

**RELACIÓN DE LAS ACCIONES DE LA
BOLSA DE VALORES DE COLOMBIA**

ANDRÉS FELIPE MUÑOZ GÓMEZ

PROYECTO DE GRADO II

PROFESORES:

JULIO CÉSAR ALONSO

BEATRIZ GALLO CÓRDOBA

UNIVERSIDAD ICESI

**FACULTAD DE CIENCIAS ADMINISTRATIVAS Y ECONÓMICAS
PROGRAMA DE ECONOMÍA Y NEGOCIOS INTERNACIONALES**

SANTIAGO DE CALI

23 DE MAYO DE 2014

CONTENIDO

1. INTRODUCCIÓN.....	4
2. DATOS	6
3. METODOLOGÍA.....	13
4. RESULTADOS: EVALUACIÓN DE PRONÓSTICOS.....	14
5. COMENTARIOS FINALES.....	18
Referencias.....	20

RESUMEN

La necesidad por hacer buenos pronósticos se ha manifestado en la forma en la sofisticación de los métodos de estimación. En la econometría por ejemplo, se intenta llegar a resultados lo más acertados posibles que permitan tomar decisiones con diferentes técnicas. El presente trabajo presenta una aproximación a lo que se debería considerar como un buen diseño de selección de modelo para hacer aproximaciones sobre los rendimientos de los activos financieros, basándose en una selección donde prime la homogeneidad entre ellos.

Palabras claves: rendimientos, acciones, BVC, VAR, ARMA.

1. INTRODUCCIÓN

El mercado accionario colombiano o de Renta Variable concentra sus actividades en la Bolsa de Valores de Colombia (BVC), y aunque ha sido foco de estudio para intentar predecir el comportamiento de los activos financieros desde distintos enfoques; que van desde lo meramente analítico sobre las condiciones fundamentales del mercado, hasta sofisticadas investigaciones con modelaciones estadísticas. Son muchos los estudios que se han encargado de mostrar evidencia a favor de la hipótesis de la eficiencia débil de mercado, que implica básicamente que los precios no contienen la suficiente información sobre el comportamiento futuro de los mismos (Fama, 1991).

Sin embargo con el paso del tiempo, las técnicas en la econometría aplicada se han ido afinando y actualmente se cuentan con herramientas sofisticadas para dar mayor validez a los resultados, por ejemplo en las estimaciones de las proyecciones del mercado bursátil. Y eso sobre este último punto que se toma como referencia para este trabajo, dejando de un lado la eficiencia débil del mercado, lo que se pretendemos es entender cuál es la mejor forma de pronosticar los retornos de acciones similares de la BVC, decidiendo si es mejor reconocer la existencia de relaciones entre ellas.

Si bien ya son varios los trabajos que intentan abordar el mismo objetivo, como por ejemplo los estudios de Montenegro (2007) y Rivera (2009) que caracterizan la modelización de los índices accionarios en Colombia a partir de particularidades exógenas como el efecto que tiene día de la semana o trabajos como los de Alonso y Gallo (2013) y Carvajal, Nieto, y Trejos (2003)) que abordan el tema de la proyección de rendimientos y sobre los cuales este trabajo toma como referencia metodológica, no presentan

de forma clara la selección que puede llegar a ser arbitraria de las series a analizar.

Para cumplir con esta serie de objetivos, el presente está estructurado de la siguiente manera: en la segunda sección se explica el manejo de los datos y el proceso de selección de un portafolio idóneo a través del análisis de *clusters*, en la tercera sección, se especifica la metodología usada de los distintos modelos considerados, posteriormente, se muestran los principales resultados de las estimaciones y finalmente, se exponen algunos comentarios con consideraciones para futuros trabajos.

