

**DESARROLLO DE UN MODELO PARA EL  
PRONÓSTICO DE LA DEMANDA POR COLOR  
EN LA EMPRESA COATS CADENA ANDINA S.A.**

Tesista:  
**CAROLINA AGUDELO RENDÓN**

**UNIVERSIDAD TECNOLÓGICA DE PEREIRA  
MAESTRIA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA  
FACULTAD INGENIERÍA INDUSTRIAL  
Pereira, Risaralda 2013**

**Desarrollo de un modelo para el pronóstico de la  
demanda por color en la empresa Coats Cadena Andina  
S.A.**



Tesista:  
**CAROLINA AGUDELO RENDÓN**

Directora:  
**LUZ MARÍA OSPINA GUTIÉRREZ**  
Magister en Investigación Operativa y Estadística  
Universidad Tecnológica de Pereira

Documento presentado como requisito para optar al título de  
**Magister en Investigación Operativa y Estadística**

**UNIVERSIDAD TECNOLÓGICA DE PEREIRA**  
**MAESTRIA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA**  
**FACULTAD INGENIERÍA INDUSTRIAL**  
**Pereira, Risaralda 2013**

Nota de Aceptación

---

---

---

---

---

---

---

Jurado

---

Jurado

---

Jurado

Pereira, 2013.

# **Agradecimientos**

A mi directora M.Sc. Luz María Ospina Gutiérrez por toda su dedicación y atención.

A los integrantes del grupo de Investigación Control e Instrumentación de la Universidad Tecnológica de Pereira, y al Ing. Alejandro Mora Barón por su colaboración.

Al señor Gildardo Herrera, gerente de manufactura de Coats Cadena Andina S.A. por su apoyo incondicional en la culminación de este trabajo.

# Tabla de contenido

<b>Introducción</b> .....	1
<b>1. Planteamiento del Problema</b> .....	2
<b>2. Justificación</b> .....	5
2.1. Pertinencia .....	5
2.2. Impacto .....	6
2.3. Viabilidad y revisión bibliográfica .....	6
<b>3. Objetivos</b> .....	12
3.1. Objetivo general .....	12
3.2. Objetivos específicos .....	12
<b>4. Marco Conceptual</b> .....	13
4.1. Series de tiempo de la demanda del color .....	13
4.2. Introducción a los Sistemas dinámicos .....	14
4.3. Series temporales y los Procesos de media móvil y ARMA .....	23
4.4. Modelos ARMA .....	37
4.5. Modelos de heterocedasticidad condicional .....	60
4.6. Método de los datos sustitutos .....	73
<b>5. Desarrollo Metodológico</b> .....	80
5.1. Metodología de desarrollo .....	80
5.2. Descripción de las bases datos utilizadas .....	81
5.3. Software .....	81
5.4. Algoritmos .....	81

<b>6.</b>	<b>Resultados y Análisis</b> .....	82
6.1.	Selección del Modelo ARMA.....	82
6.2.	Selección del modelo GARCH.....	98
6.3.	Validación de las características de la dinámica de la serie de la demanda del color café de Coats Cadena Andina S.A. ....	106
6.4.	Validación de los pronósticos respecto a las observaciones reales.....	107
<b>7.</b>	<b>Conclusiones</b> .....	110
<b>8.</b>	<b>Futuras líneas de trabajo</b> .....	112
	Anexo 1. Elección del modelo arma con base en el criterio de Akaike .....	113
	Anexo 2. Algoritmo de ajuste del mejor modelo ARMA.....	114
	Anexo 3. Algoritmo de ajuste de un Modelo Garch .....	115
	Anexo 4. Función Algoritmo SSS.....	116
	Anexo 5. Función Algoritmo TFT.....	118
	Anexo 6. Algoritmo con el que se generan los sustitutos SSS y se calculan los estadísticos AMI y CMP .....	121
	Anexo 7 Algoritmo con el que se generan los sustitutos TFT y se calculan los estadísticos AMI y CMP .....	125
	<b>Referencias</b> .....	129

# Introducción

Este trabajo presenta los resultados de investigación asociados al estudio de la dinámica de las series de la demanda del color en Coats Cadena Andina S.A.

Aquí se propone una metodología basada en el análisis de los modelos de pronóstico para series de tiempo ARMA y GARCH, para tratar de determinar cuál será la demanda futura por color de la empresa Coats Cadena Andina S.A.

En los tres primeros capítulos de este trabajo se aborda el planteamiento del problema, la justificación y los objetivos.

El cuarto capítulo muestra el comportamiento de las series de tiempo de la demanda del color en Coats Cadena Andina S.A y los dos principales enfoques del pronóstico de la demanda del color que han sido utilizados por los expertos a nivel mundial. También se encuentra toda la fundamentación teórica de los sistemas dinámicos, datos sustitutos y de los modelos ARMA y GARCH.

El capítulo 5 corresponde a la metodología del desarrollo del trabajo, la descripción de la base de datos y los softwares utilizados.

En el capítulo 6 se exponen los resultados y la discusión de datos obtenidos a través del desarrollo de esta metodología para la serie de la demanda del color en Coats Cadena Andina S.A.

# 1. Planteamiento del problema

Según Chang, Gao, Zhang, Lu y Pan (2009a), uno de los temas más relevantes para la industria textil consiste en tratar de determinar cuáles serán las tendencias futuras de la moda en cuanto al color, ya que este se ha aceptado como uno de los elementos más activos en el comercio. El mejor diseño de color así como las estrategias de mercadeo, puede estimular el comportamiento de compra de productos de manera eficiente y aumentar la utilidad de la empresa. Así, para Chang, et al. (2009a), Chang, Gao y Zhang (2009b) y Stanstield y Whitfield (2005), los colores de moda, los símbolos de moda y los gustos, juegan papeles importantes en la economía de comercialización. Por lo tanto, la idea de pronosticar cuál será el color más popular entre los consumidores ha recibido mucha atención en el mundo entero.

Para Choi, Hui, Ng y Yu (2011), y Sun, Choi, Au y Yu (2008), la industria de insumos textiles, así como la de la moda, cuentan con una larga cadena de suministros que incluyen operaciones tales como el diseño, las ventas, la fabricación, la financiación, y la marca, entre otras. Al mismo tiempo, es sabido que esta industria se enfrenta a tendencias y demandas altamente variables; por lo tanto, el uso de información es un tema importante. Asimismo, Choi, et al.; Sun, et al.; Yu, et al. (2011); y Choi, et al. (2011), señalan que la precisión en los pronósticos es verdaderamente crucial. Sin embargo, generar pronósticos de las propiedades de los productos de moda como el color es siempre un trabajo desafiante debido a que estas propiedades, por lo general, tienen ciclos de vida muy cortos.

Para Chang, et al. (2009b), el enfoque principal de las investigaciones acerca de las tendencias de la moda en cuanto al color, se ha basado en el análisis sistémico. Según el mismo Chang, et al., esta metodología cualitativa es ampliamente utilizada por las asociaciones de la moda u organismos como la Comisión Internacional para el Color en Textiles y Moda (IC), la Autoridad Internacional del Color (ICA), e Inexmoda, entre otros. Con esta metodología se analizan factores complejos relacionados con el color de moda, como el estilo de vida social, la



economía mundial, las preferencias de los consumidores en cuanto al color, y asuntos militares y políticos, entre otros. Las decisiones en cuanto al color de moda en el futuro depende de expertos, y de las intuiciones de los pronosticadores. Sin embargo, Choi, et al. (2011), plantean que estos pronósticos podrían presentar sesgos de decisión debido a los juicios subjetivos y a las diferencias entre los mercados.

La práctica mencionada anteriormente ha motivado investigaciones con el fin de desarrollar herramientas matemáticas para ser ejecutadas en la predicción de la tendencia de la demanda de color, la metodología de análisis histórico de datos, es decir, la metodología cuantitativa, tiene como objetivo investigar los posibles cambios en las normas existentes en la tendencia del color de moda (Chang, et al. 2009a; Chang, et al. 2009b; Stanstield y Whitfield, 2005; Choi, et al. 2011; Yu, et al. 2012; Wu, et al. 2010; Gu y Liu, 2011). Esta metodología de análisis estadístico-matemática fue desarrollada originalmente por la Asociación del Color de Moda de Japón (FAFCA) y desde su introducción varios métodos de pronóstico han sido estudiados y aplicados a la industria textil y de la moda cuando se dispone de abundantes datos históricos (Choi, et al., 2011). Estudios previos de las propiedades de los productos de la industria textil y de la moda como el color y el estilo, se han basado en modelos estadísticos tradicionales como los modelos de pronóstico bayesianos, Auto-Regresivos (AR), de Medias Móviles (MA) y Auto-Regresivos Integrados de Media Móvil (ARIMA), que hoy en día siguen siendo vigentes para el pronóstico de series temporales. Sin embargo, estos modelos se basan en algunos supuestos estadísticos como la estacionariedad, la estacionalidad y la linealidad del sistema. Por lo tanto, estos solo son útiles para analizar series temporales cuando dichas asunciones son válidas. En la práctica la estacionariedad, estacionalidad y la linealidad de la serie de estudio puede no ser evidente, por lo tanto los modelos tradicionales para series temporales deben ser implementados con mucho cuidado en aplicaciones reales (Choi, et al., 2011; Sun, et al., 2008; Yu, et al., 2011; y Choi, et al., 2011).

En Coats Cadena Andina S.A. durante los últimos años se ha identificado que el tamaño de lote promedio demandado por color ha disminuido en más de un 31% (de 25 kg. a 19 kg.). Actualmente el 52% de los lotes tienen un tamaño promedio menor a 3,6 kg. Este comportamiento de la demanda hace que el volumen de los lotes se concentre en un grupo de máquinas teñidoras, lo cual implica que la capacidad instalada para el volumen de lotes no

alcance a cubrir las necesidades del mercado en el tiempo requerido, obligando a la planta a producir los días domingos y festivos para mantener el servicio, sacrificando con esto el costo de la manufactura del proceso. En un periodo de casi un año la tintorería ha reproducido alrededor de 2.936 colores, de los cuales solo 761 colores representan el 80% de la producción por color. Por lo tanto, en este trabajo se pretende construir una metodología de pronóstico con el fin de determinar cuál es la dinámica de las series de la demanda de los colores y obtener el pronóstico de la siguiente observación. Para esto se espera evaluar el desempeño de diversos modelos, ya que el rendimiento de estos depende de las características y la cantidad de datos (Choi, et al., 2011; Yu, et al., 2012; Chen, et al., 2009).

## **2. Justificación**

Según Valenzuela (2012), el resultado de un hecho futuro, que involucre incertidumbre, siempre ha sido un tema de interés para los sectores económicos e industriales, ya sea para tomar decisiones, diseñar estrategias de negocios o bien para anticipar como responderán los diferentes mercados ante la influencia de ciertos fenómenos. Si bien los pronósticos podrían ser efectuados apoyándose en la intuición, la importancia que han tomado en diversas áreas productivas ha generado la necesidad de desarrollar herramientas matemáticas que sean capaces de realizar esta tarea de forma automática y que permitan procesar la información suministrada por el sistema hasta el momento de la predicción (Choi, et al., 2011).

### **2.1. Pertinencia**

Debido a la naturaleza de las series temporales de demanda de color se han realizado diferentes estudios con modelos de regresión, sin embargo, estos modelos dependen de la cantidad de datos para funcionar bien. Por lo tanto, se evidencia la necesidad de aplicar o integrar otros métodos que puedan mostrar mayor realidad en los pronósticos, cuando el número de datos históricos es limitado (Choi, et al., 2011). Adicionalmente, en Coats Cadena Andina S.A., se ha observado la necesidad de realizar un estudio que permita definir cuáles son los colores que en el futuro tendrán mayor aceptación entre los consumidores y cuáles no, para de esta forma plantear la posibilidad de incluirlos o sacarlos de la carta de colores.

## **2.2. Impacto**

Para la empresa Coats Cadena Andina S.A., realizar pronósticos acertados en cuanto al color de moda permitirá corroborar si hay tendencia a incrementarse la producción de ciertos colores y si éstos coinciden con la carta de los colores destacados de la temporada suministrada por INEXMODA para cada semestre del año, y con la carta de colores ofertada por la compañía, para de esta forma crear estrategias de mercadeo y de promoción que permitan estimular las ventas.

## **2.3. Viabilidad y revisión bibliográfica<sup>1</sup>**

La metodología actual para analizar series temporales es la confluencia de varias líneas de trabajo desarrolladas en distintos campos científicos. En el caso de las series temporales pueden identificarse cinco campos de trabajos principales.

El primero tiene sus raíces en los estudios de series astronómicas y climáticas, que dio lugar a la teoría de los procesos estocásticos estacionarios, desarrolladas por matemáticos en la primera mitad del siglo XX. El segundo, es el desarrollo de los métodos de alisado, inventados por investigadores operativos para prever series de producción y ventas en la década de 1960-70, aprovechando las facilidades de cálculo aportadas por los primeros ordenadores. El tercero, la teoría de predicción y control de sistemas lineales, desarrollada en ingeniería de control y automática en los años 70, y estimulada por el desarrollo de la ingeniería aeronáutica y espacial. El cuarto, es la teoría de procesos no estacionarios y no lineales, desarrollada por estadistas y econométricos en los últimos veinte años del siglo XX. Finalmente, el último campo son los modelos multivariantes y los métodos de reducción de la dimensión en sistemas dinámicos, que se encuentra todavía en la fase de desarrollo.

---

<sup>1</sup> Los supuestos presentados en este apartado se corresponden con lo descrito por Peña (2010), en su libro *Análisis de series temporales*. Madrid, Alianza.

De esta forma, los métodos disponibles en la actualidad para el análisis de las series temporales son deudores de las investigaciones de matemáticos, estadísticos, ingenieros, físicos y economistas durante el siglo XX para resolver problemas de predicción y control de variables y sistemas dinámicos.

Las primeras series temporales estudiadas correspondían a datos astronómicos y meteorológicos. El matemático Laplace, analizó series temporales de la presión barométrica en París, ajustándolas a una función sinusoidal, pero sus resultados no fueron correctos porque no tuvo en cuenta la dependencia temporal de las observaciones. Después de la invención del método de mínimos cuadrados por Gauss y Legendre a principios del siglo XIX, se realizan diversos ajustes de funciones sinusoidales a datos astronómicos para detectar las posibles periodicidades. El estadístico británico G. U. Yule, introductor del coeficiente de correlación múltiple y de los coeficientes de autocorrelación parcial, extendió estas ideas a las series temporales, proponiendo los procesos autoregresivos en 1927 para explicar la serie de manchas solares. Ese mismo año, el estadístico ruso E. E. Slutsky descubrió que al tomar promedios móviles en una serie se generan artificialmente periodicidades y estudió los procesos de media móvil para representar los ciclos económicos.

Estos avances llevan a la formalización del concepto de serie temporal como proceso estocástico estacionario, debido a A. N. Kolmogorov, y a Markov. La generalización del análisis de Fourier para representar procesos estacionarios es debida al matemático sueco Cramer. Los trabajos de Bartlett y de John Tukey, sentaron las bases del análisis espectral moderno. Otras contribuciones importantes en el campo espectral son debidas a H. Akaike, que propuso la estimación del espectro ajustando modelos AR largos, y a P. Whittle, que introduce los procedimientos de estimación de mínimos cuadrados basados en el espectro.

La aparición del computador impulsa el desarrollo de métodos de predicción para series reales no estacionarias basados en métodos heurísticos. Holt y Winters, introducen los métodos de alisado exponencial, que tienen mucha repercusión en las aplicaciones prácticas de la predicción. La unión de los métodos de alisado y los procesos estocásticos se establece con el trabajo de Muth, que demostró que los métodos de alisado eran óptimos para un tipo especial de proceso ARIMA, y, sobre todo, por las investigaciones en los años 60 de G. E. P. Box y G. Jenkins, que estudian el problema de la predicción y control de series temporales industriales.

Paralelamente al trabajo estadístico de Box y Jenkins en los años 60, R. Kalman y Busy desarrollan una metodología para estimar las variables de estado y prever las observaciones futuras en sistemas lineales (filtro Kalman). Este algoritmo se introduce en la estadística gracias a los trabajos de Harrison y Stevens, que presentan una forma alternativa de modelar las series temporales mediante la formulación en el espacio de los estados y el enfoque Bayesiano. Estos modelos aparecieron como una forma de modelar series temporales contrapuesta a los modelos ARIMA propuestos por Box y Jenkins, pero en la actualidad se reconoce el carácter complementario de ambas formulaciones. El ajuste de modelos autorregresivos a series reales planteó el problema de cómo seleccionar el orden del modelo, donde H. Akaike, proporciona una solución utilizando conceptos de teoría de la información (criterio AIC).

La teoría de procesos se desarrolla por J. Dickey, quién plantea los procesos integrados de primer orden y propone un contraste de raíces unitarias. La teoría de procesos integrados se ha ido completando en la última década del siglo XX, tanto desde el campo de la estadística como desde la econometría. El estudio de los procesos no estacionarios vectoriales lleva al concepto de cointegración, que se define como la presencia de combinaciones lineales de series no estacionarias que forman un proceso estacionario. Esta posibilidad fue estudiada inicialmente por Box y Tiao, y establecida por Granger y Engle en , donde R. Engle introduce en 1982 el modelo ARCH, cuya varianza futura varía en función de los valores pasados de la serie y que ha sido generalizado por numerosos autores.

Otra línea importante en los años 80 y 90 es el estudio de los modelos no lineales, donde se destaca el trabajo de H. Tong, con los autorregresivos por umbrales, Priestley con sus modelos generales, Granger et al., con los bilineales y Mandelbrot con los modelos de memoria larga introducidos en los años 60 para datos hidrológicos. Un caso especial de no linealidad es la presencia de valores atípicos, donde Fox introdujo dos tipos básicos de atípicos en series estacionarias, y G. Tiao y R. Tsay han propuesto otras clases de atípicos y diseño de métodos eficientes para modelar series con estos efectos.

El estudio de las relaciones dinámicas entre series se ha realizado con cierto retraso respecto a los análisis multivariantes. En los años 90, la introducción de retardos en las mismas relaciones dinámicas entre dos series es debida a R. A. Fisher, que estudió la dependencia de la producción del trigo y la lluvia en la estación agraria de Rothamsted, en Inglaterra. Paralelamente, Irving

Fisher introdujo los retardos para describir las relaciones entre variables económicas. La generalización de estas ideas para construir funciones polinómicas de retardos que pueden estimarse por mínimos cuadrados es debida a Shirley Almon. Un trabajo pionero sobre las series temporales multivariantes fue debido a Quenouille, que desarrolló métodos muy avanzados para su época para el estudio de series multivariantes. Hannan estudia las series temporales multivariantes y Hannan y Deistler en el espacio de los estados. Zellner y Palm relacionan la teoría de modelos ARMA vectoriales con los modelos econométricos dinámicos desarrollados años antes. Tiao y Box proponen un método general para construir modelos ARIMA.

Concatenando los avances desarrollados a la fecha para realizar el análisis de series temporales con el requerimiento puntual para la predicción del color se tiene<sup>2</sup>:

La predicción del color de moda ha sido un producto significativo en las últimas décadas. Harold Linton, profesor y director del departamento de arte y tecnología visual de la Universidad de George Mason, contribuyó principalmente a la demostración sistemática de la predicción del color y sus aplicaciones. En su libro, *Color Forecasting: A Survey of International Color Marketing* (1994), se describe el misterio y la tecnología para el pronóstico del color, donde se desatan los enigmas que preocuparon durante mucho tiempo a los clientes y a las compañías.

Este documento tuvo incidencias considerablemente importantes en posteriores investigaciones. Deng Qingzeng, profesor de la Universidad de Beijing, fue uno de los primeros investigadores, en indagar mediante el análisis estadístico en el tema de la predicción de la tendencia del color de moda. En su artículo: *Debate sobre la teoría estadística de color de moda* (1989), se considera el color de moda como un cambio de fase social determinado. Analizan estadísticamente los colores de moda mediante la ecuación de Fokker-Planck, llevada a cabo según el principio de sinergia. Finalmente muestran los resultados numéricos calculados y la explicación acerca de si un color será popular o no en el futuro.

En China, la CFCA ha hecho grandes contribuciones al desarrollo de la investigación en el tema del color y sus aplicaciones. Recientemente, la CFCA a través de Cui (2007), hizo un análisis de los colores de moda para las mujeres en la temporada primavera/verano, publicado

---

<sup>2</sup> Lo consignado en este apartado se toma del anteproyecto de grado de Agudelo, C (2012), *Desarrollo de un modelo para el pronóstico de la demanda por color en la empresa Coats Cadena Andina S.A.* (documento no publicado).

por IC de 1995 a 2005. Para este análisis se recogieron 359 muestras de demandas históricas de colores y se cuantificó mediante el sistema NCS, la distribución de los colores en términos del tono, el valor y la saturación. Se analizó el cambio en los tonos de los colores y finalmente se mostraron las tendencias futuras en cuanto al color. Esta fue la primera contribución basada en información histórica de una agencia de pronósticos.

En los últimos años diversos modelos de series temporales han ganado fuerza e interés para el pronóstico de la tendencia de la demanda o las ventas de productos relacionados con la industria textil y de la moda. Es así como: Choi, et al. (2011), consideran un modelo híbrido ARIMA estacional con la transformada Wavelet, para pronosticar las ventas diarias de un producto de moda, y comparan los resultados con los de un método clásico de descomposición temporal con extrapolación lineal y con ajuste estacional (CDS+ LESA, por sus siglas en inglés), y concluyen que el modelo CDS + LESA genera buenos resultados cuando el ruido es un factor dominante en la serie temporal, es decir, cuando la varianza del ruido es alta, pero éste no proporciona buenos resultados en comparación con el modelo ARIMA híbrido cuando la componente de ciclo estacional es significativa, que según varios estudios es el caso de las series de demanda de productos relacionados con la industria de la moda.

Choi, et al. (2011), estudian varios modelos de pronóstico para estimar las ventas de diferentes colores con una cantidad limitada de datos. De esta forma utilizan para el análisis datos históricos de cuatro años de ventas de diferentes colores de una empresa de prendas de vestir.

Entre los diferentes modelos que analizaron, concluyen que el modelo basado en la teoría Grey (GM) y el modelo de Redes Neuronales Artificiales (ANN) proporcionan los mejores resultados utilizando como referencia el error porcentual medio absoluto (MAPE). Sin embargo resaltan, que el modelo GM requiere una menor cantidad de datos históricos que el modelo de ANN para generar resultados razonables.

Adicionalmente, construyen un modelo híbrido GM + ANN mejorando con este el rendimiento en comparación con cada uno de los modelos por separado.





## **3. Objetivos**

### **3.1. Objetivo general**

Desarrollar una metodología para el pronóstico de la demanda por color en la empresa Coats Cadena Andina S.A. con el fin de determinar cuál es la dinámica de las series de la demanda de un determinado color y obtener el pronóstico de la siguiente observación.

### **3.2. Objetivos específicos**

- Construir los principales modelos utilizados en el pronóstico de la demanda del color.
- Estimar los errores en la muestra y en los pronósticos (porcentaje de error medio absoluto (MAPE) y error medio cuadrático (MSE)) de cada uno de los modelos construidos con el fin de evaluar el desempeño de estos.
- Realizar la validación del modelo a partir de la metodología de los datos sustitutos.

## **4. Marco Conceptual**

### **4.1. Series de tiempo de la demanda del color**

Las series de tiempo de la demanda del color son una secuencia cronológica tomada normalmente con una frecuencia diaria (lunes a sábado) a través del *call center express* del proceso de Servicio al Cliente de Coats Cadena Andina S.A.

Los clientes pueden realizar sus pedidos de artículos existentes que se encuentren en la carta de color o no, por medio de llamadas, correos electrónicos y faxes, los cuales son atendidos por agentes capacitados en esta labor.

Cuando el color deseado no se encuentre dentro de los ofertados por Coats Cadena Andina S.A. (colores de moda), los clientes pueden enviar a través del asesor comercial la muestra del color, que normalmente es una tela, para que en el laboratorio de la Tintorería sea desarrollado el color requerido por el cliente y de este modo satisfacer la demanda del color de las nuevas colecciones.

Esta dinámica del mercado de la moda textil ha influenciado el aumento del número de los lotes de teñido, pero ha significado la disminución del tamaño en kg de éste, lo cual es el reflejo del crecimiento de la demanda de los colores especiales para cada temporada en particular.

Tradicionalmente el estudio de las tendencias del color ha sido un tema de juicio humano, el cual se construye a través del contacto con los clientes, la asistencia a ferias de moda y el asesoramiento de instituciones expertas como InexModa, que suministran información acerca de las tendencias en color de las siguientes temporadas. Las decisiones que se tomen pueden tener un sesgo frente a la realidad del comportamiento futuro de las series de la demanda de color.

La anterior situación coincide con la metodología sistémica, que ha sido el método típico implementado en el estudio del pronóstico del color en la moda y en la industria textil y se basa en un análisis subjetivo y cualitativo que ha sido ampliamente usado por las autoridades del color a nivel mundial. Los modelos de predicción de la metodología sistémica tienen en cuenta variables como la moda, el estilo de vida social, la economía global, los colores preferidos de los clientes, y otras diferentes variables socioeconómicas que hacen que normalmente la predicción del color dependa más del experto en moda que de las percepciones del pronosticador.

En la medida en que la demanda es más variable y los clientes esperan respuesta en un determinado tiempo, el manejo de la información se vuelve una herramienta esencial, y es aquí donde entra a desempeñar un papel clave el uso de la metodología de análisis de datos históricos, que tiene en cuenta métodos cuantitativos con el fin de investigar las posibles reglas de cambio en las tendencias de la moda del color. Esto ha llevado a los investigadores a examinar diferentes modelos de predicción que van desde la utilización de modelos estadísticos básicos hasta modelos de inteligencia artificial, que han surgido a partir del avance de las herramientas computacionales como las redes neuronales artificiales o los modelos que trabajan a partir de pocos datos históricos..

## **4.2. Introducción a los Sistemas dinámicos<sup>3</sup>**

### **4.2.1. *La dinámica de los sistemas***

En la dinámica es destacable la existencia de dos tipos de sistemas, las ecuaciones diferenciales y los mapas iterados (también conocidos como ecuaciones en diferencia). Las ecuaciones diferenciales describen la evolución de los sistemas en un tiempo continuo y se utilizan ampliamente en la ciencia y la ingeniería, mientras que los mapas iterados surgen en los

---

<sup>3</sup> Los supuestos que se presentan en este capítulo corresponden al documento de Ospina G., L. M.; Soto M., J. A. & Orozco G., A. A. (2011). *Determinismo y no linealidad en señales financieras: Método de los datos sustitutos*. Pereira, Universidad Tecnológica de Pereira.

problemas donde el tiempo es discreto. Por conveniencia se considera únicamente sistemas dinámicos discretos, dada la restricción tecnológica actual (Strogatz, 1994).

Así un sistema dinámico puede ser entendido como la representación matemática de un proceso que cambia en el tiempo atendiendo a una cierta ley, que determina la evolución temporal del estado del sistema en función de su historia previa. En particular un sistema dinámico discreto viene dado por  $x_t = f(x_{t-1})$  donde  $t = 1, 2, \dots, n$  representa el tiempo, medido a intervalos regulares y  $x_t$  el estado del sistema en el instante  $t$ , determinado por el estado en el instante anterior<sup>4</sup>. Así pues, el conocimiento del estado inicial  $x_0$  basta para reconstruir toda la evolución del sistema, lo que confiere al sistema *carácter determinista*.

En un sistema aleatorio o con incertidumbre,  $x_t$  no se conoce exactamente a partir de  $x_{t-1}$ , en el mejor de los casos se conoce la probabilidad  $p(z/x_{t-1})$  de que  $x_t$  tome un valor  $z$  supuesto conocido  $x_{t-1}$  (Mantilla y Vásquez, 2000).

Cuando la función  $f$  que define el sistema es lineal, siempre es posible obtener la expresión de la órbita (sucesión infinita de puntos). Sin embargo esto no sucede en *general* para sistemas no lineales; en esta situación se debe abandonar la idea de calcular su solución para afrontar la cuestión de cómo será su comportamiento cualitativo futuro.

La solución de los sistemas lineales es sencilla de lograr, estos pueden dividirse en partes, luego cada parte es resuelta por separado y re combinada para obtener la respuesta. Esto permite una simplificación de problemas complejos y es la base de diferentes métodos tales como la transformada de Laplace, los argumentos de superposición y el análisis de Fourier. De lo anterior se concluye que la solución de sistema lineal equivale a la suma de sus partes.

Sin embargo, la vida cotidiana no es lineal, y el principio de superposición falla.

Por lo tanto, los sistemas no lineales que en su mayoría son imposibles de resolver analíticamente, en estos casos un enfoque geométrico resulta ser muy poderoso y brinda la mayor parte de la información que se quiere, a pesar de que normalmente no pueda solucionar las ecuaciones en el sentido tradicional de encontrar una fórmula para la respuesta.

---

<sup>4</sup> Esta es la expresión de un sistema dinámico, discreto, autónomo, unidimensional y de primer orden.

La solución gráfica de un sistema corresponde a un punto en movimiento a lo largo de un espacio, donde la línea trazada se denomina *trayectoria* y el espacio recibe el nombre de *espacio de fases* para el sistema. El espacio de fase está completamente lleno de trayectorias, ya que cada punto puede servir como una condición inicial.

En la mayoría de las ocasiones no se tiene un conocimiento previo de la dinámica del fenómeno ni de las ecuaciones que gobiernan su comportamiento, solo se conoce un grupo de mediciones experimentales que se denomina serie temporal. A partir de este conjunto de datos se desea saber si la dinámica del sistema es estocástica, lineal, no lineal.

A continuación se describe una técnica para distinguir entre fluctuaciones debidas al azar y aquellas que exhiben una regularidad si se les examina en forma adecuada.

## 4.2.2. *Tratamiento de Series Temporales Lineales*

Para el análisis de series de tiempo lineales, es preciso considerar:

### 4.2.2.1. **Estacionariedad**

Según Priestley (1991), los términos exponenciales usualmente llevan a no estacionariedad. En esta situación algunas características de las series temporales como la media (ecuación 5.1), y la varianza (ecuación 5.2), cambian en el tiempo:

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n, \quad (5.1)$$

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2} \quad (5.2)$$

La desviación estándar es la raíz de la varianza y mide el ancho de la distribución. Una prueba sencilla es calcular estas cantidades para la primera mitad ( $1 \leq n \leq N/2$ ) y la segunda mitad ( $N/2 + 1 \leq n \leq N$ ) en los datos. Si la media de las dos partes difiere por más que un error estándar para cada mitad, entonces es casi seguro que la estacionariedad será un problema. El error estándar es la desviación estándar dividida por la raíz cuadrada del número de datos  $\sigma/\sqrt{n}$ .

Dado que la desviación estándar involucra el segundo momento de la distribución<sup>5</sup>, es sensible a puntos lejanos de la media y en algunos casos puede que no converja a medida que se incrementa el número de puntos. Por lo tanto, una medida más robusta es la desviación promedio, dada por,

$$desviacionpromedio = \frac{1}{N} \sum_{n=1}^N |x_n - \bar{x}|, \quad (5.3)$$

La cual suministra casi la misma información que la desviación estándar, siendo más sensible a los puntos lejanos. Se dice que una serie temporal cuyos dos primeros momentos (media y varianza) son constantes exhibe una estacionariedad débil, pero esta condición es insuficiente para analizar un sistema caótico. Es posible examinar momentos superiores de los datos, pero estos tienden a ser menos robustos, y su dependencia temporal aún no garantiza la estacionariedad para un sistema caótico. Un tercer momento es el denominado asimetría, donde:

$$asimetría = \frac{1}{N} \sum_{n=1}^N \left[ \frac{x_n - \bar{x}}{\sigma} \right]^3, \quad (5.4)$$

Ésta es una medida de la simetría de los datos alrededor de la media. El cuarto momento es el conocido como curtosis:

$$curtosis = \frac{1}{N} \sum_{n=1}^N \left[ \frac{x_n - \bar{x}}{\sigma} \right]^4 - 3 \quad (5.5)$$

Que compara el pico de la distribución con respecto al de una distribución Gaussiana. Si la serie de tiempo es fractal, los momentos de la distribución puede que no existan en el sentido en que sus valores calculados dependen del tamaño de los datos adquiridos y pueden tender a infinito o a cero a medida que el número de datos se incrementa.

#### 4.2.2.2. Función de Autocorrelación

La transformada de Fourier del espectro de potencias en el dominio del tiempo, de acuerdo al teorema de Wiener-Khinchin está dada por la función de autocorrelación:

$$G(k) \simeq \frac{\sum_{n=1}^{N-k} (x_n - \bar{x})(x_{n+k} - \bar{x})}{\sum_{n=1}^{N-k} (x_n - \bar{x})^2}, \quad (5.6)$$

---

<sup>5</sup> Debido a  $((x_n - \bar{x})^2)$  en la ecuación de la varianza.

La cual mide cuan fuerte en promedio cada dato está relacionado con otro  $k$  espacios de tiempo después. En general, la función de autocorrelación pasa de 1 en  $k = 0$  a cero para  $k$  muy grande. El valor de  $k$  para el cual  $G(k)$  cae a  $1/e \approx 0,37$  es llamado tiempo de correlación  $\tau_c$ . Para  $X_n$  cuasi-periódica, la función de correlación será una oscilación decadente, en tal caso  $\tau_c$  es el tiempo para que la envolvente caiga a  $1/e$ . El espectro de potencia puede ser calculado de la autocorrelación por medio de:

$$S_m = G(0) + G(K) \cos\left(\frac{2\pi m K}{N}\right) + 2 \sum_{k=1}^{K-1} G(k) \cos\left(\frac{2\pi m k}{N}\right), \quad (5.7)$$

Donde  $K$  es el máximo valor de  $k$ , el cual según Davis (1986), se debe tomar cerca de  $N/4$ .

A partir del análisis anterior se puede concluir que el estudio lineal convencional es insuficiente para poder distinguir ruido de caos, por lo tanto es necesario introducir nuevas herramientas que permitan realizar un tratamiento eficiente de las series caóticas.

### **4.2.3. Tratamiento de Series Temporales No Lineales**

Los sistemas no lineales son sistemas que no responden a los estímulos en proporción directa. En ciertas circunstancias, los sistemas determinísticos no lineales, entran en un estado, que presentan una marcada sensibilidad a las condiciones iniciales, por lo que siendo sistemas determinísticos presentan comportamiento impredecible a largo plazo (Parker y Chua, 1987). Se ha venido observando la existencia de nuevas propiedades que aparecen en estos sistemas como resultado de la interacción entre sus partes y que no pueden explicarse a partir de las propiedades de sus elementos componentes (Solé, Bascompte, Delgado, Luque y Manrubia, 1996).

#### **4.2.3.1. Tiempo de retardo**

El estado de un sistema puede definirse como la información necesaria para predecir que hará el sistema un tiempo después. La idea es que la dinámica del sistema no se hace visible a partir de las observaciones de una sola variable, pues para esto se requiere un espacio multidimensional donde queden eliminadas todas las posibles auto-intersecciones del atractor.



Por lo tanto se debe hacer un gráfico retrasado  $y(t)$  a partir de las coordenadas retrasadas, es decir  $y(t) = [x(t), x(t - \tau_1), x(t - \tau_2), \dots]$  donde la elección  $\tau_i$ , se logra a partir de analizar cuando una función de correlación no lineal denominada información mutua promedio (Average Mutual Information, AMI) tiene su primer mínimo. Esta función da la cantidad de información acerca de cuan correlacionadas están las mediciones  $x(t - T\tau_i)$  a un tiempo  $\tau_T = T\tau_i$  respecto de la medición  $x(t + (T - 1)\tau_S)$  en un tiempo anterior. El múltiplo entero  $T$  del tiempo de muestreo es uno de los objetivos en el método.

Si los puntos elegidos  $[x(t - \tau_1), x(t - \tau_2), \dots]$  están muy cercanos, entonces el atractor reconstruido no representará toda la dinámica subyacente, por el contrario si están muy separados entonces no existirá ninguna correlación entre los dos puntos. Una buena elección del tiempo de retardo es el primer mínimo de la información mutua promedio (AMI<sup>6</sup>).

El número de coordenadas del espacio de fases es una cantidad que se determina a partir del sistema que se desea analizar. Para un sistema general es un espacio de coordenadas  $x_1, x_2, \dots, x_m$ . Para un espacio  $m$ -dimensional, se hará referencia a un sistema de  $m$  dimensiones o sistema de orden  $m$ . Así,  $m$  representa la dimensión del espacio de fases.

