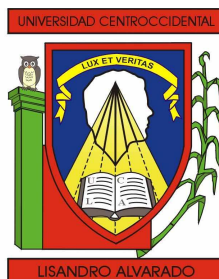


UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"

Decanato de Ciencias y Tecnología
Coordinación de Postgrado



ANÁLISIS ESPECTRAL DE OLAS MARINAS.

TRABAJO ESPECIAL DE GRADO PRESENTADO POR

Lcdo. Néstor J. Escudero M.

como requisito final
para obtener el título de Magister
en Ciencias Mención Física Matemáticas

Área de Conocimiento: Series Temporales.
Tutor: Dr. Joaquín Ortega Sánchez (CIMAT-México).
MSc. Jhonny Escalona(DCyT- UCLA)

Barquisimeto - Venezuela
Septiembre 2012

*A Dios Todopoderoso por ser mi guía, a
mi familia presente y futura...*

AGRADECIMIENTOS

A Dios todopoderoso por permitirme alcanzar otra meta en mí vida. A mis padres y a mi hermana por ser mi apoyo en todo momento y a nuestra Valeria por ser el motivo de nuestra felicidad. A mis tios, en especial Tio Ramón y Tia Lelis, a mis primos muy especialmente a Hildemar, Katty, Ángel y Hector. Gracias por el apoyo y por estar siempre presentes.

A quienes colocaron toda su confianza en mí y me dieron la oportunidad de aprender tanto de ellos, a mis tutores: Joaquín Ortega y Jhonny Escalona. Es un verdadero honor ser estudiante de ustedes, agradecido eternamente por sus conocimientos y por ser incondicionales. Tambien al Dr. Ernesto Mordecki, a quien agradezco enormemente su confianza y sus concejos.

Al Centro de Investigación en Matemáticas, AC por permitirme ser parte de esa prestigiosa casa de estudios y abrirme las puertas a continuar desarrollandome como investigador al realizar este trabajo, en especial quiero agradecer al Dr. Victor Rive-ro, quien siempre me tendió la mano y me ayudo al proceso de adaptación. Tambien quiero agradecer el apoyo, confianza y receptividad de mis compañeros y amigos: Harold, Laura, Noemí, Airam Blancas, Guillermo, al equipo de volleyball PiQ2, Anahí y Enrique. Les recuerdo siempre.

A mis compañeros y amigos de la Maestría en Ciencias del DCyT, Octalis, Marco, Gladymar, Jessica, David, Luis y Gaby. Gracias por el apoyo y la amistad de años que hemos construido. A los Profesores, Rafael Torrealba, Wilfredo Angulo, siempre agradecido por el apoyo que me brindan.

A mis amigos de siempre: Liliana, Reydis, Norelia, Francis, Yackson, Luis Alexander, Luis Alejandro, Mariangely y Luis Alcides, han sido importantes en este camino con sus concejos, comprensión y por permitirme entrar en sus familias como uno más.

A todos aquellos que de alguna u otra forma me apoyaron, muchas gracias!

RESUMEN

En este trabajo se desarrolló el análisis espectral a datos de olas del Mar del Norte recabados en el año 1.999. Forman parte de una base de datos, y corresponden a más de 15 años en los cuales se han producido una cantidad importante de tormentas, estos datos corresponden a la altura de olas tomados en la plataforma North Alwyn en el Mar del Norte y almacenados en la universidad de Heriot-Watt de Edimburgo, en este proyecto analizamos una de estas. Los datos están separados en periodos de 20 minutos y fueron registrados a una frecuencia de 5 Hz.

Para el estudio, se grafico el espectro de los 244 periodos disponibles, con el fin de analizar su comportamiento. El área total bajo la curva en cada periodo representa la energía presente en la serie de tiempo durante el periodo. En los extremos de este conjunto de datos se observó que existía información que no era necesaria tomar en cuenta, ya que representa ruido, esta información representa el 3 % y no fue tomada en cuenta para el análisis. Luego, se procedió a discretizar la energía en intervalos, escogidos de modo que capturen dentro de cada uno de ellos los puntos críticos de la energía. Esta partición se eligió de acuerdo a múltiples apreciaciones graficas, y como producto de ello, se determino que era necesaria una partición de 12 subintervalos para realizar el presente estudio. Luego de analizar las funciones de autocovarianza (ACF) y autocovarianza parcial (PACF) para el modelo autorregresivo (AR) por cada serie de tiempo y observar las ACF y PACF para distintos ordenes de estos modelos se determino que tenían asociado un modelo AR de orden uno, también apoyados en graficas para el diagnostico del modelo observando además los resultados gráficos del estadístico Ljun Box. Un análisis adicional para determinar este buen ajuste, es que al observar la grafica de cuantiles, algunos residuales de estos modelos no ajustaban con la distribución normal, lo que conlleva a usar la transformación de Box-Cox con el que se intentó corregir el problema de la no-normalidad. Con esto, se encontró un conjunto de 12 modelos autorregresivos de orden uno que posteriormente servirán para el análisis del conjunto de modelos en estudio y así estudiar la relación e interacción que existe entre ellos para poder dar explicaciones más explicitas acerca del comportamiento de estos datos.

Seguidamente, se construyeron los modelos vectoriales autorregresivos con los modelos univariados de donde resulto viable construir modelos vectoriales dos a dos, es decir, para dos modelos contiguos se obtuvo un vector autorregresivo al cual se realizó un análisis donde se estudiaron las características entre las variables del modelo en estudio, partiendo de la correlación que existe entre las variables, se determino la cointegración, también se verificó que no existe posibilidad de raíces unitarias para la estabilidad del modelo vectorial. Se determino que el orden del modelo vectorial autorregresivo es uno para cada par de variables contiguas, también se realizaron los estudios para la ACF y los pronósticos correspondientes para determinar si el modelo encontrado es capaz de aproximar a los datos originales (ver [3]).

Palabras claves: Espectro, modelos autorregresivos, transformación de Box-Cox, vectores autorregresivos.

ÍNDICE

Agradecimientos	i
Resumen	iii
Introducción	1
1. Preliminares Matemáticos	5
1.1. Series de Tiempo Estacionarias	5
1.2. Modelos ARMA	7
1.2.1. Modelos Autorregresivos	8
1.2.2. Modelos de Medias Móviles (MA)	9
1.2.3. Modelos Autorregresivos con Media Móvil	10
1.2.4. Sobre la Transformación de Box-Cox	12
1.3. Modelos Vectoriales Autorregresivos (VAR)	13
1.3.1. Procesos Multivariados ARMA	14
1.4. Análisis de Ajuste de Modelos según Box y Jenkins	15
1.4.1. Identificación y Comportamiento de la ACF PACF	18
1.5. Sobre el Teorema de Bochner y la Representación Espectral.	18
1.5.1. Función Covarianza para Procesos Reales.	19
1.5.2. Funciones Definidas No-negativas.	19
1.5.3. El Teorema de Bochner y la Distribución Espectral.	19
1.5.4. La Distribución Espectral.	19
1.5.5. Consecuencias del Teorema de Bochner	20
1.6. Características de las Olas.	22
1.6.1. Altura Significativa.	22
1.6.2. Características Basadas en Cruces Nivel Medio.	22
1.6.3. Características Basadas Extremos Sucesivos.	23
1.6.4. Momento Espectral de Orden n	24
1.6.5. Modelo de Banda Estrecha.	25

1.6.6. Modelo de Banda Ancha.	25
1.7. Hipótesis Necesarias para Olas Marinas	28
2. Análisis Espectral de los Datos y Modelos Asociados	31
2.1. Descripción de los Datos	31
2.2. Estudio del Comportamiento del Espectro	33
2.2.1. Modelos Asociados al Estudio de la Energía	33
2.2.2. Sobre la Construcción de los Modelos	34
2.3. Características de los Modelos	35
2.4. Comparación de los Modelos Obtenidos	37
2.5. Características de los Residuos	39
2.6. Residuos del Modelo Asociado al Tercer Subintervalo	40
2.7. Residuales de los Modelos	41
2.8. Pronósticos de los Modelos Asociados a cada Subintervalo de $[0,04, 0,25]$. 43	
2.8.1. Pronóstico para Subintervalos con la Transformación de Box-Cox. 43	
2.8.2. Pronóstico para los Subintervalos que no Requieren Transformación Box-Cox. 45	
3. Modelos Vectoriales Autorregresivos VAR	47
3.1. Modelos VAR Asociado al Primer y Segundo Subintervalo.	48
3.2. Modelo VAR Asociado al Segundo y Tercer Subintervalo.	52
3.2.1. Modelo Asociado al Tercer y Cuarto Subintervalo.	55
3.3. Modelo Asociado al Cuarto y Quinto Subintervalo.	58
3.4. Modelo Asociado al Quinto y sexto Subintervalo	61
3.5. Modelo Asociado al Sexto y Séptimo Subintervalo.	64
3.6. Modelo Asociado al Séptimo y Octavo Subintervalo.	67
3.7. Modelo Asociado al Octavo y Noveno Subintervalo.	70
3.8. Modelo Asociado al Noveno y Décimo Subintervalo.	75
3.9. Modelo Asociado al Décimo Primer y Décimo Segundo Subintervalo.	77
4. Conclusiones	83
A. Modelos Univariados	87
A.1. Subintervalos que requieren de la Transformación de Box Cox.	87
A.1.1. Segundo Subintervalo.	87
A.1.2. Quinto Subintervalo.	90

A.1.3. Sexto Subintervalo.	93
A.1.4. Séptimo Subintervalo.	96
A.1.5. Octavo Subintervalo.	99
A.2. Subintervalos que no Requieren la Transformación Box Cox.	102
A.2.1. Cuarto Subintervalo.	102
A.2.2. Noveno Subintervalo.	104
A.2.3. Décimo Subintervalo.	107
A.2.4. Décimo Primer Subintervalo.	109
A.2.5. Décimo Segundo Subintervalo.	111

Referencias	115
--------------------	------------

INTRODUCCIÓN

El uso de los procesos aleatorios en el estudio y la modelación de las olas marinas, se inicia en la década de los 50 con el trabajo de Longuet-Higgins, y desde entonces ha tenido un enorme desarrollo (ver por ejemplo [4]). Los modelos basados en procesos aleatorios han mostrado ser de gran utilidad permitiendo el estudio de diversas características de las olas para su estudio.

Las olas marinas, son ondas sísmicas (movimientos de un medio material) que se propagan entre dos medios materiales. En aguas profundas, las moléculas de agua regresan casi al mismo lugar donde se encontraban al principio. En una ola se distinguen dos movimientos: la oscilación del medio movido por la onda, el cual es un movimiento circular y la propagación de la onda, que se produce porque la energía se transmite con ella, trasladando el fenómeno con una dirección y una velocidad (velocidad de la onda).

Una característica muy importante de las olas marinas, es que cuanto mayor es su altura mayor es la cantidad de la energía que se produce al extraer del viento, ya que la altura depende de la velocidad, su persistencia en el tiempo y la estabilidad de la dirección de este. Por otro lado, las olas que se desplazan sobre aguas profundas disipan su energía muy lentamente, de forma que alcanzan regiones separadas de su lugar de formación. Hay tres formas que se disipe la energía: una parte puede convertirse en corriente superficial, por fricción sobre el aire, por interacción con la corteza sólida.

Una herramienta muy útil con la que se pueden estudiar datos de olas marinas, es el análisis del espectro, también conocida como análisis de dominio de frecuencia o estimación de la densidad espectral, es el proceso técnico de la descomposición de una señal compleja en partes más simples. Como se describió anteriormente, muchos de los procesos físicos que se describe mejor como una suma de muchos componentes de frecuencia individuales. Cualquier proceso que cuantifica las diversas cantidades (por ejemplo, amplitudes, potencias, intensidades, o fases), frente a la frecuencia puede ser llamado análisis de espectro. El análisis del espectro se puede realizar en toda la señal. Alternativamente, una señal puede ser dividida en segmentos cortos y el análisis del espectro se puede aplicar a estos segmentos individuales.

Los modelos usuales para olas en mares profundos y condiciones normales son los

procesos Gaussianos y centrados (ver de nuevo [4]). Estas hipótesis permiten obtener modelos manejables, que en muchas situaciones resultan adecuados. La hipótesis de estacionaridad permite el uso de técnicas de análisis espectral de Fourier para el estudio de distribución de energía de la ola en las distintas frecuencias que la componen. Este análisis espectral está asociado a distintas características de interés, como por ejemplo la altura significativa o el período dominante que puede calcularse a partir de la distribución espectral o de sus momentos. El modelo Gaussiano, además de resultar una muy buena aproximación de primer orden, permite obtener expresiones explícitas para muchas de las distribuciones de los parámetros que resultan importantes para los oceanógrafos e ingenieros. Sin embargo, ambas hipótesis; estacionaridad y Gaussianidad tienen sus limitaciones, una de tales limitaciones es que las características del mar cambian bruscamente con el tiempo, sin embargo se puede considerar que en algún período suficientemente prolongado de condiciones meteorológicas estables, la distribución de la altura del mar no cambia con el tiempo en algún punto específico del espacio. Otras características fundamentales que como hipótesis se requieren para obtener los deseados modelos manejables para datos de olas, es que se requiere que la trayectoria del proceso que sigue la ola sea continua, la ergodicidad, la cual se traduce en poder sustituir los valores esperados (teóricos) por promedios temporales (empíricos). Estas dos suposiciones se requieren también para deducir propiedades importantes del proceso que sigue la ola.

Como consecuencia de la hipótesis de estacionaridad, la función covarianza del proceso y el proceso mismo tiene una representación espectral. En el caso de la función covarianza, esta representación se obtiene a partir del Teorema de Bochner. La densidad espectral o espectro, juega un rol principal en el estudio de las olas marinas y se interpreta como la distribución de la energía por frecuencias. Esta densidad en una función simétrica y normalmente solo se considera la parte positiva. La integral de la densidad espectral es la varianza del proceso y también corresponde a la energía total presente durante el período de tiempo en consideración. Los momentos espectrales, que son los momentos calculados respecto a la densidad espectral también juegan un rol importante en esta teoría.

De acuerdo a la evolución del espectro de la ola o del conjunto de olas, es preciso detallar la técnica que se usará en este trabajo para ajustar modelos adecuados bajo las hipótesis anteriores. El procedimiento consiste en desprestigiar la energía que

se considera ruido situada en los extremos y discretizar el intervalo restante en una partición que capture la mayor cantidad posible de características en la evolución de la energía. A cada subintervalo de esta partición se le ajusta un modelo ARMA cuyo orden se determina de acuerdo a la ACF, PACF y las características asociadas a sus p valores y al valor de la AIC (Akaike Information Criteria) asociado. Seguidamente, para completar el análisis de estos modelos univariados, se construirá un *modelo vectorial autorregresivo* (VAR) el cual reúne el conjunto completo de modelos univariados tantos como intervalos halla, de tal manera que se pueda obtener información más precisa acerca de la evolución de los espectros de las olas [3].

Este trabajo se estructuró de la siguiente manera: En el Capítulo 2 se hace un resumen de la teoría necesaria para tratar con el problema en estudio, en el se da una descripción los modelos ARMA, además de sus casos particulares y se establecen las diferencias entre distintos tipos de modelos que se presentan y se pueden ajustar. También se definen matemáticamente las hipótesis necesarias para poder obtener modelos adecuados y manejables para datos de olas marinas. El Capítulo 3 se detallan los modelos autorregresivos univariados resultantes de las pruebas realizadas con los datos en estudio, aplicando la técnica antes descrita para ajustar modelos a la evolución de la energía considerando una serie de análisis para los datos de manera tal de comprobar la veracidad del orden y tipo de estos modelos haciendo uso del software R. Con esta herramienta usando las pruebas estadísticas y diversas gráficas se puede comprobar que los modelos ajustados son apropiados para dar respuesta al comportamiento de los datos en estudio para las olas marinas del Mar del Norte.

El el Capítulo 4, se incluyen los resultados más generales y relevantes en el presente trabajo, estos son los modelos vectoriales autorregresivos construidos a partir de los modelos univariados, estos modelos vectoriales se construyen dos a dos para los modelos univariados contiguos obtenidos en el capítulo anterior, estos vienen a explicar de manera más precisas el comportamiento de los datos. Tales modelos también son sometidos a pruebas graficas y estadísticas para validarlos. Finalmente, en el Capítulo ?? se presentan las conclusiones que se derivan del estudio, también se da una perspectiva del mismo y se dan las recomendaciones consideradas necesarias para ser tomadas en cuenta en trabajos futuros.

CAPÍTULO 1

PRELIMINARES MATEMÁTICOS

A continuación se presentan algunas definiciones y resultados teóricos generales sobre los modelos que estudiamos, pueden consultarse en las referencias [1], [2], [3] y [4].

1.1. Series de Tiempo Estacionarias

En este punto, se pretende dejar claras las definiciones que explican la estacionaridad de un proceso aleatorio y las definiciones que estas traen como consecuencia como parte complementaria para entender las propiedades de las olas marinas. Se introducirán los conceptos básicos necesarios para adentrar en el tema de las olas marinas desde el punto de vista de los procesos estocásticos, el cual será el enfoque que tomaremos en este trabajo, pero antes, recordaremos unos conceptos necesarios para adentrar en el contexto de las olas marinas, el de la función media, función covarianza, entre otros.

Definición 1.1. La función **media** de un proceso aleatorio x_t se define como

$$\mu_{x_t} = E(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx \quad (1.1)$$

donde E denota el operador valor esperado usual, $f_t(x)$ es la función de densidad de la variable aleatoria x_t , cuando no exista confusión alguna usaremos μ_t en vez de μ_{x_t} .

Definición 1.2. La **función autocovarianza** se define como

$$\gamma_x(s, t) = E[(x_s - \mu_s)(x_t - \mu_t)] \quad (1.2)$$

para todo t y s . De nuevo, cuando no haya confusión escribimos $\gamma(s, t)$ en vez de $\gamma_x(s, t)$.

Nótese que $\gamma_x(s, t) = \gamma_x(t, s)$ para cualesquiera t y s . La función autocovarianza mide la dependencia lineal entre dos puntos de la serie observada en tiempos diferentes. Para $s = t$ la autocovarianza se reduce a la varianza (asumiendo que es finita), dado que

$$\gamma_x(t, t) = E[(x_t - \mu_t)^2] \quad (1.3)$$

Definición 1.3. La **función autocorrelación (ACF)** se define como

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}} \quad (1.4)$$

La ACF mide la dependencia lineal de la serie en el tiempo t , es decir, la dependencia de x_t , de los valores de x_s . Un resultado clásico en estadística es que $-1 \leq \rho(s, t) \leq 1$ que se obtiene usando la desigualdad de Cauchy Schwarz. Si podemos predecir x_t perfectamente de x_s a través de la relación lineal $x_t = \beta_0 + \beta_1 x_s$, entonces la correlación será 1 si $\beta_1 > 0$ y -1 cuando $\beta_1 < 0$. Así, tenemos una medida aproximada de la capacidad de predecir la serie en el tiempo t a partir del tiempo s .

Definición 1.4. Una serie de tiempo **estrictamente estacionaria** es aquella para la cual, la distribución de probabilidad de toda colección de valores

$$\{x_{t_1}, x_{t_2}, \dots, x_{t_k}\}$$

es idéntica a la del conjunto

$$\{x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\}$$

Esto es

$$P\{x_{t_1} \leq c_1, \dots, x_{t_k} \leq c_k\} = P\{x_{t_1+h} \leq c_1, \dots, x_{t_k+h} \leq c_k\} \quad (1.5)$$

para todo $k = 1, 2, \dots$ y todo tiempo t_1, t_2, \dots, t_k , todos los números c_1, c_2, \dots, c_k y todo $h \in \mathbb{R}$.

La versión de estacionaridad en (1.5) es muy fuerte para muchas aplicaciones. Es difícil calcular la estacionaridad estricta para un sólo conjunto de datos. En vez de las condiciones impuestas sobre todas las posibles distribuciones de una serie de tiempo, usamos una versión débil que impone condiciones sólo sobre los dos primeros momentos de las series. Luego, sigue la siguiente definición.

Definición 1.5. Una serie de tiempo **débilmente estacionaria**, x_t , es un proceso con varianza finita tal que

1. la función de valor medio, μ_t definida en (1.1) es constante y no depende del tiempo t , y
2. la función covarianza, $\gamma(s, t)$ definida en (1.2) depende de s y t , solo a través de la diferencia $|s - t|$.

Una consecuencia de esta definición, es que la función covarianza depende sólo del parámetro h , es decir

Definición 1.6. La **función autocovarianza de una serie de tiempo estacionaria** se escribe como

$$\gamma(h) = E[(x_{t+h} - \mu)(x_t - \mu)] \quad (1.6)$$

Definición 1.7. La **función autocorrelación (ACF) de una serie de tiempo estacionaria** se escribe de la forma

$$\rho(h) = \frac{\gamma(t+h, t)}{\sqrt{\gamma(t+h, t+h)\gamma(t, t)}} \quad (1.7)$$

Los conceptos antes definidos serán usados para determinar características para los modelos de olas en el presente análisis. Estas características están muy conectadas con la determinación precisa de los modelos que se desean hallar, es decir, a partir de ellas estaremos en capacidad de decidir cuáles son los modelos a utilizar en el estudio.

1.2. Modelos ARMA

En esta sección se presentan algunos conceptos relativos a los distintos modelos que se pueden ajustar a una serie de tiempo. A continuación se presenta una serie de definiciones que se requieren para entender lo referente a los modelos ARMA, sus distintos tipos y casos que pueden presentarse. Los modelos que se tratan como modelos ARMA son los modelos autorregresivos (AR) los modelos de media móvil (MA) y la forma combinada de los dos primeros, los modelos ARMA. Dos herramientas fundamentales para el uso de estos modelos son la ACF y la PACF la cuales determinan el orden y el tipo de modelos que debe usarse para los datos en estudio (véase [1]).

1.2.1. Modelos Autorregresivos

Los modelos autorregresivos están basados en la idea de que los valores presentes de la serie x_t pueden ser explicados en función de sus p valores pasados $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ donde p determina el número de pasos en el pasado necesarios para explicar los valores presentes de la serie de tiempo. En esta sección se presentan los aspectos necesarios sobre los modelos autorregresivos, los cuales serán útiles en la decisión acerca del tipo de modelos que se estudian en este trabajo. La discusión precedente justifica las siguientes definiciones.

Definición 1.8. Un **modelo autorregresivo de orden p** , abreviado $AR(p)$, es una serie estacionaria de la forma

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \quad (1.8)$$

donde x_t es estacionaria, ϕ_i , $i = 1, \dots, p$, ($\phi_p \neq 0$) son constantes, asumimos w_t es un ruido blanco.

Observación 1. Si la media de x_t es distinta de cero, reemplazamos x_t por $x_t - \mu$ en la ecuación anterior, es decir,

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

o escribimos,

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \quad (1.9)$$

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

Nótese que (1.9) es similar a un modelo de regresión clásico. Una forma usual para escribir los modelos $AR(p)$ es mediante el operador **de retardo o backshift** ([1] capítulo 2) que se define de la forma

$$B(x_t) = x_{t-1} \quad (1.10)$$

usando este operador, la ecuación (1.8) se puede escribir de la forma

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)x_t = w_t \quad (1.11)$$

o en forma compacta,

$$\phi(B)x_t = w_t \quad (1.12)$$

siendo $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$. De (1.12) se pueden obtener las propiedades de $\phi(B)$, formalmente para $\phi(B)$ se tiene la siguiente definición.

Definición 1.9. El operador autorregresivo esta definido como

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (1.13)$$

Consideremos el modelo de primer orden, $AR(1)$, dado por $x_t = \phi x_{t-1} + w_t$. Iterando hacia atrás k veces, tenemos

$$\begin{aligned} x_t &= \phi x_{t-1} + w_t = \phi(\phi x_{t-2} + w_{t-1}) + w_t \\ &= \phi^2 x_{t-2} + \phi w_{t-1} + w_t \\ &= \vdots \\ &= \phi^k x_{t-k} + \sum_{j=0}^{k-1} \phi^j w_{t-j} \end{aligned}$$

Este método sugiere que, si continuamos hacia atrás y además tenemos que $|\phi| < 1$ y x_t es estacionaria, podemos representar un modelo $AR(1)$ como un proceso lineal dado por

$$x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j} \quad (1.14)$$

El $AR(1)$ definido en (1.14) es estacionario con media

$$E(x_t) = 0$$

y función de autocovarianza,

$$\gamma(h) = cov(x_{t+h}, x_t) = \frac{\sigma_w^2 \phi^h}{1 - \phi^2}, h \geq 0 \quad (1.15)$$

Para más detalles vease [1]

1.2.2. Modelos de Medias Móviles (MA)

Estos procesos son aquellos que se estructuran estableciendo una relación de dependencia entre la variable que se modela y el ruido w_t . Si existe un sólo término de ruido blanco, entonces el proceso será de orden 1, si en cambio existe una combinación lineal de q términos de ruido blanco, el proceso es denominando $MA(q)$. A continuación se presenta una pequeña reseña acerca de las definiciones para este tipo de procesos.

Definición 1.10. El **modelo de media móvil de orden q** o modelo $MA(q)$ se define como

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q} \quad (1.16)$$

donde existen q términos para la media móvil y $\theta_1, \dots, \theta_q$ ($\theta_q \neq 0$) son parámetros. El ruido w_t se asume como ruido blanco.

También podemos escribir el proceso $MA(q)$ en forma equivalente

$$x_t = \theta(B)w_t \quad (1.17)$$

usando la siguiente definición

Definición 1.11. El **operador de media móvil** es

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \quad (1.18)$$

El proceso de medias móviles es estacionario para cualquier valor de los parámetros $\theta_1, \dots, \theta_q$.

1.2.3. Modelos Autorregresivos con Media Móvil

Ahora procedemos con el desarrollo de modelos autorregresivos, media móvil, y una mezcla de media móvil autorregresivo (ARMA) para series de tiempo estacionarias.

Definición 1.12. Una serie de tiempo $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$ es $ARMA(p, q)$, si es estacionaria y,

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q} \quad (1.19)$$

con $\phi_p \neq 0$, $\theta_q \neq 0$ y $\sigma_w^2 > 0$. Los parámetros p y q son llamados órdenes autorregresivo y de media móvil, respectivamente.

Si x_t tiene media $\mu \neq 0$, ponemos $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$ y escribimos el modelo como

$$x_t = \alpha + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q} \quad (1.20)$$

A menos que se indique lo contrario, $\{w_t : t = 0, \pm 1, \pm 2, \dots\}$ es un ruido gaussiano.

Observación 2. Cuando $q = 0$ obtenemos el modelo autorregresivo de orden p , $AR(p)$, cuando $p = 0$, obtenemos el modelo de media móvil de orden q , $MA(q)$. Para investigar los modelos ARMA será muy útil usar el operador AR , (1.13), y el operador MA , (1.18), el modelo $ARMA(p, q)$ en (1.19) se escribe en forma concisa como

$$\phi(B)x_t = \theta(B)w_t$$

Existen algunas dificultades al tratar con modelos ARMA en general, los siguientes problemas se presentan a la hora de tratar con estos modelos:

1. Parámetros redundantes del modelo.
2. Modelos AR estacionarios que dependen del futuro, y
3. Modelos MA que no son únicos

Para solucionar estos problemas, requerimos algunas restricciones adicionales sobre el modelo. Primero se presentan las siguientes definiciones,

Definición 1.13. Los **polinomios AR y MA** están definidos como

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p, \quad \phi_p \neq 0 \quad (1.21)$$

y

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q, \quad \theta_q \neq 0 \quad (1.22)$$

respectivamente, donde $z \in \mathbb{C}$

Para resolver los problemas enumerados anteriormente, nos tenemos que referir al modelo $ARMA(p, q)$ en su forma simple, es decir, además de su definición original (1.19) también requerimos que al construir los polinomios $\phi(z)$ y $\theta(z)$ provenientes de x_t no tengan factor en común o que no definan un modelo $ARMA(p, q)$.

Por otro lado, para el problema de modelos dependientes del futuro, formalmente introducimos el concepto de causalidad

Definición 1.14. Un modelo $ARMA(p, q)$ $\phi(B)x_t = \theta(B)w_t$ se dice **causal**, si la serie de tiempo x_t , $t = 0, \pm 1, \pm 2, \dots$ se puede escribir como un proceso lineal,

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B)w_t \quad (1.23)$$

donde $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ y $\sum_{j=0}^{\infty} |\psi_j| < \infty$, ponemos $\psi_0 = 1$

En general, la causalidad cumple las siguientes propiedades,

Propiedad 1 (Causalidad de un modelo $ARMA(p, q)$). Un modelo $ARMA(p, q)$ es **causal** sí y solo sí $\phi(z) \neq 0$ para $|z| \leq 1$. Los coeficientes del proceso lineal dado en (1.23) pueden ser determinados resolviendo

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1$$

Una manera de expresar la propiedad anterior, es que un proceso $ARMA$ es causal solo cuando las raíces de $\phi(z)$ permanecen fuera del círculo unitario, esto es, $\phi(z) = 0$ sólo cuando $|z| > 1$. Finalmente, para resolver el problema de unicidad, escogemos el modelo que permite una representación autorregesiva infinita.

Definición 1.15. Un modelo $ARMA(p, q)$, $\phi(B)x_t = \theta(B)w_t$, se dice **invertible**, si la serie de tiempo $\{x_t\}$, $t = 0 \pm 1, \pm 2, \dots$ puede escribirse como

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t \quad (1.24)$$

donde $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ y $\sum_{j=0}^{\infty} |\pi_j| < \infty$; ponemos $\pi_0 = 1$

Analogamente a la propiedad anterior, tenemos la siguiente propiedad

Propiedad 2 (Invertibilidad de un proceso $ARMA(p, q)$). Un modelo $ARMA(p, q)$ es invertible, sí y sólo sí, $\theta(z) \neq 0$ para $|z| \leq 1$. El coeficiente π_j de $\phi(B)$ dado en (1.24) puede ser determinado resolviendo

$$\phi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1$$

1.2.4. Sobre la Transformación de Box-Cox

Muchas series temporales presentan tendencia, estacionaridad, variabilidad que crece con su nivel y componentes deterministas (valores atípicos, entre otros). Una

forma para eliminar este tipo de característica que se utiliza entre muchas otras, es la transformación de Box-Cox. Esta transformación se caracteriza por un parámetro λ que dependiendo de este valor la transformación es necesaria su aplicación o no. Cuando es aplicable, se debe a que los datos en estudio provienen de una distribución notablemente asimétrica y se aplica la transformación para inducir normalidad en los datos y reducir la desviación estándar de la serie. Otro caso en el cual la transformación de Box-Cox se aplica es cuando esta presente en el comportamiento de los cuantiles el fenómeno de *heterocedasticidad*, el cual hace que estos se comporten de forma que por un lado son cóncavos y por el otro toman un comportamiento convexo.

La transformación de potencia propuestas en 1964 por Box y Cox, así también llamada, consiste en determinar un parámetro λ para el cual se establecen ciertas condiciones, puesto que de acuerdo al valor de este parámetro se puede decidir el uso o no de la transformación en una serie de datos. Matemáticamente, la transformación de Box-Cox se define como sigue

$$y_t = \begin{cases} \frac{x_t^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln x_t, & \lambda = 0. \end{cases} \quad (1.25)$$

El valor de λ juega un papel fundamental en la ecuación anterior. Al aplicar la transformación de Box-Cox a un conjunto de datos para $\lambda \approx 1$ no es necesaria la aplicación de la transformación a estos datos, pero para $\lambda \neq 1$ es necesario aplicar la transformación a la serie de datos.

El objetivo principal en el análisis de la transformación de Box-Cox es hacer inferencia sobre el parámetro λ . Esta inferencia se hace por dos vías, el método de Máxima Verosimilitud el cual es el más común y fácil de usar y el otro método también usado es el Bayesiano.

1.3. Modelos Vectoriales Autorregresivos (VAR)

Muchas series de tiempo surgen en la práctica como componentes de alguna serie de tiempo a valores vectoriales (multivariada) $\{\mathbf{X}_t\}$ teniendo, no solo dependencia serial con cada componente de la serie $\{X_{ti}\}$ sino también interdependencia entre las diferentes componentes de la serie $\{X_{ti}\}$ y $\{X_{tj}\}$, $i \neq j$. Mucha de la teoría de las series de tiempo univariadas se extiende de manera natural al caso multivariado, sin embargo surgen algunos problemas.

