



Facultad de Ciencias  
Instituto de Estadística  
Ingeniería en Estadística

# DetECCIÓN DE PUNTOS DE CAMBIO EN SERIES DE TIEMPO UTILIZANDO ANÁLISIS WAVELET

DANIEL FABIÁN SOTO ZAPATA

11 de enero de 2022

## **Profesor Guía**

Alejandra Christen, Dra.

Instituto de Estadística, Universidad de Valparaíso

## **Profesor Co-Guía**

Daira Velandia, Dra.

Instituto de Estadística, Universidad de Valparaíso

## **Proyecto de titulación para optar al:**

grado académico de: *Licenciado en Estadística*

título profesional de: *Ingeniero en Estadística*

minor en: *Estadística Financiera*

# Resumen

Los puntos de cambio son instantes en el tiempo en donde se presentan variaciones bruscas en la media, en la varianza e incluso en los valores de los coeficientes de un modelo de regresión. Es de mucha importancia detectar rápidamente estas alteraciones a tiempo, ya que son capaces de provocar errores en las estimaciones y, posteriormente, en las predicciones de los modelos. La detección de puntos de cambio en series temporales se ha desarrollado ampliamente a través de distintos métodos, tales como, la detección de múltiples puntos de quiebre propuesta por Bai y Perron, el test de Chow, el test CUSUM y el test de MOSUM. Sin embargo, los estudios relacionados con la detección de puntos de cambio utilizando el análisis wavelet son poco recurrentes. Aun así, el análisis wavelet posee un gran potencial para la detección de puntos de cambio en series temporales, capaz de encontrar cambios donde los otros métodos no pueden o no son tomadas en cuenta por su baja perturbación dentro de una serie temporal. En esta investigación se evalúa la efectividad que tienen los métodos más conocidos en series temporales y el análisis wavelet al momento de detectar los puntos de cambio utilizando simulaciones. Posteriormente, se aplican todas las metodologías clásicas y el análisis wavelet a distintos conjuntos de datos para probar su efectividad en la detección de puntos de cambio.

# Algunas Palabras

Le doy agradecimientos a mis profesores que me han guiado desde el comienzo de mi carrera y me han formado como profesional hasta llegar a este momento, sobre todo al profesor Carlos Henríquez, su famosa manera de combinar la estadística con el humor lo hacen un profesor muy carismático y con sus enseñanzas nos ha mostrado el potencial que tenemos como profesionales estadísticos del mañana. También agradezco a mis amistades que he forjado en la carrera, a Valentina Escobar con quien comencé en este camino de conocimientos compartiendo nuestros triunfos y apoyándonos en los momentos difíciles, a Nicolás Vidal, un buen amigo que demostró ser una persona muy confiable, con mucho talento en los deportes y en las estadísticas.

Agradezco a mis amigos cercanos, Felipe Rojas y Francisca Rojas, que me han impulsado para seguir mejorando ante las dificultades de la vida y ojala pueda seguir compartiendo mis logros durante muchos años.

Agradezco a mi familia por estar a mi lado constantemente viendo mi crecimiento como profesional desde mi comienzo en la carrera entregando consejos y apoyo ante todo los desafíos que he logrado cumplir, y por sobre todo, a mi abuela quien me ha criado y me ha acompañado en todas mis etapas mostrándome hasta dónde puede llegar mis capacidades y mucho más si me esfuerzo mucho para lograrlo.

# Índice general

<b>Resumen</b>	<b>2</b>
<b>Algunas Palabras</b>	<b>3</b>
<b>1. Introducción</b>	<b>6</b>
1.1. Detección de puntos de cambio	7
1.2. Test CUSUM	8
1.3. Test MOSUM	12
1.3.1. Test de fluctuación generalizado	13
1.3.2. Proceso de fluctuación empírico	13
1.4. Test de Chow	16
1.5. Test de múltiples puntos de quiebre	18
1.6. Estado del arte	20
1.7. Planteamiento del problema	22
1.8. Objetivos	22
1.8.1. Objetivo general	22
1.8.2. Objetivos específicos	22
1.9. Hipótesis	23
<b>2. Marco teórico</b>	<b>24</b>
2.1. Series de tiempo	24
2.2. Componentes de las series de tiempo	24
2.3. Descomposición clásica de series de tiempo	25
2.4. Procesos Estocásticos	25
2.5. Procesos estacionarios	25
2.5.1. Modelos autoregresivos (AR) de orden $p$	27
2.5.2. Modelo de medias móviles (MA) de orden $p$	28
2.5.3. Modelos ARMA de orden $(p,q)$	29
2.5.4. Procesos de memoria larga ARIMA	30
2.5.5. Diferenciación en estacionalidad	30
2.5.6. Modelo multiplicativo estacional	30
2.6. Modelos heterocedásticos condicionales	30
2.6.1. Modelo ARCH( $p$ )	31
2.6.2. Modelo GARCH( $p,q$ )	31

<b>3. Wavelet</b>	<b>32</b>
3.1. Teoría Wavelet . . . . .	33
3.2. Análisis Multiresolución . . . . .	35
3.3. Familia wavelet . . . . .	36
3.4. Modelo de detección de puntos de cambio utilizando wavelets . . . . .	40
3.4.1. Identificación del valor crítico aproximado del test y la estimación de discontinuidades y cúspides . . . . .	42
<b>4. Simulaciones</b>	<b>43</b>
4.1. Resultados de las simulaciones . . . . .	46
<b>5. Aplicaciones</b>	<b>48</b>
5.1. Flujo anual del caudal del río Nilo . . . . .	48
5.2. Tasa de desempleo en Chile . . . . .	52
5.3. IPC en Chile . . . . .	57
5.4. Seeing . . . . .	60
<b>6. Conclusiones</b>	<b>67</b>
<b>A. Anexos</b>	<b>69</b>
<b>Referencias</b>	<b>73</b>

# Capítulo 1

## Introducción

Los puntos de cambio se definen como variaciones bruscas en la media, la varianza y de los valores de los coeficientes de un modelo de regresión a lo largo del tiempo (Valera, 2020). Recientemente, ha aumentado el interés por la detección de puntos de cambio en distintas áreas de las ciencias. Ignorar esos puntos de cambio puede llevar fácilmente a cometer errores en el proceso de análisis de series temporales con resultados inesperados y conclusiones incorrectas (Su, Zhu, Liu, y Hu, 2013). Siendo uno de los desafíos más grandes para analistas e investigadores. En muchas ocasiones, los analistas se encuentran en la dificultad de detectar la cantidad de puntos de cambio y las ubicaciones exactas de ellos. Los registros que se tienen ponen en evidencia que los modelos de puntos de cambio solo han desarrollado su potencial en el estudio de las distribuciones gaussianas. A causa de ello, a mediados de los años 80 su enfoque fue más allá de los modelos tradicionales y se impulsó el desarrollo de modelos con base en test de hipótesis no paramétricos (Valera, 2020). En áreas científicas aplicadas también se ha encontrado el potencial de la detección de puntos de cambio. En medicina se utilizan para estudiar las frecuencias ECG (Electrocardiograma) y EEG (Electroencefalograma), en meteorología para detectar cambios en la temperatura o la velocidad del viento, en *Machine Learning* para el análisis de voz e imágenes y en psicología para el análisis de las conductas cotidianas de las personas, entre otros (Aminikhanghahi y Cook, 2016).