Cuando la trayectoria  $x(t)$  es independiente del tiempo inicial se denomina sistema autónomo, de lo contrario es un sistema no autónomo. Para eliminar cualquier dependencia con el tiempo, se adiciona una dimensión al sistema, lo que permite visualizar un espacio de fases con trayectorias congeladas en él. De lo contrario, si se admite la dependencia temporal, las trayectorias estarán en continuo movimiento desmantelando la imagen geométrica que se está tratando de construir (Strogatz, 1994).

La dimensión del espacio de fases debe ser la necesaria para que la estructura geométrica proyectada o atractor, sea completamente revelada. Es decir, concluir si los puntos cercanos en el espacio de fases se encuentran allí debido a la dinámica del sistema o simplemente a causa de una inadecuada proyección, es decir, a una superposición de la trayectoria. Para esto se deben estudiar los puntos y su vecindad (cercanía entre ellos) en las diferentes dimensiones,  $d = 1, d = 2, \dots, d = n$ , hasta que se pueda establecer que la cercanía de los puntos de la estructura obtenida obedece a una razón dinámica y no a la de una proyección geométrica de un atractor de mayor

---

<sup>6</sup> Average Mutual Information, mayor ampliación en las invariantes del sistema.

dimensión en un espacio de fases de menor dimensión (Abarbanel, 1996). En síntesis, lo que se pretende es descubrir la geometría del atractor de forma que nada nuevo surja si se aumentan las dimensiones del espacio, quedando capturada las características dinámicas de los datos del desplazamiento en  $x$ .

Lo que asegura que se pueda ir de la observación de una variable, a un estado multivariado o espacio de fases, está contenido en un teorema geométrico denominado *teorema de embedding*.

#### 4.2.4. Características invariantes de la dinámica

Una invariante de la dinámica es una cantidad que describe el comportamiento dinámico de un sistema con la propiedad especial de que no depende del sistema de coordenadas. Es decir, el valor de dicha cantidad medida del sistema original (si fuera posible observarlo directamente) es igual al obtenido cuando es medida en el sistema reconstruido. La estimación de los invariantes es un campo de investigación abierto, pues continuamente se realizan investigaciones sobre métodos que puedan reducir el tiempo de cómputo y/o aumentar la precisión.

##### 4.2.4.1. Información Mutua Promedio (Average Mutual Information. AMI)

Una forma de evaluar la correlación o información entre mediciones está basada en la idea de Galler (1968), acerca de la información mutua entre dos mediciones  $a_i$  y  $b_j$ , de dos conjuntos  $A$  y  $B$  de posibles mediciones. La información mutua entre una medición  $a_i$  extraída de un conjunto  $A = a_i$  y una  $b_j$  extraída del conjunto  $B = b_j$  es la cantidad de información contenida en  $a_i$  acerca de la medición de  $b_j$ . Esto es:

$$I_{AB} = \log_2 \left[ \frac{P_{AB}(a_i, b_j)}{P_A(a_i)P_B(b_j)} \right] \quad (5.8)$$

Donde  $P_{AB}(a,b)$  es la densidad de probabilidad conjunta de que las mediciones de  $A$  y  $B$  resulten en valores  $a$  y  $b$ .  $P_A(a)$  y  $P_B(b)$  son las densidades de probabilidad individual para las mediciones de  $A$  y  $B$ . Si ejercer una medición sobre un valor de  $A$  cuyo resultado es  $a_i$  es

completamente independiente de haber ejercido una medición sobre un valor  $B$  que resultó en  $b_j$ , entonces la probabilidad conjunta  $PAB(a,b)$  se factoriza como  $PAB(a,b) = P_A(a)P_B(b)$  y la cantidad de información entre las dos mediciones (la información mutua) resulta ser cero. El promedio sobre todas las mediciones de esta información estadística, se la denomina información mutua promedio entre las mediciones  $A$  y las mediciones  $B$ , esto es:

$$I_{AB} = \sum_{a_i b_j} P_{AB}(a_i, b_j) \log_2 \left[ \frac{P_{AB}(a_i, b_j)}{P_A(a_i)P_B(b_j)} \right] \quad (5.9)$$

Esta cantidad no está relacionada con las reglas de evolución lineal o no-lineal de las cantidades medidas. Este es un producto netamente teórico que conecta dos conjuntos de mediciones entre ellas y establece un criterio para su dependencia mutua basado en la noción de la información compartida entre ellas. A partir de esto se puede decir que la medición  $s(t)$  en el momento  $t$  está conectada con la medición  $s(t+T)$  a un tiempo  $t+T$ . Tomando un conjunto de mediciones  $A$  como el grupo de valores observados  $s(n)$ , y un conjunto de mediciones  $B$  como los valores  $s(n+T)$ , la información mutua promedio entre estas dos mediciones es:

$$I(T) = \sum_{s(n), s(n+T)} P(s(n), s(n+T)) \log_2 \left[ \frac{P(s(n), s(n+T))}{P(s(n))P(s(n+T))} \right] \quad (5.10)$$

$I(T)$  es mayor o igual que cero. Cuando  $T$  se hace muy grande, el comportamiento de la serie hace que las mediciones  $s(n)$  y  $s(n+T)$  sean independientes en un sentido práctico, e  $I(T)$  tenderá a cero. Fraser y Swinney (1986), sugieren que la función  $I(T)$  puede pensarse como un tipo de función de autocorrelación no-lineal para determinar cuando los valores  $s(n)$  y  $s(n+T)$  son lo suficientemente independientes entre sí para ser usados como coordenadas de un vector de datos con retrasos temporales, pero no tan independientes como para no estar relacionados. La sugerencia para tomar el valor de retardo  $T$  en la reconstrucción del espacio de fases, es el valor de  $T$  para el cual ocurre el primer mínimo de  $I(T)$ .

#### 4.2.4.2. Complejidad relativa de Lempel y Ziv

La complejidad de Lempel-Ziv (1976), no es un invariante de la dinámica, pero en la práctica es bastante útil y por lo tanto es necesario considerarla.

Es una medida de la complejidad algorítmica de la serie temporal que se realiza sobre la serie temporal y no sobre el atractor. La complejidad de Lempel-Ziv o  $C(n)$  refleja el orden que se

retiene en un patrón temporal unidimensional o en una cadena de  $n$  símbolos (Lempel y Ziv, 1976; Lempel y Ziv, 1978; Kaspar y Schuster, 1997).

Para calcular la complejidad de Lempel-Ziv primero se codifica la serie temporal por medio de algún sistema de codificación, por simplicidad se puede considerar sólo cadena con ceros y unos (por ejemplo 01000110101). Para ello se convierten las series de tiempo originales en secuencias de 0 cuando la diferencia entre dos elementos sucesivos de la serie de tiempo es negativa y 1 cuando la diferencia entre los dos elementos sucesivos es positiva o nula. El algoritmo para estimar  $C(n)$  se basa en la reconstrucción de la cadena original  $S$  mediante operaciones de copia e inserción de símbolos en una nueva cadena. La  $C(n)$  aumenta conforme se necesiten nuevos símbolos o secuencias de símbolos (0,1,01,010, *etc.*) para reproducir la cadena  $S$ . Posteriormente se mide qué tan compleja es esta secuencia de símbolos, donde la complejidad máxima tiene un valor de 1 y una perfecta predictibilidad tiene un valor de 0. En este sentido el ruido tiene la máxima complejidad posible.

Para un alfabeto de  $d$  símbolos y una secuencia de tamaño  $N$  se ha demostrado que:

$$c < \frac{N}{\left(1 - 2^{\frac{(1 + \log_d \log_d(dN))}{\log_d(dN)}}\right) \log_d(dN)} \quad (5.11)$$

Dado que:

$$\left(1 - 2^{\frac{(1 + \log_d \log_d(dN))}{\log_d(dN)}}\right) \log_d(dN) \rightarrow 0 \quad (5.12)$$

Cuando  $N \rightarrow \infty$  generalmente se observa que  $c$  está limitada por  $N \log d / \log N$ . De hecho,  $N \log d / \log N$  es la complejidad de una secuencia aleatoria de tamaño  $N$  (con un alfabeto de  $d$  símbolos). Por lo tanto es más práctico considerar la complejidad normalizada:

$$C = \frac{c}{N} \log_d(N) \quad (5.13)$$

A la fecha, no se ha determinado como elegir de manera adecuada un alfabeto  $A$ , ni un esquema de codificación  $g$ . El método estándar es dividir los datos en  $d$ , usualmente de igual tamaño; para esto normalmente se elige  $d = 2$  (Lempel y Ziv, 1976).

### 4.3. Series temporales y los Procesos de media móvil y ARMA<sup>7</sup>

Los fenómenos dinámicos que se observan mediante series temporales que pueden clasificarse en dos clases:

- i) Procesos estacionarios: Son los que toman valores estables en el tiempo alrededor de un nivel constante, sin mostrar una tendencia a crecer o decrecer a largo plazo. Su estudio es útil en la práctica para representar la dependencia de los valores de una serie temporal
- ii) Procesos no estacionarios: Son aquellos que pueden mostrar tendencia, estacionalidad y otros efectos evolutivos en el tiempo.

Los modelos más simples son los autorregresivos, que generalizan la idea de la regresión para representar para representar la dependencia lineal entre dos variables aleatorias. El modelo de regresión simple explica la evolución de una variable,  $x_t$ , mediante la ecuación:

$$y_t = c + bx_t + a_t$$

Donde  $c$  y  $b$  son constantes a determinar y  $a$  es una variable aleatoria normal, con media nula y varianza constante. Si se aplica esta estructura de dependencia a las observaciones contiguas en una serie temporal,  $z_t$ , y se toma  $y_t = z_t$  y  $x_t = z_{t-1}$ , se obtiene el modelo de dependencia que se llama proceso autorregresivo de primer orden.

Los procesos autorregresivos son una clase de procesos estacionarios lineales que se caracterizan por tener muchos coeficientes de autocorrelación distintos de cero y que decrecen con el retardo. Estos procesos tienen memoria relativamente larga, ya que el valor actual está correlado con todos los anteriores, aunque con coeficientes decrecientes. Esta propiedad se traduce en que podemos escribir un proceso AR como una función lineal de todas las innovaciones que le han generado, con pesos que tienden a cero con el retardo. Los procesos AR no pueden representar series de memoria muy corta, donde el valor actual de la serie sólo

---

<sup>7</sup> Los supuestos sobre los que se basa este apartado son retomados de Peña (2010). *Análisis de series temporales*. Madrid, Alianza.

está correlado con un número pequeño de valores anteriores, de manera que la función de autocorrelación simple tenga sólo unas pocas autocorrelaciones distintas de cero. Una familia de procesos que tienen esta propiedad de «memoria muy corta», son los procesos de media móvil, o procesos MA (en inglés, *moving average*). Estos procesos de MA son función de un número finito, y generalmente pequeño, de las innovaciones pasadas.

Después se combinan las propiedades de los procesos AR y MA para definir los procesos ARMA, que dan lugar a una familia muy amplia y flexible de procesos estocásticos estacionarios útiles para representar muchas series temporales. Los procesos puros AR y MA aparecen, entonces, como casos particulares de esta representación general ARMA.

#### 4.3.1. *El proceso de media móvil de orden uno (MA(1))*

Se obtiene el proceso de media móvil de orden uno MA(1), añadiendo al proceso de ruido blanco una dependencia del valor actual de la serie de la última innovación ocurrida. De esta manera, el proceso será una combinación lineal de las dos últimas innovaciones, de acuerdo con la siguiente ecuación:

$$\tilde{z}_t = a_t - \theta a_{t-1}, \quad (6.1)$$

Donde  $\tilde{z} = z_t - \mu$ , siendo  $\mu$  la media del proceso y  $a_t$  es un proceso de ruido blanco con varianza  $\sigma^2$ . El proceso MA(1) puede escribirse con la notación de operadores:

$$\tilde{z}_t = (1 - \theta\beta)a_t. \quad (6.2)$$

Este proceso es la suma de los dos procesos estacionarios,  $a_t$  y  $-\theta a_{t-1}$  y, por tanto, siempre será estacionario, para cualquier valor del parámetro, a diferencia de los procesos AR.

En las aplicaciones de este proceso supondremos que  $|\theta| < 1$ , de manera que la innovación pasada tenga menos peso que la presente. Entonces, decimos que el proceso es *invertible* y tiene la propiedad de que el efecto de los valores pasados decrece con el tiempo. Para justificar esta propiedad, se sustituye  $a_{t-1}$  en (6.1) como función de  $z_{t-1}$ . El resultado es:

$$\tilde{z}_t = a_t - \theta(\tilde{z}_{t-1} + \theta a_{t-2}) = -\theta\tilde{z}_{t-1} - \theta^2 a_{t-2} + a_t$$

Y repitiendo esta operación para  $a_{t-2}$ :

$$\tilde{z}_t = -\theta\tilde{z}_{t-1} - \theta^2(\tilde{z}_{t-2} + \theta a_{t-3}) + a_t = -\theta\tilde{z}_{t-1} - \theta^2\tilde{z}_{t-2} - \theta^3 a_{t-3} + a_t$$

Mediante sustituciones sucesivas de  $a_{t-3}$ ,  $a_{t-4}$ , ..., etc., se obtiene:

$$\tilde{z}_t = -\sum_{i=1}^{t-1} \theta^i \tilde{z}_{t-i} - \theta^t a_0 + a_t, \quad (6.3)$$

Se observa que si  $|\theta| < 1$ , el efecto de  $\tilde{z}_{t-k}$  tiende hacia cero con  $k$  y el proceso se denomina invertible. Si  $|\theta| \geq 1$ , se daría la situación paradójica en que el efecto de las observaciones pasadas aumentaría con la distancia y, aunque el proceso seguiría siendo estacionario, parece poco adecuado para representar series reales. En adelante, se supondrá que el proceso es invertible. Entonces, como  $|\theta| < 1$ , existe el operador inverso  $(1 - \theta B)^{-1}$  y se puede escribir la ecuación (6.2) como:

$$(1 + \theta B + \theta^2 B^2 + \dots)\tilde{z}_t = a_t, \quad (6.4)$$

Que equivale a:

$$\tilde{z}_t = -\sum_{i=1}^{\infty} \theta^i \tilde{z}_{t-i} + a_t,$$

Que es la (6.3), suponiendo que el proceso comienza en el infinito pasado. Esta ecuación representa el proceso MA(1) invertible como un AR ( $\infty$ ) con coeficientes que decrecen en progresión geométrica.

La varianza del proceso se calcula a partir de (6.1). Elevando al cuadrado y tomando esperanzas, se obtiene:

$$E(\tilde{z}_t^2) = E(a_t^2) + \theta^2 E(a_{t-1}^2) - 2\theta E(a_t a_{t-1})$$

Como  $E(a_t a_{t-1}) = 0$ , por ser un proceso de ruido blanco, y  $E(a_t^2) = E(a_{t-1}^2) = \sigma^2$ , se tiene que:

$$\sigma_z^2 = \sigma^2(1 + \theta^2), \quad (6.5)$$

Esta ecuación nos indica que la varianza marginal del proceso,  $\sigma_z^2$ , es siempre mayor que la varianza de las innovaciones,  $\sigma^2$ , y tanto mayor cuanto mayor sea  $\theta^2$ .

#### 4.3.1.1. Función de autocorrelación simple y parcial

Para obtener la función de autocorrelación simple de un proceso MA(1) se comienza calculando las covarianzas. La de primer orden se obtiene multiplicando la ecuación (6.1) por  $\tilde{z}_{t-1}$  y tomando las esperanzas:

$$\gamma_1 = E(\tilde{z}_t \tilde{z}_{t-1}) = E(a_t \tilde{z}_{t-1}) - \theta E(a_{t-1} \tilde{z}_{t-1}).$$

En esta expresión el primer término  $E(a_t \tilde{z}_{t-1})$  es cero, ya que  $\tilde{z}_{t-1}$  según (6.1), depende de su innovación,  $a_{t-1}$  y de la innovación,  $a_{t-2}$ , pero no de las innovaciones futuras,  $a_t$ . Para calcular el segundo término, sustituyendo  $\tilde{z}_{t-1}$  por su expresión según (6.1), resulta:

$$E(a_{t-1} \tilde{z}_{t-1}) = E(a_{t-1}(a_{t-1} - \theta a_{t-2})) = \sigma^2,$$

Con lo que se obtiene:

$$\gamma_1 = -\theta \sigma^2.$$

De la misma forma, se calcula la autocovarianza de orden dos:

$$\gamma_2 = E(\tilde{z}_t \tilde{z}_{t-2}) = E(a_t \tilde{z}_{t-2}) - \theta E(a_{t-1} \tilde{z}_{t-2}) = 0,$$

Ya que al estar la serie incorrelada con sus innovaciones futuras, los dos términos son nulos. El mismo resultado se obtiene para las covarianzas de orden superior a dos, ya que al multiplicar la ecuación (6.1) por  $\tilde{z}_{t-j}$ , donde  $j > 1$ , tendremos productos de la serie por sus innovaciones futuras y las esperanzas de estos términos son cero. En conclusión:

$$\gamma_j = 0, \quad j > 1. \quad (6.7)$$

Dividiendo las autocovarianzas (6.6) y (6.7) por la expresión (6.5) de la varianza del proceso, se obtiene que los coeficientes de autocorrelación de un proceso MA(1) verifican:

$$\rho_1 = \frac{-\theta}{1+\theta^2}, \quad \rho_k = 0 \quad k > 1, \quad (6.8)$$



Y las *fas* (función de autocorrelación simple) tendrá únicamente un valor distinto de cero en el primer retardo. Como  $|\theta| < 1$ , el valor del coeficiente de autocorrelación en un MA(1) invertible es siempre menor a 0,5.

Este resultado demuestra que la *fas* de un proceso MA(1) tiene las mismas propiedades que la función de autocorrelación parcial (*fap*) de un proceso AR(1): existe un primer coeficiente distinto de cero y todos los demás son nulos. Esta dualidad entre AR(1) y MA(1) se presenta también en la *fap*. En efecto, según (6.4), al escribir un proceso MA(1) en forma autorregresiva hay un efecto directo de  $z_{t-k}$  sobre  $z_t$  de magnitud  $\theta^k$ , cualquiera que sea  $k$ . Por tanto, la *fap* tendrá todos los coeficientes no nulos y que decrecen geoméricamente con  $k$ . Ésta es la estructura de la *fas* de un AR(1) y, por tanto, se concluye que la *fap* de una MA(1) tiene la misma estructura que la *fas* de un AR (1).

### 4.3.2. El proceso MA( $q$ )

Generalizando la idea de un MA(1), se puede escribir procesos cuyo valor actual dependa no sólo de la última innovación, sino de las  $q$  últimas innovaciones. Se obtiene entonces el proceso MA( $q$ ), con representación general:

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}.$$

Introduciendo la notación de operadores:

$$\tilde{z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t. \quad (6.9)$$

Puede escribirse de forma más compacta como:

$$\tilde{z}_t = \theta_q(B) a_t. \quad (6.10)$$

Un proceso MA( $q$ ) es siempre estacionario, por ser la suma de procesos estacionarios. Se dice que el proceso es *invertible* si las raíces del operador  $\theta_q(B) = 0$  son, en módulo, mayores que la unidad.

Las propiedades de este proceso se obtienen con el mismo método utilizado para el MA(1). Multiplicando (6.9) por  $\tilde{z}_{t-k}$  para  $k \geq 0$  y tomando esperanzas, se obtienen las autocovarianzas:

$$\begin{aligned}\gamma_0 &= (1 + \theta_1^2 + \dots + \theta_q^2)\sigma^2 \\ \gamma_k &= (-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q)\sigma^2 \quad k=1, \dots, q, \\ \gamma_k &= 0 \quad k > q,\end{aligned}$$

Resultando que un proceso MA( $q$ ) tiene exactamente los  $q$  primeros coeficientes de la función de autocovarianzas distintos de cero. Dividiendo las covarianzas por  $\gamma_0$  y utilizando una notación más compacta, la función de autocorrelación será:

$$\begin{aligned}\rho_k &= \frac{\sum_{i=0}^{i=q-k} \theta_i \theta_{k+i}}{\sum_{i=0}^{i=q} \theta_i^2}, \quad k = 1, \dots, q \quad (6.11) \\ \rho_k &= 0, \quad k > q,\end{aligned}$$

Donde  $\theta_0 = -1$ ,  $\theta_k = 0$  para todo  $k \geq q + 1$ . En particular, para que  $q=k=1$  se obtiene la ecuación (6.8).

Para calcular la función de autocorrelación parcial de un MA( $q$ ) se expresa el proceso como un AR ( $\infty$ ):

$$\theta_q^{-1}(B)\tilde{z}_t = a_t,$$

Y llamando  $\theta_q^{-1}(B) = \pi(B)$ , donde:

$$\pi(B) = 1 - \pi_1 B - \dots - \pi_k B^k - \dots$$

Y los coeficientes de  $\pi(B)$  se obtienen imponiendo que  $\pi(B)\theta_q(B) = 1$ . Suponer que el proceso es invertible implica que las raíces de  $\theta_q(B) = 0$  están fuera del círculo de la unidad, y la serie  $\pi(B)$  será convergente. Igualando las potencias de  $B$  a cero, se obtiene que los coeficientes de  $\pi_i$  verifican la relación siguiente:

$$\pi_k = \theta_1 \pi_{k-1} + \dots + \theta_q \pi_{k-q}$$

Donde  $\pi_0 = -1$  y  $\pi_j = 0$  para  $j < 0$ . La solución de esta ecuación en diferencias será de la forma  $\sum A_i G_i^k$ , donde ahora  $G_i^{-1}$  son las raíces del operador de media móvil. Obtenidos los coeficientes  $\pi_i$  de la representación AR ( $\infty$ ) se puede escribir el proceso MA como:

$$\tilde{z}_t = \sum_{i=1}^{\infty} \pi_i \tilde{z}_{t-i} + a_t.$$

Esta expresión permite concluir que la *fap* de un MA será no nula para todo retardo, ya que existe un efecto directo de  $\tilde{z}_{t-i}$  sobre  $\tilde{z}_t$  para todo  $i$ . La *fap* de un proceso MA tendrá pues la misma estructura que la *fas* de un proceso AR del mismo orden.

Se concluye que existe una dualidad entre procesos AR y MA, de manera que la *fap* de un MA( $q$ ) tiene la estructura de la *fas* de un AR( $q$ ), y la *fas* de un MA( $q$ ) tiene la estructura de la *fap* de un AR( $q$ ).

### 4.3.3. El proceso MA( $\infty$ ). La descomposición de Wold

Los procesos autorregresivos y de media móvil estudiados son casos particulares de una representación general de procesos estacionarios obtenida por Wold (1938). Este autor demostró que todo proceso estocástico débilmente estacionario,  $z_t$ , de media finita,  $\mu$ , que no contenga componentes deterministas, puede escribirse como una función lineal de variables aleatoria incorreladas,  $a_t$ , como:

$$z_t = \mu + a_t + \varphi_1 a_{t-1} + \varphi_2 a_{t-2} + \dots = \mu + \sum_{i=0}^{\infty} \varphi_i a_{t-i}, \quad (\varphi_0 = 1) \quad (6.12)$$

Donde  $E(z_t) = \mu$ , y  $E(a_t) = 0$ ;  $Var(a_t) = \sigma^2$ ;  $E(a_t a_{t-k}) = 0, k > 1$ . Se supone que  $\sum_{i=0}^{\infty} \varphi_i^2 < \infty$  para que la serie esté bien definida. Llamando  $\tilde{z}_t = z_t - \mu$ , y utilizando el operador de retardo, se puede escribir:

$$\tilde{z}_t = \varphi(B) a_t, \quad (6.13)$$

Siendo  $\varphi(B) = 1 + \varphi_1 B + \varphi_2 B^2 + \dots$  un polinomio indefinido en el operador de retardo B. Se llama a (6.13) la representación lineal general de un proceso estacionario no determinista.

Esta representación es importante, porque garantiza que cualquier proceso estacionario admite una representación lineal. En general, las variables  $a_t$ , forman un proceso de ruido blanco, es decir, están incorreladas con media cero y varianza constante. En ciertos casos particulares el proceso puede escribirse en función de las variables  $\{a_t\}$  normales independientes. Entonces la variable  $\tilde{z}_t$  tendrá una distribución normal, y la estacionaridad débil coincide con la estricta.

La serie  $\tilde{z}_t$ , puede considerarse el resultado de pasar un proceso de impulsos  $\{a_t\}$  de variables incorreladas a través del filtro lineal  $\varphi(B)$  que determina la ponderación de cada «impulso» en la respuesta.

Las propiedades del proceso se obtienen como en los casos de un modelo MA. La varianza de  $\tilde{z}_t$  en (6.12) será:

$$\text{Var}(z_t) = \gamma_0 = \sigma^2 \sum_{i=0}^{\infty} \varphi_i^2, \quad (6.14)$$

Y para que el proceso tenga varianza finita, la serie  $\{\varphi_i^2\}$  debe ser convergente. Se observa que si los coeficientes  $\varphi_i$  se anulan a partir de un retardo  $q$ , que el modelo general se reduce a un MA( $q$ ) y la fórmula (6.14) coincide con la obtenida en el apartado anterior. Las covarianzas se obtienen con:

$$\gamma_k = E(\tilde{z}_t \tilde{z}_{t-k}) = \sigma^2 \sum_{i=0}^{\infty} \varphi_i \varphi_{i+k},$$

Que para  $k=0$  proporcionan, como un caso particular, la fórmula (6.14) para la varianza. También, si los coeficientes  $\varphi_i$  se anulan a partir de un retardo  $q$ , esta expresión proporciona las autocovarianzas de un proceso MA( $q$ ). Los coeficientes de autocorrelación están dados por:

$$\rho_k = \frac{\sum_{i=0}^{\infty} \varphi_i \varphi_{i+k}}{\sum_{i=0}^{\infty} \varphi_i^2}, \quad (6.15)$$

Que generaliza la expresión (6.11) de las autocorrelaciones de un MA( $q$ ). Una consecuencia de (6.12) es que todo proceso estacionario admite también una representación autorregresiva, que puede ser de orden infinito. Esta representación es la inversa de la de Wold, y se escribe:

$$\tilde{z}_t = \pi_1 \tilde{z}_{t-1} + \pi_2 \tilde{z}_{t-2} + \dots + a_t,$$

Que, en notación de operadores, se reduce a:

$$\pi(B)\tilde{z}_t = a_t,$$

Donde:

$$\sum_{i=0}^{\infty} \pi_i^2 < \infty.$$

La representación de AR ( $\infty$ ) es el dual de la MA( $\infty$ ) y se verifica:

$$\pi(B)\varphi(B) = 1,$$

Con lo que, igualando las potencias de  $B$  a cero, se puede obtener los coeficientes de una representación a partir de los de la otra.

#### 4.3.3.1. Los modelos AR y MA y el proceso general

Es inmediato comprobar que un proceso MA es un caso particular de la representación de Wold, como también lo son los procesos AR. Por ejemplo, el proceso AR(1):

$$(1 - \phi B)\tilde{z}_t = a_t, \quad (6.16)$$

Puede escribirse, multiplicando por el operador inverso  $(1 - \phi B)^{-1}$ :

$$\tilde{z}_t = (1 + \phi B + \phi^2 B^2 + \dots) a_t,$$

Que representa al proceso AR(1) como un caso particular de la forma MA( $\infty$ ) del proceso lineal general, con coeficientes  $\varphi_i$  que decaen con progresión geométrica. La condición de estacionaridad y varianza finita, series de coeficientes  $\varphi_i^2$  convergente, equivalen ahora a que  $|\phi| < 1$ .

Para procesos AR de orden más alto resulta más cómodo obtener los coeficientes de la representación MA( $\infty$ ) imponiendo la condición que, al ser resultado de invertir un proceso AR, el producto de ambos operadores debe ser la unidad. Es así que para un AR(2) la condición será:

$$(1 - \phi_1 B - \phi_2 B^2)(1 + \varphi_1 B + \varphi_2 B^2 + \dots) = 1.$$

Imponiendo la anulación de potencias de  $B$ , se obtienen los coeficientes:

$$\varphi_1 = \phi_1$$

$$\varphi_2 = \phi_1\varphi_1 + \phi_2$$

$$\varphi_i = \phi_1\varphi_{i-1} + \phi_2\varphi_{i-2}, \quad i \geq 2$$

Donde  $\varphi_0 = 1$ . Análogamente, para un AR(p) los coeficientes  $\varphi_i$  de la representación general se calculan mediante:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 + \varphi_1 B + \varphi_2 B^2 + \dots) = 1$$

Y para  $i \geq p$  deben verificar la condición:

$$\varphi_i = \phi_1\varphi_{i-1} + \dots + \phi_p\varphi_{i-p}, \quad i \geq p.$$

La condición de estacionaridad implica que las raíces de la ecuación característica del proceso AR(p),  $\phi_p(B) = 0$ , deben estar fuera del círculo unidad. Escribiendo el operador  $\phi_p(B)$  como:

$$\phi_p(B) = \prod_{i=1}^p (1 - G_i B),$$

Donde  $G_i^{-1}$  son las raíces de  $\phi_p(B) = 0$ , se verifica que, desarrollando en fracciones parciales:

$$\phi_p^{-1}(B) = \sum \frac{k_i}{(1 - G_i B)},$$

Será convergente si:

$$|G_i| < 1.$$

En resumen, los procesos AR pueden considerarse casos particulares de la representación del proceso lineal general que se caracterizan porque: (1) todos los  $\varphi_i$  son distintos de cero; (2) existen restricciones sobre las  $\varphi_i$ , que dependen del orden del proceso. En general, verifican la secuencia  $\varphi_i = \phi_1\varphi_{i-1} + \dots + \phi_p\varphi_{i-p}$ , con condiciones iniciales que dependen del orden del proceso.

Esta relación entre los coeficientes  $\varphi_i$  y los del proceso permiten concluir que los coeficientes  $\varphi_i$  tienen la misma estructura que los coeficientes de autocorrelación simple. En efecto, para un  $AR(p)$  se ha visto que satisfacen la ecuación de un proceso, como ocurriría con las autocorrelaciones, para un  $MA(q)$  el número de coeficientes de autocorrelación no nulos es el orden del proceso.

#### 4.3.4. *El proceso ARMA (1,1)*

Una conclusión es que los procesos AR y MA aproximan un proceso lineal  $MA(\infty)$  desde puntos de vista complementarios: los AR permiten estructura  $MA(\infty)$ , pero imponen restricciones sobre las pautas de decrecimiento de los coeficientes  $\varphi_i$ ; mientras que los MA exigen un número de términos finito, pero a cambio no imponen restricciones sobre sus valores. Desde el punto de vista de la estructura de autocorrelación, los AR permiten muchos coeficientes distintos de cero, pero con una pauta de decrecimiento fija; mientras que los MA permiten unos pocos coeficientes distintos de cero, con valores arbitrarios. Los procesos ARMA intentan combinar estas propiedades y permiten representar de forma *escueta* (utilizando pocos parámetros) procesos cuyos primeros  $q$  coeficientes son cualesquiera, mientras que los siguientes decrecen según leyes simples.

Matemáticamente, los procesos ARMA resultan de añadir estructura MA a un proceso AR o viceversa. El proceso más simple es el  $ARMA(1,1)$ , que se escribe:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + a_t - \theta_1 a_{t-1},$$

O, con notación de operadores:

$$(1 - \phi_1 B)\tilde{z}_t = (1 - \theta_1 B)a_t, \quad (6.17)$$

Donde  $|\phi_1| < 1$  para que el proceso sea estacionario, y  $|\theta_1| < 1$  para que sea invertible. Además se supone que  $\phi_1 \neq \theta_1$ . Si ambos parámetros fuesen idénticos se tiene que, multiplicando ambos miembros por el operador  $(1 - \phi_1 B)^{-1}$ , resulta que  $\tilde{z}_t = a_t$ , y el proceso

sería ruido blanco. En la formulación de los modelos ARMA se supone siempre que no hay raíces comunes en los operadores AR y MA.

Para obtener la función de autocorrelación de un ARMA (1,1), multiplicando (6.17) por  $\tilde{z}_{t-k}$  y tomando esperanzas, se obtiene:

$$\gamma_k = \phi_1 \gamma_{k-1} + E(a_t \tilde{z}_{t-k}) - \theta_1 E(a_{t-1} \tilde{z}_{t-k}). \quad (6.18)$$

Para  $k > 1$ , el ruido de  $a_t$  está incorrelado con la historia de la serie. Para comprobarlo, basta expresar el proceso como un MA( $\infty$ ) invirtiendo la parte AR:

$$\tilde{z}_t = (1 - \phi_1 B)^{-1} (1 - \theta_1 B) a_t = a_t + (\phi_1 - \theta_1) a_{t-1} + \phi_1 (\phi_1 - \theta_1) a_{t-2} + \dots +$$

Y al ser  $a_t$  ruido blanco, estará incorrelado con los valores  $a_{t-i}$  y, por tanto, con  $\tilde{z}_{t-i}$ . En consecuencia:

$$\gamma_k = \phi_1 \gamma_{k-1}, \quad k > 1. \quad (6.19)$$

Para  $k = 0$ ,  $E[a_t \tilde{z}_t] = \sigma^2$ , y:

$$E[a_{t-1} \tilde{z}_t] = E[a_{t-1} (\phi_1 \tilde{z}_{t-1} + a_t - \theta_1 a_{t-1})] = \sigma^2 (\phi_1 - \theta_1),$$

Sustituyendo estos resultados en (6.18), para  $k = 0$ :

$$\gamma_0 = \phi_1 \gamma_1 + \sigma^2 - \theta_1 \sigma^2 (\phi_1 - \theta_1). \quad (6.20)$$

Tomando (6.16)  $k = 1$ , resulta que  $E[a_t \tilde{z}_{t-1}] = 0$ ,  $E[a_{t-1} \tilde{z}_{t-1}] = \sigma^2$  y:

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma^2, \quad (6.21)$$

Resolviendo (6.20) y (6.21) se obtiene:

$$\gamma_0 = \sigma^2 \frac{1 - 2\phi_1 \theta_1 + \theta_1^2}{1 - \theta_1^2}$$

Para obtener el primer coeficiente de autocorrelación, se divide (6.21) por esta expresión:

$$\rho_1 = \frac{(\phi_1 - \theta_1)(1 - \phi_1 \theta_1)}{1 - 2\phi_1 \theta_1 + \theta_1^2}. \quad (6.22)$$



Se observa que si  $\phi_1 = \theta_1$ , esta autocorrelación es cero, porque, como se indica antes, los operadores  $(1 - \phi_1 B)$  y  $(1 - \theta_1 B)$  se cancelan y resultará un proceso de ruido blanco. En el caso frecuente en que ambos coeficientes son positivos y  $\phi_1 > \theta_1$  es fácil comprobar que la correlación aumenta con  $(\phi_1 - \theta_1)$ .

El resto de los coeficientes de autocorrelación se obtienen dividiendo (6.19) por  $\gamma_0$ , que resulta en:

$$\rho_k = \phi_1 \rho_{k-1} \quad k > 1 \quad (6.23)$$

Que indica que a partir del primer coeficiente la *fas* de un ARMA(1,1) tiene un decrecimiento exponencial determinado por el parámetro  $\phi_1$  de la parte AR. La diferencia con AR(1) es que este decrecimiento comienza a partir de  $\rho_1$ , y no de  $\rho_0 = 1$ , y este primer valor de la autocorrelación de primer orden depende de la diferencia relativa entre  $\phi_1$  y  $\theta_1$ . Se observa que si  $\phi_1 \cong 1$  y  $\phi_1 - \theta_1 = \varepsilon$ , pequeño, se puede tener muchos coeficientes distintos de cero, pero todos de pequeño tamaño.

Para calcular la *fap*, se escribe el ARMA(1,1) en la forma AR( $\infty$ ):

$$(1 - \theta_1 B)^{-1}(1 - \phi_1 B)\tilde{z}_t = a_t,$$

Y utilizando que  $(1 - \theta_1 B)^{-1} = 1 - \theta_1 B + \theta_1^2 B^2 + \dots$ , y operando, se obtiene:

$$\tilde{z}_t = (\phi_1 - \theta_1)\tilde{z}_{t-1} + \theta_1(\phi_1 - \theta_1)\tilde{z}_{t-2} + \theta_1^2(\phi_1 - \theta_1)\tilde{z}_{t-3} + \dots + a_t.$$

El efecto directo de  $\tilde{z}_{t-k}$  sobre  $\tilde{z}_t$  decrece geoméricamente con  $\theta_1^k$  y la *fap* tendrá un decrecimiento geométrico a partir de un valor inicial.

