



Universidad Autónoma de Querétaro
Facultad de Contaduría y Administración
Licenciatura en Actuaría

"Clusterización de la Bolsa Mexicana de Valores empleando análisis de series de tiempo"

Tesis Colectiva

Que como parte de los requisitos para obtener el título de Actuario.

Presenta:

Santiago Aniceto Orantes Anchondo
Tlahui Alberto Bolaños Mejía

Dirigido por:

Dra. Denise Gómez Hernández

Dra. Denise Gómez Hernández
Presidente

Firma

Dr. Michael Demmler
Secretario

Firma

Dr. Ignacio Almaraz Rodríguez
Vocal

Firma

Dra. Julia Hirsch
Suplente

Firma

Dr. Humberto Banda Ortiz
Suplente

Firma

Centro Universitario
Querétaro, Qro.
Febrero 2019



Universidad Autónoma de Querétaro
Facultad de Contaduría y Administración
Licenciatura en Actuaría

“Clusterización de la Bolsa Mexicana de Valores empleando análisis de series de tiempo”

Tesis Colectiva

Que como parte de los requisitos para obtener el título de Actuario.

Presenta:

Santiago Aniceto Orantes Anchondo
Tlahui Alberto Bolaños Mejía

Dirigido por:

Dra. Denise Gómez Hernández

Centro Universitario
Querétaro, Qro.
Febrero 2019

RESUMEN

El presente trabajo compara la caracterización en forma de clústeres de 25 acciones de la Bolsa Mexicana de Valores utilizando el método clásico de media-varianza y la autocorrelación en sus series de tiempo. Para esto se emplea la clusterización jerárquica optimizada a través del índice de ancho de silueta con la información de media-varianza (clusterización 1); y media-varianza y autocorrelación (clusterización 2). Se experimentó con distintos valores para el número de clústeres y de lags, obteniéndose mejores resultados empleando entre 13 y 14 clústeres y 6 lags. Se analiza cualitativamente la representación gráfica de los clústeres, y se concluye que, al ser sustancialmente diferentes, la información adicional de autocorrelación es relevante. Finalmente se realiza un ejercicio de optimización de portafolios dentro de cada clúster y posteriormente entre los clústeres optimizados. Los resultados describen una mayor estabilidad en rendimientos altos en la clusterización media-varianza, mientras que la clusterización media-varianza y autocorrelación se comporta de una manera más apegada a las acciones individuales en rendimientos altos.

(Palabras clave: clusterización, series de tiempo, media-varianza, acciones)

SUMMARY

This paper compares the characterization in the form of clusters of 25 shares of the Mexican Stock Exchange using the classic method of mean-variance and autocorrelation in their time series. For this, Hierarchical clustering is used, optimized through the Silhouette Index, with the mean-variance (clustering 1); and mean-variance and autocorrelation (Clustering 2) information. Different values for the number of clusters and lags were tested, obtaining better results by using between 13 and 14 clusters and 6 lags. The graphical representation of the clusters is analyzed qualitatively, concluding that both methods are substantially different, thus proving the additional information provided by autocorrelation is relevant. Finally, portfolio optimization is carried out within each cluster and subsequently among the optimized clusters. The results convey a better stability in high returns in the mean-variance clustering, while the mean-variance and autocorrelation clustering behaves closer to the individual stocks in high returns.

(Key words: clustering, time series, mean-variance, shares)

AGRADECIMIENTOS

Primeramente, nos gustaría agradecer el apoyo brindado por nuestras familias, Bolaños Mejía y Orantes Anchondo, en especial a nuestras mamás, papás y hermanos, que nos han guiado y acompañado no sólo en esta tesis, sino a lo largo de nuestras vidas, y sin cuya presión constante este proyecto se habría prolongado indefinidamente.

Es importante también reconocer todo el esfuerzo por parte de nuestros mentores y profesores, que nos brindaron todas las bases necesarias para el desarrollo de este proyecto. En especial nos gustaría mencionar a la Dra. Denise Gómez, quien dirigió esta tesis; al Dr. Felipe Pérez, director de la Licenciatura en Actuaría, quien siempre se distinguió por su disponibilidad; y al Maestro Antonio Villanueva, profesor y amigo cuyos comentarios siempre fueron enriquecedores.

Finalmente, le damos las gracias a nuestros amigos y compañeros que nos acompañaron a lo largo de nuestra carrera. En particular a Mariana, Saavedra, Watson, Sergio y Marisa, sin cuya amistad, consejos y críticas este proyecto habría sido mucho más complejo, técnica y administrativamente.

DEDICATORIAS

Este trabajo está dedicado a nuestras familias y amigos, gracias por el apoyo constante.

“Life would be tragic if it weren't funny.”

-Stephen Hawking

Dirección General de Bibliotecas UAQ

TABLA DE CONTENIDOS

1. INTRODUCCIÓN	10
1.1 ANTECEDENTES	10
1.2 PLANTEAMIENTO DEL PROBLEMA	11
1.3 OBJETIVO GENERAL	12
1.4 OBJETIVOS PARTICULARES	12
1.5 HIPÓTESIS	13
2. MARCO TEÓRICO	14
2.1 BOLSA MEXICANA DE VALORES	14
2.2 TEORÍA DE PORTAFOLIOS Y ANÁLISIS DE ACCIONES	15
2.3 ACCIONES COMO SERIES DE TIEMPO	20
2.4 CLUSTERIZACIÓN COMO HERRAMIENTA DE ANÁLISIS	25
3. METODOLOGÍA	33
4. RESULTADOS Y DISCUSIÓN	37
5. CONCLUSIONES	57
6. REFERENCIAS	60
7. ANEXOS	63

ÍNDICE DE FIGURAS

Figura 1. Frontera eficiente. Fuente: elaboración propia.	18
Figura 2. Dendograma. Fuente: elaboración propia.	27
Figura 3. Dispersión de las acciones. Fuente: elaboración propia.	37
Figura 4. Función de autocorrelación (ACF), por sus siglas en inglés, del rendimiento del promedio de las acciones. Fuente: elaboración propia.	38
Figura 5. Medida m1 para K-medias con media y desviación estándar. Fuente: elaboración propia.	40
Figura 6. Medida m1 para jerárquica con media y desviación estándar. Fuente: elaboración propia.	41
Figura 7. Medida m2 para K-medias con media, desviación estándar y autocorrelación con lag 22. Fuente: elaboración propia.	42
Figura 8. Medida m2 para K-medias con media, desviación estándar y autocorrelación con lag 6. Fuente: elaboración propia.	43
Figura 9. Medida m2 para jerárquica con media, desviación estándar y autocorrelación con lag 22. Fuente: elaboración propia.	44
Figura 10. Medida m2 para jerárquica con media, desviación estándar y autocorrelación con lag 6. Fuente: elaboración propia.	45
Figura 11. Clusterización media y desviación estándar. Fuente: elaboración propia.	47
Figura 12. Clusterización media, desviación estándar y autocorrelación. Fuente: elaboración propia.	48

Figura 13. Clusterización media, desviación estándar y autocorrelación redimensionado con análisis de componentes principales. Fuente: elaboración propia..... 49

Figura 14. Frontera eficiente de los portafolios. Fuente: elaboración propia. 55

Dirección General de Bibliotecas UAQ

1. INTRODUCCIÓN

1.1 Antecedentes

Los mercados financieros son considerados sistemas complejos que están compuestos por muchos actores que interactúan entre ellos en una forma no lineal (Bonanno, Lillo, & Mantegna, 2001). Los diferentes factores macroscópicos de las finanzas, como los tipos de cambio o los precios de las acciones, son determinados por el comportamiento colectivo de empresas e individuos. Una de las grandes ventajas de los mercados financieros es la enorme colección de información histórica que existe; esta cantidad de información que se tiene de los mercados financieros alrededor del mundo, lo hacen uno de los sistemas complejos mejor documentados y permiten un análisis estadístico detallado de sus características (Rosén, Correlation based clustering of the Stockholm Stock Exchange, 2006).

Los mercados de valores en México tienen sus inicios en 1894 cuando se funda la Bolsa Nacional, antecedente directo de la BMV (Pérez Rojas & Ramírez Galindo, 2009). El año inmediato posterior, 1895, se funda la Bolsa de México cuya coincidencia de objetivos y vinculaciones entre miembros de la Bolsa Nacional y la Bolsa de México, los llevó a buscar la fusión de ambas en septiembre del mismo año, conservando el nombre de Bolsa de México S.A. (Grupo BMV, 2018).

Fue hasta el año 1928 que la bolsa queda bajo regulación de la Comisión Nacional Bancaria. En 1946, se crea la Comisión Nacional de Valores que posteriormente se fusionaría con la anterior comisión para formar la actual Comisión Nacional Bancaria y de Valores (Pérez Rojas & Ramírez Galindo, 2009). En el año de 1975 cuando se expide la Ley del Mercado de Valores y se le da el nombre actual de Bolsa Mexicana de Valores.

En junio de 2008, la BMV concluyó su proceso de desmutualización, convirtiéndose en una empresa pública cuyas acciones son susceptibles a negociarse

en el mercado de valores, llevando a cabo el 13 de junio de 2008 la oferta pública inicial de las acciones representativas de su capital social (Grupo BMV, 2018).

El movimiento de los precios de las acciones y la correlación que existe entre ellos ha sido el enfoque de muchos estudios en diferentes partes del mundo (Rosén, Correlation based clustering of the Stockholm Stock Exchange, 2006). La importancia de la correlación en la optimización de portafolios fue introducida por primera vez por Harry Markowitz en 1952 y ha sido utilizada en el desarrollo de nuevas técnicas en la teoría financiera como lo es el *Value at Risk* o la teoría del arbitraje (Campbell, Lo, & MacKinlay, 1996).

Diversos estudios han mostrado que las acciones tienen un comportamiento colectivo y que existe una alta correlación entre ciertos grupos de acciones lo cual indica que el mercado financiero se ve afectado por factores económicos comunes. Desde el punto de vista económico, puede ser de interés tener una clasificación o clusterización de acciones basada en la correlación que tienen entre ellas; en el caso de la diversificación podría ser más útil una clusterización de este tipo que una basada únicamente en el sector al que pertenecen las empresas en cuestión (Rosén, Correlation based clustering of the Stockholm Stock Exchange, 2006).

1.2 Planteamiento del Problema

Una de las mayores críticas de la optimización de portafolios basada en medias y varianzas, es que tiende a magnificar el efecto del error estándar, es decir, este proceso de optimización es extremadamente sensible a las variaciones en el rendimiento esperado y la matriz de covarianzas; lo anterior se traduce en una mala diversificación del portafolio (Ren, Portfolio Construction Using Clustering methods, 2005).

Una de las soluciones más populares a este problema de la sensibilidad a la media y varianza es la clusterización de acciones; es mucho más fácil estimar la correlación entre acciones que estimar el rendimiento esperado. Los valores de la matriz de

covarianza son más acertados y confiables que aquellos de los rendimientos esperados (Sherer, 2002). Las acciones suelen presentar fuertes correlaciones entre unas y otras, especialmente si pertenecen al mismo sector, esto permite agrupar a las acciones con alta correlación en un mismo grupo y de esta manera obtener varios grupos de acciones con baja correlación entre ellos. Si se piensa en estas agrupaciones (clústeres) como acciones individuales, se tiene un grupo de acciones con una baja correlación a partir del cual se puede encontrar un portafolio eficiente con la metodología de Markowitz.

Otra herramienta utilizada para el análisis de acciones es el pronóstico de los precios mediante series de tiempo. Existe mucha teoría referente a este tipo de modelos, no obstante, suelen ser complejos de implementar y optimizar en la práctica, además de que tienen resultados moderadamente precisos (Ladokhin, 2009). En cuanto al análisis descriptivo de series de tiempo, la evidencia empírica y teórica apunta a que la aplicación de la autocorrelación en los modelos de análisis del mercado accionario puede tener fuertes implicaciones en las regulaciones, el diseño de los mecanismos de intercambio, eficiencia de mercado, entre otros (Cerqueira, 2006).

1.3 Objetivo General

Comparar la caracterización en forma de clústeres de 25 acciones de la Bolsa Mexicana de Valores utilizando el método de media-varianza y la autocorrelación en sus series de tiempo.

1.4 Objetivos Particulares

1. Determinar el número adecuado de clústeres para las 25 acciones de la Bolsa Mexicana de Valores utilizando el método de media-varianza.
2. Determinar el número adecuado de clústeres para las 25 acciones de la Bolsa Mexicana de Valores utilizando la función de autocorrelación de sus series de tiempo.
3. Comparar los resultados de ambos modelos e identificar sus diferencias.

1.5 Hipótesis

La clusterización de acciones de la Bolsa Mexicana de Valores utilizando el análisis clásico de acciones (media-varianza); produce resultados fundamentalmente distintos a los obtenidos con la clusterización de acciones de la Bolsa Mexicana de Valores tomando en cuenta su función de autocorrelación.

Dirección General de Bibliotecas UAQ

2. MARCO TEÓRICO

2.1 Bolsa Mexicana de Valores

El mercado de valores es un tipo de mercado financiero en el que se realiza la compraventa de títulos representativos de los activos financieros de las empresas cotizadas en Bolsa. Una acción, siendo el principal instrumento en el mercado de valores, se define como el título emitido que representa una parte del capital social de una empresa la cual otorga al inversionista los derechos de un socio (Pérez Rojas & Ramírez Galindo, 2009). De acuerdo con la clasificación realizada por la Comisión Nacional Bancaria y de Valores, se puede dividir a las acciones en 4 tipos (Pérez Rojas & Ramírez Galindo, 2009):

- Al portador: Son acciones suscritas nominalmente en las cuales ejerce el derecho el portador de la misma.
- Preferentes: Gozan de ciertos derechos y prioridades principalmente en el pago en caso de liquidación, así como el cobro de dividendos antes de que sean cobrados por otros accionistas.
- Comunes u ordinarias: son las acciones que no gozan de preferencia, a menos que haya utilidades y que se realice el pago de dividendos los cuales son pagados después de haber cumplido con el pago de preferentes.
- Convertibles: Son aquellas que tienen la capacidad de convertirse en bonos en un tiempo predeterminado.

Los principales actores en el proceso de participación en el mercado de valores son las entidades emisoras, el público inversionista y los intermediarios; en México, la Ley del Mercado de Valores contempla dos tipos de intermediarios: las casas de bolsa y los especialistas bursátiles (Rueda, 2005).

La Bolsa Mexicana de Valores (BMV), es una entidad financiera que opera bajo concesión de la Secretaría de Hacienda y Crédito Público, con apego a la Ley del

Mercado de Valores (Grupo BMV, 2018). La BMV, es el foro en el que se llevan a cabo las operaciones del mercado de valores organizado en México, siendo su objeto el facilitar las transacciones con valores y procurar el desarrollo del mercado, fomentar su expansión y competitividad (Grupo BMV, 2018).

2.2 Teoría de portafolios y análisis de acciones

Realizar estudios financieros con acciones implica, la mayoría de las veces, recurrir a los rendimientos en lugar de centrarse en sus precios ya que, para el inversionista promedio, los mercados pueden llegar a considerarse perfectamente competitivos, por lo que una variación en el monto de la inversión no cambia el precio de la acción misma haciendo que la escala única de los rendimientos surja como una importante ventaja. Además, los rendimientos tienen propiedades estadísticas como la estacionariedad y ergodicidad que facilitan su manejo (Campbell, Lo, & MacKinlay, 1996). Entendiendo a la estacionariedad como un proceso en el cual sus propiedades estadísticas no se ven afectadas por cambios de origen temporal y la ergodicidad como un proceso en el cual el promedio temporal es igual a la media.

Definiendo P_t como el precio de una acción al tiempo t , entonces se puede obtener el rendimiento mediante las ecuaciones (1), (2) y (3).

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

$$R_t^B = 1 + R_t \quad (2)$$

$$R_t^B(k) = (1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1}) = \frac{P_t}{P_{t-1}} \cdot \frac{P_{t-1}}{P_{t-2}} \cdot \dots \cdot \frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}} \quad (3)$$

Con:

- R_t el rendimiento simple neto en el periodo de tiempo $(t - 1, t)$.
- R_t^B el rendimiento simple bruto en el periodo de tiempo $(t - 1, t)$.
- $R_t^B(k)$ el rendimiento simple bruto en k periodos $(t - k, t)$.

A los rendimientos de múltiples periodos se les denota rendimientos compuestos (Campbell, Lo, & MacKinlay, 1996) como lo mostró la ecuación (3). Sin embargo, la dificultad de trabajar con medias geométricas como la anterior, motiva a utilizar otro acercamiento para modelar los rendimientos de las acciones como lo es el rendimiento logarítmico r_t de una acción definido en la ecuación (4).

$$r_t = \log(1 + R_t) = \log\left(\frac{P_t}{P_{t-1}}\right) \quad (4)$$

La ventaja de utilizar rendimientos logarítmicos sobre los simples es que, al trabajar con periodos múltiples, el rendimiento total es la suma de los rendimientos individuales. Esto también facilita el modelaje del comportamiento estadístico, ya que es considerablemente más sencillo deducir las propiedades de la serie de tiempo de un proceso aditivo que de un proceso multiplicativo. No obstante, debe tenerse presente que los rendimientos logarítmicos no son una medida directa del cambio en la riqueza de un inversionista en un periodo de tiempo. (Campbell, Lo, & MacKinlay, 1996).

Para poder entender el comportamiento de las acciones como un conjunto, se deben establecer las relaciones y medidas estadísticas que definan su interacción en un portafolio (Markowitz, 1952). Sean S y Y variables aleatorias discretas, se definen las ecuaciones (5), (6), (7) y (8) de la siguiente manera (Rincón, 2014):

$$E[S] = \sum_s sP(S = s) \quad (5)$$

$$V[S] = E[S^2] - E[S]^2 \quad (6)$$

$$Cov[S, Y] = E[SY] - E[S]E[Y] \quad (7)$$

$$\rho[S, Y] = \frac{Cov[S, Y]}{\sqrt{V(S)V(Y)}} \quad (8)$$

Con:

- $E[S]$ el valor esperado de S .
- $V[S]$ la varianza de S .

- $Cov[S, Y]$ la covarianza entre S y Y .
- $\rho[S, Y]$ la correlación entre S y Y .

El análisis clásico de acciones para la conformación de portafolios está basado en la regla de media-varianza ($E-V$) la cual dicta que el inversionista preferirá los portafolios en los que se encuentre una mínima varianza dada una cierta media, o bien, una media máxima dada una cierta varianza. Esta regla está diseñada para cumplir con tres supuestos principales: Maximización de ganancias, aversión al riesgo y aleatoriedad en las acciones (Markowitz, 1952).

Tomando en cuenta un portafolio X , conformado por las acciones X_1, X_2, \dots, X_N , cada una en proporción $\alpha_1, \alpha_2, \dots, \alpha_N$ con $0 \leq \alpha_i \leq 1$ y $\sum_{i=1}^N \alpha_i = 1$; se pueden definir la media (ecuación 9) y la varianza (ecuación 10) del portafolio como (Markowitz, 1952):

$$E[X] = \sum_{i=1}^N \alpha_i \mu_i \quad (9)$$

$$V[X] = \sum_{i=1}^N \sum_{j=1}^N \sigma_{ij} \alpha_i \alpha_j \quad (10)$$

Con:

- $\mu_i = E[X_i]$ la media de la acción i .
- $\sigma_{ij} = Cov[X_i, X_j]$ la covarianza entre la acción i y la acción j .

Las distintas combinaciones de α_i resultarán en los posibles portafolios de acciones, mismos que se pueden observar de manera gráfica en el espacio (E, V) en la Figura 1. De todas las combinaciones, los que cumplan con los supuestos de maximización de ganancias y aversión al riesgo serán llamados "Frontera eficiente" (Markowitz, 1952).

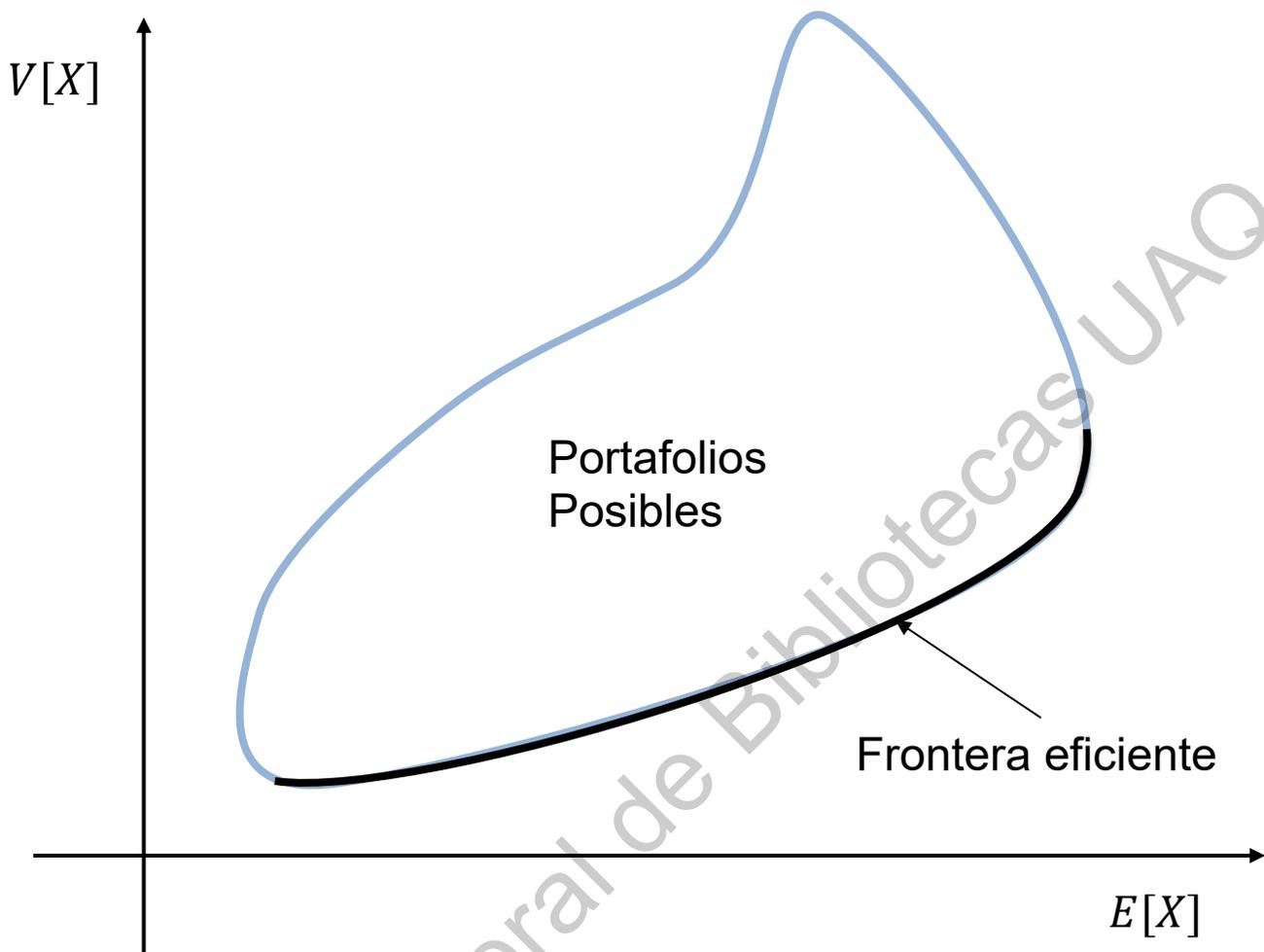


Figura 1. Frontera eficiente. Fuente: elaboración propia.

Esta metodología desarrollada por Markowitz, conocida como la teoría moderna de portafolios, introdujo una regla óptima para la distribución de la riqueza en diferentes activos de riesgo en un solo periodo. Sin embargo, este procedimiento es la segunda etapa del proceso de selección de un portafolio, la primera etapa consiste en encontrar los primeros 2 momentos centrales (media y varianza) de los activos de riesgo que conformaran el portafolio (Navarro, 2013). Como estos 2 momentos son desconocidos, deben de ser estimados.

El método más sencillo y utilizado es el de la media-varianza muestral, el cual, como su nombre lo indica, consiste en obtener la media y varianza muestrales a partir de los datos históricos de los activos de riesgo y posteriormente utilizarlos en el modelo

de Markowitz (Navarro, 2013). No obstante, este método conlleva errores de estimación y se ha mostrado que pequeños cambios en los rendimientos esperados pueden ocasionar grandes cambios en las posiciones del portafolio (Batsleer, 2016).

Existen varias alternativas para la estimación de los parámetros cuyo enfoque es el teorema de Bayes, estas metodologías se basan en el nivel de incertidumbre que existe sobre los parámetros desconocidos (media y varianza) de los activos de riesgo (Batsleer, 2016). El inversionista considera la distribución de estos parámetros tomando en cuenta el teorema de Bayes y las distribuciones a priori de estos parámetros. Las diferentes metodologías dependen de la distribución a priori elegida, varias metodologías se basan en distribuciones a priori difusas (Bawa, Brown, & Klein, 1979). Estas distribuciones le asignan la misma probabilidad a cada posible valor por lo que son consideradas distribuciones a priori no informativas (Navarro, 2013). Otros autores utilizan distribuciones a priori basadas en distintos métodos de valoración de activos como el CAPM u otras combinaciones de modelos.

Otros autores se han enfocado en las varianzas de los estimadores y han mostrado que la matriz de covarianzas puede ser estimada con mayor precisión al no basarse exclusivamente en los rendimientos esperados (Kempf & Memmel, 2003). El método más conocido de este tipo es el del portafolio de mínima varianza, en el cual, se eligen los pesos de los activos de riesgo sin tomar en cuenta los rendimientos esperados. De esta manera se reduce el número de parámetros a estimar y con ello el error de estimación, pero a costa de una pérdida de información (Navarro, 2013). Para encontrar el portafolio en el cual la varianza total es minimizada, basta con encontrar el punto con la menor varianza dentro de la frontera eficiente (Figura 1) (Batsleer, 2016).

2.3 Acciones como series de tiempo

Una serie de tiempo es una colección de observaciones ordenadas de forma secuencial y, típicamente, se encuentran espaciadas en el tiempo de manera uniforme (Green, 2011).

El objetivo principal del análisis de series de tiempo consiste en desarrollar modelos matemáticos que ayuden a describir una muestra de datos (Shumway & Stoffer, 2015). Una característica especial del análisis de series de tiempo es el hecho de que debe de tomar en cuenta el orden del tiempo, ya que las observaciones sucesivas usualmente no son independientes, mientras que, en la mayoría de la teoría estadística se trabaja con muestras aleatorias de observaciones independientes (Green, 2011).

Los métodos de análisis de series de tiempo constituyen un área importante en la estadística y, a pesar de que existen muchas aplicaciones para este tipo de análisis, todas pueden ser clasificadas como descriptivas, explicativas, predictivas o de control (Bowerman & O'Connell, 1987). Usualmente, las series de tiempo son estudiadas con el propósito de descubrir patrones históricos que puedan ser utilizados para hacer proyecciones. Para poder identificar estos patrones, resulta conveniente pensar en una serie de tiempo como un conjunto compuesto de varios componentes (Green, 2011).

Los componentes principales de una serie de tiempo son tendencia, variación estacional, cambios cíclicos y factores irregulares (Green, 2011). La tendencia es el cambio en la media a largo plazo y usualmente se interpreta como el componente de crecimiento (decrecimiento) de la serie. La variación estacional se refiere a las fluctuaciones periódicas que existen en la serie y suelen ser atribuidas a factores sociales o climáticos. Los cambios cíclicos dentro de una serie de tiempo son similares a los estacionales en el sentido de que siguen un patrón parecido al de una onda; estos cambios son variaciones que existen en un periodo de tiempo fijo, debido a factores distintos a los estacionales. Una vez que se toman en cuenta la tendencia y las variaciones cíclicas, el movimiento restante es atribuido a fluctuaciones irregulares y estos datos son considerados como una serie de residuos, los cuales deben de ser

analizados para determinar si efectivamente todas las variaciones cíclicas han sido removidas (Green, 2011).