En esta sección generalizaremos los conceptos de la Sección 1.2, de manera que podamos ver ahora los modelos antes definidos de manera vectorial, además de ver la manera en que las componentes vectoriales interactúan. Permitámonos ahora introducir algunas definiciones necesarias para precisar algunas otras ideas a continuación, para mayores detalles véase [2].

Definición 1.16. Una serie multivariada $\{\mathbf{W}_t\}$ es llamada **ruido blanco con media 0 y matriz de covarianza** Σ , escribimos

$$\{\mathbf{W}_t\} \sim iid(0, \Sigma)$$

sí los vectores aleatorios $\{\mathbf{W}_t\}$ son independientes e idénticamente distribuidos con media 0 y varianza Σ

1.3.1. Procesos Multivariados ARMA

Como en el caso univariado, podemos definir una clase muy usada de procesos estacionarios multivariados $\{\mathbf{X}_t\}$, requiriendo que este satisfaga un conjunto de ecuaciones en diferencia lineales con coeficientes constantes. La serie multivariada de la Definición 1.16 es un elemento fundamental para la construcción de estos procesos *ARMA*

Definición 1.17. $\{\mathbf{X}_t\}$ es un proceso *ARMA*(p, q), si $\{\mathbf{X}_t\}$ es estacionaria y si para todo t ,

$$\mathbf{X}_t - \Phi_1 \mathbf{X}_{t-1} - \dots - \Phi_p \mathbf{X}_{t-p} = \mathbf{W}_t + \Theta_1 \mathbf{W}_{t-1} + \dots + \Theta_q \mathbf{W}_{t-q} \quad (1.26)$$

donde $\{\mathbf{W}_t\} \sim N(\mathbf{0}, \Sigma)$. Análogamente a los procesos univariados $\{\mathbf{X}_t\}$ es un proceso *ARMA*(p, q) con media μ si $\{\mathbf{X}_t - \mu\}$ es un proceso *ARMA*(p, q)

La ecuación (1.26) puede ser escrita en forma compacta como

$$\Phi(B)\mathbf{X}_t = \Theta(B)\mathbf{W}_t, \quad \mathbf{W}_t \sim N(0, \Sigma) \quad (1.27)$$

donde $\Phi(z) := I - \Phi_1 z - \dots - \Phi_p z^p$ y $\Theta(z) := I + \Theta_1 z + \dots + \Theta_q z^q$ son matrices con valores polinomiales, I es una matriz identidad de $m \times m$ y B denota el operador usual backshift o de retardo. (Cada componente de las matrices $\Phi(z)$ y $\Theta(z)$ es un polinomio con coeficientes reales y grado menor o igual a p, q respectivamente).

Para dar una idea de este tipo de modelos presentamos explícitamente a continuación el modelo *AR*(1) en dimensión 2

Ejemplo 1.1. Haciendo $p = 1$ y $q = 0$ en (1.26) se obtiene la siguiente ecuación

$$\mathbf{X}_t = \Phi \mathbf{X}_{t-1} + \mathbf{W}_t, \quad \mathbf{W}_t \sim N(0, \Sigma) \quad (1.28)$$

donde,

$$\mathbf{X}_t = \begin{pmatrix} y_t \\ z_t \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}, \quad \mathbf{W}_t = \begin{pmatrix} w_{y,t} \\ w_{z,t} \end{pmatrix} \quad (1.29)$$

En este caso, usando el operador backshift la ecuación (1.26) se puede escribir como

$$(I - \Phi B)\mathbf{X}_t = \Phi(B)\mathbf{X}_t = \mathbf{W}_t \quad (1.30)$$

donde Φ es una matriz polinomial de orden 1 y I es la matriz identidad de orden 2×2 . Un proceso $VAR(1)$ puede ser extendido de manera natural a un proceso $VAR(p)$ para n series de tiempo puede definirse de manera análoga de la forma (1.29) en el cual I será entonces una matriz identidad de orden $n \times n$, Φ es un polinomio de matrices $n \times n$ de parámetros, \mathbf{X}_t es una matriz de $n \times 1$ de variables de las series de tiempo y \mathbf{W}_t es un ruido blanco multivariado. Para un modelo VAR, la ecuación característica está dada por el determinante de una matriz. Análogo a lo que ocurre con los modelos AR, un modelo $VAR(p)$ es estacionario, si las raíces del determinante $|\Phi(x)|$ todas exceden la unidad en valor absoluto. Para un modelo $VAR(1)$ en dimensión 2 el determinante esta dado por:

$$\det \begin{pmatrix} 1 - \phi_{11}x & -\phi_{12}x \\ -\phi_{21}x & 1 - \phi_{22}x \end{pmatrix} = (1 - \phi_{11}x)(1 - \phi_{22}x) - \phi_{12}\phi_{21}x^2 \quad (1.31)$$

Fácilmente podemos extender este procedimiento a un modelo $VAR(p)$, para más detalles consultese alguna de las referencias citadas, por ejemplo [3].

1.4. Análisis de Ajuste de Modelos según Box y Jenkins

El enfoque que proporcionan Box, Jenkins y Reinsel (1994) para la construcción de modelos para series temporales puede resumirse en el siguiente esquema

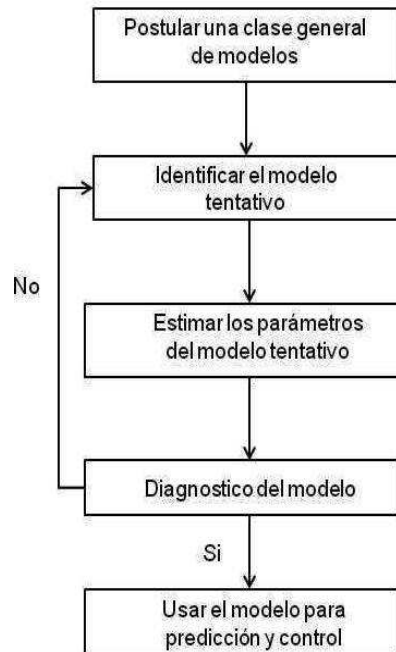


Figura 1.1: Esquema de Box, Jenkins y Reinsel.

La tarea es entonces identificar la clase de submodelo ARMA que pueda ser usado para representar una serie temporal. Para ello, se pueden seguir los siguientes pasos:

1. Diferenciar $\{X_t\}$ tantas veces como sea necesario para producir estacionaridad, con lo cual esperamos reducir el modelo en estudio a un modelo ARMA.
2. Identificar el proceso ARMA resultante.

Las principales herramientas para este fin serán la función de autocorrelación y la función de autocorrelación parcial estimadas:

$$r_k = \frac{\sum_{t=k+1}^n (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}$$

Si la serie es aleatoria, la distribución de referencia de r_k puede ser derivada (usando la distribución normal como distribución de los residuos, suponiendo que estos corresponden a ruido blanco).

3. Estimar los parámetros del modelo elegido para los modelos ARIMA, la estimación se hace usualmente por máxima verosimilitud.
4. Diagnóstico del modelo. Para ello, se emplean los residuos del ajuste. En efecto, si el modelo es adecuado, esperaríamos que los residuos se comporten como un

ruido blanco gaussiano. La presencia de autocorrelación o de falta de estacionariedad en los residuos implica que el modelo no es adecuado.

En general, el análisis de los residuos incluye:

- Gráfico de los residuos vs. el tiempo. Este gráfico permite identificar valores atípicos, falta de homogeneidad en la varianza y la presencia de estructuras en el tiempo.
- Análisis de la función de autocorrelación de los residuos. Sean $\hat{\rho}_k$ las autocorrelaciones estimadas para los residuos. Si el modelo es adecuado, los $\hat{\rho}_k$ deberían ser no correlacionados y aproximadamente normales con media cero y varianza n^{-1} . Por lo tanto, la presencia de autocorrelaciones grandes indica que el modelo puede ser inadecuado. Sin embargo es necesario tener cautela a la hora de analizar estas autocorrelaciones ya que puede probarse que para retardo pequeño n^{-1} puede sobrestimar la varianza real, con lo cual señales de falta de ajuste pudieran pasar desapercibidas
- Prueba de hipótesis sobre las autocorrelaciones. En lugar de examinar los $\hat{\rho}_k$ individualmente, es útil basar una prueba de diagnóstico en un conjunto de autocorrelaciones (típicamente entre 10 y 20). Se define el estadístico de Box-Pierce (también llamado *portmanteau*) como

$$Q = n \sum_{k=1}^K \hat{\rho}_k^2$$

donde K es el máximo número de retardos y n es el número de observaciones usadas para el cálculo de las verosimilitudes (nótese que n se reduce en 1 por cada diferencia). Si el modelo ajustado es apropiado, este estadístico se distribuye aproximadamente χ_{K-r}^2 donde r es el número de parámetros ajustados por el modelo.

Sin embargo, evidencias empíricas señalan que los valores de Q tienden a ser algo más pequeños que lo esperados bajo la distribución χ^2 . Por esta razón Ljung y Box proponen la forma modificada del estadístico:

$$\tilde{Q} = n(n+2) \sum_{k=1}^K (n-K)^{-1} \hat{\rho}_k^2$$

1.4.1. Identificación y Comportamiento de la ACF PACF

En lo que sigue, se establece una manera de identificar el tipo de modelo ARMA haciendo uso de la función de autocorrelación y autocorrelación parcial, las cuales son una herramienta útil en la detección del tipo de modelo que es adecuado en un estudio

	ACF	PACF
$AR(p)$	Muchos coeficientes no nulos que decrecen con la distancia y los retardos como suma de exponenciales y senos	p primeros coeficientes no nulos y el resto cero.
$MA(q)$	q primeros coeficientes no nulos y el resto cero.	Muchos coeficientes no nulos que decrecen con la distancia y los retardos como suma de exponenciales y senos
$ARMA(p, q)$	Decrecimiento lento hacia cero.	Decrecimiento lento hacia cero.

1.5. Sobre el Teorema de Bochner y la Representación Espectral.

En esta sección introduciremos conceptos asociados a la representación espectral de procesos aleatorios, en particular definiremos los conceptos necesarios para procesos débilmente estacionarios, estos definidos en la Sección 2.1 de este capítulo. Para procesos Gaussianos a valores reales la función media y covarianza definen todas las distribuciones finito-dimensionales y así la distribución completa del proceso. Sin embargo, la representación espectral requiere procesos a valores complejos, por lo que se necesita definir la estructura de correlación entre la parte real e imaginaria del proceso. Nos dedicaremos a definir la representación espectral para procesos reales puesto que estos son de interés en este trabajo, las distintas interpretaciones que pueden darse al espectro o representación espectral pueden hallarse con detalles en [4], [6]. Mostraremos las más importantes para dar una idea de lo que se puede hacer en el estudio de olas conociendo la distribución espectral.

1.5.1. Función Covarianza para Procesos Reales.

Para un proceso a valores reales, la función covarianza $r(\tau)$ esta determinada por todas las covarianzas entre $x(t_1), \dots, x(t_n)$

$$\Sigma(t_1, \dots, t_n) = \begin{pmatrix} r(0) & r(t_1 - t_2) & \dots & r(t_1 - t_n) \\ \vdots & \vdots & \ddots & \vdots \\ r(t_n - t_1) & r(t_n - t_2) & \dots & r(0) \end{pmatrix} \quad (1.32)$$

1.5.2. Funciones Definidas No-negativas.

Esta es una propiedad característica única de una función covarianza que es no-negativa en el siguiente sentido: Sean t_1, \dots, t_n un conjunto finito de puntos en el tiempo y se toman números complejos arbitrarios a_1, \dots, a_n . Entonces, por simplicidad asumiendo que $E(x_t) = 0$,

$$\sum_{j,k}^n a_j \overline{a_k} r(t_j - t_k) = E(a_j x(t_j) \overline{a_k x(t_k)}) \quad (1.33)$$

$$= E \left| \sum_{j=1}^n a_j x(t_j) \right|^2 \geq 0 \quad (1.34)$$

Teorema 1.1. *Cualquier función definida no-negativa, posiblemente compleja $r(\tau)$ es la función covarianza para un proceso Gaussiano estrictamente estacionario. Así, la clase de funciones covarianza es igual a la clase de funciones definidas no-negativas.*

La prueba de este teorema esta disponible en [6]

1.5.3. El Teorema de Bochner y la Distribución Espectral.

1.5.4. La Distribución Espectral.

Hemos visto que la función covarianza para procesos estacionarios es caracterizada por propiedades definidas no-negativas. También conocemos que la función covarianza es la transformada de Fourier de su *distribución espectral*, ahora formulamos esto en un teorema, la prueba puede verse también en [6]

Teorema 1.2 (Teorema de Bochner). *Una función continua $r(t)$ es definida no-negativa y así una función covarianza, si y solo si, existe una función real $F(\omega)$*

no-decreciente, continua por la derecha y acotada, tal que,

$$r(t) = \int_{-\infty}^{\infty} e^{i\omega t} dF(\omega) = \int_{-\infty}^{\infty} \cos(\omega t) dF(\omega)$$

La función $F(\omega)$ es la función de distribución espectral del proceso y tiene todas las propiedades de una distribución de probabilidad excepto que $F(+\infty) - F(-\infty) = r(0)$ no necesariamente es igual a 1. La función $F(\omega)$ es definida salvo por una constante aditiva y usualmente se toma $F(-\infty)$

La prueba este teorema se puede hallar en [6].

En lo que sigue, mostramos las consecuencias que traen en Teorema de Bochner para el estudio de olas marinas y sobre todo los múltiples usos que tiene el espectro en el estudio de olas marinas. Daremos un resumen de las características más importantes que se obtienen a través del espectro, como la altura significativa, modelos de banda ancha y estrecha, entre otros, además destacamos la importancia de los momentos espectrales para dar más fuerza a la existencia de la regularidad de la trayectoria, estudiada en la próxima sección. Para más detalles se pueden consultar en [5], [4].

1.5.5. Consecuencias del Teorema de Bochner

Si la función covarianza es integrable, entonces la formula anterior es invertible

$$F(\omega) = \frac{2}{\pi} \int_0^{\infty} \cos(\omega t) r(t) dt$$

Usando la representación espectral tenemos

$$r'(t) = \int_{-\infty}^{\infty} -\omega \sin(\omega t) dF(\omega)$$

$$r''(t) = \int_{-\infty}^{\infty} -\omega^2 \cos(\omega t) dF(\omega)$$

en particular

$$r''(0) = - \int_{-\infty}^{\infty} \omega^2 dF(\omega)$$

Esta integral se conoce como el *segundo momento espectral*:

$$m_2 = \int_{-\infty}^{\infty} \omega^2 dF(\omega) = -r''(0)$$

Si r no es diferenciable dos veces, entonces

$$m_2 = \int_{-\infty}^{\infty} \omega^2 dF(\omega) = \infty$$

La existencia del segundo momento espectral está asociado a la regularidad de la trayectoria que sigue el proceso. Cuando $m_2 < \infty$ la función covarianza tiene el siguiente desarrollo cerca del origen

$$r(h) = \sigma^2 - \frac{m_2 h^2}{2} + \mathcal{O}(h^4)$$

Es posible mostrar que $m_2 < \infty$ sí y sólo sí el proceso aleatorio M es diferenciable en media cuadrática, es decir, si existe un proceso $M'(t)$ tal que

$$\frac{M(t+h) - M(t)}{h} \rightarrow M'(t) \text{ en } L^2$$

y entonces

$$E(M'(t)) = 0 \text{ y } Var(M'(t)) = -r''(0) = m_2$$

$M'(t)$ es Gaussiano, además independiente de $M(t)$ y tiene función covarianza

$$Cov(M'(t), M'(t+h)) = -r''(h)$$

La densidad conjunta de M y M' es

$$p(u, z) = \frac{1}{2\pi\sigma\sqrt{m_2}} \exp\left\{-\frac{1}{2}\left(\frac{u^2}{\sigma^2} + \frac{z^2}{m_2}\right)\right\}$$

En el estudio de olas marinas el espectro juega un rol fundamental y se interpreta como la distribución de la energía, observemos que

$$r(0) = Var(X(t)) = \int_{-\infty}^{\infty} f(\omega) d\omega$$

de modo que la varianza del proceso, representa la energía total. El espectro es simétrico y sólo se considera la parte positiva y se renormaliza multiplicandolo por 2.

Para estudiar los espectros, lo usual es utilizar la densidades espectrales *típicas* que dependen de algunos parámetros que deben ser estimados en cada caso. Esto permite comparar resultados obtenidos bajo diferentes hipótesis y una ventaja es que los

calculos involucrados se simplifican. Los dos tipos de espectros más comunes son el de Pierson-Mostkowitz

$$f(\omega) = \frac{g^2 \alpha}{\omega^5} e^{-\beta(\frac{\omega_p}{\omega})^4} \quad \text{para } 0 \leq \omega \leq \omega_{max}$$

con frecuencia modal ω_p . Y el espectro JONSWAP (JOint North-Sea WAve Project) definido por

$$f(\omega) = \frac{g^2 \alpha}{\omega^5} e^{-\beta(\frac{\omega_p}{\omega})^4} \times \gamma^{\exp[-(1-\omega/\omega_p)^2/2\sigma_\omega^2]}$$

donde $\sigma_\omega = 0,07$, para $\omega < \omega_p$, $\sigma_\omega = 0,09$ para $\omega > \omega_p$.

En muchas ocasiones estas formulas no describen con mucha precisión a los espectros reales, por lo que muchas veces es conveniente combinar estas para poder lograr espectros más realistas. Esto es lo que ocurre cuando el viento local influye en las olas, que es la principal fuente de irregularidad, este fenomeno se conoce como *swell*, en este caso el espectro es la mezcla de los espectros individuales.

1.6. Características de las Olas.

1.6.1. Altura Significativa.

La medida más importante de la severidad del mar es la *altura significativa*. Esta medida intenta hallar la altura de las olas más altas que pueden hallar en un periodo de tiempo razonable. El término significativa significa que esta altura es suficientemente alta como para surtir efectos sobre barcos o estructuras colocadas en el océano.

Definición 1.18. La altura significativa de un estado del mar se define como

$$H_{m_0} = 4\sqrt{\text{Var}(M(t))}$$

Cabe destacar que han habido muchos intentos en encontrar una definición apropiada para la altura significativa del mar, comenzando con observaciones visuales desde barcos hasta medidas mas precisas realizadas con boyas marinas. La definición anterior, es la más aceptada.

1.6.2. Características Basadas en Cruces Nivel Medio.

Esta característica también tiene muchas definiciones. La definición más usual esta basada en los cruces del cero. Consideremos $M(t)$ como el proceso que modela

las olas en un punto del espacio; $M(t)$ representa la altura sobre el nivel medio del mar. Supongamos que $M(t)$ cruza este nivel medio hacia abajo en los instantes t_1, t_2, \dots, t_n . El tiempo entre dos cruces sucesivos del nivel medio hacia abajo definen el período de la ola, usaremos la notación $T_{d,k}$:

$$T_{d,k} = t_{k+1} - t_k$$

La distancia vertical entre el máximo y el mínimo valor del proceso en este intervalo se define como la altura de la ola. Usamos la notación $H_{d,k}$.

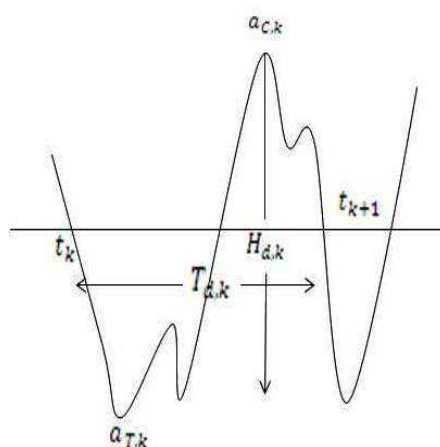


Figura 1.2: Altura media del mar

Una cresta a_C es el máximo valor de M para t en un intervalo de tiempo entre dos cruces sucesivos hacia abajo del nivel medio $t_k < t < t_{k+1}$. De manera similar un seno a_S es el (valor absoluto del) valor mínimo en el mismo intervalo de tiempo. Para distinguir diferentes crestas y senos en intervalos sucesivos se usa la notación $(a_{C,k}, a_{S,k})$.

1.6.3. Características Basadas Extremos Sucesivos.

Supongamos que la ola tiene una sucesión de mínimos locales $(M_{m,k})$ y máximos locales consecutivos $(M_{M,k})$. Una ola *min-max* es el par $(M_{m,k}, M_{M,k})$ de valores mínimos y máximos consecutivos. La altura de la ola min-max es la diferencia entre estos valores mientras que el período min-max es la diferencia de tiempos correspondientes. De manera similar se definen las olas max-max y min-min.

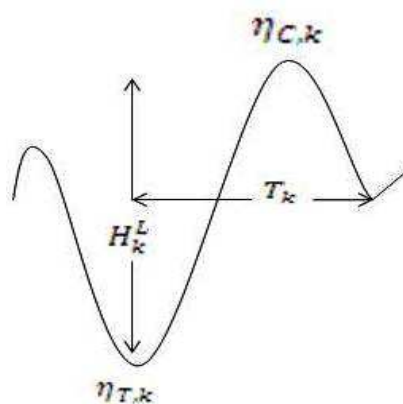


Figura 1.3: Extremos Sucesivos.

1.6.4. Momento Espectral de Orden n

El momento espectral de orden n se define como

$$m_n = \int_0^{\infty} \omega^n f(\omega) d\omega.$$

En los párrafos anteriores hemos dicho que la existencia del segundo momento espectral está relacionado con la regularidad de la trayectoria para procesos Gaussianos. En general, la existencia de los momentos de orden superior está asociada a una mayor regularidad de la trayectoria.

Los momentos espectrales no son los únicos parámetros asociados a la densidad espectral de interés en el estudio del mar. En términos de los momentos espectrales la altura significativa es

$$H_{m_0} = 4\sqrt{\text{Var}(M(t))} = 4\sqrt{m_0}.$$

A partir del espectro, vemos que la frecuencia media está dada por

$$\frac{m_1}{m_0}$$

Si el espectro está concentrado alrededor de una frecuencia dominante, la frecuencia media da el período medio. Otra característica de interés son

$$\omega_{i,j} = \sqrt{\frac{m_i}{m_j}}$$

y el ancho espectral que se define como

$$\epsilon = \sqrt{1 - \frac{m_2^2}{m_0 m_4}} = \sqrt{1 - \left(\frac{\omega_{2,0}}{\omega_{4,2}}\right)^2}$$

Recordemos que σ^2 es el área bajo la densidad espectral, de modo que si conocemos la densidad espectral podemos hacer predicción sobre la altura de la ola.

1.6.5. Modelo de Banda Estrecha.

Supongamos que el proceso que representa la altura de la ola es aproximadamente sunisoidal:

$$X(t) = a(t) \cos(\omega_0 t + \varepsilon(t))$$

donde $a(t)$ es la amplitud, $\varepsilon(t)$ es la fase y ω_0 es la frecuencia donde se concentra el espectro. Tanto $a(t)$ como $\varepsilon(t)$ varían aleatoriamente con el tiempo t pero lentamente; la velocidad es mucho menor que ω_0 . Así

$$X'(t) = a'(t) \cos(\omega_0 t + \varepsilon(t)) - a(t)(\omega_0 + \varepsilon'(t)) \sin(\omega_0 t + \varepsilon(t)) \approx -a(t)\omega_0 \sin(\omega_0 t + \varepsilon(t))$$

Por lo tanto, la distribución conjunta de X y X'/ω_0 es

$$f\left(x, \frac{x'}{\omega_0}\right) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2}\left(x^2 + \frac{x'}{\omega_0}\right)^2\right\}$$

A partir de aquí, podemos obtener la densidad conjunta de a y ε

$$f(a, \varepsilon) = \frac{1}{2\pi\sigma} a \varepsilon^{-a^2/2\sigma^2} \quad 0 \leq a \leq \infty, \quad 0 \leq \varepsilon \leq 2\pi$$

1.6.6. Modelo de Banda Ancha.

Para los proceso de banda estrecha, la trayectoria tiene un sólo máximo cada medio ciclo el cual determina la amplitud. Si el proceso no es de banda estrecha puede haber varios extremos durante el medio ciclo determinado por cruces sucesivos de nivel cero. Si eliminamos la restricción de que el proceso es de banda estrecha, podemos tener ahora máximos positivos y negativos.

Supongamos que el proceso tiene trayectorias continuamente diferenciables. El proceso X tiene un máximo en t_0 sí y sólo si, X' tiene un cruce hacia abajo en t_0 , una cantidad de resultados pueden obtenerse a partir de los resultados para cruces hacia abajo.

Para asegurar que X tiene sólo un número finito de máximos en un intervalo finito de tiempo basta pedir que

$$m_4 = \int_{-\infty}^{\infty} x^4 dF(x) = r^4(0) < \infty$$

Si $m_4 < \infty$ tiene segunda derivada (en media cuadrática) y $X(t)$, $X'(t)$, $X''(t)$ son conjuntamente Gaussianos y con matriz de covarianza

$$\begin{pmatrix} m_0 & 0 & -m_2 \\ 0 & m_2 & 0 \\ -m_2 & 0 & m_4 \end{pmatrix}$$

Si $m_4 < \infty$ tenemos

$$\text{Cov}(X'(t), X'(t+h)) = -r''(h) = m_2 - \frac{m_4}{2}h^2 + o(h^2) \quad (h \rightarrow 0)$$

y si normalizamos

$$\text{Cov}\left(\frac{X'(t)}{\sqrt{m_2}}, \frac{X'(t+h)}{\sqrt{m_2}}\right) = -r''(h) = 1 - \frac{m_4}{2m_2}h^2 + o(h^2) \quad (h \rightarrow \infty)$$

Llamamos $N'(t)$ al número de máximos locales en $[0, \infty)$ para $X(t)$, que es igual al número de cruces hacia abajo del nivel cero de $X'(t)$ y a partir de la *formula de Rice* [5] obtenemos que

$$E(N'(t)) = \frac{T}{2\pi} \left(\frac{m_4}{m_2}\right)^2$$

Sea ahora $N'_u(T)$ el número de máximos locales de $X(t)$, $0 < t < T$ con altura mayor que u .

Proposición 1.1. *Sea $X(t)$ con trayectorias continuamente diferenciables, con segunda derivada $X''(t)$ en media cuadrática que satisface $\text{Var}(X''(t)) = m_4 < \infty$ y tal que $X(t)$, $X'(t)$ y $X''(t)$ tienen distribución no singular, entonces*

$$E(N'_u(T)) = T \int_u^{\infty} \int_{-\infty}^{\infty} |z| p(x, 0, z) dz dx$$

donde $p(x, y, z)$ es la densidad conjunta de $X(t)$, $X'(t)$ y $X''(t)$. Además si X esta nomalizada

$$E(N'_u(T)) = \frac{T}{2\pi} \left[\left(\frac{m_4}{m_2}\right)^{1/2} \left(1 - \phi\left(u\left(\frac{m_4}{D}\right)^{1/2}\right) + (2\pi)^{1/2} \phi(u) \Phi\left(\frac{um_2}{D^{1/2}}\right)\right) \right]$$

Llamamos $s_1, s_2, \dots, s_{N'}$ los instantes en los cuales ocurren los $N' = N'(T)$ máximos locales en $[0, \infty]$ y sean s'_i los instantes de los mínimos, indexados de modo que $s_i < s'_i < s_{i+1}$ los valores $X(s_1), X(s'_1), X(s_2), \dots$ representan las amplitudes aparentes $X(s_1) - X(s'_i)$ en las alturas de la ola aparentes.

Observemos que

$$\max(X(s_1), X(s_2), \dots, X(s_{N'})) = \sup\{X(t), s_1 < t < s_{N'}\}$$

que es igual a $M(t) = \sup\{X(t), t < 0 < T\}$ si $M(t)$ no se alcanza en 0 ni en T .

Bajo las hipótesis anteriores, si $\text{Var}(X(t)) = m_0$, haciendo una normalización en la ecuación anterior y usando la siguiente notación

$$\begin{aligned} \nu &= \frac{1}{2\pi} \left(\frac{m_2}{m_0} \right)^{1/2} & \nu' &= \frac{1}{2\pi} \left(\frac{m_4}{m_2} \right)^{1/2} \\ \varepsilon &= \left(1 - \frac{\nu}{\nu'} \right)^{1/2} = \left(1 - \frac{m_2^2}{m_0 m_4} \right)^{1/2} \end{aligned}$$

obtenemos

$$E(N'_u(T)) = T \left[\nu' \left(1 - \Phi \left(\frac{u}{\varepsilon \sqrt{m_0}} \right) \right) + \nu \exp \left(\frac{u^2}{2m_0} \right) \Phi \left(\frac{u \nu m_2}{\varepsilon \nu' \sqrt{m_0}} \right) \right]$$

El número esperado de máximos es $T\nu'$ y podemos considerar el cociente $E(N'_u(T))/T\nu'$ como una aproximación de la probabilidad de que un máximo local esté por encima del nivel $u > 0$. Definimos la función

$$F_{max}(u) = 1 - \frac{E(N'_u(1))}{\nu'}$$

que es una función de distribución. La densidad correspondiente es

$$f_{max}(u) = \frac{\varepsilon}{\sqrt{m_0}} \phi \left(\frac{u}{\varepsilon \sqrt{m_0}} \right) + \left(\frac{1 - \varepsilon^2}{m_0} \right)^{1/2} \frac{u}{\sqrt{m_0}} \exp \left(\frac{-u^2}{2m_0} \right) \Phi \left(\frac{u(1 - \varepsilon^2)^{1/2}}{\varepsilon \sqrt{m_0}} \right)$$

Si el proceso es ergódico

$$\frac{N'_u(T)}{N'(T)} \rightarrow 1 - F_{max}(u)$$

con probabilidad 1 de modo que $F_{max}(u)$ es el límite de la distribución empírica de las amplitudes aparentes. La forma de la distribución depende del ancho espectral $\varepsilon = (1 - \frac{m_2}{m_0 m_4})^{1/2}$. Valores pequeños de ε corresponden a espectros de banda estrecha. En particular, si $\varepsilon = 0$ obtenemos la distribución de Rayleigh

$$\frac{u}{m_0} \exp \left\{ \frac{-u^2}{2m_0} \right\}$$

mientras que para $\varepsilon = 1$ obtenemos la distribución Gaussiana truncada en 0

$$\sqrt{\frac{2}{m_0}} \exp\left\{-\frac{u^2}{2m_0}\right\}$$

1.7. Sobre las Hipótesis Necesarias para Obtener Modelos Adecuados para el Análisis Espectral de Olas.

En el análisis de olas desde el punto de vista estocástico, las olas en aguas profundas se modelan como procesos aleatorios gaussianos, de modo que la distribución de probabilidad de la altura respecto al valor medio del mar es una distribución normal. Cuando se consideran olas en aguas poco profundas los efectos no lineales cobran mayor importancia y se requieren modelos no gaussianos.

En este sentido, cuando se modelan olas en aguas profundas, teniendo en cuenta que se modelan como procesos gaussianos, es necesario obtener algunas propiedades las cuales garantizan la obtención de modelos manejables para el estudio. La primera de estas propiedades es la estacionaridad, la cual tiene que ver con la permanencia en el tiempo de las condiciones del mar, que para procesos aleatorios en general el término se refiere no a la ola sino a sus propiedades estadísticas. Otra propiedad necesaria es la continuidad de la trayectoria que sigue la ola, la cual depende del tiempo y de un parámetro aleatorio sea continua. También se necesita la propiedad de ergodicidad, la cual es quien permite sustituir valores esperados (teóricos) por promedios temporales (empíricos), para más detalles véase [5]. Por otro lado, las aguas en poco profundas, donde se requieren modelos que no son gaussianos, es necesario el teorema del límite central para el desarrollo de las propiedades que se necesitan para tratar con este tipo de modelos [4]. En esta sección mostramos la teoría necesaria para aguas profundas, la cual es de interés en este trabajo.

Consideremos un proceso aleatorio M , que depende de la posición en el espacio x y el tiempo t , además de un parámetro aleatorio ω que pertenece al espacio de probabilidad $(\Omega, \mathcal{F}, \mathcal{P})$, $M(t, x, \omega)$. Por costumbre se omite el parámetro ω . Aquí usaremos $M(t)$ cual denota la altura del nivel del mar en un punto fijo, como función del tiempo.