En conclusión, en un proceso ARMA(1,1) la *fas* y la *fap* tienen una estructura similar: un valor inicial, cuya magnitud depende de  $\phi_1 - \theta_1$ , seguido de decrecimiento geométrico. La tasa de decrecimiento en la *fas* depende de  $\phi_1$ , mientras que en la *fap* depende de  $\theta_1$ .

#### 4.3.5. Procesos ARMA (p,q)

El proceso ARMA( $p, q$ ) será:

$$(1 - \phi_1 B - \dots - \phi_p B^p) \tilde{z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (6.24)$$

O, en notación compacta:

$$\phi_p(B) \tilde{z}_t = \theta_q(B) a_t.$$

El proceso será estacionario si las raíces de  $\phi_p(B) = 0$  están fuera del círculo unidad, e invertible si lo está las  $\theta_q(B) = 0$ . Se supone, además, que no hay raíces comunes que pueden cancelarse entre los operadores AR y MA.

Para obtener los coeficientes  $\varphi_i$  de la representación general del modelo MA( $\infty$ ) se escribe:

$$\tilde{z}_t = \phi_p(B)^{-1} \theta_q(B) a_t = \varphi(B) a_t$$

Y se igualan las potencias de B en  $\varphi(B)\phi_p(B)$  a las  $\theta_q(B)$ . Análogamente, se puede representar un ARMA( $p, q$ ) como un modelo AR( $\infty$ ) haciendo:

$$\theta_q^{-1}(B) \phi_p(B) \tilde{z}_t = \pi(B) \tilde{z}_t = a_t$$

Y los coeficientes  $\pi_i$  resultarán de  $\phi_p(B) = \theta_q(B)\pi(B)$ .

Para calcular las autocovarianzas, multiplicando (6.24) por  $\tilde{z}_{t-k}$  y tomando esperanzas:

$$\begin{aligned} \gamma_k - \phi_1 \gamma_{k-1} - \dots - \phi_p \gamma_{k-p} = \\ = E[a_t \tilde{z}_{t-k}] - \theta_1 E[a_{t-1} \tilde{z}_{t-k}] - \dots - \theta_p E[a_{t-p} \tilde{z}_{t-k}] \end{aligned}$$

Para  $k > q$  todos los términos de la derecha se anulan, y dividiendo por  $\gamma_0$ :

$$\rho_k - \phi_1 \rho_{k-1} - \dots - \phi_p \rho_{k-p} = 0,$$

Es decir:

$$\phi_p(B) \rho_k = 0 \quad k > q, \quad (6.25)$$

Y se concluye que los coeficientes de autocorrelación para  $k > q$  seguirán un decrecimiento determinado únicamente por la parte autorregresiva. Los primeros  $q$  coeficientes dependen de

los parámetros MA y AR y de ellos  $p$  proporcionarán los valores iniciales para el decrecimiento posterior (para  $k > q$ ) según (6.25). Por tanto, si  $p > q$  toda las *fas* mostrará un decrecimiento dictado por (6.25). En resumen, la *fas*:

1. Tendrá  $q - p + 1$  valores iniciales con una estructura que depende de los parámetros AR y MA;
2. Decrecerá a partir del coeficiente  $q - p$  como una mezcla de exponenciales y sinusoides, determinada exclusivamente por la parte autorregresiva.

## 4.4. Modelos ARMA<sup>8</sup>

### 4.4.1. *La función de verosimilitud de un proceso ARMA*

Se supone que tenemos un proceso ARMA y se desea estimar los parámetros por máxima verosimilitud. Para ello se debe escribir la función de densidad conjunta y maximizarla respecto a los parámetros, considerando a los datos como fijos. Para escribir la densidad conjunta de dos variables cualesquiera,  $\mathbf{x}, \mathbf{y}$ , escalares o vectoriales, puede siempre escribirse como el producto de dos distribuciones: la marginal de la primera variable y la distribución de la segunda condicionada a los valores de la primera. Es decir:

$$f(x, y) = f(x)f(x|y). \quad (7.1)$$

Esta expresión sigue siendo cierta si todas las funciones de densidad van condicionadas a otra variable  $\mathbf{z}$ , de manera que:

$$f(x, y | z) = f(x | z)f(y | x, z). \quad (7.2)$$

Consideraremos la función de densidad conjunta de las  $T$  observaciones  $\omega_t$ . Tomando en (7.1)  $x = \omega_1$  y  $y = (\omega_2, \dots, \omega_T)$ , se puede escribir:

---

<sup>8</sup> Los supuestos sobre los que se basa este apartado son retomados de Peña (2010). *Análisis de series temporales*. Madrid, Alianza.

$$f(\omega_T) = f(\omega_1)f(\omega_2, \dots, \omega_T | \omega_1)$$

Y descomponiendo el segundo término, con (7.2) haciendo  $z = \omega_1, x = \omega_2$  y  $y = (\omega_3, \dots, \omega_T)$ , resulta:

$$f(\omega_T) = f(\omega_1)f(\omega_2 | \omega_1)f(\omega_3, \dots, \omega_T | \omega_1, \omega_2)$$

Y repitiendo este proceso, se obtiene finalmente:

$$f(\omega_T) = f(\omega_1)f(\omega_2 | \omega_1)f(\omega_3 | \omega_2, \omega_1) \dots f(\omega_T | \omega_{T-1}, \dots, \omega_1). \quad (7.3)$$

Esta expresión permite escribir la función de densidad conjunta de las  $T$  variables como producto de  $T$  distribuciones univariantes. La diferencia entre esta representación y la que se obtiene con datos independientes es que en lugar de tener el producto de las marginales de cada dato tenemos la marginal de la primera y el producto de las condicionadas de cada dato, dados los anteriores.

La descomposición (7.3) permite escribir la verosimilitud de un modelo ARMA, ya que si se supone normalidad conjunta, todas las distribuciones condicionadas serán normales. Su esperanza, es la predicción a un paso, que minimiza el error cuadrático de predicción, y se escribe:

$$E(\omega_t | \omega_{t-1}, \dots, \omega_1) = \hat{\omega}_{t-1}(1) = \omega_t |_{t-1}$$

La notación  $\hat{\omega}_{t-1}(1)$  es conveniente cuando se piensa en predicciones a distintos horizontes, porque aparece claramente el origen de la predicción y el horizonte. Se consideran únicamente las predicciones a un paso, pues es más simple utilizar  $\omega_t |_{t-1}$  para la esperanza de la variable  $\omega_t$  cuando disponemos de los valores previos de una realización finita,  $\omega_{t-1}, \dots, \omega_1$  y conocemos los parámetros del proceso. Esto es, de nuevo, una simplificación de la notación  $e_{t-1}$ . Entonces,  $e_t = \omega_t - \omega_t |_{t-1}$ .

Estos errores de predicción están muy relacionados con las innovaciones del proceso, aunque no coinciden exactamente con ellas por un problema de valores iniciales. Para ver esta diferencia, se supone una serie,  $\omega_1, \dots, \omega_T$ , de media cero y tamaño  $T$ , generada por el proceso

AR(1) con parámetro  $\phi$  conocido. Entonces, como  $\omega_t|_{t-1} = \phi\omega_{t-1}$  para  $t = 2, \dots, T$ , pero  $E(\omega_1) = \phi E(\omega_0) = 0$ , se puede calcular los errores de predicción a un paso mediante:

$$e_1 = \omega_1 - \phi E(\omega_0) = \omega_1$$

$$e_2 = \omega_2 - \phi\omega_1$$

$$\vdots \quad \vdots \quad \vdots$$

$$e_T = \omega_T - \phi\omega_{T-1}$$

Se ve que los errores de predicción,  $e_2, \dots, e_T$ , coinciden con las innovaciones del modelo,  $a_2, \dots, a_T$ , donde  $\omega_t = \phi\omega_{t-1} + a_t$ , para  $t = 2, \dots, T$ . La diferencia aparece en el primer error de predicción a un paso,  $e_1$ , que no es igual a la innovación en ese punto,  $a_1 = \omega_1 - \phi\omega_0$ , sino a su estimación  $\omega_1$ . Esto hace que para  $t = 2, \dots, T$  la varianza de los errores de predicción a un paso sea  $\sigma^2$ , la innovación, mientras que para  $t = 1$  sea distinta. En general, se puede escribir:

$$Var(\omega_t|\omega_{t-1}, \dots, \omega_1) = \sigma^2 v_t|_{t-1}$$

Donde para AR(1)  $v_t|_{t-1} = 1$ , para  $t = 2, \dots, T$ , pero como  $Var(\omega_1) = E(\omega_1^2) = \sigma^2(1 - \phi^2)^{-1}$ . Se ha introducido la constante  $\sigma^2 = var(a_t)$  en la expresión de las varianzas condicionadas porque sale como un factor común y simplifica la presentación.

Con esta notación, la función de densidad conjunta de la muestra para un proceso ARMA general puede escribirse como:

$$f(\omega_T) = \prod_{t=1}^T \sigma^{-1} v_t|_{t-1}^{-1/2} (2\pi)^{-1/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^T \frac{\omega_t - \omega_t|_{t-1}}{v_t|_{t-1}} \right\}$$

Y tomando logaritmos, llamando  $\beta = (\mu, \phi_1, \dots, \theta_q, \sigma^2)$  al vector de parámetros, la función de verosimilitud o la función soporte es:

$$L(\beta) = -\frac{T}{2} \ln \sigma^2 - \frac{1}{2} \sum_{t=1}^T \ln v_t|_{t-1} - \frac{1}{2\sigma^2} \sum_{t=1}^T \frac{e_t^2}{v_t|_{t-1}} \quad (7.4)$$

Donde tanto las varianzas condicionales,  $v_{t|t-1}$ , como los errores de predicción a un paso,  $e_t$ , dependen de los parámetros. Por tanto, evaluar la función de verosimilitud se reduce a calcular los errores de predicción a un paso de cada observación, dadas las anteriores y sus varianzas.

La maximización de la función de verosimilitud exacta se efectúa con un algoritmo de optimización no lineal. Se utiliza mucho basado en el algoritmo de Gauss-Newton con una modificación debida a Marquardt.

## 4.4.2. Procesos AR

### 4.4.2.1. Procesos AR(1)

Como ilustración, se considera el proceso AR(1) de media  $\mu$ ,  $\omega_t = c + \phi\omega_{t-1} + a_t$ , con  $c = \mu(1 - \phi)$ . En este caso:

$$E(\omega_1) = \mu \quad (7.5)$$

Y

$$Var(\omega_1) = E(\omega_1 - \mu)^2 = \frac{\sigma^2}{1 - \phi^2}. \quad (7.6)$$

Entonces, con la notación anterior,  $\omega_{1|0} = \mu$  y  $v_{1|0} = (1 - \phi^2)^{-1}$ . Para  $\omega_2$  los momentos de la distribución condicionada a  $\omega_1$  son:  $\omega_{2|1} = E(\omega_2 | \omega_1) = c + \phi\omega_1$  y,  $Var(\omega_2 | \omega_1) = E[(\omega_2 - c - \phi\omega_1)^2] = E(a_2^2) = \sigma^2$

Con lo que  $v_{2|1} = 1$ . De la misma forma, se comprueba que:

$$\omega_{t|t-1} = E(\omega_t | \omega_{t-1}) = c + \phi\omega_{t-1}, \quad t = 2, \dots, T$$

$$Var(\omega_t | \omega_{t-1}) = \sigma^2 v_{t|t-1} = \sigma^2, \quad t = 2, \dots, T.$$

En consecuencia, la función de verosimilitud será:

$$f(\omega_T) = f(\omega_1) \prod_{t=2}^T \sigma^{-1} (2\pi)^{-1/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^T (\omega_t - c - \phi\omega_{t-1})^2 \right\} \quad (7.7)$$

Se toman logaritmos y se utiliza que  $f(\omega_1)$  es normal, con parámetros dados por (7.6) y (7.5), resulta la función soporte:

$$L(\emptyset, \sigma^2 | \omega_T) = -\frac{T}{2} \ln \sigma^2 + \frac{1}{2} \ln(1 - \emptyset^2) - \frac{(1 - \emptyset^2)(\omega_1 - \mu)^2}{2\sigma^2} - \frac{1}{2\sigma^2} \sum_{t=2}^T (\omega_t - c - \emptyset \omega_{t-1})^2 \quad (7.8)$$

Para obtener el estimador de  $\emptyset$  se tiene que derivar respecto a este parámetro e igualar a cero el resultado. Se obtiene una ecuación cúbica que tiene tres raíces, y la raíz que maximice la función de verosimilitud es el estimador MV.

La expresión (7.8) muestra que, si se prescinde del primer término, la función soporte tiene la expresión habitual de un modelo lineal. Si condicionamos a la primera observación se tiene que:

$$f(\omega_2, \dots, \omega_T | \omega_1) = \prod_{t=2}^T \sigma^{-1} (2\pi)^{-1/2} \times \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^T (\omega_t - c - \emptyset \omega_{t-1})^2 \right\}.$$

Se define la verosimilitud condicionada como la asociada a esta función de densidad conjunta:

$$L_C(\emptyset, \sigma^2 | \omega_1) = \frac{-(T-1)}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=2}^T (\omega_t - c - \emptyset \omega_{t-1})^2$$

Y el estimador del parámetro  $\emptyset$  que maximiza esta verosimilitud condicionada se obtiene minimizando la suma de los cuadrados:

$$\sum_{t=2}^T (\omega_t - c - \emptyset \omega_{t-1})^2 = \sum_{t=2}^T (\tilde{\omega}_t - \emptyset \tilde{\omega}_{t-1})^2$$

Donde  $\tilde{\omega}_t = \omega_t - \mu$ . Derivando e igualando a cero, dado que el estimador  $\bar{\omega} = \hat{\mu} = \sum_{t=1}^T \omega_t / T$ , para la media obtenida, el estimador de  $\emptyset$  es:

$$\hat{\emptyset} = \frac{\sum_{t=2}^T (\omega_t - \bar{\omega})(\omega_{t-1} - \bar{\omega})}{\sum_{t=2}^T (\omega_{t-1} - \bar{\omega})^2},$$

Que es el de la pendiente en un modelo de regresión de  $\omega_t$  respecto a  $\omega_{t-1}$ . El Estimador MV condicionado de la varianza será:

$$\hat{\sigma}^2 = \frac{\sum_{t=2}^T (\omega_t - \hat{c} - \hat{\phi}\omega_{t-1})^2}{T-1},$$

Donde  $\hat{c} = \bar{\omega}(1 - \hat{\phi})$ .

Si se condiciona al primer término y se escribe la verosimilitud de las observaciones desde la 2 a la T, se tiene un modelo lineal en los parámetros. La diferencia entre el estimador obtenido con la verosimilitud condicionada y la exacta será en general pequeña, y despreciable para muestras grandes.

#### 4.4.2.2. Procesos AR(p)

Se considera un proceso AR(p) general. La esperanza condicionada de  $\omega_t$ , para  $t = p + 1, \dots, T$  dados los datos previos,  $\omega_{t-1}, \dots, \omega_1$  será, utilizando la ecuación del AR(p):

$$E[\omega_t | \omega_{t-1}, \dots, \omega_1] = \mu + \phi_1(\omega_{t-1} - \mu) + \dots + \phi_p(\omega_{t-p} - \mu)$$

Y su varianza condicionada será:

$$Var(\omega_t | \omega_{t-1}, \dots, \omega_1) = Var(a_t) = \sigma^2.$$

En consecuencia, todas las distribuciones condicionadas para  $t = p + 1, \dots, T$  son normales, con media igual a la predicción a un paso y varianza  $\sigma^2$ . La función soporte condicionada se obtendrá a partir de la densidad conjunta de las observaciones  $\omega_{p+1}, \dots, \omega_T$  condicionadas a las  $p$  primeras. Su expresión es:

$$L_C(\mu, \phi, \sigma^2 | \omega_1, \dots, \omega_p) = -\frac{T-p}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=p+1}^T (\omega_t - \mu - \sum_{i=1}^p \phi_i(\omega_{t-i} - \mu))^2 \quad (7.9)$$

Maximizar esta función respecto a  $\mu$  y  $\phi$  equivale a minimizar la suma de los cuadrados de los errores de predicción a un paso, que, al no depender de las condiciones iniciales, serán iguales a las innovaciones y se puede escribir:

$$S = \sum_{t=p+1}^T a_t^2 = \sum_{t=p+1}^T (\omega_t - \mu - \sum_{i=1}^p \phi_i(\omega_{t-i} - \mu))^2 \quad (7.10)$$



Donde  $a_t = (\omega_t - \mu - \sum_{i=1}^p \phi_i(\omega_{t-i} - \mu))$ . Por tanto, maximizar el soporte condicional equivale a mínimos cuadrados. El estimador de  $\mu$  se obtiene derivando e igualando a cero:

$$\sum_{t=p+1}^T \left( \omega_t - \mu - \sum_{i=1}^p \phi_i(\omega_{t-i} - \mu) \right) = 0$$

Y suponiendo que  $\sum_{t=p+1}^T \omega_t \simeq \sum_{t=p+1}^T \omega_{t-i}$ , lo que será aproximadamente cierto si T es grande, se obtiene que el estimador de la media es la media muestral de las observaciones consideradas:

$$\hat{\mu} = \frac{\sum_{t=p+1}^T \omega_t}{T - p}.$$

Un estimador mejor de la media muestral de todas las observaciones  $\mu$  es  $\bar{\omega} = \sum_{t=1}^T \omega_t / T$ . Ambos estimadores son centrados, pero el calculado con toda la muestra tiene menor varianza y será el que se utilice. Esto equivale a estimar inicialmente la media con todos los datos y después escribir la verosimilitud para las variables en desviaciones a la media.

Para obtener el estimador  $\hat{\phi}$ , sustituyendo  $\mu$  por  $\bar{\omega}$  en (7.10) y llamando  $X'_t = (\omega_{t-1} - \bar{\omega}, \dots, \omega_{t-p} - \bar{\omega})$ , se obtiene el estimado habitual de mínimos cuadrados en modelos de regresión:

$$\hat{\phi} = (\sum_{t=p+1}^T X_t X'_t)^{-1} (\sum_{t=p+1}^T X_t (\omega_t - \bar{\omega})). \quad (7.11)$$

Esta expresión para muestras grandes será aproximadamente:

$$\hat{\phi} = \hat{\Gamma}_p^{-1} \hat{\gamma}_p, \quad (7.12)$$

Donde:

$$\hat{\Gamma}_p = \begin{bmatrix} \hat{\gamma}_0 & \dots & \hat{\gamma}_{p-1} \\ \dots & \hat{\gamma}_0 & \dots \\ \hat{\gamma}_{p-1} & \dots & \hat{\gamma}_0 \end{bmatrix}, \quad \hat{\gamma}_p = \begin{bmatrix} \hat{\gamma}_1 \\ \dots \\ \hat{\gamma}_p \end{bmatrix},$$

Que son las ecuaciones de Yule-Walker. Sin embargo, en muestras pequeñas ambos estimadores son diferentes y tanto más diferentes cuanto mayor sea el orden del proceso. Los

estimadores de mínimos cuadrados consideran la muestra de  $T - p$  observaciones y todos los coeficientes de (7.11) se calculan como cocientes de sumas de  $T - p$  términos en el numerador y en el denominador. Los estimadores de Yule-Walker utilizan distintos términos en cada sumando:  $\hat{\gamma}_0$  se calcula con los  $T$  datos pero  $\hat{\gamma}_k$  con  $T - k$ . Este desequilibrio introduce sesgos importantes y puede demostrarse (Tjostheim y Paulse, 1983) que los estimadores de mínimos cuadrados son más precisos que los Yule-Walker.

La estimación por la máxima verosimilitud exacta requiere calcular las esperanzas y varianzas condicionadas para las  $p$  primeras observaciones. Esto puede hacerse de manera similar a como se hizo para un AR(1).

### 4.4.3. *Estimación de modelos MA y ARMA*

#### 4.4.3.1. **Estimación Máxima Verosimilitud condicional**

La estimación de modelos con componentes MA y mixtos es más complicada que la de los AR por dos razones. En primer lugar, la función de verosimilitud, tanto la condicional como la exacta, es siempre no lineal en los parámetros. En segundo, el procedimiento de condicionar a ciertos valores iniciales, que lleva a resultados simples en los AR, es más complicado para procesos MA y ARMA, haciendo el cálculo de las esperanzas y varianzas condicionadas más difícil. Para ilustrar estos problemas, se considera el caso de un MA(1):

$$\omega_t = a_t - \theta a_{t-1}$$

Con esperanza marginal cero. La esperanza de  $\omega_t$  condicionada a sus valores previos ya no es inmediata, como en los AR, y para obtenerla tenemos que expresar  $\omega_t$  en función de los valores anteriores. Comenzando con  $t = 2$ , como  $\omega_2 = a_2 - \theta a_1$ , y  $a_1 = \omega_1 + \theta a_0$ , se tiene que:

$$\omega_2 = -\theta \omega_1 + a_2 - \theta^2 a_0$$

Y tomando esperanzas en esta expresión y suponiendo  $E(a_0|\omega_1) = 0$ , se deduce que la esperanza de la distribución condicionada es:

$$E(\omega_2|\omega_1) = -\theta\omega_1,$$

Y la varianza:

$$var(\omega_2|\omega_1) = \sigma^2(1 + \theta^4).$$

Procediendo de esta forma para  $t = 3, 4, \dots$ , se obtiene que:

$$\omega_t = -\theta\omega_{t-1} - \theta^2\omega_{t-2} - \dots - \theta^{t-1}\omega_1 + a_t - \theta^t a_0,$$

Que conduce a:

$$E(\omega_t|\omega_{t-1}, \dots, \omega_1) = -\theta\omega_{t-1} - \theta^2\omega_{t-2} - \dots - \theta^{t-1}\omega_1,$$

Y:

$$var(\omega_t|\omega_{t-1}, \dots, \omega_1) = \sigma^2(1 + \theta^{2t}).$$

Estas expresiones son no lineales en los parámetros y pesadas de calcular por este método de sustitución en procesos MA(q) generales.

Un enfoque alternativo es condicionar también en las primeras innovaciones no observadas.

Se observa que para cada valor los parámetros  $\theta$ :

$$a_t = \omega_t + \theta a_{t-1} \quad (7.13)$$

Permite calcular recursivamente las perturbaciones  $a_t$ , condicionadas a un valor inicial  $a_0$ .

Tomando  $a_0 = 0$  se puede calcular todas las restantes perturbaciones utilizando  $\omega_t$ . Entonces, como  $E(a_t|\omega_{t-1}, \dots, \omega_1, a_0) = E(a_t) = 0$  y  $E(a_{t-1}|\omega_{t-1}, \dots, \omega_1, a_0) = a_{t-1}$ , se tiene que:

$$E(\omega_t|\omega_{t-1}, \dots, \omega_1, a_0) = -\theta a_{t-1}$$

Y:

$$var(\omega_t|\omega_{t-1}, \dots, \omega_1, a_0) = E[(\omega_t + \theta a_{t-1})^2] = E[a_t^2] = \sigma^2$$

Que conduce a la verosimilitud condicionada en logaritmos o soporte condicional:

$$L_C(\theta, \omega_1 | a_0) = \frac{-(T-1)}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=2}^T a_t^2.$$

La estimación condicionada de modelos ARMA( $p, q$ ) se realiza siguiendo los mismo principios. Llamando  $r = \max(p, q)$  y  $\boldsymbol{\beta} = (\mu, \phi_1, \dots, \theta_q, \sigma^2)$  al vector de parámetros, la función soporte condicional es:

$$L_C(\boldsymbol{\beta} | \mathbf{a}_0, \boldsymbol{\omega}_p) = \frac{-(T-r)}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=r+1}^T a_t^2 \quad (7.14)$$

Donde se tiene que  $a_t^2 = a_t^2(\boldsymbol{\beta} | \mathbf{a}_0, \boldsymbol{\omega}_p)$ , para resaltar que las innovaciones se calculan a partir de unos vectores  $\mathbf{a}_0$  y  $\boldsymbol{\omega}_p$  de valores iniciales. Estas innovaciones estimadas o residuos se calculan recursivamente mediante:

$$\hat{a}_t = \omega_t - c - \phi_1 \omega_{t-1} - \dots - \phi_p \omega_{t-p} + \theta_1 \hat{a}_{t-1} + \dots + \theta_q \hat{a}_{t-q} \quad t = r+1, \dots, T \quad (7.15)$$

Donde  $c = \mu(1 - \phi_1 - \dots - \phi_p)$  y se supone que los primeros  $r$  residuos son cero. La maximización de (7.4) requiere un valor inicial de los parámetros.

#### 4.4.3.2. Estimación MV exacta

La estimación exacta requiere calcular media y varianza de las distribuciones de cada observación condicionada a las precedentes,  $\omega_t |_{t-1}$  y  $\sigma^2 v_t |_{t-1}$ , y la función a maximizar viene dada por (7.4).

Es importante resaltar que la descomposición de la función de densidad conjunta como producto de condicionadas es un caso particular de una descomposición que se utiliza en muchas aplicaciones estadísticas y que se conoce como factorización de Cholesky. Observemos que llamando  $\boldsymbol{\mu} = (\mu, \dots, \mu)$  al vector de medias marginales y  $\mathbf{M}_T$  a la matriz  $T \times T$  de covarianzas entre las variables  $\boldsymbol{\omega}_T$  es, en logaritmos:

$$\ln f(\boldsymbol{\omega}_T) = -\frac{T}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} |\mathbf{M}_T| - \frac{1}{2\sigma^2} (\boldsymbol{\omega}_T - \boldsymbol{\mu})' \mathbf{M}_T^{-1} (\boldsymbol{\omega}_T - \boldsymbol{\mu}) \quad (7.16)$$

Si se compara (7.16) y (7.4), se vé que en (7.16) se ha sustituido la forma cuadrática  $(\boldsymbol{\omega}_T - \boldsymbol{\mu})' \mathbf{M}_T^{-1} (\boldsymbol{\omega}_T - \boldsymbol{\mu})$  por la expresión más simple  $\mathbf{e}' \mathbf{D} \mathbf{e}$ , donde  $\mathbf{D}$  es una matriz diagonal

con términos  $v_{t|t-1}^{-1}$  y  $\mathbf{e} = (e_1, \dots, e_T)'$  que es el vector de errores de predicción a un paso. Además, en lugar de  $\ln|\mathbf{M}_T|$ , en (7.4), aparece el término  $\sum_{t=1}^T \ln v_{t|t-1}$ . Estos cambios pueden interpretarse como la realización de una transformación:

$$\mathbf{e} = \mathbf{L}(\boldsymbol{\omega}_T - \boldsymbol{\mu})$$

Donde la matriz  $\mathbf{L}$  es triangular inferior y tiene unos en la diagonal, de manera que  $e_t$  sea sólo función de los valores actuales y previos,  $(\omega_t, \dots, \omega_1)$  y no los posteriores  $(\omega_{t+1}, \dots, \omega_T)$ . De esta manera:

$$\mathbf{e}'\mathbf{D}\mathbf{e} = (\boldsymbol{\omega}_T - \boldsymbol{\mu})'\mathbf{L}'\mathbf{D}\mathbf{L}(\boldsymbol{\omega}_T - \boldsymbol{\mu}) = (\boldsymbol{\omega}_T - \boldsymbol{\mu})'\mathbf{M}_T^{-1}(\boldsymbol{\omega}_T - \boldsymbol{\mu})$$

Y se ha descompuesto la matriz  $\mathbf{M}_T^{-1}$  como producto de dos matrices triangulares  $\mathbf{T}'\mathbf{T}$  donde  $\mathbf{T} = \mathbf{D}^{1/2}\mathbf{L}$  y donde  $\mathbf{D}^{1/2}$  es la matriz diagonal que contiene los términos  $v_{t|t-1}^{-1/2}$ . De la misma forma, el determinante de  $\mathbf{M}_T$  es el producto de los términos diagonales de  $\mathbf{D}$ , ya que el determinante de la matriz triangular con unos en la diagonal,  $\mathbf{L}$ , es la unidad.

La clave de esta descomposición es la factorización de la matriz definida positiva  $\mathbf{M}_T^{-1}$  como producto de dos matrices triangulares, y se conoce como factorización de Cholesky.

#### 4.4.4. *Estimación recursiva del filtro de Kalman*

El filtro de Kalman es un procedimiento recursivo computacionalmente muy rápido que tiene muchas aplicaciones en series temporales. En particular, permite evaluar rápidamente la función de verosimilitud de cualquier modelo ARMA calculando los errores de predicción a un paso y sus varianzas. El filtro fue inicialmente diseñado para resolver un problema más general: la estimación de sistemas de control en el espacio de los estados, como se explica a continuación. El filtro tiene una inmediata interpretación Bayesiana, por lo que se ha utilizado mucho en la formulación de series temporales mediante este enfoque.

#### 4.4.4.1. Modelos en el espacio de los estados

Se supone que se observa un sistema que puede representarse mediante una ecuación de observación:

$$\mathbf{z}_t = \mathbf{H}_t \boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t \quad (7.17)$$

Donde  $\mathbf{z}_t$  es un vector de observaciones de dimensiones  $k \times 1$ ,  $\mathbf{H}_t$  es una matriz  $k \times p$  que se supone conocida para toda  $t$ ,  $\boldsymbol{\alpha}_t$  un vector de variables de estado de dimensión  $p \times 1$  que no se observa y  $\boldsymbol{\epsilon}_t$  es un proceso de ruido blanco que se supone tiene distribución  $N(\mathbf{0}, \mathbf{V}_t)$ . Además, la descripción del sistema incluye una ecuación que describe la evolución dinámica de las variables de estado,  $\boldsymbol{\alpha}_t$ , llamada ecuación de estado.

$$\boldsymbol{\alpha}_t = \boldsymbol{\Omega}_t \boldsymbol{\alpha}_{t-1} + \mathbf{u}_t \quad (7.18)$$

Donde  $\boldsymbol{\Omega}_t$  es una matriz conocida de dimensión  $p \times p$  y  $\mathbf{u}_t$  otro proceso de ruido blanco, independiente del anterior, que tiene distribución  $N_p(\mathbf{0}, \mathbf{R}_t)$ . Se supone que son las matrices del sistema  $\mathbf{H}_t$  y  $\boldsymbol{\Omega}_t$  se conocen para todos los instantes.

Por ejemplo, se supone que se observa en cada instante  $t$  la posición en el espacio de un satélite mediante las variables,  $\mathbf{z}_t$ , que será un vector de dimensión tres (las coordenadas de la posición del satélite en el espacio). Se supone que las medidas de la posición del satélite dependen de un conjunto de variables de estado,  $\boldsymbol{\alpha}_t$ , que no son observables, pero que están relacionadas con las observaciones a través de la matriz  $\mathbf{H}_t$ , que es conocida. Por ejemplo, si la posición del satélite depende de su velocidad y su aceleración, el vector de estado tiene dimensión dos. Se supone que existe en cada instante un error de observación, representado por el vector  $\boldsymbol{\epsilon}_t$ . Por otro lado, la ecuación de estado indica que el vector de variables de estado en el instante  $t$  depende de la situación de las variables de estado en el instante  $t-1$  más un error de medida.

La representación de un sistema mediante las ecuaciones (7.17) y (7.18) no es única. Siempre es posible aumentar la dimensión del vector de estado poniendo ceros en las matrices que le multiplican. Se dice que el vector de estado tiene dimensión mínima cuando no es posible representar el sistema con menos de  $p$  variables de estado. Una vez fijada la dimensión, el vector de estado tampoco es único, ya que dado un vector de estado  $\boldsymbol{\alpha}_t$ , el sistema puede igualmente

representarse con el vector de estado  $\alpha_t^* = \mathbf{A}\alpha_t$ , donde  $\mathbf{A}$  es cualquier matriz cuadrada no singular. En efecto, se escribe la ecuación de observación como:

$$\mathbf{z}_t = \mathbf{H}_t \mathbf{A}^{-1} \mathbf{A} \alpha_t + \epsilon_t = \mathbf{H}_t^* \alpha_t^* + \epsilon_t$$

Y la evolución del estado como:

$$\alpha_t^* = \mathbf{\Omega}_t^* \alpha_{t-1}^* + \mathbf{u}_t^*$$

Donde ahora  $\mathbf{\Omega}_t^* = \mathbf{A}\mathbf{\Omega}_t\mathbf{A}^{-1}$  y  $\mathbf{u}_t^* = \mathbf{A}\mathbf{u}_t$ . En adelante se supondrá que el sistema tiene dimensión mínima.

Cualquier modelo ARMA( $p, q$ ) puede escribirse en esta formulación como sigue. Se define  $m = \max(p, q + 1)$  y se llama  $\alpha_t = (\alpha_{1,t}, \alpha_{2,t}, \dots, \alpha_{m,t})'$  al vector de variables de estado, que seguirá la ecuación:

$$\begin{bmatrix} \alpha_{1,t} \\ \alpha_{2,t} \\ \dots \\ \alpha_{m,t} \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 & \dots & 0 \\ \phi_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & 1 \\ \phi_m & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \alpha_{1,t-1} \\ \alpha_{2,t-1} \\ \dots \\ \alpha_{m,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ -\phi_1 \\ \dots \\ -\phi_m \end{bmatrix} a_t. \quad (7.19)$$

Se observa que en esta ecuación la matriz de estado,  $\mathbf{\Omega}_t$ , tiene la forma:

$$\mathbf{\Omega}_t = \begin{bmatrix} \phi_{m-1} & \mathbf{I} \\ \phi_m & \mathbf{0}' \end{bmatrix}$$

Donde  $\phi_{m-1}$  es un vector columna de dimensión  $m - 1$ ,  $\mathbf{I}$  es la matriz de identidad y  $\mathbf{0}'$  es un vector de ceros. Por otro lado, el vector de innovaciones en esta ecuación es:

$$\mathbf{u}_t = \boldsymbol{\theta} a_t$$

Donde  $\boldsymbol{\theta}' = (1, -\theta_1, \dots, -\theta_m)$ . La matriz de varianzas y covarianzas de  $\mathbf{u}$  es

$$\mathbf{R}_t = \boldsymbol{\theta}\boldsymbol{\theta}'\sigma^2.$$

Se vá a comprobar que sustituyendo en la primera ecuación sucesivamente las variables de estado se obtiene la presentación del proceso ARMA. La primera ecuación es:

$$\alpha_{1,t} = \phi_1 \alpha_{1,t-1} + \alpha_{2,t-1} + a_t, \quad (7.20)$$

Y la segunda:

$$\alpha_{2,t} = \phi_2 \alpha_{1,t-1} + \alpha_{3,t-1} - \theta \alpha_t. \quad (7.21)$$

Sustituyendo  $\alpha_{2,t-1}$  en (7.20) de acuerdo con la expresión (7.21), se tiene que:

$$\alpha_{1,t} = \phi_1 \alpha_{1,t-1} + \phi_2 \alpha_{1,t-2} + \alpha_{3,t-2} + \alpha_t - \theta_1 \alpha_{t-1}. \quad (7.22)$$

La tercera ecuación es:

$$\alpha_{3,t} = \phi_3 \alpha_{1,t-1} + \alpha_{4,t-1} - \theta_2 \alpha_t$$

Y sustituyendo ahora en (7.22)  $\alpha_{3,t-2}$  por su expresión anterior, se va recuperando el proceso ARMA en la variable  $\alpha_{1,t}$ . La ecuación de observación sirve simplemente para hacer la variable observada,  $z_t$ , que es escalar, igual a la primera componente del vector de estado:

$$z_t = (1, 0, \dots, 0) \alpha_t. \quad (7.23)$$

Las ecuaciones (7.19) y (7.23) constituyen una forma de representar el modelo ARMA en el espacio de los estados. Se observa que son un caso particular de las (7.17) y (7.18). En la ecuación de la observación (7.17) el vector de datos es ahora un escalar, el valor de la serie observada en cada instante; el vector de estado es un vector de dimensiones  $m = \max(p, q + 1)$ .

La matriz  $\mathbf{H}_t$  es siempre el vector  $(1, 0, \dots, 0)$  y no existe error de medida o ruido en la matriz de observación. En la ecuación de estado la matriz  $\mathbf{\Omega}_t$  es univariante en el tiempo y la matriz de covarianzas de  $\mathbf{u}_t$  es singular de rango uno.

La representación de un modelo ARMA en el espacio de los estados no es única, véase Box, Jenkins y Reinsel (1994), Brockwell y Davies (1996) y Peña, Tiao y Tsay (2001) para representaciones alternativas.