Para poder trabajar con series de tiempo es necesario establecer una definición completa y formal. Si se tiene una colección de n variables aleatorias correspondientes a los tiempos t_1, t_2, \dots, t_n , para cualquier n entero positivo, se define una serie de tiempo como la función de distribución conjunta representada en la ecuación (11) (Shumway & Stoffer, 2015):

$$F(c_1, c_2, \dots, c_n) = P(x_{t_1} \leq c_1, x_{t_2} \leq c_2, \dots, x_{t_n} \leq c_n). \quad (11)$$

Desafortunadamente, la función de distribución multivariada de la ecuación (11) no es fácil de obtener a menos que las variables aleatorias tengan una distribución normal conjunta. Es por esto que a pesar de que la función de distribución conjunta describe a los datos de manera completa, es inusual utilizarla para analizar a las series de tiempo. La función de distribución debe ser evaluada como una función de n argumentos por lo que se definen las siguientes ecuaciones como las medidas de dependencia necesarias para poder examinar el comportamiento de la serie (Shumway & Stoffer, 2015):

$$F_t(x) = P(x_t \leq x) \quad (12)$$

$$f_t(x) = \frac{\partial F_t(x)}{\partial x} \quad (13)$$

$$\mu_{x_t} = E[x_t] = \int_{-\infty}^{\infty} x f_t(x) dx \quad (14)$$

$$\gamma_x(s, t) = Cov(x_s, x_t) = E[(x_s - \mu_{x_s})(x_t - \mu_{x_t})] \quad (15)$$

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}} \quad (16)$$

$$\gamma_{xy}(s, t) = Cov(x_s, y_t) = E[(x_s - \mu_{x_s})(y_t - \mu_{y_t})] \quad (17)$$

$$\rho_{xy}(s, t) = \frac{\gamma_{xy}(s, t)}{\sqrt{\gamma_x(s, s)\gamma_y(t, t)}} \quad (18)$$

Con:

- $F_t(x)$ la función de distribución marginal
- $f_t(x)$ la función de densidad marginal
- $\mu_{x_t} = E[x_t]$ la función de la media o esperanza
- $\gamma_x(s, t)$ la función de autocovarianza
- $\rho(s, t)$ la función de autocorrelación
- $\gamma_{xy}(s, t)$ la función de covarianza cruzada
- $\rho_{xy}(s, t)$ la función de autocorrelación cruzada

Las definiciones anteriores de las funciones de media y autocovarianza, ecuaciones (14) y (15), son completamente generales. A pesar de que no se ha hecho ningún tipo de suposición acerca del comportamiento de las series de tiempo, se puede imaginar la posibilidad de que un tipo de regularidad exista en el tiempo sobre el comportamiento de alguna serie de tiempo. A este concepto de regularidad en el tiempo se le llama estacionariedad (Shumway & Stoffer, 2015).

Una serie de tiempo estrictamente estacionaria es aquella en que el comportamiento probabilístico de cualquier colección de valores, $x_{t_1}, x_{t_2}, \dots, x_{t_k}$, es idéntico al comportamiento de, $x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}$, esto se muestra en la ecuación (19) (Shumway & Stoffer, 2015):

$$P(x_{t_1} \leq c_1, x_{t_2} \leq c_2, \dots, x_{t_k} \leq c_k) = P(x_{t_1+h} \leq c_1, x_{t_2+h} \leq c_2, \dots, x_{t_k+h} \leq c_k) \quad (19)$$

Una serie débilmente estacionaria, x_t , es una serie de tiempo con varianza finita tal que

- La función de la media, μ_t , es constante y no depende del tiempo t
- La función de autocovarianza, $\gamma(s, t)$, depende en s y t únicamente a través de su diferencia $|s - t|$.

Se define el Ruido Blanco, W_t , como una sucesión de variables aleatorias independientes e idénticamente distribuidas con media 0 y varianza σ_w^2 (Shumway & Stoffer, 2015) como se muestra en la ecuación (20),

$$W_t \sim wn(0, \sigma_w^2) \quad (20)$$

Para definir una regresión lineal en el contexto de series de tiempo, se asume que una serie de tiempo dependiente, x_t , para $t = 1, 2, \dots, n$, está siendo influenciada por una colección de series independientes, $z_{t_1}, z_{t_2}, \dots, z_{t_q}$ de tal manera que se puede expresar esta relación a través del modelo de regresión lineal expresado en la ecuación (21) (Shumway & Stoffer, 2015):

$$x_t = \beta_0 + \beta_1 z_{t_1} + \beta_2 z_{t_2} + \dots + \beta_q z_{t_q} + w_t \quad (21)$$

Con:

- $\beta_0, \beta_1, \dots, \beta_q$ coeficientes fijos y desconocidos de la regresión
- $w_t \sim wn(0, \sigma_w^2)$ ruido blanco

La regresión clásica suele ser insuficiente para explicar toda la dinámica de una serie de tiempo. Si se introduce la correlación como un fenómeno que depende de la separación o diferencia de tiempo, *lag*, de una relación lineal; se pueden proponer modelos autorregresivos y modelos autorregresivos de media móvil (Shumway & Stoffer, 2015).

Los modelos autorregresivos se basan en la idea de que los valores actuales de una serie de tiempo, x_t , pueden ser explicados como una función de p valores anteriores, $x_{t-1}, x_{t-2}, \dots, x_{t-p}$, donde p determina el número de observaciones pasadas necesarias para predecir el valor actual. La ecuación (22) muestra la definición de un modelo autorregresivo de orden p , $AR(p)$ (Shumway & Stoffer, 2015),

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \quad (22)$$

Con:

- x_t débilmente estacionario
- $\phi_1, \phi_2, \dots, \phi_p$ son constantes y $\phi_p \neq 0$
- $w_t \sim wn(0, \sigma_w^2)$

Si la media, μ del proceso anterior, x_t , es diferente de cero, se debe reemplazar x_t por $x_t - \mu$ como se muestra en las ecuaciones (23) y (24),

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \dots + \phi_p(x_{t-p} - \mu) + w_t \quad (23)$$

O implícitamente

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \quad (24)$$

Donde $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$

Diferentes autores han utilizado a las series de tiempo para analizar el comportamiento de las acciones y principalmente para realizar predicciones. Muchas series de datos, como es el caso de las acciones, se comportan como si no tuvieran una media fija; sin embargo, suelen exhibir homogeneidad en el sentido de que una parte de la serie de tiempo se comporta de manera muy similar a cualquier otra parte de la serie de tiempo (Green, 2011). Para describir este tipo de modelos que presentan una homogeneidad no estacionaria, suele asumirse que alguna diferencia del proceso es estacionaria; tal es el caso de los modelos ARIMA.

Desafortunadamente, a pesar de que es posible proyectar valores a través de los modelos ARIMA, no se recomienda utilizarlos como una herramienta de inversión ya que suelen tener errores grandes de estimación (Green, 2011). Cuando se trabaja con un modelo de predicción que está basado únicamente en los precios anteriores de una acción, por más correlacionados que estos estén, se ven afectados por múltiples factores externos que no pueden ser predichos por los precios pasados.

A pesar de que este tipo de modelos no son los adecuados para realizar proyecciones sobre los precios de las acciones, tiene sentido utilizar a las series de tiempo para modelar este tipo de datos (Green, 2011).

El estudio de acciones como series de tiempo realizado por Green en 2011, revela que las acciones no siempre se mueven conforme a las acciones del sector al que pertenecen; sino que su comportamiento está más relacionado con los factores particulares de la misma.

2.4 Clusterización como herramienta de análisis

La clusterización (derivado del término *clustering* en inglés) se refiere a diversas técnicas estadísticas para encontrar subgrupos dentro de una colección de datos. En estos procesos se debe cumplir que las observaciones que se encuentren en el mismo clúster sean parecidos entre sí, mientras que las observaciones catalogadas en clústeres diferentes tendrán características diferentes (James, Witten, Hastie, & Tibshirani, 2017).

Este tipo de herramientas son catalogadas como *aprendizaje estadístico no supervisado*, es decir, que como su objetivo es encontrar una estructura previamente desconocida en una colección de datos, no existen datos para comparar los resultados que arroje el proceso (James, Witten, Hastie, & Tibshirani, 2017).

El primer método por definir es K-medias. Este método pretende agrupar la base de datos en K clústeres no superpuestos. El número de clústeres, K , deberá ser especificado de manera externa al proceso (James, Witten, Hastie, & Tibshirani, 2017).

Definimos C_1, C_2, \dots, C_K como los conjuntos que contienen las observaciones. Los conjuntos deben cumplir que (James, Witten, Hastie, & Tibshirani, 2017):

1. $C_1 \cup C_2 \cup \dots \cup C_K = \{1, \dots, n\}$. Todas las observaciones deben de estar asignadas a algún clúster.
2. $C_k \cap C_{k'} = \emptyset$ para todo $k \neq k'$. Ninguna observación puede estar asignada a dos clústeres diferentes.

K-medias pretende clusterizar bajo la premisa de que es necesario minimizar la variación inter-clúster como se aprecia en la ecuación (26). Dicha variación puede ser definida para cualquier clúster C_k como una medida $W(C_k)$. Hay muchas maneras de definir la medida W , pero la más común es la *Distancia Euclidiana Cuadrática* que se muestra en la ecuación (25) (James, Witten, Hastie, & Tibshirani, 2017):

$$DEC = \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \quad (25)$$

$$\underset{C_1, C_2, \dots, C_k}{\text{minimizar}} \left\{ W(C_k) = \frac{1}{|C_k|} \sum_{i, i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \right\} \quad (26)$$

Con:

- DEC la Distancia Euclidiana Cuadrática.
- $|C_k|$ el número de observaciones en el clúster k .
- p el número de dimensiones del modelo.
- x_{ij} el valor de la observación i en la dimensión j .

Debido a que existen casi K^n formas de distribuir n observaciones en K clústeres, se debe recurrir a algoritmos computacionales para realizar esta tarea. El más común para encontrar mínimos locales es (James, Witten, Hastie, & Tibshirani, 2017):

Algoritmo 1 Clusterización K-medias

1. Asignar aleatoriamente a cada una de las observaciones un número del 1 al K .
2. Iterar los siguientes pasos hasta que los clústeres dejen de cambiar:
 - a. Calcular el centroide para cada uno de los clústeres. El centroide se define como la media de las observaciones para las p dimensiones.
 - b. Redistribuir las observaciones para asignarlas a los clústeres más cercanos, considerando la distancia Euclidiana.

Una de las desventajas más grandes de K-medias es que requiere la especificación previa del número de clústeres. Una alternativa que carece de este

problema es la *Clusterización Jerárquica*. Existen distintos tipos de clusterización jerárquica, pero en esta sección se describirá la aglomerante, la cual tiene una representación gráfica muy útil llamada dendograma (James, Witten, Hastie, & Tibshirani, 2017).

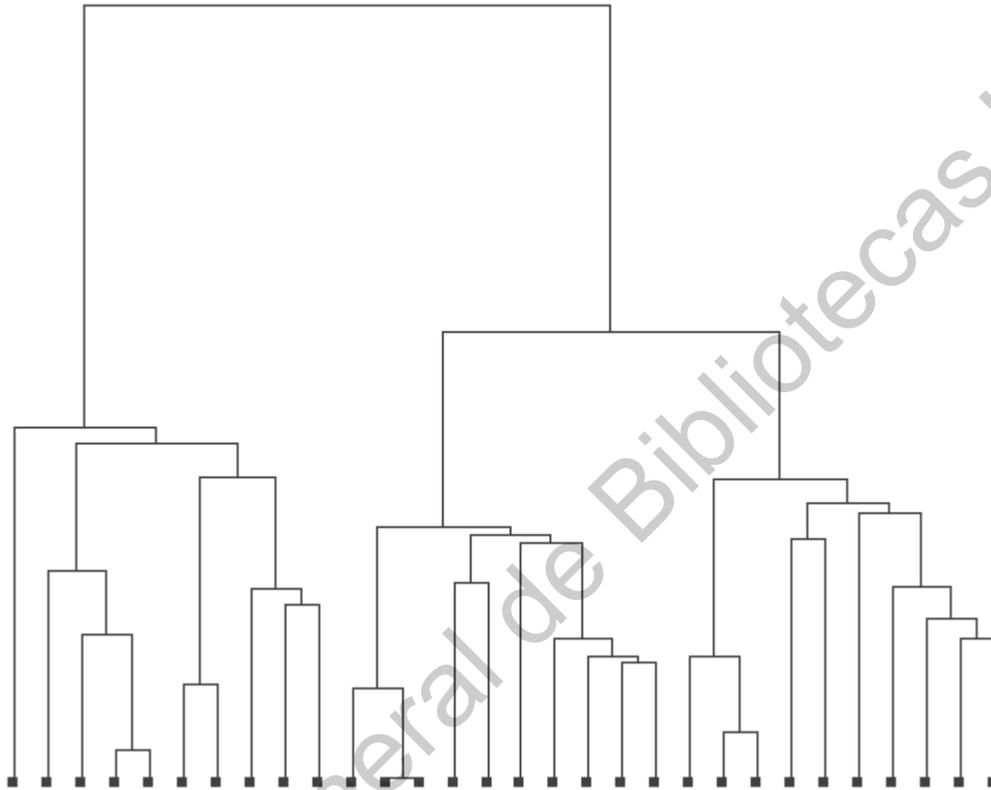


Figura 2. Dendograma. Fuente: elaboración propia.

La Figura 2 se visualiza de abajo hacia arriba, los primeros puntos son las observaciones de la colección de datos. Avanzando un poco, cada una de estas ramas se van uniendo, de acuerdo a que tan similares son, formando así los clústeres. Se puede observar que, dependiendo de la altura que se elija en la gráfica, se puede crear una clusterización de las observaciones con cualquier número de clústeres entre 1 y n . Cabe aclarar que en la parte inferior de la gráfica los clústeres que se forman tienen observaciones muy parecidas entre sí, mientras que en la parte superior las observaciones tienen más variación (James, Witten, Hastie, & Tibshirani, 2017).

Para definir el algoritmo de clusterización es necesaria una medida de disimilitud. Supondremos, al igual que en K-medias, que la medida es la distancia Euclidiana (ecuación 26), entonces el proceso de clusterización jerárquica sería (James, Witten, Hastie, & Tibshirani, 2017):

Algoritmo 2 Clusterización Jerárquica

1. Identificar cada una de las observaciones como su propio clúster (la base del dendograma).
2. Para $i = n, n - 1, \dots, 2$:
 - a. Calcular la disimilitud que existe entre cada una de las parejas posibles de clústeres, y fusionar las más parecidas. La disimilitud entre estos dos clústeres indicará la altura que tendrá la unión en el dendograma.
 - b. Como ahora se fusionaron dos clústeres, se tienen $i - 1$ clústeres restantes, así que debe repetirse el proceso hasta contar con únicamente dos.

Dentro del algoritmo se puede notar que la disimilitud es directa cuando el número de observaciones y de clústeres es el mismo, ya que se utiliza la información entre observaciones. En el momento en que un clúster tiene más de una observación, se deberán establecer criterios de disimilitud, para así determinar el papel que juegan las observaciones dentro del clúster. Los criterios de disimilitud más comunes son (James, Witten, Hastie, & Tibshirani, 2017):

Criterio	Descripción
Completa	Disimilitud máxima. Se calculan las disimilitudes de todos los pares de observaciones (a, b) con $a \in \text{Clúster } A$ y $b \in \text{Clúster } B$ y se elige la máxima.

Individual	Disimilitud mínima. Se calculan las disimilitudes de todos los pares de observaciones (a, b) con $a \in \text{Clúster } A$ y $b \in \text{Clúster } B$ y se elige la mínima.
Promedio	Disimilitud promedio. Se calculan las disimilitudes de todos los pares de observaciones (a, b) con $a \in \text{Clúster } A$ y $b \in \text{Clúster } B$ y se calcula el promedio.
Centroide	Se calculan los centroides de los clústeres A y B , posteriormente se calcula su disimilitud.

Tabla 1. Criterios de disimilitud (Handl, Knowles, & Kell, 2005).

Siendo la clusterización una rama relativamente nueva de las aplicaciones estadísticas, no existe un consenso sobre el mejor método de clusterización para una cierta colección de datos (ya que las características varían según las observaciones que se tengan), por lo que es muy común que los investigadores utilicen diferentes métodos para posteriormente evaluarlos visualmente. Esta forma de validación es subjetiva, así que se tiene el peligro de preferir resultados que refuercen las hipótesis, ignorando los que muestran situaciones inesperadas (Handl, Knowles, & Kell, 2005).

Se distinguen dos tipos de medidas para la validación de clústeres: las medidas externas comprenden a todas aquellas que funcionan conociendo la clasificación correcta de las observaciones, mientras que las internas utilizan únicamente la información contenida en la colección de datos (Halkidi, Batistakis, & Vazirgiannis, 2001).

Tipo	Medida	Descripción
Externa	Unaria	Compara un resultado de clusterización contra los grupos reales. Se compara la pureza (fracción del clúster que forma parte de la agrupación real

		predominante) y la completitud (fracción de la parte real predominante contenida en el clúster).
Externa	Binaria	Compara la similitud entre varios resultados de clusterización.
Interna	Compacidad	Compara la homogeneidad entre clústeres con la variación inter-clúster.
Interna	Conectividad	Mide a que nivel una partición toma en cuenta las densidades locales y agrupa las observaciones con sus vecinos más cercanos.
Interna	Separación	Cuantifica el nivel de separación entre clústeres.
Interna	Poder/estabilidad de predicción	Remuestran y perturban las observaciones y repiten la clusterización para medir su estabilidad.
Interna	Paridad entre partición y distancia	Miden que tanto la información de distancia entre las observaciones originales se conserva en los clústeres.
Interna	Correlación	Explotan las redundancias y correlación entre las observaciones.

Tabla 2. Tipos de medidas (Handl, Knowles, & Kell, 2005).

La mayoría de las medidas internas son utilizadas para determinar el número de clústeres en el proceso. Al realizar la clusterización con diferentes parámetros y comparando su efectividad con las medidas, se puede llegar a un “número óptimo” de clústeres (Handl, Knowles, & Kell, 2005).

El ancho de silueta es una medida interna de conectividad de clústeres, calculada como el promedio de las siluetas de las observaciones individuales, y valúa el grado de confianza en la asignación de una observación a un clúster. Para interpretar el ancho de

silueta, se desea que el valor sea lo más cercano a 1 posible. Se define la silueta de la observación i en la ecuación (27) como (Brock, Pihur, Datta, & Datta, 2008):

$$S(i) = \frac{b_i - a_i}{\max(b_i, a_i)} \quad (27)$$

Con:

- a_i la distancia media entre la observación i y todas las demás observaciones en el clúster.
- $b_i = \min_{C_k \in \mathcal{C} \setminus \mathcal{C}(i)} \sum_{j \in C_k} \frac{\text{dist}(i, j)}{n(C_k)}$ la distancia promedio entre la observación i y las observaciones del clúster más cercano.
- $\mathcal{C}(i)$ el clúster que contiene la observación i .
- \mathcal{C} el conjunto de clústeres.
- $\text{dist}(i, j)$ la distancia entre i y j .
- $n(C_k)$ la cardinalidad de C_k .

Entonces se puede obtener la media de los $S(i)$ para calcular el ancho de silueta.

Una de las mayores críticas de la optimización de portafolios basada en medias y varianzas es que tiende a magnificar el efecto del error estándar. Este proceso de optimización es extremadamente sensible a las variaciones del rendimiento esperado y la matriz de covarianzas, lo cual se traduce en una mala diversificación del portafolio. La capacidad que tiene la clusterización para encontrar estructura oculta en una colección de datos hace que sea una herramienta útil para lidiar con este problema (Ren, Portfolio Construction Using Clustering Methods, 2005).

Las acciones suelen presentar fuertes correlaciones entre unas y otras, especialmente si pertenecen al mismo sector económico. Esto permite agrupar a las acciones con alta correlación en un mismo clúster y de esta manera obtener clústeres con baja correlación entre ellos. Si se piensa en estos clústeres como acciones individuales, se tiene un grupo de acciones con una baja correlación a partir del cual se

puede encontrar un portafolio eficiente con el método de Markowitz (Ren, Portfolio Construction Using Clustering Methods, 2005).

Un ejemplo práctico de esto es un trabajo que presenta una clasificación topológica de las 30 acciones con mayor movimiento de la Bolsa de Valores de Estocolmo basándose en el co-movimiento entre acciones individuales y compara los resultados obtenidos con el *Standard & Poor's Global Industry Classification Standard* (GICS). Se trabajó bajo la hipótesis de que a través de un espacio ultramétrico es posible crear grupos homogéneos de acciones. Se realizó una clusterización jerárquica a través de la matriz de correlación de los pares de las 30 acciones de la SSE considerando sus rendimientos logarítmicos diarios. Utilizando el criterio de disimilitud promedio se obtuvieron cinco clústeres principales los cuales, en su mayoría, resultaron homogéneos con respecto a su sector y en varios casos con respecto a su sub-industria, según la clasificación del GICS (Rosén, Correlation based clustering of the Stockholm Stock Exchange, 2006).

Los resultados muestran que es posible obtener grupos representativos y homogéneos basándose únicamente en la correlación entre acciones individuales, sin la necesidad de hacer suposiciones acerca de las actividades económicas de las empresas. Los clústeres obtenidos indican que existen factores económicos que afectan ciertos grupos de acciones independientemente de su clasificación según el GICS (Rosén, Correlation based clustering of the Stockholm Stock Exchange, 2006).

3. METODOLOGÍA

Debido a que el objetivo de este trabajo es comparar la caracterización en forma de clústeres de 25 acciones de la Bolsa Mexicana de Valores utilizando el método de media-varianza y la autocorrelación en sus series de tiempo, será necesario llevar en paralelo los procesos homólogos para los dos métodos en la presente metodología.

Se empieza por definir la colección de datos para analizar. Se trata de la información de precios mensuales de las acciones de la Bolsa Mexicana de Valores (BMV) de enero del 2003 a diciembre del 2016. Las 25 acciones utilizadas se muestran en la siguiente tabla:

ALFAA	CEMEXCPO	GFINBURO	KUOB	GMD
ALSEA	CMOCTEZ	GFNORTEO	MEXCHEM	CIDMEGA
ARA	CULTIBAB	GMEXICOB	PE&OLES	GCC
BACHOCOB	CYDSASAA	ICHB	SORIANAB	HERDEZ
BIMBOA	ELEKTRA	KIMBERA	TLEVISACPO	WALMEX

Tabla 3. Acciones para analizar. Fuente: elaboración propia.

La información se delimita de esta manera para facilitar el manejo de la base de datos. A pesar de que se cuenta con el registro de 60 acciones desde enero del 2000 a mayo del 2018, dicha información es incompleta o incongruente, de forma que la muestra seleccionada proporciona mayor confianza que la base completa.

Posteriormente, se calculan los rendimientos logarítmicos mensuales por acción usando la ecuación (4), obteniendo la información de todas las acciones en la misma escala. Esta pequeña modificación servirá como base para los estudios de media-varianza y análisis de series de tiempo.

A continuación; para cada acción se calcula la media, varianza y autocorrelación utilizando las ecuaciones (9), (10) y (16) hasta un lag $p = |s - t|$ que permita caracterizar cada acción como una serie de tiempo, es decir, bajo el supuesto de un modelo autorregresivo de orden p tal como se muestra en la ecuación (24). Además, es importante mantener la posible estructura estacional derivada de algún fenómeno anual en las acciones sin saturar el cálculo de clústeres. Esto se logra eligiendo el mayor

número posible de lags p sin repercutir en la formación de clústeres evaluados por su ancho de silueta con la ecuación (27); esto se puede medir a través de las ecuaciones (28) y (29):

$$m_1 = \ln(k) \bar{S}^{\frac{1}{3}} \quad (28)$$

$$m_2 = \ln(k * p) \bar{S}^{\frac{1}{3}} \quad (29)$$

Con:

- m_i medida de optimización
- k el número de clústeres
- \bar{S} el promedio del ancho de silueta (ecuación 27)
- p el número de lags $|s - t|$

Para una mejor referencia, la metodología de este trabajo se describe a continuación en forma de apartados como sigue:

- I. Tomando en cuenta la información anterior se procede a utilizar la metodología de clusterización descrita en la sección 2.4, algoritmos (1) y (2), junto con la distancia euclidiana en la ecuación (26) mediante el siguiente proceso:
 - a. Se calcula la media, varianza y autocorrelación de la acción mediante las ecuaciones (9), (10) y (16) hasta un lag que permita caracterizar cada acción como una serie de tiempo usando la ecuación (24).
 - b. Se realiza la clusterización con la información de media y varianza.
 - i. Se elige un número de clústeres (entre 2 y 24).
 - ii. Se realiza el algoritmo (1) de clusterización K-medias.
 - iii. Se realiza el algoritmo (2) de clusterización Jerárquica.
 - iv. Se calcula la medida m_1 para ambas clusterizaciones mediante la ecuación (28).

- v. Se repiten los pasos del numeral (i) al (iv) para todos los números posibles de clústeres, y se selecciona el que maximice m_1 .
- c. Se realiza la clusterización con la información de media, varianza y autocorrelación.
 - i. Se elige un número de lags.
 - ii. Se elige un número de clústeres (entre 2 y 24).
 - iii. Se realiza el algoritmo (1) de clusterización K-medias.
 - iv. Se realiza el algoritmo (2) de clusterización Jerárquica.
 - v. Se calcula la medida m_2 para ambas clusterizaciones mediante la ecuación (29).
 - vi. Se repiten los pasos del numeral (i) al (v) para todas las combinaciones posibles de lags y número de clústeres, y se selecciona la que maximice m_2 .
- II. Una vez realizadas las dos clusterizaciones con los parámetros óptimos, se realiza el siguiente proceso:
 - a. Se realiza una validación cualitativa de los clústeres usando gráficas de dispersión y en caso de ser necesario utilizar el análisis de componentes principales para reducir las dimensiones del modelo (Smith, I.; 2002).
 - b. Posteriormente se ejecuta el algoritmo de análisis de portafolios aislando las acciones por clúster.
 - i. Para cada clúster, se calcula el rendimiento promedio suponiendo una ponderación uniforme en las acciones que lo constituyen.
 - ii. Se minimiza la desviación del portafolio formado por las acciones del clúster, y se calculan los rendimientos mensuales de este.
 - c. Una vez teniendo los rendimientos mensuales de los portafolios, se fija un rendimiento objetivo, y se minimiza la desviación de un portafolio compuesto de los portafolios anteriores.

Finalmente, al comparar cualitativamente los clústeres creados por ambas metodologías, se espera comprobar que la información obtenida a través de la caracterización de las acciones como series de tiempo es fundamentalmente distinta a la recuperada por la metodología media-varianza.

Dirección General de Bibliotecas UAQ

4. RESULTADOS Y DISCUSIÓN

En el presente capítulo se muestran los resultados obtenidos a partir de la metodología anteriormente descrita. Para mayor claridad, cada vez que se efectúe un paso del capítulo anterior, aparecerá entre paréntesis el inciso al que hace referencia.

Como un primer análisis se muestra en la figura 3 la dispersión obtenida de las 25 acciones analizadas a partir de la media y desviación estándar de sus rendimientos logarítmicos (*l.a*).

Dispersión de las acciones

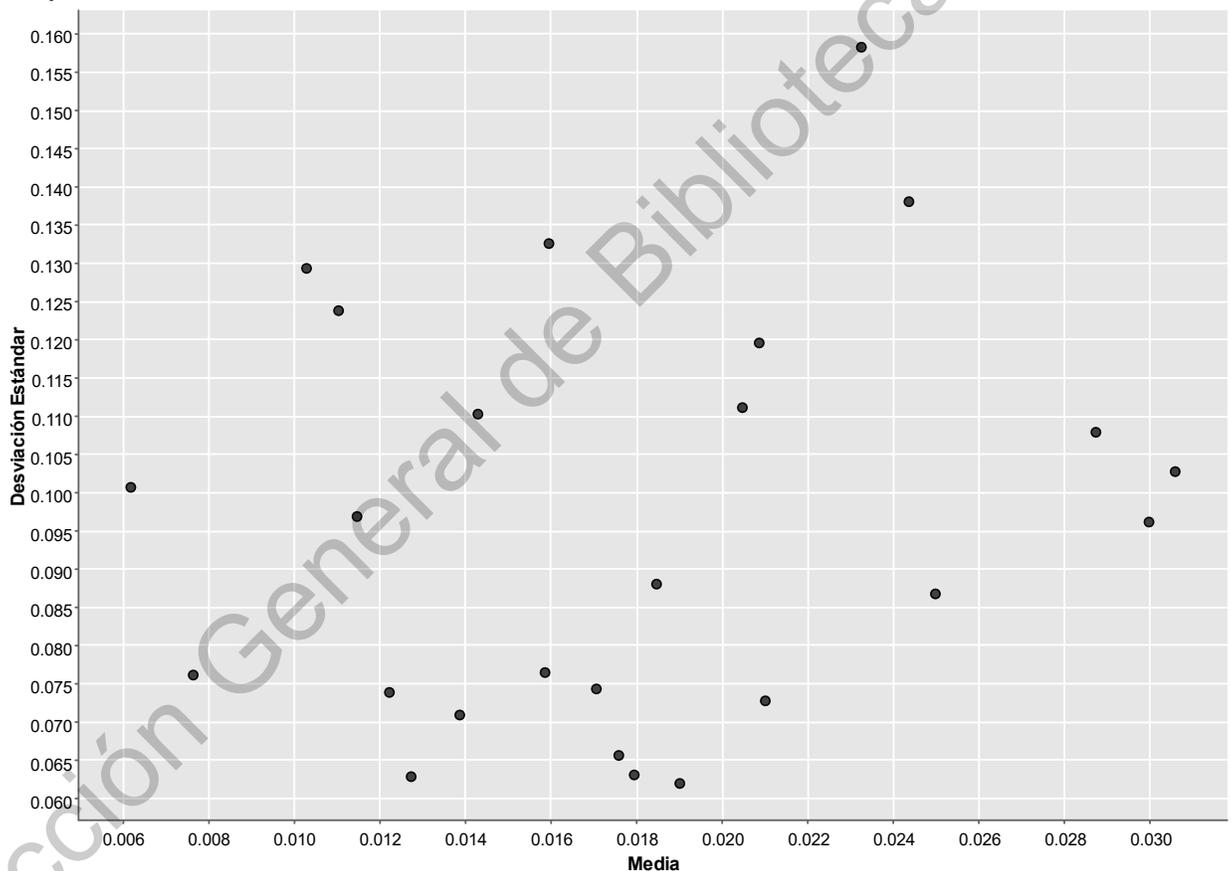


Figura 3. Dispersión de las acciones. Fuente: elaboración propia.