2. DATOS

Como se aclara en la sección anterior, siguiendo la línea de la mayoría de estudios financieros, se realiza un análisis a partir de los rendimientos (o retornos) en lugar de los precios de los activos, definiendo como retorno simple neto de un periodo (o tasa de variación de los precios) como sigue:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Como se expuso en la sección anterior, la selección de acciones para desarrollar trabajos de investigación sobre proyecciones no son muy claros y, aunque este trabajo no explore criterios formales, se presenta una primera aproximación a lo que se debería considerar. Para hablar de acciones representativas se debe considerar un grupo de características que puedan servir como indicadores de concentración. La hipótesis que los rendimientos de las acciones tienden a depender del comportamiento de las otras puede examinarse incluso con pruebas no paramétricas.

La figura 1 muestra las correlaciones de los rendimientos a lo largo del periodo de análisis (Enero 04 de 2010-Diciembre 31 de 2013), las cuales nos permiten determinar qué tan fuerte es la relación lineal entre los retornos de las acciones. A simple vista no se expone con claridad una fuerte dependencia entre activos, hay algunos pares que presenta correlaciones superiores al 0.5, provenientes de activos con sectores diferentes. Para citar un par de ejemplos, está el caso la relación lineal más fuerte entre PFGRUPOARG y el EXITO con una correlación del 0.61, y PFDAVVNDA con GRUPOARGOS y a su vez con GRUPOSURA con 0.6.

Dada la importancia de escoger un adecuado grupo de activos como aspecto trascendental para proyectar mejor sus rendimientos dentro de

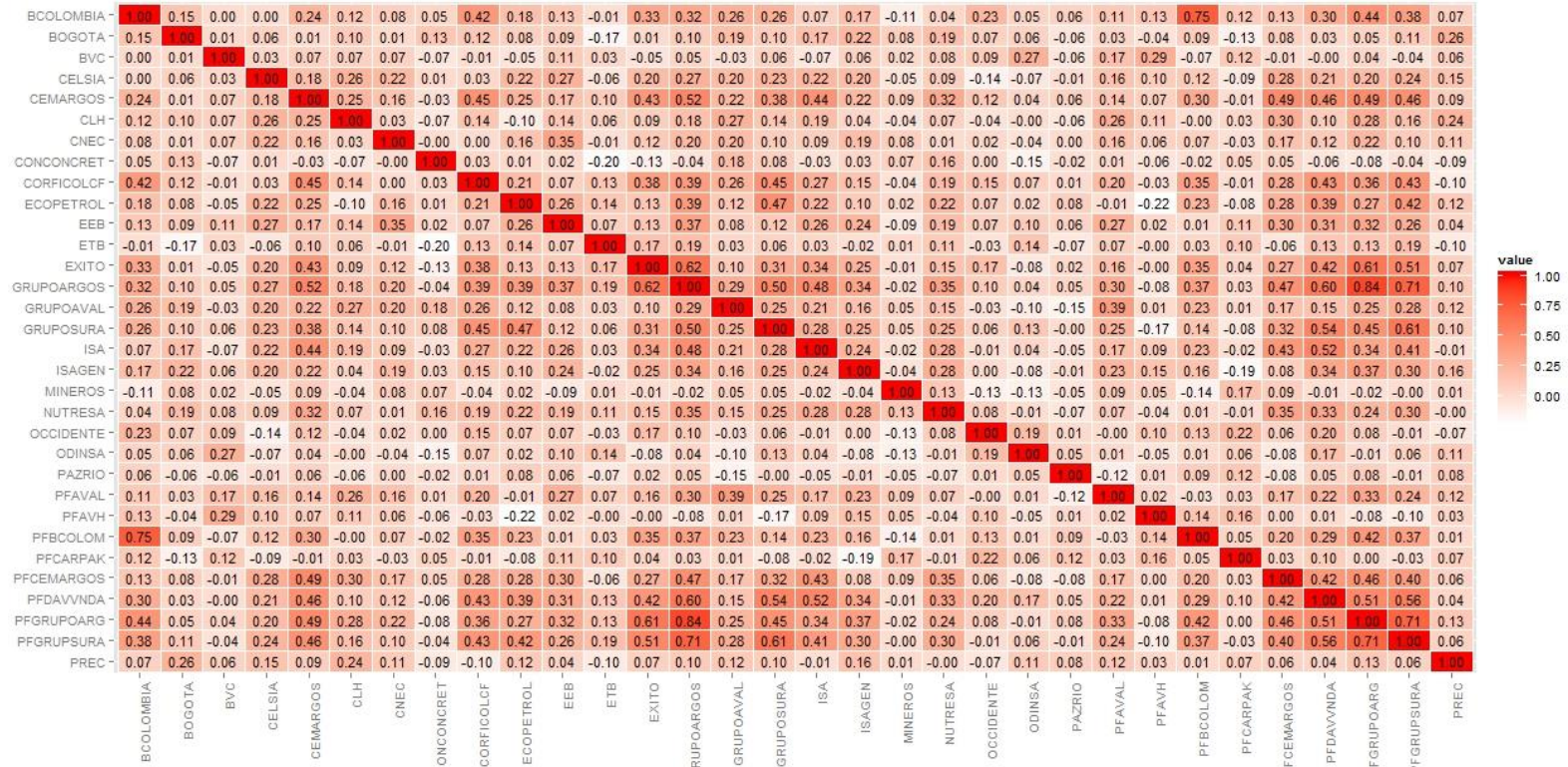
un portafolio, a sabiendas que al considerar las acciones individualmente podría causar inconvenientes dada a la relación que existe de interrelación que se ha evidenciado empíricamente.