Una forma de escribir modelos de series temporales en el espacio de los estados es mediante los modelos estructurales estudiados por West and Harrison (1989) y Harvey (1989), entre otros. En estos modelos la serie se expresa como suma de componente asociados a la tendencia y estacionalidad más un ruido blanco, como:



$$z_t = \mu_t + S_t + a_t.$$

Una formulación habitual de la tendencia es:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + v_t \quad (7.24)$$

$$\beta_t = \beta_{t-1} + u_t \quad (7.25)$$

Donde  $\beta_t$  representa el ritmo del crecimiento de la tendencia y  $u_t$  y  $v_t$  son procesos de ruido blanco independientes con covarianzas  $\sigma_u^2$  y  $\sigma_v^2$ . Si  $\sigma_u^2$  la varianza  $u_t$  en (7.25), es cero, la tendencia es constante o determinista. Si además  $\beta_t$  es cero, la serie no tiene tendencia. Si la varianza  $\sigma_u^2$  es grande, la tendencia puede cambiar mucho de un período al siguiente. La varianza  $\sigma_v^2$  determina cómo se mueve el nivel de la serie de un período al siguiente.

La estacionalidad se modela con coeficientes estacionales constantes que pueden ser arbitrarios o seguir una estructura trigonométrica. Suponiendo una estacionalidad constante, los coeficientes  $S_t$  deben verificar:

$$\sum_{j=1}^s S_{t-j} = \epsilon_t, \quad (7.26)$$

Las variables  $v_t$ ,  $u_t$  y  $\epsilon_t$  son ruidos blancos independientes. Este modelo tiene un valor de estado de dimensión 13, formado por las variables de estado  $\mu_t$ ,  $\beta_t$  y los 11 coeficientes estacionales. Definiendo:

$$\alpha_t = (\mu_t, \beta_t, S_t, S_{t-1}, \dots, S_{t-10})'$$

El modelo se escribe en el espacio de los estados con una ecuación de observación:

$$z_t = (1, 0, 1, 0, \dots, 0)\alpha_t + a_t$$

Y una ecuación de estado:

$$\begin{bmatrix} \mu_t \\ \beta_t \\ S_t \\ \dots \\ S_{t-9} \\ S_{t-10} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & -1 & \dots & -1 & -1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ S_{t-1} \\ \dots \\ S_{t-10} \\ S_{t-11} \end{bmatrix} + \begin{bmatrix} v_t \\ v_t \\ \epsilon_t \\ \dots \\ 0 \\ 0 \end{bmatrix}$$

Las tres primeras ecuaciones del vector de estado son las (7.24), (7.25) y (7.26). Las restantes son identidades, del tipo  $S_{t-j} = S_{t-j}$ .

#### 4.4.4.2. El filtro de Kalman

El filtro de Kalman es un algoritmo recursivo de predicción. Se va a presentar en su formulación general y después se indicará su particularización para calcular la función de verosimilitud de un proceso ARMA. El algoritmo funciona en tres pasos. En el primero, se estima el estado futuro a partir de la información del estado actual. En el segundo, se predice nuevas observaciones. En el tercero, que se realiza cuando llega una nueva observación al sistema, se revisa la estimación del estado en ese instante a la vista de la nueva información. Se va a revisar estas tres etapas.

La primera etapa es la estimación de estado futuro a partir de una estimación del estado actual. Suponiendo que se dispone de los datos  $\mathbf{Z}_{t-1} = \{z_1, \dots, z_{t-1}\}$  y un estimador del vector de estado,  $\hat{\boldsymbol{\alpha}}_{t-1}$ , y se desea estimar  $\hat{\boldsymbol{\alpha}}_{t|t-1}$ , el próximo valor del estado utilizando los datos observados,  $\mathbf{Z}_{t-1}$ . Esta estimación se calcula tomando las esperanzas en (7.18) condicionadas a  $\mathbf{Z}_{t-1}$  y se obtiene:

$$\hat{\boldsymbol{\alpha}}_{t|t-1} = \boldsymbol{\Omega}_t \hat{\boldsymbol{\alpha}}_{t-1} \quad (7.27)$$

Donde se ha utilizado la notación  $\hat{\boldsymbol{\alpha}}_{t|t-1} = \hat{\boldsymbol{\alpha}}_{t-1}$ . Se llama  $\mathbf{S}_{t|t-1}$  a la matriz de covarianzas de esta estimación:

$$\mathbf{S}_{t|t-1} = \left[ (\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1}) (\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1})' \mid \mathbf{Z}_{t-1} \right].$$

Para calcularla, se utiliza que restando el (7.18) de la ecuación (7.27):

$$\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1} = \boldsymbol{\Omega}_t (\boldsymbol{\alpha}_{t-1} - \hat{\boldsymbol{\alpha}}_{t-1}) + \mathbf{u}_t,$$

Y sustituyendo esta expresión en la definición de  $\mathbf{S}_{t|t-1}$  y llamando  $\mathbf{S}_{t-1} = \mathbf{S}_{t-1|t-1}$ , se obtiene que:

$$\mathbf{S}_{t|t-1} = \boldsymbol{\Omega}_t \mathbf{S}_{t-1} \boldsymbol{\Omega}_t' + \mathbf{R}_t. \quad (7.28)$$

Esta ecuación tiene una clara interpretación intuitiva: la incertidumbre al estimar el nuevo estado con información hasta  $t - 1$ , es la suma de la incertidumbre que se tenía respecto al estado anterior, medida por  $\mathbf{S}_{t-1}$ , y la incertidumbre del ruido en la ecuación del estado,  $\mathbf{R}_t$ . La matriz  $\mathbf{\Omega}_t$  aparece como en (7.18) para relacionar los componentes del estado en el instante  $t - 1$  y en el instante  $t$ . Si esta matriz fuese la identidad, el estado evoluciona como en un paseo aleatorio,  $\mathbf{\Omega}_t = \mathbf{I}$ , y la incertidumbre de estimación del estado aumentaría continuamente a través de la suma de la matriz  $\mathbf{R}_t$ . Como en general esta matriz no es la identidad, el aumento de la incertidumbre depende de su estructura. Por ejemplo, se supone un AR(1). Entonces, el vector de estado escalar, y  $\Omega_t = \phi < 1$ . La varianza de la estimación sigue el proceso:

$$s_{t|t-1} = \phi^2 s_{t-1} + \sigma^2$$

Y sólo una parte de la incertidumbre en  $t - 1$  se traslada al instante  $t$ .

El segundo paso del filtro es la predicción de la nueva observación  $\mathbf{z}_t$  dada la información hasta  $t - 1$ . Esta predicción se calcula de nuevo con la esperanza condicional dada  $\mathbf{Z}_{t-1}$  y se obtiene:

$$\hat{\mathbf{z}}_{t|t-1} = E(\mathbf{z}_t | \mathbf{Z}_{t-1}) = \mathbf{H}_t \hat{\boldsymbol{\alpha}}_{t|t-1}. \quad (7.29)$$

Esta predicción tendrá una incertidumbre que se medirá por la matriz de varianzas y covarianzas de los errores de predicción:

$$\mathbf{e}_t = \mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1},$$

Definida por:

$$\mathbf{P}_{t|t-1} = E[\mathbf{e}_t \mathbf{e}'_t].$$

Para calcular esta matriz, restando la predicción (7.29) de la ecuación de observación (7.17), se tiene que:

$$\mathbf{e}_t = \mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1} = \mathbf{H}_t (\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1}) + \boldsymbol{\epsilon}_t \quad (7.30)$$

Y sustituyendo esta expresión en la definición de  $\mathbf{P}_{t|t-1}$ , se obtiene que:

$$\mathbf{P}_{t|t-1} = \mathbf{H}_t \mathbf{S}_{t|t-1} \mathbf{H}'_t + \mathbf{V}_t. \quad (7.31)$$

Esta ecuación indica que la incertidumbre de la predicción acumula la incertidumbre sobre el estado y la del error de medida de las observaciones. El error de predicción que viene de la estimación del estado se modula dependiendo de la matriz  $\mathbf{H}_t$ . Si esta matriz es la identidad, lo que supone que las observaciones  $\mathbf{z}_t$  son mediciones de las variables de estado más un error aleatorio, al error de las variables de estado se añade el error de medición de las observaciones.

El tercer y último paso del filtro es revisar la estimación del estado a la vista de la nueva información. Supongamos que se ha observado  $\mathbf{z}_t$  con lo que la información disponible para ser  $\mathbf{Z}_t = (\mathbf{Z}_{t-1}, \mathbf{z}_t)$ . La nueva estimación del estado,  $\boldsymbol{\alpha}_t = \hat{\boldsymbol{\alpha}}_{t|t} = E(\boldsymbol{\alpha}_t | \mathbf{Z}_t)$ , se calcula por regresión (o con las propiedades de la normal multivariante) con:

$$E(\boldsymbol{\alpha}_t | \mathbf{Z}_{t-1}, \mathbf{z}_t) = E(\boldsymbol{\alpha}_t | \mathbf{Z}_{t-1}) + \text{cov}(\boldsymbol{\alpha}_t, \mathbf{z}_t | \mathbf{Z}_{t-1}) \text{var}(\mathbf{z}_t | \mathbf{Z}_{t-1})^{-1} (\mathbf{z}_t - E(\mathbf{z}_t | \mathbf{Z}_{t-1})). \quad (7.32)$$

En esta ecuación las esperanzas  $E(\boldsymbol{\alpha}_t | \mathbf{Z}_{t-1}) = \hat{\boldsymbol{\alpha}}_{t|t-1}$  y  $E(\mathbf{z}_t | \mathbf{Z}_{t-1}) = \hat{\mathbf{z}}_{t|t-1}$  son conocidas, así como la matriz  $\text{var}(\mathbf{z}_t | \mathbf{Z}_{t-1}) = \mathbf{P}_{t|t-1}$ . Lo único que queda por calcular es la covarianza entre el estado y la nueva observación, que viene dada por:

$$\text{cov}(\boldsymbol{\alpha}_t, \mathbf{z}_t | \mathbf{Z}_{t-1}) = E \left[ (\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1}) (\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})' \right] = E \left[ (\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1}) \mathbf{e}'_t \right]$$

Y sustituyendo (7.30),

$$\text{cov}(\boldsymbol{\alpha}_t, \mathbf{z}_t | \mathbf{Z}_{t-1}) = E \left[ (\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1}) ((\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1})' \mathbf{H}'_t + \boldsymbol{\epsilon}'_t) \right] = \mathbf{S}_{t|t-1} \mathbf{H}'_t. \quad (7.33)$$

Ya que el error de observación  $\boldsymbol{\epsilon}'_t$  es ruido blanco e independiente de  $\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1}$ . Sustituyendo esta covarianza en (7.32), podemos escribir:

$$\hat{\boldsymbol{\alpha}}_t = \hat{\boldsymbol{\alpha}}_{t|t-1} + \mathbf{K}_t (\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1}). \quad (7.34)$$

Donde  $\mathbf{K}_t$  es la matriz de coeficientes de regresión que se llama la ganancia del filtro, y viene dada por:

$$\mathbf{K}_t = \mathbf{S}_{t|t-1} \mathbf{H}'_t \mathbf{P}_{t|t-1}^{-1}$$

La ecuación (7.34) indica que la revisión que hacemos de la estimación previa del estado depende del error de predicción,  $\mathbf{e}_t = \mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1}$ . Si este error es cero, no modificamos la estimación; en otro caso, hacemos una modificación de la estimación del estado que depende del cociente entre el error en la estimación del estado,  $\mathbf{S}_{t|t-1}$ , y el error de predicción  $\mathbf{P}_{t|t-1}^{-1}$ . La matriz  $\mathbf{H}'_t$  permite comparar estas matrices. Una forma equivalente de escribir la ecuación (7.34) es:

$$\hat{\boldsymbol{\alpha}}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \hat{\boldsymbol{\alpha}}_{t|t-1} + \mathbf{K}_t \mathbf{z}_t,$$

Que indica que la estimación del estado es una combinación lineal de las dos fuentes de información de que se dispone. Por un lado, la estimación previa,  $\hat{\boldsymbol{\alpha}}_{t|t-1}$ , y por otra, la observación  $\mathbf{z}_t$  que también aporta información sobre el estado. Puede demostrarse que las ponderaciones de las dos fuentes de información son iguales a su precisión relativa. La matriz de covarianzas de esta estimación será:

$$\mathbf{S}_t = E[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_t)(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_t)' | \mathbf{Z}_t],$$

Y sustituyendo  $\hat{\boldsymbol{\alpha}}_t$  por su expresión en la ecuación (7.34), tenemos que:

$$\mathbf{S}_t = E\left[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1} - \mathbf{K}_t \mathbf{e}_t)(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t-1} - \mathbf{K}_t \mathbf{e}_t)' | \mathbf{Z}_t\right]$$

Y utilizando (7.33) y operando se obtiene finalmente que:

$$\mathbf{S}_t = \mathbf{S}_{t|t-1} - \mathbf{S}_{t|t-1} \mathbf{H}'_t \mathbf{P}_{t|t-1}^{-1} \mathbf{H}_t \mathbf{S}_{t|t-1}. \quad (7.35)$$

Las ecuaciones (7.27), (7.28), (7.29), (7.31), (7.34) y (7.35) constituyen el filtro de Kalman. Bajo la hipótesis de normalidad el filtro proporciona estimaciones y predicciones óptimas. La aplicación del filtro para obtener la función de verosimilitud del modelo ARMA requiere escribir el modelo en los espacios de los estados, como se ha visto, y calcular los errores de predicción  $\mathbf{e}_t = \mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1}$ , que en este caso son escalares y sus varianzas,  $\mathbf{P}_{t|t-1}$ . En este caso la matriz  $\mathbf{V}_t$  es cero. Para comenzar el filtro hace falta indicar un valor inicial para las variables

de estado,  $\alpha_0$ , y para su matriz de covarianzas  $\mathbf{S}_0$ . Estos valores iniciales no son cruciales, porque el filtro depende poco de las condiciones iniciales.

#### 4.4.5. *Propiedades de los estimadores*

Puede demostrarse que las propiedades asintóticas del método de máxima verosimilitud son válidas, en condiciones de regularidad generales, para los estimadores MV de los modelos ARMA. Estas condiciones exigen que el proceso sea estacionario y que el modelo ARMA que se estima no contenga factores comunes con su parte AR y MA. Si un proceso AR(1) tiene el parámetro igual a la unidad, el valor verdadero está en la frontera del intervalo paramétrico  $[-1,1]$  y no se cumplen las condiciones habituales de regularidad necesaria para obtener las propiedades asintóticas del estimador MV.

Para procesos estacionarios en muestras grandes los estimadores MV tendrán distribución aproximadamente normal y serán centrados y eficientes. En particular la matriz de segundas del soporte en su máximo proporciona directamente las varianzas y covarianzas de los estimadores:

$$\mathbf{Var}(\hat{\beta}_{MV}) = - \left[ \frac{\partial^2 L(\hat{\beta}_{MV})}{\partial \beta \partial \beta'} \right]^{-1}$$

Este resultado se basa en que la función de verosimilitud es aproximadamente cuadrática en el máximo. Para comprobarlo, conviene siempre que sea posible estudiar la forma de la función soporte alrededor del estimador MV, calculando numéricamente y dibujando curvas de nivel.

La condición de que no existan factores comunes en la parte AR y MA es importante. Por ejemplo, si  $\omega_t$  es ruido blanco y se estima el modelo:

$$(1 - \phi\beta)\omega_t = (1 - \theta\beta)a_t$$

Todos los valores de los parámetros con la condición  $\phi = \theta$  son compatibles con los datos y puede demostrarse que la varianza de los estimadores es infinita. En general, si el modelo está sobreparametrizado y contiene simultáneamente factores redundantes AR y MA tendremos una situación de fuerte multicolinealidad que puede dar lugar, primero, a múltiples máximos en la

función de verosimilitud y, segundo, a que la aproximación cuadrática en el máximo sea inadecuada.

#### 4.4.6. *Criterios de selección de modelos*

Se supone que se ha estimado un conjunto de modelos,  $M_1, \dots, M_m$ , y se desea seleccionar el que mejor explica la serie observada. Seleccionar el modelo por su ajuste a una muestra dada no resulta adecuado, ya que el modelo con más parámetros siempre conducirá a una mayor verosimilitud y a una menor suma de cuadrados de los errores dentro de la muestra. Por ejemplo, si se compara un  $AR(p)$  con un  $AR(p+1)$ , el ajuste del  $AR(p+1)$  no puede ser peor que el de  $AR(p)$ , y siempre se escoge el modelo más complejo. Por tanto, para seleccionar entre modelos se debe acudir a otros principios.

El problema puede verse como un problema de discriminación: tenemos modelos  $M_i$  y una serie estacionaria,  $\omega_T = (\omega_1, \dots, \omega_T)$  y se quiere seleccionar un modelo más compatible con la serie observada, y se puede abordar desde el punto de vista clásico o Bayesiano. Comenzando con el enfoque clásico, no es útil comparar la verosimilitud de distintos modelos, porque siempre el modelo con más parámetros tendrá mayor verosimilitud. Se puede, sin embargo, calcular el valor esperado de la verosimilitud para cada uno de los modelos, es decir, el valor que se espera obtener para la verosimilitud sobre muchas realizaciones del proceso si en cada una se estima los parámetros por máxima verosimilitud, y seleccionar aquel modelo que produzca un valor esperado más alto de esta verosimilitud esperada. Este es el enfoque que conduce al criterio de Akaike, que se expone a continuación.

Si se dispone de probabilidades a priori para cada modelo,  $P(M_i)$ , se podría utilizar el enfoque Bayesiano y seleccionar el modelo que tenga probabilidad máxima dados los datos. Es decir, se calcula para  $i = 1, \dots, m$  a posteriori.

$$P(M_i|\omega_T) = \frac{P(\omega_T|M_i)P(M_i)}{\sum_{j=1}^m P(\omega_T|M_j)P(M_j)}. \quad (7.36)$$

Y se selecciona el modelo más probable a la vista de los datos. Se observa que ese planteamiento no requiere que la serie sea estacionaria, por lo que puede aplicarse para comparar modelos con distinto número de diferencias y utilizarse para decidir como alternativa a los contrastes de raíces unitarias. Si se supone que las probabilidades a priori de todos los modelos son las mismas, este enfoque conduce al criterio de información Bayesiana (BIC).

#### 4.4.6.1. El criterio AIC de Akaike

La función de verosimilitud viene dada por (7.4). Multiplicando por -2 y tomando esperanzas en esta expresión se tiene que:

$$E(-2L(\beta)) = T \ln \sigma^2 + \sum_{t=1}^T \ln v_t |_{t-1} + E \left[ \sum_{t=1}^T \frac{e_t^2}{\sigma^2 v_t |_{t-1}} \right].$$

Se demuestra que si: (1) se supone que los parámetros se estiman con los datos; (2) se calcula la esperanza de esta verosimilitud respecto a futuras observaciones, se obtiene que:

$$AIC = E(-2L(\beta)) = T \ln \hat{\sigma}_{MV}^2 + 2k, \quad (7.37)$$

Donde  $T$  es el tamaño muestral utilizado para estimar el modelo,  $\hat{\sigma}_{MV}^2$ , el estimador MV de la varianza de las innovaciones y  $k$  el número de parámetros estimados para calcular las predicciones a un paso. Por tanto, seleccionar el modelo con verosimilitud esperada máxima equivale a escoger el que minimiza la verosimilitud con signo negativo dado por (7.37). Este criterio se conoce como AIC, y es debido a Akaike.

El problema con AIC es que tiende a sobreestimar el número de parámetros en el modelo y en ese efecto puede ser muy grande en pequeñas muestras. Una alternativa que corrige esta sobreestimación es el criterio AIC corregido, AICC, dado por:

$$AICC = T \ln \hat{\sigma}_{MV}^2 + T \frac{(1 + k/T)}{1 - (k + 2)/T}.$$

Si se utiliza el criterio AIC para comparar modelos ajustados a una serie, es importante que  $T$ , el número efectivo de observaciones utilizado para estimar el modelo, sea el mismo para todos ellos. Este número es el de datos estacionarios, que es igual a los datos originales menos



$d - sD$ , siendo  $d$  el número de diferencias regulares,  $s$  el periodo estacional y  $D$  el número de diferencias estacionales. Si se estima el modelo por máxima verosimilitud exacta se puede calcular residuos en todos los puntos y por tanto el número de datos efectivo es  $T - d - sD$ . Si se considera modelos con distinto número de diferencias y llamamos  $d_{max}$  y  $D_{max}$  los grados de diferenciación más altos de los modelos que se comparan, entonces el número de datos efectivos es:

$$T = T_0 - d_{max} - sD_{max}. \quad (7.38)$$

Sin embargo, si se utilizara un programa que realiza estimación condicionada para calcular los residuos, no se tienen estos datos sino un número menor, ya que, por ejemplo, para un AR(1), como se condiciona al primer dato el residuo para esa observación, no puede calcularse. En general, si se hace estimación condicional y se llama  $r = \max(p, q)$ , y  $R = \max(P, Q)$ , el número efectivo de observaciones con las que se puede calcular los residuos en un modelo dado es  $T - d - sD - r - R$ , y puede variar mucho de unos modelos a otros. Esta es una razón adicional de utilizar máxima verosimilitud exacta, de manera que siempre se considera los residuos en las mismas  $T$  observaciones en las comparaciones entre modelos. Si no se hace esto, es posible que parte de las diferencias observadas entre los modelos se deban a los distintos residuos calculados.

## 4.5. Modelos de heterocedasticidad condicional<sup>9</sup>

### 4.5.1. Modelos ARCH

Se supone que una serie de ruido blanco,  $e_t$ , que puede ser la serie de los residuos de un modelo ARMA. Si las variables  $e_t$  fuesen normales, la incorrelación garantiza la independencia, y no puede existir información en el pasado de la serie que sea útil para explicar su futuro. Los modelos ARCH relajan la hipótesis de normalidad y permiten que se tengan procesos de ruido blanco formados por variables dependientes. Una clase de modelos con esta propiedad es:

$$e_t = \sigma_t \epsilon_t \quad (8.1)$$

Donde  $\epsilon_t$  y  $\sigma_t$  son dos procesos estacionarios independientes entre sí. El proceso  $\epsilon_t$  es de ruido blanco normal estandarizado, es decir, formado por variables normales independientes de media cero y varianza unidad. El proceso  $\sigma_t$  es estacionario, pero tiene estructura dinámica, siendo su valor en  $t$  función del conjunto  $\mathbf{e}_{t-1} = (e_{t-1}, \dots, e_1)$  de los valores de la serie previos a  $t$ . La condición de independencia entre  $\epsilon_t$  y  $\sigma_t$  garantiza que la serie  $e_t$  tenga media marginal a cero, ya que:

$$E(e_t) = E(\sigma_t)E(\epsilon_t) = 0,$$

Y también media condicional nula, ya que:

$$E(e_t | \mathbf{e}_{t-1}) = E(\sigma_t | \mathbf{e}_{t-1})E(\epsilon_t) = 0.$$

Como el proceso  $e_t$  es estacionario, tendrá una varianza marginal constante, que se llamará  $\sigma^2$ . Esta varianza se calcula mediante:

---

<sup>9</sup> Los supuestos sobre los que se basa este apartado son retomados de Peña (2010). *Análisis de series temporales*. Madrid, Alianza.

$$E(e_t^2) = E(\sigma_t^2)E(\epsilon_t^2) = E(\sigma_t^2) = \sigma^2 \quad (8.2)$$

Y coincide con la varianza del proceso  $\sigma_t$ , ya que  $E(\epsilon_t^2) = 1$ . Sin embargo, el proceso  $e_t$  tendrá una varianza condicionada que no es constante. En efecto:

$$\text{Var}(e_t^2 | \mathbf{e}_{t-1}) = E(\sigma_t^2 | \mathbf{e}_{t-1})E(\epsilon_t^2) = \sigma_t^2$$

Dado que  $E(\epsilon_t^2 | \mathbf{e}_{t-1}) = E(\epsilon_t^2) = 1$ . Por tanto, el proceso  $\sigma_t^2$  tiene una interesante interpretación: representa la varianza condicionada de la serie en cada instante, que va variando en el tiempo con cierta estructura estacionaria.

La independencia entre los procesos  $\epsilon_t$  y  $\sigma_t$  garantiza que la serie  $e_t$  carece de autocorrelación y forma un proceso de ruido blanco. En efecto, las autocovarianzas de la serie son:

$$E(e_t e_{t-k}) = E(\sigma_t \epsilon_t \sigma_{t-k}) = E(\epsilon_t)E(\sigma_t \sigma_{t-k}) = 0,$$

Ya que el proceso  $\epsilon_t$  es independiente de su pasado y de los valores previos y actuales  $\sigma_t$ . Por tanto, el proceso  $e_t$  tendrá autocovarianzas nulas. Se observa, sin embargo, que, la serie  $e_t$  no es de variables independientes. En efecto,  $e_t$  y  $e_{t-1}$  son función de valores comunes pasados de la serie, aunque estén incorrelados serán, en general, dependientes.

En resumen, los modelos ARCH generalizan las hipótesis habituales de un ruido blanco permitiendo dependencia, que se introduce haciendo que la varianza condicionada no sea constante.

#### 4.5.1.1. El modelo ARCH(1)

En el modelo ARCH(1) se supone que la varianza condicional,  $\sigma_t^2$ , tiene una estructura similar a un AR(1), y depende sólo del último valor observado, mediante la ecuación:

$$E(e_t^2 | \mathbf{e}_{t-1}) = \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2, \quad (8.3)$$

Donde  $\alpha_0 > 0$  y  $\alpha_1 \geq 0$ , para asegurar que esta varianza es positiva. Esta ecuación establece que si el valor de  $e_{t-1}^2$  es alto, la varianza  $\sigma_t^2$  de la siguiente observación condicionada a este valor de  $e_{t-1}^2$  será también alta, haciendo más probable que el valor siguiente,  $e_t^2 = \sigma_t^2 \epsilon_t^2$ , sea alto. En efecto va a producir correlación entre los cuadrados de la serie, que se traducirá en

rachas de valores de magnitud relativamente elevada, o con mayor varianza. Por otro lado, como la media marginal y condicionada es cero, aunque la varianza condicionada sea alta, siempre es posible que aparezca un valor pequeño de  $e_t^2$ , disminuirá la varianza condicionada de la observación siguiente y facilitará que la siguiente sea pequeña en valor absoluto. De esta manera, la serie puede representar rachas de valores altos, pero globalmente seguirá un proceso estacionario.

Es intuitivo que la varianza marginal de la serie, que es en definitiva según (8.2) el promedio de las varianzas condicionales, debe ser mayor que  $\alpha_0$  y será tanto mayor cuanto mayor sea el coeficiente  $\alpha_1$  que transmite el efecto de la última observación. En efecto, llamando  $\sigma^2 = E(e_t^2)$  a la varianza marginal, se puede escribir:

$$\sigma^2 = E[E(e_t^2 | \mathbf{e}_{t-1})] = \alpha_0 + \alpha_1 E(e_{t-1}^2), \quad (8.4)$$

Se puede obtener la esperanza de una variable calculando: primero, la esperanza condicionada a otra variable (que puede ser vectorial como en este caso ya que  $\mathbf{e}_{t-1}$  contiene todas las variables previas) y segundo tomando la esperanza del resultado respecto a la variable condicionante. La esperanza  $E(e_t^2 | \mathbf{e}_{t-1})$  es, por (8.3), igual a  $\alpha_0 + \alpha_1 e_{t-1}^2$  y su esperanza respecto a la distribución de la variable  $\mathbf{e}_{t-1} = (e_{t-1}, \dots, e_1)$  es la misma que respecto a la distribución de  $e_{t-1}$ , ya que las restantes variables no intervienen en  $\alpha_0 + \alpha_1 e_{t-1}^2$ . La esperanza de la constante  $\alpha_0$  es ella misma y la esperanza de  $e_{t-1}^2$  respecto a su distribución es su varianza marginal. Como  $e_{t-1}$  es un proceso estacionario, se tiene que  $E(e_{t-1}^2) = E(e_t^2) = \sigma^2$ , como sustituyendo este resultado en (8.4), se obtiene que:

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1} \quad (8.5)$$

Y por, lo tanto,  $0 \leq \alpha_1 < 1$ .

#### **4.5.1.1.1. La autocorrelación de los cuadrados**

Se va a comprobar que el modelo ARCH(1) establece una dependencia de tipo AR(1) entre los cuadrados de las observaciones. Para ello, se define un nuevo proceso de ruido blanco,  $v_t$ , formado por variables estacionarias incorreladas de media cero y varianza marginal constante:

$$v_t = e_t^2 - \sigma_t^2 = \sigma_t^2(\epsilon_t^2 - 1), \quad (8.6)$$

Las variables  $v_t$  tienen media cero, ya que:

$$E(v_t) = E(e_t^2) - E(\sigma_t^2) = \sigma^2 - \sigma^2 = 0,$$

Y varianza constante:

$$var(v_t) = E(e_t^2 - \sigma_t^2)^2 = E(e_t^4) + E(\sigma_t^4) - 2E(e_t^2 \sigma_t^2). \quad (8.7)$$

Para comprobar que esta expresión es una constante, se observa primero que como  $e_t = \sigma_t \epsilon_t$ , y los procesos  $\sigma_t$  y  $\epsilon_t$ , son independientes:

1. La esperanza de  $E(e_t^4) = E(\epsilon_t^4 \sigma_t^4)$  será un producto de las esperanzas de  $E(\epsilon_t^4)$  y  $E(\sigma_t^4)$ . Como el proceso  $\epsilon_t$  es de variables normales e independientes se verifica que  $E(\epsilon_t^4) = 3$ , y  $E(e_t^4) = 3E(\sigma_t^4)$ .
2. La esperanza de  $E(e_t^2 \sigma_t^2)$ , sustituyendo  $e_t = \sigma_t \epsilon_t$ , es igual a  $E(\epsilon_t^2 \sigma_t^4) = E(\epsilon_t^2)E(\sigma_t^4) = E(\sigma_t^4)$ , ya que  $\epsilon_t$  tiene varianza unidad.

Sustituyendo estos resultados en (8.7), resulta que:

$$var(v_t) = 2E(\sigma_t^4),$$

Y como  $\sigma_t$  es un proceso estacionario, sus momentos son constantes, y se ha demostrado que el proceso  $v_t$  es homocedástico. Sin embargo, el proceso  $v_t$  es condicionalmente heterocedástico y está formado por variables incorreladas, pero no independientes. Para comprobarlo, se observa que tienen media cero y su covarianza será, utilizando (8.6):

$$E(v_t v_{t-k}) = E[\sigma_t^2 (\epsilon_t^2 - 1) \sigma_{t-k}^2 (\epsilon_{t-k}^2 - 1)],$$

Y, por la independencia entre los procesos  $\epsilon_t^2$  y  $\sigma_t^2$ , y aplicando que al tener  $\epsilon_t$  varianza unidad  $E(\epsilon_{t-k}^2 - 1) = 0$ ,

$$E(v_t v_{t-k}) = E(\epsilon_t^2 - 1) E(\sigma_t^2 \sigma_{t-k}^2 (\epsilon_{t-k}^2 - 1)) = 0$$

Se ha demostrado que el proceso  $v_t$  es de ruido blanco, formado por variables incorreladas de media cero y varianza marginal constante.

Escribiendo ahora, por (8.6),

$$e_t^2 = \sigma_t^2 + v_t,$$

Y sustituyendo  $\sigma_t^2$  por su expresión (8.3), se tiene la siguiente relación entre los cuadrados de las observaciones de la serie:

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + v_t, \quad (8.8)$$

Que establece una dependencia markoviana de primer orden, similar a un proceso AR(1), sobre los cuadrados. La ecuación (8.8) descompone la variable  $e_t^2$  en dos componentes ortogonales. Sin embargo, presenta dos diferencias principales respecto a un proceso AR(1) de variables normales. Con variables normales las innovaciones son independientes entre sí y, además, la innovación en el instante t es independiente de los valores pasados de la variable. Sin embargo, la ecuación (8.8):

1. Las variables  $v_t$  están incorreladas, pero no son independientes entre sí.
2. Las variables  $v_t$  están incorreladas, pero no son independientes, de las variables retardadas de la ecuación,  $e_{t-1}^2$ .

En efecto las variables  $v_t$  no va a ser independientes ya que, por definición (8.6),  $v_t$  es una función de  $e_t^2$  y  $\sigma_t^2$ , y  $v_{t-1}$  lo será de  $e_{t-1}^2$  y  $\sigma_{t-1}^2$ . Como  $\sigma_t^2$  depende de  $e_{t-1}^2$ , esto genera independencia entre ellas. Para comprobar que  $v_t$  depende de  $e_{t-1}^2$ , se observa que en (8.8)  $e_t^2$  debe ser no negativo, por lo que la suma de  $\alpha_0 + \alpha_1 e_{t-1}^2$  que siempre es positiva y  $v_t$ , que puede ser positiva o negativa, tiene que ser no negativa. En consecuencia, la variable  $v_t$  tendrá que satisfacer que:

$$v_t \geq -\alpha_0 - \alpha_1 e_{t-1}^2.$$

Sin embargo, es fácil comprobar que, aunque  $v_t$  no es independiente de  $e_{t-1}^2$ , está incorrelada con esta variable. Su covarianza es:

$$E(v_t e_{t-1}^2) = E(\sigma_t^2 (\epsilon_t^2 - 1) \sigma_{t-1}^2 \epsilon_{t-1}^2) = E(\epsilon_t^2 - 1) E(\sigma_t^2 \sigma_{t-1}^2) = 0.$$

El modelo AR(1) en los cuadrados establece esta estructura de autocorrelación entre los cuadrados de las variables. Llamando  $\rho_c(k)$  a la función de autocorrelación de los cuadrados de la serie, donde el subíndice c se refiere a los cuadrados, se obtiene:

$$\rho_c(k) = \alpha_1 \rho_c(k-1),$$

Que indica que las autocorrelaciones de los cuadrados de las series tienen la estructura de un AR(1) con parámetro  $\alpha_1$ , de acuerdo con la ecuación (8.8).

#### 4.5.1.2. El modelo ARCH(r)

El modelo anterior puede generalizarse permitiendo una dependencia de la varianza condicional con  $r$  retardos. Se obtiene así el proceso ARCH( $r$ ), para  $e_t = \sigma_t \epsilon_t$ , donde ahora:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_r e_{t-r}^2. \quad (8.9)$$

En este proceso las posibilidades de rachas de alta volatilidad dependen de los  $r$  últimos valores. La varianza marginal será:

$$\text{Var}(e_t) = E(e_t^2) = E[(e_t^2 | \mathbf{e}_{t-1})] = \alpha_0 + \sum \alpha_i E(e_{t-i}^2),$$

Que implica,

$$\text{Var}(e_t) = \frac{\alpha_0}{1 - \alpha_1 - \dots - \alpha_r}$$

Lo que exige la restricción  $\sum_{i=1}^r \alpha_i < 1$ .

Si se introduce, la variable  $v_t = e_t^2 - \sigma_t^2$ , podemos expresar la dependencia de los cuadrados de las observaciones como un proceso AR( $r$ ):

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_r e_{t-r}^2 + v_t, \quad (8.10)$$

Donde la secuencia de los  $v_t$  tiene las mismas propiedades que en el modelo ARCH(1): variables incorreladas de media cero, varianza constante e incorreladas con los regresores. Por otro lado, estas variables no son independientes entre sí ni los regresores, ya que la positividad de  $e_t^2$  exige que:

$$v_t > -(\alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_r e_{t-r}^2)$$

La ecuación (8.10) indica que existirá dependencia entre los cuadrados de las observaciones, y se podría haber utilizado esta expresión en lugar de (8.9) para definir el proceso ARCH( $r$ ). En

efecto, tomando esperanzas en (8.10) condicionadas a los valores observados,  $e_{t-1} = e_{t-1}, \dots, e_1$  se obtiene (8.9).

## 4.5.2. Modelos GARCH

La identificación de los modelos ARCH conduce generalmente a modelos de orden alto. Una idea natural es aproximar un proceso AR alto mediante una media móvil en las varianzas. Esto conduce al modelo GARCH o ARCH generalizado. Se va a estudiar las propiedades de este modelo en el caso más simple. Se supondrá que la serie estudiada se ha generado mediante  $e_t = \sigma_t \epsilon_t$  donde  $\epsilon_t$  y  $\sigma_t$  son dos procesos estacionarios independientes entre sí y con las mismas propiedades anteriores.

### 4.5.2.1. El modelo GARCH(1,1)

Se supone que las varianzas condicionales siguen la ecuación:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (8.11)$$

Con  $\alpha_1$  y  $\beta_1$  coeficientes positivos. En este modelo un valor alto para  $e_t^2$  da lugar a una mayor varianza en el periodo siguiente y el término  $\beta_1$  obliga a que la varianza cambie con cierta inercia, lo que produce rachas de mayor variabilidad.