En la figura 3 se puede apreciar visualmente cuáles acciones presentan la mayor o menor desviación estándar, a través del eje vertical y las que presentan los mayores y menores rendimientos mensuales en promedio a través del eje horizontal. Se puede esperar que

con el método de clusterización de media-varianza (*I.b.*), las acciones que presentan la menor distancia entre sí en la figura 3 formen parte de un mismo clúster. De manera alternativa, si estos grupos de acciones que pertenecieron a los mismos clústeres con el método de media-varianza (*I.b.*) formaran parte de diferentes clústeres a la hora de hacer la clusterización con el método de media, varianza y autocorrelación (*I.c.*); pudiera ser un indicio de que la caracterización de las acciones con ambas metodologías es fundamentalmente distinta.

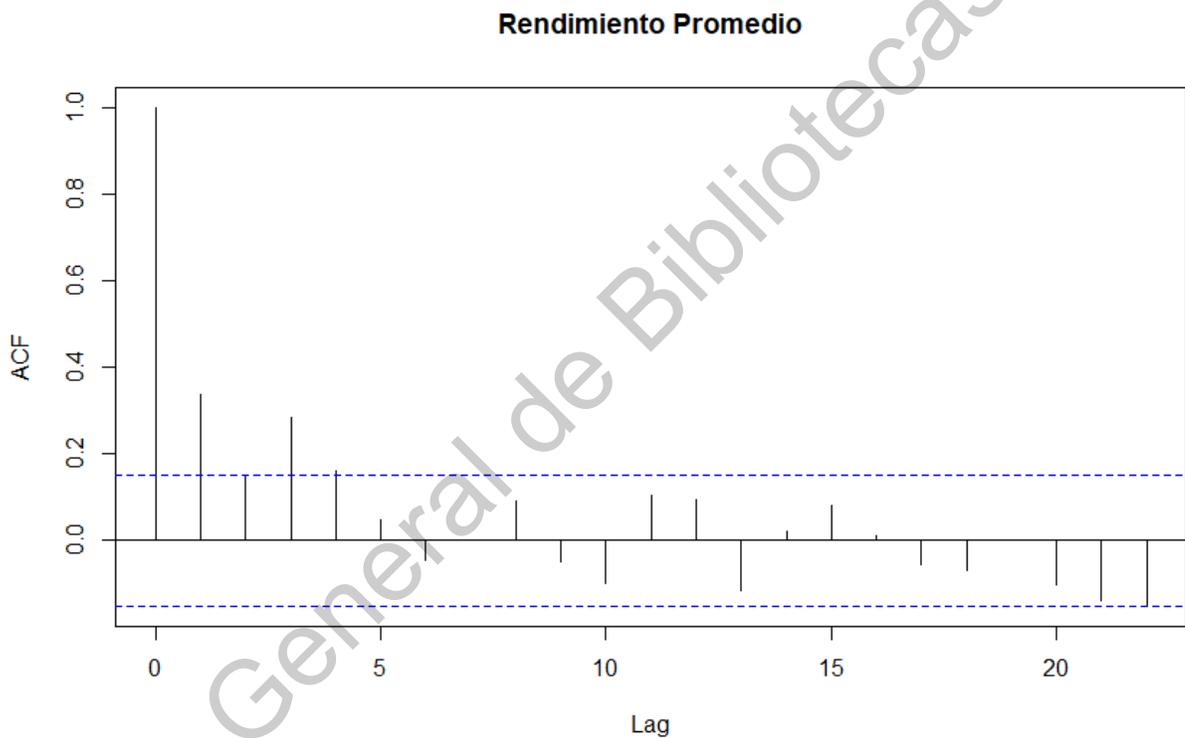


Figura 4. Función de autocorrelación del rendimiento del promedio de las acciones. Fuente: elaboración propia.

En la Figura 4 se muestran los primeros 22 lags obtenidos a partir de la función de autocorrelación del promedio de los rendimientos logarítmicos de las 25 acciones analizadas (*I.a*). Se muestran únicamente los primeros 22 debido al relativamente corto tiempo de observación de las acciones. En la figura se puede apreciar una autocorrelación significativa en los lags 1 y 3 debido a que sobrepasan el intervalo de 2

errores estándar, lo que sugiere que la serie de tiempo no es un ruido blanco (ecuación 20), permitiendo la implementación de un modelo autorregresivo. Se muestran los primeros 22 lags usando el promedio de las acciones de manera ilustrativa. Sin embargo, lo que se utiliza son los lags de las acciones individuales, los cuales se muestran en el anexo A.

Si se revisa con detenimiento las distintas gráficas del anexo A, se puede apreciar que las acciones se ajustan a distintos modelos y parámetros de series de tiempo. Derivado de que para el objetivo del presente trabajo únicamente es necesario comparar las características de las acciones entre sí, y no proyectar ni describir las mismas, la caracterización por medio de lags es suficiente, evitando así el extenso trabajo requerido para ajustar cada acción a su modelo óptimo de series de tiempo.

Las Figuras 5 y 6 presentan en el eje vertical la medida m_1 (*l.b.iv*) de cada uno de los números de clústeres posibles (*l.b.i*), mostrados en el eje horizontal, para el método de clusterización K-medias (*l.b.ii*) y jerárquico (*l.b.iii*), respectivamente.

Ancho de silueta

K-medias

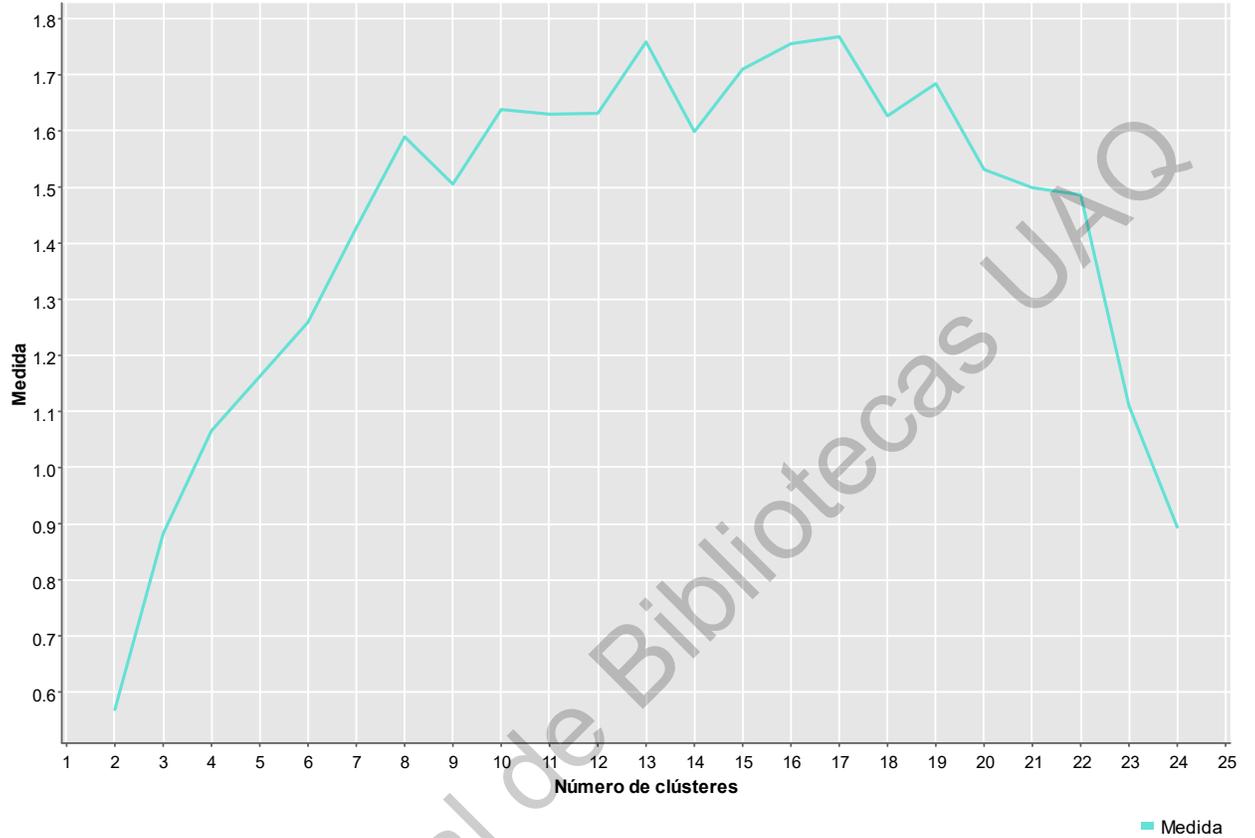


Figura 5. Medida m_1 para K-medias con media y desviación estándar. Fuente: elaboración propia.

Ancho de silueta

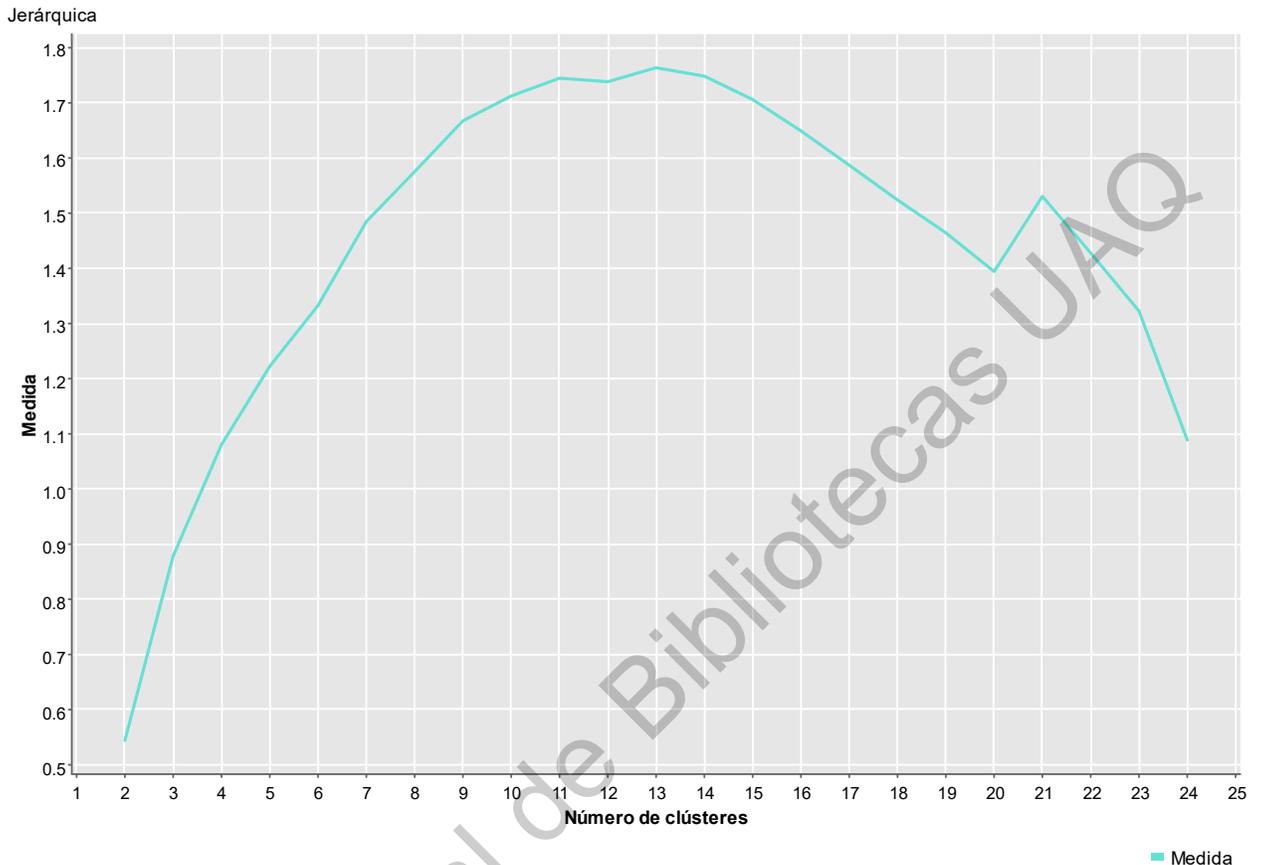


Figura 6. Medida m_1 para jerárquica con media y desviación estándar. Fuente: elaboración propia.

En las figuras 5 y 6 se puede notar que al utilizar un número muy bajo o muy alto de clústeres se obtuvo un valor bajo de la medida m_1 , esto se debe a 2 factores principalmente, por un lado, a la composición del ancho de silueta (ecuación 27) y por otro lado a la propia composición de la medida m_1 . La medida m_1 está en función de 2 variables: el número de clústeres k y el promedio del ancho de silueta \bar{S} ; mientras mayor sea k mayor será el valor de la medida m_1 , por lo tanto, el haber obtenido valores bajos de m_1 con un número alto de clústeres para las clusterizaciones de k-medias y jerárquica con media y desviación estándar se debe a que al hacer muchos clústeres se obtuvieron en promedio valores bajos del ancho de silueta.

Las Figuras 7 y 8 presentan en el eje vertical la medida m_2 (I.c.v) de cada uno de los números de clústeres posibles (I.c.ii), mostrados en el eje horizontal, para el método de clusterización K-medias (I.c.iii). Se muestran únicamente los lags 22 y 6 (I.c.i), dado que son las configuraciones con mayor información disponible y la que maximiza la medida m_2 , respectivamente. Sin embargo, se hicieron las pruebas con cada uno de los lags disponibles (de 1 a 22) y se encuentran en el anexo B.

Ancho de silueta

K-medias con ACF

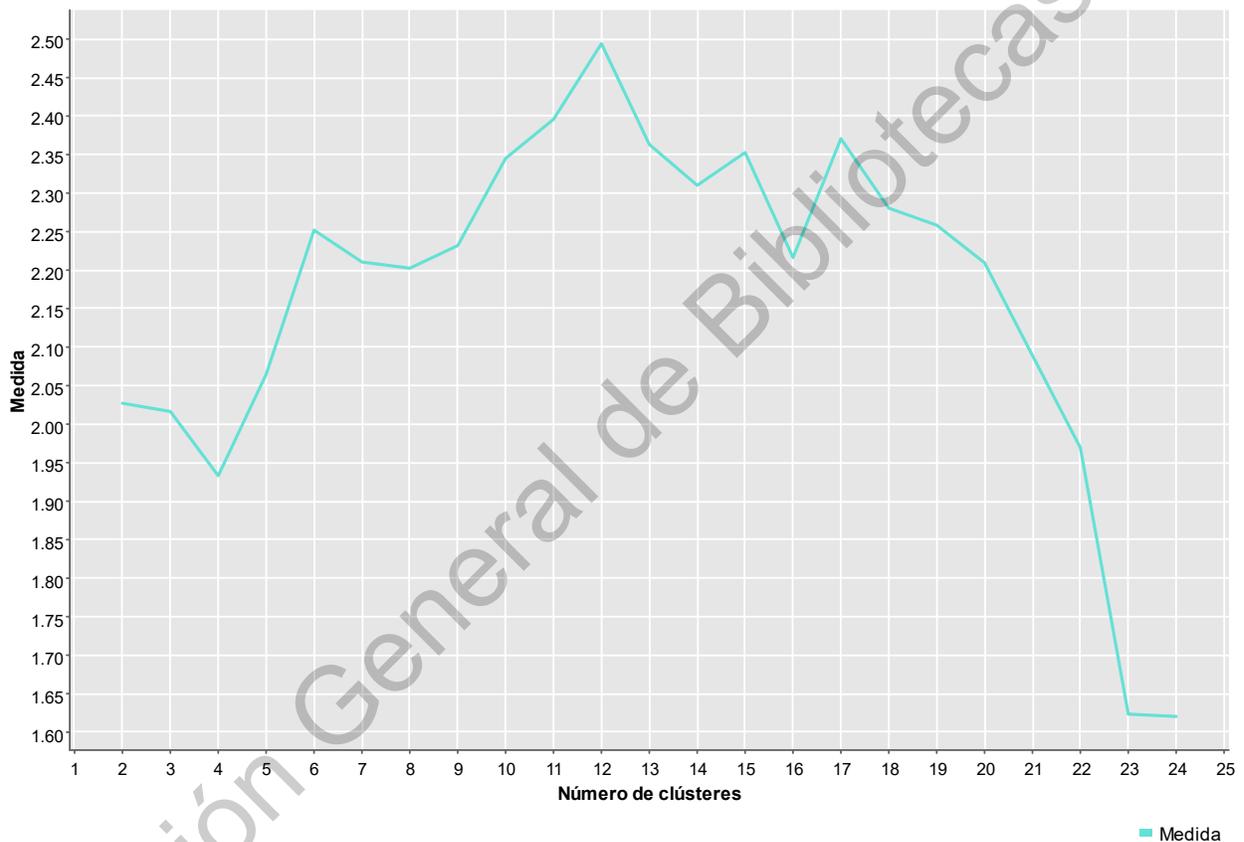


Figura 7. Medida m_2 para K-medias con media, desviación estándar y autocorrelación con lag 22. Fuente: elaboración propia.

Ancho de silueta

K-medias con ACF

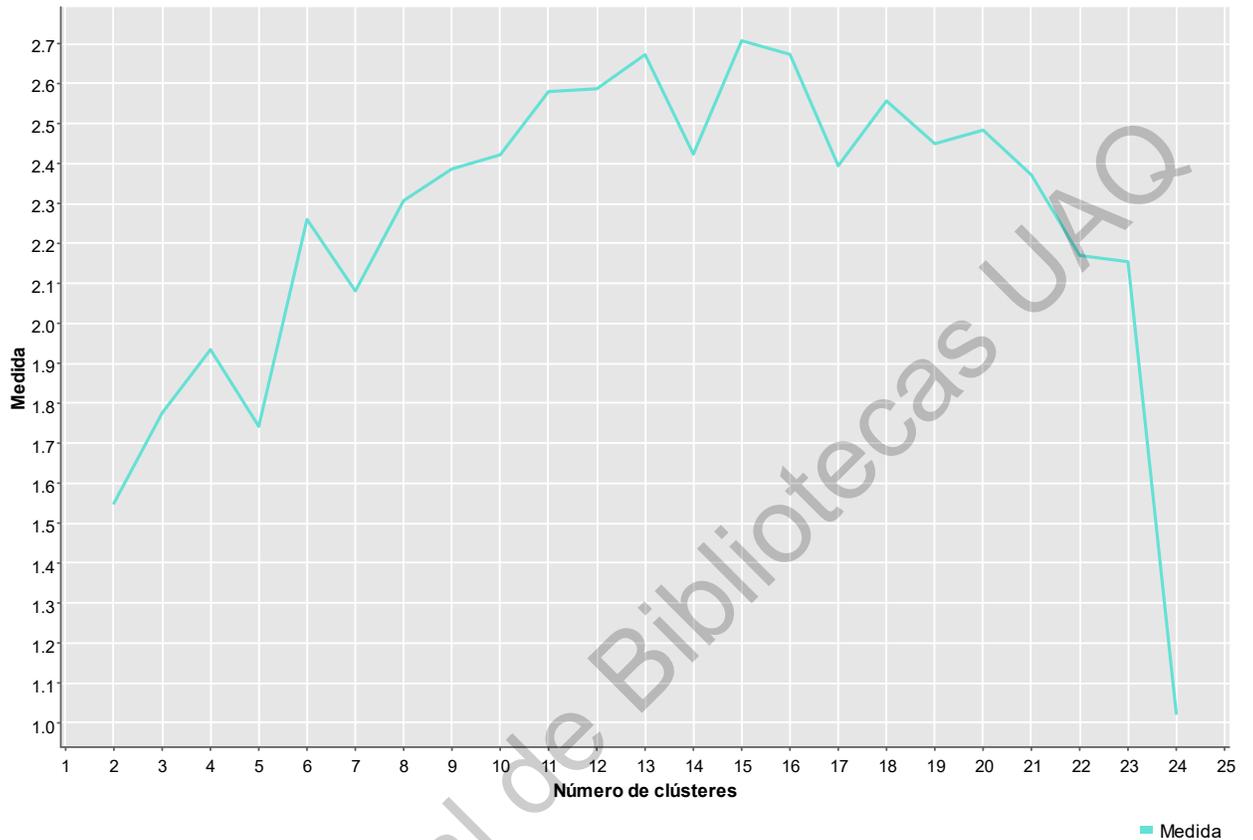


Figura 8. Medida m_2 para K-medias con media, desviación estándar y autocorrelación con lag 6. Fuente: elaboración propia.

La medida m_2 depende de una variable adicional a las de la medida m_1 que es el número de lags p . Como se puede ver en las figuras del Anexo B, al clusterizar con el método de k-medias con la información de media, desviación estándar y función autocorrelación, se obtuvieron valores más bajos de la medida m_2 con lags altos. Esto quiere decir que al utilizar muchos lags en el modelo, se obtienen valores bajos del promedio de ancho de silueta y se está saturando el modelo.

Las Figuras 9 y 10 presentan en el eje vertical la medida m_2 (I.c.v) de cada uno de los números de clústeres posibles (I.c.ii), mostrados en el eje horizontal, para el método de clusterización jerárquica (I.c.iv). Se muestran únicamente los lags 22 y 6 (I.c.i), dado que son las configuraciones con mayor información disponible y la que maximiza la medida m_2 , respectivamente. Sin embargo, se hicieron las pruebas con cada uno de los lags disponibles (de 1 a 22), ver anexo C.

Ancho de silueta

Jerárquica con ACF

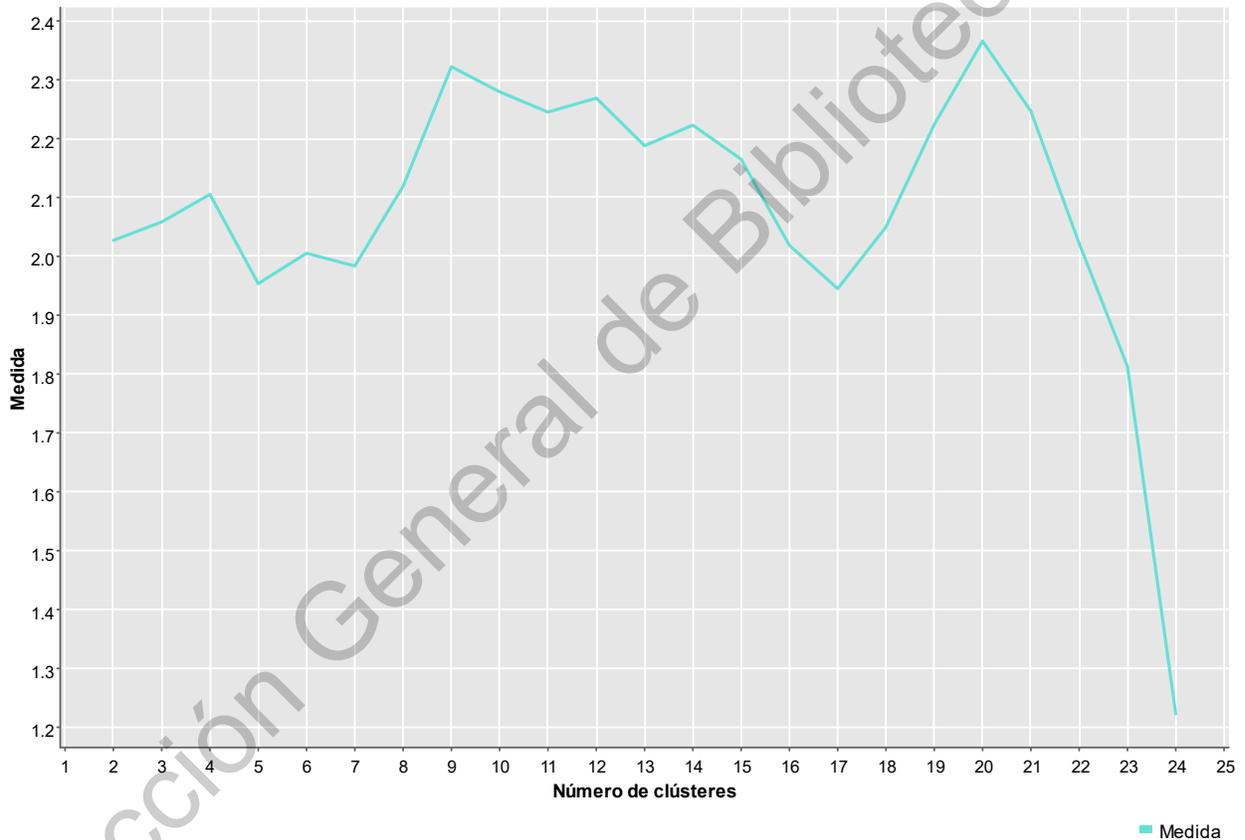


Figura 9. Medida m_2 para jerárquica con media, desviación estándar y autocorrelación con lag 22. Fuente: elaboración propia.

Ancho de silueta

Jerárquica con ACF

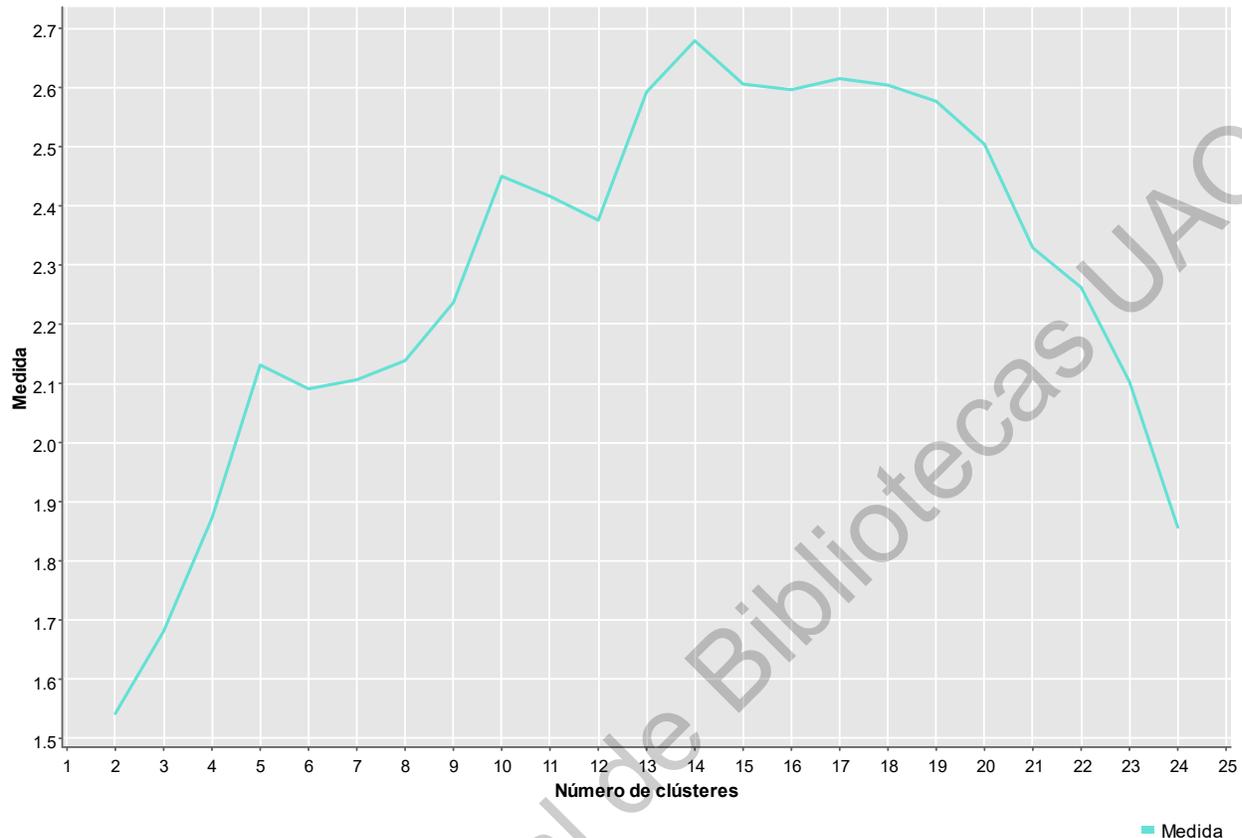


Figura 10. Medida m_2 para jerárquica con media, desviación estándar y autocorrelación con lag 6. Fuente: elaboración propia.