La *estacionaridad* del proceso, se define como en la Definición 1.5, la cual requiere que la media del proceso no dependa del valor de t , además de que la función covarianza dependa sólo de la diferencia en valor absoluto entre s y t . Es sabido que

las condiciones del mar cambian con el tiempo y así los parámetros de la distribución estadística de la altura de la ola, la longitud de la misma, periodos, entre otros, por lo que la hipótesis de estacionaridad se cumple *parcialmente*, es decir la distribución de las olas cambian con el tiempo. Tiene sentido entonces hablar de altura media, longitud media, periodo medio, y otras características estadísticas en un punto del espacio.

Esta hipótesis también implica que la distribución del proceso $M(t+h)$ es el mismo para cualquier valor, en particular es siempre igual a la de $M(0)$, aunado a esto, para cualquier n y cualesquiera tiempos t_1, t_2, \dots, t_n la distribución del vector $(M(t_1+h), M(t_2+h), \dots, M(t_n+h))$ es independiente del valor h . Otra hipótesis que se supone es que la altura media del mar es 0 y se medirán las variaciones respecto a él, esta hipótesis quiere decir que el proceso se considera *centrado*, es decir $E[M(t)] = 0$

Cuando nos referimos a la *continuidad* de la trayectoria que sigue la ola, nos referimos a considerar M como una función sobre el espacio producto $[0, \infty] \times \Omega$, con la propiedad que para cualquier t fijo en $[0, \infty]$, $M(t, \cdot)$ es medible. Si fijamos $\omega \in \Omega$ obtenemos una función

$$M(\cdot, \omega) : [0, \infty] \rightarrow \mathbb{R}$$

que se conoce como una trayectoria del proceso. Pedimos que para casi todo $\omega \in \Omega$, sea decir con probabilidad 1, esta función sea continua.

La hipótesis que define al proceso como *ergódico*, es aquella que permite sustituir los valores esperados (teóricos) por promedios temporales (empíricos), esto es

$$\begin{aligned} E(M(t)) &\equiv \int_{\Omega} M(t, \omega) dP(\omega) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^{\infty} M(u) du \\ \text{Cov}(M(t), M(t+h)) &\equiv \int_{\Omega} M(t, \omega) M(t+h, \omega) dP(\omega) \\ &= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t M(u) M(u+h) du \end{aligned}$$

Para la hipótesis de *Gaussianidad*, la distribución de la altura de la ola en un punto dado y en un instante de tiempo t tiene la función de distribución:

$$P(M(t) \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} dx$$

Adicionalmente para cualquier valor de n y cualesquiera sean los instantes de tiempo t_1, t_2, \dots, t_n la distribución del vector $(M(t_1+h), M(t_2+h), \dots, M(t_n+h))$

tiene densidad Gaussiana

$$f_{t_1, t_2, \dots, t_n}(u_1, \dots, u_n) = \frac{1}{(2\pi)^{n/2}(\Sigma)^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{u}'\Sigma^{-1}\mathbf{u}\right\}$$

donde $\mathbf{u} = (u_1, u_2, \dots, u_n)$, $\Sigma = Cov(M(t_i), M(t_j))$.

CAPÍTULO 2

ANÁLISIS ESPECTRAL DE LOS DATOS Y MODELOS ASOCIADOS

2.1. Descripción de los Datos

En este trabajo, se realiza un análisis espectral de 244 periodos de una tormenta en el Mar del Norte ocurrido en 1,999. Estos datos fueron registrados a una frecuencia de $5Hz$. Para el estudio de estos datos, se usó el software R, una herramienta muy útil en el análisis estadístico, además que proporciona información más completa y detallada sobre los elementos necesarios para el análisis espectral de datos.

Los espectros obtenidos a partir de los datos de la tormenta, pueden apreciarse en la Figura 2.1, donde se pueden observar importantes fluctuaciones por cada periodo, el área total bajo la curva representa la energía en cada uno de los intervalos de tiempo, la gráfica está hecha en base al espectro en función de la frecuencia, de donde luego se obtendrá importante información para su posterior estudio. El espectro proporciona información diferente de las propiedades estadísticas de la serie de tiempo, ya que da la distribución de la varianza de la serie con la frecuencia. Como muestran Jenkins y Watts (1968) [2] el espectro puede ser graficado en función de la frecuencia en el rango comprendido entre 0 y $1/2$. Por tanto, cuando se trabaja con el espectro, se dice que se está trabajando en el *dominio de frecuencia*, mientras que si se investiga la ACF, nos referimos al estudio en el *dominio de tiempo*. En nuestro caso, el comportamiento de los espectros frente al a frecuencia se muestra a continuación en la Figura 2.1.

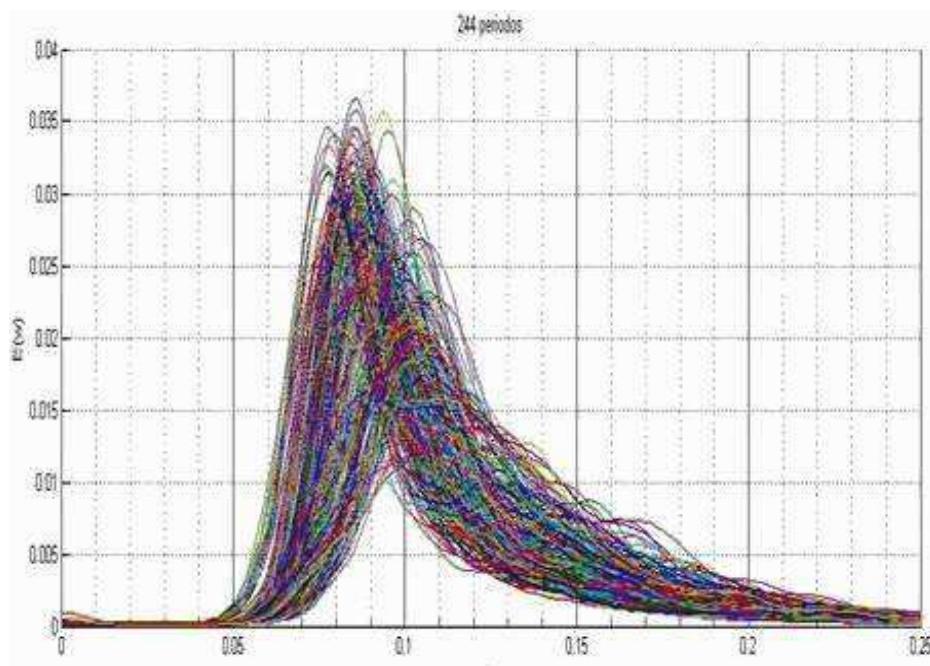


Figura 2.1: Espectros de los 244 periodos en estudio.

La técnica utilizada para el ajustar modelos estocásticos para la evolución espectral fue la siguiente: Los registros de altura del mar están divididos en intervalos de 20 minutos y para cada uno de ellos se calculo la densidad espectral usando el software WAFO. Inicialmente se determinó el *rango útil* de frecuencias común a todos los espectros y para esto se seleccionaron empíricamente valores a los extremos del rango de modo que no se observaron características importantes fuera de este rango y que la energía total presente fuera del rango fuese pequeña.

El rango obtenido luego de este procedimiento se dividió en intervalos de igual tamaño y se calculo la energía presente en cada intervalo por cada uno de los espectros, de modo que se tiene una sucesión de valores de la energía presentes en cada banda de frecuencias a lo largo de la tormenta. A cada una de estas series se le ajusto un modelo ARMA cuyo orden se determinó usando las herramientas habituales como las gráficas de la ACF y PACF, el p -valor de los coeficientes y el criterio de información de Akaike.

Para verificar el ajuste se hizo un estudio de los residuales de los modelos obtenidos usando la gráfica de cuantiles para comparar el ajuste con una distribución normal. En caso afirmativo los residuales deben colocarse aproximadamente en una línea recta. Para algunos modelos se encontró que este no era el caso y se recurrió a la

transformación de Box-Cox sobre los datos con el fin de lograr que la hipótesis de Gaussianidad se cumpla, al menos aproximadamente.

También se realizó el ajuste de modelos vectoriales autorregresivos. En estos modelos se incluye la interacción entre las distintas series agrupadas como componente de un vector que evoluciona en el tiempo a lo largo de la duración de la tormenta, estos modelos serán detallados en el Capítulo 3.

2.2. Estudio del Comportamiento del Espectro

Como mencionamos anteriormente, el primer paso fue considerar la colección de los 244 espectros con el fin de determinar el *rango útil* de las frecuencias. Toda estimación incluye un error o ruido y en caso de las densidades espectrales lo que se observa en muy bajas o muy altas frecuencias es esencialmente ruido. El objetivo es reducir el rango de frecuencias dejando fuera aquellas frecuencias en las que se observa principalmente ruido. Luego de examinar las gráficas y de calcular el porcentaje de la energía que queda fuera del intervalo decidimos tomar el intervalo $[0,04, 0,25]$, dentro del cual se encuentra aproximadamente el 97 % de la energía total para todos los períodos.

2.2.1. Modelos Asociados al Estudio de la Energía

En esta sección, nos proponemos discretizar el intervalo $[0,04, 0,25]$ en subintervalos apropiados, de tal manera que estos capturen de la mejor manera posible la variabilidad de la energía a lo largo de la tormenta, pero teniendo en cuenta que un número elevado de intervalos va a producir dificultades en la modelación y su posible interpretación. Luego de observaciones gráficas, se decidió dividir el intervalo $[0,04, 0,25]$ en 12 subintervalos, la gráfica a continuación (Figura 2.2) muestra la partición para la densidad espectral, la cual tendrá 12 series de tiempo que se obtienen a partir de la división del rango de frecuencias.

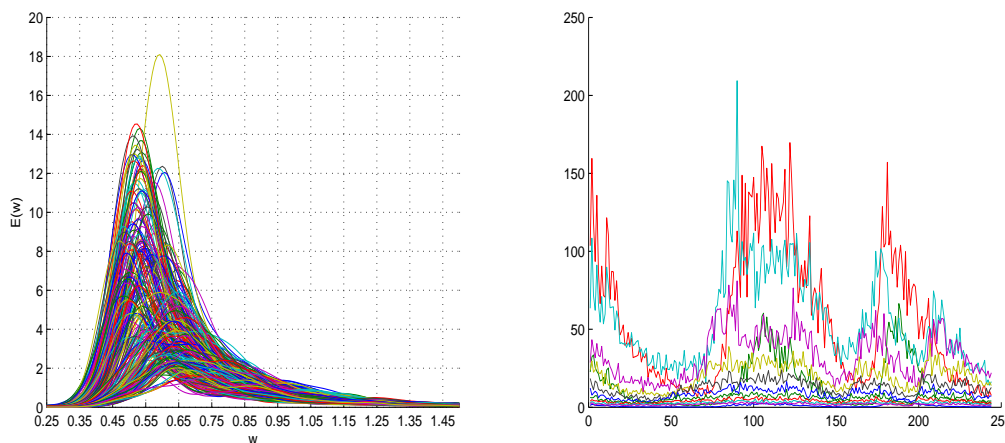


Figura 2.2: Izquierda: Evolución de la energía para cada uno de los subintervalos, derecha: Gráfico del intervalo $[0,04,0,25]$, la cual representará el 97% de los datos a estudiar

Por cada partición haremos el estudio del modelo ARMA apropiado que corresponde a cada uno de estos subintervalos de frecuencia. Esto último, se realizará estudiando su función de autocovarianza (ACF) y su función de autocovarianza parcial (PACF), parte importante en este punto también será el estudio del comportamiento de los residuales del modelo respecto a la gráfica de cuantiles para verificar el ajuste con una distribución normal. Otro análisis determinante en la escogencia y bondad de los modelos en estudio, el análisis de los p -valores mediante el *estadístico de Ljun Box* (ver por ejemplo [1] y [2]) de acuerdo al cual podemos decidir si el ajuste del modelo resulta apropiado.

2.2.2. Sobre la Construcción de los Modelos

De ahora en adelante, el análisis se centrará en la construcción y ajuste de modelos a partir de las 12 divisiones del intervalo de frecuencias que se ha considerado en este trabajo. Esto se hará de acuerdo a las características que poseen la ACF y la PACF, y a partir de allí se obtendrá respuesta sobre el tipo de modelos y orden de los mismos que se usarán para el presente análisis.

A continuación se presenta un análisis sobre los elementos que se han mencionado anteriormente para la escogencia de los modelos adecuados por cada partición, por ejemplo, para el primer subintervalo, gráficamente se obtuvo lo siguiente

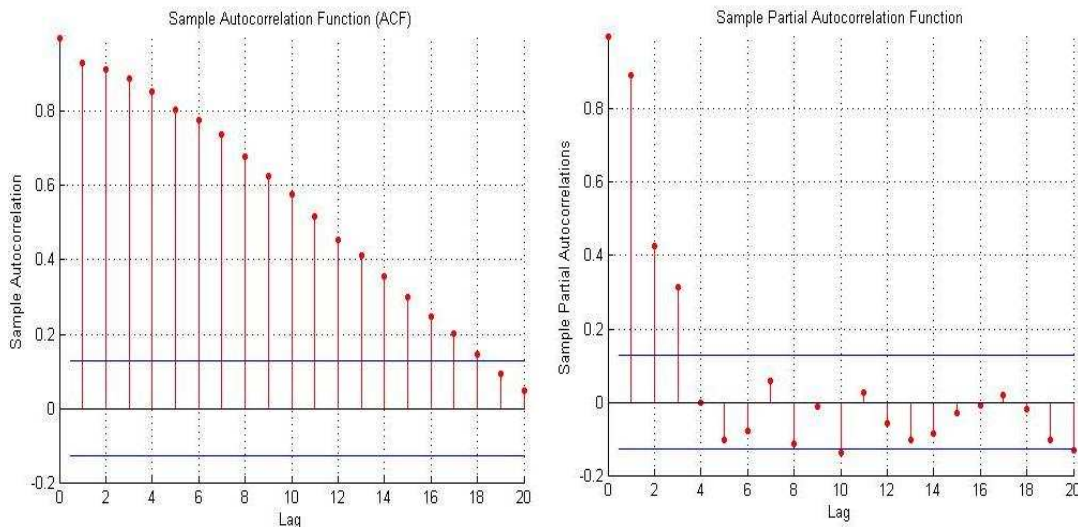


Figura 2.3: Información sobre el modelo asociado al primer suintervalo. Izquierda: Función de autocorrelación, derecha: Función de autocorrelación parcial

Las PACF y ACF para la serie de tiempo correspondiente primer subintervalo se ve en la Figura 2.3, donde se evidencia un decaimiento de la ACF hasta el número 18. Por otro lado, la ACP muestra tener los primeros 4 valores fuera del intervalo de confianza, por lo que estas características sugieren un modelo autorregresivo de orden 4. Evidencias de este tipo se analizaron para los demás subintervalos y se obtuvieron órdenes entre 4 y 6. Pero este análisis no es suficiente, ya que no hemos considerado todos los componentes estadísticos que posee el modelo. Haciendo un análisis más exhaustivo y experimentando con los órdenes para el modelo ARMA correspondiente a cada uno de los subintervalos, se encontró que el orden para cada uno de estos modelos puede reducirse, viendo los coeficientes del modelo que algunos pueden ser despreciados por su proximidad al cero y además observando el menor valor para el AIC (Akaike Information Criteria) y el estudio de los p -valores de cada modelo encontrado.

2.3. Características de los Modelos

Para dar respuesta a un estudio de olas marinas usando modelos ARMA, es necesario tomar en cuenta algunas consideraciones sobre los parámetros que están involucrados en el análisis. Usualmente, la metodología usada es estimar el modelo,

calcular su ACF y PACF de manera que se pueda determinar el orden y tipo de modelo que estaremos estudiando de acuerdo a la tabla resumida en la Sección 1.4.1. Seguidamente observar el comportamiento de las gráficas de diagnóstico, es decir, el comportamiento de los p -valores mediante el estadístico de Ljung-Box, el ACF para los residuales y los residuales estandarizados, esto con el fin de determinar y verificar el orden del modelo que estamos tratando, en este punto es conveniente también calcular el coeficiente AIC (Akaike Information Criteria) y compararlo con otros modelos que se consideren convenientes para constatar que el modelo en estudio es el que mejor ajusta a los datos, el modelo que posea el menor AIC será el más conveniente en el análisis [1]. Es útil también graficar los cuantiles del modelo para verificar la normalidad en los datos, si estos cuantiles no siguen un comportamiento aproximadamente con una línea recta entonces es conveniente usar la transformación de Box-Cox explicada en la Sección 1.2.4 para tratar de mejorar la aproximación lineal que se desea para que así los cuantiles logren la normalidad deseada. Por último, realizar los pronósticos al modelo y hacer la comparación entre este y los datos de donde proviene para constatar que el modelo es capaz de reproducir los datos en estudio.

A continuación, se da la justificación de la elección del orden y tipo de los modelos, además se muestra justificación de la escogencia de este orden mostrando la forma estadística de algunas de las componentes del modelo, por ejemplo, el valor AIC asociado al modelo, las características de sus residuales, su contraste con los cuantiles normales, aunado a esto, como ya se planteó, las funciones de autocorrelación y autocorrelación parcial además del estadístico Ljung Box. En este punto, se han evidenciado algunas particularidades en los residuos, como es el cambio brusco de su comportamiento, esto dado que hay cambios muy marcados positiva y negativamente en los residuales y casi de manera instantánea los siguientes residuos se aproximan a cero, este fenómeno se conoce como *heterocedasticidad*. Este tipo de fenómenos aporta a la no normalidad de los datos porque hay indicios de colas pesadas, lo que impide que los residuos se ajusten a la línea recta deseada, más bien, lo que se vuelve es una función cóncava hacia arriba por un lado y luego cambia a cóncava hacia abajo por otro lado o viceversa. Aunado a esto, la evidencia de que algunos valores de la autocorrelación muestral fallan dentro del intervalo de confianza que considera

el estadístico de Ljun Box, en este punto ya el modelo se considera poco viable y es necesario considerar otras alternativas para detectar la bondad del modelo y no rechazarlo.

2.4. Comparación de los Modelos Obtenidos y Selección de los Modelos con Mejor Ajuste.

De acuerdo a las características de los modelos obtenidos; en esta sección se compararán las componentes de los estos modelos y de acuerdo a criterios ya mencionados y establecidos se escogerá el que mejores características estadísticas posee. Tales características fueron observadas haciendo cambios en el orden de los modelos descritos en la Sección 2.2.2. En este sentido, para lograr el orden adecuado a los modelos se observan la ACF y el comportamiento de los p -valores usando el estadístico de Ljun Box, además del comportamiento de los residuales, también es necesario calcular el valor AIC para estos modelos. El modelo con mejores características estadísticas en cuanto al comportamiento de sus componentes y además con el menor valor para AIC, será el modelo que mejor ajuste a los datos en estudio.

A continuación se reflejan gráficamente las componentes de los modelos más adecuados que se han estimado, haciendo cambios ligeros en el orden del modelo ARMA, como consecuencia se observan notablemente como los p -valores para uno de los modelos entran dentro de la banda de confianza para el estadístico de Ljun Box, además los residuales reflejan aleatoriedad, estos son los criterios entre otros, que se han considerado para la elección de los modelos a estudiar. Como una muestra comparamos el modelo AR de orden 2 y de orden 1 para el primer subintervalo, de tal forma que se puedan apreciar las diferencias planteadas sobre el comportamiento de sus componentes estadísticas, explícitamente tenemos que los modelos son

$$AR(2): x_t = 0,006 + 0,71x_{t-1} + 0,25x_{t-2} + \omega_t$$

$$AR(1): x_t = 0,006 + 0,94x_{t-1} + \omega_t$$

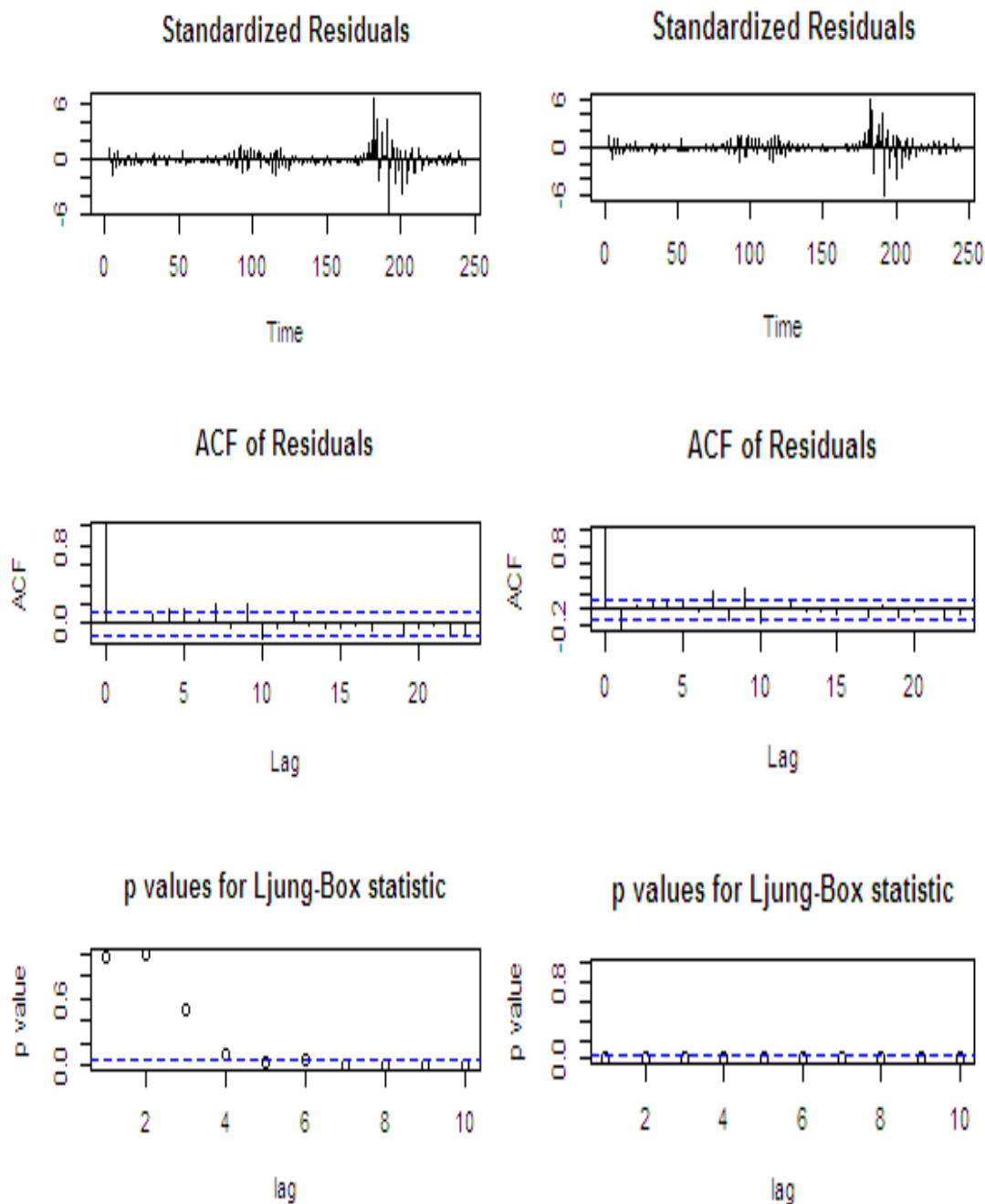


Figura 2.4: Componentes para el primer subintervalo. Izquierda: ARMA(2,0), AIC=-2308.13, derecha:ARMA(1,0), AIC=-2294.97

En la Figura 2.4, se pueden apreciar los modelos más apropiados que hemos hallado para el primer subintervalo, en este punto se debe decidir sobre cuál de estos modelos es el que mejor ajusta los datos, para ello notamos que en la gráfica de

la izquierda se observa que los residuos normalizados muestran aleatoriedad en su comportamiento, esto da indicios de que no existe una tendencia específica en el comportamiento de los datos, pero al ver el comportamiento de los p -valores, se observa que algunos de los valores permanecen fuera del intervalo de confianza considerado apropiado para estos, caso contrario ocurre con la figura de la derecha, en la cual los residuales normalizados muestran aún más indicios de aleatoriedad y la prueba para los p -valores indica que estos todos permanecen dentro del intervalo de confianza para el estadístico que se está usando, por lo que el modelo que se considerará para el estudio de este subintervalo es un modelo $AR(1)$.

De esta misma forma y usando argumentos similares se determinó el orden para los siguiente 11 subintervalo, para los cuales también se obtuvo un $AR(1)$. Mayores detalles de estos modelos se muestran sumariados en Apendice A

2.5. Características de los Residuos y Gráfica de Cuantiles para los modelos.

Componentes importantes en el análisis de datos para conocer y decidir sobre modelos ARMA en general, es el estudio exhaustivo de los residuales de estos modelos, ya que dan un panorama claro sobre el comportamiento normal o no de los estos, esto último, clave en el presente estudio, ya que en definitiva estas gráficas hablan sobre el buen ajuste de los modelos que se están considerando. En este sentido, es conveniente mencionar que en algunas situaciones se presentan particularidades en el comportamiento de los residuales y una de estas, por lo menos la más común, es que se evidencia *heterocedasticidad*, en presencia de este fenómeno la varianza cambia de manera brusca y las oscilaciones se hacen pronunciadas, por lo que al observar la gráfica de cuantiles la distribución de estos es Gaussiana, es decir, los cuantiles deben aproximarse a una línea recta; en caso contrario, si se observa algún comportamiento cóncavo o convexo se hace uso de la *transformación de Box-Cox*.

A continuación, mostraremos el estudio para los subintervalos que no requieren de esta transformación, dentro de estos están el tercer, cuarto, noveno, decimo, decimo primer y decimo segundo subintervalo en el orden, en particular, haremos el análisis para el tercer subintervalo de forma precisa y amplia, para luego mostrar a los que

sí necesitan de la transformación, los cuales son el primero, segundo, quinto, sexto, séptimo y octavo subintervalo en el orden que están en la partición, en este caso mostraremos el análisis para el primer subintervalo que también pretendemos ampliar su estudio para dejar un panorama claro de lo que se pretende hacer con los restantes. De esta manera observar y entender la técnica empleada para el análisis. Mostraremos también las gráficas de cuantiles, quienes nos proporcionarán la información necesaria para aplicar o no la transformación de Box-Cox. La transformación servirá para aproximar la línea recta para obtener la normalidad de los residuos del modelo en el caso que el comportamiento sea el antes descrito, cóncavo o convexo. Se mostrarán también el comportamiento de los residuales del modelo con la transformación de Box-Cox y sin ella, de manera tal que se pueda observar las diferencias que existen en el comportamiento de los residuos luego de aplicar la transformación.

2.6. Residuales del Modelos Asociado al Tercer Subintervalo que Aproximan a la Distribución Normal.

En esta parte, mostraremos aquellas gráficas de cuantiles de los modelos asociados a los subintervalos que no requieren la aplicación de la transformación Box Cox, esto quiere decir como ya se señaló, que el parámetro λ que se requiere para aplicar dicha transformación es aproximadamente 1, entonces los modelos asociados a cada uno de estos subintervalos proporcionan un buen ajuste y se consideran apropiados para el estudio en cuestión. Particularmente, se hará el análisis completo para el tercer subintervalo, que como ya se dijo posee estas características.

A continuación, la expresión para el modelo $AR(1)$ asociado al tercer subintervalo se muestra específicamente y seguidamente, la ya mencionada gráfica de cuantiles.

$$x_t = 0,9194 + 0,2236x_{t-1} + \omega_t$$

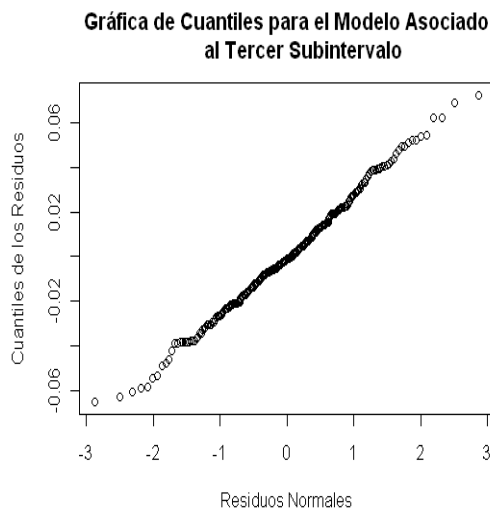


Figura 2.5: Gráficas de Cuantiles para el tercer subintervalo.

En este gráfico se presentan los residuos frente a los cuantiles de la distribución. Si los residuos tienen distribución Normal deberían quedar todos los puntos aproximadamente sobre una recta, este es el caso *ideal*, dado que es lo que se desea en el comportamiento de los residuales y obtener la gaussianidad del modelo. Además al calcular el parámetro λ que define la transformación de Box-Cox se aproxima a uno, por lo que como se mencionó en la Sección 1.2.4 entonces no es necesario aplicar la transformación a los datos.

De manera análoga este estudio se puede realizar para los demás subintervalo que poseen esta característica, es decir el cuarto, noveno, decimo, decimo primer y decimo segundo subintervalo los cuales no requieren de la transformación de Box-Cox y el comportamiento de los residuales es el mismo que el que se muestra en la Figura 2.5.

2.7. Residuales de los Modelos que Requieren de la Transformación de Box-Cox.

Daremos muestra en esta sección de los residuales de los modelos que requieren la aplicación de la transformación de Box-Cox, esto es aquellos para los cuales la gráfica de cuantiles de los residuos no se aproxima a una línea recta sino más bien que posee comportamiento cóncavo o convexo. En este sentido, usaremos el primer

subintervalo de la partición para hacer el estudio de la gráfica de cuantiles. Por un lado se muestran los residuales originales y por el otro luego de la transformación de tal manera de ver los efectos que surte esta transformación a la hora de obtener la aproximación lineal buscada.

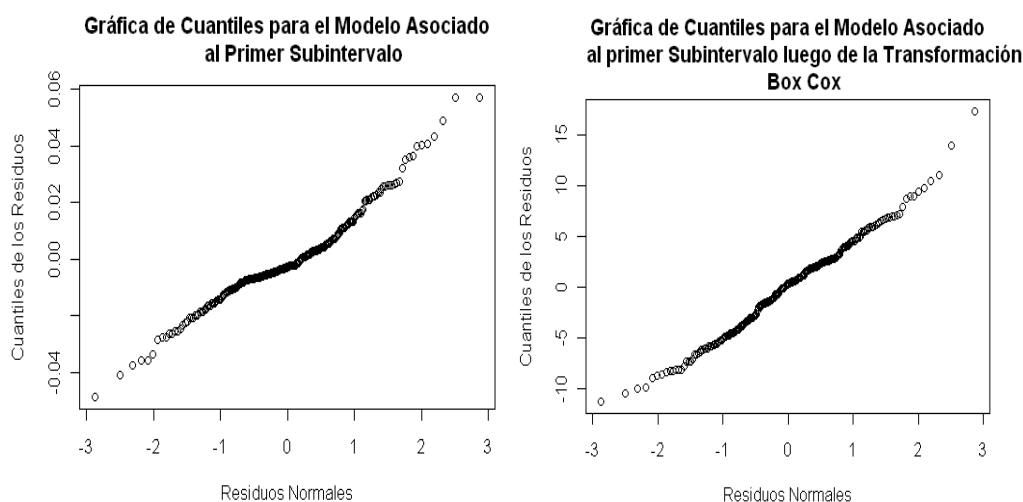


Figura 2.6: Gráficas de Cuantiles. A izquierda, el subintervalo original; a derecha, luego de la transformación Box-Cox.

En la figura anterior se pueden notar las mejoras para el primer subintervalo, observemos que en el gráfico de la izquierda el desarrollo de los cuantiles para el subintervalo original pasa de tener un comportamiento no lineal, a un comportamiento aproximadamente lineal, que es lo que se espera al hacer uso de la transformación Box-Cox. En este caso tenemos que el parámetro λ al aplicar la transformación para este subintervalo es $\lambda = -0,5520897$, por lo que la aplicación de la transformación tiene sentido desde el punto de vista teórico.