Respecto al desempeño en la detección de puntos de cambio, en ocasiones algunos métodos presentan resultados deficientes dado que poseen limitaciones en su capacidad para detectar los cambios abruptos, así como lo explican Xie, Gu, Sang, Wu, y Singh (2019), “Todos los métodos para la detección de puntos de cambio en series de tiempo tienen ventajas y desventajas (Lloyd et al., 2014; Jeon et al., 2016; Turner et al., 2016), es posible que no sean suficientes para satisfacer las necesidades prácticas”. No resulta sencillo cumplir con todas las condiciones necesarias para el uso de ciertos métodos de detección de puntos de cambio. Uno de los supuestos más relevantes para analizar las series de tiempo es la estacionaridad. Por tal motivo, cuando no se cumple este supuesto es necesario el uso de técnicas o métodos para determinar el momento de cambio bajo esta circunstancia (Olsen, Chaudhuri, y Godtlielsen, 2008).

Si bien se han utilizado muchos métodos para la detección de puntos de cambio en series de tiempo, la literatura solo ha logrado entregar un debate poco desarrollado en los desempeños relativos (Sastri, Flores, y Valdés, 1989). De esto surge la pregunta, ¿Existirá un mejor desempeño utilizando otras metodologías poco recurrentes en series de tiempo?. En este trabajo se pretende dar una respuesta parcial a esta pregunta.

Los estudios que implementan el análisis wavelet para la detección de puntos de cambios demuestran una interesante manera de obtener información del entorno que, en algunas ocasiones, pasa desapercibida o no es tomada en cuenta en los métodos convencionales. Es muy eficaz en presencia de series de tiempo no estacionarias. Se puede utilizar incluso con series de tiempo no estacionarias sin ningún tipo de transformación previa, lo que ofrece una ventaja sobre los métodos que solo se pueden aplicar a series estacionarias.

Este trabajo de investigación se encuentra dividido en los siguientes capítulos, en el capítulo 1 se presenta el contexto de los puntos de cambio, en el estado del arte se plantean las investigaciones que se han desarrollado con el análisis wavelet, tanto en la práctica como en la teoría, luego se propone el planteamiento del problema de esta investigación con los objetivos, la hipótesis y la pregunta de investigación. Posteriormente, se presentan los tests y métodos clásicos utilizados para la detección de puntos de cambio de esta investigación y los procesos que utilizan para llevar a cabo su función. En el capítulo 2 se presentan los fundamentos del modelamiento estadístico de series temporales, así como una descripción de modelos de series temporales. En el capítulo 3 se introducen los conceptos básicos de la teoría wavelet, sus principales características y sus alcances. En el capítulo 4 se compara la efectividad que poseen los tests a través de simulaciones numéricas realizadas a distintos modelos de series temporales para la detección de puntos de cambio. En el capítulo 5 se aplican los tests de detección de puntos de cambio a los datos del flujo anual del caudal del río Nilo, a los datos de la tasa de desempleo en Chile, a los datos del índice de precios al consumidor general en Chile y a los datos del Seeing. Se realiza una discusión a partir de los resultados obtenidos en cada una de los tests logrando un acuerdo con los diferentes fenómenos ocurridos. En el capítulo 6 se presentan las conclusiones del trabajo.

## 1.1. Detección de puntos de cambio

Los diferentes tests de detección de puntos de cambio se enfocan en determinar el cambio de nivel que se puede presentar en un momento que es desconocido para los analistas, estos cambios son detectados con facilidad si su efecto de cambio es fuerte, aunque es posible que detecten los cambios de nivel débiles.

En las siguientes secciones se van a presentar los tests de detección de puntos de cambio desarrollados en el artículo de [Brown, Durbin, y Evans \(1975\)](#) construyendo las bases teóricas de los test CUSUM y MOSUM. Otros investigadores como [Chu, Hornik, y Kuan \(1995\)](#) quienes proponen ideas para mejorar el test MOSUM a partir de los estudios de [Brown y cols. \(1975\)](#) y de [Ploberger y Krämer \(1992\)](#). Además, se presentan las características del test de Chow bajo los análisis de [Banda-Ortiz y González, 2020](#) quienes utilizan el test de Chow para estudiar el impacto de las crisis financieras para la fecundidad. Otros estudios implementan los métodos clásicos para detectar cambios en los índices bursátiles de países que pertenezcan al MILA (Mercado Integrado Latino Americano) entre los años 2008 y 2018 ([Morales, 2019](#)).

## 1.2. Test CUSUM

Los mejores modelos de regresión son aquellos que cumplen ciertos supuestos, como la media de los residuos sea constante, la varianza de los residuos es finita, entre otros. Sin embargo, existen otros supuestos que son interesantes de estudiar, como el supuesto de que los coeficientes de un modelo de regresión sean iguales o constantes en el tiempo. Existen debates acerca de este supuesto ya que en ocasiones se considera ambiguo, por lo que es necesario un estudio con diferentes criterios para disminuir esa incertidumbre, más aún si el modelo será utilizado para hacer pronósticos (Brown y cols., 1975).

El modelo de regresión que utilizan Ploberger y Krämer (1992) y Brown y cols. (1975) para los test CUSUM y MOSUM es el siguiente:

$$y_t = x_t' \beta_t + u_t, \quad t = 1, \dots, T, \quad (1.1)$$

donde  $t$  representa el tiempo de la serie,  $y_t$  son las observaciones de la variable dependiente,  $x_t = [1, x_{t,2}, \dots, x_{t,K}]'$  es el vector columna de las  $K + 1$  observaciones de la variable independiente con el valor a la unidad en el primer componente,  $\beta$  es el vector columna de los  $K + 1$  coeficientes y los errores  $u_t$  son iid (no necesariamente  $N(0, \sigma^2)$ ) (Ploberger y Krämer, 1992).

En la literatura se ha expresado el rol que poseen los cambios en los residuos para las estimaciones en series temporales. Sin embargo, no todos los métodos tienen la capacidad de detectar todos los cambios que se producen a través del tiempo. Algunos de los métodos más utilizados para las estimaciones son la suma de cuadrados recursivos y la suma de mínimos cuadrados residuales u ordinarios (en esta investigación se profundiza en la suma de los mínimos cuadrados ordinarios). Con respecto a la suma de cuadrados ordinarios, el método no resulta ser sensible para detectar cambios leves o de lento crecimiento de los coeficientes en el tiempo. Por otro lado, posee mejor desempeño que otros métodos utilizados en el área industrial detectando cambios leves en la media (Brown y cols., 1975).

Para poder observar de manera general los mínimos cuadrados residuales individuales  $z_t$  en los gráficos Brown y cols. (1975) proponen que

$$Z_r = \frac{1}{\hat{\sigma}_r} \sum_1^r z_t, \quad r = 1, \dots, T,$$

donde  $z_t$  son residuos de aplicar el método de mínimos cuadrados y  $\hat{\sigma}_r$  es el estimador de la desviación estándar para eliminar el factor de escala irrelevante que depende del valor de  $r$ .