Al igual que en la clusterización con el método de k-medias, se puede apreciar que con la clusterización jerárquica con media, desviación estándar y función de autocorrelación, se obtuvieron valores más altos de la medida m_2 al utilizar un número de lags bajo.

Una vez calculadas las medidas m_1 y m_2 para todas las configuraciones posibles del modelo (*l.b* y *l.c*), se obtuvo que los resultados para el método de K-medias fueron altamente variables derivado de que es un proceso de optimización que depende de una semilla aleatoria inicial lo cual se puede notar de mejor manera en las figuras 5 y 6, al ser la figura 6 (metodología jerárquica) más estable y suave que la figura 5 (metodología

k-medias). Este fenómeno es un poco más difícil de apreciar en las figuras 7 a 10, pero si se repite el proceso varias veces para k-medias comienzan a observarse resultados distintos para las mismas configuraciones, exponiendo la mayor estabilidad de la metodología jerárquica. Con base en estos resultados se decidió clusterizar utilizando únicamente el método jerárquico. Además, utilizando lags altos en la función de autocorrelación se encontraron valores pequeños de la medida m_2 , como se muestra en el anexo C, por lo que se utilizó una dimensión del modelo considerando únicamente 6 lags.

En resumen, se decide clusterizar utilizando la información de media y desviación estándar para la metodología jerárquica con $K = 13$ y con la información de media, desviación estándar y función de autocorrelación para la metodología jerárquica con $K = 14$, ; debido a que ambos valores de K maximizan el valor de m_1 y m_2 respectivamente (*I.b.v* y *I.c.vi*). Con esta información se presentan los clústeres obtenidos de las 25 acciones analizadas a través de las metodologías y parámetros seleccionados. Los diferentes colores representan los diferentes clústeres obtenidos.

En la Figura 11 se pueden observar los 13 clústeres obtenidos con el método jerárquico tomando en cuenta la media y desviación estándar de los rendimientos logarítmicos de las 25 acciones seleccionadas.

Jerárquica

Media y desviación

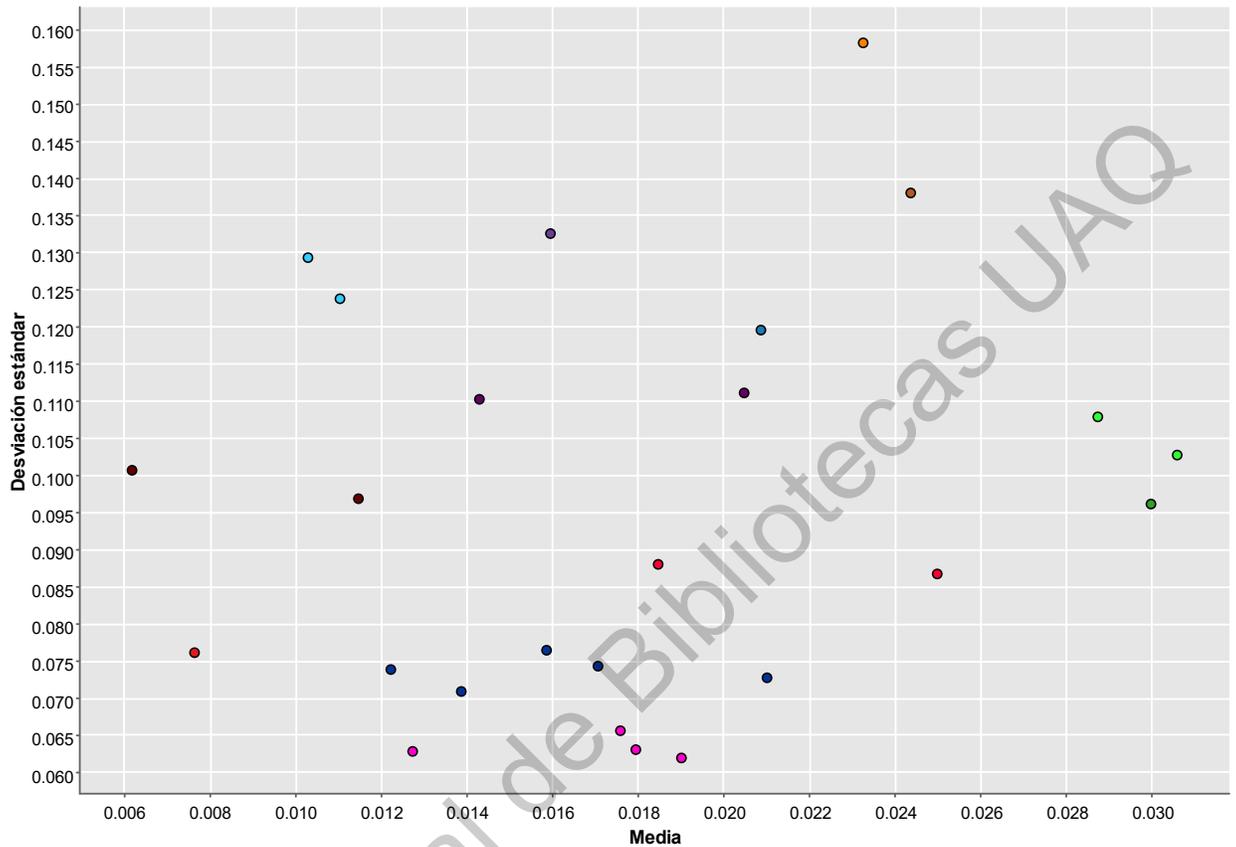


Figura 11. Clusterización media y desviación estándar. Fuente: elaboración propia.

La figura 11 es muy predecible ya que la información graficada es exactamente la misma utilizada para clusterizar, entonces los clústeres describen áreas bien definidas en el plano sin que haya superposición de clústeres. Además, existe la posibilidad de que las diferentes escalas entre las variables tengan un impacto en la clusterización, siendo las distancias entre desviaciones estándar relativamente más grandes que las distancias entre medias. Esto se puede observar en el efecto de “laminado” que presentan los clústeres, donde se prefiere el eje horizontal (media) al eje vertical (desviación).

La Figura 12 muestra los 14 clústeres obtenidos con el método jerárquico tomando en cuenta la media, desviación estándar y función de autocorrelación de los rendimientos

logarítmicos de las 25 acciones seleccionadas. Así mismo, la Figura 13 muestra una representación gráfica de los mismos 14 clústeres, obtenidos con la metodología jerárquica tomando en cuenta la media, desviación estándar y función de autocorrelación, a través del análisis de componentes principales (//.a). Con este análisis, se puede apreciar de una manera más clara la estructura de los clústeres.

Jerárquico

Media, desviación y autocorrelación

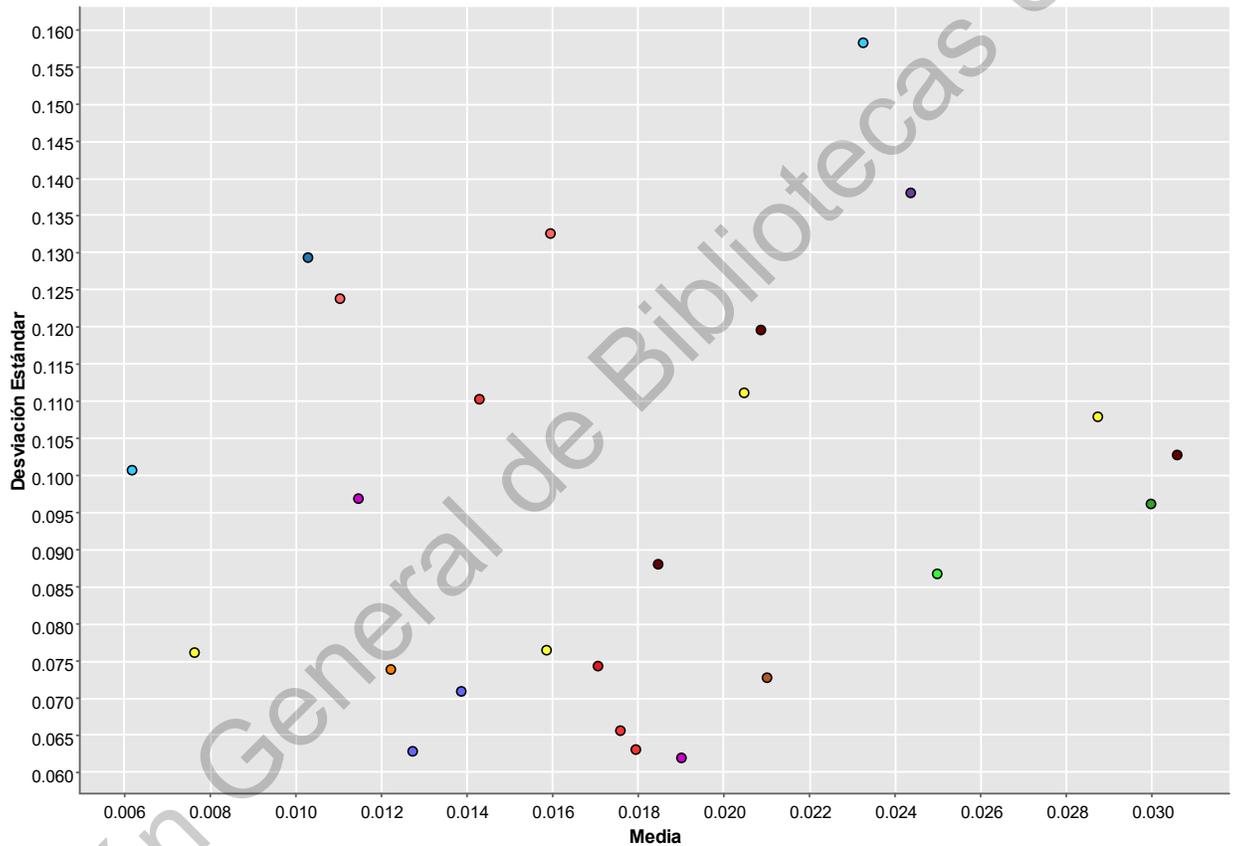


Figura 12. Clusterización media, desviación estándar y autocorrelación. Fuente: elaboración propia.

Jerárquico

Media, desviación y autocorrelación

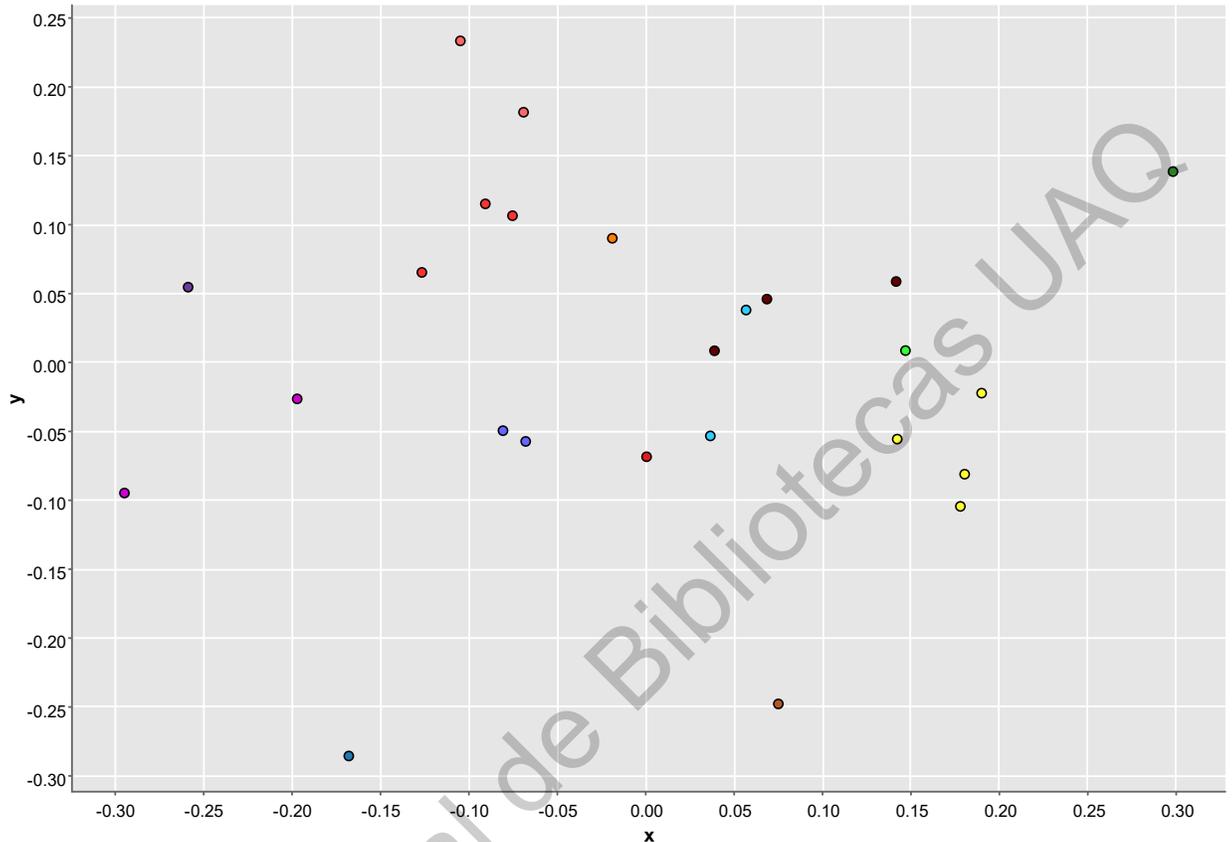


Figura 13. Clusterización media, desviación estándar y autocorrelación redimensionado con análisis de componentes principales. Fuente: elaboración propia.

Se puede notar desde la Figura 12 que, al representar gráficamente los clústeres obtenidos con la información de media, desviación estándar y autocorrelación utilizando únicamente la media y desviación estándar como ejes, no se explica la totalidad de la información contenida en las series de tiempo de las acciones. Es por esto que la figura muestra una asignación aparentemente aleatoria de clústeres.

Es necesaria una herramienta como el Análisis de Componentes Principales (Figura 13), que reduce las dimensiones de los datos mientras mantiene la mayor cantidad de información intacta, para poder representar los clústeres de manera que sean visualmente interpretables.

De esta forma, como se menciona en la etapa final de la metodología, se validan de manera cualitativa los resultados de las clusterizaciones con y sin las variables de autocorrelación (//.a). En la tabla 4 se muestra a qué clúster pertenece cada una de las 25 acciones analizadas con los dos métodos seleccionados: Jerárquico con media, desviación estándar y función de autocorrelación (Clúster con ACF) y jerárquico con media y desviación estándar (Clúster sin ACF).

Acción	Clúster con ACF	Clúster sin ACF
ALFAA	Clúster 12	Clúster 10
ALSEA	Clúster 1	Clúster 1
ARA	Clúster 11	Clúster 11
BACHOCOB	Clúster 2	Clúster 13
BIMBOA	Clúster 14	Clúster 9
CEMEXCPO	Clúster 8	Clúster 8
CIDMEGA	Clúster 7	Clúster 12
CMOCTEZ	Clúster 3	Clúster 13
CULTIBAB	Clúster 12	Clúster 2
CYDSASAA	Clúster 4	Clúster 3
ELEKTRA	Clúster 8	Clúster 4
GCC	Clúster 12	Clúster 13
GFINBURO	Clúster 13	Clúster 13
GFNORTEO	Clúster 9	Clúster 12
GMD	Clúster 11	Clúster 6
GMEXICOB	Clúster 12	Clúster 7
HERDEZ	Clúster 14	Clúster 9
ICHB	Clúster 14	Clúster 10
KIMBERA	Clúster 10	Clúster 9
KUOB	Clúster 5	Clúster 8
MEXCHEM	Clúster 9	Clúster 7
PE&OLES	Clúster 9	Clúster 5
SORIANAB	Clúster 6	Clúster 13
TLEVISACPO	Clúster 13	Clúster 9
WALMEX	Clúster 10	Clúster 11

Tabla 4. Resultados de la clusterización. Fuente: elaboración propia.

La tabla 5 presenta de manera resumida los primeros 3 niveles de la clasificación sectorial de las acciones analizadas según el esquema utilizado por la BMV; este esquema permite establecer segmentos y grupos homogéneos de acciones para facilitar su estudio y análisis comparativo.

Acción	Sector	Subsector	Ramo
ALFAA	Industrial	Bienes de equipo	Controladoras
ALSEA	Servicios y bienes de consumo no básico	Servicios al consumidor	Hoteles, restaurantes y esparcimiento
ARA	Industrial	Construcción	Construcción de viviendas
BACHOCOB	Productos de consumo frecuente	Alimentos, bebidas y tabaco	Alimentos
BIMBOA	Productos de consumo frecuente	Alimentos, bebidas y tabaco	Alimentos
CEMEXCPO	Materiales	Materiales	Materiales de construcción
CIDMEGA	Servicios y bienes de consumo no básico	Servicios al consumidor	Hoteles, restaurantes y esparcimiento
CMOCTEZ	Materiales	Materiales	Materiales de construcción
CULTIBAB	Productos de consumo frecuente	Alimentos, bebidas y tabaco	Alimentos
CYDSASAA	Materiales	Materiales	Productos químicos
ELEKTRA	Servicios y bienes de consumo no básico	Venta al por menor	Ventas especializadas
GCC	Materiales	Materiales	Materiales de construcción
GFINBURO	Servicios financieros	Entidades financieras	Grupos financieros
GFNORTEO	Servicios financieros	Entidades financieras	Grupos financieros
GMD	Industrial	Construcción	Construcción e ingeniería
GMEXICOB	Materiales	Materiales	Metales y minería
HERDEZ	Productos de consumo frecuente	Alimentos, bebidas y tabaco	Alimentos

ICHB	Materiales	Materiales	Fabricación y comercialización de materiales
KIMBERA	Productos de consumo frecuente	Productos domésticos y personales	Productos domésticos
KUOB	Industrial	Bienes de equipo	Controladoras
MEXCHEM	Materiales	Materiales	Productos químicos
PE&OLES	Materiales	Materiales	Metales y minería
SORIANAB	Productos de consumo frecuente	Venta de productos de consumo frecuente	Venta de productos de consumo frecuente
TLEVISACPO	Servicios de telecomunicaciones	Medios de comunicación	Medios de comunicación
WALMEX	Productos de consumo frecuente	Venta de productos de consumo frecuente	Venta de productos de consumo frecuente

Tabla 5. Descripción de las empresas por sector, subsector y ramo. Fuente: elaboración propia.

En la tabla 6 se muestra el desglose de los 14 clústeres obtenidos con el método Jerárquico con media, desviación estándar y función de autocorrelación (Clúster ACF) según el número de empresas que contienen de cada sector.

Clúster con ACF	Sector	No. de empresas
Clúster 1	Servicios y bienes de consumo no básico	1
Clúster 2	Productos de consumo frecuente	1
Clúster 3	Materiales	1
Clúster 4	Materiales	1
Clúster 5	Industrial	1
Clúster 6	Productos de consumo frecuente	1
Clúster 7	Servicios y bienes de consumo no básico	1
Clúster 8	Materiales	1
	Servicios y bienes de consumo no básico	1
Clúster 9	Materiales	2
	Servicios financieros	1
Clúster 10	Productos de consumo frecuente	2
Clúster 11	Industrial	2
Clúster 12	Industrial	1
	Materiales	2

	Productos de consumo frecuente	1
Clúster 13	Servicios de telecomunicaciones	1
	Servicios financieros	1
Clúster 14	Materiales	1
	Productos de consumo frecuente	2
Total general		25

Tabla 6. Clústeres con ACF por sector. Fuente: elaboración propia.

En la tabla 7 se muestra un desglose de los 13 clústeres obtenidos con el método Jerárquico con media y desviación estándar (Clúster sin ACF) según el número de empresas que contienen de cada sector.

Clúster sin ACF	Sector	No. de empresas
Clúster 1	Servicios y bienes de consumo no básico	1
Clúster 2	Productos de consumo frecuente	1
Clúster 3	Materiales	1
Clúster 4	Servicios y bienes de consumo no básico	1
Clúster 5	Materiales	1
Clúster 6	Industrial	1
Clúster 7	Materiales	2
Clúster 8	Industrial	1
	Materiales	1
	Productos de consumo frecuente	3
Clúster 9	Servicios de telecomunicaciones	1
	Industrial	1
Clúster 10	Materiales	1
	Industrial	1
Clúster 11	Productos de consumo frecuente	1
	Servicios financieros	1
Clúster 12	Servicios y bienes de consumo no básico	1
	Materiales	2
	Productos de consumo frecuente	2
Clúster 13	Servicios financieros	1
Total general		25

Tabla 7. Clústeres sin ACF por sector. Fuente: elaboración propia.

Como primer resultado se puede observar que los clústeres obtenidos con los dos métodos, Clúster con ACF y Clúster sin ACF, contienen empresas de diferentes sectores, no se obtuvo ningún clúster que contenga a todas las empresas de un sector particular.

De esta manera se puede señalar que dentro de cada sector existen empresas que presentan un comportamiento o caracterización diferente al de otras empresas pertenecientes al mismo sector en términos tanto de su media y desviación estándar, así como de su función de autocorrelación. Los clústeres obtenidos indican que existen condiciones de mercado que afectan a ciertos grupos de acciones independientemente del sector, subsector o ramo al que pertenezcan.

Las acciones correspondientes a los sectores de “Materiales” y “Productos de Consumo Frecuente” presentaron la mayor relación entre acciones del mismo sector en lo que a su media, desviación estándar y función de autocorrelación respecta; tanto en el Clúster con ACF como en el Clúster sin ACF se obtuvieron clústeres conformados por 2 o más acciones de estos sectores.

Posteriormente se sigue el algoritmo de optimización de portafolios descrito en el inciso (II.b) y (II.c) de la metodología tanto para la clusterización con ACF como para la clusterización sin ACF con distintos valores objetivos de media, construyendo así una frontera eficiente para cada metodología:

Portafolios vs acciones

Frontera Eficiente

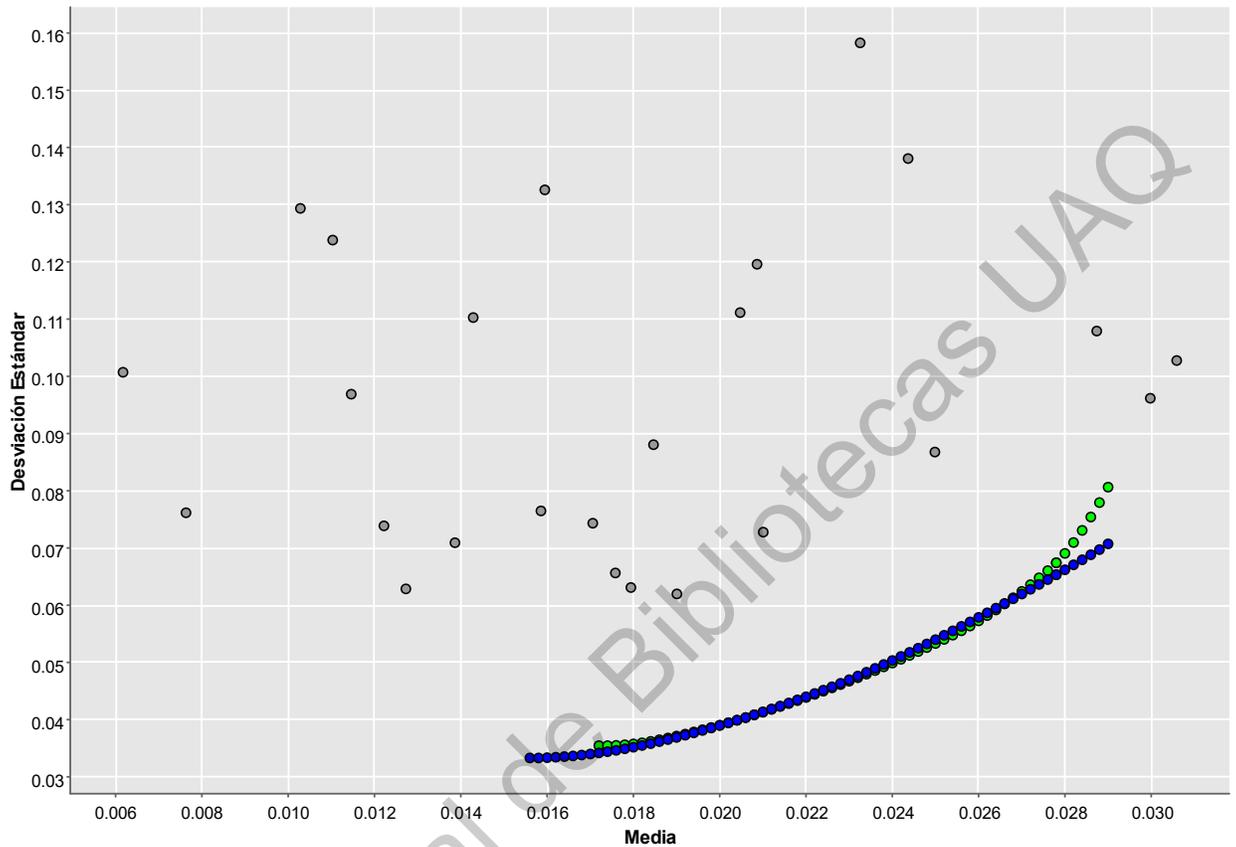


Figura 14. Frontera eficiente de los portafolios. Fuente: elaboración propia.

En la Figura 16 se pueden apreciar gráficamente los resultados obtenidos con el algoritmo, con la frontera eficiente para los portafolios tomando en cuenta la función de autocorrelación (verde) y sin tomar en cuenta la función de autocorrelación (azul). También se muestran las 25 acciones que conforman los portafolios (gris) para contraste.

Un resultado que se observa en las comparaciones de portafolios es que la frontera eficiente es similar para ambos casos, separándose al final ya que la clusterización con autocorrelación comienza a ganar más volatilidad que su contraparte. Esto es de esperarse ya que la optimización de portafolios está basada en la media y desviación estándar, así que tendría que hacerse un análisis más profundo para garantizar que en verdad el portafolio sin autocorrelación tuvo un mejor desempeño. Sin

embargo, el portafolio de la autocorrelación tiende de manera más directa a los valores más extremos de las acciones individuales, lo que podría indicar que el método E-V está subvaluando la varianza.

Dirección General de Bibliotecas UAQ

5. CONCLUSIONES

La presente tesis tuvo como objetivo comparar la caracterización en forma de clústeres de 25 acciones de la Bolsa Mexicana de Valores utilizando el método de media-varianza y la autocorrelación en sus series de tiempo, de tal manera que se pudiera observar si se producen resultados fundamentalmente distintos entre ambas metodologías.

Para efectos de esta comparación, primero se analizó el número óptimo de clústeres para cada metodología, así como el número óptimo de lags para la metodología con función de autocorrelación tomando en cuenta el ancho de silueta a través de las ecuaciones (28) y (29). Dichas ecuaciones redistribuyen el peso de cada una de las variables en la priorización de los resultados, estabilizando la función y facilitando la distinción de *outliers* que podrían indicar algún tipo de relación entre las acciones de esa clusterización en particular.

Posteriormente, al analizar las clusterizaciones obtenidas con los métodos jerárquico y k-medias, se concluyó que los resultados obtenidos con el método de k-medias fueron altamente variables debido a la asignación aleatoria inicial de las acciones a cada clúster con lo que se decidió clusterizar únicamente utilizando el método jerárquico tanto para el análisis de media-varianza como para el análisis de media-varianza y función de autocorrelación. Con base en lo anterior se pudo determinar que el número óptimo de clústeres fue de entre 13 y 14 mientras que el número óptimo de lags fue de 6.

A continuación, al estudiar cualitativamente las gráficas obtenidas por la clusterización jerárquica (Figuras 11 y 12), se observó que los clústeres formados con el método de media-varianza fueron significativamente diferentes a aquellos obtenidos con el método de media-varianza y función de autocorrelación; indicando de esta manera que sí existe una clara diferencia a la hora de incluir la función de autocorrelación de las acciones en el modelo.

Al analizar a detalle la distribución de las acciones en cada clúster de manera individual (Tabla 4) se corroboró que en efecto las acciones contenidas en los clústeres de media-varianza variaron significativamente a las contenidas en los clústeres de media-varianza y función de autocorrelación. La única similitud destacable que se presentó en la clusterización de ambas metodologías fue la agrupación de las acciones correspondientes a los sectores de “Materiales” y “Productos de Consumo Frecuente”, las cuales, presentaron la mayor relación entre acciones del mismo sector.

Al realizar el ejercicio de optimización de portafolios, se encontró que los resultados basados en las dos clusterizaciones fueron similares, divergiendo ligeramente en los casos con la media más elevada. Podría pensarse que esto contradice la hipótesis de este trabajo, pero si se evalúa que en esta prueba la metodología era la misma y la clusterización solamente tenía un efecto sobre el orden en el que se efectuaba esta, y no sobre la información empleada, es razonable concluir que la mínima diferencia en resultados es consecuencia de una diferencia importante en estructura.