En resumen, como ya se había mencionado, la transformación Box-Cox sirve como herramienta para mejorar el comportamiento de la gráfica de cuantiles de los modelos en estudio, además si observamos y comparamos la evolución de los residuales normalizados, la ACF y los p -valores mediante el estadísticos de Ljung Box observamos que sobre todo los residuales normalizados muestran el aspecto y comportamiento aleatorio que se espera. Estos resultados serán de mucha utilidad para construir el modelo vectorial autorregresivo, de manera que el estudio anterior juega un papel fundamental en las características de modelos más generales que se propongan. Si-

milaramente, se puede observar que se logra la aproximación lineal que se desea hallar para los subintervalos segundo, quinto, sexto, séptimo y octavo, ver Apendice A.

2.8. Pronósticos de los Modelos Asociados a cada Subintervalo de $[0,04, 0,25]$.

Al elaborar pronósticos, el objetivo fundamental es predecir valores futuros de una serie de tiempo, basados en datos del presente. El objetivo del pronóstico en este trabajo es explorar que tan bien ajustan los modelos en estudio. En este caso, lo que se hará es predecir valores ya conocidos de la serie, es decir, puesto que conocemos 244 periodos de la tormenta y no podemos conocer más datos a futuro con precisión exacta, el procedimiento que usaremos consiste en trabajar con 234 periodos de los 244 y ajustar el modelo $AR(1)$ asociado para luego hacer el pronóstico de los 10 periodos que restan con una banda de confianza para la predicción hasta 95 %. Si los datos están dentro de esta banda de confianza, entonces esto es evidencia de un buen ajuste del modelo, además estos datos pronosticados deben aproximar a los valores originales que anteriormente fueron removidos.

En esta sección se muestra la evidencia gráfica para algunos de los subintervalos en estudio. Para el caso de los subintervalos donde se usó la transformación de Box-Cox presentaremos ambos modelos para el primer subintervalo, el de los datos originales y al que se le aplicó la transformación de Box-Cox en conjunto con sus respectivos pronósticos. El análisis para los subintervalos restantes que ya hemos mencionados se hace de la misma forma. Los modelos fueron obtenidos en base a la suma de los datos para el correspondiente intervalo con y sin la transformación de Box-Cox, según sea el caso y los residuos asociados al modelo $AR(1)$. De manera similar, para los subintervalos que no requieren de la transformación de Box-Cox se muestra el análisis para el tercer subintervalo con su respectivo pronóstico con el que se muestra el buen ajuste para el subintervalo. También para los subintervalos que tienen esta característica, el estudio es análogo.

2.8.1. Pronóstico para Subintervalos con la Transformación de Box-Cox.

El pronóstico se muestra en la Figura A.27, en donde al hacer la restricción de los 10 datos (en rojo) se tiene que el pronóstico realizado se encuentra dentro

de la banda de confianza (líneas punteadas) hasta 95 %, se muestra el pronóstico anteriormente comentado, frente a los datos originales (en azul), donde se evidencia una gran similitud entre el modelo encontrado y los datos originales, además los datos originales se hallan dentro de la banda de confianza considerada en este caso. En general, para aquellos subintervalos donde se aplicó la transformación de Box-

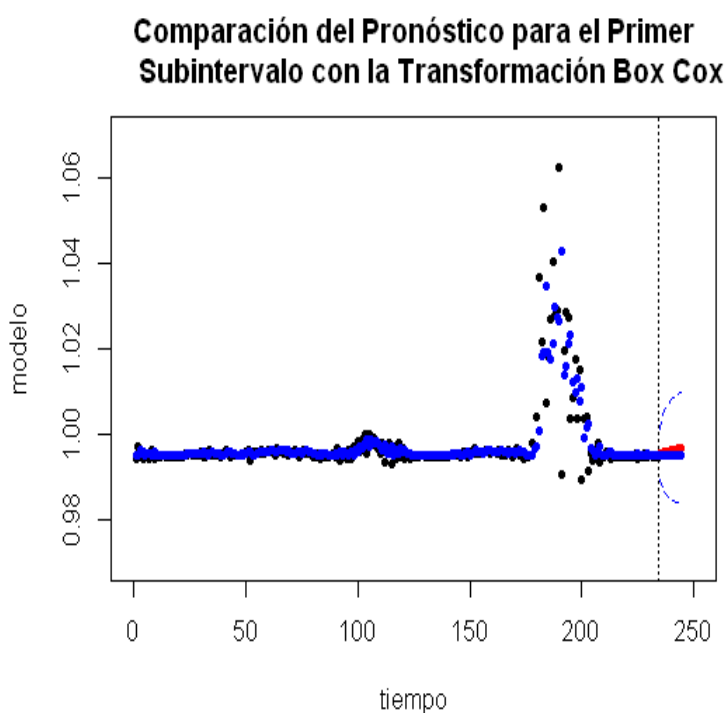


Figura 2.7: Primer subintervalo: El subintervalo (en azul) frente al pronóstico para el modelo(negro).

Cox se obtuvo que el pronóstico realizado para el respectivo modelo es bastante confiable en el sentido de que los 10 datos que hemos reproducido como se explicó anteriormente están dentro de la banda de confianza considerada casi en su totalidad, lo que indica que estos pronósticos son capaces de *imitar* el comportamiento de los datos originales.

2.8.2. Pronóstico para los Subintervalos que no Requieren Transformación Box-Cox.

En esta sección, se muestran aquellos subintervalos que no requieren de la transformación de Box-Cox, es decir aquellos para los cuales el parámetro $\lambda \approx 1$, además los cuantiles para el modelo asociado a este subintervalo aproxima a una distribución normal. De la misma manera que la sección anterior procedemos a dar la prueba gráfica del pronóstico para estos subintervalo, en particular para el tercer subintervalo tenemos los siguientes pronósticos y además la comparación del modelo con los datos originales, de lo que se puede apreciar el buen ajuste del modelo con los datos de partida que hemos considerado. Además como se observó en la Figura A.23 los residuales para este modelo, al igual que para los demas subintervalos que ya se han mencionado que no requieren de la transformación, aproximan la línea recta que se requiere para obtener normalidad en los cuantiles. La Figura 2.8 a continuación muestra el pronóstico para el tercer subintervalo.

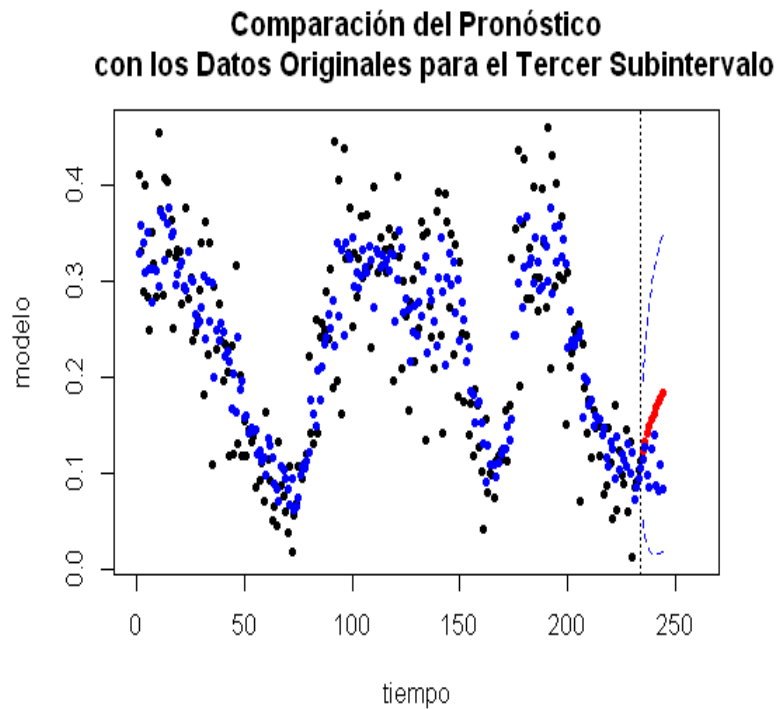


Figura 2.8: Pronóstico para el tercer subintervalo; el subintervalo (en azul) frente al pronóstico para el modelo (negro).

De esta manera, hemos obtenido un buen pronóstico tanto para los subintervalos con la transformación de Box-Cox, como para los que no requieren de esta. De esta manera, podemos construir modelos un poco más complejos para explicar el comportamiento de la tormenta en estudio, dichos modelos son lo ya mencionados Modelos Vectoriales Autorregresivos que vienen a explicar en conjunto los modelos AR que hemos encontrado de forma que se pueda lograr una mejor interpretación acerca del comportamiento de esta tormenta.

CAPÍTULO 3

MODELOS VECTORIALES AUTORREGRESIVOS VAR

Al construir modelos vectoriales autorregresivos podemos extender la teoría conocida sobre los modelos autorregresivos *AR*. Con estos vectores se pretende investigar la interacción dinámica entre las variables que intervienen en el estudio. Esta interacción involucra algunas consideraciones para ser estudiada y así obtener los modelos vectoriales que expliquen de forma adecuada relación entre las variables con las que se trabaja. El estudio de la cointegración entre las variables que interaccionan, es una de las condiciones que se debe investigar, de estar presente asegura la estacionaridad de la serie que resulta de la combinación de las variables. Otra de las condiciones que se requiere y que es fundamental en el estudio para construir vectores autorregresivos es la existencia de raíces unitarias o no. En presencia de raíces que superan la unidad en valor absoluto, indica que el modelo es estable, quiere decir que se puede considerar aceptable el modelo que se analiza. En particular y además es el caso del presente trabajo, los modelos VAR resultan de orden uno, para este caso, la estacionaridad o no se verifica con la ecuación 1.31, caso contrario, cuando las raíces de la ecuación característica que proviene de estos modelos no supera la unidad entonces decimos que el modelo es inestable. Otro aspecto que se debe considerar para la aceptación de modelos VAR es el pronóstico del modelo que resulta de la interacción de las variables, como siempre en conjunto con la ACF para cada submodelo de la combinación nos proporcionan herramientas para aceptar o reconsiderar los modelos que se estudian.

En este sentido, para los efectos del presente estudio con la ayuda del software R, se construirán los modelos VAR correspondientes al conjunto de subintervalos que han resultado de la partición del intervalo $[0,04, 0,25]$. Primero, se construyó el modelo para el conjunto de los 12 subintervalos que se están estudiando, el resultado es un modelo de orden 18 pero que resulta difícil de manejar, ya que no existe una forma computacional de dominar este tipo de modelos. La metodología a emplear entonces

para los 12 subintervalos establecidos es relacionar estos dos a dos, es decir obtener un modelo VAR para el primer subperiodo con el segundo, luego relacionar un modelo del mismo tipo para el segundo subperiodo con el tercero y así sucesivamente para los subintervalos restantes. Este análisis se realiza bajo la suposición de que existe correlación entre las variables que se están estudiando, de otro modo entraríamos en la discusión de la causalidad entre las variables y en este caso por tener las variables el mismo origen, se supone correlación.

3.1. Modelos VAR Asociado al Primer y Segundo Subintervalo.

El modelo asociado a los dos primeros subintervalos tiene orden 12, el cual es un orden complicado de manejar, sin embargo las técnicas usadas para determinar el orden adecuado nos proporcionaran la información necesaria para aceptar este orden del modelo o adoptar otro. Por otro lado, la correlación entre los dos subintervalos es 0,88, La grafica de estos dos subintervalos se ve en la Figura 3.1, donde se evidencia la fuerte correlación para la mayoría de los datos, mientras nos alejamos del 1 empiezan a ser un poco dispersos.

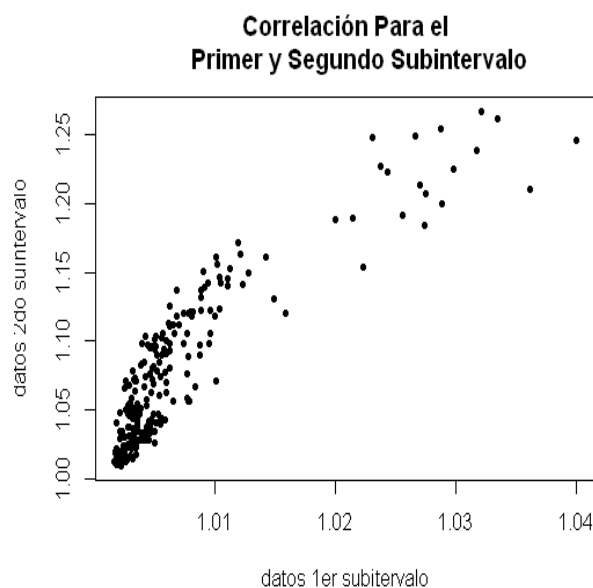


Figura 3.1: Correlación entre el primer y segundo subintervalo.

Por otro lado, al aplicar el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene el 5% para el primer subintervalo, y para el segundo subintervalo se obtiene el 6%. De esta forma se puede afirmar que no existe evidencia para rechazar la hipótesis para raíces unitarias, concluimos entonces que tanto el primer como el segundo subintervalo poseen raíces unitarias.

Para verificar si estos dos subperíodos están o no cointegrados usamos el test de Phillips-Ouliaris, el resultado es que para ambos subintervalos el 1% de los p -valores rechaza la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, para este caso el modelo de regresión tiene orden 4 por lo que el primer orden que podríamos intentar para el modelo ARMA es 4, y así variar los parámetros p y q para el cálculo del modelo en R comenzando con $p = 4$. A continuación se muestran los distintos AIC calculados,

- Para un modelo $AR(1)$ el AIC es $-2300,905$
- Para un modelo $ARMA(1, 1)$ el AIC es $-2352,228$
- Para un modelo $AR(2)$ el AIC es $-2332,188$

Por lo que, comparando los AIC para estos tres modelos, tenemos que el $AR(1)$ es el de menor valor. Dicho modelo esta dado por

$$\begin{array}{rcc} & x_1 & x_2 \\ x_1 & 0,85 & 0,013 \\ x_2 & 1,05 & 0,842 \end{array}$$

x_1 y x_2 son el primer y segundo subintervalo respectivamente. Por otro lado, las raíces del determinante que produce la ecuación característica asociada a este modelo son 1,1 y 1,3 las cuales exceden la unidad, por lo que se puede asegurar que el modelo $AR(1)$ considerado para la combinación de estos dos subintervalos es estacionario.

El modelo Vectorial Autorregresivo asociado a la combinación de estos subintervalos es un VAR(1) cuyos coeficientes estan dados a continuación. Para x_1 tenemos

	Estimate	Std. Error	t value	$Pr(> t)$
$x_1.l_1$	1,002	$2,99e^{-03}$	335,37	$1,19e^{-322}$
$x_2.l_1$	$-2,46e^{-03}$	$2,80e^{-03}$	-0,88	$3,81e^{-01}$
trend	$-1,69e^{-07}$	$2,09e^{-06}$	-0,081	$9,4e^{-01}$

Para las variables x_2 tenemos que

	Estimate	Std. Error	t value	$Pr(> t)$
$x_1.l_1$	$6,26e^{-02}$	$2,40e^{-02}$	2,61	$9,72e^{-03}$
$x_2.l_1$	$9,41e^{-01}$	$2,25e^{-02}$	41,78	$4,24e^{-112}$
trend	$2,35e^{-06}$	$1,68e^{-05}$	0,14	$8,89e^{-01}$

Con respecto a los residuales, se muestra a continuación en la Figura 3.2, el comportamiento de la ACF para cada uno de los subintervalos

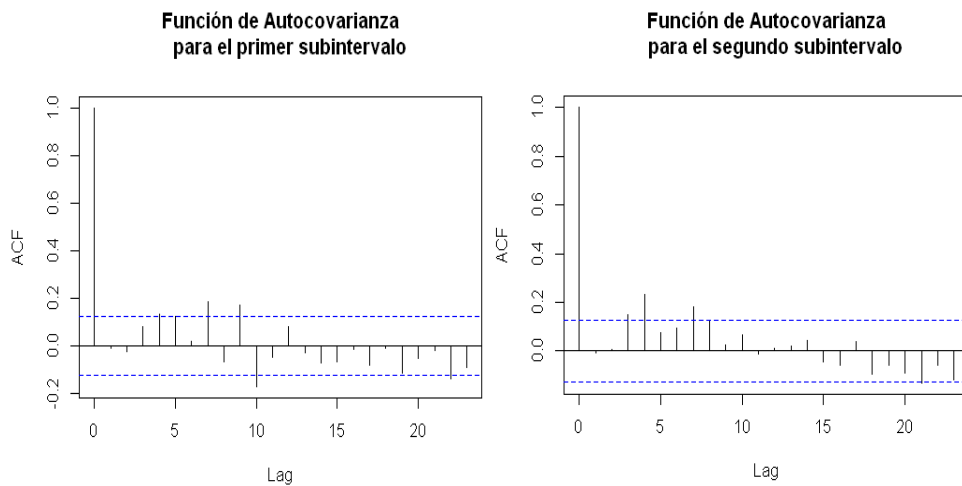


Figura 3.2: A izquierda, la ACF para el primer subintervalo; a derecha, la ACF para el segundo subintervalo.

Haciendo los pronósticos respectivos se obtiene la siguiente información para cada submodelo, donde se han eliminado los últimos diez datos del subintervalo original para construir el modelo VAR y reproducir estos, de manera tal que se pueda verificar si con esto los modelos son capaces de comportarse similar a los datos de donde se hallaron. Gráficamente para los dos modelos asociados a estos subintervalos y

comparandolos con el subintervalo original se ven en la Figura 3.3 correspondiente obtenemos, Para el segundo subintervalo, obtenemos para los datos originales y para

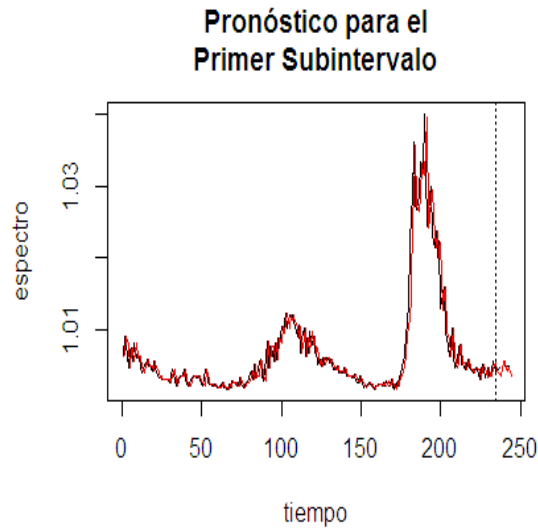


Figura 3.3: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

el pronóstico respectivo, se puede apreciar en la Figura 3.4

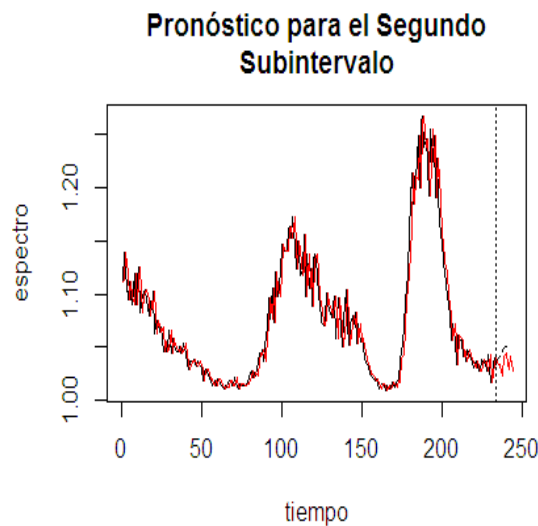


Figura 3.4: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.2. Modelo VAR Asociado al Segundo y Tercer Subintervalo.

Siguiendo la misma metodología, para el modelo VAR asociado al segundo y tercer subintervalo se obtuvo la siguiente información. La correlación entre estos dos subintervalos es 0,7633449, la Figura 3.5 muestra gráficamente la correlación entre estos subintervalos.

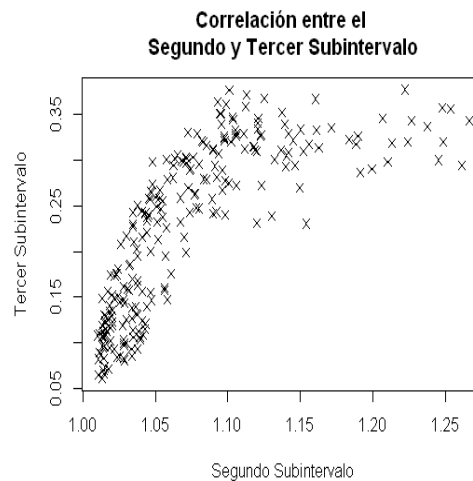


Figura 3.5: Correlación entre el segundo y tercer subintervalo.

El modelo originalmente tiene orden 13. Luego de aplicar los test correspondientes respectivamente el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el tercer subintervalo se obtiene el 4%, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1% de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos órdenes del modelo ARMA calculado, de donde se obtiene que:

- Para un $ARMA(12, 0)$ el AIC es $-1340,165$.
- Para un $ARMA(2, 0)$ el AIC es $-1327,307$.

- Para un $ARMA(1,0)$ el AIC es $-1268,254$.

Al comparar las AIC para distintos ordenes del modelo para los residuos asociados a estos dos subintervalo, se determino que el orden de dicho modelo es uno. El modelo para la combinación de estos dos subintervalos esta dado por

$$\begin{matrix} & x_2 & x_3 \\ x_2 & 0,93 & 0,02 \\ x_3 & 0,1 & 0,87 \end{matrix}$$

Donde x_3 es el tercer subintervalo. Ademas al calcular las raíces asociadas al polinomio característico asociado al modelo tenemos que estas raíces son 1,1 y 1,2 por lo que el modelo es estacionario. El Modelo Vectorial Autorregresivo para el segundo subintervalo (x_2)

	Estimate	Std. Error	t value	$Pr(> t)$
$x_2.l_1$	$1,0e^{04}$	$4,8e^{-03}$	210,8	$2,16e^{-274}$
$x_3.l_1$	$1,72e^{-02}$	$1,58e^{-02}$	-1,08	$2,78e^{-01}$
trend	$-4,80e^{-06}$	$1,84e^{-05}$	-0,26	$7,94e^{-01}$

Para el tercer subintervalo x_3

	Estimate	Std. Error	t value	$Pr(> t)$
$x_2.l_1$	$2,83e^{-02}$	$9,36e^{-03}$	3,02	$2,8e^{-03}$
$x_3.l_1$	$8,89e^{-01}$	$3,11e^{-02}$	28,58	$3,13e^{-79}$
trend	$-4,99e^{-05}$	$3,63e^{-05}$	-1,38	$1,7e^{-01}$

Las ACF para cada submodelo se muestra en la Figura 3.6

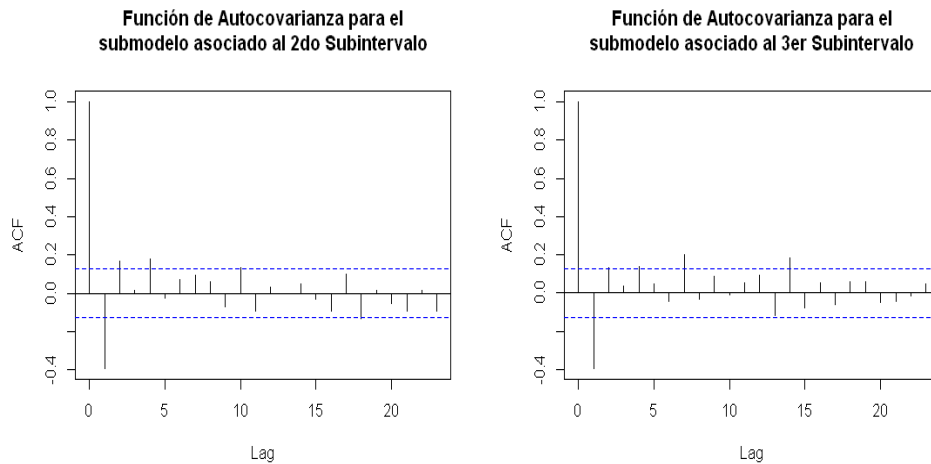


Figura 3.6: Función de Autocovarianza asociada al segundo (izquierda) y tercer (derecha) subintervalo.

Ademas para los pronósticos asociados a cada submodelo junto al subintervalo original se muestran a continuación en la Figura 3.7 muestra el pronóstico para el segundo subintervalo

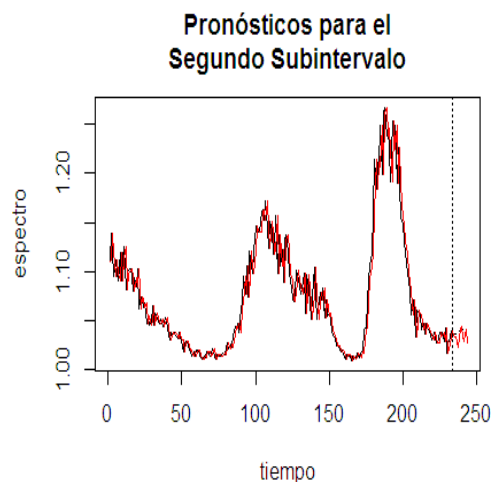


Figura 3.7: subintervalo (rojo), modelo (negro), pronóstico (punteadas)

Ahora para el tercer subintervalo se tiene para el pronóstico del submodelo asociado en la siguiente figura (Figura 3.8)

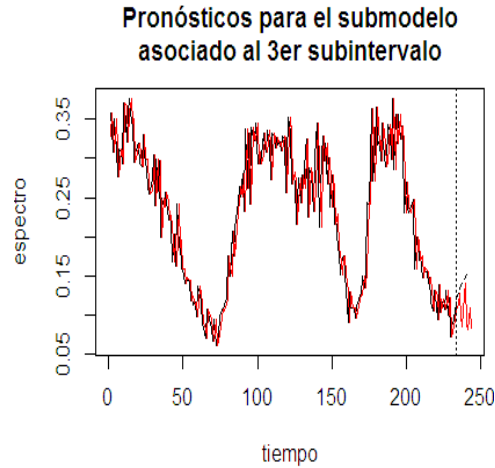


Figura 3.8: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.2.1. Modelo Asociado al Tercer y Cuarto Subintervalo.

Para el tercer y cuarto subintervalo, se tiene que la correlación existente entre estos es $-0,031$ la gráfica de la correlación se muestra a continuación en la Figura 3.9

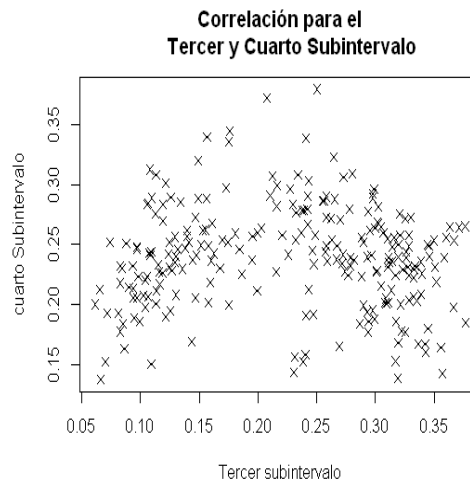


Figura 3.9: Correlación entre el tercer y cuarto subintervalo.

El modelo originalmente tiene orden 3, el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el cuarto subintervalo se obtiene el 1%, por lo que se puede afirmar que no existe la posibilidad de tener

raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 15 % de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, de donde se obtiene que:

- $ARMA(2,0)$ tiene AIC igual a $-967,2571$
- $ARMA(1,1)$ tiene AIC igual a $-967,3863$
- $ARMA(1,0)$ tiene AIC igual a $-915,3804$

Por lo que podemos afirmar que el modelo tiene orden 1. Los coeficientes del modelo son los siguientes

$$\begin{array}{ccc} & x_3 & x_4 \\ x_3 & 0,92 & 0,1 \\ x_4 & -0,02 & 0,71 \end{array}$$

Donde x_4 es el cuarto subintervalo. Además al calcular las raíces asociadas al polinomio característico asociado al modelo tenemos que estas raíces son 1,22 y 1,22 por lo que el modelo es estacionario. El Modelo Vectorial Autorregresivo para el tercer subintervalo

	Estimate	Std. Error	t value	$Pr(> t)$
$x_3.l_1$	$9,21e^{-01}$	$2,33e^{-02}$	39,51	$5,01e^{-107}$
$x_4.l_1$	$7,59e^{-02}$	$2,70e^{-02}$	2,80	$5,41e^{-03}$
trend	$-1,20e^{-05}$	$2,91e^{-05}$	-0,41	$6,80e^{-01}$

Para la variable x_4 tenemos

	Estimate	Std. Error	t value	$Pr(> t)$
$x_3.l_1$	$3,24e^{-02}$	$2,1e^{-02}$	1,55	$1,21e^{-01}$
$x_4.l_1$	$9,5e^{-01}$	$2,42e^{-02}$	39,16	$3,29e^{-106}$
trend	$3,1e^{-05}$	$2,60e^{-05}$	1,19	$2,35e^{-01}$

Las funciones de autocorrelación para cada subintervalo se muestran a continuación en la Figura 3.10

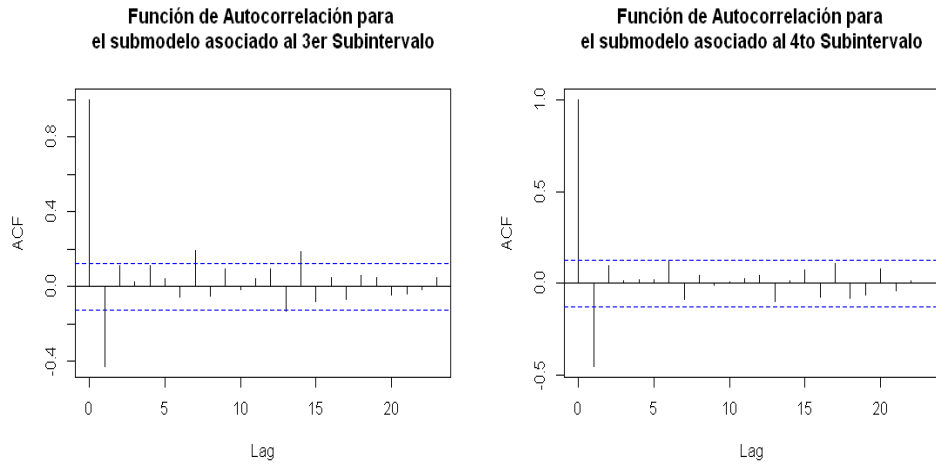


Figura 3.10: Función de autocovarianza asociado al submodelo para el 3er subintervalo a izquierda y para el 4to subintervalo a derecha.

Ademas los pronósticos para cada submodelo asociado son, la Figura 3.11 para el tercer submodelo

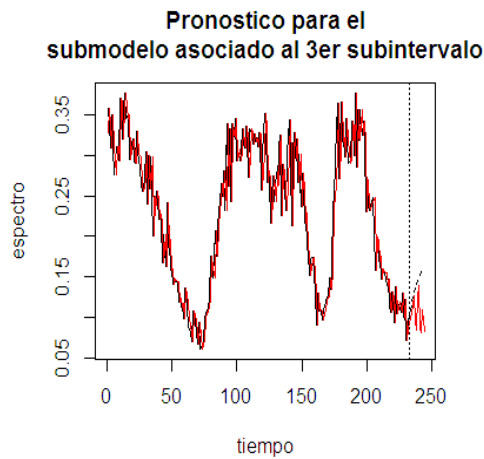


Figura 3.11: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

Similarmente para el cuarto subintervalo se tiene en la Figura 3.12

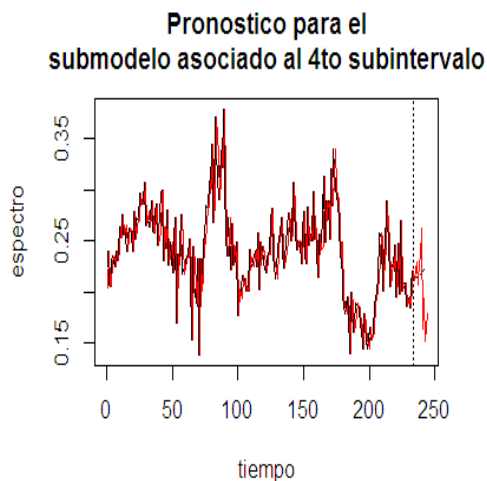


Figura 3.12: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.3. Modelo Asociado al Cuarto y Quinto Subintervalo.

se tiene en principio que la correlación que existe entre estos dos subintervalos es 0,1740726 graficamente se puede observar a continuación en la Figura 3.13

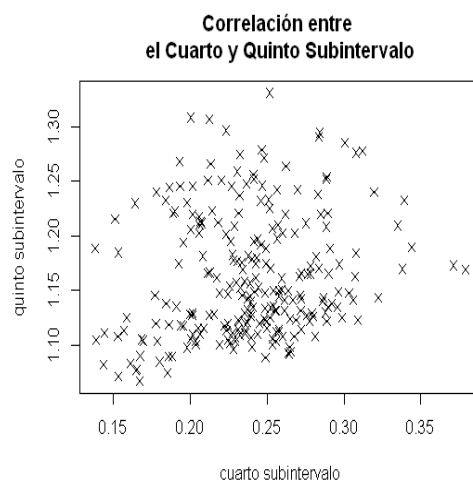


Figura 3.13: Correlación entre el cuarto y quinto subintervalo.