Se establece como hipótesis nula que los coeficientes son constantes en el tiempo

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_T = \beta.$$

Si se asume que  $H_0$  es verdadero, entonces  $b_r$  será la estimación de mínimos cuadrados ordinarios (MCO) de  $\beta$  bajo las primeras  $r$  observaciones, ver definición (1.1.1).

**Definición 1.2.1** Sea  $b_r = (X_r' X_r)^{-1} X_r' Y_r$ , en el cual se tiene una matriz conformada por  $X_r' X_r$  sin singularidad y sea  $w_r$  definido de la siguiente manera:

$$w_r = \frac{y_r - x_r' b_{r-1}}{\sqrt{(1 + x_r'(X_{r-1}' X_{r-1})^{-1} x_r)}}, \quad r = k + 1, \dots, T, \quad (1.2)$$

donde  $X_{r-1}' = [x_1, \dots, x_{r-1}]$ ,  $Y_r = [y_1, \dots, y_r]$  y  $k$  es la cantidad de regresores.

**Definición 1.2.2** Sea  $S_r$  la suma de cuadrados residuales luego de ser ajustado el modelo a las  $r$  primeras observaciones, teniendo en cuenta que  $H_0$  es verdadero, entonces  $S_r = (Y_r X_r b_r)' (Y_r X_r b_r)$ .

En muchas ocasiones se presenta una constante en el modelo de regresión y coincide que una de las variables regresoras es una constante para las primeras  $r_1$  observaciones, donde  $r_1 \geq k$ . La idea principal es que la variable regresora constante inicial se eliminará al comienzo de las recursiones disminuyendo los regresores a  $k - 1$  (Brown y cols., 1975).

Según los investigadores Brown y cols. (1975) para el test CUSUM se debe asumir que  $\beta_t$  es constante hasta el tiempo  $t = t_0$  el cual cambia de valor después de ese tiempo. Entonces, los  $w_r$  tienen medias iguales a 0 en  $r = k + 1, \dots, t_0$ , y medias diferentes de 0 luego de la aplicación en  $r = t_0, \dots, T$ . De esta manera se estudian los gráficos que se utilizan para dar información acerca de las desviaciones de las medias de los  $w_r$  desde el origen mientras se mueve a lo largo del trayecto de la serie en el tiempo.

El primer gráfico es la cantidad acumulada:

$$W_r = \frac{1}{\hat{\sigma}} \sum_{k+1}^r w_j. \quad (1.3)$$

Con  $r$  tomando nuevos valores  $r = k + 1, \dots, T$ , donde  $\hat{\sigma}$  corresponde a la desviación estándar estimada determinada por  $\hat{\sigma}^2 = S_T / (T - k)$ . Posteriormente, se necesita un método para evidenciar la relevancia de la desviación del trayecto muestral de  $W_r$  del valor medio  $E(W_r) = 0$ . Por lo tanto, es necesario encontrar un par de líneas que se ubiquen simétricamente por arriba y por abajo de la línea de  $W_r = 0$ , de tal forma que la probabilidad asociada al cruce de ambas líneas o rectas sea igual a  $\alpha$ , el nivel de significación (Brown y cols., 1975).

Utilizando la propiedad de los  $w_r$  bajo  $H_0$ , las  $W_{k+1}, \dots, W_T$  se definen como una secuencia de variables aproximadamente normal, tales que:

$$E(W_r) = 0 \quad V(W_r) = r - k \quad y \quad C(W_r, W_s) = \min(r, s) - k, \quad \text{para } r, s \geq k$$

Para derivar el test CUSUM, los  $W_r$  se aproximan usando un proceso Gaussiano continuo  $\{Z_t, k \leq t \leq T\}$  que se denomina movimiento Browniano con inicio en  $Z_t = 0$  en el tiempo  $t = k$ . Por lo tanto,  $Z_t$  tiene media igual a cero y desviación estándar de  $\sqrt{t - k}$ . La línea recta a escoger se obtiene de 2 maneras. Si se quiere encontrar una curva de manera tal que bajo la hipótesis nula la probabilidad de que la ruta de la muestra por arriba de la curva en un punto arbitrario entre  $t = k$  y  $t = T$  sea constante, se escogen curvas mediante la cantidad  $\pm \lambda \sqrt{(t - k)}$  donde  $\lambda$  es constante. Sin embargo, el objetivo central es probar que los  $W_r$  son capaces de traspasar las rectas, por lo que la probabilidad de cruce no puede ser constante en todo el recorrido de  $t$ , para solucionar este

problema se elige una familia de rectas tangentes a las curvas  $\pm\lambda\sqrt{(t-k)}$  en los puntos intermedios entre  $t = k$  y  $t = T$ . Esto nos deja con una familia de pares de rectas que se encuentran entre los puntos  $\{k, \pm a\sqrt{(T-k)}\}, \{T, \pm 3a\sqrt{(T-k)}\}$ , donde  $a$  es un parámetro. Por ende, dentro de esta familia se debe encontrar un miembro cuya probabilidad de que la muestra  $Z_t$  traspase los pares de rectas es de  $\frac{1}{2}\alpha$  (Brown y cols., 1975).

Ploberger y Krämer (1992) buscan la forma de mejorar los métodos propuestos por Brown y cols. (1975) y proponen un nuevo enfoque como se observa en la ecuación (1.4)

$$\hat{u}_t^{(T)} = y_t - x_t' \hat{\beta}^{(T)}, \quad (1.4)$$

donde  $\hat{\beta}^{(T)} = (\sum_{t=1}^T x_t x_t')^{-1} \sum_{t=1}^T x_t y_t$  es la estimación de los coeficientes por MCO y  $(T)$  denota la dependencia de las observaciones de la muestra ( $t = 1, \dots, T$ ). Por lo que poseen una gran ventaja contra los residuos recursivos, ya que los residuos van cambiando a media que se ingresan más observaciones a la muestra. También, están correlacionados y son heterocedásticos bajo  $H_0$ . Una de las razones de utilizar los MCO es que poseen una mejor aproximación a los verdaderos cambios abruptos bajo  $H_0$ , es decir, es posible detectar los cambios abruptos con mayor facilidad si se presenta una desviación en el modelo (Ploberger y Krämer, 1992).

La estimación de mínimos cuadrados ordinarios es un método con ventajas considerables al momento de determinar cambios abruptos. Aunque, no es totalmente fiable, su desventaja se produce cuando se presentan intersecciones en el modelo de regresión provocando que la suma de residuos sea cero. Por lo que es difícil verificar que ha ocurrido una desviación del modelo después del cambio estructural. Además, el tiempo transcurrido al detectar el cambio estructural es irrelevante, ya que tarde o temprano los residuos de mínimos cuadrados ordinarios volverán a cero (Ploberger y Krämer, 1992).

La ecuación (1.4) da cuenta de las ventajas para detectar los cambios estructurales, por lo que Ploberger y Krämer (1992) proponen una mejora para la ecuación (1.3),

**Definición 1.2.3** Sea  $B(z)^{(T)}$  el puente Browniano estándar que se compone de la suma acumulada de los mínimos cuadrados residuales  $\hat{u}_t^{(T)}$ . Entonces,

$$B(z)^{(T)} = \frac{1}{\hat{\sigma}\sqrt{T}} \sum_{t=1}^{T_z} \hat{u}_t^{(T)}, \quad (1.5)$$

donde  $T_z$  es el momento en que se desvanece uniformemente en  $z$  y se definen las rutas de las muestras para que sean elementos aleatorios del espacio de probabilidad subyacente dentro de  $D[0, 1]$  de todas las funciones reales en el intervalo  $[0, 1]$  que son continuas por la derecha y tienen límites por la izquierda, es decir, que haya convergencia entre las funciones (Ploberger y Krämer, 1992).