Se considera el análisis *cluster*, que es una técnica multivariante que tiene como objetivo agrupar elementos (en nuestro caso los activos financieros), tratando de lograr la máxima homogeneidad en cada grupo y a su vez, la mayor diferencia. Sin embargo, es necesaria una representación gráfica para determinar con mayor facilidad los distintos grupos que se forman; por eso, se usa el dendrograma para interpretar estos resultados.

Para el análisis *cluster* se consideraron las siguientes variables:

- Rendimiento medio (\bar{R}) de los retornos de las acciones
- La desviación (SD_r) de los retornos
- Una variable dummy que toma el valor de 1 si el activo perteneció al COLCAP histórico y 0 en caso contrario.
- Un índice de variación, denominado como la razón entre la desviación y el rendimiento medio (SD_r / \bar{R}). Buscamos medir cuántas veces la media está en una desviación estándar. Entre más alto este índice, mayor dispersión se puede encontrar entre los rendimientos.

Figura 1: Correlaciones de los rendimientos de las acciones (2010-2013)



Se recurre al método de aglomeración donde la función de enlace que especifica la distancia entre las dos agrupaciones a través de la distancia media, se puede describir con la siguiente expresión (Díaz y Morales, 2012):

$$D(X, Y) = \frac{1}{N_x \times N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} d(x_i, y_j); \forall x_i \in X, y_j \in Y$$

donde:

$d(x, y)$ es la distancia entre los objetos $x \in X$ y $y \in Y$

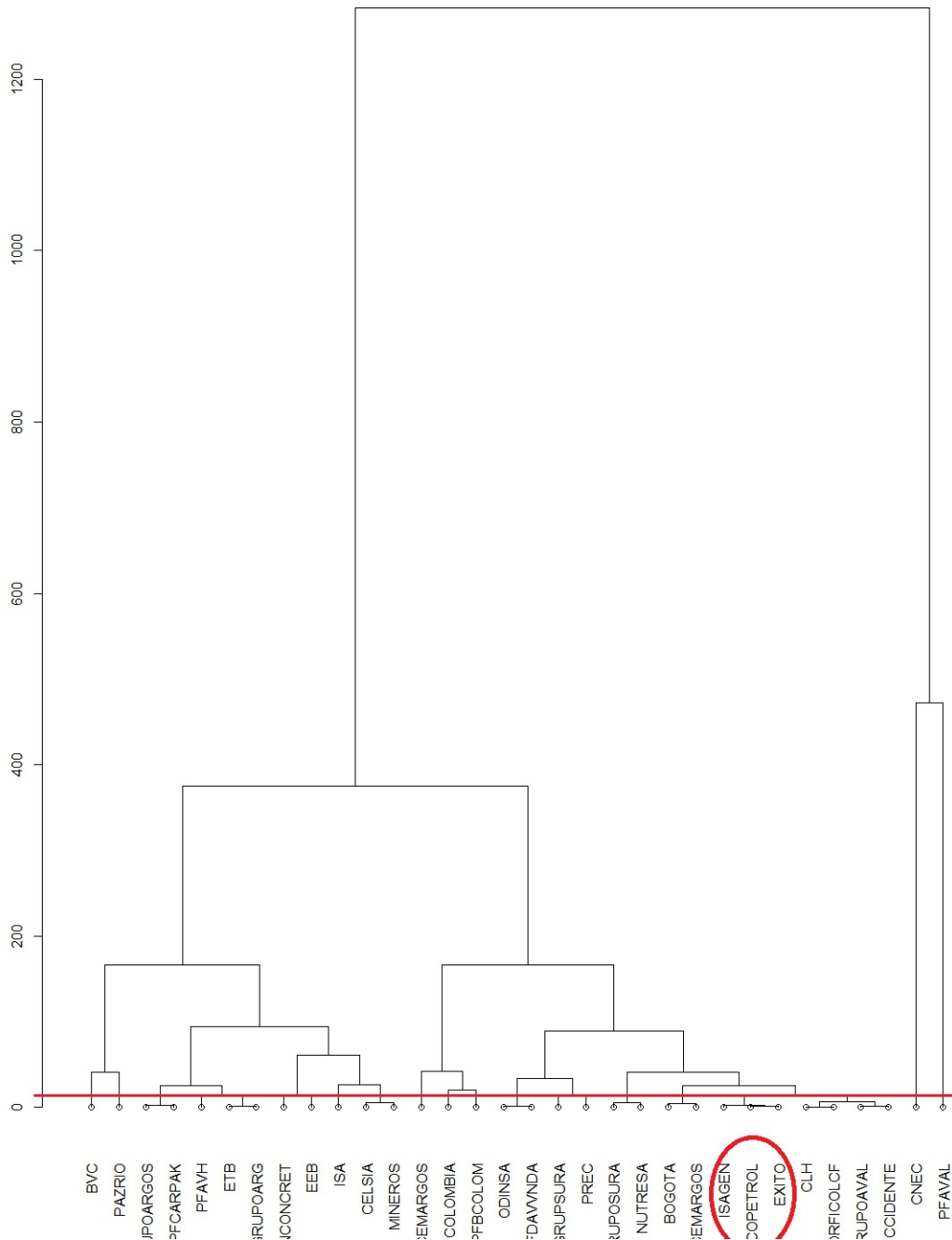
X e Y representan las serie de objetos (clusters)

N_x y N_y es el total de objetos en los clusters X e Y , respectivamente.

El dendrograma resultante (ver figura 2) se realiza para el grupo de 32 activos de la Bolsa de Valores de Colombia. Fueron descartados aproximadamente 85 activos por varias razones, entre ellas porque no se registró información a partir de un periodo, o porque la acción no presenta transacciones con regularidad. Cabe destacar el hecho que las series de NUTRESA, CELSIA y PFAVH fueron empalmadas porque antes se registraban con los nemotécnicos de CHOCOLATES, COLINVERS y PFAVTA, respectivamente.

Se elige el grupo formado por las acciones registradas en el cuadro 1 intentando conformar un grupo de mínimo tres acciones que a su vez fuera lo más homogéneo posible. En el dendrograma se observa la línea divisora donde se traza la marca de referencia, con una distancia cercana a cero.