Es importante tomar en cuenta que los métodos aquí descritos muestran de manera cualitativa la diferencia de enfoques en la selección de portafolios, y que la comparación de resultados obtenidos está realizada únicamente con la metodología de media y varianza, por lo que los resultados pueden estar sesgados a favor de esta. Sería óptimo el desarrollar una nueva metodología de valuación neutral que midiera la efectividad de ambos acercamientos, pero lamentablemente eso queda fuera del alcance de esta investigación.

Finalmente, y como propuesta para futuros trabajos, se puede explotar la flexibilidad que ofrecen las series de tiempo para hacer portafolios temporalmente dinámicos. Esto es, considerar que las acciones no necesariamente deben de adquirirse al mismo tiempo y que, por lo tanto, existe la posibilidad de encontrar correlaciones más bajas que las del método clásico. Una idea interesante sería mantener la caracterización por lags de las acciones, aprovechando la generalización que surge de esta, para así entrenar alguna herramienta de aprendizaje supervisado que permita el análisis de la

interacción de diversas acciones con distintos horizontes temporales, sin entrar en la complicación de elegir modelos de series de tiempo que se ajusten a las acciones individualmente.

Dirección General de Bibliotecas UAQ

6. REFERENCIAS

- Batsleer, T. (2016). *Alternatives to the Markowitz Rule*. Gante, Bélgica: Universiteit Gent.
- Bawa, V. S., Brown, S. J., & Klein, R. W. (1979). *Estimation risk and optimal portfolio choice*. Amsterdam: North-Holland Pub. Co.
- Bonanno, G., Lillo, F., & Mantegna, R. N. (2001). Levels of complexity in financial markets. *Physica*.
- Bowerman, B., & O'Connell, R. (1987). *Time Series Forecasting: Unified Concepts and Computer Implementation*. Boston: Duxbury Press.
- Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). cValid, an R package for cluster validation. *Journal of Statistical Software*.
- Campbell, J. Y., Lo, A. W., & MacKinlay, C. (1996). *The Econometrics of Financial Markets*. Nueva Jersey: Princeton University Press.
- Cerqueira, A. (2006). Autocorrelation in Daily Stock Prices. *Working Paper, 4th PFN Conference*.
- Green, S. (2011). *Time Series Analysis of Stock Prices Using the Box-Jenkins Approach*. Georgia: Georgia Southern University.
- Grupo BMV. (2018, 06 16). *Grupo BMV*. Retrieved from <http://www.bmv.com.mx/es/grupo-bmv/acerca-de>
- Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, 107-145.
- Handl, J., Knowles, J., & Kell, D. B. (2005). *Computational cluster validation in post-genomic data analysis*. Manchester: University of Manchester.
- Smith, L. (2002, Febrero 26). A tutorial on Principal Components Analysis.

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An Introduction to Statistical Learning with Applications in R*. New York: Springer.
- Kempf, A., & Memmel, C. (2003). On the Estimation of the Global Minimum Variance Portfolio. *SSRN Electronic Journal*.
- Ladokhin, S. (2009). Forecasting Volatility in the Stock Market. *Business Mathematics and Informatics*.
- Markowitz, H. (1952, Marzo). Portfolio Selection. *The Journal of Finance*, pp. 77-91.
- Navarro, F. J. (2013). *Out-of-Sample Performance of Mean-Variance Strategies: Is Active Portfolio Management Worth the Effort in Europe?* Valencia: Universidad de Valencia.
- Pérez Rojas, E. A., & Ramírez Galindo, R. (2009). *Financiamiento bursátil para la mediana empresa: Dos casos de éxito*. Cholula, Puebla.
- Ren, Z. (2005). *Portfolio Construction Using Clustering methods*. Worcester Polytechnic institute .
- Ren, Z. (2005). *Portfolio Construction Using Clustering Methods*. Worcester: Worcester Polytechnic Institute.
- Rincón, L. (2014). *Introducción a la probabilidad*. México DF: Facultad de Ciencias UNAM.
- Rosén, F. (2006). *Correlation based clustering of the Stockholm Stock Exchange*. Stockholm: Stockholm University.
- Rosén, F. (2006). *Correlation based clustering of the Stockholm Stock Exchange*. Stockholm University.
- Rueda, A. (2005). *Para entender la bolsa: financiamiento e inversión en el mercado de valores*. México: Thomson Learning.

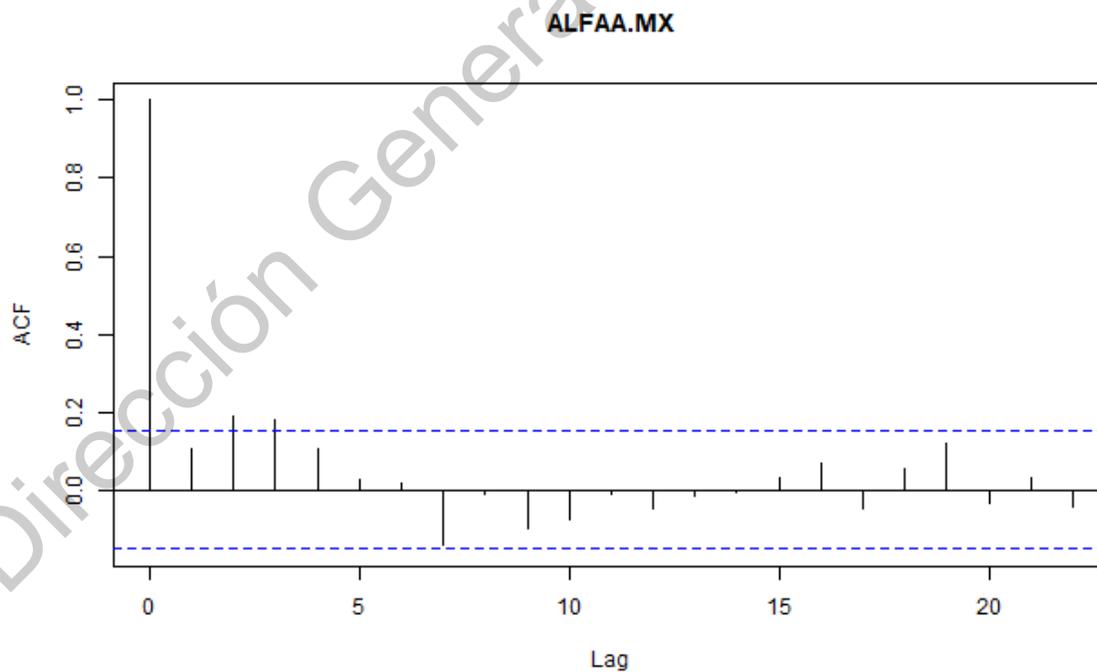
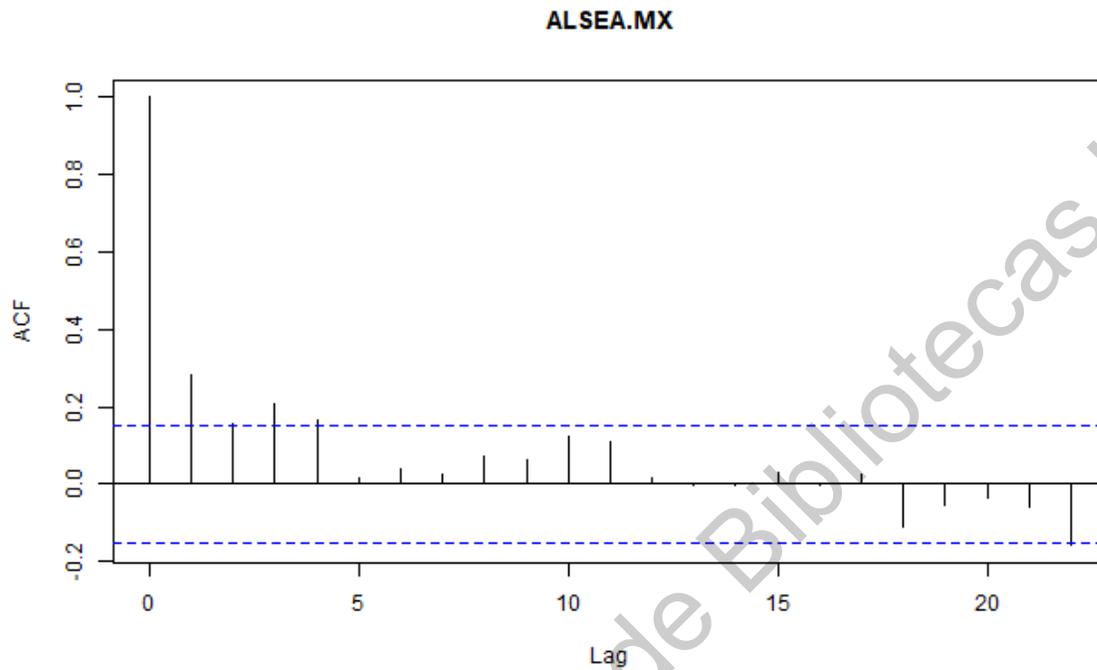
Sherer, B. (2002). *Portfolio Construction and Risk Budgeting*. Londres: Risk Books.

Shumway, R. H., & Stoffer, D. S. (2015). *Time Series Analysis and Its Applications*. Davis, California: Springer.