Para evaluar la convergencia del proceso definido por la ecuación (1.5) se deben tener en cuenta 3 supuestos:

- **Supuesto 1:** Los regresores  $x_t$  y los errores  $u_t$  se definen como un espacio de probabilidad tal que

$$\lim_{T \rightarrow \infty} \sup \frac{1}{T} \sum_{t=1}^T \|x_t\|^{2+\gamma} < \infty \quad \text{para algún } \gamma > 0,$$

- **Supuesto 2:** Sean  $x_t$  el vector de regresores y  $x_t'$  el vector de regresores inversos tales que

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T x_t x_t' = R,$$

para alguna matriz  $R$  sin singularidad y no estocástica definida en  $(K \times K)$ , donde se asume que no hay pérdida de generalidad y el modelo se ha reparametrizado tal que

$$R = \begin{bmatrix} 1 & 0 \\ 0 & R^* \end{bmatrix},$$

En este caso especial, se tiene que

$$c \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T x_t = [1, 0, \dots, 0]',$$

- **Supuesto 3:** Sean los errores  $u_t$  son estacionarios y ergódicas, con

$$E(u_t | \Upsilon_t) = 0, \quad E(u_t^2 | \Upsilon_t) = \sigma^2,$$

donde  $\Upsilon_t$  es el espacio de  $\sigma$  generado por  $\{y_{t-s}, x_{t-s+1}, u_{t-s} | s \geq 1\}$ .

Una vez cumplido con los supuestos, bajo  $H_0$ ,

$$B^{(T)}(z) \xrightarrow{d} B(z) \quad T \rightarrow \infty, \quad (1.6)$$

donde  $\xrightarrow{d}$  es la convergencia en distribución.

El estadístico de prueba para el test CUSUM es

$$\sup_{0 \leq z \leq 1} |B^{(T)}(z)|, \quad (1.7)$$

dado que  $\sup_{0 \leq z \leq 1} |B^{(T)}(z)|$  es una función continua de  $B(z)$ . Se tiene lo siguiente

$$\sup_{0 \leq z \leq 1} |B^{(T)}(z)| \xrightarrow{d} \sup_{0 \leq z \leq 1} |B(z)|, \quad (1.8)$$

Considerando la ecuación (1.8) es posible construir la probabilidad de que  $B(z)$  cruce ambas bandas de confianza lo que lleva a

$$P\left(\sup_{0 \leq z \leq 1} |B^{(T)}(z)| > a\right) \longrightarrow P\left(\sup_{0 \leq z \leq 1} |B(z)| > a\right),$$

donde  $a$  pertenece a todos los reales y  $B(z)$  tiene media igual a 0. Además, dado que  $T \rightarrow \infty$ , [Ploberger y Krämer \(1992\)](#) proponen un valor para determinar que los residuos han cruzado ambas bandas de confianza igual a

$$2 \sum_{j=1}^{\infty} (-1)^{j+1} \exp(-2j^2 a^2).$$

La varianza del puente Browniano a través del intervalo  $[0,1]$  está dado por

$$V(B(z)) = z. \quad (1.9)$$

el resultado equivale a que la probabilidad de rechazar la hipótesis nula va variando a través del intervalo  $[0,1]$ , es decir, que la probabilidad de que  $B(z)$  traspase una de las líneas rectas paralelas al eje  $z$  es de un máximo para  $z = 1/2$ .

### 1.3. Test MOSUM

El test MOSUM (*Moving Sum*) posee características similares al test CUSUM, ya que se utilizan los residuos recursivos o los mínimos cuadrados residuales de un proceso Wiener estándar para determinar el cruce de las bandas de confianza en valor absoluto. Sin embargo, se debe incluir un componente adicional al test MOSUM, que es una ventana de movimiento que se construye a partir de una cierta cantidad de observaciones de la muestra (Chu y cols., 1995), la cual se explica a continuación.

El test MOSUM de mínimos cuadrados residuales es descrito por Chu y cols. (1995) de la siguiente manera. Sean  $\hat{\epsilon}_{T,n} = y_n - x_n' \hat{\beta}_T$ , con  $n = 1, \dots, T$  los mínimos cuadrados residuales,  $\hat{\beta}_n = (X_n' X_n)^{-1} X_n' Y_n$  el estimador de mínimos cuadrados ordinarios en el momento  $n$  para  $j = 0, \dots, T - [Th]$ , el  $j$ -ésimo momento de la suma móvil de los mínimos cuadrados residuales se obtiene a partir de

$$\frac{1}{\hat{\sigma}[Th]^{\frac{1}{2}}} \sum_{n=j+1}^{j+[Th]} \hat{\epsilon}_{T,n}, \quad (1.10)$$

donde  $[Th]$  es la ventana de sumas móviles de los mínimos cuadrados residuales que se mueve a través de la muestra, además,  $[Th]$  es la parte entera de  $Th$ , cuyo valor que tome ( $0 < h < 1$ ) es la proporción de residuos utilizados para conformar cada suma móvil. El estadístico de prueba para la suma móvil de los mínimos cuadrados residuales está constituido de la misma manera que el estadístico de prueba de la suma acumulativa de mínimos cuadrados residuales de Ploberger y Krämer (1992), solo que en esta ocasión se introduce un nuevo componente  $[Th]$ .

$$MS_{T,h}^o = \max_{0 \leq j \leq T-[Th]} \frac{1}{\hat{\sigma}[Th]^{\frac{1}{2}}} \left| \sum_{n=j+1}^{j+[Th]} \hat{\epsilon}_{T,n} \right|, \quad (1.11)$$

donde el símbolo  $o$  indica que el proceso browniano se construyen a partir de los mínimos cuadrados residuales.

En el teorema presentado por Chu y cols. (1995) se define el proceso browniano de la siguiente manera. Sea  $W^o$  un proceso Wiener definido en  $[0,1]$  y

$$S_h^o(t) = h^{\frac{1}{2}} \{W^o(t+h)\} \quad (0 \leq t \leq 1-h).$$

Además, la hipótesis nula para este método es

$$MS_{T,h}^o \longrightarrow \max_{0 \leq t \leq 1-h} |S_h^o(t)|.$$

Cuyos valores críticos asintóticos del test MOSUM de mínimos cuadrados están determinados por la probabilidad de que traspase las bandas de confianza de ambos lados del proceso no estandarizado  $S_h^o$ ,

$$\lim_{T \rightarrow \infty} P(MS_{T,h}^o > b) = P\{|S^o(t)| > b, \text{ para algún } 0 \leq t \leq 1-h\}. \quad (1.12)$$

Existe otra manera más eficiente de estudiar la variación en el tiempo de  $\beta_t$  mediante un ajuste a la regresión en un intervalo pequeño de  $k$  observaciones sucesivas y trasladar este intervalo a través de la serie. Los gráficos de los coeficientes obtenidos por la función en el tiempo otorgan otra prueba de medición más a las desviaciones de la estabilidad de  $\sigma^2$  (Chu y cols., 1995).