En principio el modelo para la combinación de estos subintervalos es 3. De igual forma que en los otros casos, el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el cuarto subintervalo se obtiene el

1%, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1% de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están coin-
tegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, de donde se obtiene que:

- ARMA(3, 0) tiene AIC igual a $-1058,838$
- ARMA(2, 0) tiene AIC igual a $-1063,449$
- ARMA(1, 0) tiene AIC igual a $-1017,513$

Por lo que podemos afirmar de igual forma que las anteriores que el modelo tiene orden 1. Los coeficientes del modelo son los siguientes

$$\begin{array}{ccc} & x_4 & x_5 \\ x_4 & 0,69 & 0,47 \\ x_5 & -0,11 & 0,84 \end{array}$$

Donde x_5 es el quinto subintervalo. Además, las raíces de la ecuación característica asociada a este modelo son 1,29 y 1,29, los coeficientes para el modelo vectorial autorregresivo para el cuarto subintervalo (x_4) son

	Estimate	Std. Error	t value	Pr(> t)
$x_4.l_1$	0,62	$5,02e^{-02}$	12,28	$3,20e^{-27}$
$x_5.l_1$	0,1	$1,22e^{-02}$	7,44	$1,74e^{-12}$
trend	-0,00011	$3,07e^{-05}$	-3,73	$2,43e^{-04}$

Para el quinto subintervalo tenemos el modelo x_5

	Estimate	Std. Error	t value	Pr(> t)
$x_4.l_1$	$-9,44e^{-02}$	$5,62e^{-02}$	-1,68	$9,43e^{-02}$
$x_5.l_1$	1,02	$1,37e^{-02}$	74,54	$5,84e^{-168}$
trend	$-1,19e^{-05}$	$3,44e^{-05}$	-0,35	$7,28e^{-01}$

Ademas las funciones de autocorrelación se muestran en la Figura 3.14

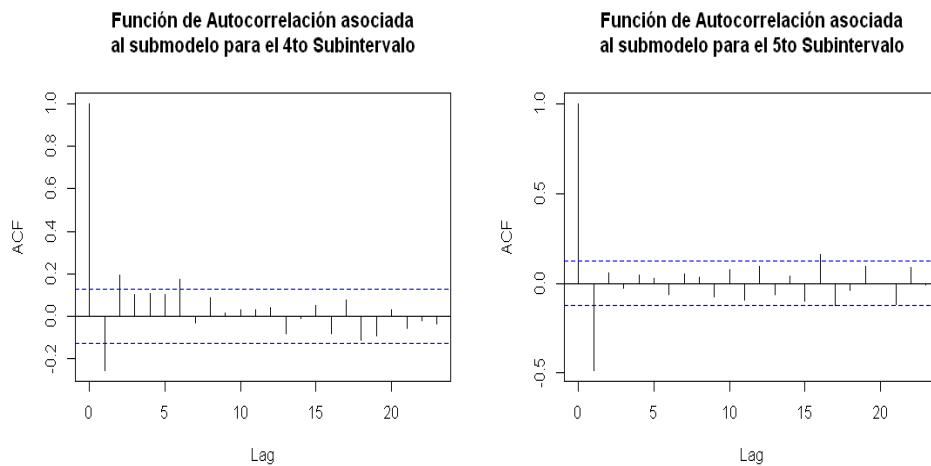


Figura 3.14: Función de autocovarianza asociada a los submodelos para el cuarto (izquierda) y quinto (derecha) subintervalo.

Los pronósticos para cada submodelo asociado se muestran a continuación. Para el cuarto subintervalo se puede observar en la Figura 3.15

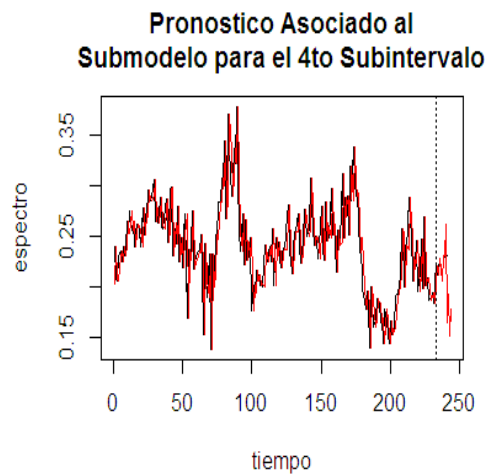


Figura 3.15: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

Similarmente para el quinto subintervalo, se aprecia la Figura 3.16

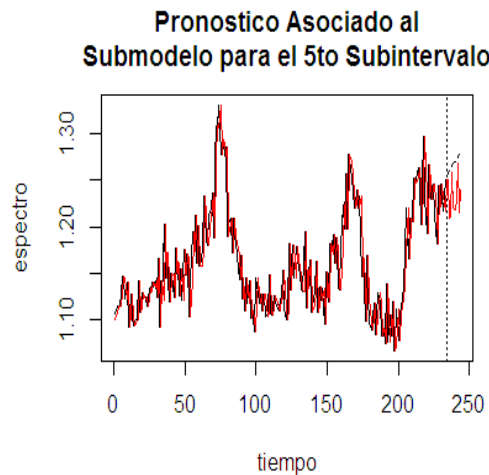


Figura 3.16: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.4. Modelo Asociado al Quinto y sexto Subintervalo

En este caso, se tiene una correlación de 0,79 la evidencia grafica se muestra a continuación, en la Figura 3.17, donde se observa un fenomeno similar a los anteriores

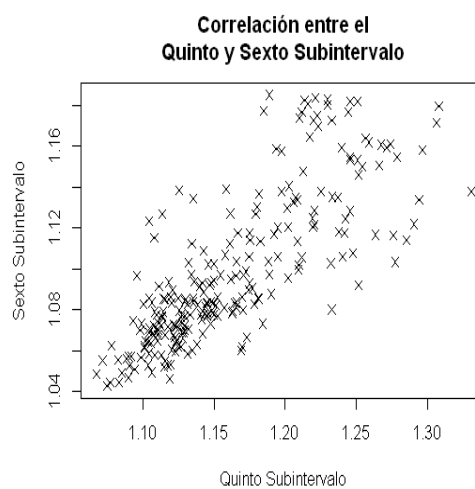


Figura 3.17: Correlación entre el Quinto y Sexto Subintervalo.

Originalmente el modelo AR asociado a estos subintervalo es 3. Luego de aplicar los test correspondientes respectivamente el test Phillips-Perron para explorar la

existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el tercer subintervalo se obtiene el 1%, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1% de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, de donde se obtiene que:

- Para un $ARMA(3, 0)$ el AIC es $-1026,65$.
- Para un $ARMA(2, 0)$ el AIC es $-1017,137$.
- Para un $ARMA(1, 0)$ el AIC es $-998,2336$.

Por lo que el modelo AR para estos subintervalo tiene orden 1. Los coeficientes para el modelo son los siguientes

$$\begin{matrix} & x_5 & x_6 \\ x_5 & 0,58 & 0,49 \\ x_6 & 0,11 & 0,67 \end{matrix}$$

Donde x_6 es el sexto subintervalo. Al construir el polinomio característico asociado al modelo para determinar la estacionaridad se obtuvieron las raíces 1,17 y 2,50 las cuales por teoría determinan que el modelo es estacionario. Los coeficientes para el modelo VAR asociado a estos dos subintervalo es

	Estimate	Std. Error	t value	$Pr(> t)$
$x_5.l_1$	$5,8e^{-01}$	$5,46e^{-02}$	10,64	$6,68e^{-22}$
$x_6.l_1$	$4,42e^{-01}$	$5,77e^{-02}$	7,66	$4,58e^{-13}$
trend	$1,70e^{-05}$	$2,78e^{-05}$	0,61	$5,41e^{-01}$

Para el sexto subintervalo tenemos el modelo x_6

	Estimate	Std. Error	t value	$Pr(> t)$
$x_5.l_1$	$7,22e^{-02}$	$4,26e^{-02}$	1,70	$9,15e^{-02}$
$x_6.l_1$	$9,22e^{-01}$	$4,50e^{-02}$	20,49	$1,24e^{-54}$
trend	$8,30e^{-06}$	$2,17e^{-05}$	0,38	$0,24e^{-01}$

Por otro lado, la función de autocovarianza para cada subintervalo se puede notar en la Figura 3.18

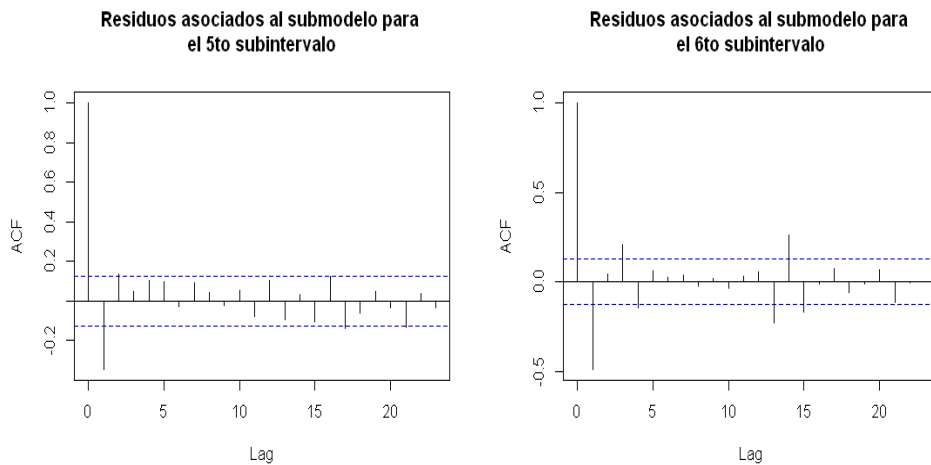


Figura 3.18: Función de Autocovarianza para los residuales asociados a los submodelos para el quinto subintervalo(izquierda) y sexto subintervalo (derecha).

Los pronósticos asociados a cada subintervalo, evidencian que, para el quinto subintervalo se tienen las siguientes apreciaciones en la Figura 3.19

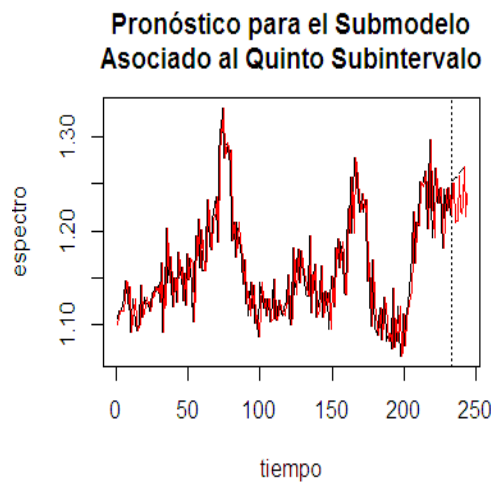


Figura 3.19: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

Ahora para el sexto subintervalo, la Figura 3.20 da la evidencia gráfica para este

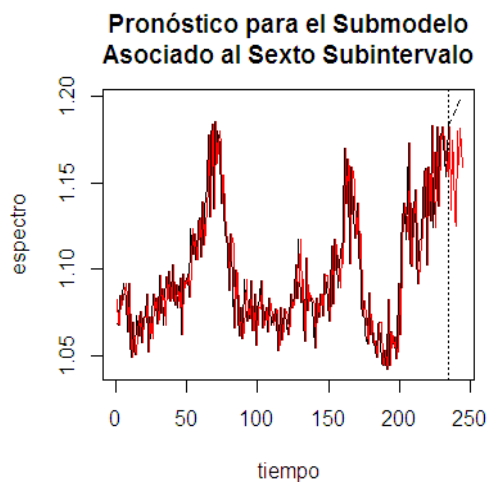


Figura 3.20: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.5. Modelo Asociado al Sexto y Séptimo Subintervalo.

Para el modelo VAR asociado al sexto y séptimo subintervalo, se tiene una correlación entre ellos de 0,86 cuya grafica se observa a continuación en la Figura 3.21

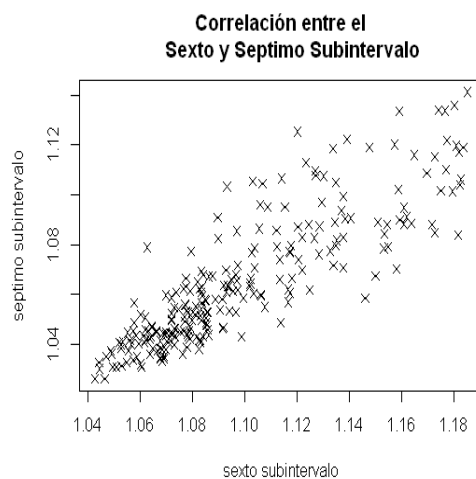


Figura 3.21: Correlación entre el sexto y séptimo subintervalo.

El modelo AR para la combinación de estos dos subintervalos tiene orden 3. Luego de aplicar los test correspondientes respectivamente el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el tercer

subintervalo se obtiene el 1 %, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1 % de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, de donde se obtiene que:

- ARMA(5, 0) el AIC es $-1303,892$.
- ARMA(2, 0) el AIC es $-1289,499$.
- ARMA(1, 0) el AIC es $-1273,505$.

Se obtiene que por el valor de la AIC el orden del modelo sigue siendo uno, el modelo se muestra a continuación

$$\begin{matrix} & x_6 & x_7 \\ x_6 & 0,52 & 0,48 \\ x_7 & 0,13 & 0,64 \end{matrix}$$

Donde x_7 es el séptimo subintervalo. Además las raíces del polinomio característico asociado a los coeficientes de este modelo son 1,2 y 3,1 lo cual justifica que el modelo es estacionario. Además los coeficientes asociados al modelo VAR son los siguientes

	Estimate	Std. Error	t value	$Pr(> t)$
$x_6.l_1$	$4,52e^{-02}$	$7,28e^{-02}$	6,20	2,39
$x_7.l_1$	$5,59e^{-01}$	$7,45e^{-02}$	7,51	1,17
trend	$5,25e^{-05}$	$2,03e^{-05}$	2,58	$1,05e^{-02}$

Para el séptimo subintervalo tenemos el modelo para x_7

	Estimate	Std. Error	t value	$Pr(> t)$
$x_6.l_1$	$5,27e^{-02}$	$5,88e^{-02}$	0,89	$3,71e^{-01}$
$x_7.l_1$	$9,45e^{-01}$	$6,01e^{-02}$	15,72	$8,81e^{-39}$
trend	$5,55e^{-06}$	$1,64e^{-05}$	0,34	$7,36e^{-01}$

Ademas, la función de autocovarianza para los submodelos asociados a estos subintervalos se evidencian en la Figura 3.21

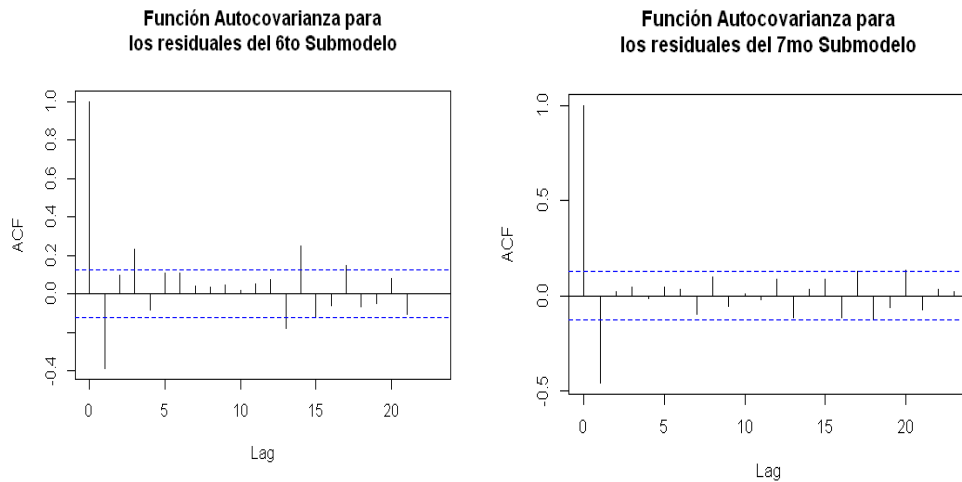


Figura 3.22: Función Autocovarianza para los residuales del sexto submodelo (izquierda) y séptimo submodelo (derecha).

Los pronósticos asociados a estos submodelos se evidencian a continuación, para el sexto subintervalo se tiene la Figura 3.23

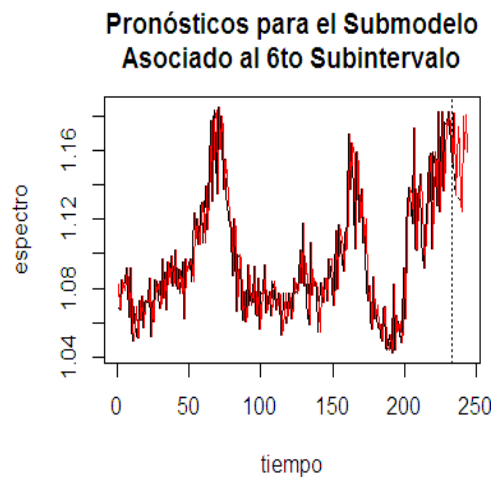


Figura 3.23: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

Para el septimo subintervalo se muestra en la Figura 3.24

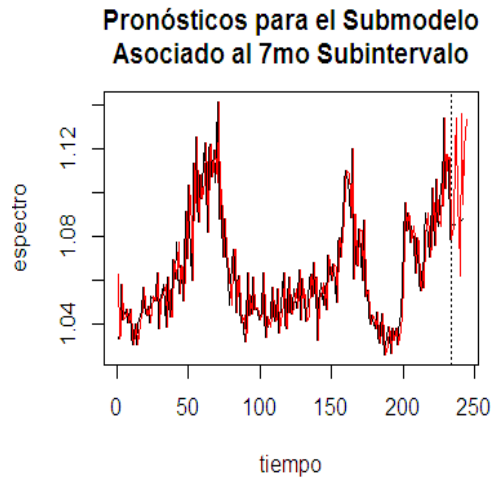


Figura 3.24: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.6. Modelo Asociado al Septimo y Octavo Subintervalo.

Para la construcción del modelo VAR asociado al octavo y noveno subintervalo, comenzamos con mostrar que existe una correlación de 0,9 la evidencia gráfica se puede notar en la Figura 3.25

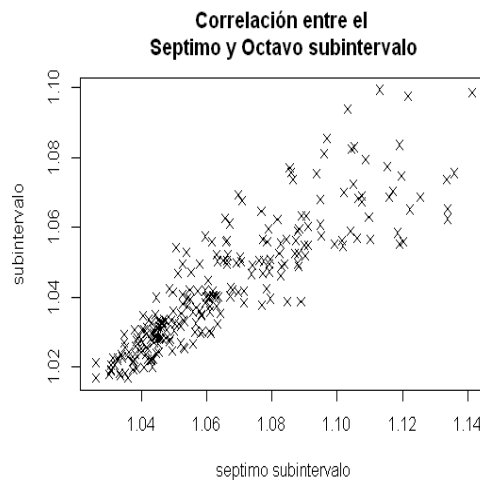


Figura 3.25: Correlación entre el septimo y octavo subintervalo.

El modelo AR para la combinación de estos dos subintervalos tiene orden 7. De

nuevo, aplicar los test correspondientes respectivamente el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el tercer subintervalo se obtiene el 1%, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1% de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos órdenes del modelo ARMA calculado, de donde se obtiene que:

- $ARMA(11, 0)$ el AIC es $-1485,707$.
- $ARMA(2, 0)$ el AIC es $-1472,154$.
- $ARMA(1, 0)$ el AIC es $-1465,835$.

De igual forma, el AIC con menor valor es para el $ARMA(1, 0)$ por lo que podemos asegurar que el modelo AR para la combinación de estos subintervalo es uno. Lo coeficientes del modelo estan dados por

$$\begin{array}{cc}
 & x_7 & x_8 \\
 x_7 & 0,40 & 0,65 \\
 x_8 & 0,10 & 0,70
 \end{array}$$

x_8 es el octavo subintervalo. Las raíces del polinomio característico asociado a los coeficientes del modelo $AR(1)$ son 1,2 y 3,81, lo que indica que el modelo es estacionario. Los coeficientes para el modelo VAR asociado son respectivamente para el septimo (x_7) y octavo (x_8) subintervalo

	Estimate	Std. Error	t value	$Pr(> t)$
$x_7.l_1$	$3,54e^{-01}$	$7,69e^{-02}$	4,61	$6,62e^{-06}$
$x_8.l_1$	$6,55e^{-01}$	$7,81e^{-02}$	8,38	$4,44e^{-15}$
trend	$4,25e^{-05}$	$1,43e^{-05}$	2,96	$3,33e^{-03}$

Para el octavo subintervalo tenemos el modelo

	Estimate	Std. Error	t value	$Pr(> t)$
$x_7.l_1$	$-1,57e^{-02}$	$5,68e^{-02}$	-0,28	$7,82e^{-01}$
$x_8.l_1$	1,01	$5,76e^{-02}$	17,61	$3,81e^{-42}$
trend	$5,13e^{-06}$	$1,05e^{-05}$	0,49	$6,27e^{-01}$

Por otro lado, las funciones de autocovarianza para los residuales asociados a cada componente del modelo se aprecian en la Figura 3.26

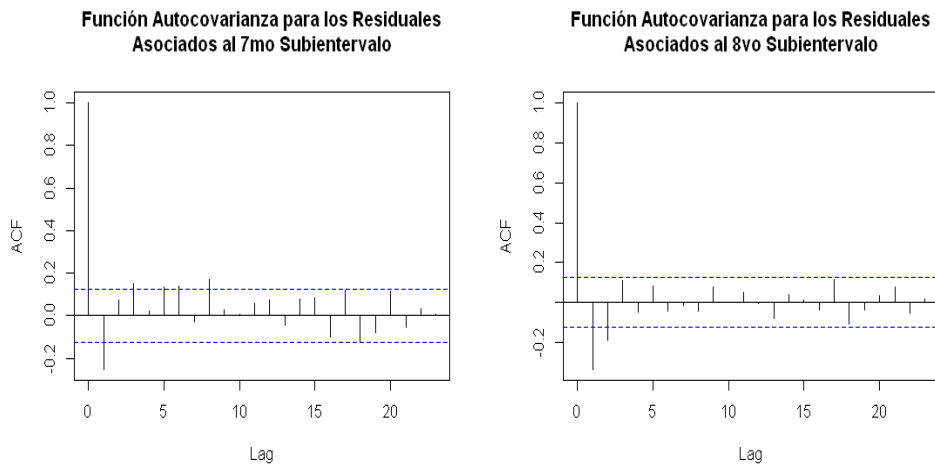


Figura 3.26: Función Autocovarianza para los residuales del séptimo submodelo (izquierda) y octavo submodelo (derecha).

Los pronósticos asociados al submodelo asociado al séptimo subintervalo se muestran en la Figura 3.27

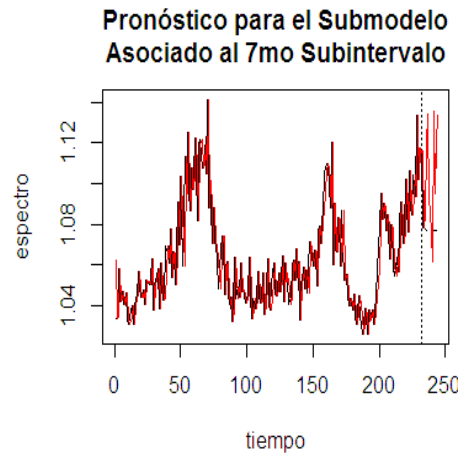


Figura 3.27: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

Para el submodelo vectorial asociado al octavo subintervalo tenemos la Figura 3.28

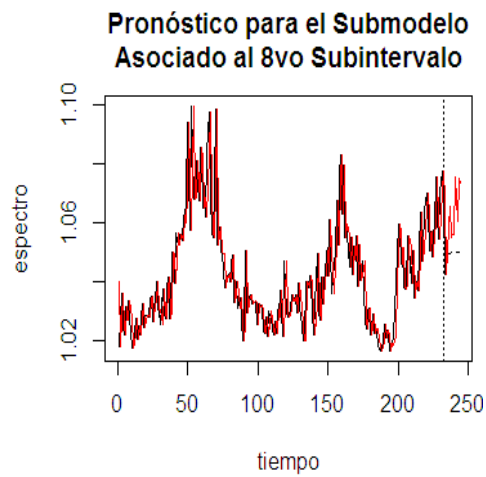


Figura 3.28: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.7. Modelo Asociado al Octavo y Noveno Subintervalo.

Para obtener el modelo vectorial asociado a los subintervalos octavo y noveno se obtuvo que la correlación entre estos subintervalos es de 0,9 gráficamente esta correlación se muestra en la Figura 3.33

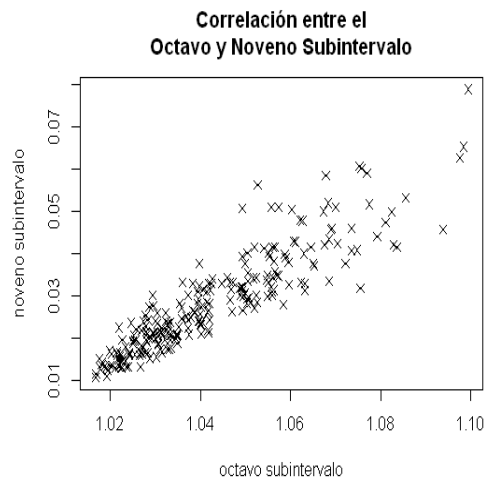


Figura 3.29: Correlación entre el octavo y noveno subintervalo.

El modelo AR para la combinación de estos dos subintervalos tiene orden 5. De nuevo, aplicar los test correspondientes respectivamente el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el tercer subintervalo se obtiene el 1%, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1% de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, de donde se obtiene que:

- $ARMA(5, 0)$ el AIC es $-1701,303$.
- $ARMA(2, 0)$ el AIC es $-1684,676$.
- $ARMA(1, 0)$ el AIC es $-1683,246$.

De igual forma se obtuvo que el modelo autorregresivo asociado a la combinación de estos subintervalos es de orden uno, los coeficientes del modelo son

$$\begin{array}{cc} & x_8 & x_9 \\ x_8 & 0,61 & 0,33 \\ x_9 & 0,34 & 0,29 \end{array}$$

x_9 se refiere al noveno subintervalo. Se puede entonces afirmar la estacionaridad del modelo ya que las raíces del polinomio característico asociado a este modelo son 1,23 y 11,63 las cuales superan la unidad, por lo que se puede decir que el modelo es estacionario. Para el modelo VAR asociado a estos se tiene que para el octavo subintervalo

	Estimate	Std. Error	t value	Pr(> t)
$x_8.l_1$	1,00	$2,11e^{-03}$	474,51	0,000
$x_9.l_1$	$-1,89e^{-01}$	$6,04e^{-02}$	-3,13	0,002
trend	$4,90e^{-06}$	$9,93e^{-06}$	0,49	0,62

Para el noveno subintervalo tenemos el modelo

	Estimate	Std. Error	t value	Pr(> t)
$x_8.l_1$	$7,17e^{-03}$	$1,57e^{-03}$	4,58	$7,33e^{-06}$
$x_9.l_1$	$7,31e^{-01}$	$4,47e^{-02}$	16,35	6,82
trend	$2,04e^{-06}$	$7,34e^{-06}$	0,28	$7,81e^{-01}$

Ademas, la función autocovarianza para los residuales asociados a cada submodelo se muestran a continuación en la Figura 3.30

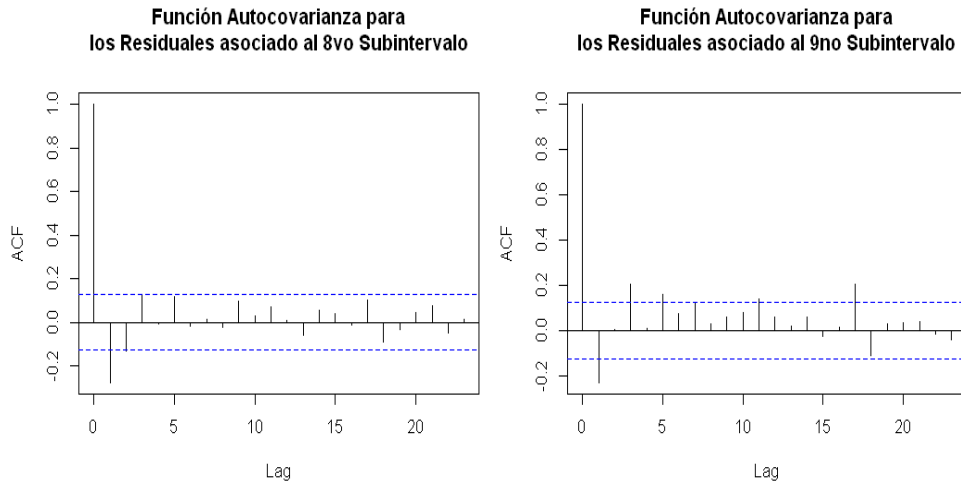


Figura 3.30: Función Autocovarianza para los residuales del octavo submodelo (izquierda) y noveno submodelo (derecha).

Para cada submodelo, los pronósticos se muestran respectivamente a continuación. Para el octavo subintervalo se muestra en la Figura 3.31

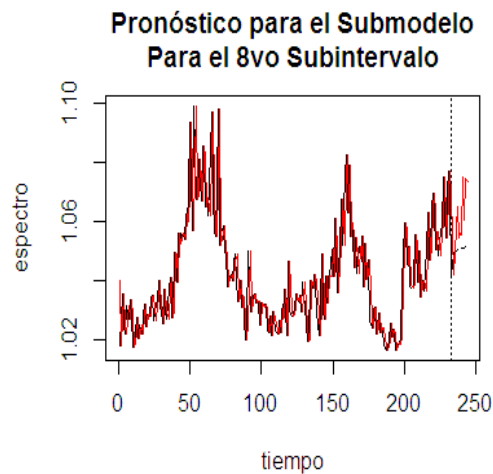


Figura 3.31: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

De igual forma, para el noveno subintervalo se evidencia la Figura 3.32

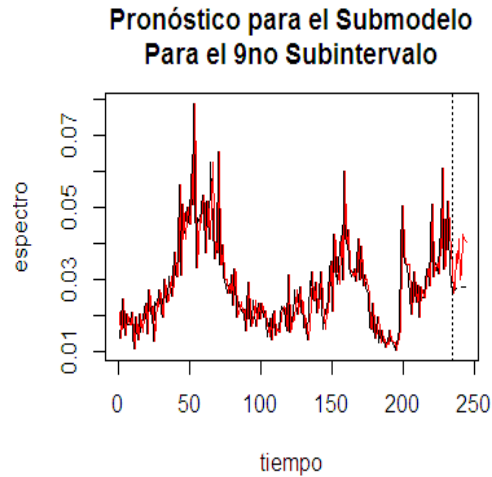


Figura 3.32: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.8. Modelo Asociado al Noveno y Decimo Subintervalo.