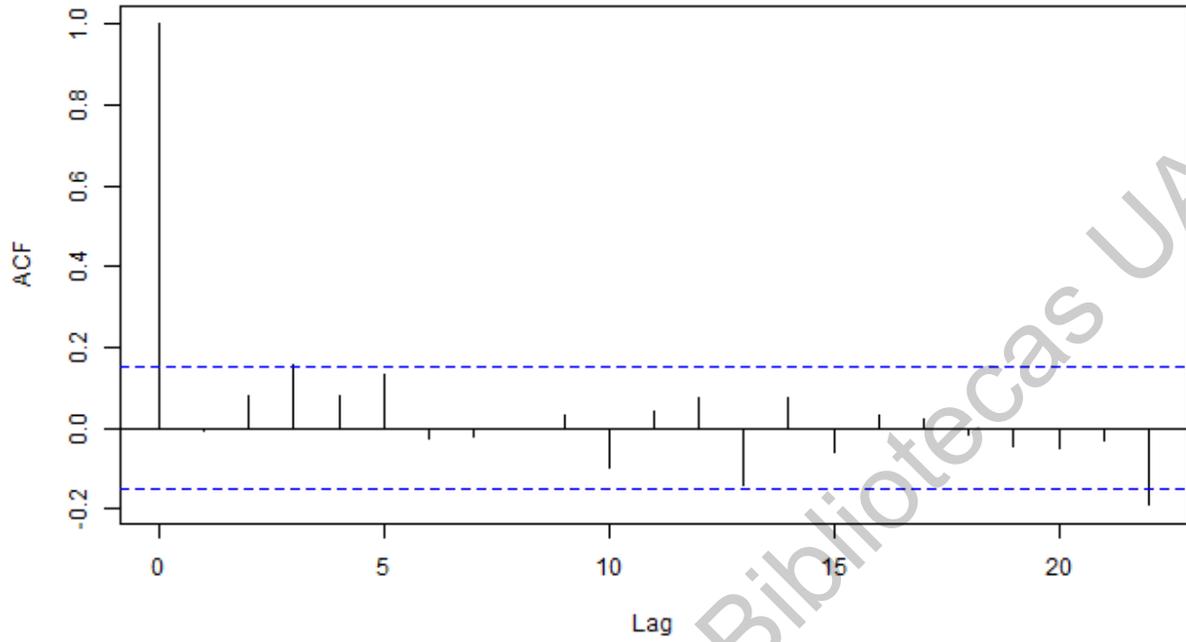
Dirección General de Bibliotecas UAQ

7. ANEXOS

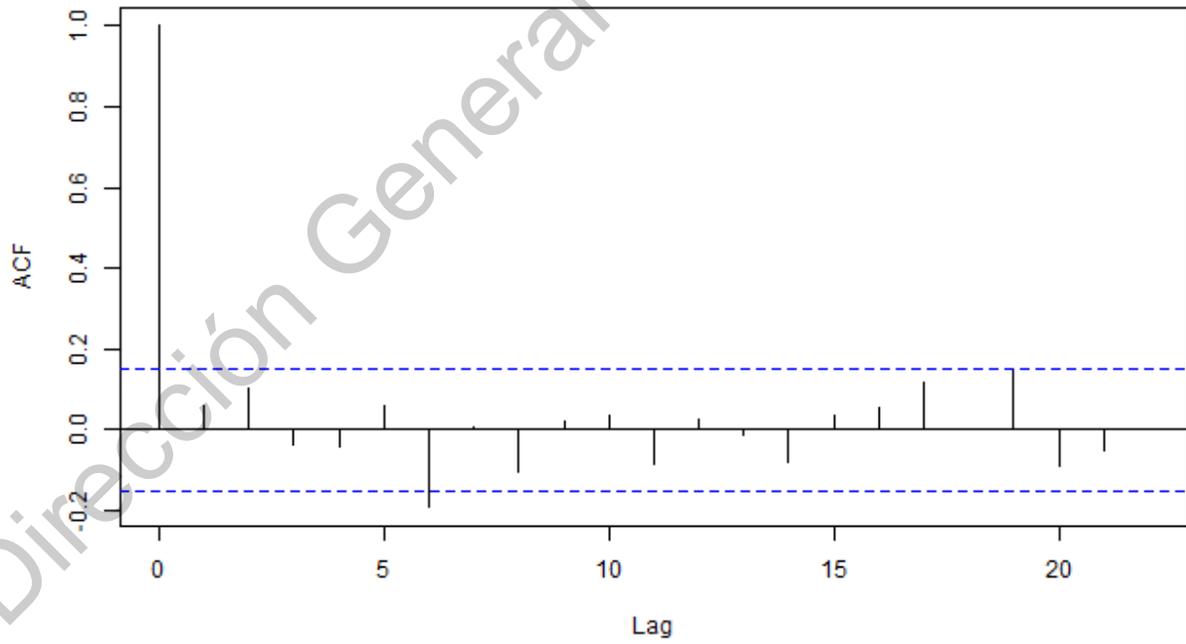
7.1 ANEXO A



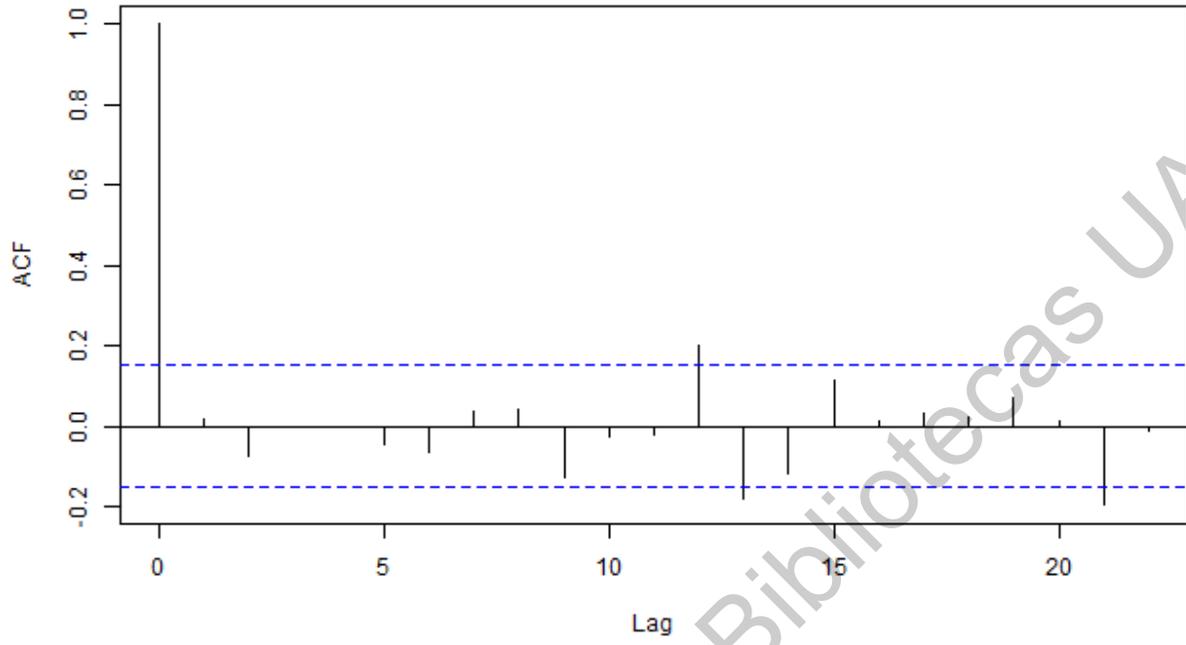
ARA.MX



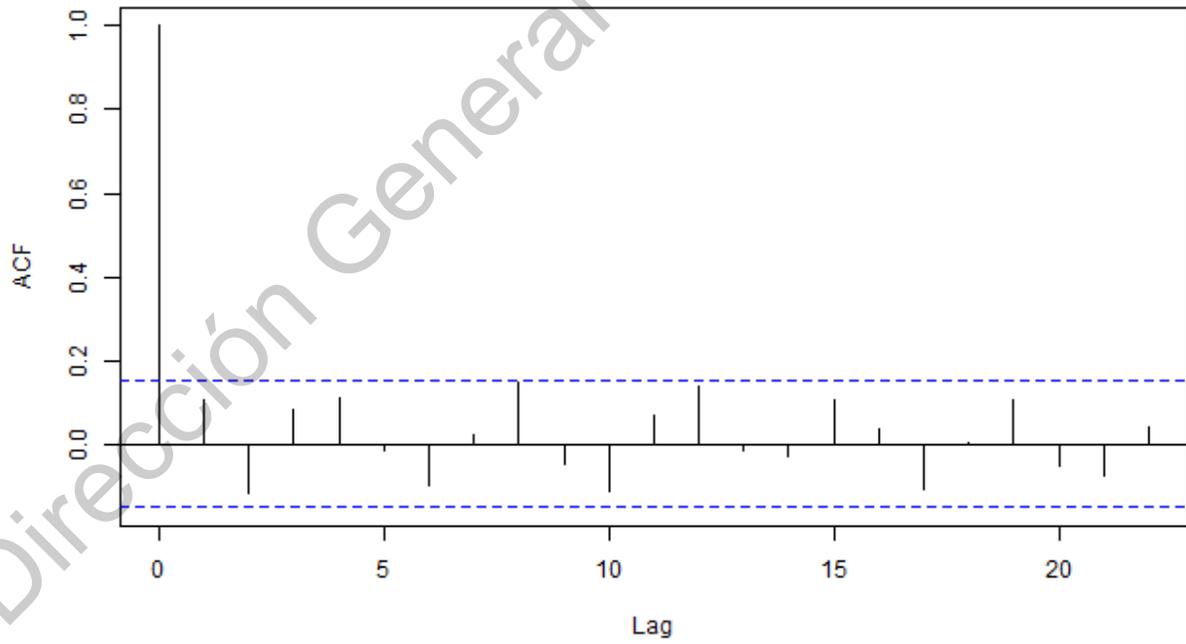
BACHOCOB.MX



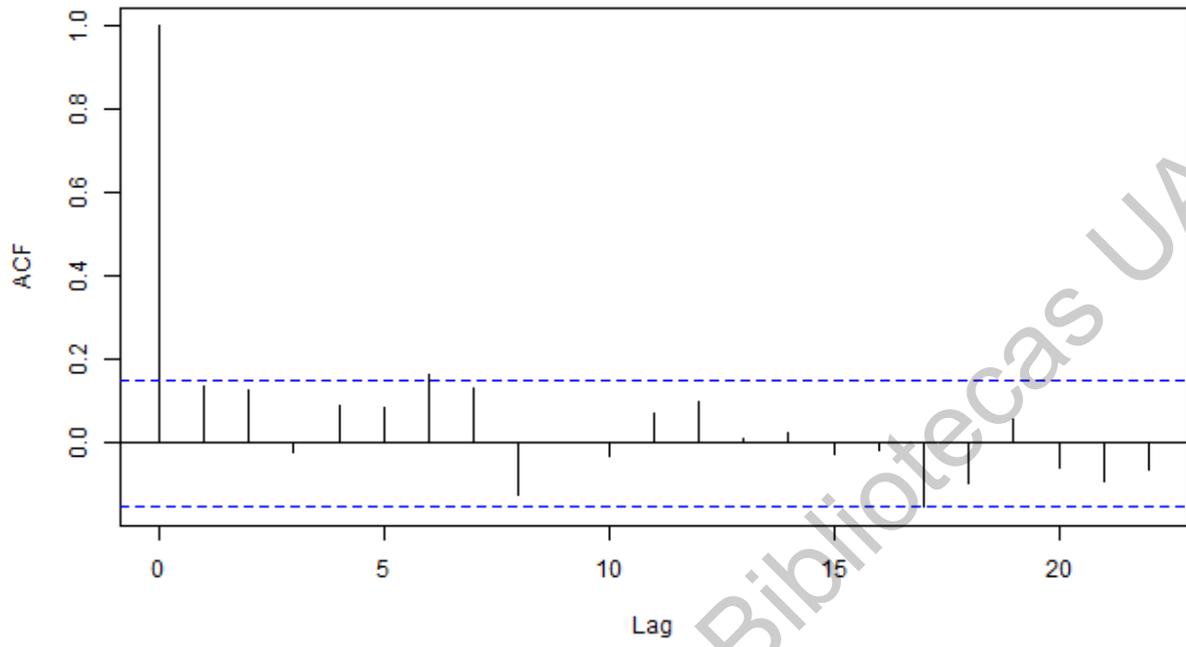
BIMBOA.MX



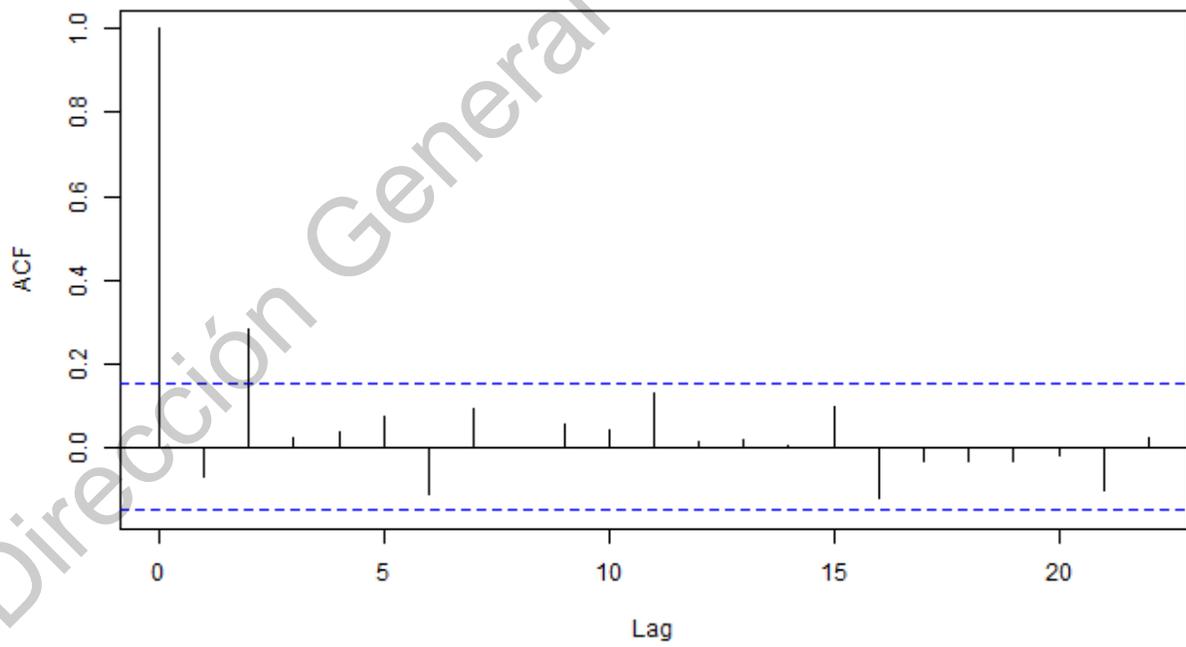
CEMEXCPO.MX



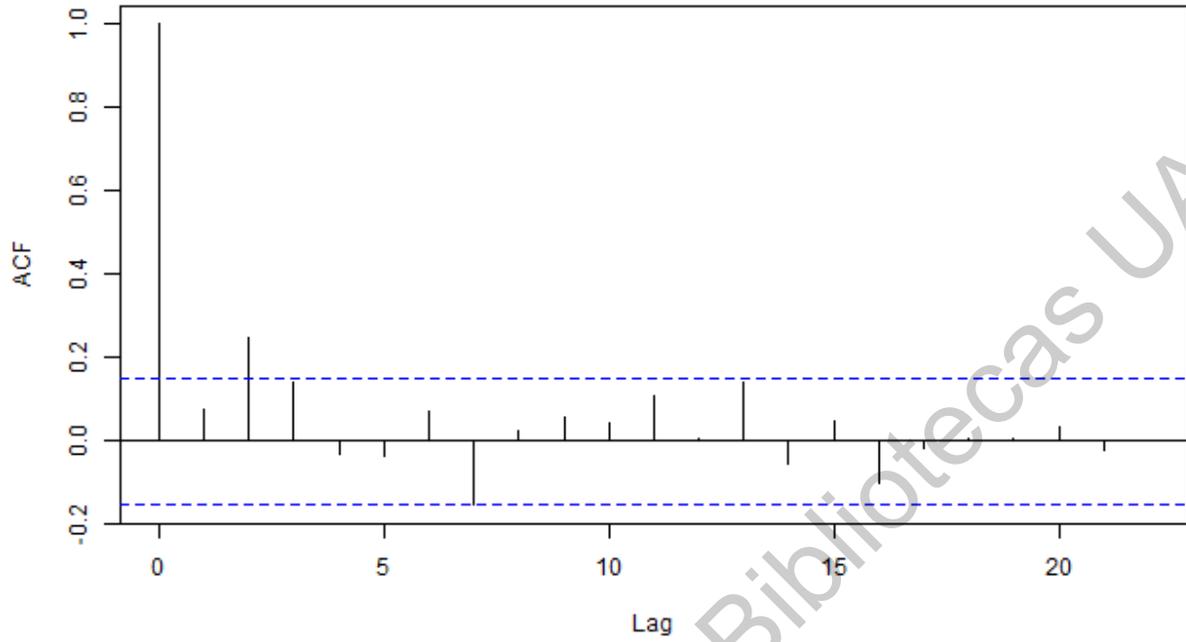
CIDMEGA.MX



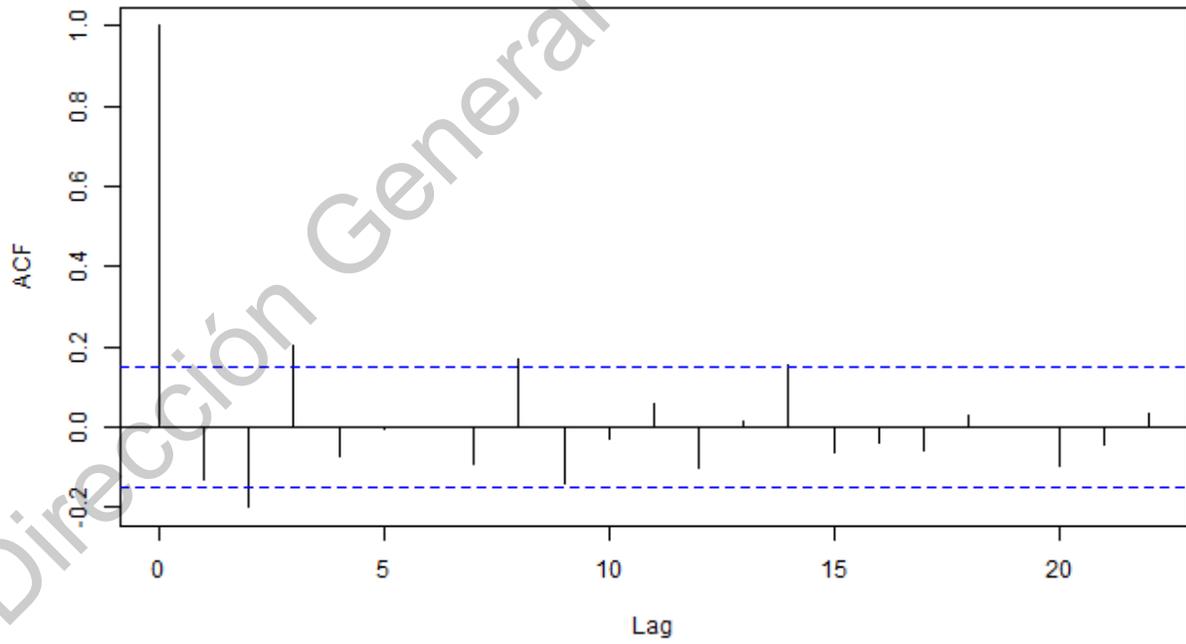
CMOCTEZ.MX



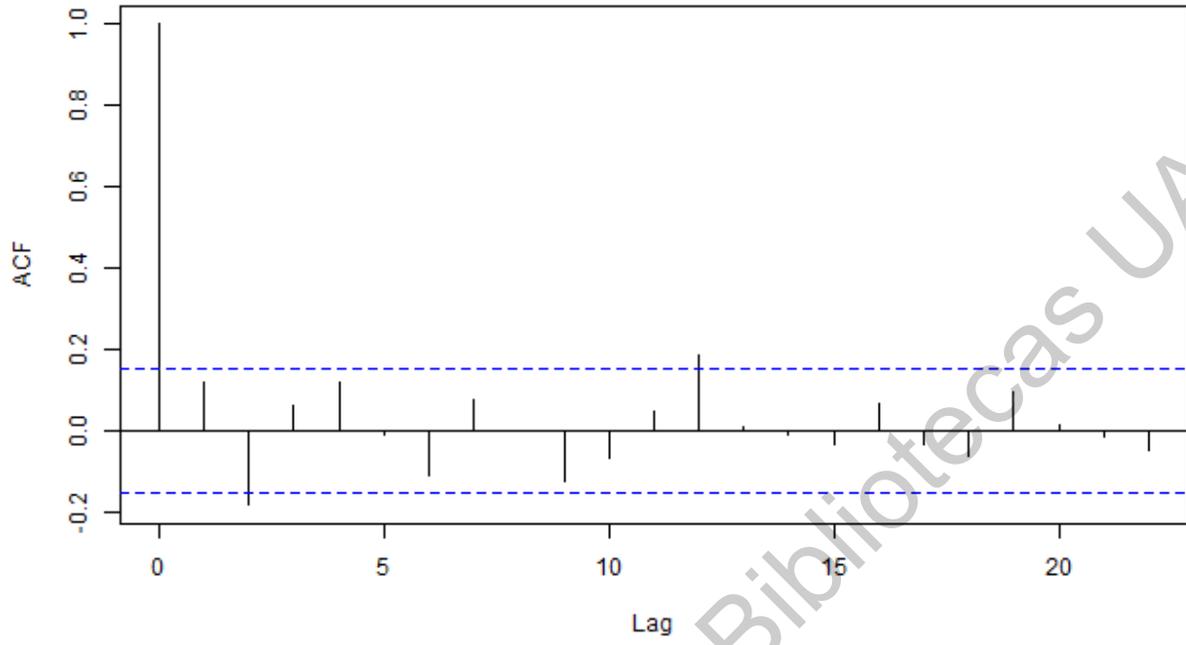
CULTIBAB.MX



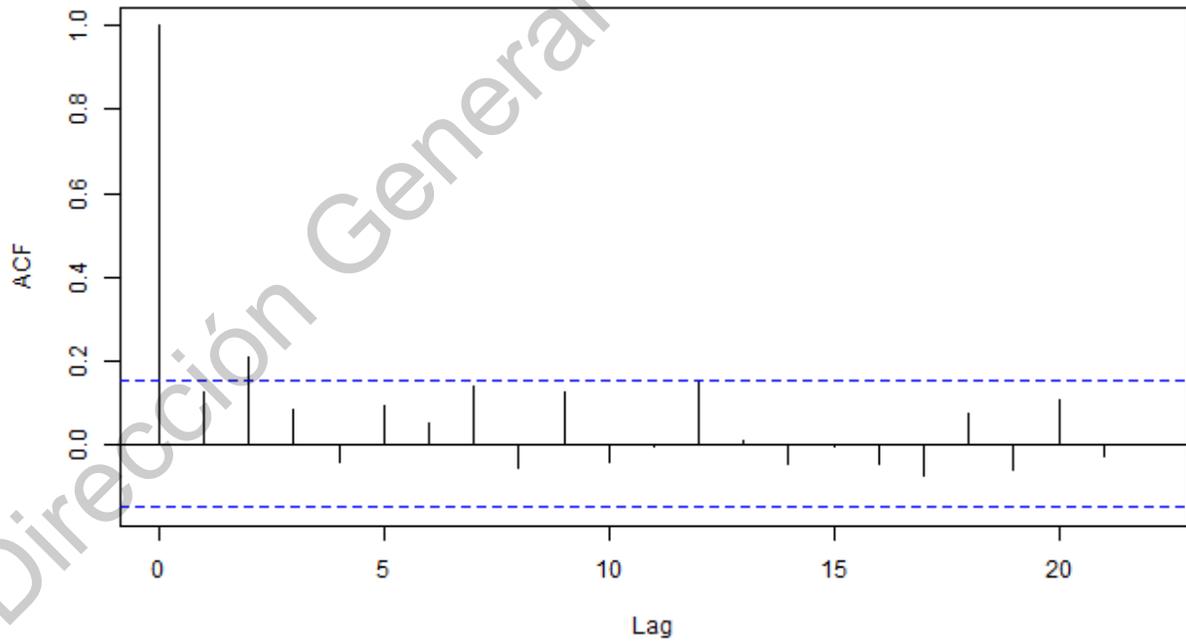
CYDSASAA.MX



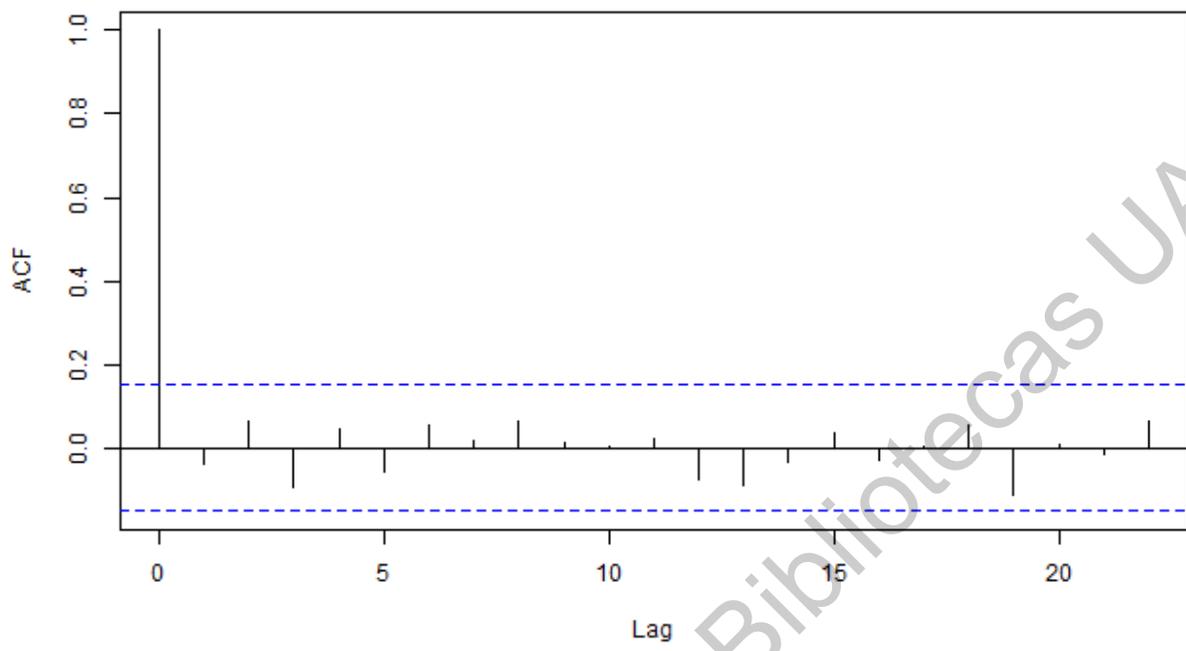
ELEKTRA.MX



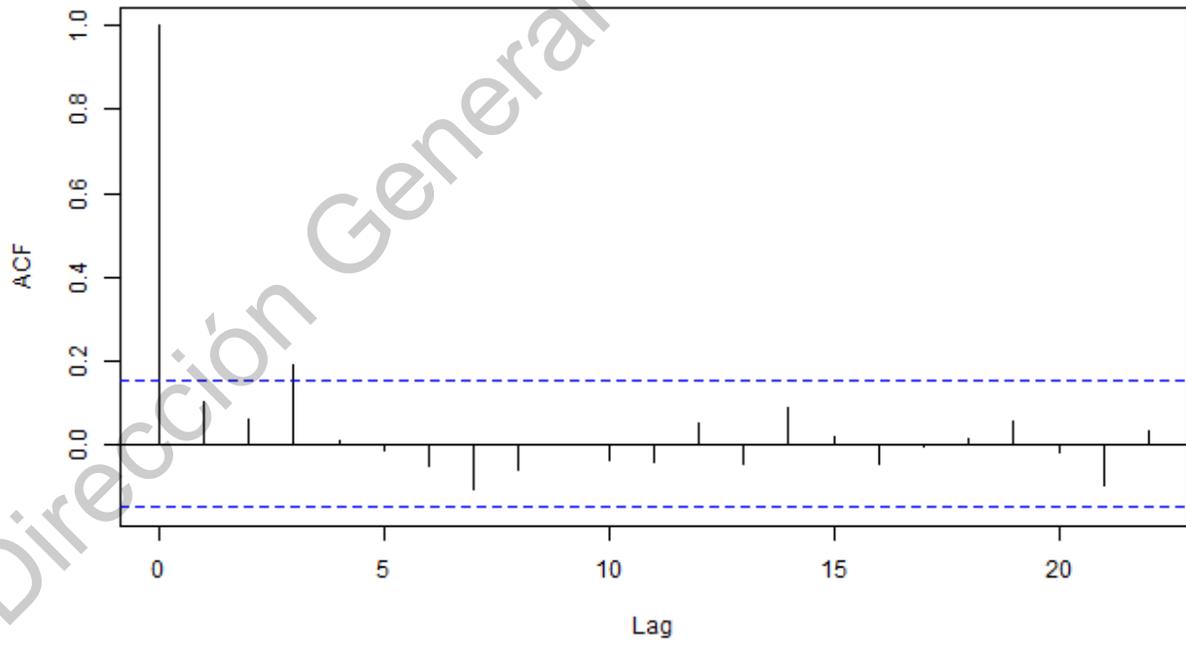
GCC.MX



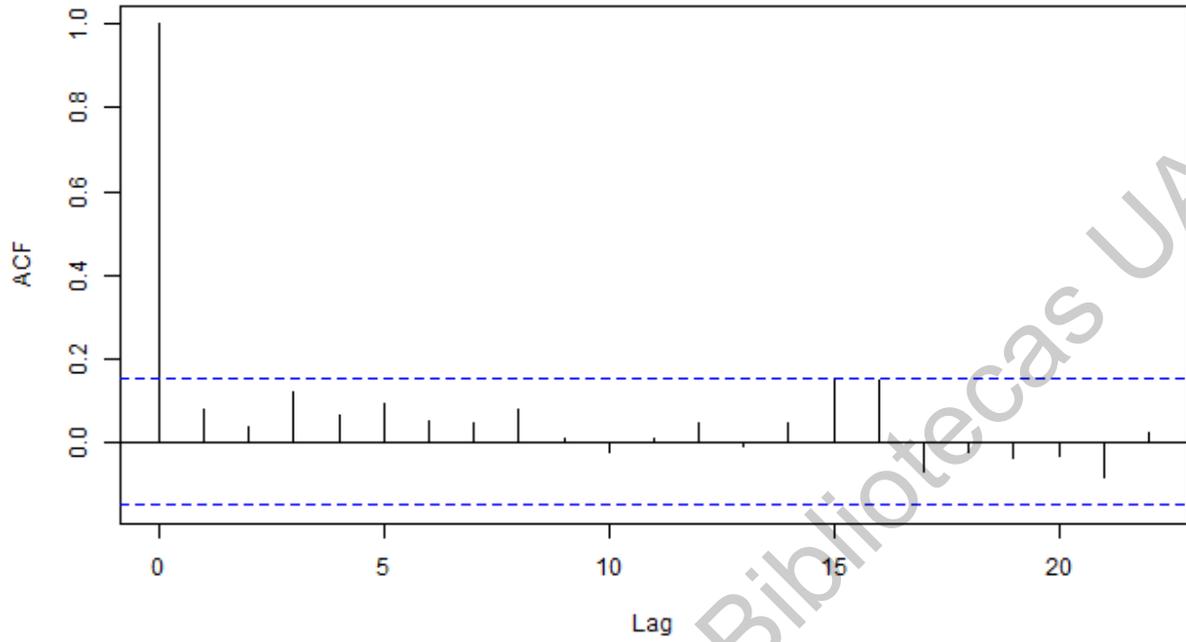
GFINBURO.MX



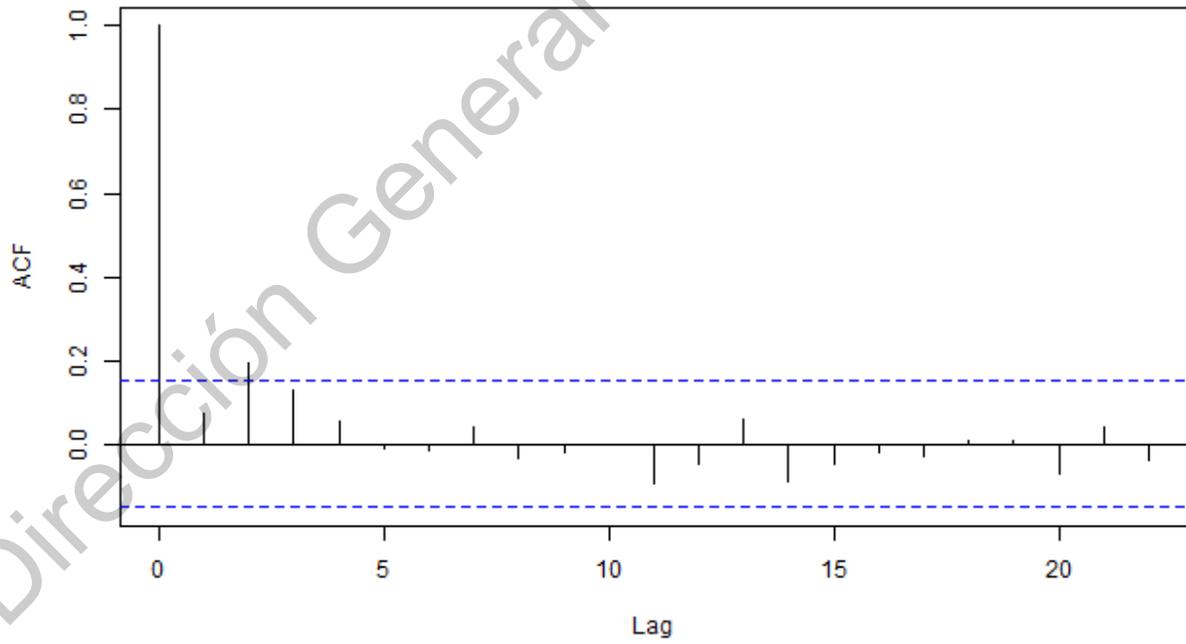
GFNORTEO.MX



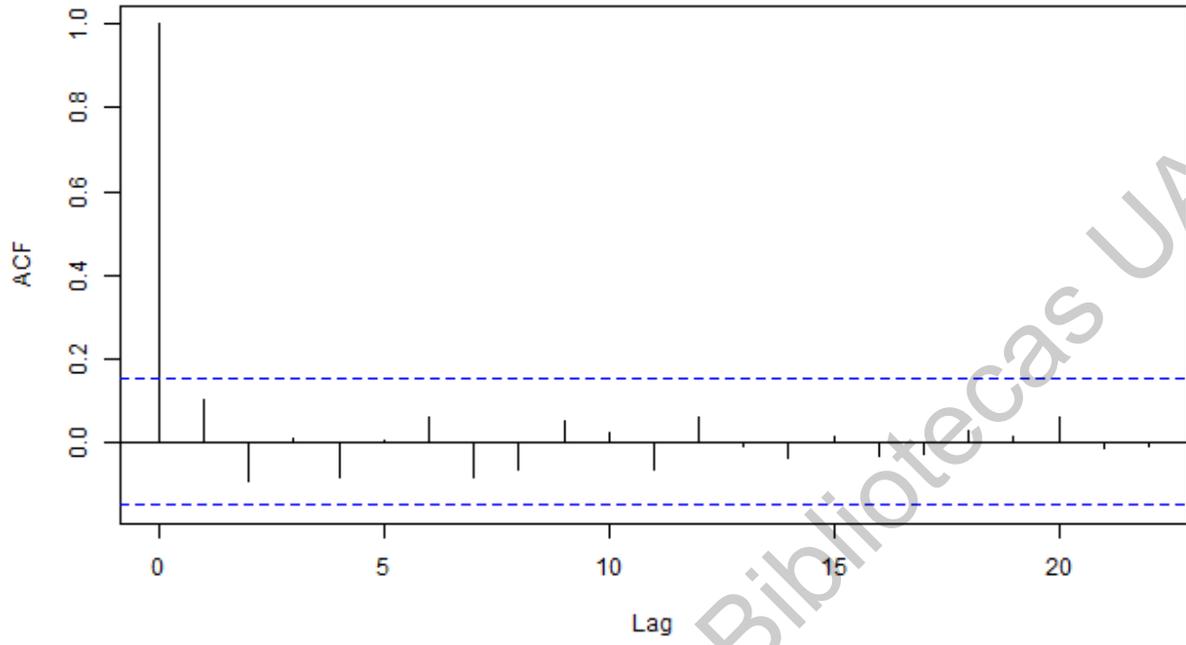
GMD.MX



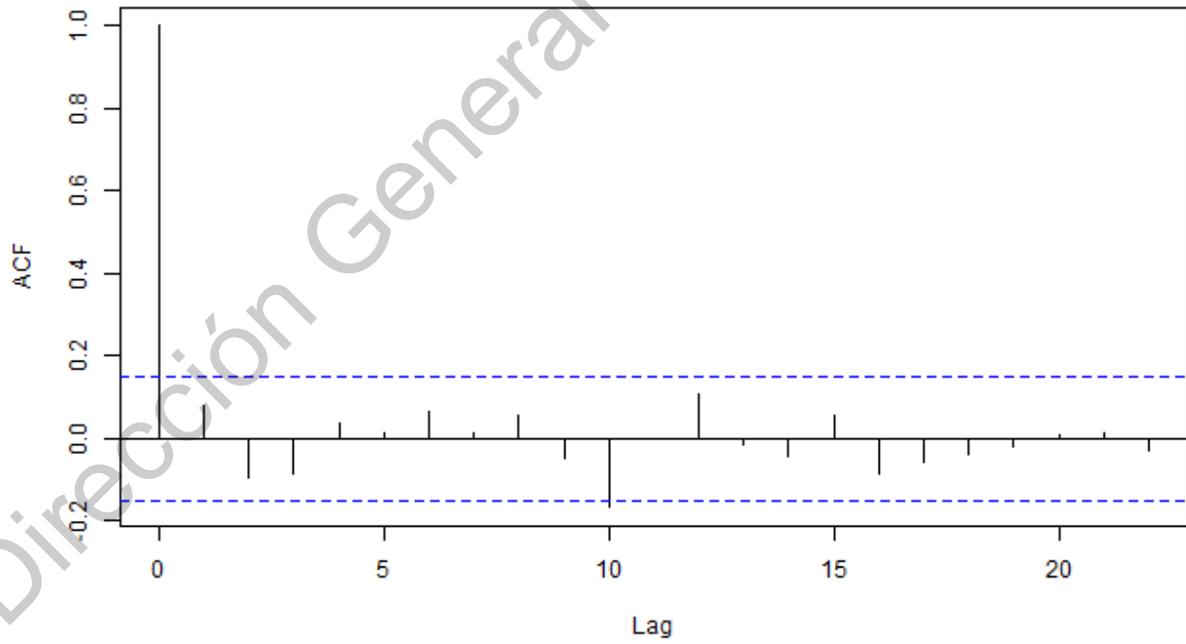
GMEXICOB.MX



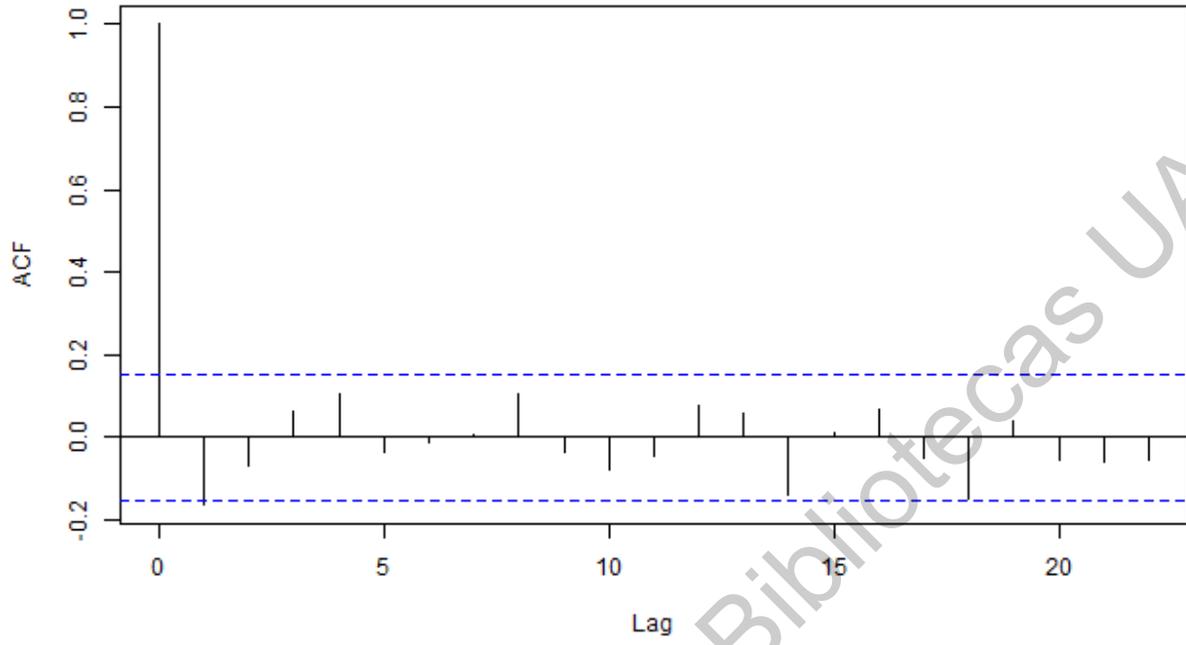
HERDEZ.MX



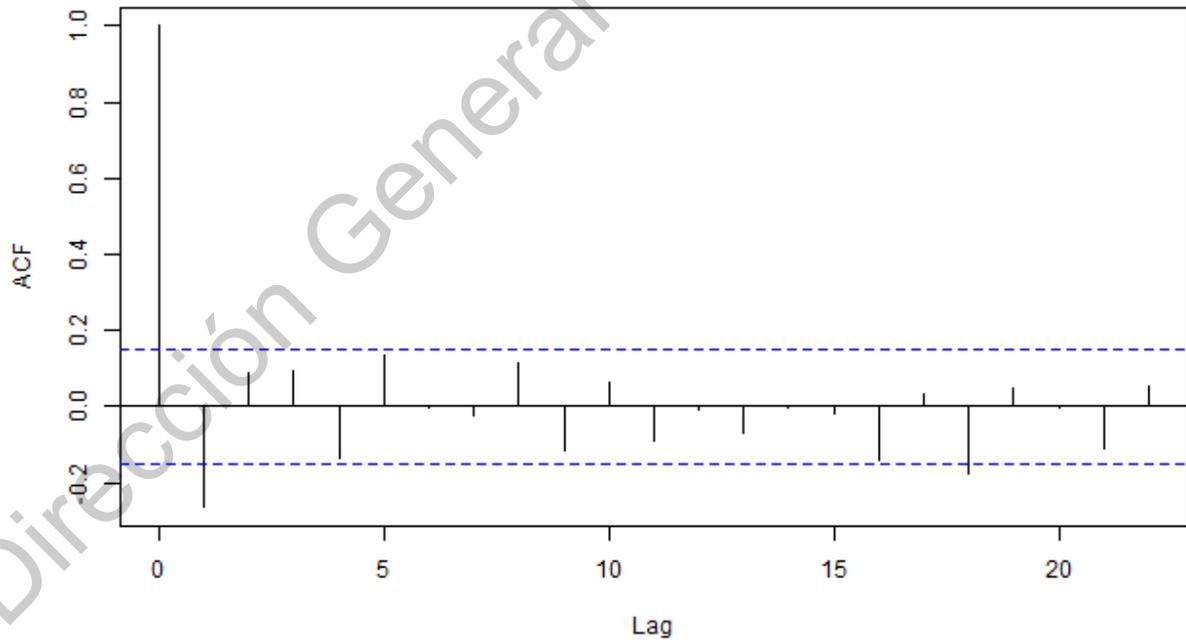
ICHB.MX



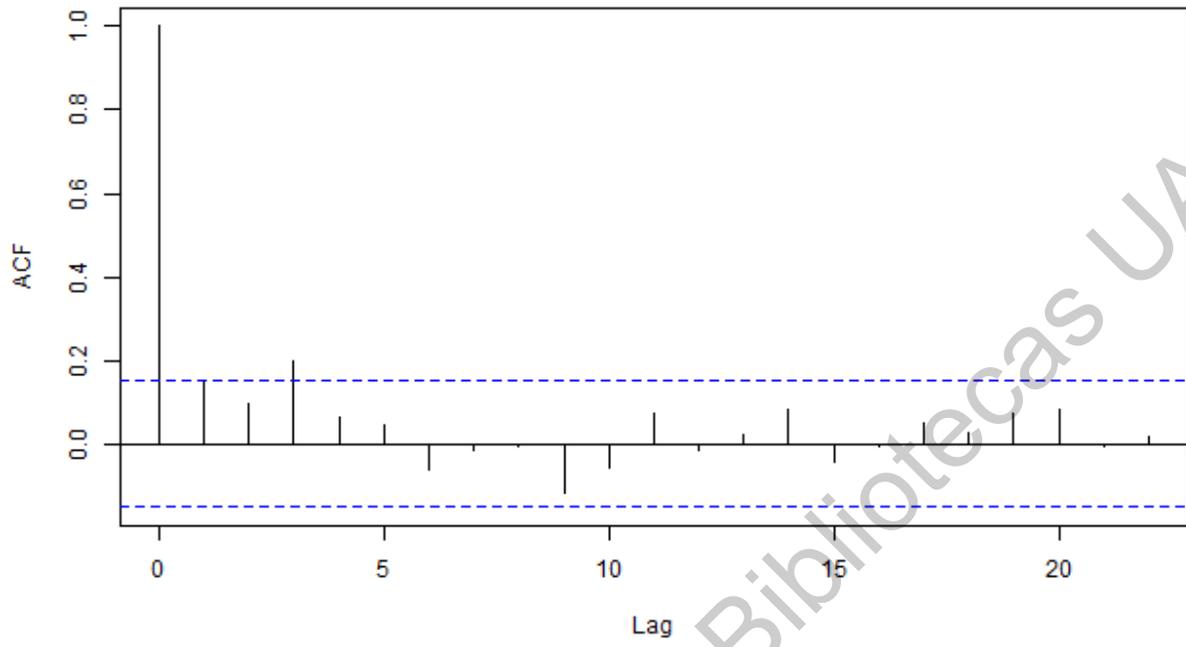
KIMBERA.MX



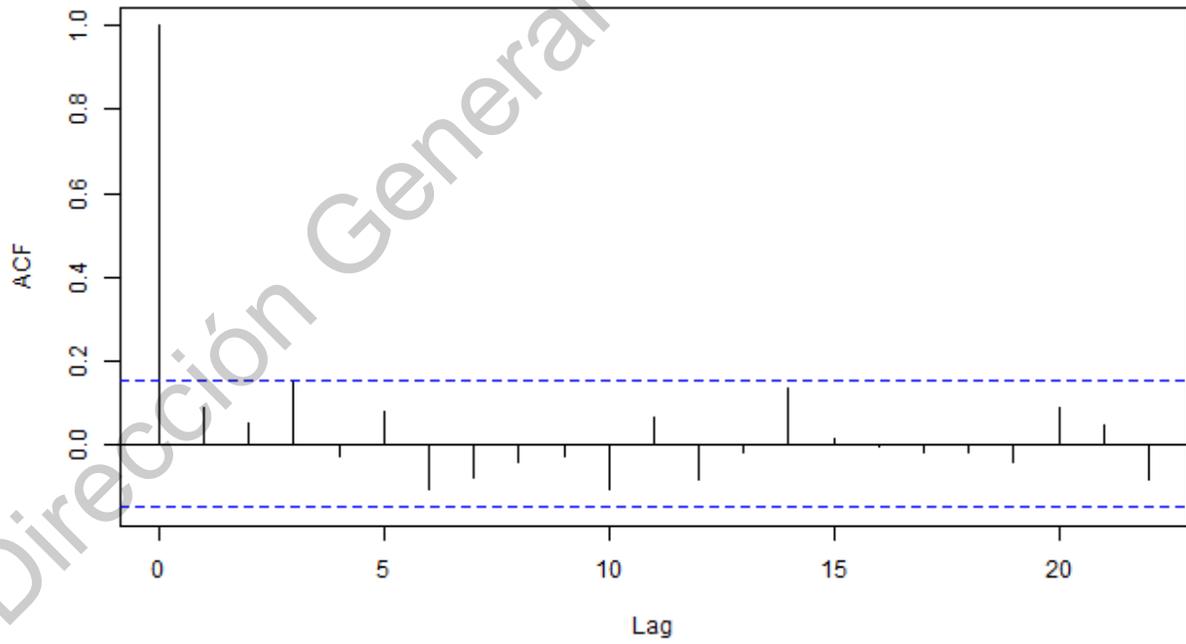
KUOB.MX