La detección de puntos de cambio está implementada en el paquete "Strucchange" del software R con distintos métodos ya conocidos, tales como el test CUSUM, el test MOSUM y el test de Chow. Los 2 primeros test utilizan el test de fluctuación generalizado y el proceso de fluctuación empírico. El desarrollo teórico y computacional es entregado por los investigadores Zeileis, Leisch, Hornik, y Kleiber (s.f.).

### 1.3.1. Test de fluctuación generalizado

Los tests de fluctuación generalizados para el proceso CUSUM y MOSUM ajustan un modelo a las observaciones y conforman un proceso empírico que toma la fluctuación en los residuos o en las estimaciones. En este proceso se conocen los límites, cuya probabilidad de que traspase las rectas bajo  $H_0$  es  $\alpha$ . Si la ruta del proceso empírico atraviesa esos límites, la fluctuación resultante será inusualmente grande y, por ende, la hipótesis nula debe rechazarse con un nivel de significancia  $\alpha$  (Zeileis y cols., s.f.).

### 1.3.2. Proceso de fluctuación empírico

En base a las fórmulas que se utilizan en un modelo de regresión lineal simple o múltiple, la función "efp" procede de un proceso de fluctuación empírico ajustado. A continuación se presentan los tipos de procesos de fluctuación más conocidos. En el paquete "strucchange" se puede utilizar la función *efp* para aplicar los test CUSUM y MOSUM con los métodos *OLS-CUSUM* o *Rec-CUSUM* y *OLS-MOSUM* o *Rec-MOSUM* en la condición *type*.

El primero es el proceso CUSUM por MCO que utiliza la suma de los mínimos cuadrados residuales el cual es descrito por [Ploberger y Krämer \(1992\)](#) con un estadístico dado por

$$W_n^o(t) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=1}^{[nt]} \hat{u}_i \quad (0 \leq t \leq 1).$$

El proceso limitante o estadístico de prueba para  $W_n^o(t)$  es el puente browniano estándar  $W^o(t) = W(t) - tW(1)$ , donde  $W(\cdot)$  detecta que hay un cambio abrupto alrededor del punto de quiebre ([Zeileis, Kleiber, Krämer, y Hornik, 2003](#)).

Para el caso del proceso MOSUM que utiliza las sumas móviles de los residuos del puente browniano estándar para detectar cambios estructurales, a diferencia del proceso CUSUM, este tiene la característica de que no acumula la suma de todos los residuos hasta llegar a un tiempo  $t$ , sino que suma un número fijo de residuos en una ventana de observaciones cuyas dimensiones están sujetas al parámetro de anchura  $h \in (0, 1)$  y que va desplazándose por todo el recorrido temporal de la muestra ([Zeileis y cols., s.f.](#)). Por todo lo anterior, el proceso MOSUM por el método de los mínimos cuadrados ordinarios (MCO) se describe de la siguiente manera:

$$\begin{aligned} M_n^o(t | h) &= \frac{1}{\tilde{\sigma}\sqrt{n}} \left( \sum_{i=[N_n t]+1}^{[N_n t]+[nh]} \hat{u}_i \right) \quad (0 \leq t \leq 1 - h), \\ &= W_n^o \left( \frac{[N_n t] + [nh]}{n} \right) - W_n^o \left( \frac{[N_n t]}{n} \right). \end{aligned} \quad (1.13)$$

donde  $N_n = (n - [nh]) / (1 - h)$ . El proceso de fluctuación empírico está constituido por un puente browniano. Del mismo modo que con el test CUSUM, si se detecta un nuevo cambio estructural en el tiempo  $t_0$ , entonces ambas rutas del test MOSUM van a tener un cambio abrupto alrededor de  $t_0$  ([Zeileis y cols., s.f.](#)).

## Límites de bondad de ajuste de los test CUSUM y MOSUM

Los límites de los procesos MOSUM son constantes,  $b(t) = \lambda$ , en este caso se asume que los procesos limitantes son estacionarios. Sin embargo, en los procesos CUSUM no es igual, ya que el movimiento browniano o puente browniano, no cumple el criterio de estacionariedad. Lo más natural es el uso de límites que sean proporcionales a la función de desviación estándar del proceso teórico escogido, es decir,

$$b(t) = \lambda \sqrt{t(1-t)} \quad (1.14)$$

Para el proceso CUSUM por MCO, donde  $\lambda$  corresponde al nivel de confianza, los límites que se utilizan usualmente son lineales, ya que se tiene conocimiento acerca de la forma cerrada para la probabilidad de que atraviese las rectas. Por ende se estandarizan sus bandas de confianza teniendo en cuenta que

$$b(t) = \lambda \quad (1.15)$$

son las rectas tangenciales a las bandas de confianza de la ecuación (1.18) en el tiempo  $t = 0,5$  ([Zeileis y cols., s.f.](#)). Lo cual coincide con los estudios realizados por [Brown y cols. \(1975\)](#) quienes proponen una familia de pares de rectas tangenciales a las curvas  $\pm \lambda \sqrt{(t-k)}$ . Sin embargo, en

algunos estudios se entregan valores críticos útiles dependiendo del valor de significancia, para  $\alpha = 1\%$ (1,63),  $\alpha = 5\%$ (1,36) y  $\alpha = 10\%$ (1,22) (Ploberger y Krämer, 1992).

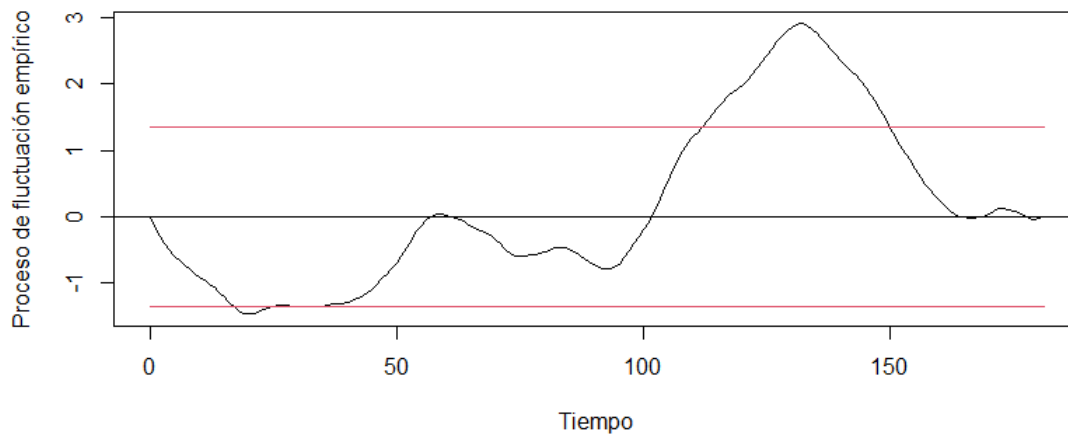


Figura 1.1: Ejemplo de un proceso de fluctuación empírico del test CUSUM por MCO junto con las rectas umbrales que permiten detectar los puntos de cambio para un nivel de significación dado.