Para obtener el modelo vectorial asociado a los subintervalos octavo y noveno se obtuvo que la correlación entre estos subintervalos es de 0,92 gráficamente esta correlación se muestra en la Figura 3.33

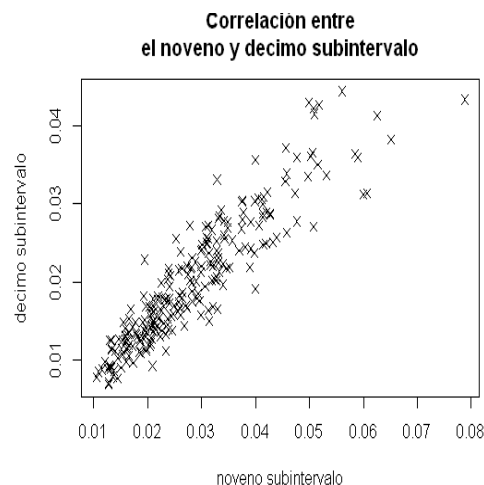


Figura 3.33: Correlación entre el noveno y decimo subintervalo.

El modelo AR para la combinación de estos dos subintervalos tiene orden 6. De nuevo, aplicar los test correspondientes respectivamente el test Phillips-Perron para explorar la existencia de raíces unitarias, concluimos que para cada subintervalo

tenemos que viendo los valores arrojados para los p valores se obtiene para el tercer subintervalo se obtiene el 1 %, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1% de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, de donde se obtiene que:

- $ARMA(5, 0)$ el AIC es $-1915,281$.
- $ARMA(1, 1)$ el AIC es $-1919,475$.
- $ARMA(1, 0)$ el AIC es $-1913,615$.

De igual forma se obtuvo que el modelo autorregresivo asociado a la combinación de estos subintervalos es de orden uno, los coeficientes del modelo son

$$\begin{matrix} & x_9 & x_{10} \\ x_9 & 0,2 & 0,86 \\ x_{10} & 0,12 & 0,61 \end{matrix}$$

x_{10} se refiere al decimo subintervalo. Se puede entonces afirmar la estacionaridad del modelo ya que las raices del polinomio característico asociado a este modelo son 2,83 y 17,67 las cuales superan la unidad, por lo que se puede decir que el modelo es estacionario. Para el modelo VAR asociado a estos se tiene que para el noveno subintervalo

	Estimate	Std. Error	t value	$Pr(> t)$
$x_9.l_1$	$1,95e^{-01}$	$1,04e^{-01}$	1,88	$6,18e^{-02}$
$x_{10}.l_1$	1,00	$1,45e^{-01}$	6,93	$3,9e^{-11}$
trend	$1,92e^{-05}$	$5,89e^{-06}$	3,26	$1,28e^{-03}$

Para el decimo subintervalo tenemos el modelo

	Estimate	Std. Error	t value	$Pr(> t)$
$x_9.l_1$	$1,32e^{-01}$	$7,46e^{-02}$	1,76	$7,91e^{-02}$
$x_{10}.l_1$	$7,31e^{-01}$	$1,04e^{-01}$	7,03	$2,16e^{-11}$
trend	$1,04e^{-05}$	$4,23e^{-06}$	2,47	$1,43e^{-02}$

Ademas, la función autocovarianza para los residuales asociados a cada submodelo se muestran a continuación en la Figura 3.34

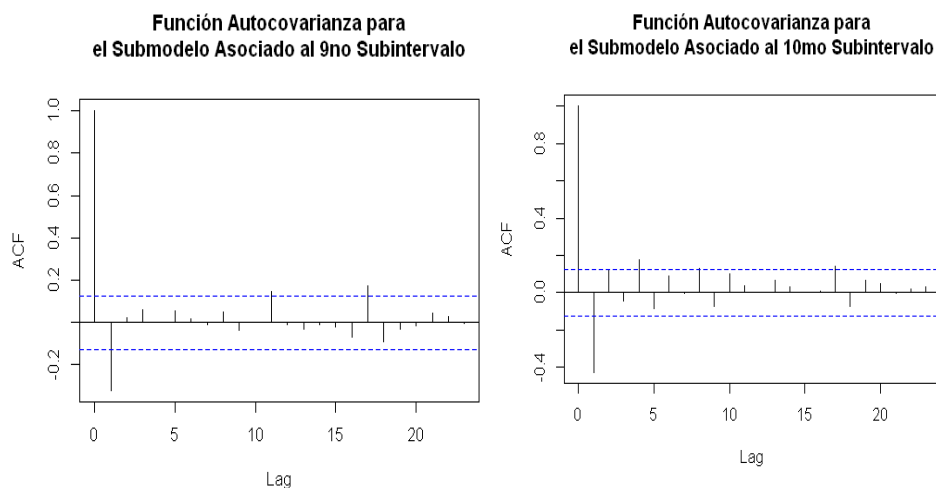


Figura 3.34: Función Autocovarianza para los residuales del octavo submodelo (izquierda) y noveno submodelo (derecha).

Para cada submodelo, los pronósticos se muestran respectivamente a continuación. Para el noveno subintervalo se muestra en la Figura 3.35

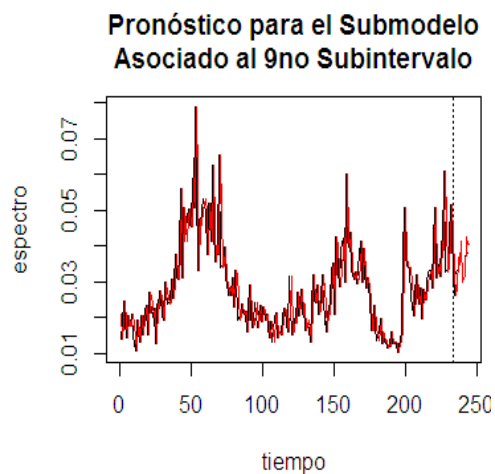


Figura 3.35: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

De igual forma, para el decimo subintervalo se evidencia

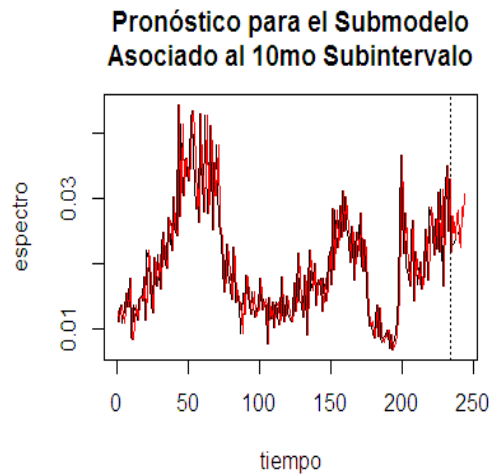


Figura 3.36: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

3.9. Modelo Asociado al Decimo Primer y Decimo Segundo Subintervalo.

Para observar el modelo vectorial que resulta de la combinación de los noveno y decimo subintervalo, se tiene que la correlación que resulta de estos dos subintervalos es 0,90, graficamente se puede observar en la Figura 3.37

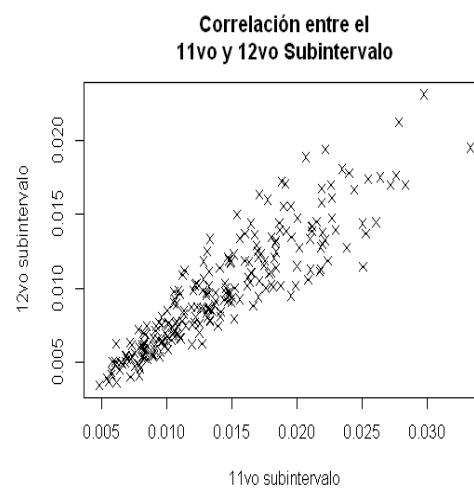


Figura 3.37: Correlación entre el decimo primer y decimo segundo subintervalo.

El modelo AR para la combinación de estos dos subintervalos tiene orden 4. De nuevo, aplicar los test correspondientes respectivamente el test Phillips-Perron para

explorar la existencia de raíces unitarias, concluimos que para cada subintervalo tenemos que viendo los valores arrojados para los p valores se obtiene para el tercer subintervalo se obtiene el 1 %, por lo que se puede afirmar que no existe la posibilidad de tener raíces unitarias. También, aplicando el test de Phillips-Ouliaris tenemos que el 1 % de los p -valores nos permiten rechazar la hipótesis nula de que estos dos subintervalos no están cointegrados.

El test aplicado en la parte anterior, nos da pie para usar modelos de regresión, y a estos ajustar un modelo ARMA para los residuales de este modelo de regresión, con esto calcular el AIC para distintos ordenes del modelo ARMA calculado, de donde se obtiene que:

- $ARMA(1, 0)$ el AIC es $-2247,71$.
- $ARMA(2, 0)$ el AIC es $-2248,127$.
- $ARMA(1, 1)$ el AIC es $-2257,094$.

Podemos afirmar que el modelo es autorregresivo de orden uno. Los coeficientes de este modelo son

$$\begin{array}{cc} & x_{11} & x_{12} \\ x_{11} & 0,56 & 0,33 \\ x_{12} & 0,35 & 0,25 \end{array}$$

siendo x_{11} el decimo primer y x_{12} el decimo segundo subintervalo. Al calcular las raíces del polinomio característico asociado a los coeficientes del modelo son 1,27 y 39,22 lo que significa que el modelo es estacionario. Los coeficientes para el modelo VAR asociado a estos subintervalos, respectovamente para cada subintervalo son, para el decimo primer subintervalo

	Estimate	Std. Error	t value	$Pr(> t)$
$x_{11}.l_1$	$6,15e^{-01}$	$9,91e^{-02}$	6,21	$2,36e^{-09}$
$x_{12}.l_1$	$4,47e^{-01}$	$1,45e^{-01}$	3,09	$2,23e^{-03}$
trend	$6,67e^{-06}$	$2,85e^{-06}$	2,34	$2,00e^{-02}$

Para el decimo segundo subintervalo tenemos el modelo

	Estimate	Std. Error	t value	Pr(> t)
$x_{11}.l_1$	$3,92e^{-01}$	$7,10e^{-02}$	5,52	$8,75e^{-08}$
$x_{12}.l_1$	$3,40e^{-01}$	$1,04e^{-01}$	3,28	$1,20e^{-03}$
trend	$5,69e^{-06}$	$2,04e^{-06}$	2,78	$5,82e^{-03}$

La función de autocovarianza asociada a los residuales para cada submodelo se aprecian en la Figura 3.38

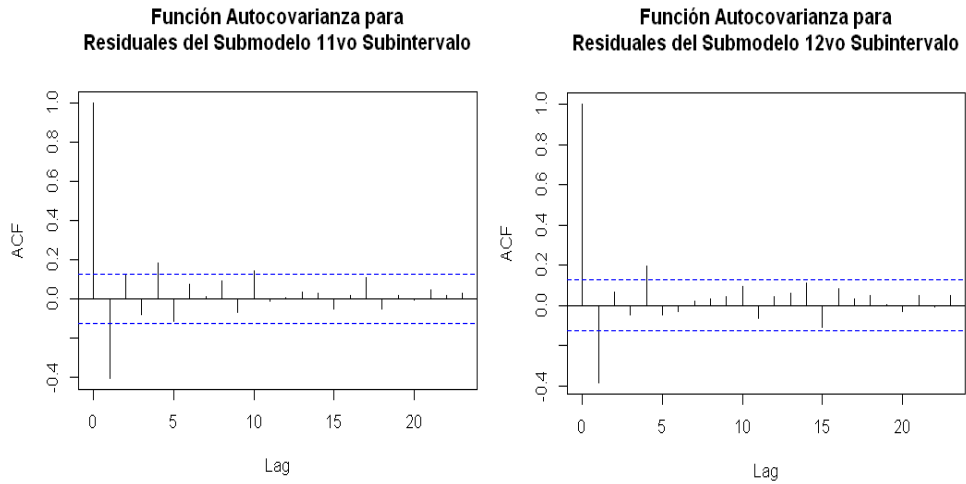


Figura 3.38: Función Autocovarianza para los residuales del decimo primer submodelo (izquierda) y decimo segundo submodelo (derecha).

Los pronósticos se muestran respectivamente a continuación. Para el decimo primer subintervalo se tiene la Figura 3.39

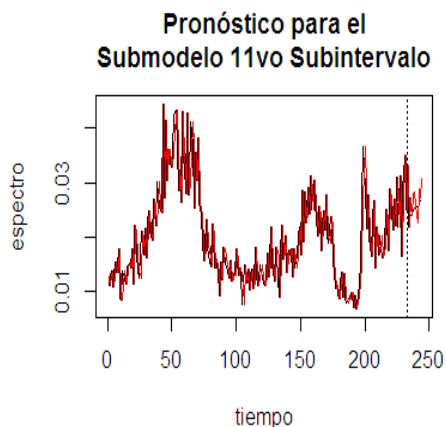


Figura 3.39: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

Para el decimo segundo subintervalo, observamos a continuación

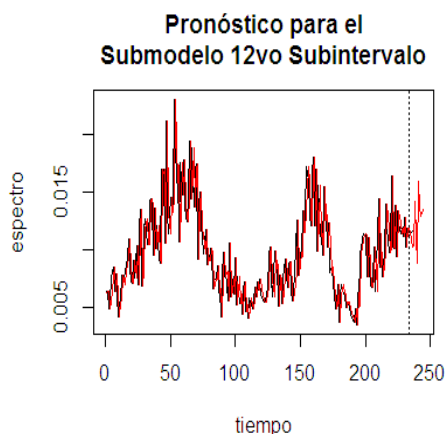


Figura 3.40: subintervalo (rojo), modelo (negro), pronóstico (punteadas).

Es así como hemos encontrado los modelos vectoriales regresivos (VAR) para los subintervalos en estudio, como ya se dijo antes estos modelos fueron hallados dos a dos para subintervalos contiguos siguiendo la metodología de [3] donde se sugiere aplicar algunos test a los datos en estudio y poder asegurar las hipótesis necesarias para que los modelos hallados sean adecuados a los datos en estudio, además poder obtener los pronósticos y comparar con los datos originales, lo cual resultó acertado en el presente estudio, dado que en la mayoría de los casos se pudo hallar el modelo

VAR con sus condiciones necesarias de existencia (cointegración, estabilidad, buen comportamiento de la ACF y pronósticos). Sin embargo, dos subintervalos contiguos para los que no se pudo encontrar el desarrollo del modelo VAR fueron el decimo y decimo primer subintervalo, quienes no lograron una cointegración necesaria, dado que al aplicar test de Phillips-Ouliaris como consecuencia tenemos que el modelo de regresión lineal que se construye según la metodología, tiene orden cero, lo cual dificulta continuar con el análisis, puesto que no se puede determinar un orden adecuado para este modelo según el valor de la AIC, así como tampoco se puede decidir sobre la construcción del modelo VAR que sigue de aplicar este test. Una causa de este hecho es la similitud que poseeb estos subintervalos, lo que trae como consecuencia que este modelo líneal no resulte tener un desarrollo.

CAPÍTULO 4

CONCLUSIONES

El presente análisis se realizó con el fin de estudiar el espectro de los 244 periodos obtenidos en el Mar del Norte en 1,999, para esto, se usó el software R con el que se logro ajustar modelos al comportamiento de la tormenta. De este análisis, se desprenden los siguientes resultados, Se graficaron los espectros calculados a partir de los datos de la tormenta, luego que se determinó el rango útil de frecuencias común a todos los espectros, se seleccionaron empíricamente valores a los extremos del rango de modo que no se observaron características importantes fuera de este rango, este rango forma el 97 % de la información total, la energía total presente fuera del rango sólo representa el 3 %. A este rango útil, se decidió dividir en subintervalos de forma tal que esta obedezca a capturar por cada subintervalo la mayor cantidad posible de cambios en la energía, esto con el objeto de explicar precisamente el comportamiento de estos cambios mediante los modelos que se obtienen luego. Como conclusión en esta parte se decidió dividirlo en 12 subintervalos de igual tamaño, seguido a esto, ajustar un modelo a cada uno de los intervalos.

A continuación, para cada uno de estos intervalos se obtuvo un modelo autorregresivo, que después de analizar el comportamiento de la ACF, PACF, los p -valores y el valor del AIC, más el estadístico de Ljung Box, se determinó que cada subintervalo tenía asociado un modelo $AR(1)$. Para algunos de estos subintervalos, fue necesario aplicar la transformación de Box-Cox, con el fin eliminar la no estacionaridad en la varianza y con esto lograr un mejor ajuste de los residuales asociado al modelo con la distribución normal. Este primer análisis, se realizó con la finalidad de obtener una primera impresión respecto al comportamiento de los datos en cuestión, debido a que el orden y tipo de modelo que resulta en esta parte, nos dará idea de lo que a continuación en los modelos VAR podríamos apreciar, sin embargo, este análisis no es suficiente para explicar el comportamiento de la tormenta, por lo que se recurre a un estudio más general de los subintervalos para dar una respuesta más acertada

sobre el comportamiento de los datos.

Seguidamente, para complementar el análisis, se construyeron los modelos vectoriales autorregresivos de forma tal que estos den información acerca del comportamiento de la tormenta de forma más precisa, puesto que este estudio aporta información sobre la relación que poseen las 12 variables en conjunto. Primero, se intentó construir un modelo vectorial con los 12 subintervalos simultáneamente, este modelo resultó difícil de manejar numéricamente, ya que por la ausencia de cointegración entre algunas de las series, el modelo en conjunto no se puede obtener. Es por esto que, se recurre a la metodología expuesta en [3] para construir vectores dos a dos para subintervalos contiguos. Estos subintervalos se suponen correlacionados debido a que provienen de la misma tormenta, para este análisis en particular la cointegración entre las series es fundamental, dado que esta aporta información sobre relación entre las variables, para obtener esta propiedad importante las series deben ser estacionarias (ya demostrada en la construcción de los modelos univariados). Esta propiedad, implica también que las series se mueven conjuntamente en el tiempo y las diferencias entre ellas son estables o estacionarias.

El caso antes expuesto, es el de la mayoría de las series bivariadas que hemos construido, sin embargo el único caso donde las condiciones para la cointegración no se cumplen, resultó ser para el decimo y decimo primer subintervalo, los cuales no mostraron evidencias de cointegración y por tanto no fue posible obtener un modelo para la combinación de estas series. A los modelos bivariados restantes, se les sometió a pruebas para la cointegración, también para probar su estabilidad, donde se evidenció que las raíces excedían la unidad por pequeñas pero significativas unidades, para algunos otros, las raíces excedían la unidad exageradamente. También, se graficaron las ACF para cada componente del modelo para obtener la certeza de que en efecto, tenían orden 1, en contraste con el valor del AIC que resultó ser el menor de los distintos ordenes a los cuales fueron construidos cada modelo bivariado.

Muy relevante en este estudio, fueron los pronósticos tanto para los modelos univariados como para los bivariados, dado que estos son los que aseguran que los modelos son capaces de reproducir los datos en estudio. La técnica usada fue la de truncar las series quitándoles 10 datos para saber si el modelo era capaz de reproducirlos, sobre el pronóstico obtenido en este trabajo se puede concluir que en algunos casos los modelos no se ajustan completamente a los datos ya que los

cambios bruscos en la energía que inicialmente se observan. Una de las causas de este problema, puede ser la velocidad del viento cuyos datos no se conocen, dado que la altura muchas depende de la cantidad de energía que puede extraerse del viento.

Es de resaltar, que estos modelos no son definitivos dado que, se deben mejorar los pronósticos y la heterocedasticidad que presentan los residuales, de forma tal que se pueda controlar la variabilidad pronunciada de la energía que se presenta al observar los residuos y también en los pronósticos.

APÉNDICE A

MODELOS UNIVARIADOS

En este capítulo, se incluye el material suplementario acerca de los modelos tratados en el Capítulo 2. Se muestran los subintervalos restantes para los modelos univariados, se mostrarán los coeficientes para los modelos, así como también las ACF, PACF, los gráficos para el comportamiento de los cuantiles, de diagnósticos y pronósticos para sustentar el presente análisis. Dividiremos el capítulo en dos secciones, los que requieren de la transformación de Box Cox, el segundo, quinto, sexto, séptimo y octavo en el orden; y los que no requieren esta transformación cuarto, noveno, decimo, decimo primer y segundo subintervalos.

A.1. Subintervalos que requieren de la Transformación de Box Cox.

A.1.1. Segundo Subintervalo.

El modelo tiene la siguiente expresión:

$$x_t = 0,01 + 0,94x_{t-1} + w_t$$

Para el cual cambiando los distintos ordenes del modelo y así escoger el que mejor ajusta a los datos en estudio de acuerdo a lo establecido en el Capítulo 2, se muestran por ejemplo los siguientes gráficos de diagnóstico para dos tipos de modelos un $AR(1)$ y $AR(2)$ (Figura A.1) así observar tanto el comportamiento de los residuales, la ACF, así como también el comportamiento de los p -valores quienes proporcionan información relevante en este análisis para decidir sobre los modelos, a continuación en la Figura A.1

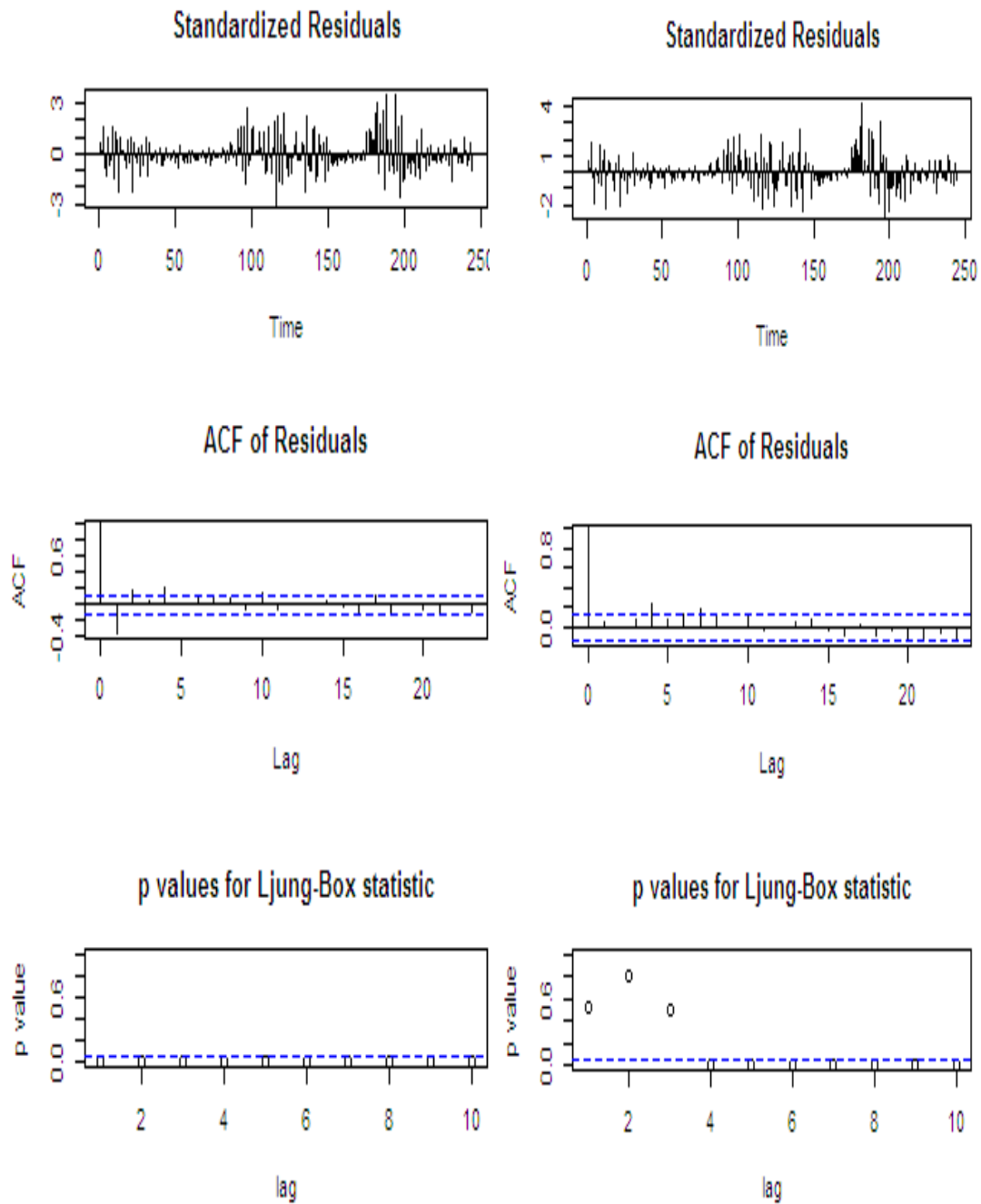


Figura A.1: Gráficas de diagnóstico. A izquierda, $AR(1)$; a derecha, $AR(2)$ para el segundo subintervalo.

Donde se obtuvo, para el modelo $AR(1)$ el AIC es $-1309,252$ mientras que para el $AR(2)$ el AIC obtenido fue $-1344,71$, así que se puede afirmar que el modelo apropiado según estos valores es $AR(1)$. Seguidamente se muestran las gráficas de cuantiles antes después de la transformación de Box Cox se puede apreciar en la Figura A.2.

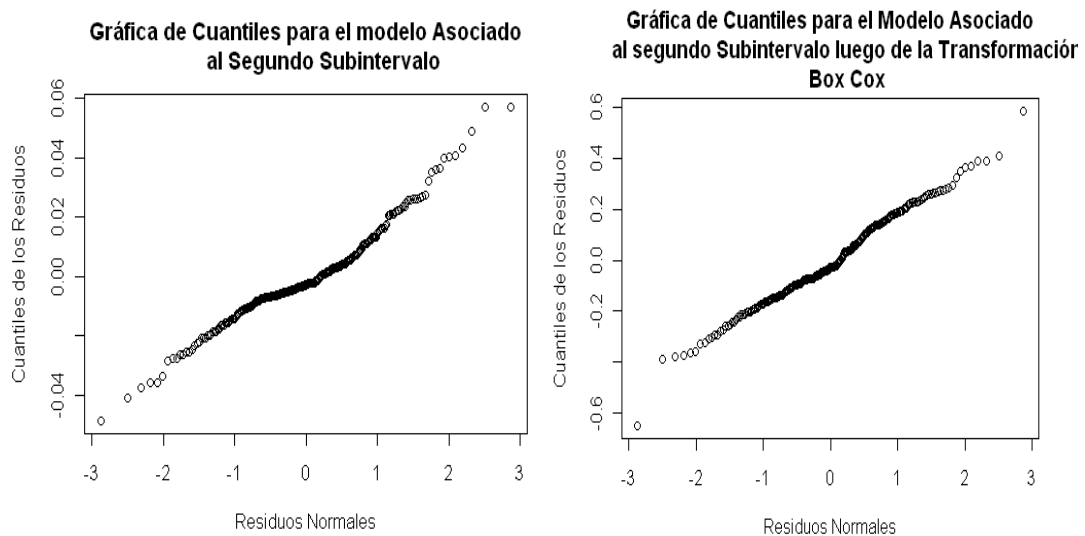


Figura A.2: Gráficas de cuantiles para el segundo subintervalo. A izquierda, los datos originales; a derecha, con la transformación de Box Cox para $\lambda = 0,08$.

Donde se puede observar notoriamente como luego de aplicar la transformación con el parámetro λ obtenido se obtiene la mejoría deseada en cuanto a la aproximación lineal de los residuales del modelo. A continuación, también se da evidencia del comportamiento de los pronósticos que como ya se dijo son parte importante en el análisis de los modelo ya que estos dan información sobre el buen comportamiento del mismo cuando son capaces de reproducir por sí mismos los datos en estudio, a continuación los pronósticos en la Figura A.3

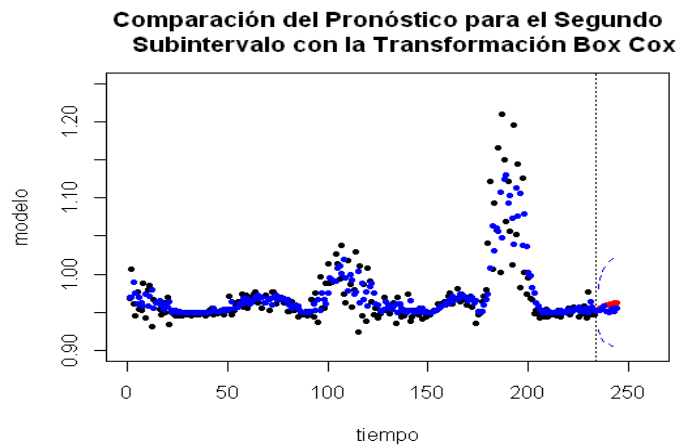


Figura A.3: Pronóstico para el modelo asociado al segundo subintervalo. Los datos del segundo subintervalo (en azul) frente al pronóstico para el modelo (negro).

A.1.2. Quinto Subintervalo.

El modelo está dado de la siguiente forma

$$x_t = 0,83 + 0,15x_{t-1} + w_t$$

Para el cual cambiando los distintos órdenes del modelo y así escoger el que mejor ajusta a los datos en estudio de acuerdo a lo establecido en el Capítulo 2, se muestran por ejemplo los siguientes gráficos de diagnóstico para dos tipos de modelos un $AR(1)$ y $AR(2)$ así observar tanto el comportamiento de los residuales, la ACF, así como también el comportamiento de los p -valores quienes proporcionan información relevante en este análisis para decidir sobre los modelos, los diagnósticos se pueden apreciar en la Figura A.4

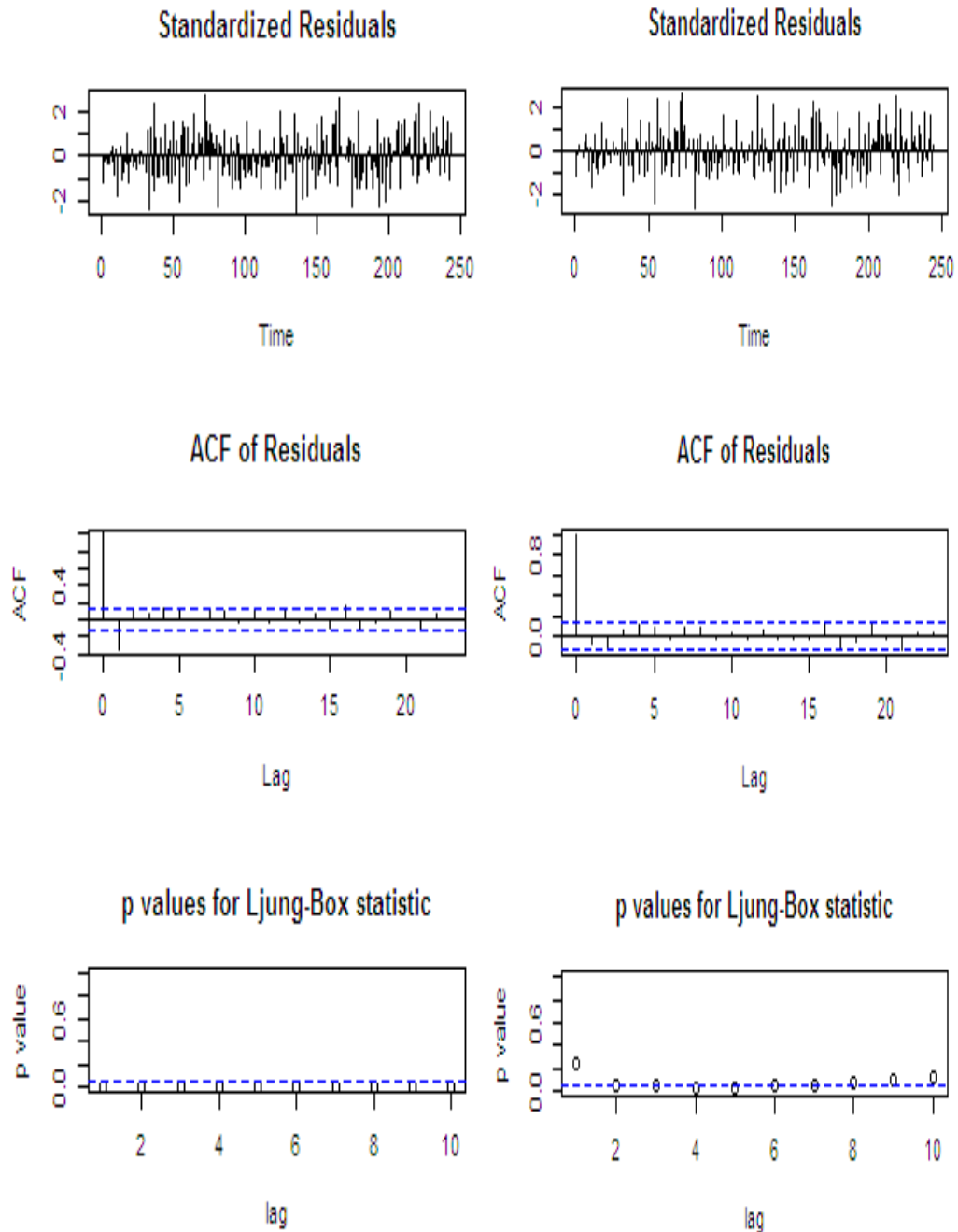


Figura A.4: Gráficas de diagnóstico para el modelos asociado al quinto subintervalo. A izquierda, $AR(1)$; a derecha, $AR(2)$.

