  assign(paste(name,"ES", sep=""),subconjunto)
}

```

```

## Generación de los conjuntos de sectores a partir de las series de rendimientos
## tickers sector 1

```

```
tsectores1<-  
c("IBE","REP","UNF","REE","GAS","IBR","ENG","ELE","SLR","AGS","CEP","ENC","FR  
S","MTB")
```

```
sectores1<-merge(IBEES,REPES,UNFES,REEES,GASES,IBRES,ENGES,ELEES,SLRES,  
AGSES,CEPES,ENCES,FRSES,MTBES)
```

```
sectores1<-na.omit(sectores1)
```

```
length(sectores1)/length(tsectores1)
```

```
#tickers del sector 2
```

```
tsectores2<-c("GAM","ACS","ANA","FER","ACX","FCC","TRE","SCYR",  
"ABG","OHL","SED","TUB","ZOT","CAF","GALQ",  
"ERC","CPL","TRG","MDF","FDR","CIE","AZK","CCMC","ESZ")
```

```
sectores2<-merge(GAMES,ACSES,ANAES,FERES,ACXES,FCCES,TREES,SCYRES,  
ABGES,OHLES,SEDES,TUBES,ZOTES,CAFES,GALQES,  
ERCES,CPLES,TRGES,MDFES,FDRES,CIEES,AZKES,CCMCES,ESZES)
```

```
sectores2<-na.omit(sectores2)
```

```
length(sectores2)/length(tsectores2)
```

```
tsectores3<-c("ITX","GRF","ZEL","ALM","EBRO","OLE","FAE","VIS","PAC",  
"MLX","CPF","BDL","ROVI","SNC","VID","PVA","NAT")
```

```
sectores3<-merge(ITXES,GRFES,ZELES,ALMES,EBROES,OLEES,FAEES,WISES,PACES,  
MLXES,CPFES,BDLES,ROVIES,SNCES,VIDES,PVAES,NATES)
```

```
sectores3<-na.omit(sectores3)
```

```
length(sectores3)/length(tsectores3)
```

```
tsectores4<-
```

```
c("ABE","CIN","IBLA","TL5","PRS","A3M","MEL","PSG","NH","VOC","VLG",  
"SPS","CDR","ITI","DERM","CBAV","VER")
```

```
sectores4<-
```

```
merge(ABEES,CINES,IBLAES,TL5ES,PRSES,A3MES,MELES,PSGES,NHES,VOCES,VL  
GES,
```

```
SPSES,CDRES,ITIES,DERMES,CBAVES,VERES)
```

```
sectores4<-na.omit(sectores4)
```

```
length(sectores4)/length(tsectores4)
```

```
tsectores5<-c("SAN","BBVA","POP","SAB","MAP","BME","BKT","CABK","COL",
             "BTO","MVC","ALB","PAS","BVA","GCO","QBT","REY","REN","RLIA","PSL",
             "UBS","GUI","R4")
```

```
sectores5<-merge(SANES,BBVAES,POPES,SABES,MAPES,BMEES,BKTES,CABKES,
                COLES,BTOES,MVCES,ALBES,PASES,BVAES,GCOES,QBTES,REYES,RENES,
                RLIAES,PSLES,UBSES,GUIES,R4ES)
```

```
sectores5<-na.omit(sectores5)
```

```
length(sectores5)/length(tsectores5)
```

```
tsectores6<-c("TEF","IDR","EZE","AMP","TEC")
```

```
sectores6<-merge(TEFES,IDRES,EZEES,AMPES,TECES)
```

```
sectores6<-na.omit(sectores6)
```

```
length(sectores6)/length(tsectores6)
```

Contenido del vector de tickers que se utiliza en la mayor parte de los cálculos secuenciales para todas las series. Se carga desde un xls. Inicialmente se carga con sufijo en algunas de ellas .MC, que posteriormente en las versiones finales se limpia.