**Fuente:** Elaboración propia.

En la figura 1.1 se observa que las rectas paralelas a la media de los residuos del proceso de fluctuación empírico son las bandas de confianza, el proceso de fluctuación empírico va aumentando su altura a medida que transcurre el tiempo hasta traspasar una de las bandas de confianza y señalar que hay un punto de cambio, al llegar al punto de cambio el proceso comienza a descender hacia la media igual a cero.

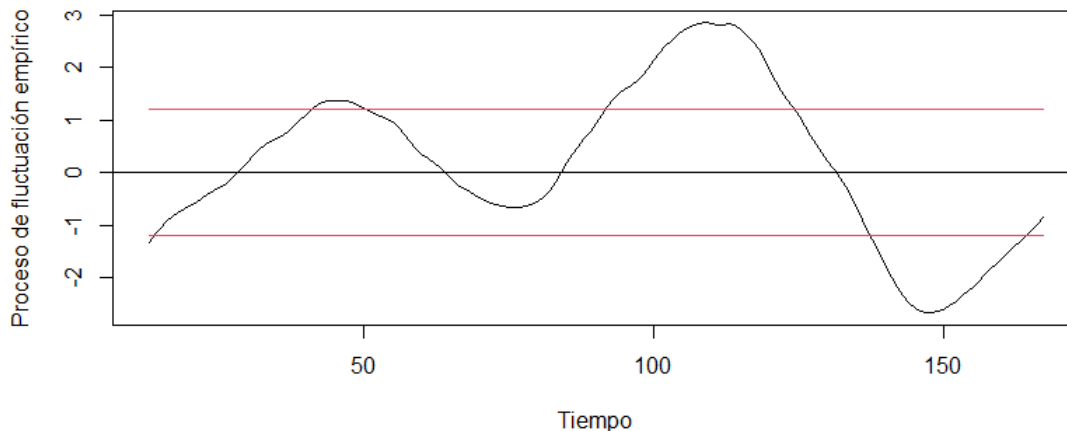


Figura 1.2: Ejemplo de un proceso de fluctuación empírico del test MOSUM por MCO junto con las rectas umbrales que permiten detectar los puntos de cambio para un nivel de significación dado.

**Fuente:** Elaboración propia.

Al igual que en la figura 1.1, en la figura 1.2 se observa que las rectas paralelas a la media de los residuos del proceso de fluctuación empírico son las bandas de confianza, el proceso de fluctuación empírico va aumentando su altura a medida que transcurre el tiempo hasta traspasar una de las bandas de confianza y señalar que hay un punto de cambio, al llegar al punto de cambio el proceso comienza a descender hacia la media igual a cero.

## 1.4. Test de Chow

El test de Chow fue desarrollado por el economista Gregory Chow (1960) y tiene como función probar si los coeficientes de 2 modelos de regresión lineal son iguales. El test de Chow consiste en utilizar 2 submuestras de la muestra original y, posteriormente, se contrastan por medio de la suma de los residuos al cuadrado (Morales, 2019). Siguiendo con la propuesta de la hipótesis nula de que los parámetros del modelo son iguales o que no existe cambios estructurales, se construye el estadístico de la siguiente manera:

$$F = \frac{(SRC_R - SRC_{NR})/k}{SRC_{NR}/(n_1 + n_2 - 2k)} \sim F_{[k, (n_1 + n_2 - 2k)]} \quad (1.16)$$

donde  $SRC_R$  corresponde a la suma de residuos al cuadrado restringido o de la muestra original;  $SRC_{NR}$  son las sumas de los residuos al cuadrado no restringidos o de las submuestras;  $k$  corresponde al número de regresores y  $n_1, n_2$  corresponden al tamaño de las 2 submuestras, el estadístico de prueba sigue una distribución F con  $(n_1 + n_2 - 2k)$  grados de libertad. Por consiguiente, si el valor del estadístico de prueba es mayor al valor crítico F obtenido con un nivel de significancia  $\alpha$ , se rechaza la hipótesis nula y se concluye de que hay evidencia suficiente para asegurar que existe un cambio estructural en uno de los parámetros del modelo (Banda-Ortiz y González, 2020).

Para obtener el cambio estructural, en primera instancia, se realizan iteraciones de los valores que puede tomar el estadístico F entre los posibles cambios estructurales de la muestra en intervalos fijos.

$$SupF = \sup_{l_1 \leq l \leq l_2} F(l), \quad (1.17)$$

donde  $l$  son los puntos que dividen la muestra total para cada submuestra,  $(l_1, l_2)$  son las iteraciones que marcan el intervalo de interés y  $F(l)$  es el valor del estadístico F más grande de todos los valores que puede tomar el estadístico F entre  $(l_1, l_2)$ . Además, el margen de iteraciones está delimitado entre los percentiles 15 y 85 de la muestra, de esta manera se potencia la efectividad de la detección del cambio estructural. Sin embargo, no es suficiente el uso de este margen de estudio, ya que se puede estar excluyendo potenciales cambios estructurales que sean de interés para una investigación (Morales, 2019).

Los límites del estadístico F se determinan de manera similar que los límites de los procesos de fluctuación empíricos. Bajo la hipótesis nula de que no hay cambios estructurales, la probabilidad asintótica de que el valor supremo del estadístico de prueba F atraviese esos límites es  $\alpha$  (Zeileis y cols., s.f.).

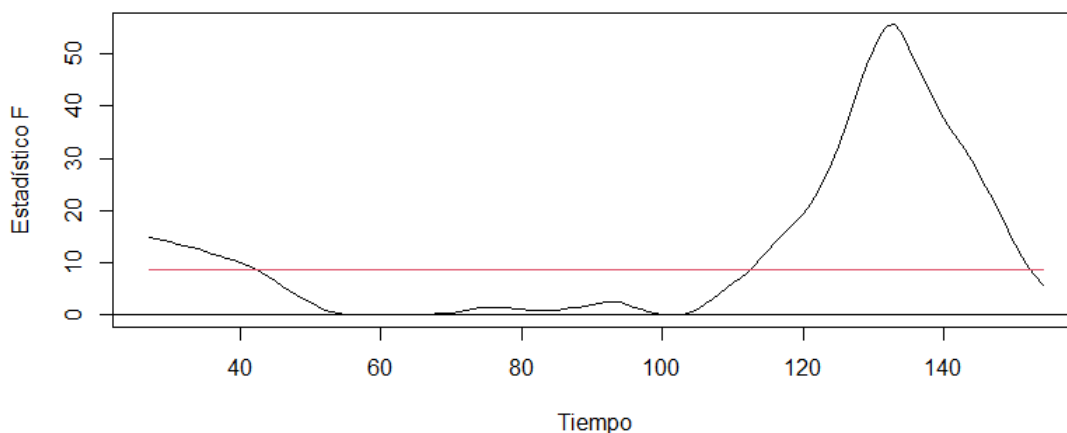


Figura 1.3: Ejemplo del gráfico del estadístico de prueba F del test de Chow junto con la recta umbral que permite detectar los puntos de cambio.

**Fuente:** Elaboración propia.

En la figura 1.3 se observan los distintos valores que toma el estadístico F en las distintas combinaciones de las submuestras cuyo valor más alto señala el punto de cambio dentro de la serie temporal que traspasa la banda de confianza.