Donde se obtuvo, para el modelo $AR(1)$ el AIC es $-1069,08$ mientras que para el $AR(2)$ el AIC obtenido fue $-1114,57$, así que se puede afirmar que el modelo apropiado según estos valores es $AR(1)$. Seguidamente se muestran las gráficas de

cuantiles antes y después de la transformación de Box Cox, cuya evidencia gráfica se ve a continuación en la Figura A.5.

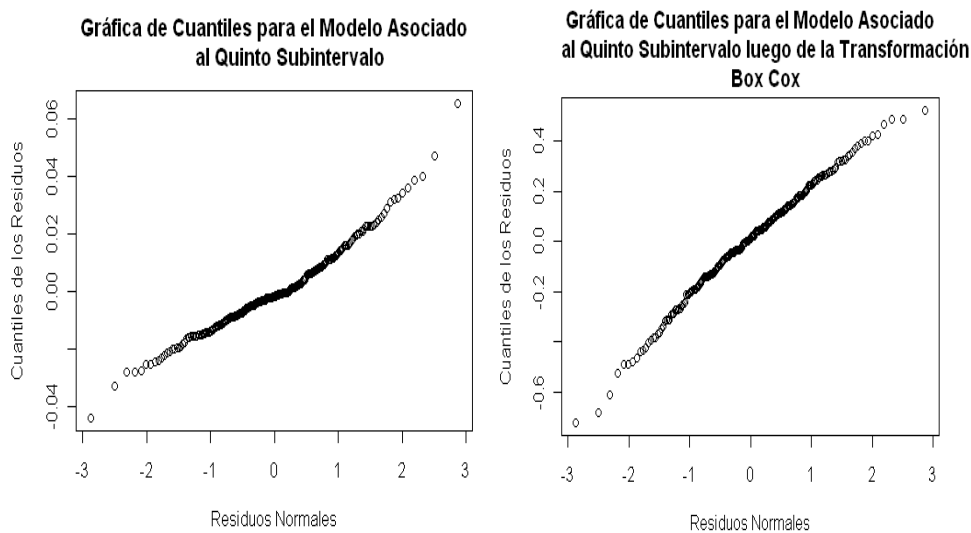


Figura A.5: Gráficas de cuantiles. A izquierda, los datos originales; a derecha, con la transformación de Box Cox para $\lambda = -0,07$.

Donde se puede observar notoriamente como luego de aplicar la transformación con el parámetro λ obtenido se obtiene la mejoría deseada en cuanto a la aproximación líneal de los residuales del modelo. A continuación, también se da evidencia del comportamiento de los pronósticos que como ya se dijo son parte importante en el análisis de los modelo ya que estos dan información sobre el buen comportamiento del mismo cuando son capaces de reproducir por sí mismos los datos en estudio, los pronósticos para el quinto subintervalo se observan en la Figura A.6

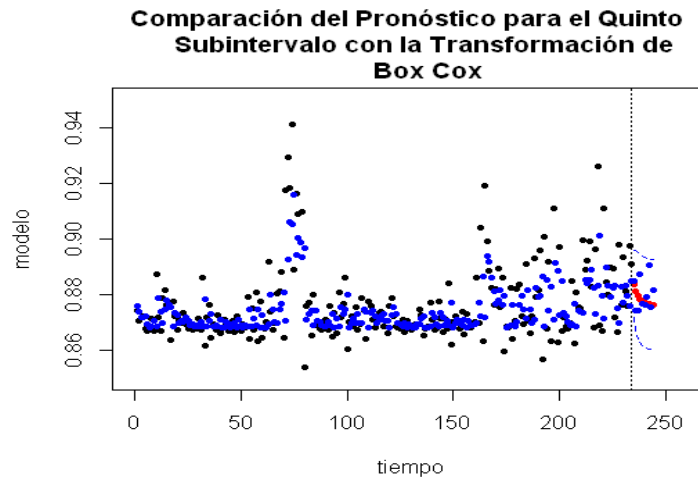


Figura A.6: Pronóstico para el modelo asociado al quinto subintervalo. El subintervalo (en azul) frente al pronóstico para el modelo (negro).

A.1.3. Sexto Subintervalo.

El modelo está dado de la siguiente forma

$$x_t = 0,80 + 0,0,1x_{t-1} + w_t$$

Para el cual cambiando los distintos órdenes del modelo y así escoger el que mejor ajusta a los datos en estudio de acuerdo a lo establecido en el Capítulo 2, se muestran por ejemplo los siguientes gráficos de diagnóstico para dos tipos de modelos un $AR(1)$ y $AR(2)$ así observar tanto el comportamiento de los residuales, la ACF, así como también el comportamiento de los p -valores quienes proporcionan información relevante en este análisis para decidir sobre los modelos, en seguida se observa la Figura A.7

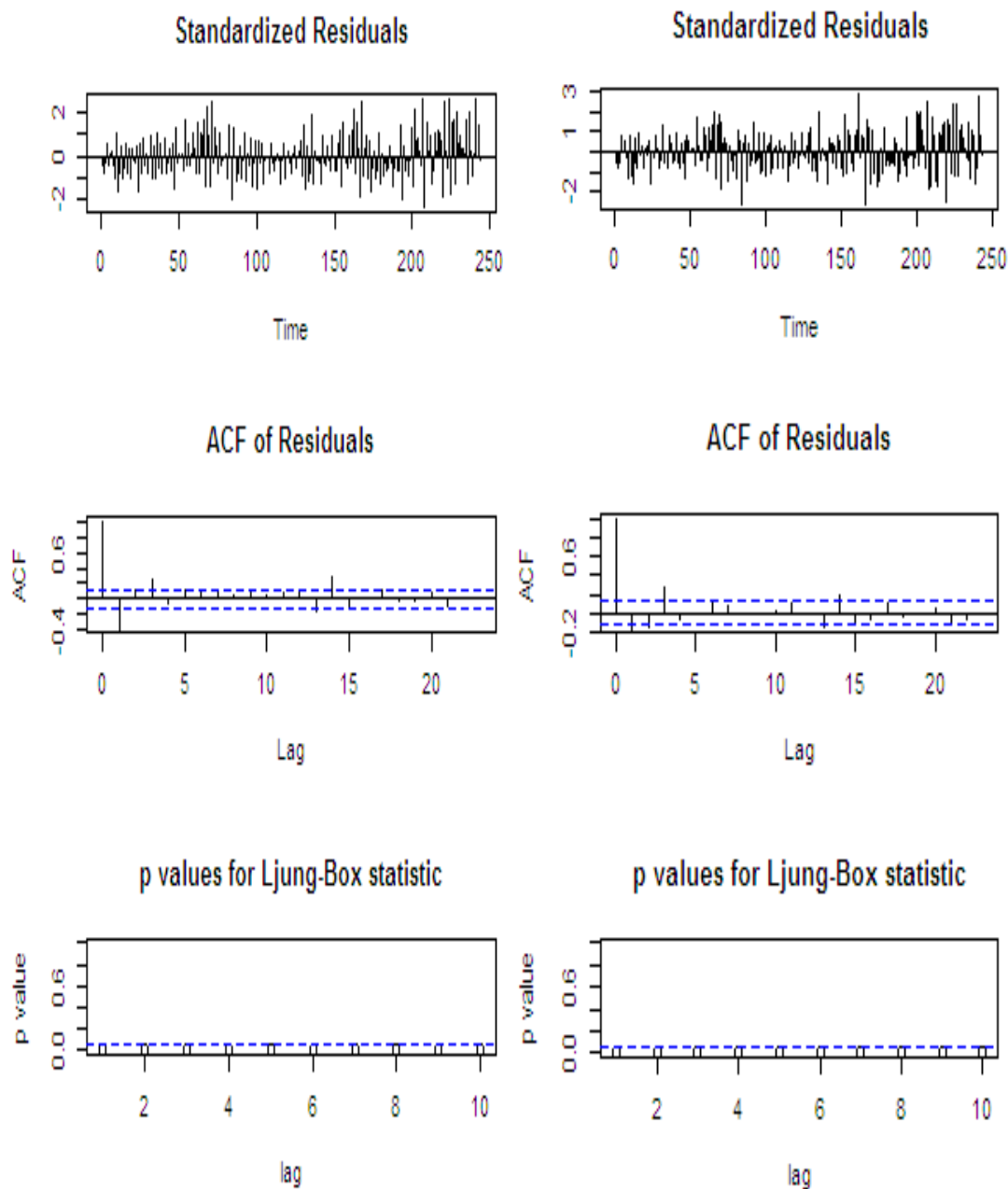


Figura A.7: Gráficas de diagnóstico para los modelos asociados al sexto subintervalo. A izquierda, $AR(1)$; a derecha, $AR(2)$.

Donde se obtuvo, para el modelo $AR(1)$ el AIC es $-1226,3$ mientras que para el $AR(2)$ el AIC obtenido fue $-1287,03$, así que se puede afirmar que el modelo apropiado según estos valores es $AR(1)$. Seguidamente se muestran las gráficas de cuantiles antes después de la transformación de Box Cox (ver Figura A.8).

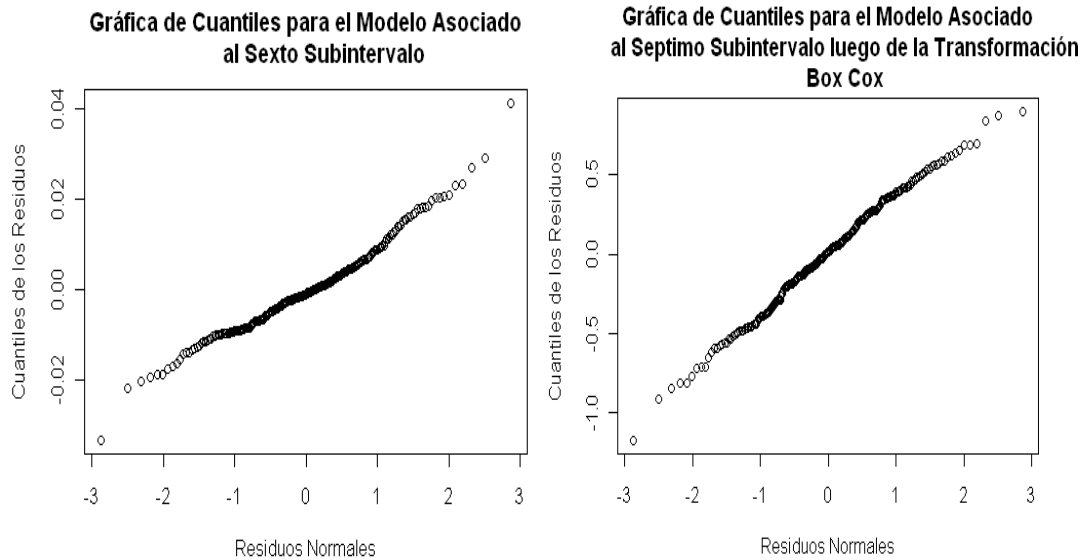


Figura A.8: Gráficas de cuantiles. A izquierda, los datos originales; a derecha, con la transformación de Box Cox para $\lambda = -0,22$.

Donde se puede observar notoriamente como luego de aplicar la transformación con el parámetro λ obtenido, se obtiene la mejoría deseada en cuanto a la aproximación lineal que requieren los residuales del modelo. A continuación, también se da evidencia del comportamiento de los pronósticos que como ya se dijo, son parte importante en el análisis de los modelo ya que estos dan información sobre el buen comportamiento del mismo cuando son capaces de reproducir por sí mismos los datos en estudio. Para los pronósticos asociados al modelo asociado al sexto subintervalo ver la Figura A.9

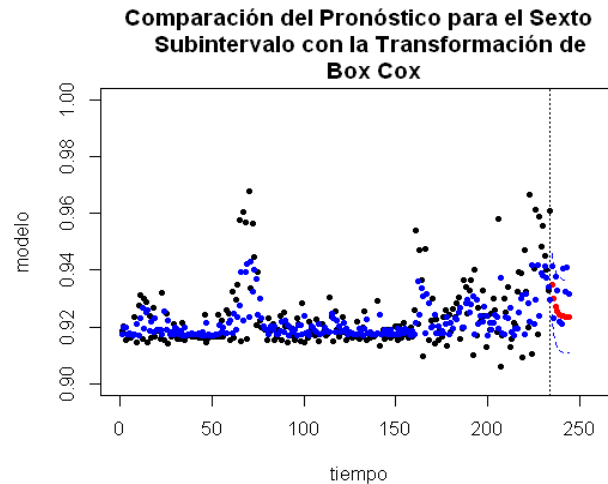


Figura A.9: .Pronóstico para el modelo asociado al sexto subintervalo.El subintervalo (en azul) frente al pronostico para el modelo(negro).

A.1.4. Séptimo Subintervalo.

El modelo esta dado de la siguiente forma

$$x_t = 0,79 + 0,1x_{t-1} + w_t$$

Para el cual cambiando los distintos ordenes del modelo y así escoger el que mejor ajusta a los datos en estudio de acuerdo a lo establecido en el Capítulo 2, se muestran por ejemplo los siguientes gráficos de diagnóstico para dos tipos de modelos un $AR(1)$ y $AR(2)$ así observar tanto el comportamiento de los residuales, la ACF, así como también el comportamiento de los p -valores quienes proporcionan información relevante en este análisis para decidir sobre los modelos. El diagnóstico a continuación (ver Figura A.10) compara dos modelos AR.

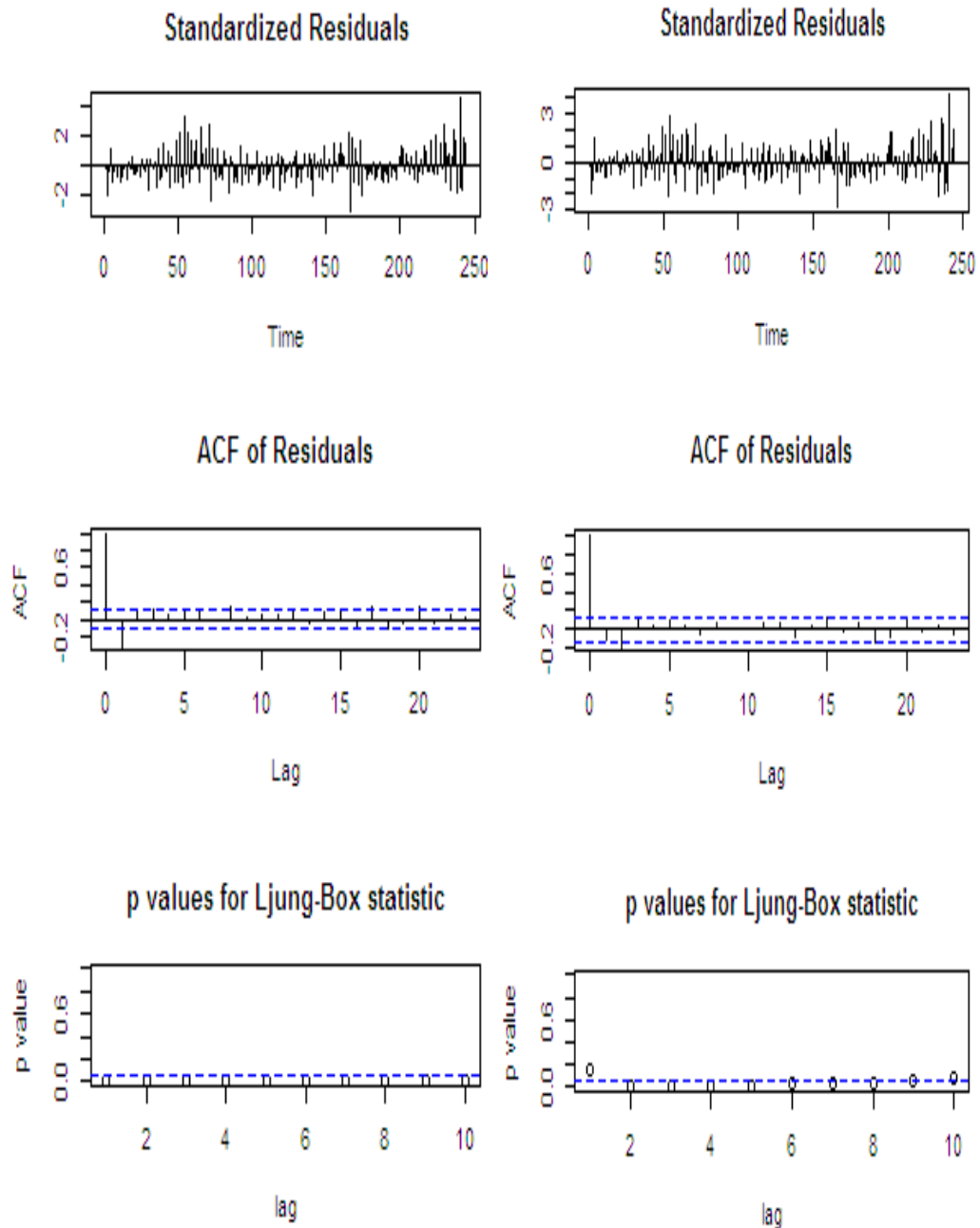


Figura A.10: Gráficas de diagnóstico. A izquierda, $AR(1)$; a derecha, $AR(2)$.

Donde se obtuvo, para el modelo $AR(1)$ el AIC es $-1364,17$ mientras que para el $AR(2)$ el AIC obtenido fue $-1406,41$, así que se puede afirmar que el modelo apropiado según estos valores es $AR(1)$. Seguidamente se muestran las gráficas de cuantiles antes después de la transformación de Box-Cox. A continuación las gráficas

de cuantiles para el séptimo subintervalo (ver Figura A.10)

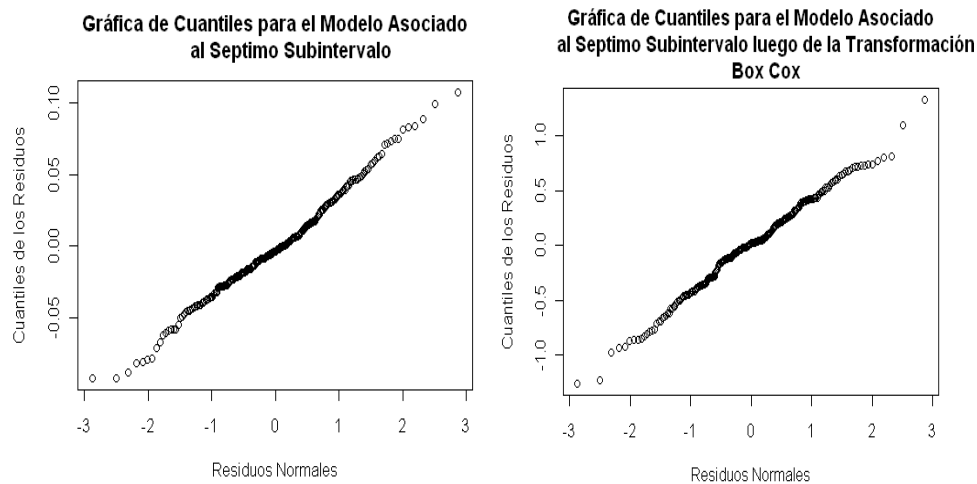


Figura A.11: Gráficas de cuantiles para el modelo asociado al séptimo subintervalo. A izquierda, los datos originales; a derecha, con la transformación de Box Cox para $\lambda = -0,23$.

Donde se puede observar notoriamente como luego de aplicar la transformación con el parámetro λ obtenido se obtiene la mejoría deseada en cuanto a la aproximación lineal de los residuales del modelo. A continuación, también se da evidencia del comportamiento de los pronósticos que como ya se dijo son parte importante en el análisis de los modelo ya que estos dan información sobre el buen comportamiento del mismo cuando son capaces de reproducir por sí mismos los datos en estudio (ver Figura A.12)

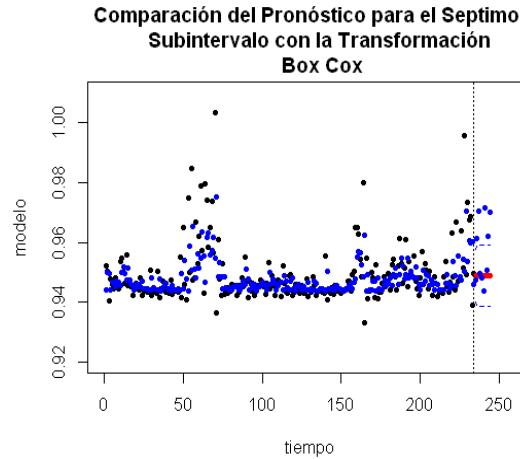


Figura A.12: Pronóstico para el modelo asociado al séptimo subintervalo. El subintervalo (en azul) frente al pronóstico para el modelo (negro).

A.1.5. Octavo Subintervalo.

El modelo está dado de la siguiente forma

$$x_t = 0,81 + 0,04x_{t-1} + w_t$$

Para el cual cambiando los distintos órdenes del modelo y así escoger el que mejor ajusta a los datos en estudio de acuerdo a lo establecido en el Capítulo 2, se muestran por ejemplo los siguientes gráficos de diagnóstico para dos tipos de modelos un $AR(1)$ y $AR(2)$ así observar tanto el comportamiento de los residuales, la ACF, así como también el comportamiento de los p -valores quienes proporcionan información relevante en este análisis para decidir sobre los modelos, la Figura A.13 compara dos modelos AR con distintos órdenes para este subintervalo.

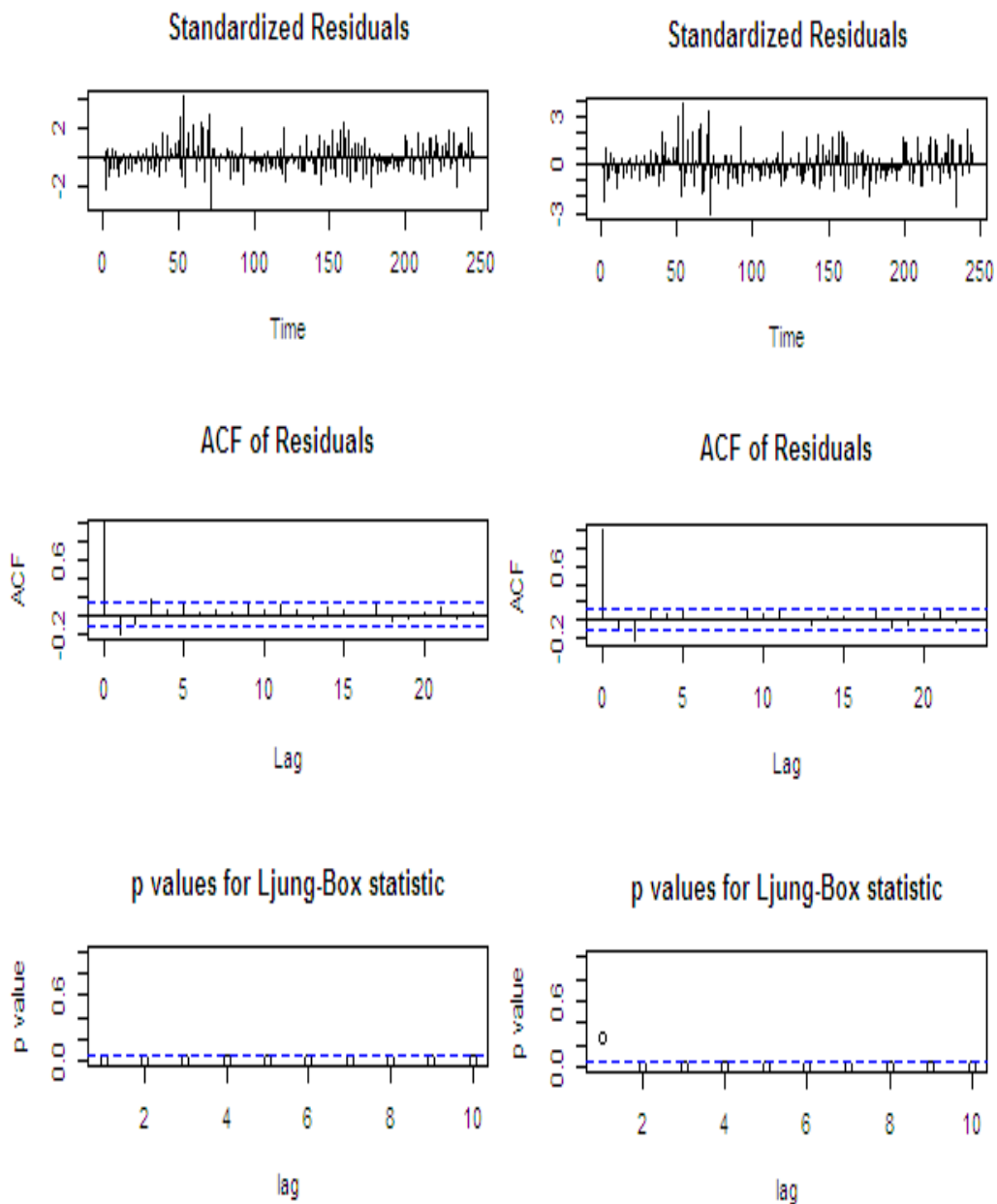


Figura A.13: Gráficas de diagnóstico para los modelos asociados el octavo subintervalo. A izquierda, $AR(1)$; a derecha, $AR(2)$.

Donde se obtuvo, para el modelo $AR(1)$ el AIC es $-1554,13$ mientras que para el $AR(2)$ el AIC obtenido fue $-1569,6$, así que se puede afirmar que el modelo apropiado según estos valores es $AR(1)$. Seguidamente se muestran las gráficas de cuantiles antes después de la transformación de Box Cox (ver Figura A.14).

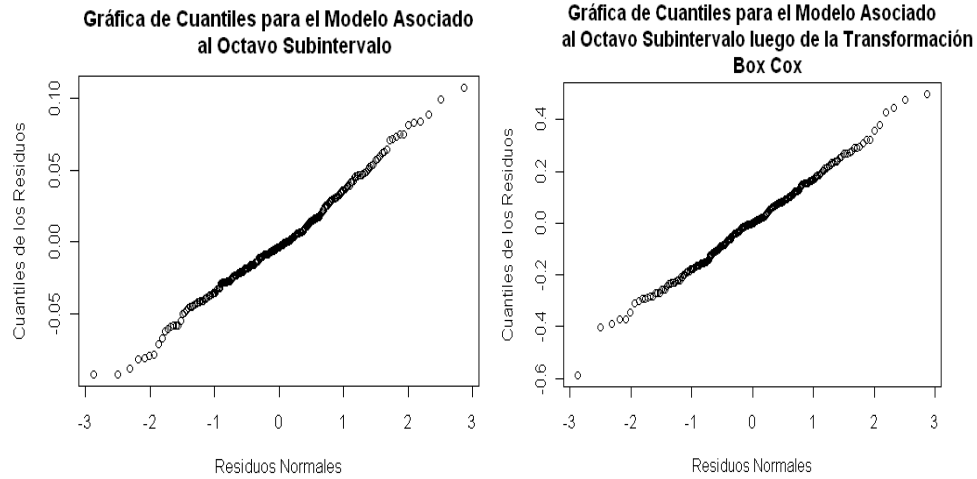


Figura A.14: Gráficas de cuantiles. A izquierda, los datos originales; a derecha, con la transformación de Box Cox para $\lambda = -0,23$.

Donde se puede observar notoriamente como luego de aplicar la transformación con el parámetro λ obtenido se obtiene la mejoría deseada en cuanto a la aproximación lineal de los residuales del modelo. A continuación, también se da evidencia del comportamiento de los pronósticos que como ya se dijo son parte importante en el análisis de los modelo ya que estos dan información sobre el buen comportamiento del mismo cuando son capaces de reproducir por sí mismos los datos en estudio, la Figura A.15 muestra los pronósticos para el $AR(1)$ asociado a este subintervalo.

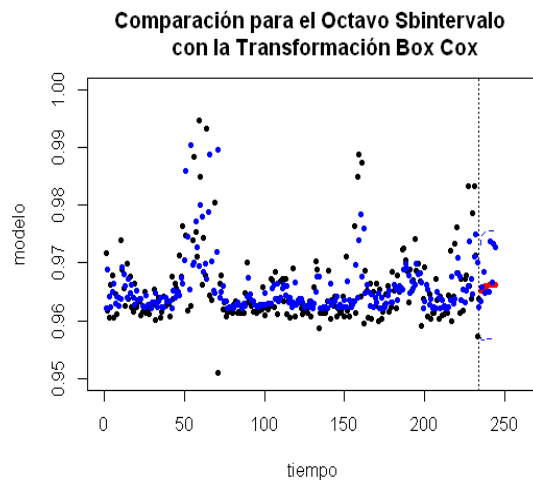


Figura A.15: Pronóstico para el octavo subintervalo. El subintervalo (en azul) frente al pronóstico para el modelo (negro).

A.2. Subintervalos que no Requieren la Transformación Box Cox.

En esta sección se muestran los subintervalos donde al calcular el parámetro λ que define la transformación de Box Cox se aproxima a la unidad, lo que teóricamente quiere decir que dicha transformación no es necesaria aplicarse a los datos en estudio.

Los subintervalos que poseen esta característica en el presente estudio son el tercero (estudiado en el Capítulo 1) cuarto, noveno, decimo, decimo primer y segundo subintervalo. A continuación se da un breve resumen de los resultados obtenidos para estos subintervalos, de forma tal que se complete el estudio en cuestión de los modelos univariados.

A.2.1. Cuarto Subintervalo.