Figura 2: Dendrograma de las acciones en estudio



Dado que los precios de cierre de los activos son proceso $I(1)$ y no presentan evidencia de cointegración, se utilizan las series de los retornos de las acciones del portafolio escogido. El cuadro 2 muestra las diversas pruebas que se realizaron para determinar la presencia de raíces unitarias para los retornos. Las pruebas que se realizaron fueron:

- Dickey-Fuller Aumentada (ADF) (Dickey, 1984)
- Phillips y Perron (Phillips & Perron, 1988)
- KPSS (Kwiatkowski et al., 1992)
- Breitung (Breitung, 2002)
- Zivot y Andrews (Zivot & Andrews, 2002)

Para todas las pruebas se rechaza la hipótesis de existencia de al menos una raíz unitaria, y aunque para la prueba KPSS en Ecopetrol (señalada de color rojo en el cuadro 2) también se rechace la hipótesis nula de estacionariedad. Se concluye a partir de la prueba de Dickey-Fuller debido a que ésta prueba concluye la no existencia de raíces unitarias.

Cuadro 2: Prueba de raíces unitarias para los retornos de las acciones

	Rezagos	Estadístico	
ECOPETRO			
L			
ADF-BIC	23	-29.2537	***
ADF-AIC	29	-16.3965	***
ADF-Significancia		-29.6668	***
Phillips y Perron		-29.6287	***
KPSS		0.5698	**
Breitung		0.0006	
Zivot y Andrews	29	-5.8898	***
ISAGEN			
ADF-BIC	13	-31.6818	***
ADF-AIC	29	-15.8011	***
ADF-Significancia		-15.9152	***
Phillips y Perron		-32.2400	***
KPSS		0.1705	
Breitung		0.0001	
Zivot y Andrews	29	-7.3551	
ÉXITO			
ADF-BIC	23	-31.4365	***
ADF-AIC	29	-31.2277	***
ADF-Significancia		-31.7471	***
Phillips y Perron		-31.7908	***
KPSS		0.1015	
Breitung		0.0001	
Zivot y Andrews	29	-6.3288	

Notas: ***: valor-p <0.01, **: valor-p <0.05, *: valor-p <0.1

3. METODOLOGÍA

Previamente se muestra que los retornos de los activos son estacionarios, esto implica que puede estimarse a través de un modelo Vector Autorregresivo de orden p que cumple la función de identificar las interrelaciones entre las variables estudiada con variables exógenas ($VAR(p)$), s, que en este caso corresponderá a los días de la semana, con $j = \text{martes, miércoles, jueves y viernes}$, en el modelo como el siguiente

$$y_t = \alpha + \sum_{i=1}^p \beta_i y_{t-i} + \sum_{j=\text{martes}}^{\text{viernes}} \delta_j D_{jt} + v_t \quad (1)$$

Del modelo (1) se resalta que α recoge el efecto para el día lunes, δ_j la diferencia del día j respecto al lunes, y donde el efecto del rezago de los retornos entre unos y otros activos se recogen en la matriz de coeficientes β_j . Ahora bien, para la estimación del modelo se tiene en cuenta los criterios de información de Akaike y Schwartz y de Hannan-Quinn.

Otra alternativa que se considera para modelar el comportamiento de los rendimientos, es el modelo autorregresivo de media móvil ($ARM A(p, q)$) que a diferencia del modelo $VAR(p)$, no tiene en cuenta el comportamiento de las otras series en estudio. El modelo para este caso corresponde al siguiente:

$$y_t = \beta + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (2)$$

Para encontrar el mejor ajuste en la estimación de (2) se aplica la metodología de Box-Jenkins.