Llamando como antes  $\sigma^2$  a la varianza marginal del proceso  $e_t$ :

$$Var(e_t) = E(e_t^2) = E(\sigma_t^2) = \sigma^2,$$

Y tomando esperanzas en (8.11), se tiene que:

$$\sigma^2 = \alpha_0 + \alpha_1 E(e_{t-1}^2) + \beta_1 E(\sigma_{t-1}^2),$$

Y al ser  $e_t$  estacionario,  $Var(e_t) = Var(e_{t-1})$ , que implica  $E(\sigma_t^2) = E(\sigma_{t-1}^2)$ , se tiene que:

$$Var(e_t) = \sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1},$$



Y para que el proceso sea estacionario  $\alpha_1 + \beta_1 < 1$ . Es frecuente en las aplicaciones a series de rendimientos financieros que la suma de estos dos parámetros esté cerca de la unidad.

#### 4.5.2.1.1. Dependencia de los cuadrados

Se va a escribir este modelo en la representación de los cuadrados de las variables introduciendo como se hizo en los procesos ARCH, el proceso de ruido blanco:

$$v_t = e_t^2 - \sigma_t^2,$$

Que, por construcción, es inmediato que es un proceso estacionario (por ser la diferencia de procesos estacionarios) de media cero. Se comprueba, igual que en los modelos ARCH, que las variables  $v_t$ , aunque incorreladas, no son independientes. Sustituyendo  $\sigma_t^2 = e_t^2 - v_t$  en la ecuación (8.11), se puede escribir:

$$e_t^2 = \alpha_0 + (\alpha_1 + \beta_1)e_{t-1}^2 + v_t - \beta_1 v_{t-1}, \quad (8.12)$$

Que muestra que los cuadrados de las observaciones siguen una estructura de dependencia similar a un ARMA(1,1).

Al coeficiente  $\lambda = (\alpha_1 + \beta_1)$  se le denomina *persistencia* y es frecuente en las aplicaciones a series financieras, que se estime próximo a la unidad. Se demuestra que, llamando como en la ecuación anterior  $\rho_c(k)$  a los coeficientes de autocorrelación al cuadrado de la serie, se obtiene:

$$\rho_c(1) = \frac{\alpha_1(1 - \alpha_1\beta_1 - \beta_1^2)}{(1 - 2\alpha_1\beta_1 - \beta_1^2)}$$

Mientras que:

$$\rho_c(k) = (\alpha_1 + \beta_1)^{k-1} \rho_c(1), \quad k > 1$$

Que indica que el decrecimiento de estos coeficientes depende de la persistencia  $\lambda = (\alpha_1 + \beta_1)$ , y será lenta si este coeficiente es próximo a la unidad. Se observa que en el caso particular  $\beta = 0$  obtenemos la estructura ARCH(1).

#### 4.5.2.2. El modelo GARCH general

El modelo anterior puede generalizarse permitiendo simultáneamente una estructura AR de orden  $r$  en la dependencia de la varianza de los cuadrados de la serie y una estructura MA de orden  $s$  en la varianza condicional donde la secuencia  $\epsilon_t$  tiene las mismas propiedades que en los procesos ARCH. Los parámetros de esa ecuación deben verificar ciertas restricciones para que la varianza sea positiva y existan los momentos de orden superior. En particular, como la varianza debe ser positiva,  $\alpha_0 > 0, \alpha_i \geq 0, \beta_i \geq 0$ , y  $\sum_{i=1}^{\max(r,s)} (\alpha_i + \beta_i) < 1$ .

Definiendo como antes las variables incorreladas de media cero y varianza constante  $v_t = e_t^2 - \sigma_t^2$ , se puede escribir la dependencia de los cuadrados como:

$$e_t^2 = \alpha_0 + \sum_{i=1}^{\max(r,s)} (\alpha_i + \beta_i) e_{t-i}^2 + \eta_t - \sum_{j=1}^s \beta_j \eta_{t-j}, \quad (8.14)$$

Ecuación que representa un proceso ARMA en los cuadrados de las observaciones. Se espera, en consecuencia, que los cuadrados de las observaciones presenten estructura dinámica similar a la de un proceso ARMA. Por otro lado, tomando esperanzas en esta ecuación, se obtiene inmediatamente que:

$$E(e_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(r,s)} (\alpha_i + \beta_i)}$$

Que implica restricciones sobre los parámetros del proceso para que sea estacionario.

Cuando el operador AR de la representación de los cuadrados tiene una raíz en el círculo unidad, se obtienen procesos con varianza condicional que siguen un proceso integrado no estacionario, o procesos IGARCH. En particular el IGARCH(1,1) es, haciendo  $\alpha_1 + \beta_1 = 1$  en (8.12):

$$\nabla e_t^2 = \alpha_0 + v_t - \beta_1 v_{t-1},$$

Con varianza condicional, poniendo  $\alpha_1 = 1 - \beta_1$  en (8.11), dada por:

$$\sigma_t^2 = \alpha_0(1 - \beta_1)e_{t-1}^2 + \beta_1\sigma_{t-1}^2.$$

La varianza condicional en este caso es una combinación lineal del cuadrado de la variable y de la varianza condicional anterior.

### 4.5.3. Construcción de los modelos ARCH y GARCH

La construcción de estos modelos implica primero identificar la estructura; segundo, estimar los parámetros; y tercero, comprobar su adecuación a los datos. Vamos a revisar estos tres aspectos.

#### 4.5.3.1. Identificación

La identificación de efectos ARCH se efectúa normalmente después de ajustar un modelo ARMA a la serie para eliminar la dependencia en la media. Si existen efectos ARCH, los residuos del modelo ARMA estarán incorrelados pero serán independientes y este efecto será visible en la función de autocorrelación de los residuos al cuadrado, que mostrarán correlación serial. Además, si se calculan los coeficientes de autocorrelación parcial de los residuos al cuadrado y el modelo para los residuos es ARCH puro, el número de términos distintos de cero nos indicará, aproximadamente, el orden del proceso.

#### 4.5.3.2. Estimación

Se va a explicar la estimación de estos modelos en el caso de un ARCH(1). Las ideas se generalizan fácilmente para otros modelos más complejos. Los datos de la serie verifican la ecuación (8.8):

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + v_t, \quad (8.15)$$

Y se podría estimar esta ecuación por mínimos cuadrados. El ruido de esta ecuación,  $v_t$ , está incorrelado con el regresor, por lo que la estimación proporcionará estimadores consistentes. Sin embargo, la varianza condicionada en esta ecuación no es constante y la regresión es heterocedástica, por lo que no serán eficientes. En efecto, utilizando, que según (8.6):

$$v_t^2 = \sigma_t^4 (\epsilon_t^2 - 1)^2$$

Se tiene que:

$$\text{var}(v_t | e_{t-1}^2) = E(\sigma_t^4 | e_{t-1}^2) E(\epsilon_t^4 + 1 - 2\epsilon_t^2) = 2\sigma_t^4,$$

Ya que:

$$E(\epsilon_t^4 + 1 - 2\epsilon_t^2) = 3 + 1 - 2 = 2,$$

Y  $E(\sigma_t^4 | e_{t-1}^2)$  es una constante igual a  $\sigma_t^4$  ya que  $\sigma_t^4$  sólo depende de  $e_{t-1}^2$ , que aparece condicionado. Al depender  $\sigma_t^4$  de  $e_{t-1}^2$ , que es el regresor de la ecuación, la varianza condicionada no es constante y se tendría que estimar la ecuación (8.15) por mínimos cuadrados generalizados. Para ello, se necesita un estimador de la desviación típica,  $\sigma^2\sqrt{2}$ . Entonces, dada una estimación de  $\sigma_t^2$ , que se llamará  $\hat{\sigma}_t^2$ , se obtiene una innovación de varianza condicional constante. Llamando  $y_t = e_t^2 / \hat{\sigma}_t^2$ ,  $z_t = 1 / \hat{\sigma}_t^2$ ,  $x_t = e_{t-1}^2 / \hat{\sigma}_t^2$  y  $u_t = v_t / \hat{\sigma}_t^2$ , se puede escribir (8.15) como:

$$y_t = \alpha_0 z_t + \alpha_1 x_t + u_t \quad (8.16)$$

Y en esta ecuación el residuo  $u_t$  será aproximadamente homocedástico.

Se va a comprobar que éste es precisamente el resultado de aplicar máxima verosimilitud a un modelo ARCH(1). La función de verosimilitud es:

$$f(e_1, \dots, e_T | \alpha_0, \alpha_1) = f(e_1) f(e_2 | e_1) \dots f(e_T | e_{T-1}).$$

Donde las funciones de densidad  $f(e_T | e_{T-1})$  son normales. En efecto, como  $e_t = \sigma_t \epsilon_t$ , si se condiciona a  $e_{t-1}$ , el valor de  $\sigma_t$  es una constante y la única variable es  $\epsilon_t$  que tiene distribución normal. La media condicionada de la distribución  $f(e_t | e_{t-1})$  es cero, y la varianza  $\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2$ . La función de verosimilitud condicionada a  $e_1$  en logaritmos será:

$$L(e_2, \dots, e_T | \alpha_0, \alpha_1, e_1) = -\frac{1}{2} \sum_{t=2}^T \ln(\alpha_0 + \alpha_1 e_{t-1}^2) - \frac{1}{2} \sum_{t=2}^T \frac{e_t^2}{\alpha_0 + \alpha_1 e_{t-1}^2},$$

Y derivando respecto a los parámetros, llamando  $\hat{\sigma}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1 e_{t-1}^2$  e igualando a cero, se obtienen las ecuaciones:

$$\sum \frac{e_t^2}{\hat{\sigma}_t^4} = \sum \frac{1}{\hat{\sigma}_t^2} \quad (8.17)$$

$$\sum \frac{e_{t-1}^2}{\hat{\sigma}_t^2} = \sum \frac{e_{t-1}^2 e_t^2}{\hat{\sigma}_t^4} \quad (8.18)$$

Para resolverlas, multipliquemos y dividamos el primer miembro por  $\hat{\alpha}_0 + \hat{\alpha}_1 e_{t-1}^2$ , y se escribe (8.17) como:

$$\hat{\alpha}_0 \sum \frac{1}{\hat{\sigma}_t^4} + \hat{\alpha}_1 \sum \frac{e_{t-1}^2}{\hat{\sigma}_t^4} = \sum \frac{e_t^2}{\hat{\sigma}_t^4}, \quad (8.19)$$

Y (8.18) como:

$$\hat{\alpha}_0 \sum \frac{e_{t-1}^2}{\hat{\sigma}_t^4} + \hat{\alpha}_1 \sum \frac{e_{t-1}^4}{\hat{\sigma}_t^4} = \sum \frac{e_{t-1}^2 e_t^2}{\hat{\sigma}_t^4},$$

Que son las ecuaciones de mínimos cuadrados para obtener parámetros en la regresión (8.16).

Este resultado sugiere el siguiente procedimiento iterativo para maximizar la función de verosimilitud. Estimar los mínimos cuadrados la regresión (8.15) y obtener unos valores iniciales de los parámetros. Calcular con estos valores  $\hat{\sigma}_t^2$  y resolver las ecuaciones (8.18) y (8.19) para obtener nuevos estimadores de  $\hat{\alpha}_0$  y  $\hat{\alpha}_1$ . Con ellos, calcular de nuevo  $\hat{\sigma}_t^2$  y volver a las ecuaciones (8.18) y (8.19) para obtener un nuevo estimador. Iterar hasta convergencia.

En general, la función de verosimilitud de un proceso estacionario  $w_t$  cuyas variables siguen un proceso GARCH( $r,s$ ), donde se supone  $r \geq s$  es llamado  $W_t = (w_1, \dots, w_t)$ .

$$\begin{aligned} f(w_1, \dots, w_n | \alpha) &= \\ &= f(w_n | W_{n-1}) f(w_{n-1} | W_{n-2}) \dots f(w_{r+1} | W_r) f(w_1, \dots, w_r | \alpha) = \\ &= \prod_{t=r+1}^n \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp \left[ -\frac{(w_t - E(w_t | W_{t-1}))^2}{2\sigma_t^2} \right] \times f(w_1, \dots, w_r | \alpha) \end{aligned}$$

Ya que la varianza condicional de las variables es  $\sigma_t^2$ , condicionando a las primeras  $r$  observaciones, que tienen una distribución más complicada, la función soporte condicionada es:

$$L(w_{r+1}, \dots, w_n | \alpha, W_r) = -\frac{1}{2} \sum_{t=r+1}^n \ln \sigma_t^2 - \frac{1}{2} \sum_{t=r+1}^n \frac{(w_t - E(w_t | W_{t-1}))^2}{\sigma_t^2} \quad (8.20)$$

Que puede maximizarse con un algoritmo de optimización no lineal para obtener los estimadores de los parámetros que aparecen en la media condicional y en la varianza condicional.

La estimación puede realizarse en dos etapas ya que la correlación entre los parámetros ARMA y la de los GARCH suele ser pequeña. En la primera etapa se estiman los parámetros de la media condicional, es decir, el modelo ARMA, y se construyen las innovaciones  $e_t = w_t - E(w_t | W_{t-1})$ . En la segunda se estiman los parámetros de la varianza condicional maximizando la verosimilitud de los residuos. Alternativamente, es posible estimar las ecuaciones de la media condicional y la varianza condicional conjuntamente, con lo que se obtiene una estimación más precisa.

#### 4.5.3.3. Diagnosis

Si se llama  $e_t$  a los residuos y  $\hat{\sigma}_t$  a las varianzas condicionadas estimadas, los residuos estandarizados,  $e_t/\hat{\sigma}_t$ , deben seguir un proceso de ruido blanco normal .

Un punto importante a tener en cuenta en la diagnosis de estos modelos es la posible confusión entre valores atípicos y heterocedasticidad condicional. Los valores atípicos pueden interpretarse como un aumento de la varianza en ese instante y, especialmente si aparecen en rachas, pueden confundirse con efectos de heterocedasticidad condicional. Por otro lado, una serie que sigue un modelo ARCH puede mostrar muchos atípicos si se analiza como si tuviese varianza constante y siguiese un modelo ARMA. Por lo tanto, es importante diferenciar ambos fenómenos. Un procedimiento simple que funciona bien en la práctica es limpiar la serie inicialmente de las observaciones que presentan residuos tan grandes que no pueden ser debidas a heterocedasticidad condicional y que son muy probablemente valores atípicos. Una regla simple y efectiva es tener en cuenta que es muy poco probable que la heterocedasticidad condicional pueda generar observaciones con residuos mayores de siete desviaciones típicas, y considerar estos datos como atípicos. A continuación se estiman los efectos ARCH o GARCH con la serie corregida y se construye la serie de innovaciones estandarizadas,  $e_t/\sigma_t$ . Sobre esta serie de innovaciones se puede, de nuevo, buscar valores atípicos y, si se detectan, corregirlos y volver a modelar las varianzas condicionales. El proceso se itera hasta convergencia.

#### 4.5.4. Modelos de volatilidad estocástica

Un procedimiento alternativo para modelar heterocedasticidad es suponer que la varianza condicional depende de factores no observados. Se obtienen así los modelos de volatilidad estocástica, o modelos SV (Stochastic volatility) que han sido estudiados por Harvey, Ruiz y Shephard (1994) y Jacquier, Polson y Rossi (1994). La ecuación estructural de estos modelos es idéntica a la de los modelos GARCH,  $e_t = \sigma_t \epsilon_t$ , donde los errores  $\epsilon_t$  son iid  $N(0,1)$ . Sin embargo, a diferencia de los GARCH donde la varianza depende de los factores observables, en este modelo se supone que los logaritmos de las varianzas condicionadas siguen un proceso  $AR(p)$ :

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^p \alpha_i \ln(\sigma_{t-i}^2) + \xi_t, \quad (8.21)$$

Donde el término de error en esta ecuación,  $\xi_t$ , es iid  $N(0, \sigma_\xi^2)$  e independiente de  $\epsilon_t$ . El coeficiente  $\alpha_0$  es una constante y los parámetros del proceso verifican las condiciones estándar de estacionaridad. Es decir, los ceros del polinomio  $1 - \sum_{i=1}^p \alpha_i B^i$  deben estar fuera del círculo unidad.

#### 4.6. Método de los datos sustitutos

El método de los datos sustitutos establece un régimen para poner a prueba hipótesis específicas sobre la dinámica del sistema, el método procede a comparar el valor de las estadísticas (no lineales) de los datos y la distribución aproximada de las diversas clases de sistemas lineales; de este modo prueba si los datos tienen algunas características que son distintas de sistemas estocásticos lineales. Desafortunadamente, lo que la mejor técnica de datos sustitutos puede hacer es mostrar que los datos estadísticamente pueden no ser coherentes con determinados sistemas.

En primer lugar, se suponen las siguientes hipótesis: 0) Ruido independiente e idénticamente distribuido (i.i.d); 1) ruido filtrado linealmente; y 2) transformación no lineal estática monótona del ruido filtrado linealmente.

El principio de los datos sustitutos es el siguiente: un primer supuesto es que los datos provienen de alguna clase de proceso dinámico específico que podría corresponder a un modelo paramétrico de los datos. Entonces se generan datos sustitutos de ese proceso hipotético y se calculan varias estadísticas a los sustitutos y a los datos originales. Los datos sustitutos darán la distribución esperada de los valores estadísticos y uno puede comprobar que los datos originales tienen un valor típico. Si los datos originales tienen estadísticas atípicas, se rechaza la hipótesis de que el proceso que genera los datos originales es de la clase asumida. Siempre se progresa desde suposiciones simples y específicas a un modelo mayor y más sofisticado.

En cualquier caso, para comprobar si hay alguna evidencia de que los datos son inconsistentes con la hipótesis, se comparan los datos con los sustitutos. Al elegir un estadístico de prueba apropiado, se puede comparar el valor del estadístico de prueba para los datos y el conjunto de valores obtenidos para los sustitutos. Si son diferentes, se puede concluir que los sustitutos, y por lo tanto la hipótesis subyacente, no son representación de los datos y se calculan varias estadísticas a los sustitutos y a los datos originales. Los datos sustitutos darán la distribución esperada de los valores estadísticos y uno puede comprobar que los datos originales tienen un valor típico. Si los datos originales tienen estadísticas atípicos, se rechaza la hipótesis de que el proceso que genera los datos originales es de la clase asumida. Siempre se progresa desde suposiciones simples y específicas a un modelo mayor y más sofisticado.

#### **4.6.1. *Fundamentación***

Con base en lo desarrollado por Theiler, et al. (1992) y Theiler (1995), y mejorado posteriormente por Small, Yu y Harrison (2001), Small y Tse (2003) y Small (2005), se tiene:

Sea  $\phi$  una hipótesis específica y  $F_\phi$  el conjunto de todos los procesos (o sistemas) consistentes con la hipótesis. Sea  $\bar{Z} \in R^N$  la serie de tiempo (que consiste en  $N$  medidas escalares) a



considerar, y sea  $T: R^N \rightarrow \mu$  una estadística cualquiera usada para probar la hipótesis  $\phi$  de modo que  $\bar{Z}$  es generada por algún proceso  $F \in F \phi$ . El conjunto de datos sustitutos  $\bar{S}_i$ ,  $i = 1, 2, \dots, n$  son generados desde  $\bar{Z}$  (y son de la misma longitud de  $\bar{Z}$  y consistentes con la hipótesis  $\phi$  puesta a prueba). Generalmente  $\mu \subseteq R$  y se puede discriminar entre el dato  $\bar{Z}$  y el sustituto  $\bar{S}_i$  consistente con la hipótesis, dada la densidad de probabilidad aproximada  $P_{T,F}(t) = P(T(S^{-i}))$  es decir, la densidad de probabilidad de  $T$  dado  $F$ .

En una segunda parte de su trabajo, Theiler sugiere que hay dos tipos diferentes de estadísticos de prueba: fundamentales y no fundamentales.

- *Definición 1:* Un estadístico de prueba  $T$  es fundamental si la densidad de probabilidad  $P_{T,F}$  es la misma para todos los procesos  $F$  consistentes con la hipótesis, de lo contrario no es fundamental.

Una elección prudente de un estadístico de prueba (es decir, un estadístico de prueba fundamental) puede aliviar algo de la tensión de garantizar que el algoritmo de generación de sustitutos ha creado sustitutos *derechos* (Small, 2005). Claramente la densidad de probabilidad  $P_{T,F}$  depende de la elección de  $T$  y  $F$ . Sin embargo, se desea  $P_{T,F}(t) = P(T(\bar{Z} < t | \bar{Z} \text{ es consistente con } \phi))$ . Para más consideraciones precisas de esta situación, Theiler también diferencié dos diferentes tipos de hipótesis: Hipótesis Simples e Hipótesis Compuestas.

- *Definición 2:* Una hipótesis es simple si el conjunto de todos los procesos consistentes con la hipótesis  $F \phi$  es unitario.
- *Definición 3:* Si una hipótesis no es simple es compuesta. Para una hipótesis simple el método de datos sustitutos es trivial, solo se necesita comparar varias realizaciones del sistema particular con los datos observados y decidir si es típica o no. Cuando se ha formulado una hipótesis de un problema, no solo se han generado sustitutos consistentes con  $F$  (de un proceso particular) también se estima  $F \in F \phi$ . Theiler argumenta que es conveniente utilizar estadísticos de prueba fundamentales si la hipótesis es compuesta. En el caso de que la hipótesis sea compuesta se debe especificar a  $F$ , a menos de que el estadístico de prueba  $T$  sea fundamental. En este caso  $P_{T,F}$  es el mismo para todo  $F \in F \phi$ . En los casos en que los estadísticos de prueba no fundamentales no son aplicables a

hipótesis compuestas, Theiler sugiere que un esquema de realización más limitado sea empleado.

- *Definición 4:* Sea  $F \in F \phi$  el proceso estimado de los datos  $\bar{Z}$  y sea  $\bar{S}_i$  el conjunto de datos sustitutos generados por  $F_i \in F \phi$ . Sea  $\hat{F}_i \in F \phi$  el proceso estimado de  $\bar{S}_i$ , un sustituto  $\bar{S}_i$  es una realización limitada de  $\hat{F}_i = \hat{F}$ .
- *Definición 5:* Si  $\hat{F}_i \neq \hat{F}$ , el sustituto  $\bar{S}_i$  es no limitado.
- *Definición 6:* Un algoritmo que genera sustitutos limitados es limitado. Uno que no, es no limitado. Es decir, así como la generación de sustitutos son realizaciones típicas de un modelo de los datos, hay que asegurarse de que los sustitutos son realizaciones de un proceso que da estimaciones de los parámetros idénticos (de ese proceso) a las estimaciones de los parámetros de los datos. Por ejemplo, sea  $\phi$  la hipótesis de que  $\bar{Z}$  es generado por ruido linealmente filtrado (independiente e idénticamente distribuido). Los sustitutos para  $\bar{Z}$  pueden ser generados por estimación del mejor modelo lineal (a partir de  $\bar{Z}$ ) y generando realizaciones de este modelo asumido. Estos sustitutos serán realizaciones no limitadas. Limitadas realizaciones sustitutas pueden ser generadas al aleatorizar las fases de la transformada de Fourier de los datos (esto produce un conjunto de datos aleatorios con los mismos espectros de potencia, y por lo tanto con la misma autocorrelación que los datos originales).

Es decir, así como la generación de sustitutos son realizaciones típicas de un modelo de los datos, hay que asegurarse de que los sustitutos son realizaciones de un proceso que da estimaciones de los parámetros idénticos (de ese proceso) a las estimaciones de los parámetros de los datos. Por ejemplo, sea  $\phi$  la hipótesis de que  $\bar{Z}$  es generado por ruido linealmente filtrado (independiente e idénticamente distribuido). Los sustitutos para  $\bar{Z}$  pueden ser generados por estimación del mejor modelo lineal (a partir de  $\bar{Z}$ ) y generando realizaciones de este modelo asumido. Estos sustitutos serán realizaciones no limitadas. Limitadas realizaciones sustitutas pueden ser generadas al aleatorizar las fases de la transformada de Fourier de los datos (esto produce un conjunto de datos aleatorios con los mismos espectros de potencia, y por lo tanto con la misma autocorrelación que los datos originales).

## 4.6.2. *Sustitutos Lineales*

### 4.6.2.1. **Algoritmo de Aleatorización a Pequeña Escala (SSS)<sup>10</sup>**

Para investigar si las fluctuaciones irregulares, son aleatorias o deterministas lo que se busca es destruir las estructuras locales o correlaciones en las fluctuaciones irregulares (variabilidad de corto plazo) y preservar los comportamientos globales. Para generar datos sustitutos que pueden cumplir esas condiciones conflictivas, se baraja el índice de un conjunto de datos dados en una «pequeña» (local) escala: esto está en contraste con el método sustituto RW, donde el índice se baraja en «gran» (global) escala y cualquier estructura de los datos originales es completamente destruida (Small en Ospina, et al., 2011; y Torkamani, Mahmoodzadeh, Pourroostaei y Lucas en Ospina, et al., 2011).

T. Nakamura y M. Small (Ospina, et al., 2011) proponen un nuevo método para la generación de datos sustitutos denominado, sustitutos de aleatorización a pequeña escala (SSS), la hipótesis nula dirigida por este algoritmo es que las fluctuaciones irregulares son variables aleatorias distribuidas independientemente. El método SSS es básicamente una extensión del algoritmo de Theiler para la generación de sustitutos aleatorios i.i.d. para los datos no estacionarios.

Los datos sustitutos son generados por el siguiente proceso [41]:

1. Empezar con la serie original  $x[t]$ ,  $t = 1, \dots, N$ .
2. Obtener el índice  $i[t]$  de  $x[t]$ , (Eso es,  $i[t] = t$ )
3. Obtener  $i'[t] = i[t] + Ag[t]$  donde  $g[t]$  son números aleatorios gaussianos y  $A$  es su amplitud.
4. Ordenar  $i'[t]$  y dejar que el índice  $i'[t]$  sea  $\hat{i}[t]$
5. Los datos sustitutos de  $x[t]$  son dados por  $x'[t] = x[\hat{i}[t]]$

Los datos sustitutos del algoritmo SSS son influenciados principalmente por la amplitud  $A$ . En términos generales, si  $A$  es demasiado pequeño, los datos se barajan muy poco o nada, y entonces los sustitutos son casi idénticos a los datos originales. Por el contrario si  $A$  es demasiado grande, los datos se barajan mucho, y los datos sustitutos son casi aleatorios al igual

---

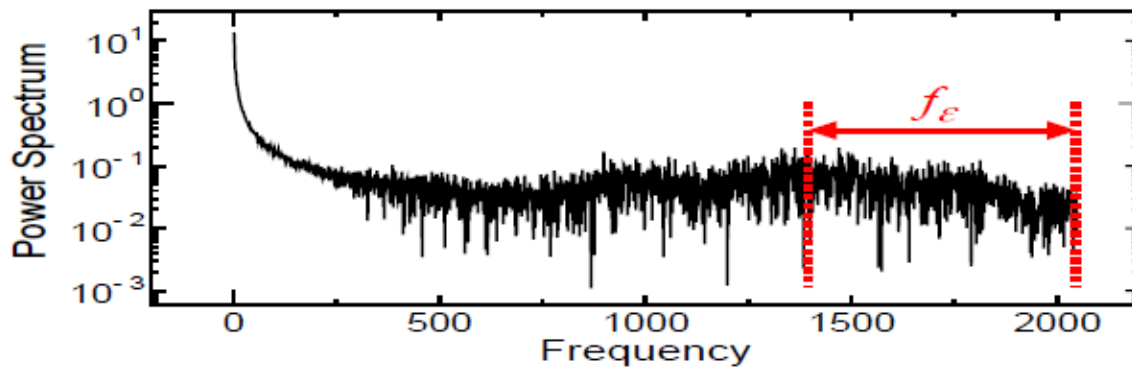
<sup>10</sup>Los supuestos sobre los que se basa este apartado corresponden a Tovar L., J. E. (2011). *Desarrollo de una metodología que permita La caracterización de la dinámica intrínseca de series de tiempo financieras, empleando el método de los datos Sustitutos (aplicación al precio del Commodity de la energía eléctrica)*.

que los datos sustitutos del algoritmo RW. En otras palabras, los valores más pequeños son mejores para la preservación de cualquier estructura y la correlación en los datos originales, sin embargo no son eficaces en la destrucción de las estructuras locales. Los valores más altos son mejores para destruir cualquier estructura y la correlación de los datos originales, sin embargo, no son eficaces en la preservación del comportamiento de largo plazo.

La influencia del  $A$  se ha investigado y se encuentra que  $A = 1.0$  es apropiado y más que adecuado para casi todos los propósitos (Theiler, Galdrikian, Longtin, Eubank y Farmer en Ospina, et al., 2011). Además, se observa que los sustitutos son muy similares a los datos originales.

#### 4.6.2.2. Algoritmo de la Transformada Truncada de Fourier (TFT)

Para investigar si existe no linealidad en las fluctuaciones irregulares de series temporales que presentan tendencias de largo plazo, lo que se quiere es destruir la no linealidad y preservar los comportamientos globales. Cuando los datos muestran fluctuaciones irregulares y tendencias de largo plazo el espectro de potencia suele ser como el de la siguiente figura [16].



Gráfica 1. Espectro de la potencia de una serie temporal.

Esta figura indica que los datos tienen grandes picos de potencia en el dominio de baja frecuencia, mientras que, en el dominio de alta frecuencia la potencia es casi como ruido blanco. Por lo tanto de esta figura se concluye que el dominio de alta frecuencia está probablemente dominada por fluctuaciones irregulares. Esto implica que si se aleatorizan las fases en el dominio de alta frecuencia  $f_e$  la influencia de la tendencia de largo plazo no será significativa. Por lo tanto, al aleatorizar las fases sólo en el dominio de alta frecuencia y no alterar las fases de baja frecuencia se conservará la tendencia de los datos mientras se destruye la no linealidad. Este

nuevo método se denomina Transformada Truncada de Fourier (TFT). Dado que algunas fases están intactas, los datos sustitutos del algoritmo TFT aún puede tener no linealidad. Sin embargo, es posible discriminar entre los datos lineales y no lineales porque el principio de superposición es válido sólo para datos lineal.

Por lo tanto, la hipótesis nula (NH) dirigida por nuestro algoritmo es que los datos de prueba son producidos por un proceso lineal estocástico en la forma de un modelo ARMA con parámetros desconocidos, lo cual es un filtro lineal del ruido i.d [16, 40].

Los datos sustitutos son generados por el siguiente proceso:

1. Empezar con la serie original  $x[t]$ ;  $t: 1, \dots, N$ .
2. Calcular  $z[n]$ , la transformada de Fourier de  $x[t]$ .
3. Aleatorizar las fases:  $z'[n] = z[n]e^{i\Phi[n]}$ . Donde,
4.  $\Phi[n] \in N(0, \pi)$  si  $n > f_c$
5.  $\Phi[n] = 0$  si  $n \leq f_c$
6. Simetrizar  $z'[n]$ .
7. Obtener  $x'[t]$ , la transformada inversa de Fourier  $z'[n]$ .
8.  $x'[t]$ , es la serie de datos sustitutos de  $x[t]$ .

6.3.3 Cuadro Resumen de los Resultados de las Hipótesis Nulas (Algoritmos para series con tendencia).

	Estadístico AMI	Estadístico CMP
Algoritmo SSS	Rechazo	Rechazo
Algoritmo TFT	Rechazo	Rechazo

**Tabla 1.** Decisión de rechazo o no rechazo de las hipótesis nulas a partir de los estadísticos (algoritmos para series con tendencia).

En algunos casos en las series analizadas se encuentra evidencia de que las señales pueden tener comportamiento determinista. Estos hallazgos se pueden detectar con mayor facilidad con las series de mayor longitud y respalda las premisas del análisis técnico (Ospina, et al., 2011).

## 5. Desarrollo metodológico

El método empleado es el siguiente:

### 5.1. Metodología de desarrollo

#### 5.1.1. *Fase 1: Análisis de la serie temporal y definición del proceso ARMA*

- Análisis gráfico de la serie de la demanda para evidenciar dependencia temporal, condición necesaria para la selección del modelo ARMA.
- Cálculo del criterio de Akaike para una combinación de modelos ARMA de ordenes  $p$  de 0-10 y  $q$  de 0-10.
- Selección del modelo ARMA, teniendo en cuenta el menor valor del criterio de Akaike.
- Estimación de los parámetros del modelo ARMA seleccionado, mediante el uso del filtro de Kalman (máxima verosimilitud exacta).
- Construcción de la ecuación pronóstico del modelo ARMA seleccionado

#### 5.1.2. *Fase 2: Análisis de la serie temporal y definición del proceso GARCH*

- Análisis de los residuos cuadrados de los efectos ARCH, para la selección del orden del modelo GARCH y cálculo del criterio de Akaike para una combinación de modelos AR ordenes  $p$  de 0-10 y posterior selección del modelo AR
- Cálculo de los retornos de la serie de la demanda, la estimación de su media y varianza condicional.
- Estimación de los parámetros del modelo seleccionado.
- Construcción de las ecuaciones de pronósticos del modelo AR( $p$ ) GARCH ( $p,q$ ) seleccionado.

- Presentación de resultados y análisis para una serie de demanda.

### **5.1.3. Fase 3: Metodología de validación (Surrogate data)**

- Validación de la dinámica de la serie de la demanda a través del método de los datos sustitutos.

## **5.2. Descripción de las bases datos utilizadas**

En este trabajo se analiza la serie de la demanda del color café de Coats Cadena Andina S.A., que cuenta con 104 mediciones, que han sido obtenidas desde ARP SAP y corresponden a observaciones de los años 2010,2011 y 2012.

## **5.3. Software**

Para el desarrollo de este trabajo se hizo uso de los siguientes softwares: MatLab R2010a, Microsoft Visual Studio y Gretel.

## **5.4. Algoritmos**

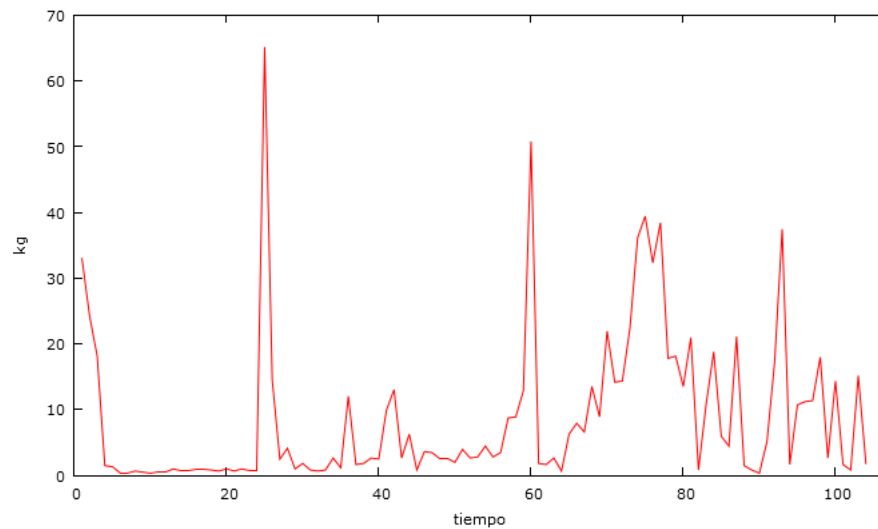
Los algoritmos utilizados en MatLab durante el desarrollo de este trabajo se presentan en los anexos (1 a 7).

## 6. Resultados y análisis

En esta sección se muestra los resultados del ajuste de los modelos ARMA y GARCH a la serie de demanda del color café de la empresa Coats Cadena Andina S.A.