El modelo tiene la expresión

$$x_t = 0,42 + 0,71x_{t-1} + \omega_t$$

Se realizaron varios cambios en el orden de los modelos ARMA, como ejemplo se muestra un $AR(1)$ y un $AR(2)$ para los cuales se muestran los comportamientos de la ACF de los residuales normalizados, como ya se mencionó en los capítulos anteriores, estos diagnósticos darán fuerza a la decisión de escoger un modelo adecuado para ajustar los datos, a continuación la Figura A.16

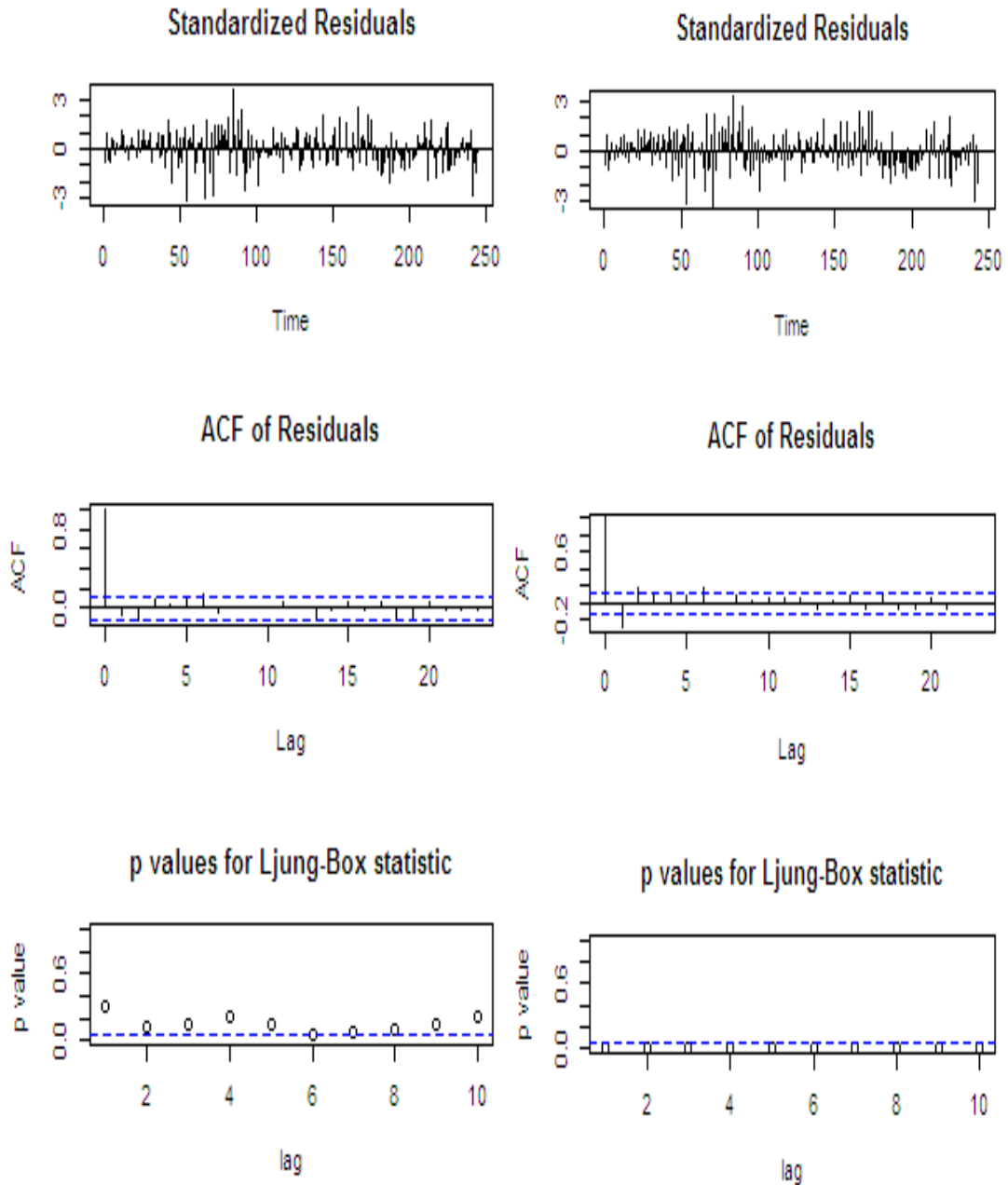


Figura A.16: Diagnóstico para el cuarto subintervalo. Izquierda: ARMA(2,0), derecha: ARMA(1,0)

Para estos modelos se obtuvo que para el $AR(1)$ el AIC es $-1008,91$ y para el $AR(2)$ este valor está dado por $-1049,92$ por lo que análogamente los otros casos podemos afirmar y aceptar que el modelo que mejor puede ajustar los datos es el $AR(1)$. Además, la Figura A.17 muestra la gráfica de cuantiles para los residuos

asociados a este modelo.

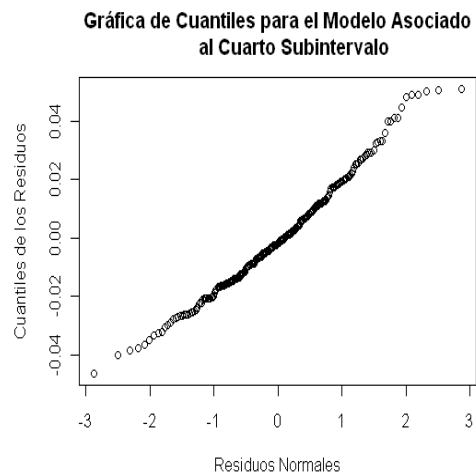


Figura A.17: Gráficas de Cuantiles para el cuarto subintervalo.

Los pronósticos que apoyan la elección de este modelo, los cuales reproducen de manera *optima* los datos, se muestran a continuación en la Figura A.18

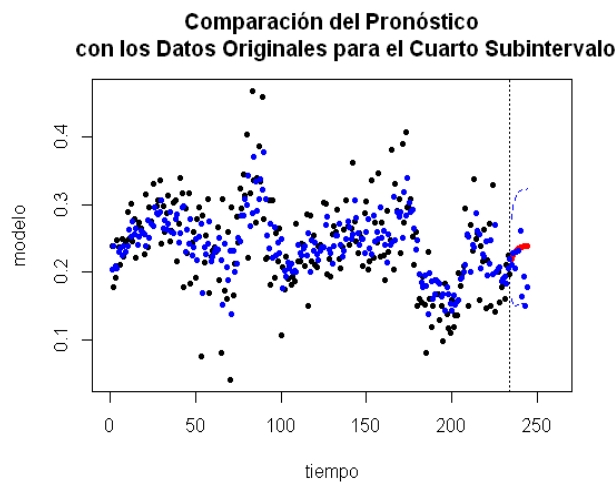


Figura A.18: El subintervalo (en azul) frente al pronóstico para el modelo(negro).

A.2.2. Noveno Subintervalo.

El modelo está dado por

$$x_t = 0,74 + 0,03x_{t-1} + \omega_t$$

Se realizaron varios cambios en el orden de los modelos ARMA, como ejemplo se muestra un $AR(1)$ y un $AR(2)$ para los cuales se muestran los comportamientos de la ACF de los residuales normalizados, como ya se mencionó en los capítulos anteriores, estos diagnósticos darán fuerza a la decisión de escoger un modelo adecuado para ajustar los datos (ver Figura A.19)

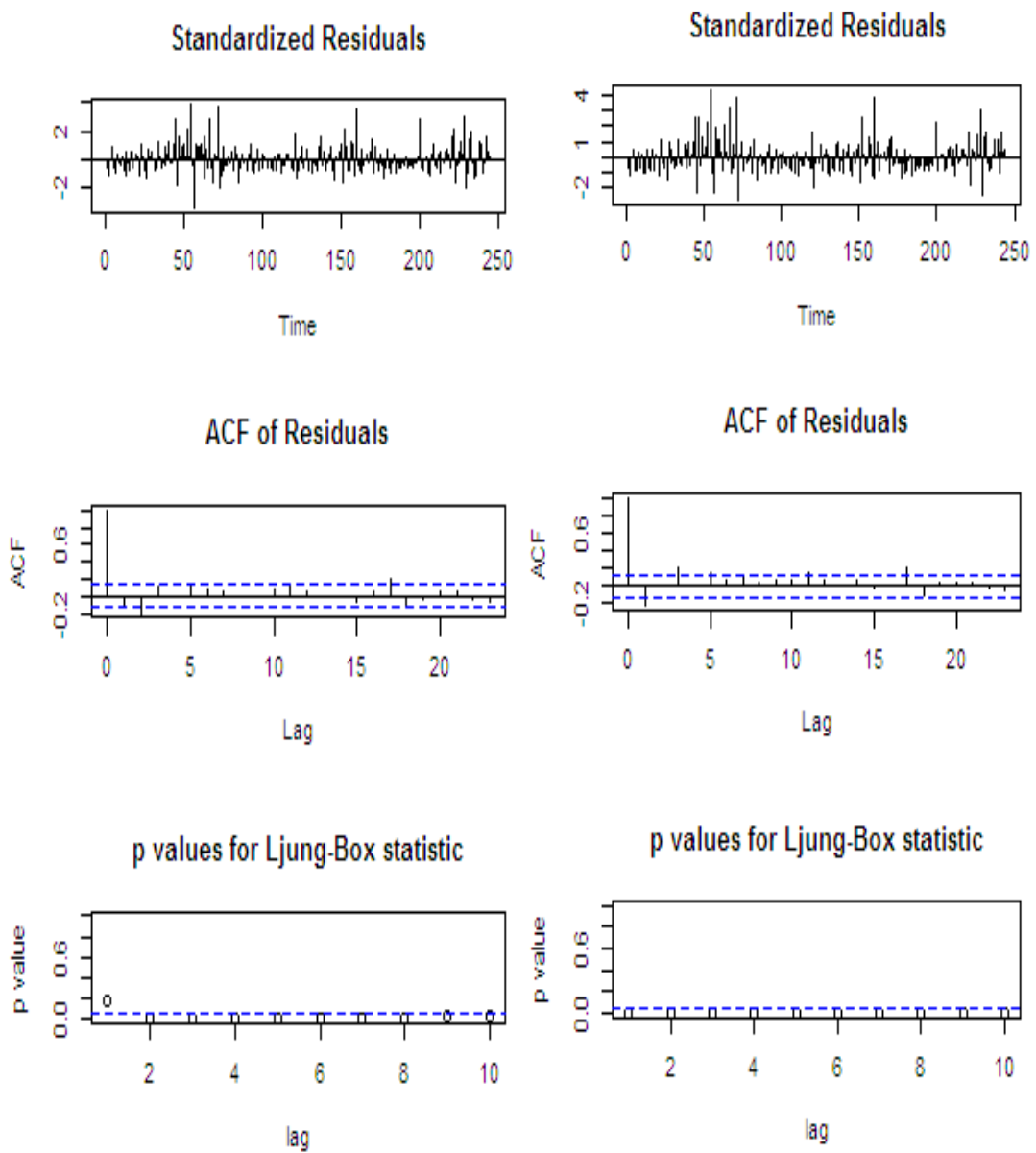


Figura A.19: Diagnóstico para el cuarto subintervalo. Izquierda: ARMA(2,0), derecha: ARMA(1,0)

Para estos modelos se obtuvo que para el $AR(1)$ el AIC es $-1659,83$ y para el $AR(2)$ este valor está dado por $-1682,86$ por lo que análogamente los otros casos podemos afirmar y aceptar que el modelo que mejor puede ajustar los datos es el $AR(1)$. Además, la Figura A.20 muestra la gráfica de cuantiles para los residuos asociados a este modelo.

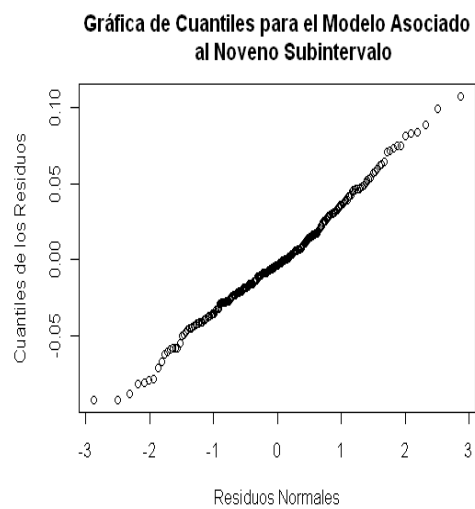


Figura A.20: Gráficas de Cuantiles para el tercer subintervalo.

Los pronósticos que apoyan la elección de este modelo, los cuales reproducen de manera *optima* los datos, se muestran a continuación en la Figura A.21

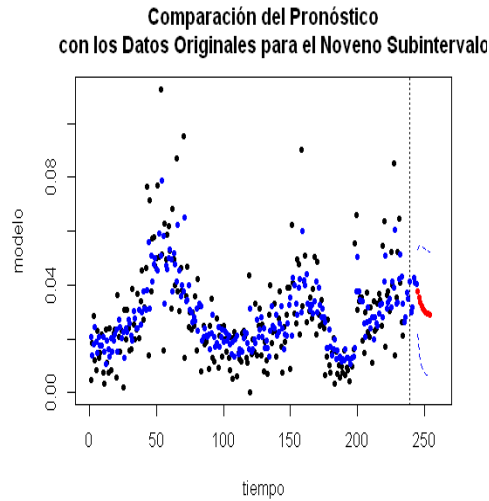


Figura A.21: Pronóstico para el modelo asociado al noveno subintervalo. El subintervalo (en azul) frente al pronóstico para el modelo (negro).

A.2.3. Decimo Subintervalo.

El modelo está dado por

$$x_t = 0,74 + 0,03x_{t-1} + \omega_t$$

Se realizaron varios cambios en el orden de los modelos ARMA, como ejemplo se muestra un $AR(1)$ y un $AR(2)$ para los cuales se muestran los comportamientos de la ACF de los residuales normalizados, como ya se mencionó en los capítulos anteriores, estos diagnósticos darán fuerza a la decisión de escoger un modelo adecuado para ajustar los datos

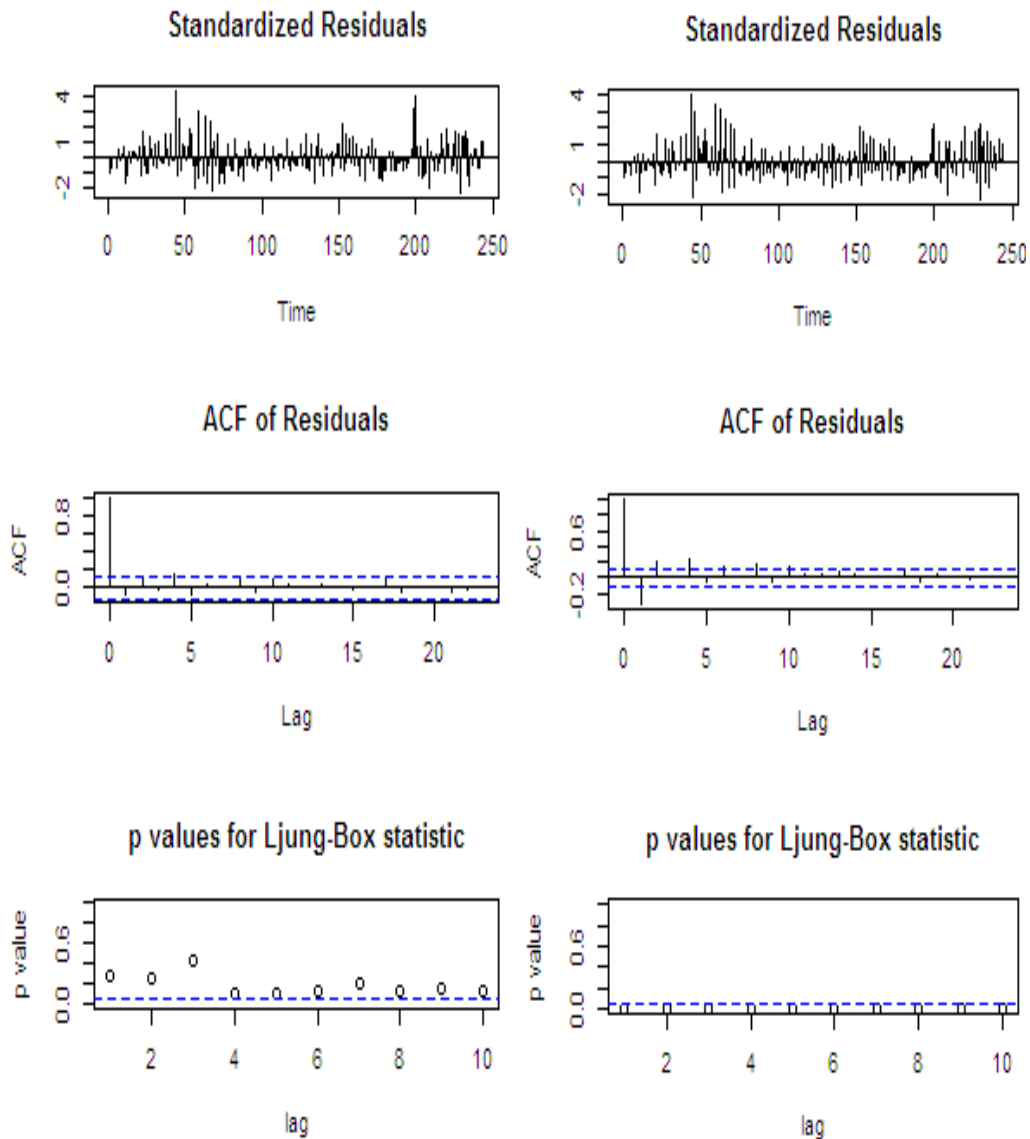


Figura A.22: Diagnóstico para el cuarto subintervalo. Izquierda: ARMA(2,0), derecha: ARMA(1,0)

Para estos modelos se obtuvo que para el $AR(1)$ el AIC es $-1659,83$ y para el $AR(2)$ este valor está dado por $-1682,86$ por lo que análogamente los otros casos podemos afirmar y aceptar que el modelo que mejor puede ajustar los datos es el $AR(1)$. Además, la gráfica de cuantiles para los residuos asociados a este modelo son

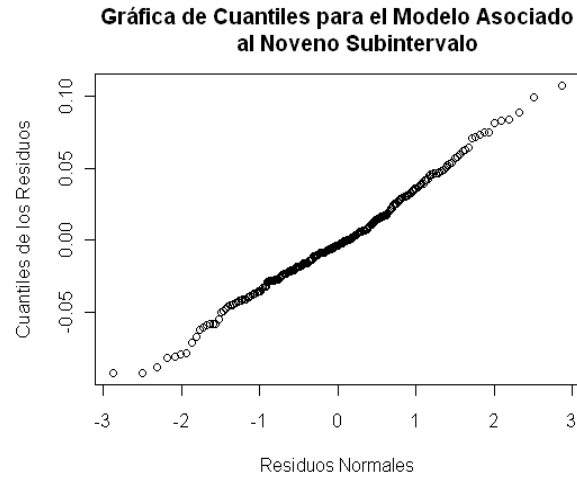


Figura A.23: Gráficas de Cuantiles para el tercer subintervalo.

Los pronósticos que apoyan la elección de este modelo, los cuales reproducen de manera *óptima* los datos, se muestran a continuación

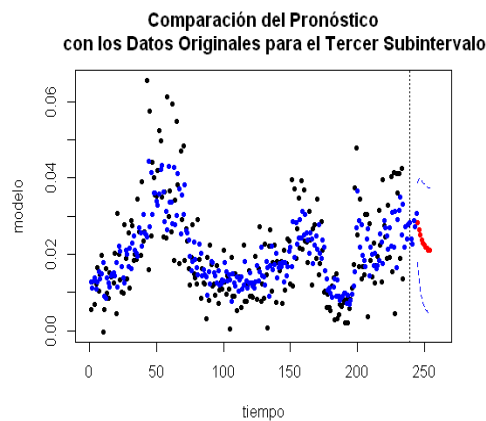


Figura A.24: El subintervalo (en azul) frente al pronóstico para el modelo(negro).

A.2.4. Decimo Primer Subintervalo.

El modelo está dado por

$$x_t = 0,76 + 0,01x_{t-1} + \omega_t$$

Se realizaron varios cambios en el orden de los modelos ARMA, como ejemplo se muestra un $AR(1)$ y un $AR(2)$ para los cuales se muestran los comportamientos de la

ACF de los residuales normalizados, como ya se menciona en los capítulos anteriores, estos diagnósticos darán fuerza a la decisión de escoger un modelo adecuado para ajustar los datos, a continuación ver la Figura A.25 para comparar dos modelos

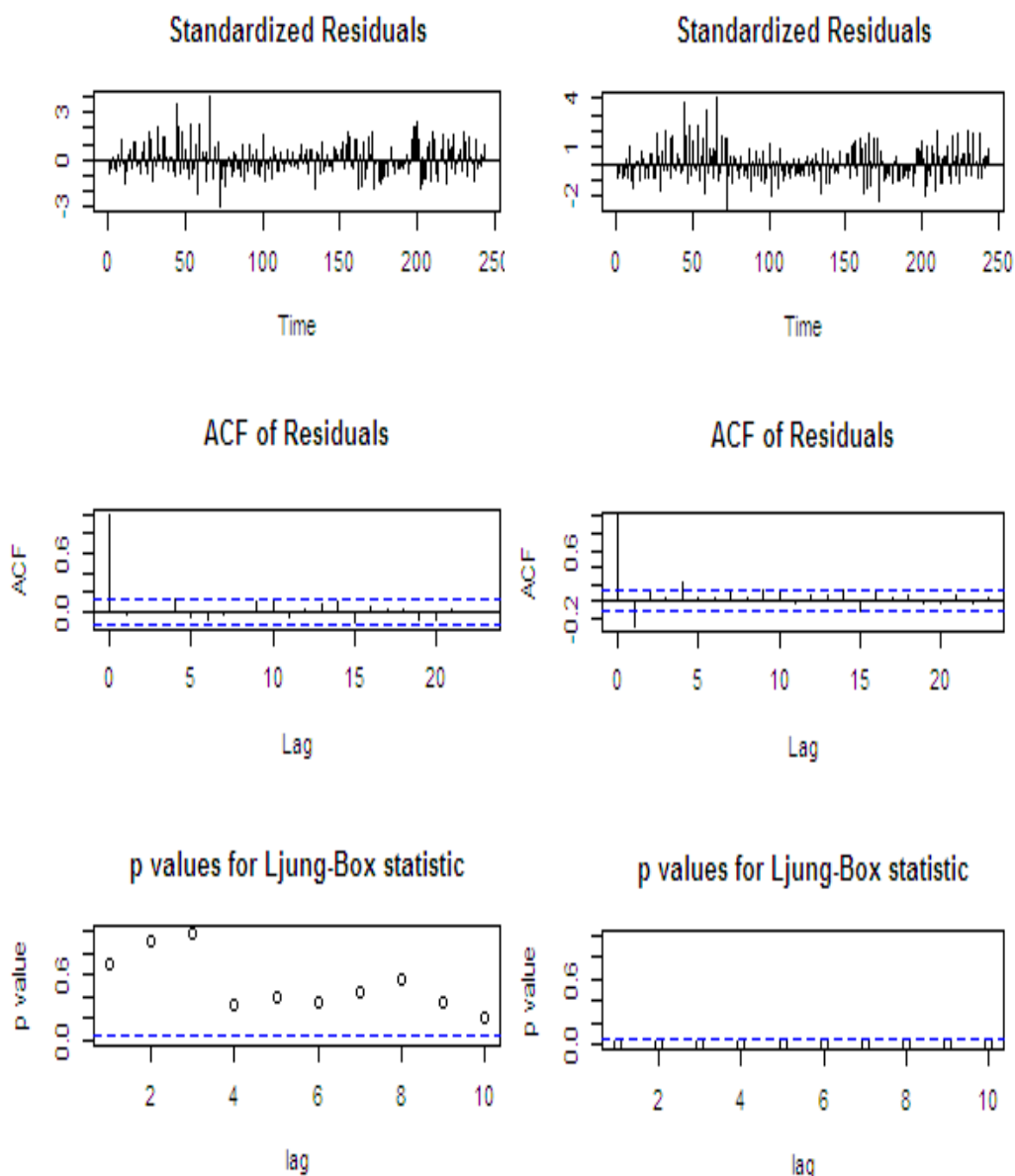


Figura A.25: Diagnóstico para el cuarto subintervalo. Izquierda: ARMA(2,0), derecha: ARMA(1,0)

Para estos modelos se obtuvo que para el $AR(1)$ el AIC es $-2039,82$ y para el $AR(2)$ este valor está dado por $-2077,1$ por lo que análogamente los otros casos podemos afirmar y aceptar que el modelo que mejor puede ajustar los datos es el

$AR(1)$. Además, la gráfica de cuantiles para los residuos asociados a este modelo se pueden ver en la Figura A.26

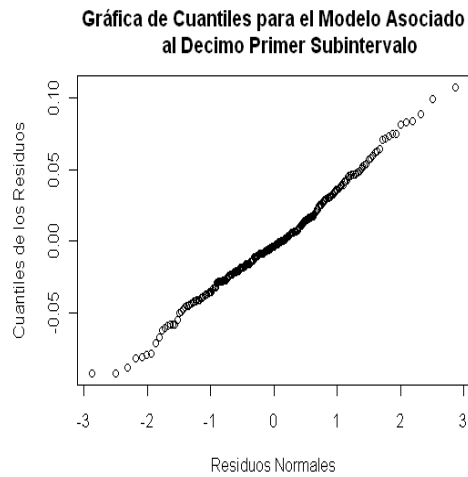


Figura A.26: Gráficas de Cuantiles para el tercer subintervalo.

Los pronósticos que apoyan la elección de este modelo, los cuales reproducen de manera *optima* los datos, se muestran a continuación en la Figura A.27

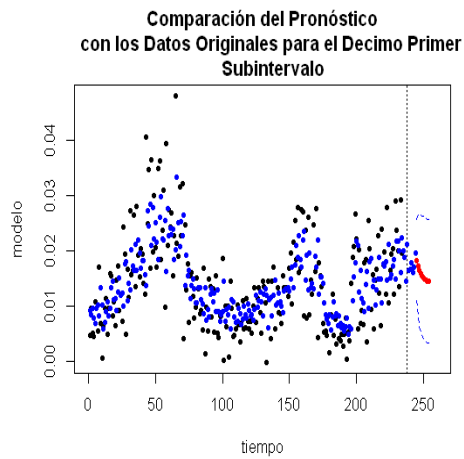


Figura A.27: El subintervalo (en azul) frente al pronóstico para el modelo(negro).

A.2.5. Decimo Segundo Subintervalo.

El modelo está dado por

$$x_t = 0,71 + 0,01x_{t-1} + \omega_t$$

Se realizaron varios cambios en el orden de los modelos ARMA, como ejemplo se muestra un $AR(1)$ y un $AR(2)$ para los cuales se muestran los comportamientos de la ACF de los residuales normalizados, como ya se mencionó en los capítulos anteriores, estos diagnósticos darán fuerza a la decisión de escoger un modelo adecuado para ajustar los datos, a continuación se aprecian en la Figura A.28

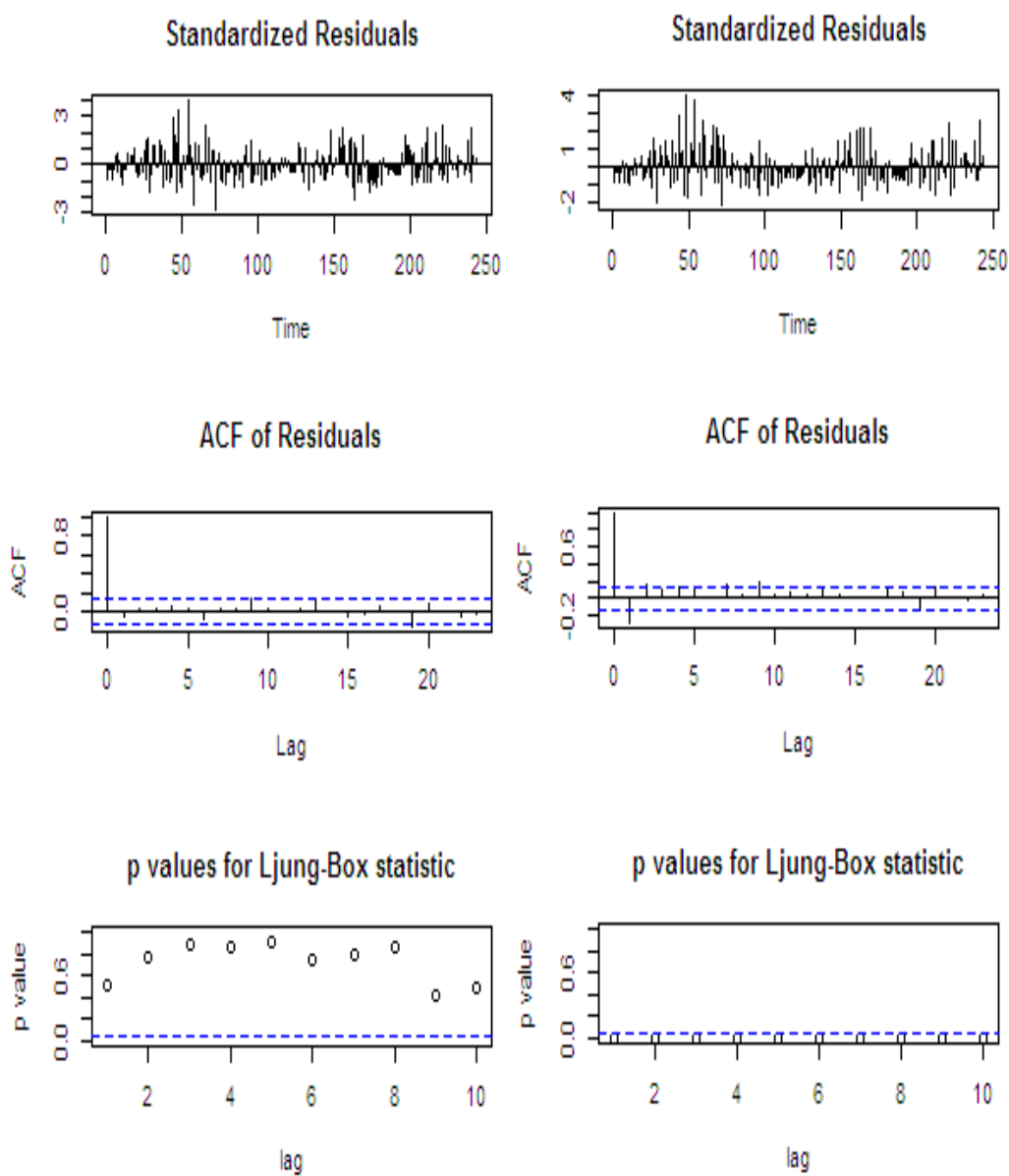


Figura A.28: Diagnóstico para el cuarto subintervalo. Izquierda: ARMA(2,0), derecha: ARMA(1,0)

Para estos modelos se obtuvo que para el $AR(1)$ el AIC es $-2188,12$ y para el

$AR(2)$ este valor esta dado por $-2227,92$ por lo que análogamente los otros casos podemos afirmar y aceptar que el modelos que mejor puede ajustar los datos es el $AR(1)$. Además, la gráfica de cuantiles para los residuos asociados a este modelo se aprecian en la Figura A.29

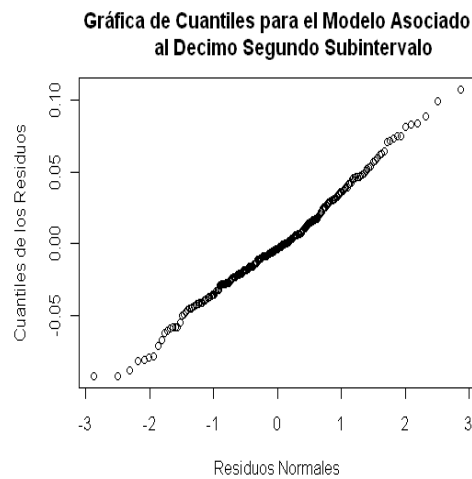


Figura A.29: Gráficas de Cuantiles para el tercer subintervalo.

Los pronósticos que apoyan la elección de este modelo, los cuales reproducen de manera *optima* los datos, se muestran a continuación en la Figura A.30

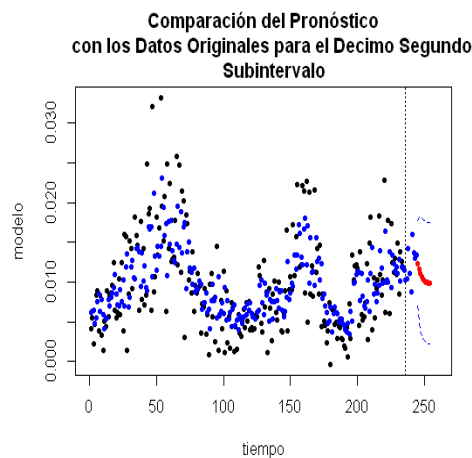


Figura A.30: El subintervalo (en azul) frente al pronóstico para el modelo(negro).

REFERENCIAS

- [1] ROBERT H. SHUMWAY Y DAVID S. STOFFER, *Time Series Analysis and Its Applications With R Examples*, Springer (2006).
- [2] PETER J. BROCKWELL Y RICHARD A. DAVIS, *Time Introduction to Time Series and Forecasting*, Springer (2002).
- [3] PAUL S.P. COWPERTWAIT Y ANDREW V. METCALFE, *Introductory Time Series Analysis With R*, Springer (2008).
- [4] MICHEL K. OCHI, *Ocean Waves. The Stochastic Approach*, Cambridge University Press (2005).
- [5] JOAQUÍN ORTEGA, *Estudio de Algunas Propiedades del Mar Usando Modelos Aleatorios*, Boletín de la Asociación Matemática Venezolana, Vol. VIII, N° 2(2001).
- [6] GEORG LINDGREN , *Lectures on Stationary Stochastic Processes. A course for PhD student in Mathematical Statistic and Other Fields*, Lund University. Centre for Mathematical Science. Mathematical Statistic, October 2006.