4. RESULTADOS: EVALUACIÓN DE PRONÓSTICOS

Al estimar el modelo (1) se concluye que el día de la semana tiene un efecto significativo sobre los retornos de los activos financieros. Como se había dicho en la sección anterior, se usan los criterios de información de Akaike y Schwartz y de Hannan-Quinn, específicamente, se selecciona el modelo que mínima cada criterio de información.

Se encuentra que el mejor modelo Vector Autorregresivo para estimar los retornos considerando las variables exógenas es un VAR(2). El Cuadro 3 muestra los resultados, usando Mínimos Cuadrados Ordinarios (MCO) para estimar cada ecuación.

A partir de estos resultados, se puede concluir que los retornos de las acciones están inter-relacionados, además, que los días de la semana sí influyen. Por ejemplo, el rendimiento de ECOPETROL puede verse afectado positivamente de tratarse del día martes o miércoles. Por otro lado, un aumento en el rendimiento de ISAGEN puede explicarse por un aumento en el rendimiento del día anterior de ECOPETROL o EXITO. Además martes, miércoles o viernes también explicarían un aumento en los retornos de ISAGEN. Por último, los rendimientos del EXITO solo se ven afectados si hubo un cambio del rendimiento de ISAGEN el día anterior

Cuadro 3: Estimación del modelos VAR

	rECOPETROL		rISAGEN		rEXITO
$rECOPETROL_{t-1}$	0.064667	*	0.08191	***	-0.0225374
$rISAGEN_{t-1}$	-0.018525		-0.06377	*	0.0735299 *
$rEXITO_{t-1}$	-0.005112		0.05262	*	-0.0272022
$rECOPETROL_{t-2}$	-0.005831		-0.02786		0.0532483
$rISAGEN_{t-2}$	0.051561		-0.05348		0.0073621
$rEXITO_{t-2}$	-0.05283		-0.01906		-0.0312023
<i>Constante</i>	-0.08539		-0.15214		0.0005272
<i>Jueves</i>	0.023431		0.12239		0.0077596
<i>Martes</i>	0.249374	*	0.27639	**	0.0540905
<i>Miercoles</i>	0.260429	*	0.2934	**	0.1563692
<i>Viernes</i>	0.07456		0.24121	*	-0.0220993
R^2	0.0975		0.1289		0.08521
$R^2 - ajustado$	0.0789		0.4385		0.079
F_{Global}	1.489		2.287		1.563

Notas: ***: valor-p <0.01, **: valor-p <0.05, *: valor-p <0.1

Ahora, consideramos las estimaciones para el modelo bajo el supuesto que no hay relación entre los rendimientos de las acciones para comparar con los resultados anteriores. Para esto, se calculan los errores de predicción para los modelos $ARMA(p, q)$, usando el error absoluto medio (MAE), el error absoluto porcentual medio (MAPE) y el error cuadrático medio (MSE) como medida para evaluar el pronóstico *one - step - ahead* de los modelos para los últimos quince periodos (al interior de la muestra) de los modelos (1) y (2). El cuadro 4 se exponen los modelos $ARMA(p,q)$ obtenidos según la metodología Box-Jenkins.

Cuadro 4: Modelos $ARMA(p,q)$ para los rendimientos de las acciones estudiadas

Acción	Modelo
ECOPETROL	$ARMA(0,1)$ o $ARMA(0,0)$
ISAGEN	$ARMA(0,4)$ o $ARMA(0,0)$
ÉXITO	$ARMA(0,0)$

Los resultados del cuadro anterior implican que: para ECOPETROL, se encontraron dos modelos que no tienen problemas de autocorrelación y que minimizan el criterio de información Akaike y Schwartz ($ARMA(0,1)$), y el criterio de información bayesiano ($ARMA(0,0)$). A su vez, para ISAGEN se encuentra que los modelos $ARMA(0,4)$ y $ARMA(0,0)$ minimizan los criterios mencionados, respectivamente. Por tal motivo, para estos dos activos, se relacionan los criterios de evaluación, MAE, MAPE y MSE. El cuadro 5 muestra los resultados.