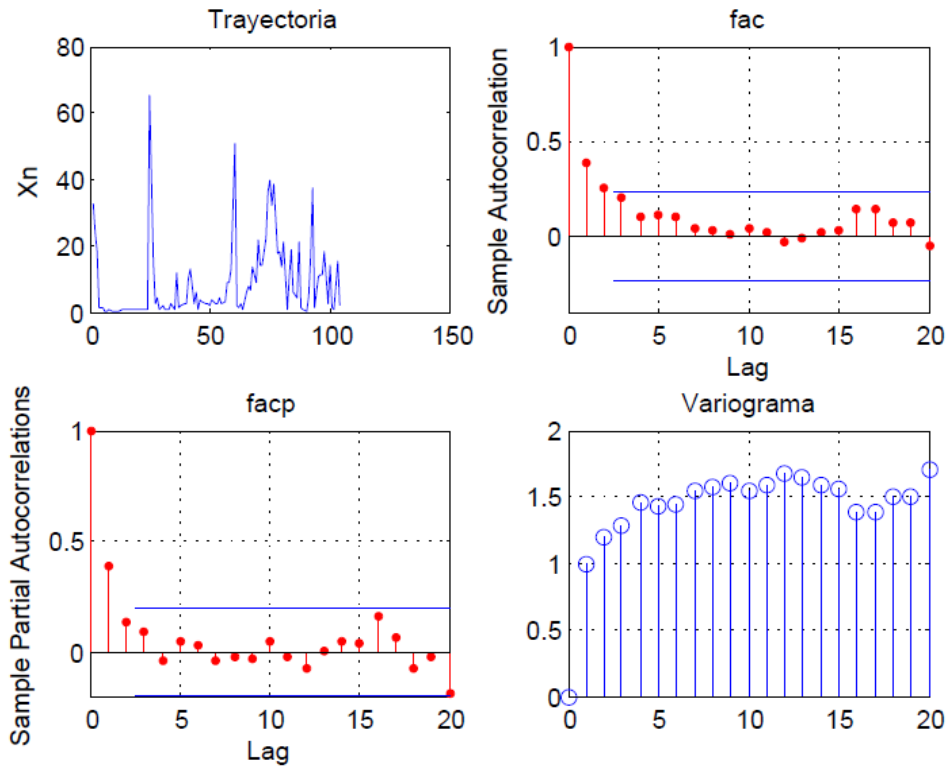
### 6.1. Selección del Modelo ARMA

#### 6.1.1. Modelo 1 ARMA (4,4)



**Gráfica 2.** Serie de demanda del color café en Coats Cadena Andina S.A.





**Gráfica 3.** Serie de la demanda del color café en Coats Cadena Andina S.A., su función de auto correlación, función de auto correlación parcial y variograma.

En la gráfica trayectoria, la serie de la demanda es volátil por tramos, y tiene unos picos muy altos. En la función autocorrelación y autocorrelación parcial, se observa que la correlación no es cero para algunos tiempos de retardo. Esto quiere decir que hay evidencia de dependencia temporal, por lo tanto se puede ajustar un modelo ARMA a la serie.

Para la selección del modelo ARMA se utilizó el criterio de selección de Akaike (AIC). En la tabla 2 se muestran los resultados del cálculo de este criterio para una combinación de modelos ARMA de ordenes  $p$  de 0-10 y  $q$  de 0-10.

AIC : 5.12632 NEWAIC : 5.12595 (p,q)=(0,1) aic=5.12632 fpe=168.365	AIC : 4.89813 NEWAIC : 4.89398 (p,q)=(2,9) aic=4.89813 fpe=131.335	AIC : 4.91814 NEWAIC : 4.90746 (p,q)=(5,6) aic=4.91814 fpe=133.832	AIC : 4.88337 NEWAIC : 4.86574 (p,q)=(8,3) aic=4.88337 fpe=129.093
AIC : 5.06485 NEWAIC : 5.0637 (p,q)=(0,2) aic=5.06485 fpe=158.235	AIC : 4.91604 NEWAIC : 4.91152 (p,q)=(2,10) aic=4.91604 fpe=133.227	AIC : 4.93741 NEWAIC : 4.92575 (p,q)=(5,7) aic=4.93741 fpe=135.917	AIC : 4.92496 NEWAIC : 4.90573 (p,q)=(8,4) aic=4.92496 fpe=134.035
AIC : 4.9421 NEWAIC : 4.94097 (p,q)=(0,3) aic=4.9421 fpe=139.831	AIC : 4.93518 NEWAIC : 4.93227 (p,q)=(3,0) aic=4.93518 fpe=138.853	AIC : 4.9543 NEWAIC : 4.94167 (p,q)=(5,8) aic=4.9543 fpe=137.671	AIC : 4.92281 NEWAIC : 4.90197 (p,q)=(8,5) aic=4.92281 fpe=133.173
AIC : 4.94742 NEWAIC : 4.94591	AIC : 4.79521 NEWAIC : 4.79292	AIC : 4.95001 NEWAIC : 4.93641	AIC : 4.92933 NEWAIC : 4.90689

(p,q)=(0,4) aic=4.94742 fpe=140.4	(p,q)=(3,1) aic=4.79521 fpe=120.57	(p,q)=(5,9) aic=4.95001 fpe=136.489	(p,q)=(8,6) aic=4.92933 fpe=133.433
AIC : 4.9455 NEWAIC : 4.94362	AIC : 4.92674 NEWAIC : 4.9198	AIC : 4.99836 NEWAIC : 4.9838	AIC : 4.97445 NEWAIC : 4.95041
(p,q)=(0,5) aic=4.9455 fpe=139.909	(p,q)=(3,2) aic=4.92674 fpe=137.245	(p,q)=(5,10) aic=4.99836 fpe=142.598	(p,q)=(8,7) aic=4.97445 fpe=138.918
AIC : 4.93303 NEWAIC : 4.93076	AIC : 4.81046 NEWAIC : 4.80703	AIC : 4.88934 NEWAIC : 4.88228	AIC : 5.04431 NEWAIC : 5.01867
(p,q)=(0,6) aic=4.93303 fpe=137.91	(p,q)=(3,3) aic=4.81046 fpe=121.987	(p,q)=(6,0) aic=4.88934 fpe=131.948	(p,q)=(8,8) aic=5.04431 fpe=148.214
AIC : 4.9451 NEWAIC : 4.94246	AIC : 4.84229 NEWAIC : 4.83829	AIC : 4.86849 NEWAIC : 4.86025	AIC : 5.05573 NEWAIC : 5.02848
(p,q)=(0,7) aic=4.9451 fpe=139.277	(p,q)=(3,4) aic=4.84229 fpe=125.648	(p,q)=(6,1) aic=4.86849 fpe=128.917	(p,q)=(8,9) aic=5.05573 fpe=149.119
AIC : 4.94788 NEWAIC : 4.94486	AIC : 4.84213 NEWAIC : 4.83757	AIC : 4.86183 NEWAIC : 4.85241	AIC : 4.89174 NEWAIC : 4.86289
(p,q)=(0,8) aic=4.94788 fpe=139.314	(p,q)=(3,5) aic=4.84213 fpe=125.308	(p,q)=(6,2) aic=4.86183 fpe=127.714	(p,q)=(8,10) aic=4.89174 fpe=125.863
AIC : 4.96241 NEWAIC : 4.95902	AIC : 4.86187 NEWAIC : 4.85672	AIC : 4.90858 NEWAIC : 4.89799	AIC : 4.98108 NEWAIC : 4.96468
(p,q)=(0,9) aic=4.96241 fpe=140.958	(p,q)=(3,6) aic=4.86187 fpe=127.441	(p,q)=(6,3) aic=4.90858 fpe=133.424	(p,q)=(9,0) aic=4.98108 fpe=143.325
AIC : 4.96817 NEWAIC : 4.9644	AIC : 4.89196 NEWAIC : 4.88625	AIC : 4.89721 NEWAIC : 4.88543	AIC : 4.96019 NEWAIC : 4.94197
(p,q)=(0,10) aic=4.96817 fpe=141.335	(p,q)=(3,7) aic=4.89196 fpe=130.923	(p,q)=(6,4) aic=4.89721 fpe=131.478	(p,q)=(9,1) aic=4.96019 fpe=139.871
AIC : 4.96631 NEWAIC : 4.96593	AIC : 4.9104 NEWAIC : 4.90411	AIC : 4.85866 NEWAIC : 4.84571	AIC : 4.94768 NEWAIC : 4.92764
(p,q)=(1,0) aic=4.96631 fpe=143.469	(p,q)=(3,8) aic=4.9104 fpe=132.906	(p,q)=(6,5) aic=4.85866 fpe=126.052	(p,q)=(9,2) aic=4.94768 fpe=137.607
AIC : 4.88954 NEWAIC : 4.8884	AIC : 4.90957 NEWAIC : 4.90271	AIC : 4.8129 NEWAIC : 4.79878	AIC : 4.92762 NEWAIC : 4.90575
(p,q)=(1,1) aic=4.88954 fpe=132.791	(p,q)=(3,9) aic=4.90957 fpe=132.309	(p,q)=(6,6) aic=4.8129 fpe=119.948	(p,q)=(9,3) aic=4.92762 fpe=134.321
AIC : 4.88949 NEWAIC : 4.88718	AIC : 4.94993 NEWAIC : 4.9425	AIC : 4.9674 NEWAIC : 4.9521	AIC : 4.96101 NEWAIC : 4.93732
(p,q)=(1,2) aic=4.88949 fpe=132.656	(p,q)=(3,10) aic=4.94993 fpe=137.218	(p,q)=(6,7) aic=4.9674 fpe=139.409	(p,q)=(9,4) aic=4.96101 fpe=138.274
AIC : 4.80205 NEWAIC : 4.80055	AIC : 4.85671 NEWAIC : 4.85364	AIC : 4.95362 NEWAIC : 4.93714	AIC : 4.87781 NEWAIC : 4.85231
(p,q)=(1,3) aic=4.80205 fpe=121.405	(p,q)=(4,0) aic=4.85671 fpe=128.21	(p,q)=(6,8) aic=4.95362 fpe=136.896	(p,q)=(9,5) aic=4.87781 fpe=126.644
AIC : 4.79851 NEWAIC : 4.79663	AIC : 4.81475 NEWAIC : 4.8109	AIC : 4.96304 NEWAIC : 4.94538	AIC : 4.96338 NEWAIC : 4.93605
(p,q)=(1,4) aic=4.79851 fpe=120.784	(p,q)=(4,1) aic=4.81475 fpe=122.738	(p,q)=(6,9) aic=4.96304 fpe=137.549	(p,q)=(9,6) aic=4.96338 fpe=137.28
AIC : 4.81811 NEWAIC : 4.81584	AIC : 4.82937 NEWAIC : 4.82476	AIC : 4.93978 NEWAIC : 4.92094	AIC : 5.00036 NEWAIC : 4.97122
(p,q)=(1,5) aic=4.81811 fpe=122.939	(p,q)=(4,2) aic=4.82937 fpe=124.301	(p,q)=(6,10) aic=4.93978 fpe=133.73	(p,q)=(9,7) aic=5.00036 fpe=141.717
AIC : 4.83655 NEWAIC : 4.83391	AIC : 4.84802 NEWAIC : 4.84263	AIC : 4.91907 NEWAIC : 4.90936	AIC : 5.07454 NEWAIC : 5.04356
(p,q)=(1,6) aic=4.83655 fpe=124.95	(p,q)=(4,3) aic=4.84802 fpe=126.349	(p,q)=(7,0) aic=4.91907 fpe=135.58	(p,q)=(9,8) aic=5.07454 fpe=151.802
AIC : 4.84913 NEWAIC : 4.84611	AIC : 4.78235 NEWAIC : 4.77619	AIC : 4.89965 NEWAIC : 4.88855	AIC : 5.03696 NEWAIC : 5.00417
(p,q)=(1,7) aic=4.84913 fpe=126.214	(p,q)=(4,4) aic=4.78235 fpe=118.009	(p,q)=(7,1) aic=4.89965 fpe=132.605	(p,q)=(9,9) aic=5.03696 fpe=145.378
AIC : 4.86868 NEWAIC : 4.86529	AIC : 4.82012 NEWAIC : 4.8132	AIC : 4.90202 NEWAIC : 4.88953	AIC : 4.98601 NEWAIC : 4.95139
(p,q)=(1,8) aic=4.86868 fpe=128.346	(p,q)=(4,5) aic=4.82012 fpe=122.198	(p,q)=(7,2) aic=4.90202 fpe=132.512	(p,q)=(9,10) aic=4.98601 fpe=137.342
AIC : 4.88828 NEWAIC : 4.88451	AIC : 4.90787 NEWAIC : 4.90018	AIC : 4.92697 NEWAIC : 4.91309	AIC : 5.00491 NEWAIC : 4.98445
(p,q)=(1,9) aic=4.88828 fpe=130.483	(p,q)=(4,6) aic=4.90787 fpe=132.979	(p,q)=(7,3) aic=4.92697 fpe=135.401	(p,q)=(10,0) aic=5.00491 fpe=146.211
AIC : 4.90667 NEWAIC : 4.90253	AIC : 4.91949 NEWAIC : 4.91102	AIC : 4.877 NEWAIC : 4.86174	AIC : 4.99215 NEWAIC : 4.96964

(p,q)=(1,10) aic=4.90667 fpe=132.462	(p,q)=(4,7) aic=4.91949 fpe=134.067	(p,q)=(7,4) aic=4.877 fpe=128.331	(p,q)=(10,1) aic=4.99215 fpe=143.797
AIC : 4.9306 NEWAIC : 4.92945	AIC : 4.86793 NEWAIC : 4.8587	AIC : 4.89569 NEWAIC : 4.87903	AIC : 5.00797 NEWAIC : 4.98342
(p,q)=(2,0) aic=4.9306 fpe=138.356	(p,q)=(4,8) aic=4.86793 fpe=126.854	(p,q)=(7,5) aic=4.89569 fpe=130.235	(p,q)=(10,2) aic=5.00797 fpe=145.48
AIC : 4.87637 NEWAIC : 4.87406	AIC : 5.04052 NEWAIC : 5.03052	AIC : 4.83319 NEWAIC : 4.81515	AIC : 4.98797 NEWAIC : 4.96138
(p,q)=(2,1) aic=4.87637 fpe=130.926	(p,q)=(4,9) aic=5.04052 fpe=150.149	(p,q)=(7,6) aic=4.83319 fpe=121.829	(p,q)=(10,3) aic=4.98797 fpe=141.964
AIC : 4.88514 NEWAIC : 4.88125	AIC : 4.95831 NEWAIC : 4.94754	AIC : 4.96577 NEWAIC : 4.94634	AIC : 4.96159 NEWAIC : 4.93295
(p,q)=(2,2) aic=4.88514 fpe=131.899	(p,q)=(4,10) aic=4.95831 fpe=137.713	(p,q)=(7,7) aic=4.96577 fpe=138.479	(p,q)=(10,4) aic=4.96159 fpe=137.612
AIC : 4.90654 NEWAIC : 4.90066	AIC : 4.86967 NEWAIC : 4.86481	AIC : 4.98106 NEWAIC : 4.96024	AIC : 4.91471 NEWAIC : 4.88402
(p,q)=(2,3) aic=4.90654 fpe=134.514	(p,q)=(5,0) aic=4.86967 fpe=129.656	(p,q)=(7,8) aic=4.98106 fpe=139.946	(p,q)=(10,5) aic=4.91471 fpe=130.654
AIC : 4.81901 NEWAIC : 4.81674	AIC : 4.83931 NEWAIC : 4.83348	AIC : 4.99928 NEWAIC : 4.97708	AIC : 4.87401 NEWAIC : 4.84128
(p,q)=(2,4) aic=4.81901 fpe=123.049	(p,q)=(5,1) aic=4.83931 fpe=125.525	(p,q)=(7,9) aic=4.99928 fpe=141.81	(p,q)=(10,6) aic=4.87401 fpe=124.783
AIC : 4.81476 NEWAIC : 4.81212	AIC : 4.8567 NEWAIC : 4.8499	AIC : 4.88458 NEWAIC : 4.86099	AIC : 4.98242 NEWAIC : 4.94764
(p,q)=(2,5) aic=4.81476 fpe=122.257	(p,q)=(5,2) aic=4.8567 fpe=127.429	(p,q)=(7,10) aic=4.88458 fpe=125.782	(p,q)=(10,7) aic=4.98242 fpe=138.304
AIC : 4.85398 NEWAIC : 4.85096	AIC : 4.83813 NEWAIC : 4.83036	AIC : 4.94923 NEWAIC : 4.93641	AIC : 4.91219 NEWAIC : 4.87536
(p,q)=(2,6) aic=4.85398 fpe=126.828	(p,q)=(5,3) aic=4.83813 fpe=124.752	(p,q)=(8,0) aic=4.94923 fpe=139.312	(p,q)=(10,8) aic=4.91219 fpe=128.182
AIC : 4.86073 NEWAIC : 4.85734	AIC : 4.8784 NEWAIC : 4.86966	AIC : 4.92843 NEWAIC : 4.91401	AIC : 4.9836 NEWAIC : 4.94473
(p,q)=(2,7) aic=4.86073 fpe=127.33	(p,q)=(5,4) aic=4.8784 fpe=129.494	(p,q)=(8,1) aic=4.92843 fpe=136.017	(p,q)=(10,9) aic=4.9836 fpe=136.846
AIC : 4.86372 NEWAIC : 4.85995	AIC : 4.90307 NEWAIC : 4.89336	AIC : 4.93114 NEWAIC : 4.91511	AIC : 4.88196 NEWAIC : 4.84104
(p,q)=(2,8) aic=4.86372 fpe=127.318	(p,q)=(5,5) aic=4.90307 fpe=132.297	(p,q)=(8,2) aic=4.93114 fpe=135.917	(p,q)=(10,10) aic=4.88196 fpe=122.851

**Tabla 2.** Criterio AIC para la selección del mejor modelo ARMA.

En la tabla anterior, en la celda sombreada se observa que el modelo que presenta menor criterio de Akaike es el modelo ARMA(4,4) con un valor AIC de 4.78235 (salida de MatLab), por lo tanto a continuación se muestran los resultados de los parámetros estimados para este modelo.

Coeficiente	Desv.	Típica	z	Valor p	
const	8,96286	1,24051	7,225	5,01E-13	***
phi_1	0,334104	0,167732	1,992	0,0464	**
phi_2	0,658079	0,0799585	8,23	1,87E-16	***
phi_3	0,705887	0,0784879	8,994	2,39E-19	***
phi_4	-0,727753	0,166734	-4,365	1,27E-05	***
theta_1	0,0304482	0,232949	0,1307	0,896	
theta_2	-0,550352	0,142125	-3,872	0,0001	***
theta_3	-0,851077	0,114864	-7,409	1,27E-13	***
theta_4	0,370981	0,242497	1,53	0,1261	

**Tabla 3.** Parámetros estimados para modelo ARMA(4,4).

<b>Media de la vble. dep.</b>	8,745	<b>D.T. de la vble. dep.</b>	11,81271
<b>Media innovaciones</b>	-0,483495	<b>D.T. innovaciones</b>	10,14609
<b>Log-verosimilitud</b>	-390,8964	<b>Criterio de Akaike</b>	4,78235
<b>Criterio de Schwarz</b>	828,2367	<b>Crit. de Hannan-Quinn</b>	812,506

**Tabla 4.** Estadísticos del ajuste de los parámetros del modelo ARMA(4,4).

La tabla 3 muestra los parámetros del modelo ARMA (4,4) estimados mediante el uso del filtro de Kalman (Máxima verosimilitud exacta). Además se pueden apreciar los parámetros significativos del modelo, señalados con asteriscos (Salida Gretel). En la tabla 4 se observa el cálculo de diferentes estadísticos que miden el ajuste de parámetros al modelo. El criterio de Akaike es el tenido en cuenta para la selección del mejor modelo, como se evidencia en la tabla 2.

<b>Observaciones</b>	<b>Color Café</b>	<b>Estimada</b>	<b>Residuo</b>
1	33	8,963	24,037
2	24,255	18,512	5,743
3	18,15	17,799	0,351
4	1,485	16,227	-14,742
5	1,32	7,972	-6,652
6	0,33	6,412	-6,082
7	0,33	4,023	-3,693
8	0,66	3,834	-3,174
9	0,495	4,374	-3,879
10	0,33	3,332	-3,002
11	0,495	4,207	-3,712
12	0,495	4,045	-3,55
13	0,99	3,621	-2,631
14	0,825	4,787	-3,962
15	0,825	4,07	-3,245
16	0,99	4,379	-3,389
17	0,99	4,923	-3,933
18	0,825	4,202	-3,377
19	0,66	4,931	-4,271
20	0,99	4,717	-3,727
21	0,66	4,464	-3,804
22	0,99	5,178	-4,188
23	0,825	4,721	-3,896
24	0,825	4,881	-4,056

25	65,175	5,356	59,819*
26	14,52	27,081	-12,561
27	2,475	17,409	-14,934
28	4,125	13,37	-9,245
29	0,99	5,102	-4,112
30	1,815	6,971	-5,156
31	0,825	6,936	-6,111
32	0,66	2,178	-1,518
33	0,825	7,561	-6,736
34	2,64	4,043	-1,403
35	1,155	4,242	-3,087
36	12,045	8,096	3,949
37	1,65	6,801	-5,151
38	1,815	7,408	-5,593
39	2,64	8,024	-5,384
40	2,475	3,245	-0,77
41	9,9	8,51	1,39
42	13,035	8,563	4,472
43	2,64	9,003	-6,363
44	6,27	10,974	-4,704
45	0,825	6,312	-5,487
46	3,63	6,251	-2,621
47	3,465	9,006	-5,541
48	2,64	3,938	-1,298
49	2,64	8,543	-5,903
50	1,98	6,852	-4,872
51	3,96	4,19	-0,23
52	2,64	9,822	-7,182
53	2,805	5,074	-2,269
54	4,455	6,589	-2,134
55	2,805	9,554	-6,749
56	3,465	4,497	-1,032
57	8,745	8,944	-0,199
58	8,91	9,52	-0,61
59	12,87	7,925	4,945
60	50,82	14,311	36,509*
61	1,815	24,188	-22,373
62	1,65	12,335	-10,685
63	2,64	12,253	-9,613
64	0,66	3,594	-2,934
65	6,27	7,955	-1,685

66	7,92	9,068	-1,148
67	6,6	5,43	1,17
68	13,53	12,609	0,921
69	8,91	9,881	-0,971
70	21,945	9,144	12,801
71	14,19	18,693	-4,503
72	14,355	9,966	4,389
73	22,44	15,063	7,377
74	36,135	17,362	18,773
75	39,435	18,099	21,336
76	32,34	28,417	3,923
77	38,445	21,871	16,574
78	17,82	23,361	-5,541
79	18,15	20,968	-2,818
80	13,53	12,334	1,196
81	20,955	13,45	7,505
82	0,825	15,872	-15,047
83	10,725	4,227	6,498
84	18,81	11,904	6,906
85	5,94	10,723	-4,783
86	4,455	6,918	-2,463
87	21,12	10,314	10,806
88	1,485	8,816	-7,331
89	0,825	7,732	-6,907
90	0,33	7,172	-6,842
91	5,115	0,166	4,949
92	17,16	8,605	8,555
93	37,455	9,684	27,771*
94	1,65	16,899	-15,249
95	10,725	13,068	-2,343
96	11,22	7,204	4,016
97	11,385	9,057	2,328
98	17,985	12,113	5,872
99	2,64	8,581	-5,941
100	14,355	9,033	5,322
101	1,65	10,553	-8,903
102	0,825	2,975	-2,15
103	15,18	8,06	7,12
104	1,815	7,273	-5,458

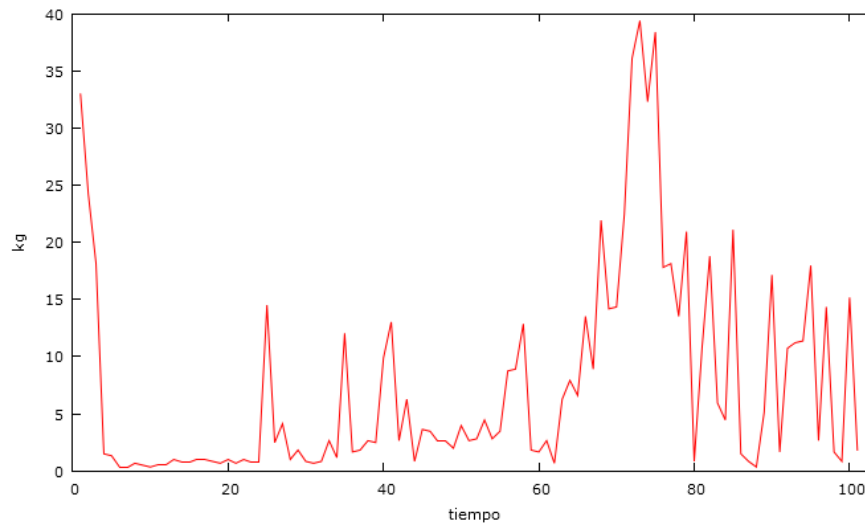
**Tabla 5.** Variable observada, variable estimada y residuo del color café en Coats Cadena Andina S.A.

Estadístico de evaluación de la predicción	
Error medio	-0,48349
Error cuadrático medio	108,59
Raíz del error cuadrático medio	10,42
Error absoluto medio	6,7782
Porcentaje de error medio	-284,13
Porcentaje de error absoluto medio	308,67

**Tabla 6.** Errores del modelo ARMA(4,4).

La tabla 5 muestra los datos de la variable original, variable estimada y residuo de la demanda del color café de Coats Cadena Andina S.A. Los valores en asterisco denotan valores atípicos ya que son superiores a 2,5 desviaciones típicas (Salida Gretel). Éstos se eliminan de la serie y se procede de nuevo con el proceso de modelado. En la tabla 6 se aprecian los errores del modelo.

### 6.1.2. Modelo 2 ARMA (8,9)



**Gráfica 4.** Serie de demanda del color café en Coats Cadena Andina S.A., sin datos atípicos.

Para la selección del nuevo modelo ARMA se utiliza el criterio de Akaike (AIC). En la tabla 7 se muestran los resultados del cálculo de este criterio para una combinación de modelos ARMA de ordenes  $p$  de 0-10 y  $q$  de 0-10.

AIC : 4.5007 NEWAIC : 4.5003	AIC : 3.8057 NEWAIC : 3.8033	AIC : 3.73868 NEWAIC : 3.7297	AIC : 3.64042 NEWAIC : 3.62041	AIC : 3.77401 NEWAIC : 3.7527
(p,q)=(0,1) aic=4.5007 fpe=90.0622	(p,q)=(2,4) aic=3.8057 fpe=44.652	(p,q)=(4,7) aic=3.73868 fpe=41.1117	(p,q)=(6,10) aic=3.64042 fpe=36.3752	(p,q)=(9,2) aic=3.77401 fpe=42.491

AIC : 4.17893 NEWAIC : 4.17813 (p,q)=(0,2) aic=4.17893 fpe=65.244	AIC : 3.76504 NEWAIC : 3.76224 (p,q)=(2,5) aic=3.76504 fpe=42.7723	AIC : 3.72878 NEWAIC : 3.71898 (p,q)=(4,8) aic=3.72878 fpe=40.5459	AIC : 3.78093 NEWAIC : 3.77061 (p,q)=(7,0) aic=3.78093 fpe=43.416	AIC : 3.78881 NEWAIC : 3.76556 (p,q)=(9,3) aic=3.78881 fpe=42.9372
AIC : 3.9667 NEWAIC : 3.9655 (p,q)=(0,3) aic=3.9667 fpe=52.7166	AIC : 3.82146 NEWAIC : 3.81826 (p,q)=(2,6) aic=3.82146 fpe=45.1348	AIC : 3.71295 NEWAIC : 3.70234 (p,q)=(4,9) aic=3.71295 fpe=39.7411	AIC : 3.80209 NEWAIC : 3.79029 (p,q)=(7,1) aic=3.80209 fpe=44.2149	AIC : 3.74529 NEWAIC : 3.72011 (p,q)=(9,4) aic=3.74529 fpe=40.9184
AIC : 3.9495 NEWAIC : 3.9479 (p,q)=(0,4) aic=3.9495 fpe=51.7485	AIC : 3.79872 NEWAIC : 3.79512 (p,q)=(2,7) aic=3.79872 fpe=43.9899	AIC : 3.67511 NEWAIC : 3.66368 (p,q)=(4,10) aic=3.67511 fpe=38.0939	AIC : 3.71555 NEWAIC : 3.70228 (p,q)=(7,2) aic=3.71555 fpe=40.4176	AIC : 3.77107 NEWAIC : 3.74395 (p,q)=(9,5) aic=3.77107 fpe=41.7802
AIC : 3.86574 NEWAIC : 3.86374 (p,q)=(0,5) aic=3.86574 fpe=47.5111	AIC : 3.80648 NEWAIC : 3.80248 (p,q)=(2,8) aic=3.80648 fpe=44.1887	AIC : 3.72992 NEWAIC : 3.72476 (p,q)=(5,0) aic=3.72992 fpe=41.4647	AIC : 3.70294 NEWAIC : 3.6882 (p,q)=(7,3) aic=3.70294 fpe=39.7687	AIC : 3.76617 NEWAIC : 3.73711 (p,q)=(9,6) aic=3.76617 fpe=41.3599
AIC : 3.84693 NEWAIC : 3.84453 (p,q)=(0,6) aic=3.84693 fpe=46.5315	AIC : 3.83973 NEWAIC : 3.83533 (p,q)=(2,9) aic=3.83973 fpe=45.5212	AIC : 3.73843 NEWAIC : 3.73225 (p,q)=(5,1) aic=3.73843 fpe=41.7299	AIC : 3.72891 NEWAIC : 3.71269 (p,q)=(7,4) aic=3.72891 fpe=40.6568	AIC : 3.77928 NEWAIC : 3.74829 (p,q)=(9,7) aic=3.77928 fpe=41.6769
AIC : 3.83764 NEWAIC : 3.83484 (p,q)=(0,7) aic=3.83764 fpe=45.9934	AIC : 3.64684 NEWAIC : 3.64204 (p,q)=(2,10) aic=3.64684 fpe=37.3927	AIC : 3.76624 NEWAIC : 3.75902 (p,q)=(5,2) aic=3.76624 fpe=42.8	AIC : 3.71883 NEWAIC : 3.70113 (p,q)=(7,5) aic=3.71883 fpe=40.0808	AIC : 3.73062 NEWAIC : 3.69768 (p,q)=(9,8) aic=3.73062 fpe=39.4701
AIC : 3.86154 NEWAIC : 3.85834 (p,q)=(0,8) aic=3.86154 fpe=46.9807	AIC : 3.78011 NEWAIC : 3.77829 (p,q)=(3,0) aic=3.78011 fpe=43.7419	AIC : 3.78683 NEWAIC : 3.77857 (p,q)=(5,3) aic=3.78683 fpe=43.5674	AIC : 3.76067 NEWAIC : 3.7415 (p,q)=(7,6) aic=3.76067 fpe=41.6071	AIC : 3.7746 NEWAIC : 3.73974 (p,q)=(9,9) aic=3.7746 fpe=40.9988
AIC : 3.77194 NEWAIC : 3.76834 (p,q)=(0,9) aic=3.77194 fpe=42.8273	AIC : 3.776 NEWAIC : 3.77357 (p,q)=(3,1) aic=3.776 fpe=43.5033	AIC : 3.7191 NEWAIC : 3.70981 (p,q)=(5,4) aic=3.7191 fpe=40.5869	AIC : 3.72692 NEWAIC : 3.70627 (p,q)=(7,7) aic=3.72692 fpe=40.0357	AIC : 3.73565 NEWAIC : 3.69885 (p,q)=(9,10) aic=3.73565 fpe=39.1875
AIC : 3.64765 NEWAIC : 3.64365 (p,q)=(0,10) aic=3.64765 fpe=37.699	AIC : 3.78492 NEWAIC : 3.78189 (p,q)=(3,2) aic=3.78492 fpe=43.8179	AIC : 3.70396 NEWAIC : 3.69365 (p,q)=(5,5) aic=3.70396 fpe=39.84	AIC : 3.74976 NEWAIC : 3.72764 (p,q)=(7,8) aic=3.74976 fpe=40.7558	AIC : 3.81733 NEWAIC : 3.79557 (p,q)=(10,0) aic=3.81733 fpe=44.5324
AIC : 3.8741 NEWAIC : 3.8737 (p,q)=(1,0) aic=3.8741 fpe=48.1296	AIC : 3.75864 NEWAIC : 3.755 (p,q)=(3,3) aic=3.75864 fpe=42.5934	AIC : 3.73776 NEWAIC : 3.72641 (p,q)=(5,6) aic=3.73776 fpe=41.0556	AIC : 3.67068 NEWAIC : 3.64708 (p,q)=(7,9) aic=3.67068 fpe=37.4587	AIC : 3.8326 NEWAIC : 3.80866 (p,q)=(10,1) aic=3.8326 fpe=45.0317
AIC : 3.75759 NEWAIC : 3.75679 (p,q)=(1,1) aic=3.75759 fpe=42.8109	AIC : 3.7583 NEWAIC : 3.75405 (p,q)=(3,4) aic=3.7583 fpe=42.4774	AIC : 3.71948 NEWAIC : 3.70711 (p,q)=(5,7) aic=3.71948 fpe=40.1501	AIC : 3.60733 NEWAIC : 3.58226 (p,q)=(7,10) aic=3.60733 fpe=34.9655	AIC : 3.80082 NEWAIC : 3.77471 (p,q)=(10,2) aic=3.80082 fpe=43.4302
AIC : 3.76064 NEWAIC : 3.75944 (p,q)=(1,2) aic=3.76064 fpe=42.9	AIC : 3.84045 NEWAIC : 3.8356 (p,q)=(3,5) aic=3.84045 fpe=45.9898	AIC : 3.64459 NEWAIC : 3.63118 (p,q)=(5,8) aic=3.64459 fpe=37.0929	AIC : 3.81361 NEWAIC : 3.79998 (p,q)=(8,0) aic=3.81361 fpe=44.7154	AIC : 3.71264 NEWAIC : 3.68435 (p,q)=(10,3) aic=3.71264 fpe=39.5766
AIC : 3.77013 NEWAIC : 3.76853 (p,q)=(1,3) aic=3.77013 fpe=43.2516	AIC : 3.76656 NEWAIC : 3.76111 (p,q)=(3,6) aic=3.76656 fpe=42.5854	AIC : 3.46427 NEWAIC : 3.44983 (p,q)=(5,9) aic=3.46427 fpe=30.8314	AIC : 3.82014 NEWAIC : 3.80481 (p,q)=(8,1) aic=3.82014 fpe=44.859	AIC : 3.69426 NEWAIC : 3.6638 (p,q)=(10,4) aic=3.69426 fpe=38.6612
AIC : 3.78949 NEWAIC : 3.78749 (p,q)=(1,4) aic=3.78949 fpe=44.023	AIC : 3.76339 NEWAIC : 3.75733 (p,q)=(3,7) aic=3.76339 fpe=42.3103	AIC : 3.64534 NEWAIC : 3.62986 (p,q)=(5,10) aic=3.64534 fpe=36.7732	AIC : 3.74329 NEWAIC : 3.72626 (p,q)=(8,2) aic=3.74329 fpe=41.3894	AIC : 3.71179 NEWAIC : 3.67915 (p,q)=(10,5) aic=3.71179 fpe=39.1362



AIC : 3.79679 NEWAIC : 3.79439 (p,q)=(1,5) aic=3.79679 fpe=44.2559	AIC : 3.82515 NEWAIC : 3.81848 (p,q)=(3,8) aic=3.82515 fpe=44.8438	AIC : 3.75791 NEWAIC : 3.7504 (p,q)=(6,0) aic=3.75791 fpe=42.5442	AIC : 3.75313 NEWAIC : 3.73439 (p,q)=(8,3) aic=3.75313 fpe=41.6335	AIC : 3.69886 NEWAIC : 3.66405 (p,q)=(10,6) aic=3.69886 fpe=38.4186
AIC : 3.80448 NEWAIC : 3.80168 (p,q)=(1,6) aic=3.80448 fpe=44.4929	AIC : 3.75191 NEWAIC : 3.74464 (p,q)=(3,9) aic=3.75191 fpe=41.5154	AIC : 3.77655 NEWAIC : 3.76779 (p,q)=(6,1) aic=3.77655 fpe=43.2349	AIC : 3.73995 NEWAIC : 3.71951 (p,q)=(8,4) aic=3.73995 fpe=40.9137	AIC : 3.73629 NEWAIC : 3.6993 (p,q)=(10,7) aic=3.73629 fpe=39.6511
AIC : 3.79985 NEWAIC : 3.79665 (p,q)=(1,7) aic=3.79985 fpe=44.1699	AIC : 3.65042 NEWAIC : 3.64254 (p,q)=(3,10) aic=3.65042 fpe=37.3536	AIC : 3.68477 NEWAIC : 3.67476 (p,q)=(6,2) aic=3.68477 fpe=39.3304	AIC : 3.72803 NEWAIC : 3.70588 (p,q)=(8,5) aic=3.72803 fpe=40.2448	AIC : 3.80649 NEWAIC : 3.76732 (p,q)=(10,8) aic=3.80649 fpe=42.2757
AIC : 3.81942 NEWAIC : 3.81582 (p,q)=(1,8) aic=3.81942 fpe=44.9097	AIC : 3.71533 NEWAIC : 3.71207 (p,q)=(4,0) aic=3.71533 fpe=40.94	AIC : 3.74417 NEWAIC : 3.73291 (p,q)=(6,3) aic=3.74417 fpe=41.6044	AIC : 3.85392 NEWAIC : 3.83007 (p,q)=(8,6) aic=3.85392 fpe=45.4237	AIC : 3.70645 NEWAIC : 3.6651 (p,q)=(10,9) aic=3.70645 fpe=38.0089
AIC : 3.79 NEWAIC : 3.786 (p,q)=(1,9) aic=3.79 fpe=43.4664	AIC : 3.72222 NEWAIC : 3.71814 (p,q)=(4,1) aic=3.72222 fpe=41.1508	AIC : 3.78522 NEWAIC : 3.77271 (p,q)=(6,4) aic=3.78522 fpe=43.1959	AIC : 3.7294 NEWAIC : 3.70385 (p,q)=(8,7) aic=3.7294 fpe=39.9012	AIC : 3.60866 NEWAIC : 3.56514 (p,q)=(10,10) aic=3.60866 fpe=34.2415
AIC : 3.62689 NEWAIC : 3.62249 (p,q)=(1,10) aic=3.62689 fpe=36.7943	AIC : 3.74087 NEWAIC : 3.73597 (p,q)=(4,2) aic=3.74087 fpe=41.8377	AIC : 3.70146 NEWAIC : 3.6877 (p,q)=(6,5) aic=3.70146 fpe=39.5743	AIC : 3.77051 NEWAIC : 3.74326 (p,q)=(8,8) aic=3.77051 fpe=41.3527	
AIC : 3.76414 NEWAIC : 3.76333 (p,q)=(2,0) aic=3.76414 fpe=43.0921	AIC : 3.76148 NEWAIC : 3.75577 (p,q)=(4,3) aic=3.76148 fpe=42.6051	AIC : 3.72413 NEWAIC : 3.70912 (p,q)=(6,6) aic=3.72413 fpe=40.3157	AIC : 3.421 NEWAIC : 3.39204 (p,q)=(8,9) aic=3.421 fpe=28.9914	
AIC : 3.77686 NEWAIC : 3.77566 (p,q)=(2,1) aic=3.77686 fpe=43.6018	AIC : 3.77554 NEWAIC : 3.76901 (p,q)=(4,4) aic=3.77554 fpe=43.089	AIC : 3.78376 NEWAIC : 3.7675 (p,q)=(6,7) aic=3.78376 fpe=42.6058	AIC : 3.44641 NEWAIC : 3.41575 (p,q)=(8,10) aic=3.44641 fpe=29.5632	
AIC : 3.72723 NEWAIC : 3.72563 (p,q)=(2,2) aic=3.72723 fpe=41.4351	AIC : 3.79515 NEWAIC : 3.7878 (p,q)=(4,5) aic=3.79515 fpe=43.8073	AIC : 3.58191 NEWAIC : 3.5644 (p,q)=(6,8) aic=3.58191 fpe=34.6563	AIC : 3.83611 NEWAIC : 3.81867 (p,q)=(9,0) aic=3.83611 fpe=45.5653	
AIC : 3.78617 NEWAIC : 3.78417 (p,q)=(2,3) aic=3.78617 fpe=43.877	AIC : 3.7314 NEWAIC : 3.72323 (p,q)=(4,6) aic=3.7314 fpe=40.9632	AIC : 3.73377 NEWAIC : 3.71501 (p,q)=(6,9) aic=3.73377 fpe=40.1416	AIC : 3.8275 NEWAIC : 3.80812 (p,q)=(9,1) aic=3.8275 fpe=45.0069	

**Tabla 7.** Criterio AIC para la selección del mejor modelo ARMA.

En la tabla 7, se observa en la celda sombreada que el modelo que presenta menor criterio de Akaike es el modelo ARMA(8,9) con un valor de 3.421 (Salida MatLab), por lo tanto a continuación se muestran los resultados de los parámetros estimados para este modelo.