```
tickers<-Lista_valores_completa_2008_migration_yahoo$TICKER
```

```
tickers
 [1] "SAN" "TEF" "BBVA" "IBE" "REP" "UNF" "POP" "ITX" "GAM" "ACS"
"REE"
 [12] "ANA" "GAS" "FER" "IBR" "ABE" "ENG" "SAB" "ACX" "MAP" "ELE"
"FCC"
 [23] "BME" "GRF" "CIN" "IBLA" "BKT" "CABK" "IDR" "TL5" "TRE" "COL"
"SCYR"
 [34] "SLR" "PRS" "AGS" "BTO" "ABG" "OHL" "SED" "MVC" "ZEL" "A3M"
"MEL"
 [45] "TUB" "ALM" "ALB" "PAS" "EZE" "EBRO" "ZOT" "CEP" "OLE" "BVA"
"PSG"
 [56] "NH" "GCO" "QBT" "CAF" "FAE" "VIS" "ENC" "VOC" "GALQ" "REY"
"PAC"
 [67] "ERC" "REN" "CPL" "VLG" "TRG" "SPS" "FRS" "MLX" "RLIA" "CPF"
"CDR"
 [78] "AMP" "BDL" "PSL" "ITI" "DERM" "ROVI" "SNC" "UBS" "MDF" "CBAV"
"VID"
 [89] "PVA" "MTB" "FDR" "NAT" "CIE" "GUI" "R4" "AZK" "VER" "TEC"
"CCMC"
[100] "ESZ"
```

## Generación de series sectoriales

sector1O, sector2O, sector3O, etc, contienen las series originales de cotización entre marzo y agosto de 2008.

```
##Creación de series sintéticas sectoriales
```

```
sector1<-apply.daily(sector1O,function(a){prod(a)^(1/length(a))})
sector2<-apply.daily(sector2O,function(a){prod(a)^(1/length(a))})
sector3<-apply.daily(sector3O,function(a){prod(a)^(1/length(a))})
sector4<-apply.daily(sector4O,function(a){prod(a)^(1/length(a))})
sector5<-apply.daily(sector5O,function(a){prod(a)^(1/length(a))})
sector6<-apply.daily(sector6O,function(a){prod(a)^(1/length(a))})
```

```
##Creación de series de rendimientossintéticas sectoriales
```

```
sectores1REN<-diff(log(sector1))
sectores2REN<-diff(log(sector2))
sectores3REN<-diff(log(sector3))
sectores4REN<-diff(log(sector4))
sectores5REN<-diff(log(sector5))
sectores6REN<-diff(log(sector6))
```

```
sectorial<-
merge(sectores1REN,sectores2REN,sectores3REN,sectores4REN,sectores5REN,sectores6REN)
```

## Estimación de modelo DCC

En este caso para el sector 6, por ser el más pequeño. La diferencia radica en que para cada valor se carga el número de especificaciones necesario, todas ellas individuales donde se indica el modelo a usar de los previamente seleccionado.

```
## Sectores conjunto 6
```

```
#TEF IDR EZE AMP TE
```

```
TEFGARCH = ugarchspec(mean.model= list(armaOrder= c(0,0)),
```

```

variance.model= list(garchOrder= c(1,1), model = "sGARCH"),
distribution.model= "norm")
IDRGARCH = ugarchspec(mean.model= list(armaOrder= c(0,0)),
variance.model= list(garchOrder= c(1,1), model = "sGARCH"),
distribution.model= "norm")
EZEGARCH = ugarchspec(mean.model= list(armaOrder= c(0,0)),
variance.model= list(garchOrder= c(1,1), model = "sGARCH"),
distribution.model= "norm")
AMPGARCH = ugarchspec(mean.model= list(armaOrder= c(0,0)),
variance.model= list(garchOrder= c(1,1), model = "sGARCH"),
distribution.model= "norm")
TECGARCH = ugarchspec(mean.model= list(armaOrder= c(0,0)),
variance.model= list(garchOrder= c(1,2), model = "sGARCH"),
distribution.model= "std")
dcc.sector6 = dccspec(uspec=
multispec( c(TEFGARCH,IDRGARCH,EZEGARCH,AMPGARCH,TECGARCH)),
dccOrder= c(1,1),distribution = "mvnorm")
dccfitsectores6 = dccfit(dcc.sector6, data=sectores6)

```

### **Preparación de las matrices de correlación.**

Se realiza la misma operativa para cada sector.

```

MatrizCMSectores1 <- list()
MatrizCMSectores1 = rcor(dccfitsectores1)
matriz<-matrix(0,length(tsectores1),length(tsectores1))
for(i in 1:length(tsectores1)){
  for(j in 1:length(tsectores1)){
    print(paste(i,":",j))
    matriz[i,j]<-prod(MatrizCMSectores1[i,j,])^(1/length(MatrizCMSectores1[i,j,]))
  }
}
matrizS1<-matriz
rownames(matrizS1)<-tsectores1

```

```
colnames(matrizS1)<-tsectores1
corrplot(matrizS1, order = "hclust", type="upper",addrect=2,cl.lim = c(0, 1),is.corr = TRUE)
corrplot(matrizS1, order = "hclust", addrect = 2)
#"ward", "single", "complete", "average", "mcquitty", "median" or "centroid".
corrplot(matrizS1, order = "hclust", hclust.method="ward",cl.lim = c(0, 1), addrect = 4)
```