Los resultados implican que, para ECOPETROL y EXITO, el modelo $VAR(p)$ es el mejor para las proyecciones de los rendimientos fuera de muestra, considerando las variables exógenas de días, en cambio para ISAGEN, el mejor modelo resulta ser el ARMA(0,0).

Cuadro 5: Evaluación de pronósticos *one-step-ahead*

	MAE	MAPE	MSE	MAE	MAPE	MSE
ARMA(p,q)						
ECOPETROL (ARMA (0,1) - ARMA (0,0))	1.4679	93.0021	4.8691	1.5717	100.2015	5.5825
ISAGEN (ARMA (0,4) - ARMA (0,0))	0.9677	-	1.6043	0.9303	-	1.5158
EXITO ARMA(0,0)	0.9288	-	1.5136			
VAR(2)						
ECOPETROL	1.4207	87.389	4.6973			
ISAGEN	1.0509	-	1.7581			
EXITO	0.9023	-	1.3584			

5. COMENTARIOS FINALES

En el presente trabajo intenta solucionar varias inquietudes, unas metodológicas como la de la importancia de la selección de una buena canasta de series para analizar, hasta la de encontrar el modelo que mejor se ajuste a las diferentes series tanto de forma individual como interrelacionadas.

A partir de un proceso de análisis de *cluster*, considerando varias características como el rendimiento medio, la variación y la permanencia histórica al COLCAP, se crearon las distancias con del método del promedio y así, lograr agruparlas en una escala de homogeneidad. Tras la depuración de datos, y con un análisis gráfico con ayuda de un dendrograma, se concluye que el grupo de acciones que debían considerarse eran: ECOPETROL, ISAGEN y EXITO.

Con este portafolio de activos, se realizaron las estimaciones a partir de las que se concluyó que la estimación del modelo VAR(2) es el mejor para representar las series interrelacionadas entre sí. Además, se evidenció que la inclusión de variables exógenas significativas como los días de la semana al menos para activos como ECOPETROL y EXITO, y también, que sólo para este primero se podría atribuirle algún efecto sobre los demás activos tras el análisis de la función impulso-respuesta. En cambio, las conclusiones a partir de la evaluación *one - step - ahead* afirmaron que solo para los retornos de ISAGEN, el modelo ARMA (0,0) resulta ser el más apropiado.

Por último, se deja claridad que, si bien, el documento no explora formas de modelización de la varianza, ni considera un estudio para el cálculo del Valor en Riesgo, se deja como guía para ampliaciones y posteriores trabajos considerando por ejemplo la influencia no sólo del día de la semana, sino también del mes como otra variable exógena.

Referencias

- Alonso, J., y Gallo, B. (2013). Relación entre los retornos de cuatro acciones de la bolsa de valores de Colombia. *XXIII Simposio Internacional de Estadística*.
- Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. *Journal of Econometrics*, 108(2), 343-363.
- Carvajal, P., Nieto, S., y Trejos. (2003). Modelo de predicción del precio de la acción ordinaria cementos argos. *Scientia et Technica*, 23.
- Díaz, L., y Morales, M. (2012). *Análisis estadístico de datos multivariados* (1.ª ed.).
- Dickey, D., y Said, E. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71, 599-607.
- Fama, E. (1991). Efficient capital markets:ii. *Journal of Business*, 46(5), 1575-1617.
- Kwiatkowski, P. P. S. P., D., y Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 57, 159-178.
- Montenegro, L. (2007). El efecto del día en la bolsa de valores de Colombia. *Technical report, Universidad Javeriana*.
- Phillips, P., y Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75, 335-346.
- Rivera, D. M. (2009). Modelación del efecto del día de la semana para los índices accionarios de Colombia mediante un modelo STAR GARCH. *Revista de Economía del Rosario*, 12(1), 1-24.
- Zivot, E., y Andrews, D. (2002). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, 20(1), 25-44.