Coeficiente	Desv.	Típica	$z$	Valor $p$	
const	8,08865	1,16852	6,922	4,45E-12	***
phi_1	-0,455369	0,285578	-1,595	0,1108	
phi_2	0,104047	0,248393	0,4189	0,6753	
phi_3	0,813299	0,290899	2,796	0,0052	***

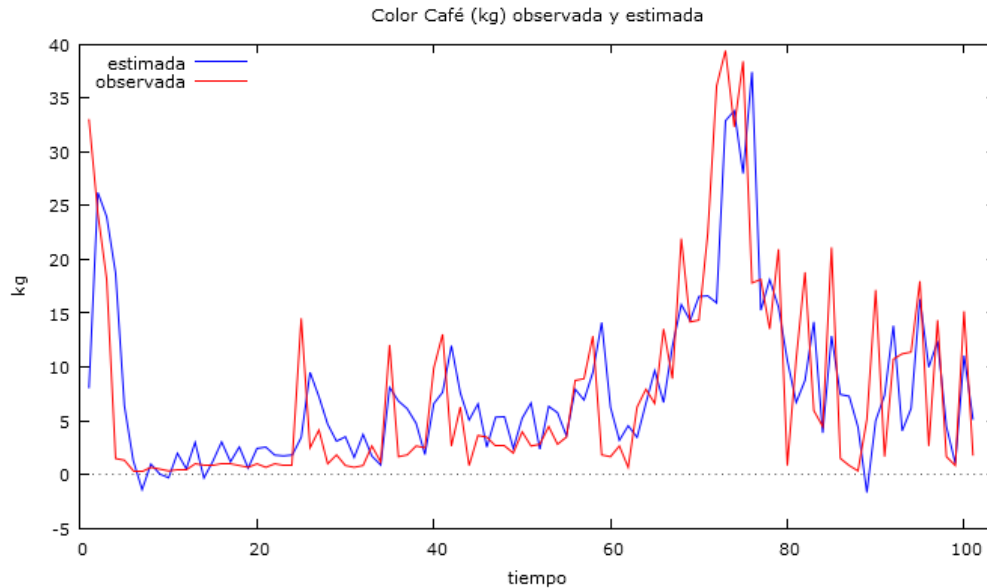
phi_4	0,531739	0,373684	1,423	0,1547	
phi_5	0,708123	0,307156	2,305	0,0211	**
phi_6	-0,099841	0,1971	-0,5066	0,6125	
phi_7	-0,324635	0,235694	-1,377	0,1684	
phi_8	-0,434419	0,168438	-2,579	0,0099	***
theta_1	1,0933	0,296705	3,685	0,0002	***
theta_2	0,804891	0,396715	2,029	0,0425	**
theta_3	-0,0883879	0,42122	-0,2098	0,8338	
theta_4	-0,412735	0,469257	-0,8796	0,3791	
theta_5	-1,10413	0,391918	-2,817	0,0048	***
theta_6	-0,754106	0,367439	-2,052	0,0401	**
theta_7	-0,353888	0,353289	-1,002	0,3165	
theta_8	0,115441	0,355198	0,325	0,7452	
theta_9	-0,300391	0,282542	-1,063	0,2877	

**Tabla 8.** Parámetros estimados para modelo ARMA(8,9).

<b>Media de la vble. dep.</b>	7,485446	<b>D.T. de la vble. dep.</b>	9,183394
<b>Media innovaciones</b>	-0,318634	<b>D.T. innovaciones</b>	5,437611
<b>Log-verosimilitud</b>	-320,63	<b>Criterio de Akaike</b>	3,73062
<b>Criterio de Schwarz</b>	728,9472	<b>Crit. de Hannan-Quinn</b>	699,3748

**Tabla 9.** Estadísticos del ajuste de los parámetros del modelo ARMA(8,9).

La tabla 8 muestra los parámetros del modelo ARMA (8,9) estimados mediante el uso del filtro de Kalman (Máxima verosimilitud exacta). Además se pueden apreciar los parámetros del modelo, con los cuales se calcula la ecuación para el pronóstico, lo más significativos están señalados con asteriscos. En la tabla 9 se observa el cálculo de diferentes estadísticos que miden el ajuste de parámetros al modelo, en este caso el criterio de Akaike es tenido en cuenta para la selección del mejor modelo. (Salida Gretel).



**Gráfica 5.** Serie observada y serie estimada de la demanda del color café en Coats Cadena Andina S.A.

En la gráfica 5, se puede evidenciar claramente que la serie estimada con el modelo ARMA (8,9) capta el comportamiento de la serie original, hasta en aquellas observaciones que presentan cambios fuertes, picos altos.

Observaciones	Color Café	Estimada
1	33	8,089
2	24,255	26,226
3	18,15	23,964
4	1,485	18,729
5	1,32	6,386
6	0,33	1,308
7	0,33	-1,426
8	0,66	0,962
9	0,495	0,027
10	0,33	-0,306
11	0,495	1,97
12	0,495	0,524
13	0,99	2,982
14	0,825	-0,335
15	0,825	1,215
16	0,99	3,026
17	0,99	1,176
18	0,825	2,55
19	0,66	0,582

20	0,99	2,415
21	0,66	2,535
22	0,99	1,801
23	0,825	1,722
24	0,825	1,804
25	14,52	3,439
26	2,475	9,499
27	4,125	7,24
28	0,99	4,661
29	1,815	3,089
30	0,825	3,499
31	0,66	1,58
32	0,825	3,722
33	2,64	1,702
34	1,155	0,875
35	12,045	8,091
36	1,65	6,833
37	1,815	6,11
38	2,64	4,765
39	2,475	1,856
40	9,9	6,587
41	13,035	7,65
42	2,64	11,973
43	6,27	7,591
44	0,825	5,066
45	3,63	6,564
46	3,465	2,543
47	2,64	5,327
48	2,64	5,364
49	1,98	2,417
50	3,96	5,225
51	2,64	6,639
52	2,805	2,345
53	4,455	6,33
54	2,805	5,74
55	3,465	3,6
56	8,745	7,927
57	8,91	6,925
58	12,87	9,507
59	1,815	14,134
60	1,65	6,245

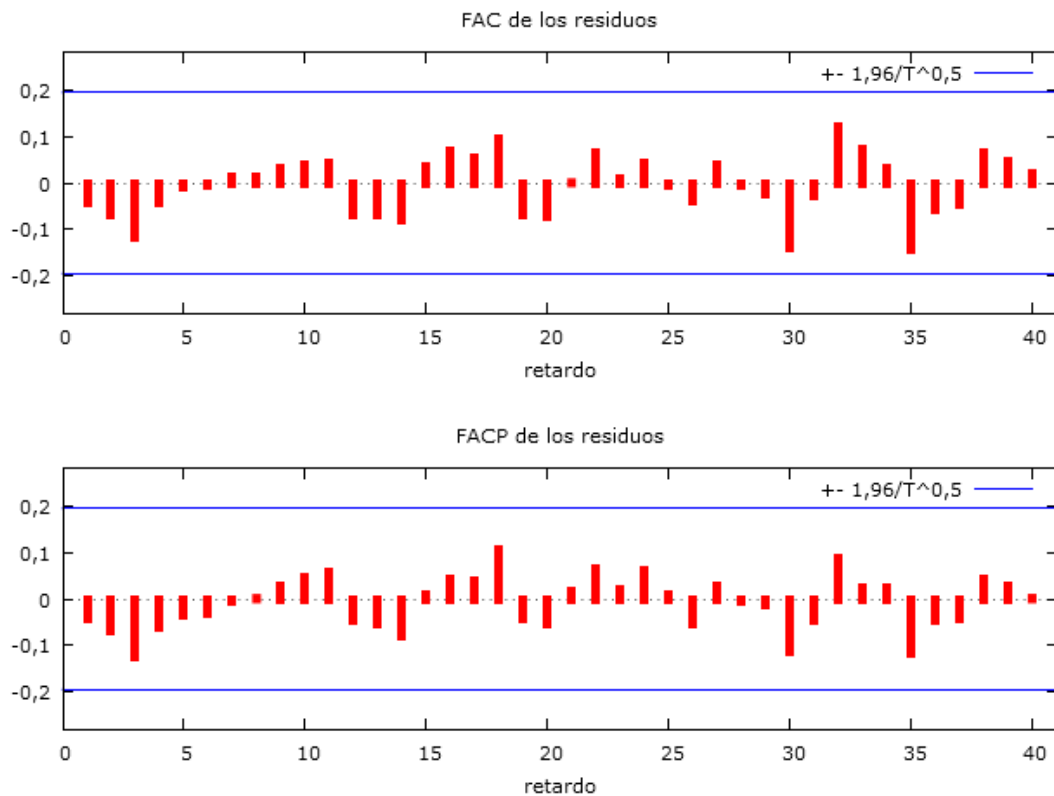
61	2,64	3,174
62	0,66	4,528
63	6,27	3,453
64	7,92	6,609
65	6,6	9,658
66	13,53	6,696
67	8,91	11,931
68	21,945	15,806
69	14,19	14,329
70	14,355	16,547
71	22,44	16,616
72	36,135	15,96
73	39,435	32,903
74	32,34	33,838
75	38,445	28,001
76	17,82	37,449
77	18,15	15,276
78	13,53	18,117
79	20,955	15,579
80	0,825	10,609
81	10,725	6,703
82	18,81	8,747
83	5,94	14,192
84	4,455	3,852
85	21,12	12,88
86	1,485	7,437
87	0,825	7,269
88	0,33	4,406
89	5,115	-1,686
90	17,16	5,005
91	1,65	7,352
92	10,725	13,833
93	11,22	4,05
94	11,385	6,129
95	17,985	16,312
96	2,64	9,976
97	14,355	12,397
98	1,65	4,452
99	0,825	1,029
100	15,18	11,061
101	1,815	5,153

**Tabla 10.** Variable observada y variable estimada del color café en Coats Cadena Andina S.A.

Error medio	-0,31863
Error cuadrático medio	36,679
Raíz del error cuadrático medio	6,0563
Error absoluto medio	4,1308
Porcentaje de error medio	-116,61
Porcentaje de error absoluto medio	160,79

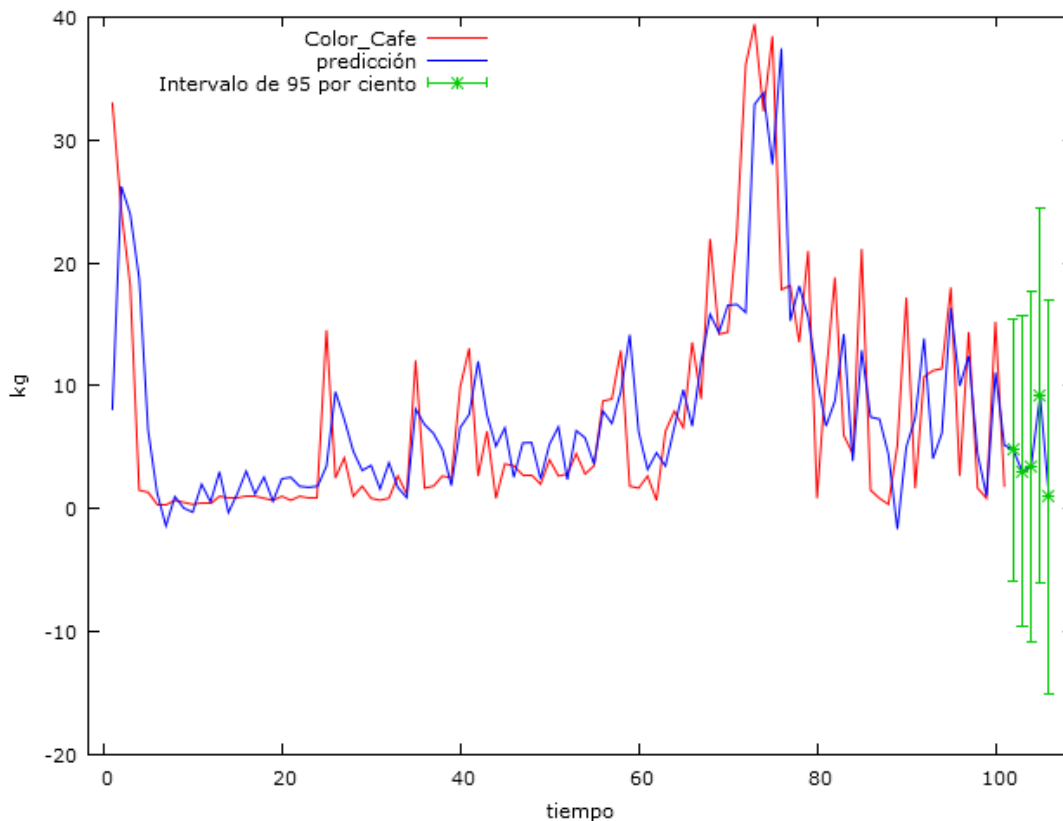
**Tabla 11.** Errores del modelo ARMA (8,9).

La tabla 10 muestra los datos de la variable original, variable estimada y residuo para el modelo ARMA (8,9). En la tabla 6, para el modelo ARMA(4,4) se observa un error cuadrático medio de 108, mientras que para el modelo nuevo ARMA(8,9) de la tabla 11, este error se reduce hasta 36, luego de la eliminación de los datos atípicos, esto evidencia el por qué el modelo ARMA (8,9) se ajusta de una mejor manera al comportamiento de la serie original de la demanda del color café de Coats Cadena Andina S.A



**Gráfica 6.** Función de autocorrelación y de autocorrelación parcial del modelo ARMA(8,9).

En la gráfica 6, la autocorrelación y autocorrelación parcial de los residuos para todos los retardos, se puede considerar estadísticamente como cero, debido a que los valores de éstas se encuentran dentro de los límites de confianza. Estos residuos son ruido blanco independientemente distribuido, lo que indica que el modelo se ajusta a los datos. Por lo tanto, el modelo ARMA (8,9) se utilizará para hacer pronósticos fuera de la muestra.



**Gráfica 7.** Pronósticos de la demanda del color café en Coats Cadena Andina S.A.

En la gráfica 7 se muestran los pronósticos de las cinco observaciones siguientes con su respectivo intervalo de confianza al 95%. La ecuación del modelo ARMA(8,9) para realizar los pronósticos es la siguiente:

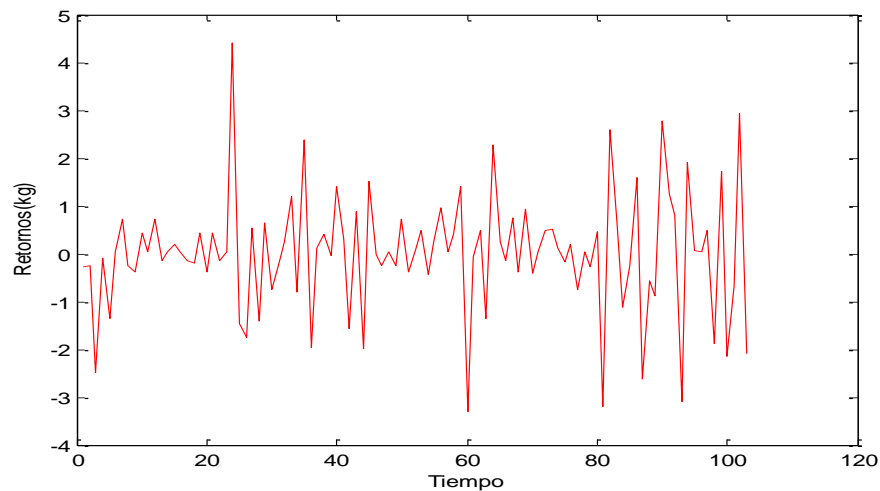
$$\begin{aligned} \hat{x}_t = & 8.088 - 0.455x_{t-1} + 0.104x_{t-2} + 0.813x_{t-3} + 0.532x_{t-4} + 0.708x_{t-5} - 0.0998x_{t-6} \\ & - 0.325x_{t-7} - 0.434x_{t-8} + 1.093a_{t-1} + 0.805a_{t-2} - 0.088a_{t-3} \\ & - 0.412a_{t-4} - 1.104a_{t-5} - 0.754a_{t-6} - 0.354a_{t-7} + 0.115a_{t-8} \\ & - 0.301a_{t-9} \end{aligned}$$

Luego de la eliminación de los datos atípicos de la serie original de la demanda del color café de Coats Cadena Andina S.A, el modelo seleccionado es el ARMA(8,9) , ya que presenta el

menor criterio AIC, el error cuadrático medio del modelo disminuye a 36 frente al del ARMA (4,4) que es 108 y además se evidencia que los residuos de los retardos son ruido blanco lo que indica que los datos se ajusta a este modelo seleccionado.

## 6.2. Selección del modelo GARCH

Para seleccionar el modelo GARCH primero se debe ajustar un modelo AR a la serie de retornos de la demanda del color café. En la selección del orden del modelo se utiliza el criterio de Akaike. Para la selección del orden GARCH se analizan los residuos cuadrados ya que los efectos ARCH serán visibles en la función de autocorrelación de los residuos al cuadrado que mostrarán correlación serial, además se calcula coeficientes de autocorrelación parcial de los residuos al cuadrado, donde el número de términos distintos de cero indica aproximadamente el orden del proceso.



**Gráfica 8.** Serie de retornos de la demanda del color café en Coats Cadena Andina S.A.

---

AIC : 0.343623 NEWAIC : 0.343238  
 (p,q)=(1,0) aic=0.343623 fpe=1.40977

---

AIC : 0.33026 NEWAIC : 0.329095  
 (p,q)=(2,0) aic=0.33026 fpe=1.39025

---

AIC : 0.360917 NEWAIC : 0.357945  
 (p,q)=(3,0) aic=0.360917 fpe=1.43206

---



AIC : 0.355035 NEWAIC : 0.349371  
 (p,q)=(4,0) aic=0.355035 fpe=1.42154

AIC : 0.271857 NEWAIC : 0.266904  
 (p,q)=(5,0) aic=0.271857 fpe=1.30601

AIC : 0.28953 NEWAIC : 0.282323  
 (p,q)=(6,0) aic=0.28953 fpe=1.32638

AIC : 0.283524 NEWAIC : 0.273613  
 (p,q)=(7,0) aic=0.283524 fpe=1.31498

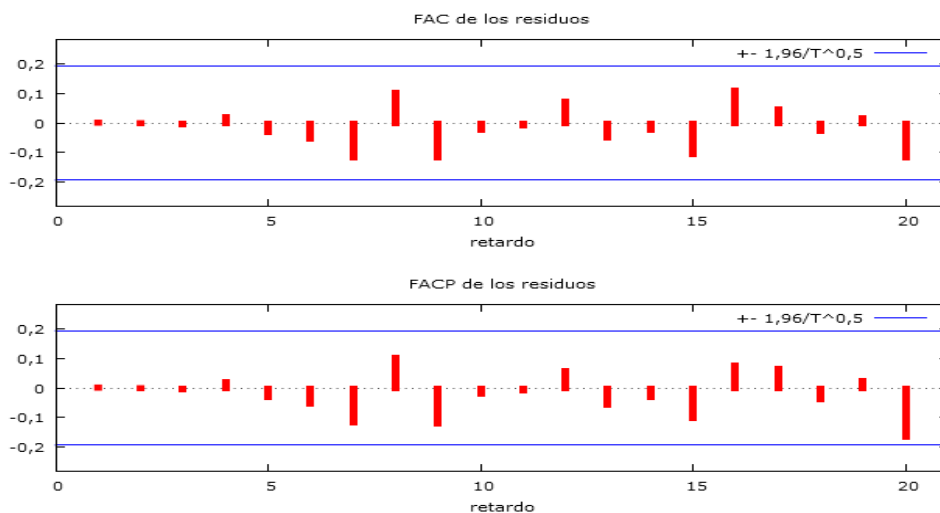
AIC : 0.307475 NEWAIC : 0.294394  
 (p,q)=(8,0) aic=0.307475 fpe=1.34273

AIC : 0.316661 NEWAIC : 0.299929  
 (p,q)=(9,0) aic=0.316661 fpe=1.35037

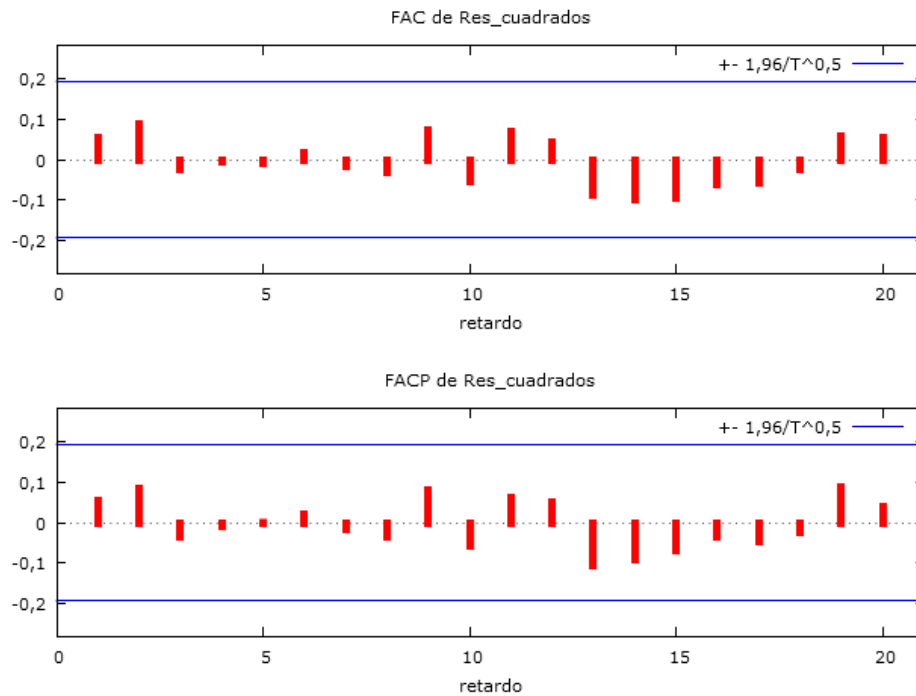
AIC : 0.336553 NEWAIC : 0.315674  
 (p,q)=(10,0) aic=0.336553 fpe=1.37202

**Tabla 12.** Criterio AIC para la selección del mejor modelo AR.

En la tabla 12 se observa que el modelo que presentó menor criterio de Akaike fue el modelo AR(5), con un valor de 0.271 (Salida MatLab). En la gráfica 9 y en la gráfica 10 se muestran las funciones de autocorrelación y autocorrelación parcial de los residuos y de los residuos al cuadrado del modelo AR(5) respectivamente, donde se puede observar que ninguno de los casos presenta correlación serial diferente de cero ya que se encuentran dentro de los límites de confianza, por lo tanto no se puede apreciar efectos ARCH en esta serie, que permitan establecer el orden de un modelo GARCH, pues cuando hay correlación en los residuos el orden del modelo GARCH se dá con el número de retardos con correlación.



**Gráfica 9.** Función de autocorrelación y de autocorrelación parcial del modelo AR (5).



**Gráfica 10.** Función de autocorrelación y de autocorrelación parcial de los residuos al cuadrado del modelo AR (5).

Sin embargo, se utilizará un modelo AR(5) en la media y GARCH (1,1) en varianza para estimar los retornos y la varianza condicional de la serie del color café de Coats Cadena Andina S.A. y así indicar la metodología del uso del modelo GARCH. Los parámetros estimados mediante máxima verosimilitud se muestran en la siguiente tabla:

Retornos de la demanda del color café (observada)	Estimación de la media condicional	Estimación de la varianza condicional	Residuo
-0,280	-0,316	1,137	0,036
-0,262	-0,440	1,028	0,178
-2,475	-2,706	1,020	0,231
-0,090	-1,454	1,573	1,364
-1,358	-2,056	1,283	0,698
0,028	-0,907	1,394	0,935
0,721	-0,117	1,142	0,838
-0,260	-0,039	1,021	-0,220
-0,377	-0,666	0,999	0,289
0,434	0,181	1,040	0,253
0,028	0,264	1,006	-0,236
0,721	0,728	1,004	-0,006
-0,154	0,125	1,049	-0,279
0,028	0,159	1,005	-0,131

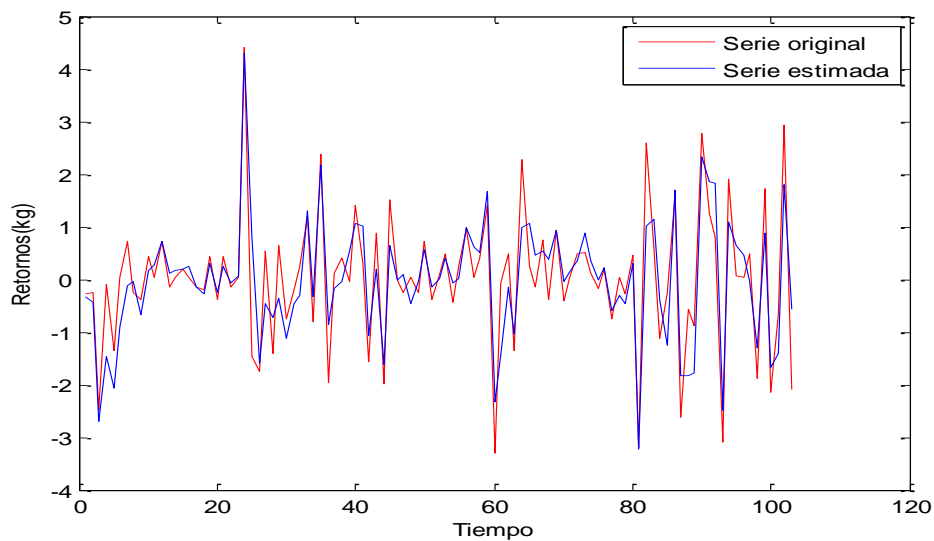
0,210	0,204	0,999	0,007
0,028	0,245	1,000	-0,217
-0,154	-0,153	1,002	-0,001
-0,195	-0,288	0,999	0,093
0,434	0,310	1,004	0,124
-0,377	-0,237	1,006	-0,140
0,434	0,259	1,003	0,174
-0,154	-0,072	1,003	-0,082
0,028	0,076	0,997	-0,048
4,398	4,292	0,996	0,106
-1,473	0,804	2,170	-2,278
-1,741	-1,584	1,301	-0,157
0,539	-0,462	1,267	1,001
-1,399	-0,718	1,062	-0,681
0,634	-0,349	1,056	0,983
-0,760	-1,126	1,017	0,366
-0,195	-0,459	1,120	0,264
0,251	-0,300	1,036	0,551
1,191	1,316	1,011	-0,124
-0,799	-0,337	1,160	-0,462
2,373	2,159	1,033	0,213
-1,960	-0,854	1,394	-1,106
0,123	-0,181	1,133	0,304
0,403	-0,027	1,021	0,429
-0,036	0,534	0,999	-0,570
1,414	1,074	1,024	0,341
0,303	1,005	1,110	-0,702
-1,569	-1,058	1,110	-0,511
0,893	0,201	1,120	0,693
-2,000	-1,621	1,019	-0,379
1,510	0,647	1,236	0,863
-0,018	-0,001	1,076	-0,018
-0,244	0,103	1,008	-0,347
0,028	-0,458	0,998	0,486
-0,260	-0,044	1,017	-0,215
0,721	0,557	0,999	0,164
-0,377	-0,154	1,027	-0,224
0,089	0,009	1,003	0,080
0,491	0,402	0,996	0,088
-0,434	-0,077	1,012	-0,358
0,239	0,025	0,998	0,215
0,954	0,985	0,996	-0,031
0,047	0,626	1,089	-0,579

0,396	0,521	1,049	-0,125
1,402	1,673	1,031	-0,271
-3,304	-2,341	1,252	-0,963
-0,067	-1,431	1,478	1,364
0,498	-0,141	1,260	0,639
-1,358	-1,043	1,042	-0,316
2,279	0,993	1,107	1,287
0,262	1,065	1,107	-0,803
-0,154	0,472	1,120	-0,626
0,746	0,537	1,037	0,209
-0,390	0,389	1,031	-0,779
0,930	0,927	1,016	0,003
-0,408	-0,047	1,082	-0,361
0,040	0,132	1,009	-0,092
0,475	0,346	0,999	0,129
0,505	0,884	1,008	-0,379
0,116	0,367	1,073	-0,251
-0,170	-0,007	1,021	-0,163
0,201	0,229	0,999	-0,028
-0,741	-0,598	1,001	-0,143
0,047	-0,302	1,032	0,348
-0,266	-0,464	1,010	0,199
0,466	0,305	1,019	0,161
-3,207	-3,213	1,008	0,007
2,593	1,021	1,755	1,572
0,590	1,135	1,234	-0,545
-1,125	-0,372	1,154	-0,753
-0,260	-1,253	1,035	0,993
1,584	1,705	1,149	-0,120
-2,627	-1,831	1,276	-0,796
-0,560	-1,820	1,329	1,260
-0,888	-1,764	1,334	0,876
2,769	2,341	1,319	0,428
1,239	1,855	1,487	-0,617
0,809	1,821	1,369	-1,012
-3,094	-2,487	1,340	-0,607
1,900	1,101	1,537	0,799
0,073	0,646	1,203	-0,573
0,043	0,472	1,070	-0,430
0,485	-0,040	1,029	0,526
-1,891	-1,306	1,001	-0,585
1,721	0,868	1,156	0,853
-2,135	-1,678	1,094	-0,457

-0,665	-1,408	1,261	0,743
2,941	1,806	1,218	1,135
-2,096	-0,566	1,313	-1,530

**Tabla 13.** Variable observada, estimación de la media, varianza condicional y residuos.

La tabla 13 contiene los valores de los retornos de la variable observada y la estimación de la media condicional (salida MatLab), dichos valores se pueden ver en la gráfica 11, en ésta se puede percibir que la serie estimada capta el comportamiento de la serie original, pero en los picos no se alcanza ajustar acertadamente a ésta.

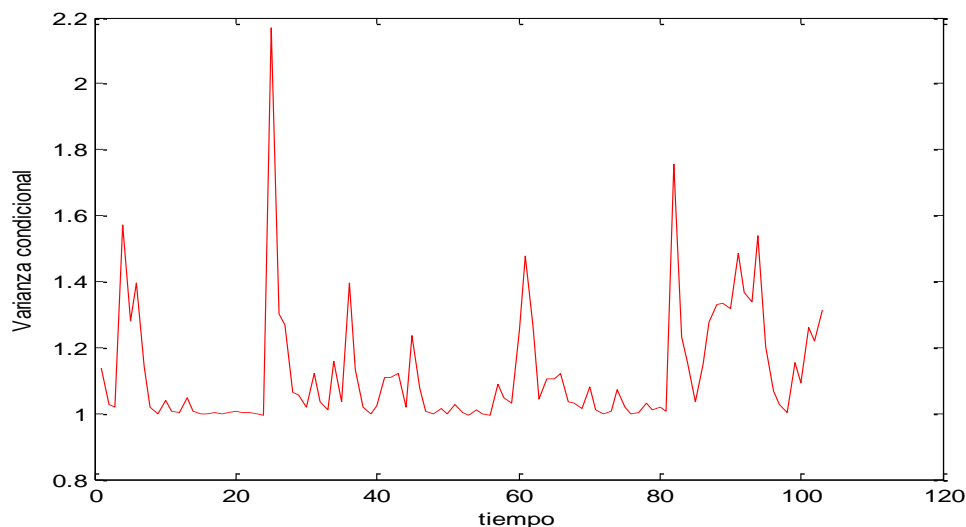


**Gráfica 11.** Serie observada y serie estimada de los retornos de la demanda del color café en Coats Cadena Andina S.A.

Error medio	0,0240
Error cuadrático medio	0,4028
Raíz del Error cuadrático medio	0,6347
Error absoluto medio	0,4789
Porcentaje de error medio	2,2257
Porcentaje de error absoluto medio	227,0436

**Tabla 14.** Errores del modelo GARCH (1,1).

La tabla 14 muestra los errores arrojados para el modelo GARCH (1,1), a pesar de que los valores de algunos de estos errores son pequeños, no se debe concluir acerca del modelo, ya que como se menciona al inicio de esta sección, se trata de indicar la metodología del uso del modelo GARCH.

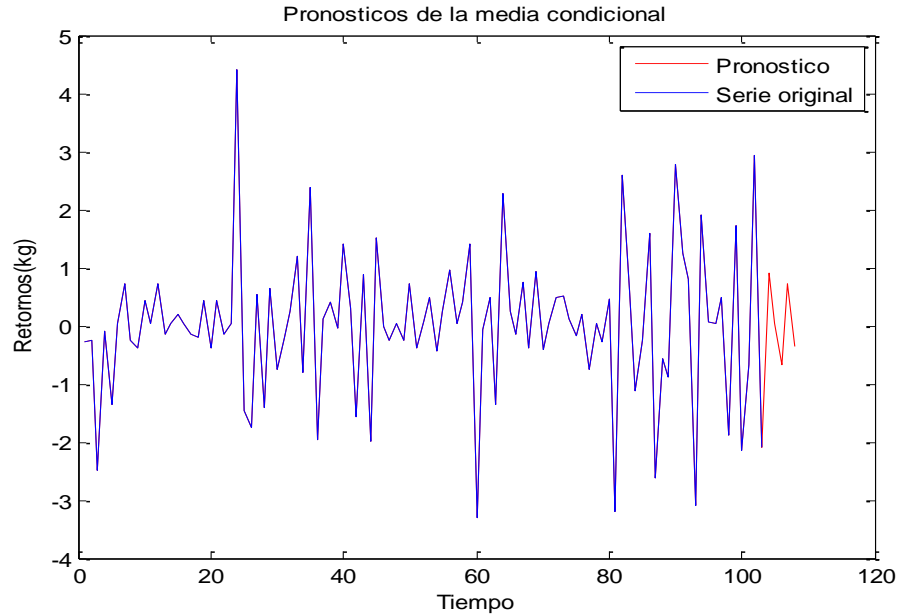


**Gráfica 12.** Varianza condicional estimada de los retornos de la demanda del color café en Coats Cadena Andina S.A.

En la gráfica 12 se muestra el comportamiento de la varianza condicional estimada para los retornos de la serie de la demanda de color café en Coats Cadena Andina S.A., la cual indica el riesgo de que posiblemente exista una diferencia significativa entre la serie real de los retornos de la demanda y la serie estimada de los retornos de la demanda, debido a los picos que se evidencian en el gráfico señalado.