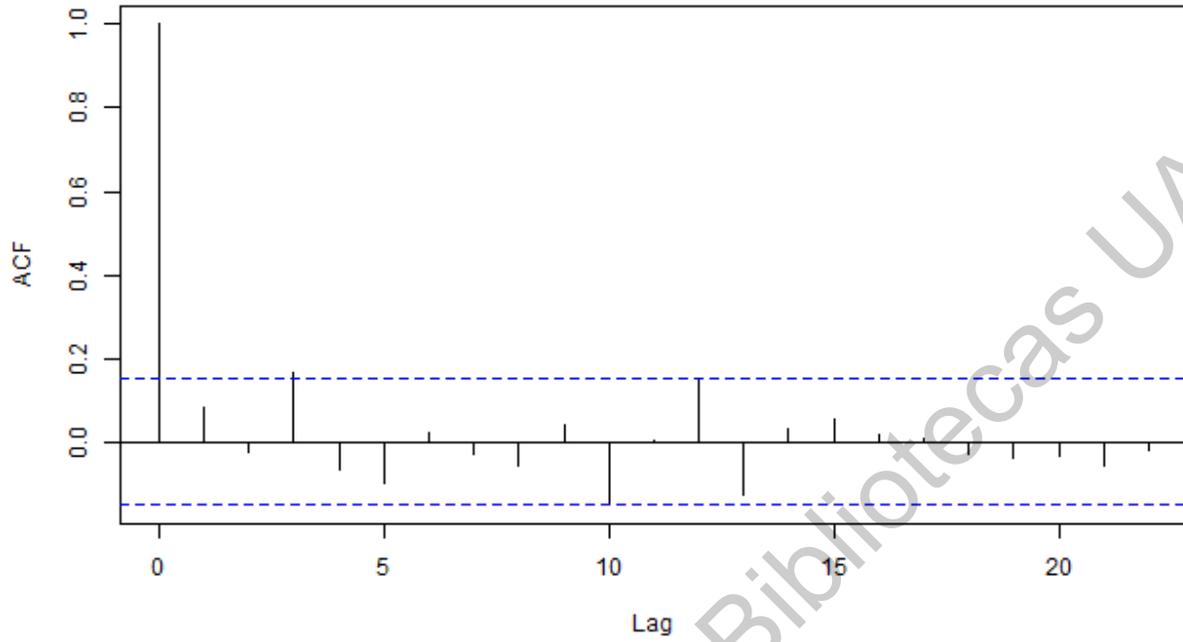
MEXCHEM.MX



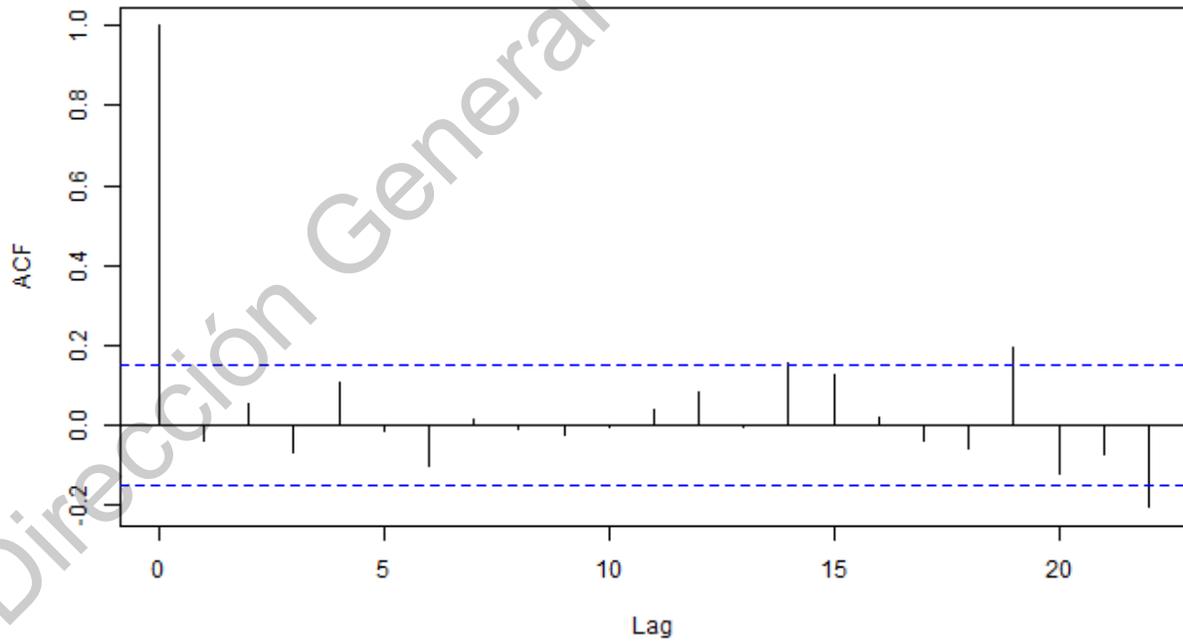
PE&OLES.MX



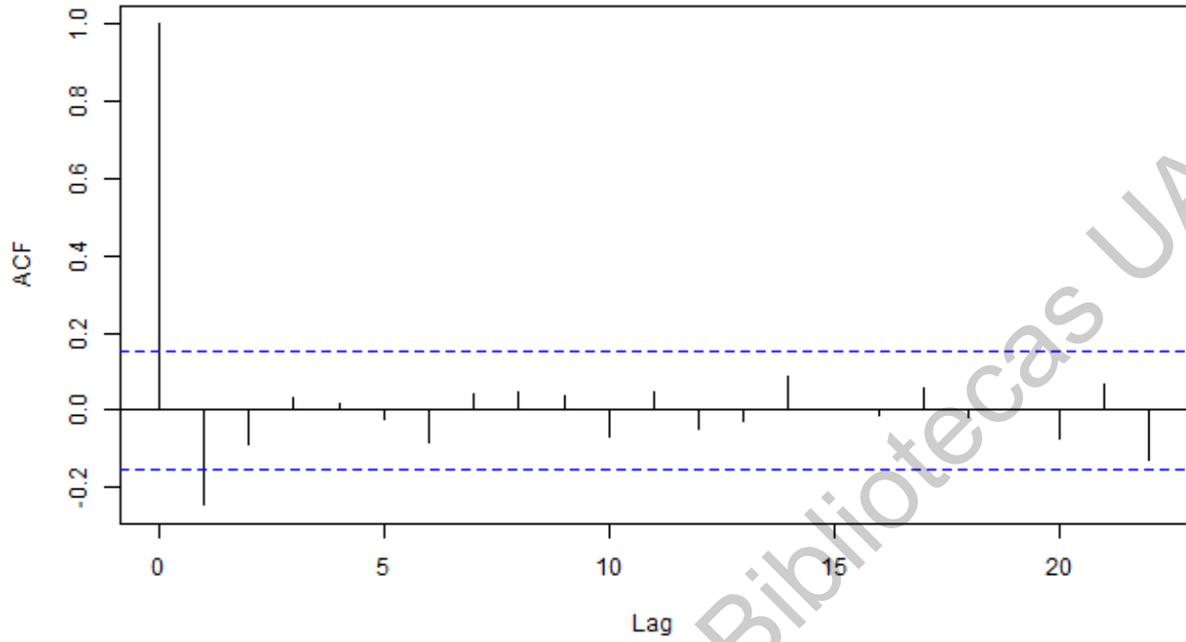
SORIANAB.MX



TLEVISACPO.MX



WALMEX.MX

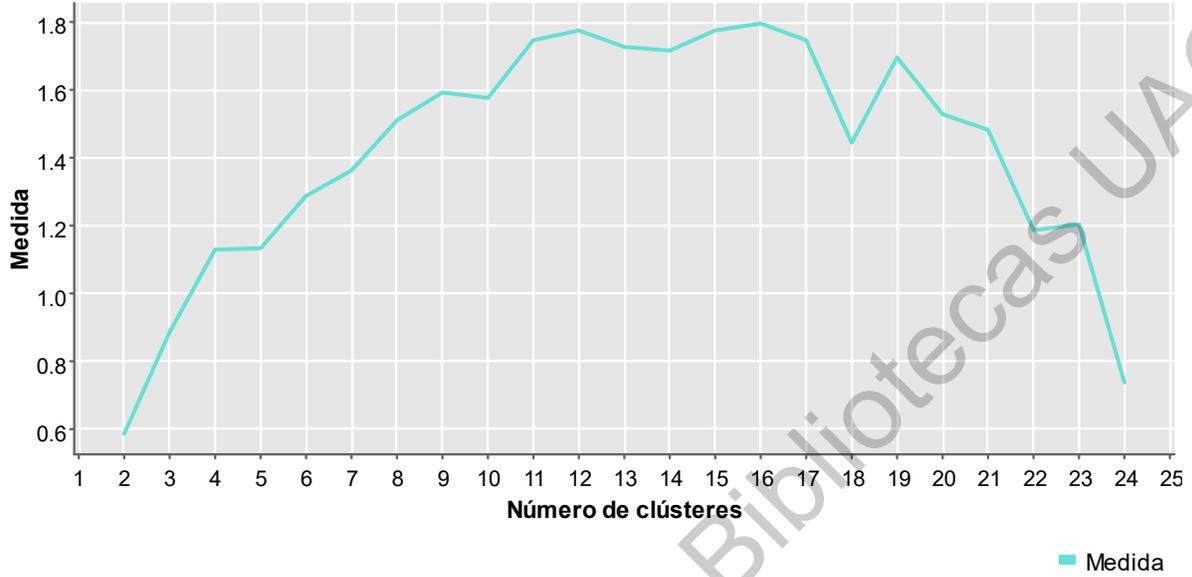


Dirección General de Bibliotecas UAQ

7.2 ANEXO B

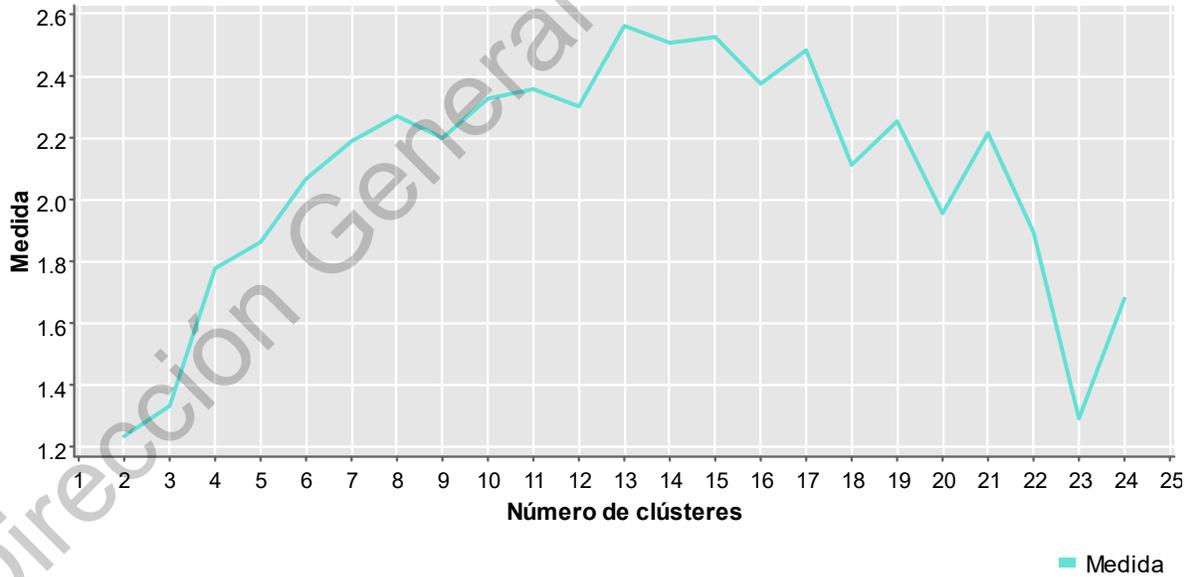
Ancho de silueta

K-medias con ACF y 1 lag



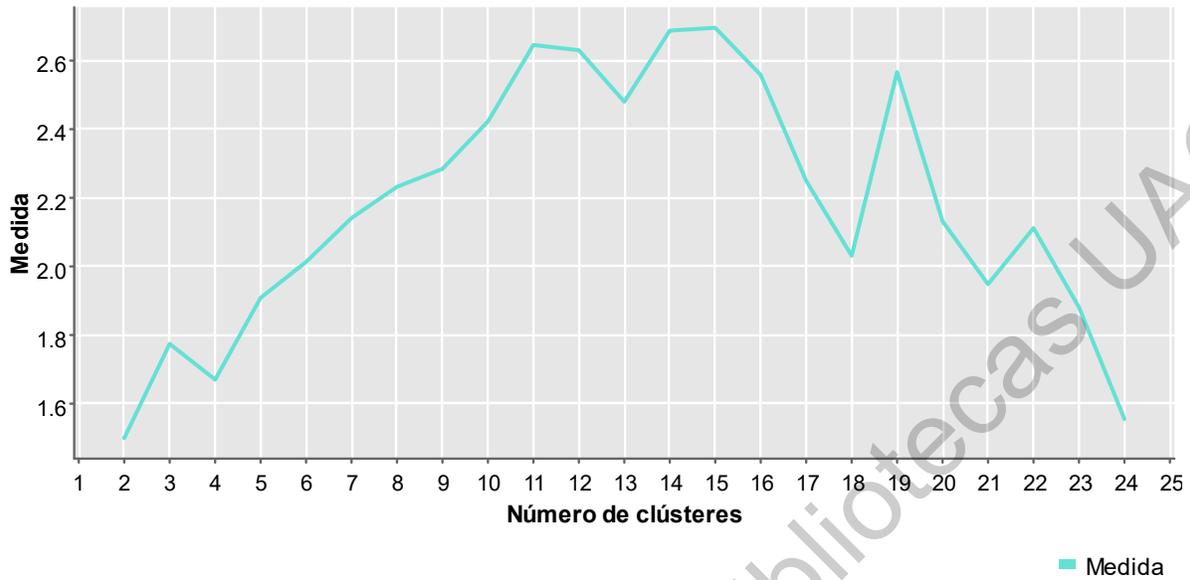
Ancho de silueta

K-medias con ACF y 3 lags



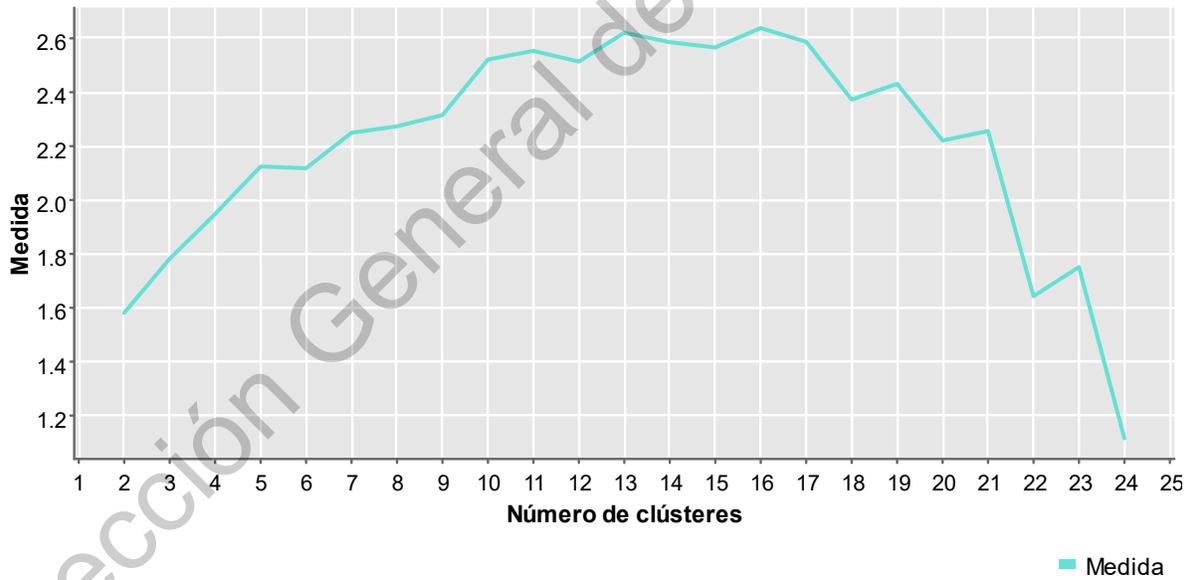
Ancho de silueta

K-medias con ACF y 5 lags



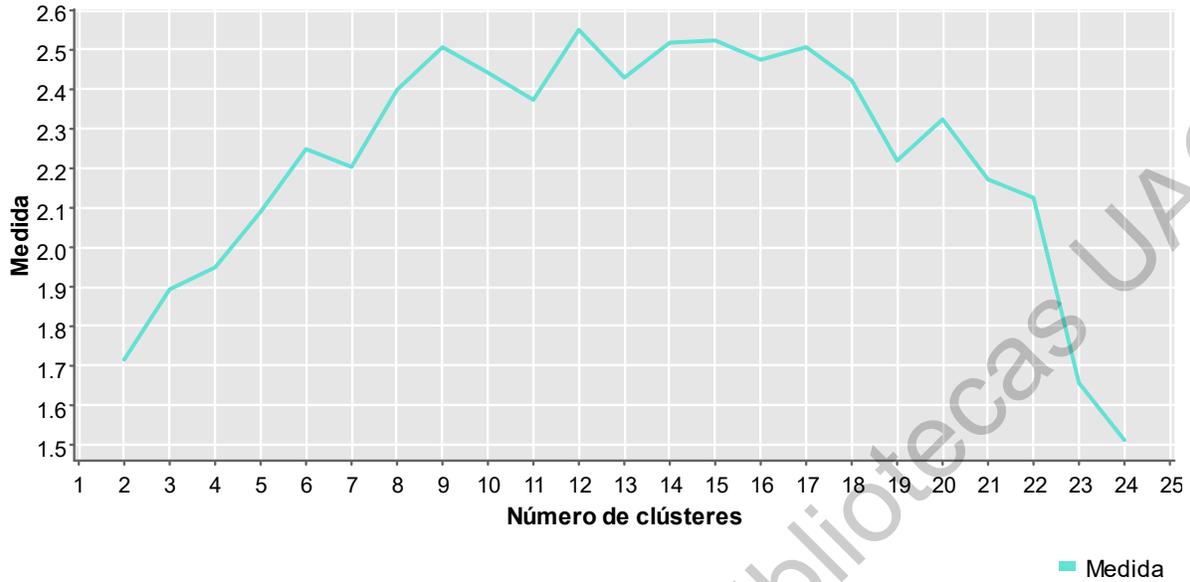
Ancho de silueta

K-medias con ACF y 7 lags



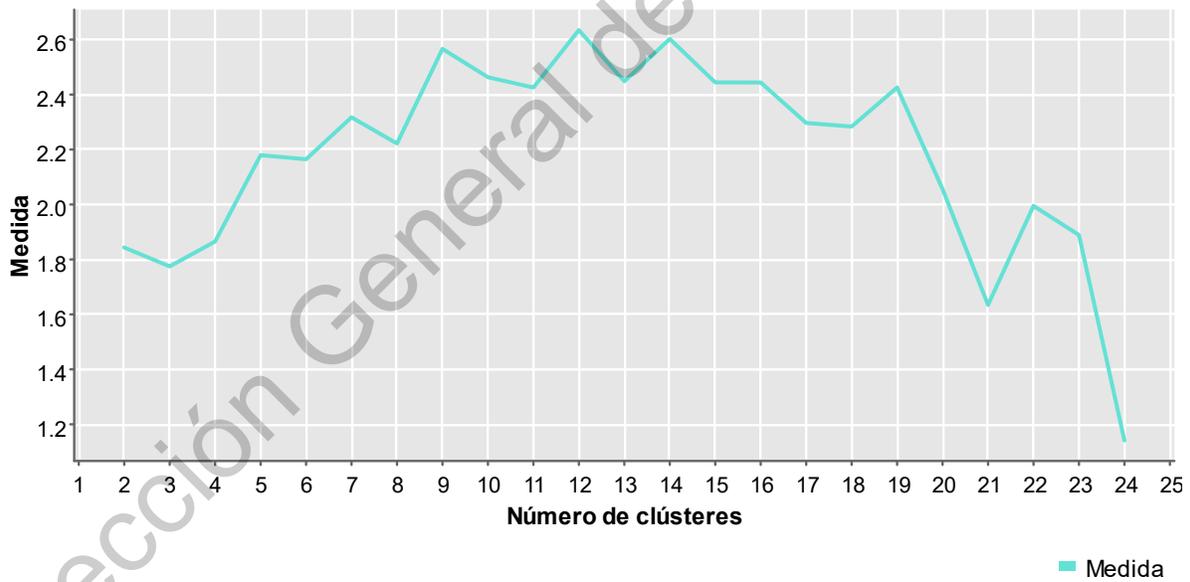
Ancho de silueta

K-medias con ACF y 9 lags



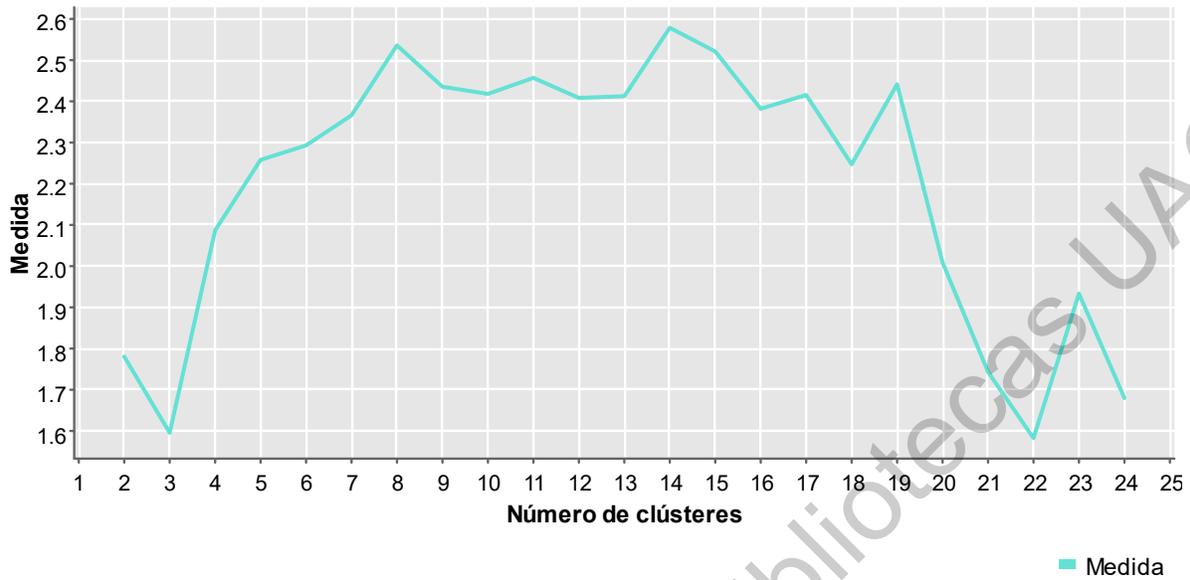
Ancho de silueta

K-medias con ACF y 11 lag



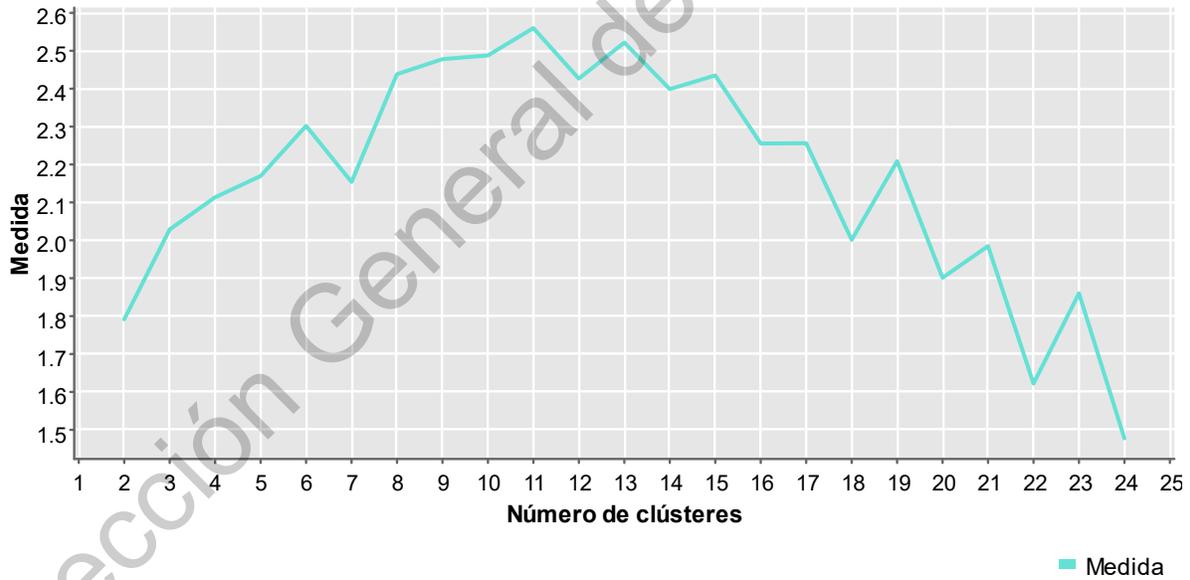
Ancho de silueta

K-medias con ACF y 13 lags



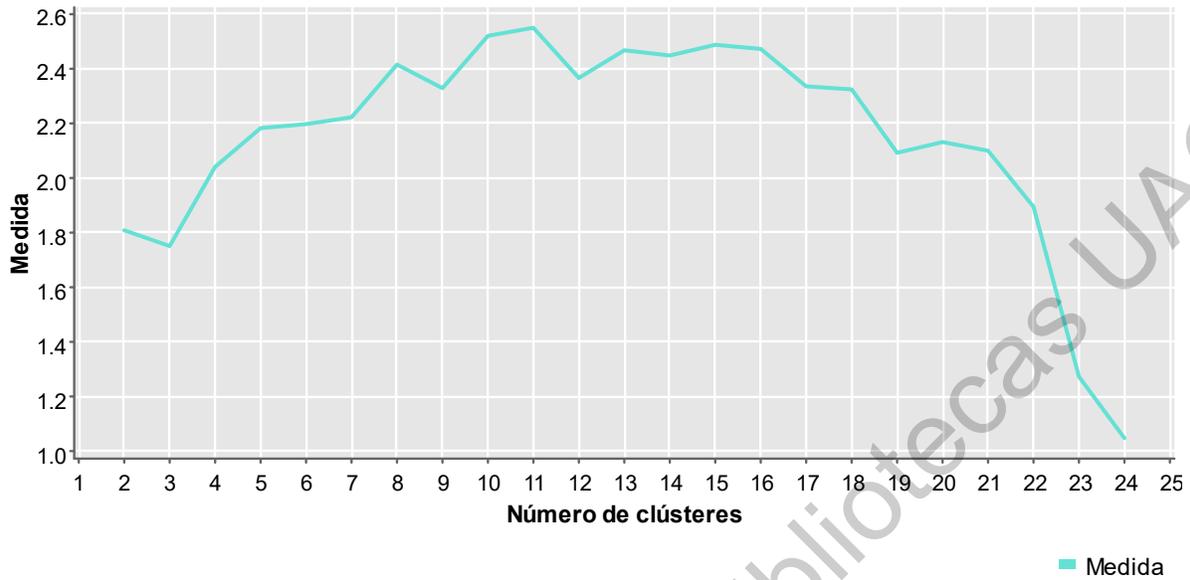
Ancho de silueta

K-medias con ACF y 15 lags



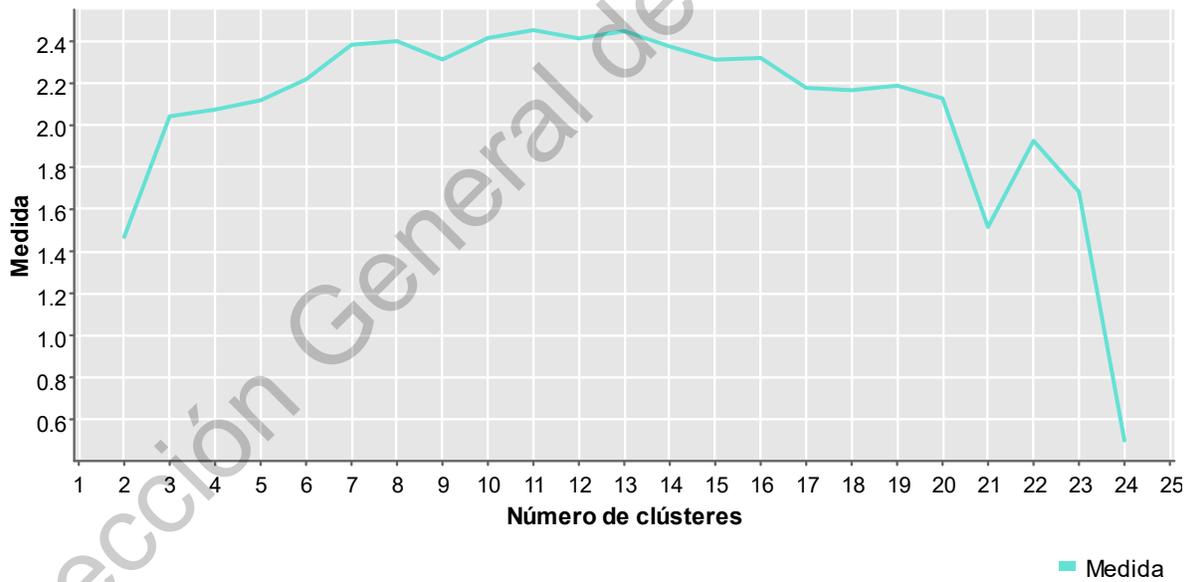
Ancho de silueta

K-medias con ACF y 17 lags



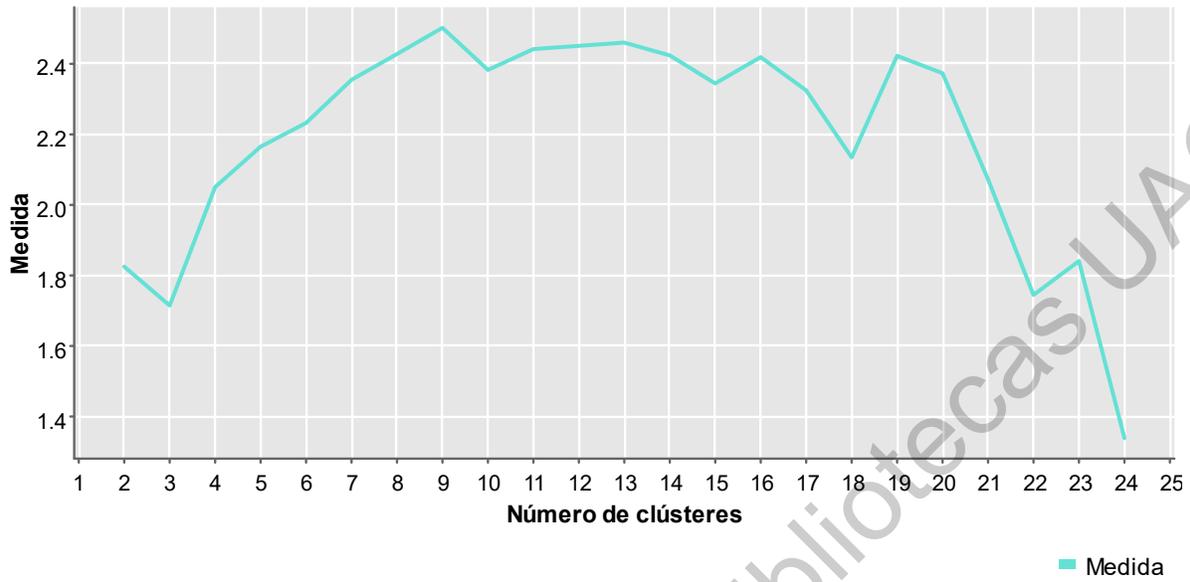
Ancho de silueta

K-medias con ACF y 19 lags