## 1.5. Test de múltiples puntos de quiebre

En la literatura y en múltiples estudios se han abordado las técnicas para estimar un punto de cambio o cambio estructural que es desconocido en un modelo de regresión lineal, pero no son muchos los estudios donde se estimen los múltiples cambios estructurales y la estimación de las fechas en que se detectan. Sin embargo, [Bai y Perron \(1998\)](#) buscaron la manera de establecer los métodos de estimación para ambas interrogantes, los múltiples puntos de quiebre y las fechas de esos puntos de quiebre. Para su estudio consideran un modelo de regresión múltiple con  $m$  puntos de quiebres (o  $m + 1$  quiebres),

$$y_t = x_t' \beta + z_t' \delta_j + u_t \quad t = (T_{j-1} + 1, \dots, T_j), \quad (1.18)$$

para  $j = 1, \dots, m+1$ , se utiliza la convención que  $T_0 = 0$  y  $T_{m+1} = T$ ,  $y_t$  es la variable dependiente observada,  $x_t(p \times 1)$  y  $z_t(q \times 1)$  son los vectores covariantes,  $\beta$  con  $\delta_j$  son los vectores de los coeficientes y  $u_t$  es el ruido. Además, los subíndices de  $(T_1, \dots, T_m)$  indican los puntos de quiebres que se consideran desconocidos. En síntesis, se estiman los coeficientes de regresión desconocidos en compañía de los puntos de quiebre cuando las  $T$  observaciones en  $(x_t, y_t, z_t)$  estén disponibles.

El modelo de regresión lineal múltiple expresado de forma matricial es

$$Y = X\beta + \bar{Z}\delta + U, \quad (1.19)$$

donde  $Y = (y_1, \dots, y_T)'$ ,  $X = (x_1, \dots, x_T)'$ ,  $U = (u_1, \dots, u_T)'$ ,  $\delta = (\delta_1', \dots, \delta_{m+1}')'$ ,  $\beta$  es el vector de coeficientes y  $\bar{Z}$  es la matriz que se separa diagonalmente  $Z$  en los  $m$ -quiebres  $(T_1, \dots, T_m)$ , es decir,  $\bar{Z} = \text{diag}(Z_1, \dots, Z_{m+1})$  con  $Z_i = (z_{T_{i-1}+1}, \dots, z_{T_i})'$ . Además, el índice 0 indica el verdadero valor del parámetro. En específico,  $\delta^0 = (\delta_1^0, \dots, \delta_{m+1}^0)'$  y  $(T_1^0, \dots, T_m^0)$  son los verdaderos valores de los parámetros  $\delta$  y los puntos de quiebre. La matriz  $\bar{Z}^0$  divide de manera diagonal  $Z$  en las  $m$ -quiebres  $(T_1^0, \dots, T_m^0)$ . Dado lo anterior, el nuevo modelo para estimar múltiples puntos de quiebres es

$$Y = X\beta^0 + \bar{Z}^0\delta^0 + U. \quad (1.20)$$

Cuyo principal objetivo es estimar los coeficientes desconocidos  $(\beta^0, \delta_1^0, \dots, \delta_m^0, T_1^0, \dots, T_m^0)$ , para los cuales se asume que  $\delta_i^0 \neq \delta_{i+1}^0$  ( $1 \leq k \leq m$ ).

Para realizar la estimación [Bai y Perron \(1998\)](#) utilizan el método de los mínimos cuadrados residuales. Entonces, para cada  $m$ -vector  $(T_1, \dots, T_m)$  se denota como  $\{T_j\}$ , la estimación de mínimos cuadrados de  $\beta$  y los  $\delta_j$  se obtiene minimizando la suma de los cuadrados residuales  $\sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} [y_t - x_t'\beta - z_t'\delta_i]^2$ . Sean  $\hat{\beta}(T_j)$  y  $\hat{\delta}(T_j)$  los resultados de las estimaciones. Luego se reemplaza en la función matricial y se obtienen los resultados de la suma de cuadrados residuales como  $S_T(S_1, \dots, S_m)$ , los estimadores de los puntos de quiebre  $(\hat{T}_1, \dots, \hat{T}_m)$  se obtienen

$$(\hat{T}_1, \dots, \hat{T}_m) = \text{argmin}_{T_1, \dots, T_m} S_T(T_1, \dots, T_m), \quad (1.21)$$

donde la minimización se toma sobre todas las particiones  $(T_1, \dots, T_m)$  tal que  $T_i - T_{i-1} \geq q$ . Por consiguiente, los estimadores de los parámetros de regresión son las estimaciones de los mínimos cuadrados vinculados a los  $m$ -vectores o separaciones estimadas  $\{\hat{T}_j\}$ , donde  $\hat{\beta} = \hat{\beta}(\{\hat{T}_j\})$  y  $\hat{\delta} = \hat{\delta}(\{\hat{T}_j\})$ .

En lo que respecta al test, Bai y Perron (1998) consideran un estadístico de test de tipo supremo de una función F con hipótesis  $m = 0$  puntos de quiebre versus un determinado número de puntos de quiebres  $m = k$ .

Sea  $(T_1, \dots, T_k)$  un vector con  $m = k$  puntos de quiebre tal que, la fecha del puntos de quiebre  $T_i$  se obtiene a partir de la tasa de convergencia del verdadero valor de los puntos de quiebre  $T_{m+1}^0 = T$  de cada partición  $\lambda_i$ , cuyas estimaciones de  $\hat{\lambda} = (\hat{\lambda}_1, \dots, \hat{\lambda}_m) = (\hat{T}_1/T, \dots, \hat{T}_m/T)$  contienen los verdaderos valores de las particiones  $\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0)$ , es decir,  $T_i = [T\lambda_i]$  con  $i = (1, \dots, k)$ . A partir de lo anterior, se define el estadístico de múltiples puntos de quiebres de la siguiente manera:

$$F_T(\lambda_1, \dots, \lambda_k; q) = \left( \frac{T - (k + 1)q - p}{kq} \right) \frac{\hat{\delta}' R' (R(\bar{Z}' M_x \bar{Z})^{-1} R')^{-1} R \hat{\delta}}{SSR_k} \quad (1.22)$$

donde R es la matriz convencional tal que  $(R\delta)' = (\delta'_1 - \delta'_2, \dots, \delta'_k - \delta'_{k+1})$  y  $M_x = I - X(X'X)^{-1}X'$ . Además,  $SSR_k$  es la suma de cuadrados residuales que depende del segmento  $(T_1, \dots, T_k)$ .

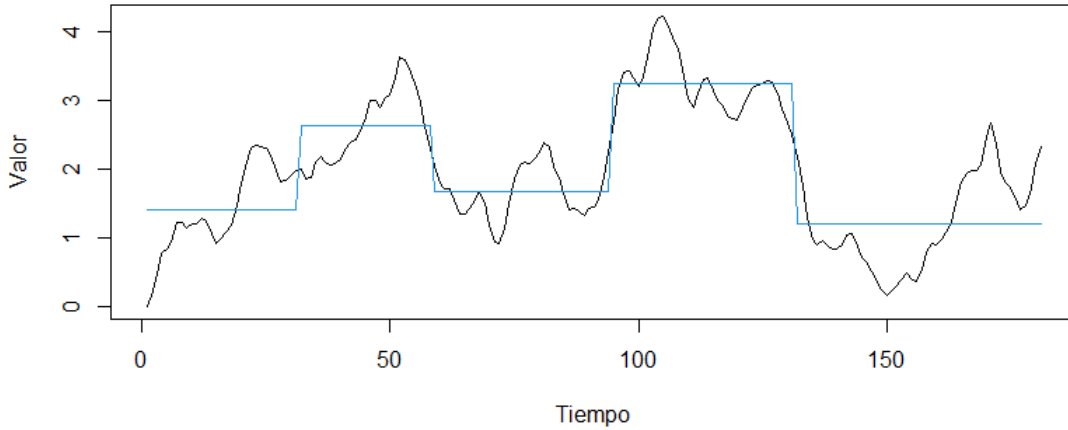


Figura 1.4: Ejemplo del gráfico con el resultados del test de múltiples puntos de quiebre.