Distribución		'Gaussiana'	
R			5
C			0.0364
AR		[-0.5072 -0.2198 -0.0541 -0.2060 -	0.0096]
Varianza del Modelo		'GARCH'	
P			1
Q			1
K			0.8382
GARCH			0.1538
ARCH			0.2018
log Verosimilitud			-1.570.160

**Tabla 15.** Estadísticos del ajuste de los parámetros del modelo GARCH(1,1).



**Gráfica 13.** Pronósticos de los retornos de la demanda del color café en Coats Cadena Andina S.A.

Desde la tabla 15 (salida Gretel) se obtienen los parámetros necesarios para la construcción de la ecuación del modelo AR (5), GARCH(1,1), la cual se escribe de la siguiente manera:

$$\hat{x}_t = 0.0364 - 0.5072x_{t-1} - 0.2198x_{t-2} - 0.0541x_{t-3} - 0.2060x_{t-4} - 0.0096x_{t-5}$$

$$\sigma_t^2 = 0.8382 + 0.1538e_{t-1}^2 + 0.2018\sigma_{t-1}^2$$

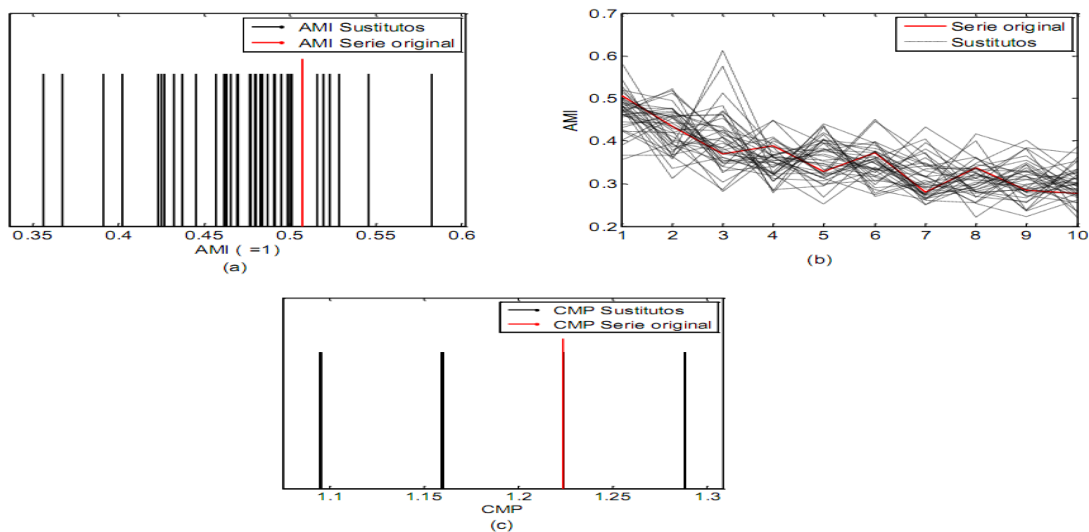
En la gráfica 13 se muestran los pronósticos de las siguientes 5 observaciones de los retornos de la serie de la demanda del color café de Coats Cadena Andina S.A. que siguen la dinámica de la serie, aunque el modelo AR(5), GARCH(1,1) no es el que mejor modelo que se ajusta a los datos.

El modelo que mejor se ajusta a los datos de la demanda del color café en Coats Cadena Andina S.A. es el ARMA(8,9). Se concluye ésto luego de estudiar en modelo ARMA(4,4) en el cual se hallaron datos atípicos y de evaluar los efectos ARCH que permitieron evidenciar que no se podía establecer el orden para un modelo GARCH.

### 6.3. Validación de las características de la dinámica de la serie de la demanda del color café de Coats Cadena Andina S.A.

En esta sección se determina si el modelo ARMA(8,9) seleccionado es adecuado de acuerdo a la dinámica intrínseca en la serie de la demanda del color café de Coats Cadena Andina S.A.

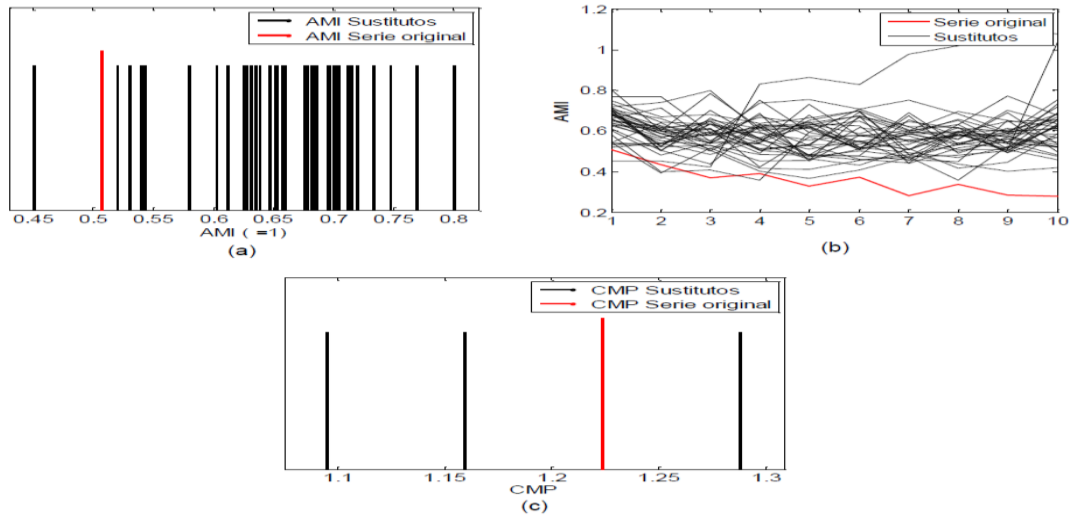
#### 6.3.1. Algoritmo SSS



Como se puede observar en el gráfico 14, los estadísticos de prueba AMI (Información Mutua Promedio) y CMP (Complejidad Relativa de Lempel-Ziv) de la serie de la demanda del color café de Coats Cadena Andina S.A. caen por dentro de la distribución de los estadísticos de los sustitutos generados por el algoritmo SSS, por lo tanto no se rechaza la hipótesis nula, esto significa que no hay evidencia de que las fluctuaciones irregulares de la serie analizada posean algún tipo de dinámica determinista.



### 6.3.2. Algoritmo TFT



**Gráfica 15.** (a) Estadístico de prueba AMI ( $\tau=1$ ). (b) Estadístico de prueba AMI  $\tau=10$ . (c) Estadístico de prueba CMP.

Como se puede observar en el gráfico 15, los estadísticos de prueba AMI y CMP de la serie de la demanda del color café de Coats Cadena Andina S.A. caen por dentro de la distribución de los estadísticos de los sustitutos generados por el algoritmo TFT, por lo tanto no se rechaza la hipótesis nula, esto significa que no hay evidencia de no linealidad en la serie analizada.

A través de los estadísticos de prueba AMI y CMP, se puede concluir la validación del modelo seleccionado ARMA(8,9), ya que los datos de la serie de la demanda del color café de Coats Cadena Andina S.A. son estocásticos lineales y se ajustan a este modelo.

## 6.4. Validación de los pronósticos respecto a las observaciones reales

La siguiente tabla muestra las observaciones reales, la estimación en la muestra y los siguientes 5 pronósticos con el fin de validar los resultados. Además se puede ver el intervalo de confianza al 95 % de los pronósticos, en donde 4 de las 5 siguientes observaciones reales están dentro este intervalo.

<b>Observaciones</b>	<b>Color café</b>	<b>Predicción</b>	<b>Desv. Típica</b>	<b>Intervalo de confianza 95%</b>
52	2.805	2.345		
53	4.455	6.33		
54	2.805	5.74		
55	3.465	3.6		
56	8.745	7.927		
57	8.91	6.925		
58	12.87	9.507		
59	1.815	14.134		
60	1.65	6.245		
61	2.64	3.174		
62	0.66	4.528		
63	6.27	3.453		
64	7.92	6.609		
65	6.6	9.658		
66	13.53	6.696		
67	8.91	11.931		
68	21.945	15.806		
69	14.19	14.329		
70	14.355	16.547		
71	22.44	16.616		
72	36.135	15.96		
73	39.435	32.903		
74	32.34	33.838		
75	38.445	28.001		
76	17.82	37.449		
77	18.15	15.276		
78	13.53	18.117		
79	20.955	15.579		
80	0.825	10.609		
81	10.725	6.703		
82	18.81	8.747		
83	5.94	14.192		
84	4.455	3.852		
85	21.12	12.88		
86	1.485	7.437		
87	0.825	7.269		
88	0.33	4.406		
89	5.115	-1.686		
90	17.16	5.005		
91	1.65	7.352		
92	10.725	13.833		

93	11.22	4.05			
94	11.385	6.129			
95	17.985	16.312			
96	2.64	9.976			
97	14.355	12.397			
98	1.65	4.452			
99	0.825	1.029			
100	15.18	11.061			
101	1.815	5.153			
102	<b>12.54</b>	4.76	5.4376	-5.897	- 15.418
103	<b>17.985</b>	3.002	6.4498	-9.639	- 15.644
104	<b>0.825</b>	3.45	7.2739	-10.807	- 17.706
105	<b>1.815</b>	9.148	7.7839	-6.108	- 24.405
106	<b>1.815</b>	0.961	8.1929	-15.096	- 17.019

Tabla 16. Observaciones reales v s pronósticos de la demanda del color café en Coats Cadena Andina S.A.

En la siguiente tabla se observan los errores de los pronósticos para los siguientes 5 observaciones. El error cuadrático medio en la muestra es 69.282, éste aumentó respecto al del modelo ARMA(8,9) el cual es 36.679, sin embargo cabe observar que los pronósticos presentan un mejor desempeño que los del modelo ARMA (4,4).

Error medio	2.732
Error cuadrático medio	69.282
Raíz del Error cuadrático medio	8.324
Error absoluto medio	6.715
Porcentaje de error medio	-105.960
Porcentaje de error absoluto medio	216.888

**Tabla 16.** Errores de los pronósticos.

## 7. Conclusiones

- A partir del análisis de la serie de la demanda del color café de Coats Cadena Andina S.A., es desarrollada una metodología paso a paso para establecer el ajuste de los datos a un modelo ARMA o GARCH, y desde el modelo seleccionado pronosticar la próxima observación de la demanda de este color. El modelo seleccionado en ese trabajo es ARMA(8,9).
- En el modelo ARMA(8,9) el error medio cuadrático (MSE) registra un menor valor al arrojado por el modelo ARMA(4,4), lo cual afianza que es uno de los criterios importantes para tener en cuenta en el análisis del desempeño del modelo.
- En la selección del modelo GARCH no se encontró evidencia de efectos ARCH en los errores cuadrados, es decir, varianza condicional no lineal, lo cual explica en cierta forma por qué mostró menos errores el ajuste del modelo ARMA.
- De acuerdo a las pruebas realizadas con el método de Datos Sustitutos, la dinámica de la serie de demanda del color café de Coats Cadena Andina S.A. es posible que sea estocástica lineal, ya que no hubo evidencia de determinismo no lineal, lo cual es consistente con las características de un modelo ARMA.
- Se comprueba que para series temporales que presentan pocas mediciones el proceso tiende a ser estocástico, por el contrario series de mayor longitud presentan comportamientos que podrían llegar a ser deterministas.

- Aunque en la literatura se habla de que los modelos de series económicas y financieras sean no lineales, en este trabajo se puede evidenciar que los modelos lineales son adecuados y vigentes para este tipo de series, especialmente series de demanda.
- Con base en las pruebas realizadas a través de este trabajo, es posible que la demanda del color café sea estocástica, lo cual indica aleatoriedad en los pedidos de éste. Dicha conclusión coincide con el planificador de fabricación actual, ya que tiene el perfil de *Especial*, o sea que solo se fabrica contra pedido de cliente y en tanda completa.
- El modelo ARMA (8,9) en su límite superior de confianza de los pronósticos, puede considerarse como un inventario límite máximo para responder instantáneamente a la demanda. Estadísticamente esto quiere decir que con un nivel de confianza de 95% el siguiente pedido no superará esta cantidad.
- Esta metodología permite modelar individualmente cada serie de color y concluir de manera casi inmediata si éste se puede incluir en la carta de colores ofertada por la empresa Coats Cadena Andina S.A., sin esperar al análisis global hecho con la herramienta OFA (Orders Frequency Analysis), el cual se realiza con una periodicidad anual.

## 8. Futuras líneas de trabajo

Como futuras líneas de investigación se puede tener:

- Automatizar el proceso de análisis gráfico para evidenciar el ajuste de los datos reales al modelo ARMA o al modelo GARCH.
- Automatizar la selección del modelo.
- Integrar el componente externo Gretel en Matlab.
- Crear una base de datos para capturar y almacenar la información de los pedidos por color.

# Anexos

## Anexo 1.

### Elección del modelo arma con base en el criterio de Akaike

```
x=data;
figure(1)
subplot(2,2,1), plot(x);
ylabel('Xn')
title('Trayectoria')
[fac_y,m]=autocorr(x,[],2);
subplot(2,2,2), autocorr(x,[],2)
title('fac');
[facp_y, mp] = parcorr(x,[],2);
subplot(2,2,3), parcorr(x,[],2)
title('facp');
v = (fac_y(1)-fac_y)/(fac_y(1)-fac_y(2));
subplot(2,2,4), stem(m,v);
grid
title('Variograma')

% explora el orden
pvec = [0 1 2 3 4 5 6 7 8 9 10];
qvec = [0 1 2 3 4 5 6 7 8 9 10];
[mbest,minaic,pbest,qbest]=armabat(x,pvec,qvec);
pbest
qbest
```

## Anexo 2.

### Algoritmo de ajuste del mejor modelo ARMA

```
data=x;
data1=iddata(data);
yp = predict(mbest,data1);

plot(data1,'r',yp,'b');
legend('Serie original','Serie estimada');
title('')
xlabel('Tiempo')
ylabel('kg')
```



## Anexo 3.

### Algoritmo de ajuste de un Modelo Garch

```
% Modelo Garch

xt = price2ret(data);

%%
spec = garchset('VarianceModel','GARCH',...
               'R',5,'M',0,'P',1,'Q',1);
spec = garchset(spec,'Display','on','Distribution','Gaussian')
%%
[coeff,errors,LLF,eFit,sFit] = ...           % cálculo de los
garchfit(spec,xt)                          % parámetros del modelo

[eFit,sFit] = garchinfer(coeff,xt); % estimación del modelo
%%

horizon = 5; % Define el horizonte de predicción

[sigmaForecast,meanForecast,sigmaTotal,meanRMSE] = ... % pronostico
garchpred(coeff,xt,horizon);
```

## Anexo 4

### Función Algoritmo SSS

```
function z = SSS(x,A,k)

% z=SSS(x,fe,k);
% Small Shuffle surrogates see
% [T. Nakamura and M. Small. "Applying the method of Small-Shuffle
% surrogate data: Testing for dynamics in fluctuating data with trends."
% International Journal of Bifurcations and Chaos 16 (2006):3581-3603]
% and [M. Small T. Nakamura and X. Luo. "Surrogate data methods for data
% that isn't linear noise." In: Nonlinear Phenomena Research Perspectives,
% ed. C.W. Wang, pages 55-81, Nova Science Publications, 2007] for more
info.
% x -> The time series
% fe -> The range of frequencies to modifie
% k -> The number of surrogates
% Implmented by Diego L. Guarin on april 2010
% dlguarin@gmail.com

% DISCLAIMER

% We make no warranties, explicit or implicit, that this program
% is free of error, or is consistent with any particular standard
% of accuracy, or that it will meet your requirements for any
% particular application.

% It should not be relied on for any purpose where incorrect
% results could result in loss of property or personal injury.
% If you do use this program for any such purpose it is at your own
% risk. The author disclaims all liability of any kind, either
% direct or consequential, resulting from your use of this program.

if nargin<3
    k=1;
end
if nargin<2
    A=1;
end
if nargin <1
    disp('Insuficientes argumentos de entrada');
    return
end

x=x(:);
z=[];
for i=1:k
    z=[z SSSalg(x,A)];
```

```

end
return
end

function xp = SSSalg(x,A)

[n,m]=size(x);
if m>1
    x=x';
    [n,m]=size(x);
    if m>1
        disp('No funciona con multiples canales');
        return
    end
end

end
x=x(:,1);

y=0+A.*randn(n,1);

i=(1:n)'+y;
[isort,iind]=sort(i);

xp=x(iind);

return
end

```

## Anexo 5

### Función Algoritmo TFT

```
function z = TFTS (x,fe,k,ptofinal)

% z=TFTS(x,fe,k,ptofinal);
% Truncate Fourier Transform Surrogate, see
% [T. Nakamura, M. Small, and Y. Hirata, "Testing for nonlinearity in
% irregular fluctuations with long-term trends," Physical Review E, vol.
% 74, p. 026205, 2006.] for more info.
% x -> The time series
% fe -> The range of frequencies to modifie
% k -> The number of surrogates
% ptofinal -> If the first and last point of x are too diferent then
%             ptofinal=-1, else ptofinal=1 (default ptofinal=1)
% Implmented by Diego L. Guarin on febraury 2010
% dlguarin@gmail.com

if nargin<4
    ptofinal=1;
end
if abs(ptofinal)~= 1
    ptofinal=1;
end
if nargin<3
    k=1;
    ptofinal=1;
end
if nargin <2
    disp('This program need at least a time series (x) and a frequency
domain(fe)');
    return
end

x=x(:);
z=[];
for i=1:k
    z=[z TFTSalg(x,fe,ptofinal)];
end
return
end

function xp = TFTSalg(x,fe,ptofinal)

n=length(x);

%doesn't work for multichannel data
```

```

x=reshape(x,[length(x) 1]);

if fe>1
    error('fe can't be greater than 1');
end

y=x(:,1);
m= mean(y);
y= y-m;

% endpoint correction
if ptofinal<0
    y0= y(1,:);
    y1= y(n,:);
    u= (0:n-1)'/(n-1);
    c= u*y1 + (1-u)*y0;
else
    c= 0;
end;

%fft
fy = fft(y-c);

n_m=ceil(n/2);
fc=n_m-ceil(n_m*fe);

%if the number of points is even
if mod(n,2)==0
    fy2=fy(1:ceil(n/2)+1);

    % random number generator
    phz= rand(ceil(n/2)+1,1);

    %truncating the higher frequencies
    phz(1:fc)=0;
    rot= exp(1) .* (2*pi*sqrt(-1)*phz);
    fyp=fy2.*rot;

    fyp=[fyp; conj(fyp(end-1:-1:2))];
%if is odd
else
    fy2=fy(1:ceil(n/2));

    % random number generator
    phz= rand(ceil(n/2),1);

    %truncating the higher frequencies
    phz(1:fc-1)=0;
    rot= exp(1) .* (2*pi*sqrt(-1)*phz);
    fyp=fy2.*rot;

    fyp=[fyp; conj(fyp(end:-1:2))];

end

```

```
xp=real(iff(y)) + c * m(ones(n,1),1);  
return  
end
```

## Anexo 6

# Algoritmo con el que se generan los sustitutos SSS y se calculan los estadísticos AMI y CMP

```
xs=sss(x,1,39);

% SSS AMI Grafico 1

A=informacion_mutua(x,1);
plot(A,'bs');
xlabel('Time lag')
ylabel('AMI')
text(1,A,[' \leftarrow = ',num2str(A)]);
hold on
max=0;
min=999999999;

figure(1)
for i=1:39
    As(i)=informacion_mutua(xs(:,i),1);
    plot(As(i),'ro');
    if As(i)> max
        max=As(i);
    end
    if As(i)< min
        min=As(i);
    end
    hold on
end
text(1,max,[' \leftarrow Maximo = ',num2str(max)]);
text(1,min,[' \leftarrow Minimo = ',num2str(min)]);
legend('Serie original','Sustitutos')

% SSS AMI Grafico 2
figure(2)
h=ones(39,1);
stem(As,h,'Marker','None','Color','Black','LineWidth',2);
hold on
stem(A,1.1,'Marker','None','Color','red','LineWidth',2)

min=999999999;
max=-999999999;
for i=1:39
    if As(i)< min
```

```

        min=As(i);
    end
    if As(i) > max
        max=As(i);
    end
end

if A < min
    xmin=A-0.02;
else
    xmin= min-0.02;
end

if A > max
    xmax=A + 0.02;
else
    xmax= max + 0.02;
end

xlim([xmin xmax]);
ylim([0 1.4]);
set(gca, 'XTick', 0:0.05:3, 'FontSize', 14);
set(gca, 'YTick', [])
xlabel('AMI (\tau=1)', 'FontSize', 14);
legend('AMI Sustitutos', 'AMI Serie original')

% SSS AMI grafico 3
figure (3)

A=informacion_mutua(x,10);
plot(A, 'r', 'LineWidth', 2);
hold on
for i=1:39
    As=informacion_mutua(xs(:,i),10);
    plot(As, 'black:');

    hold on
end
xlim([1 10])
set(gca, 'XTick', 1:1:10, 'FontSize', 14);
xlabel('\tau', 'FontSize', 14)
ylabel('AMI')

legend('Serie original', 'Sustitutos')

% SSS CMP grafico 4

figure(4)

C=complexity(x);
plot(C, 'bs');

xlabel('TIME LAG')

```



```

ylabel('CMP')
text(1,C,[' \leftarrow = ',num2str(C)]);
hold on
max=0;
min=999999999;
for i=1:39
    Cs(i)=complexity(xs(:,i));

    plot(Cs(i),'ro');
    if Cs(i)> max
        max=Cs(i);
    end
    if Cs(i)< min
        min=Cs(i);
    end
    hold on
end
text(1,max,[' \leftarrow Maximo = ',num2str(max)]);
text(1,min,[' \leftarrow Minimo = ',num2str(min)]);
legend('Serie original','Sustitutos')

% SSS CMP grafico 5

figure(5)

h=ones(39,1);
stem(Cs,h,'Marker','None','Color','Black','LineWidth',2);
hold on
stem(C,1.1,'Marker','None','Color','red','LineWidth',2)

min=999999999;
max=-999999999;
for i=1:39
    if Cs(i)< min
        min=Cs(i);
    end
    if Cs(i)> max
        max=Cs(i);
    end
end

if C < min
    xlmin=C-0.02;
else
    xlmin= min-0.02;
end

if C > max
    xlmax=C + 0.02;
else
    xlmax= max + 0.02;
end

xlim([xlmin xlmax]);
ylim([0 1.4]);
set(gca,'XTick',0:0.02:3,'FontSize',14)

```

```
set(gca, 'YTick', [])  
xlabel('CMP', 'FontSize', 14)  
legend('CMP Sustitutos', 'CMP Serie original')
```

## Anexo 7

# Algoritmo con el que se generan los sustitutos TFT y se calculan los estadísticos AMI y CMP

```
xs=TFTS(x,0.5,39,-1);

%% TFTS AMI Grafico 1

A=informacion_mutua(x,1);
plot(A,'bs');
title('ALGORITMO TFTS (COLOR CAFE)')
xlabel('Time lag')
ylabel('AMI')
text(1,A,[' \leftarrow = ',num2str(A)]);
hold on
max=0;
min=999999999;

figure(1)
for i=1:39
    As(i)=informacion_mutua(xs(:,i),1);
    plot(As(i),'ro');
    if As(i)> max
        max=As(i);
    end
    if As(i)< min
        min=As(i);
    end
    hold on
end
text(1,max,[' \leftarrow Maximo = ',num2str(max)]);
text(1,min,[' \leftarrow Minimo = ',num2str(min)]);
legend('Serie original','Sustitutos')

% TFTS AMI Grafico 2

figure(2)
h=ones(39,1);
stem(As,h,'Marker','None','Color','Black','LineWidth',2);
hold on
stem(A,1.1,'Marker','None','Color','red','LineWidth',2)

min=999999999;
max=-999999999;
```

```

for i=1:39
    if As(i) < min
        min=As(i);
    end
    if As(i) > max
        max=As(i);
    end
end

if A < min
    xmin=A-0.02;
else
    xmin= min-0.02;
end

if A > max
    xmax=A + 0.02;
else
    xmax= max + 0.02;
end

xlim([xmin xmax]);
ylim([0 1.4]);
set(gca, 'XTick', 0:0.05:3, 'FontSize', 14);
set(gca, 'YTick', [])
xlabel('AMI (\tau=1)');
ylabel('AMI (\tau=1)', 'FontSize', 14);
legend('AMI Sustitutos', 'AMI Serie original')

% TFTS AMI Grafico 3

figure (3)

A=informacion_mutua(x,10);
plot(A, 'r', 'LineWidth', 2);

hold on
for i=1:39
    As=informacion_mutua(xs(:,i), 10);
    plot(As, 'black:');

    hold on
end
xlim([1 10])
set(gca, 'XTick', 1:1:10, 'FontSize', 14);
xlabel('\tau', 'FontSize', 14)
ylabel('AMI')
legend('Serie original', 'Sustitutos')

% TFTS CMP Grafico 4

figure(4)

C=complexity(x);
plot(C, 'bs');

```

```

%title('ALGORITMO TFTS (COLOR CAFÉ)')
xlabel('TIME LAG')
ylabel('CMP')
text(1,C,[' \leftarrow = ',num2str(C)]);
hold on
max=0;
min=999999999;
for i=1:39
    Cs(i)=complexity(xs(:,i));

    plot(Cs(i),'ro');
    if Cs(i) > max
        max=Cs(i);
    end
    if Cs(i) < min
        min=Cs(i);
    end
    hold on
end
text(1,max,[' \leftarrow Maximo = ',num2str(max)]);
text(1,min,[' \leftarrow Minimo = ',num2str(min)]);
legend('Serie original','Sustitutos')

% TFTS CMP Grafico 5

figure(5)

h=ones(39,1);
stem(Cs,h,'Marker','None','Color','Black','LineWidth',2);
hold on
stem(C,1.1,'Marker','None','Color','red','LineWidth',2)

min=999999999;
max=-999999999;
for i=1:39
    if Cs(i) < min
        min=Cs(i);
    end
    if Cs(i) > max
        max=Cs(i);
    end
end

if C < min
    xmin=C-0.02;
else
    xmin= min-0.02;
end

if C > max
    xmax=C + 0.02;
else
    xmax= max + 0.02;
end

```

```
xlim([xmin xmax]);  
ylim([0 1.4]);  
set(gca, 'XTick', 0:0.02:3, 'FontSize', 14)  
set(gca, 'YTick', [])  
xlabel('CMP', 'FontSize', 14)  
legend('CMP Sustitutos', 'CMP Serie original')
```

## Referencias

- ABARBANEL, H. (1996). *Analysis of observed chaotic data*. Springer, pp. 268.
- BABAI, M. Z.; ALI, M. M.; BOYLAN, J. E. & SYNTETOS, A. A. (2011). Forecasting and inventory performance in a two-stage supply chain with ARIMA(0,1,1) demand: Theory and empirical analysis, *International Journal of Production Economics*, Available online 9 September 2011.
- BOWERMAN, B.; O'CONNELL, R. & KOEHLER, A. (2006). *Pronósticos, Series de Tiempo y Regresión*. Cengage, 4<sup>ta</sup> Edición, p. 663.
- BOX, G. E. P.; JENKINS, G. M. & REINSEL, G. (1994). *Time Series Analysis: Forecasting and Control*. Prentice Hall.
- BROCKWELL, P. J. & DAVIS, R. A. (1996). *Introduction to Time Series Analysis and Forecasting*. Springer-Verlag.
- CHANG, L.; GAO, W.; ZHANG, X.; LU, Y. & PAN, R. (2009a). Review of researches on fashion color prediction based on Grey systems theory. *Grey Systems and Intelligent Services*, 2009. GSIS 2009. IEEE International Conference on, vol., no., pp.383-387, 10-12 Nov.
- CHANG, L.; GAO, W. & ZHANG, X. (2009b). Discussion on Fashion Color Forecasting Researches for Textile and Fashion Industries. *Journal of Fiber Bioengineering and Informatics*, Review, JFBI Vol. 2 No. 1.
- CHEN, W.; WANG, B. & ZHAN, H. (2009). A Novel Grey Forecasting Model. *Fuzzy Systems and Knowledge Discovery*, 2009. FSKD '09. Sixth International Conference on, vol.1, no., pp.571-575, 14-16 Aug. 2009.
- CHOI, T. M.; YU, Y. & AU, F. (2011). A hybrid SARIMA wavelet transform method for sales forecasting. *Decision Support Systems*, Volume 51, Issue 1, April, Pages 130-140.

- CHOI, T.M.; HUI, C.; NG, S.; YU, Y. (2011). Color Trend Forecasting of Fashionable Products with Very Few Historical Data. *Systems, Man, and Cybernetics, Part C: Applications and Reviews*, IEEE Transactions on , vol.PP, no.99, pp. 1-8.
- CUI W. (2007). Analyses on 1996-2005 inter color spring/summer fashion color. *J International Fashion & Fabrics*, 3, pp. 27.
- DAVIS, J. C. (1986). *Statistics and data analysis in geology*. 2 edition, John Wiley & Sons Inc.
- DENG, Q. (1989). Statistical theory of fashionable colors. *Chinese Journal of Computational Physics*, 6, pp. 27-37.
- FRASER, A. M. & SWINNEY, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Phys A*, 33, pp. 1134-1140.
- GALLER, R. G. (1968). *Information Theory and Reliable Communication*. JohnWiley and Son.
- GRASSBERGER, P. & PROCACCIA, I (1983a). Characterization of strange attractors. *Phys A*, 50, pp. 346-349.
- \_\_\_\_\_. (1983b). Measuring the strangeness of strange attractors. *Physica, D*, 9, pp. 189-208.
- GU, W. & Liu, X. (2010). System of Color Resource Management for Fashion Color Forecasting. *Management and Service Science (MASS)*, 2010. International Conference on , vol., no., pp.1-4, 24-26 Aug. 2011.
- HARVEY, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
- HARVEY, A. C.; RUIZ, E. & SHEPHARD, N. (1994). Multivariate stochastic variance models, *Review of Economic Studies*, 61, 247-264.
- JACQUIER, E., POLSON, N. G., & ROSSI, P. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of business & Economic Statistics*, 12, 371-417.
- KASPAR, F. & SCHUSTER, H. G. (1997). Easily calculable measure for the complexity of spatio temporal patterns. *Phys A*, 36, pp. 842-848.



- KOREISHA, S. & PUKKILA, T. M. (1990). A generalized least-squares approach for estimation of autoregressive moving average models. *Journal of Time Series Analysis*, 11, pp. 139-151.
- LINTON, H. (1994). *Color forecasting: A survey of international color marketing*. New Jersey, Van Nostrand Reinhold.
- LEMPEL, A. & ZIV, J. (1976). On the complexity of finite sequences. *IEEE Transactions on Information Theory*, IT-22 (1), pp. 75-81.
- \_\_\_\_\_. (1978). Compression of individual sequences via variable-rate coding. *IEEE Transactions on Information Theory*, IT-24 (5), pp. 530-536.
- MATILLA, M. & VÁSQUEZ, F.J. (2000). Determinismo, No Linealidad y Caos comentarios y reflexiones sobre la obra Caos y Orden. *Empiria*, Vol. 3200, pp. 167-175.
- McLEOD, A. I. & LI, W. K. (1983): Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4, 269-273.
- NEWLAND, D. E. (1993). *An introduction to random vibrations, spectral and wavelet analysis*. Longman, Harlow.
- OSPINA G., L. M.; SOTO M., J. A. & OROZCO G., A. A. (2011). *Determinismo y no linealidad en señales financieras: Método de los datos sustitutos*. Pereira, Universidad Tecnológica de Pereira.
- PARKER, T. S. & CHUA, L. O. (1987). Chaos: A tutorial for engineers. *In IEEE Proceedings*, Vol. 75, pp. 982-1008.
- PEÑA, D. (2010). *Análisis de series temporales*. Madrid, Alianza.
- PEÑA, D. & RODRÍGUEZ, J. (2006): The Log of the Determinant of the Autocorrelation Matrix for Testing Goodness of Fit in Time Series. *Journal of Statistical Planning and Inference*, 136, 2706-2718.
- PEÑA, D. & SÁNCHEZ, I. (2005). Multifold Predictive Validation in Time Series Models. *The Journal of American Statistical Association*, V. 100, No.469, pp. 135-146.
- PEÑA, D.; TIAO, G. C. & TSAY, R. S. (2001). *A Course in Time Series Analysis (Wiley Series in Probability and Statistics)*. John Wiley & Sons Inc.

- PRIESTLEY, M. B. (1991). *Spectral analysis and time series*. Academic Press.
- SÁNCHEZ, I. & PEÑA, D. (2001). Properties of Predictors In Overdifferenced Nearly Nonstationary Autoregression. *The Journal of Time Series Analysis*, V. 22, pp. 45-46.
- SCHWARZ, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, pp. 461-464.
- SMALL, M. (2005). Applied NonLinear Time Series Analysis, Applications in Physics, Physiology and Finance, *World Scientific Publishing*, ISBN 981-256-117-x, pp. 241, 2005.
- SMALL, M. & TSE, C. K. (2003). Determinism in Financial Time Series. *Studies in Nonlinear Dynamics & Econometrics*, Vol. 7, Article 5.
- SMALL, M.; YU, D & R. HARRISON, R. (2001). Surrogate test for pseudoperiodic time series data. *Physical Review Letters*. Vol. 87, pp. 188-101.
- SOLÉ, R.V.; BASCOMPTE, J.; DELGADO, J.; LUQUE, B. & MANRUBIA, S. C. (1996). Complejidad en la frontera del caos. *Investigación y ciencia*, Vol. 236, pp. 14-21.
- STANSTIELD, J. & WHITE, A\*. (2005). Can Future Color Trends Be Predicted on the Basis of Past Color Trends?, An Empirical Investigation. *In Color Research & Application*, 2005, 30(3): pp.235-242.
- STROGATZ, S.H. (1994). *Nonlinear Dynamics and Chaos*. Perseus Books Publishing. pp. 496.
- SUN, Z. L.; CHOI, M.; AU, F. & YU, Y. (2008). Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Syst.*, vol. 46, no. 1, pp. 411-419, Dec.
- TAKENS, F. (1991). The Takens embedding theorem. *International Journal of Bifurcation and Chaos*, Vol. 1, pp. 867-872.
- THEILER, J.; GALDRIKIAN, B.; LONGTIN, A.; EUBANK, S. & FARMER, J. D. (1992). Testing for nonlinearity in time series: The method of surrogate data. *Physica* Vol. 58, pp. 77-94.
- TOSTTHEIM, D. & PAULSEN, J. (1983). Bias os some commonly-used time series estimates. *Biometrika*, 70, pp. 389-399.

- VALENZUELA, C.F. (2011). *Predicción de largo plazo para la generación eólica mediante Modelos Grises*, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas.
- WEST, M. & HARRISON, J. (1989). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag.
- WOLD, H. O. (1938). *A study in the Analysis of Stationary Time Series*, Almqvist and Wicksell, Uppsala.
- WU, Y. Z.; SUN, L. & LE, J. J. (2010). *G-LMBPNN: A New Fashion Color Prediction Model*, *Computational Aspects of Social Networks (CASoN)*, International Conference on, vol., no., pp.501-504, 26-28 Sept.
- YU, Y.; CHOI, T. M. & HUI, C. L. (2011). An intelligent fast sales forecasting model for fashion products. *Expert Systems with Applications*, Volume 38, Issue 6, June, Pages 7373-7379.
- YU, Y.; HUI, C. L. & CHOI, T. M. (2012). An empirical study of intelligent expert systems on forecasting of fashion color trend. *Expert Systems with Applications*, Volume 39, Issue 4, March, Pages 4383-4389.