Ancho de silueta

K-medias con ACF y 21 lag

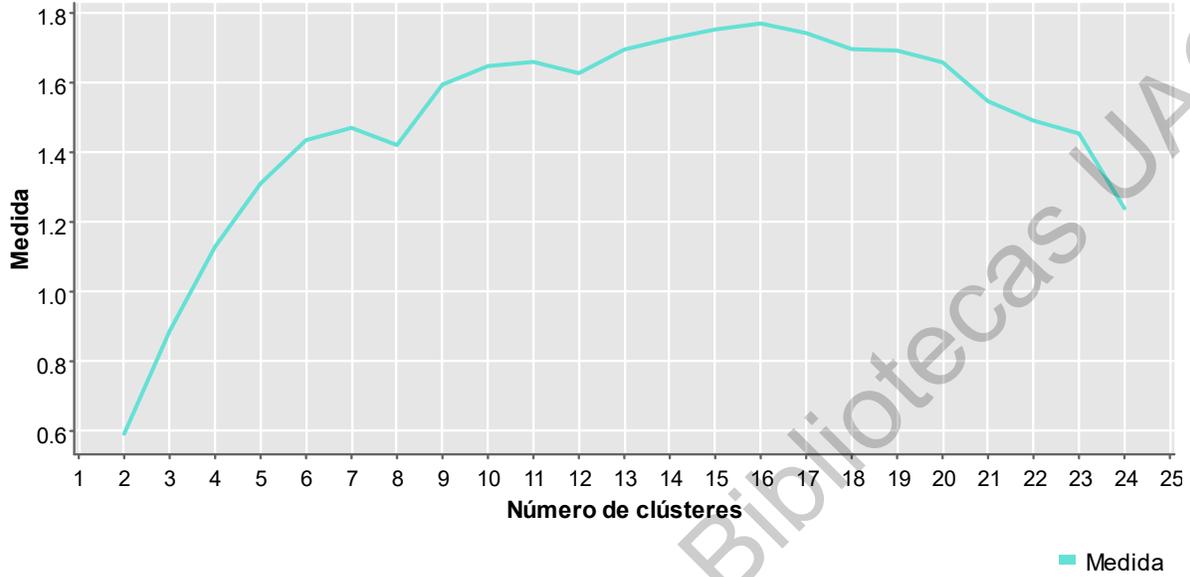


Dirección General de Bibliotecas UAQ

7.3 ANEXO C

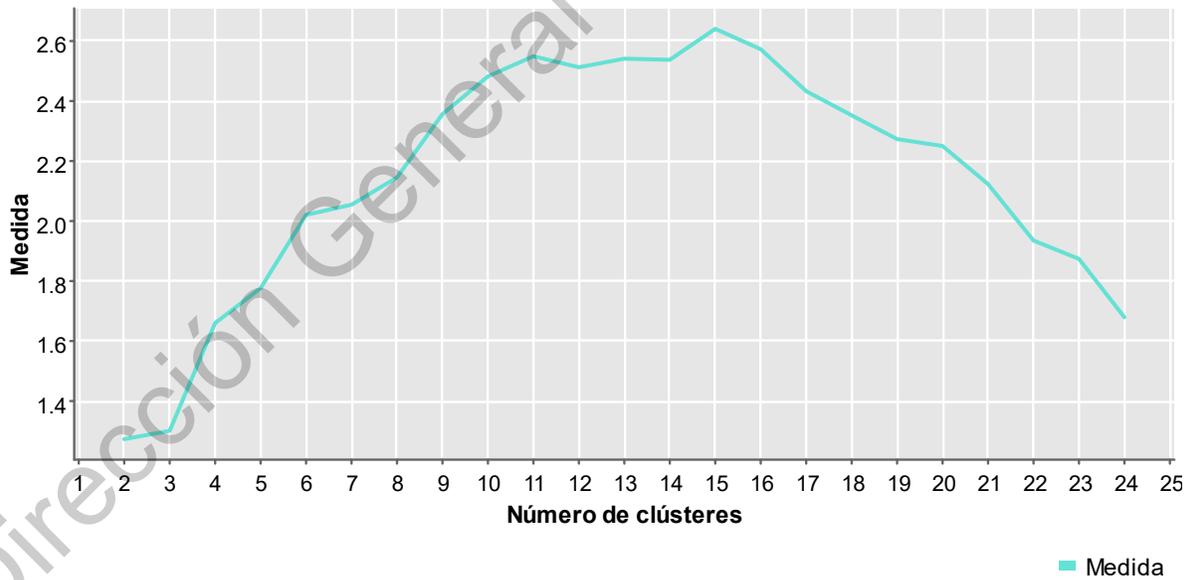
Ancho de silueta

Jerárquica con ACF y 1 lag



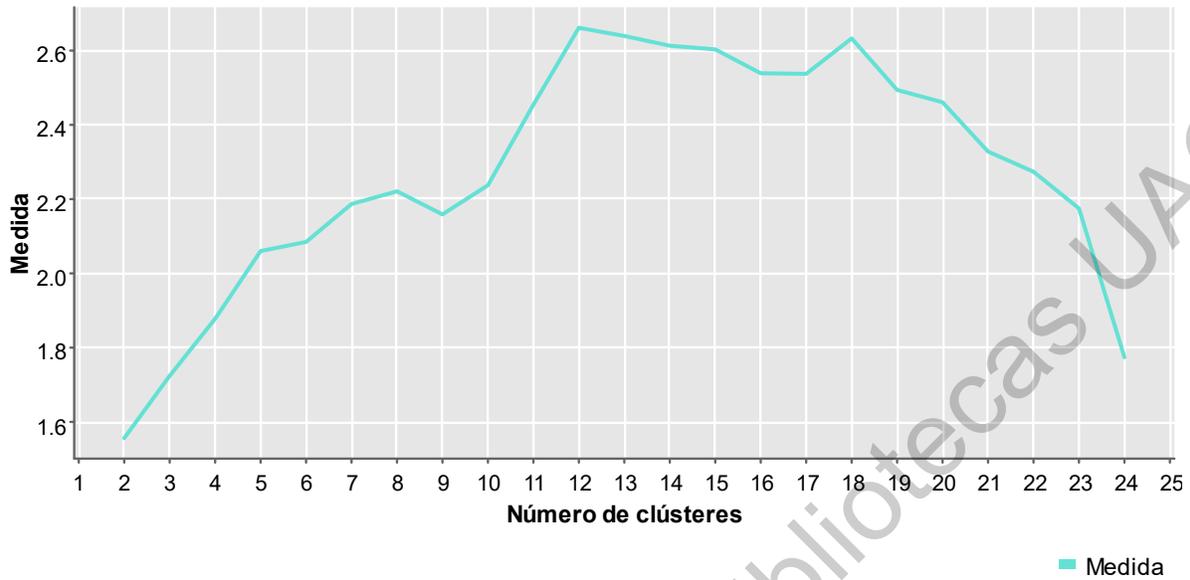
Ancho de silueta

Jerárquica con ACF y 3 lags



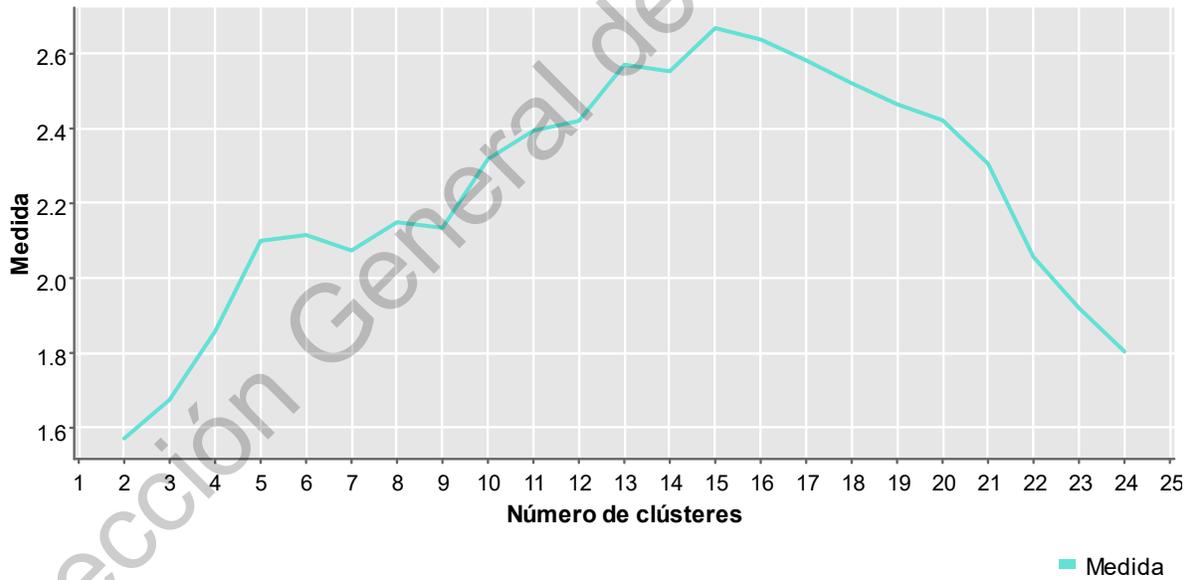
Ancho de silueta

Jerárquica con ACF y 5 lags



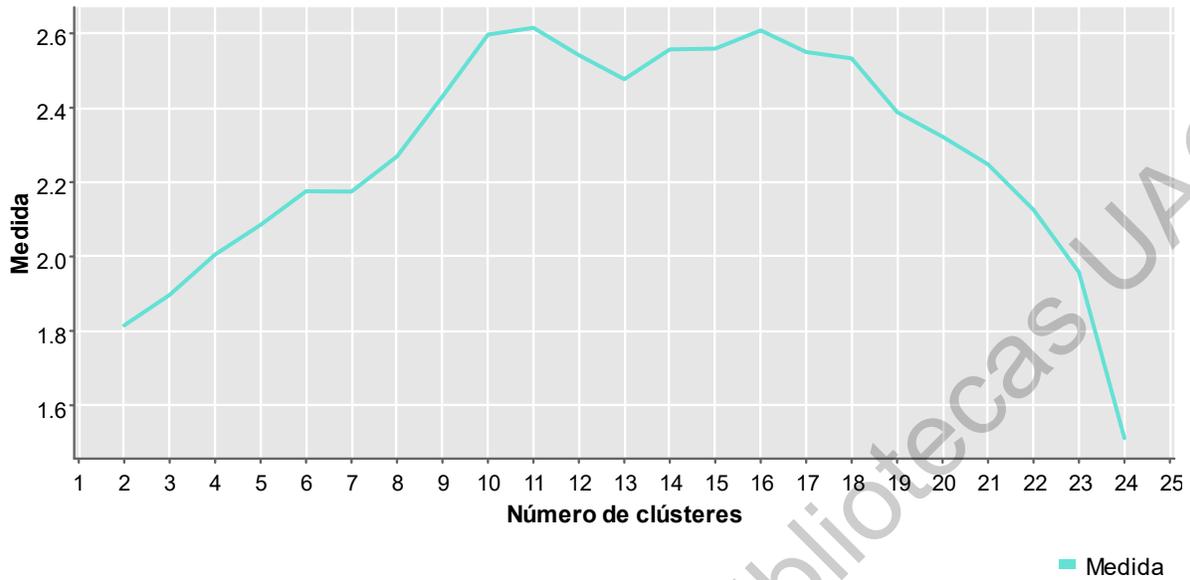
Ancho de silueta

Jerárquica con ACF y 7 lags



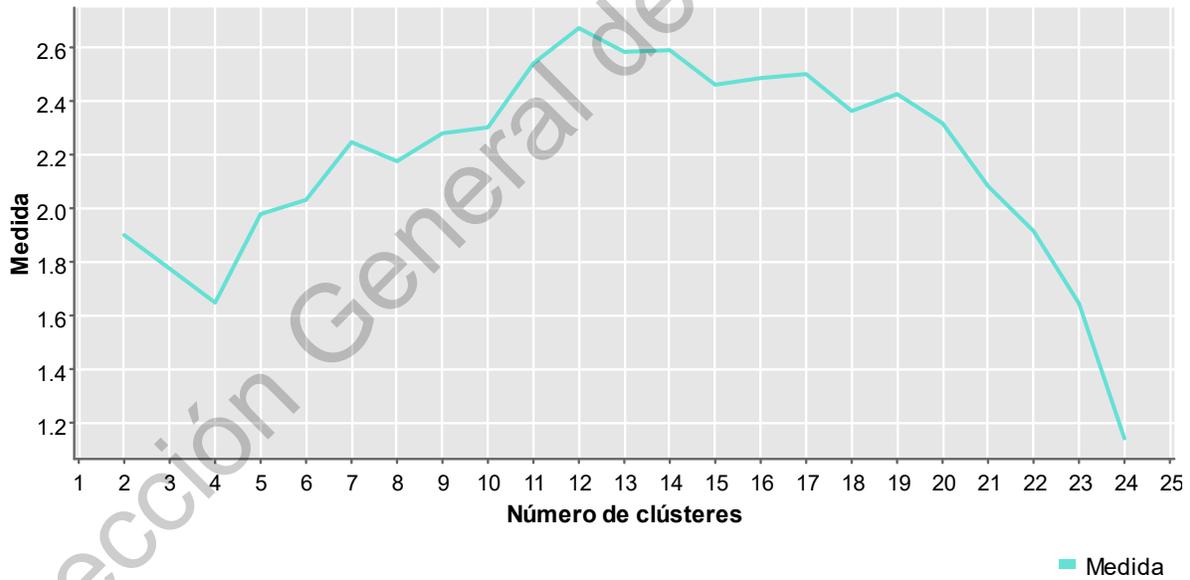
Ancho de silueta

Jerárquica con ACF y 9 lags



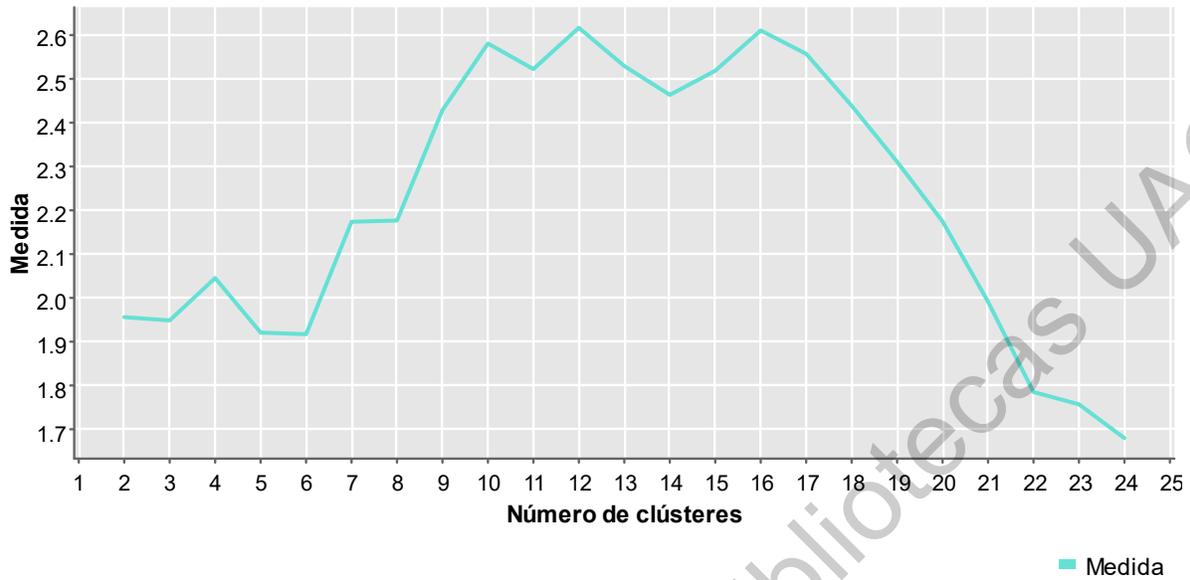
Ancho de silueta

Jerárquica con ACF y 11 lag



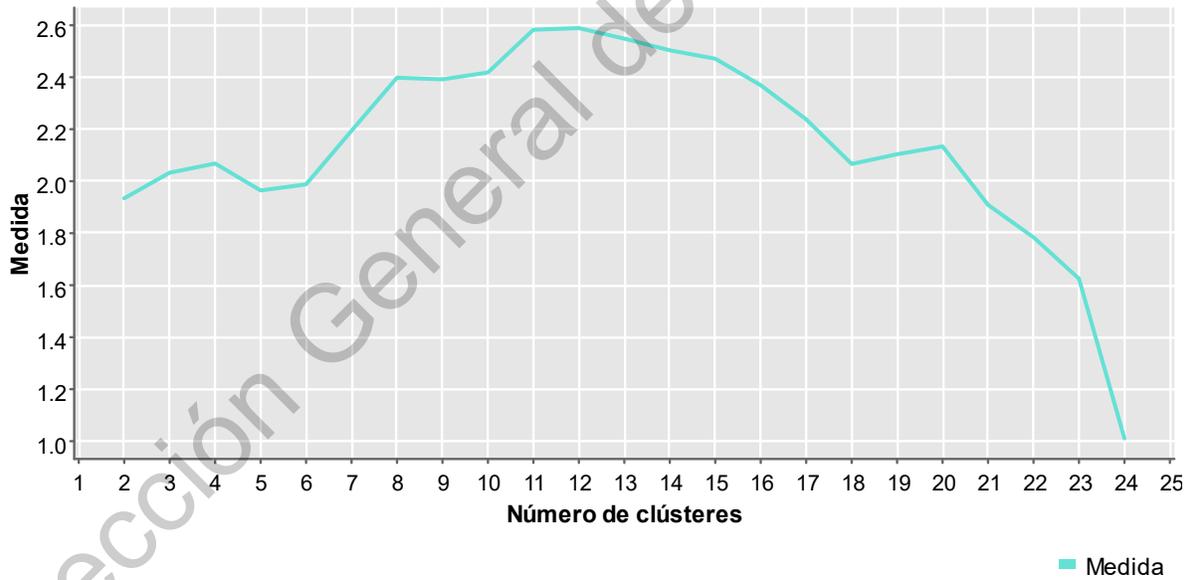
Ancho de silueta

Jerárquica con ACF y 13 lags



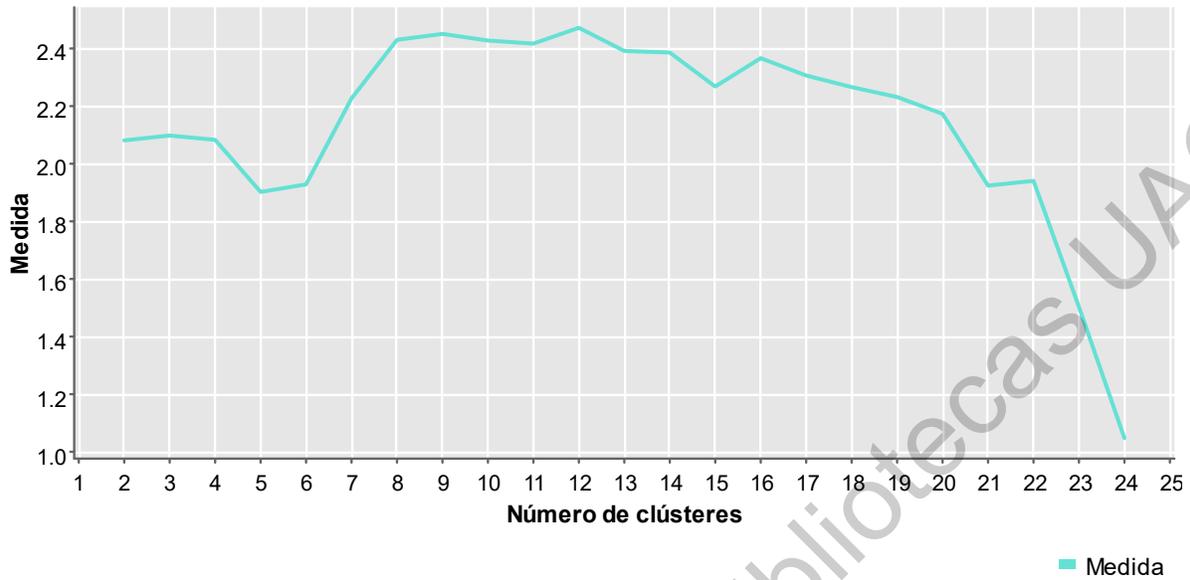
Ancho de silueta

Jerárquica con ACF y 15 lags



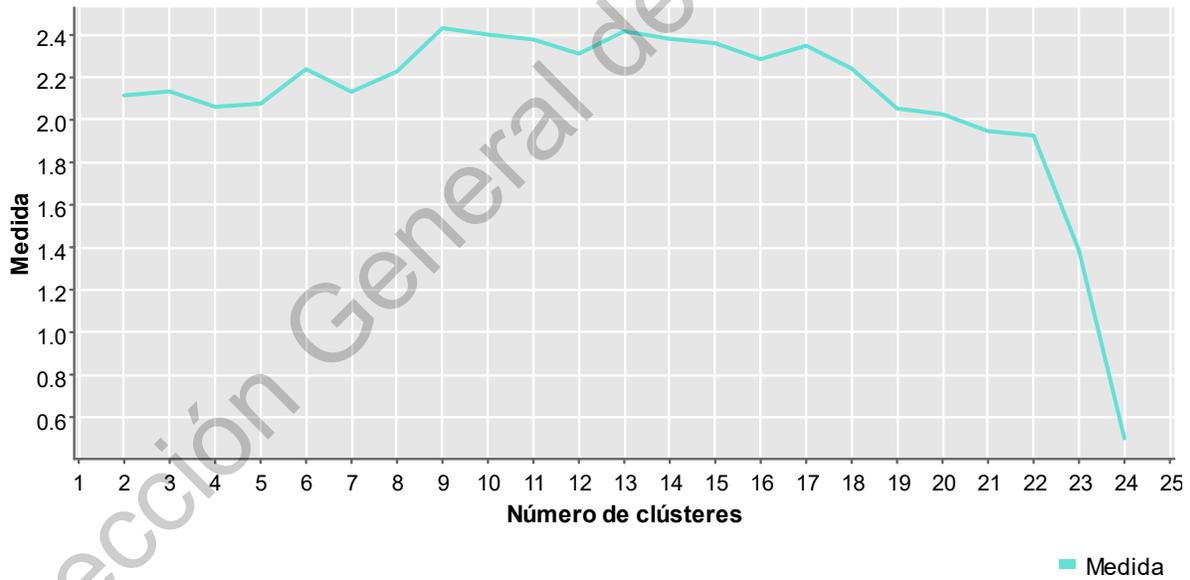
Ancho de silueta

Jerárquica con ACF y 17 lags



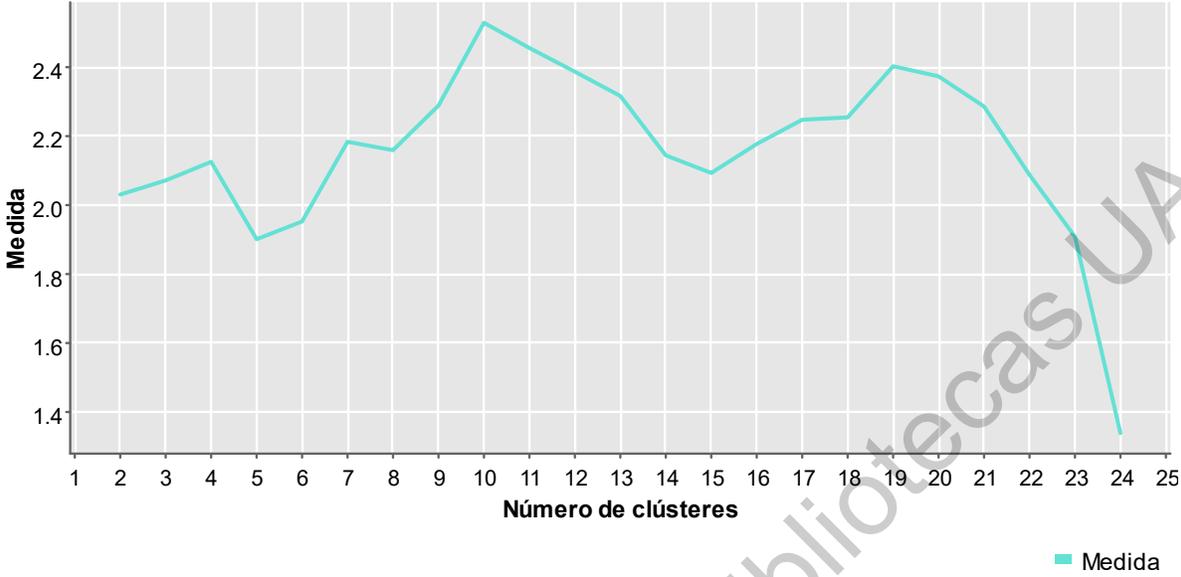
Ancho de silueta

Jerárquica con ACF y 19 lags



Ancho de silueta

Jerárquica con ACF y 21 lag



Dirección General de Bibliotecas UAQ