**Fuente:** Elaboración propia

En la figura 1.4 se puede observar la formación de segmentos a lo largo de la serie temporal con sus respectivos puntos de cambio. Sin embargo, es necesario determinar restricciones sobre los posibles valores de las fechas de los puntos de quiebre. Es decir, se debe limitar cada fecha de los puntos de quiebre para que sea asintóticamente distinta y limitada a los límites que se construyen a partir de la muestra. El conjunto que determina el efecto de un punto de quiebre es  $\Lambda_\epsilon = \{(\lambda_1, \dots, \lambda_k); |\lambda_{i+1} - \lambda_i| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_k \leq 1 - \epsilon\}$ . Finalmente, el estadístico del *SupF* se define como

$$\sup F_T(k; p) = \sup_{(\lambda_1, \dots, \lambda_k) \in \Lambda_\epsilon} F_T(\lambda_1, \dots, \lambda_k; q). \quad (1.23)$$

donde  $q$  son los grados de libertad,  $p$  es el número de coeficientes y  $\epsilon = 0,05$  es un valor crítico asintótico.

En la siguiente sección se presentan las investigaciones relacionadas con el análisis wavelet aplicadas a series temporales para la detección de puntos de cambio, tanto prácticas como teóricas.

## 1.6. Estado del arte

Existen muchos métodos para la detección de puntos de cambio en series temporales, pero es difícil encontrar un método válido para la detección de puntos de cambio implementando el análisis wavelet. Es por eso que [Chen y Gupta \(2001\)](#) proponen diferentes métodos para la detección y estimación de los puntos de cambio, los métodos que presentan son en su gran mayoría métodos utilizados en las estadísticas tales como, cambios en la media con varianza conocida y desconocida, cambios en la varianza, modelos normales multivalentes, entre otros. Por otro lado, también se encuentra la detección de puntos de cambio mediante las transformadas wavelet utilizando un test de hipótesis en el cual la hipótesis nula es un cambio suave en el modelo versus la hipótesis alternativa donde el cambio es abrupto. Este test se describe con más detalle en la sección (3.4).

Los estudios que comparan el análisis wavelet con otros métodos frecuentemente utilizados en series temporales, demuestran que el análisis wavelet posee una mayor efectividad al momento de detectar puntos de cambio. Un ejemplo de ello se encuentra en el estudio de [Alarcon-Aquino y Barria \(2009\)](#) donde utilizan la segmentación wavelet para detectar los puntos de cambio en la varianza de los coeficientes mediante la transformada wavelet discreta de máxima superposición (TWDMS) o en inglés (*MODWT*). Las wavelet utilizadas son la wavelet Haar, la familia wavelet de Daubechies y la wavelet Spline cuadrática. Los resultados obtenidos en los estudios de simulación, dan cuenta que la eficacia de detección de puntos de cambio es mayor con las wavelets que en las pruebas de divergencia y el test GLR de Brandt. Las wavelets que destacaron fueron la wavelet Spline cuadrática y la wavelet de baja asimetría que pertenece a la familia wavelet de Daubechies que utilizó un filtro igual a 10.

Muchos investigadores han buscado la manera de incorporar el análisis wavelet de tal forma que su potencia de detección disminuya lo menos posible en el tiempo. Ante ese desafío los investigadores [Yazgan y Özkan \(2015\)](#) utilizan las propiedades de la descomposición wavelet para elaborar una prueba estadística que detecte cambios en la media de un proceso distribuido independientemente, comparando el rendimiento de su prueba con otros 3 métodos de detección de puntos de cambio tales como, la prueba MOSUM (Suma acumulativa móvil), la prueba CUSUM (Suma acumulativa) y la prueba Sup-f (Supremo de la prueba F), los resultados demuestran que el algoritmo propuesto es mejor que los métodos CUSUM y Sup-f cuando existen múltiples puntos de cambio, pero si se tiene un intervalo de tiempo más extenso estos métodos no pierden tan rápidamente su potencia de detección. Por otro lado, la prueba MOSUM mantiene su potencia por una gran cantidad de puntos de cambio ya que wavelet y este método utilizan información local de la misma ventana de aprendizaje.

La búsqueda por ampliar las capacidades del análisis wavelet ha influenciado el trabajo de [Raimondo y Tajvidi \(2004\)](#) para llevar a cabo un método capaz de diferenciar un punto de cambio dentro de series con mucho ruido. Ellos implementan un método basado en el análisis wavelet y teoría de valores extremos para determinar la existencia de discontinuidades en una función con ruido. Su función es modelar los puntos críticos que sobrepasen el umbral de la transformada wavelet. Para obtener información acerca de las discontinuidades se usan las propiedades de los coeficientes wavelet, ya que tienen como función reflejar la regularidad local de la señal, es decir, donde la

función es irregular los coeficientes tienden a ser grandes y donde la función es regular o suave los coeficientes son pequeños (ver (Chen y Gupta, 2001)). De esta manera se puede identificar las discontinuidades o cambios abruptos en las señales ruidosas. El test del modelo implementado utiliza las sumas normalizadas provenientes de una aproximación de distribución Pareto generalizada. La elección del umbral es un proceso delicado para lograr el equilibrio entre el sesgo y la varianza. Si el umbral está en una posición muy alta se tendrán menos observaciones y, en consecuencia, se obtendrá una mayor variación de las estimaciones. Por otro lado, para el caso de seleccionar un umbral muy bajo las estimaciones van estar muy sesgadas. No resulta fácil determinar los límites correctos para optar por un umbral adecuado, más aún cuando se tiene un tamaño de muestra pequeño.

Otros trabajos sobre detección de puntos de cambio con análisis wavelet son los siguientes:

- El potencial del análisis wavelet se ha extendido a áreas conservadoras como la medicina cuyos resultados deben ser los más precisos y asumiendo los menores riesgos para sus pacientes. [Keković y Sekulić \(2019\)](#) utilizan las capacidades de detección del análisis wavelet para elaborar un algoritmo basado en filtros de medias móviles con la transformada wavelet para detectar puntos de cambio en las frecuencias de un electroencefalograma (EEG). Los resultados obtenidos muestran una efectividad del 90 % en la detección de puntos de cambio en filtros de medias móviles con transformadas wavelet en frecuencias inferiores a 40 Hz.
- En el estudio de [Chan y cols. \(2021\)](#) desarrollan un algoritmo de detección de puntos de cambio llamado WAvelet KNEe (WANE) para estimar el tiempo de llegada de las señales vibratorias en las redes de distribución de agua. Este algoritmo es comparado con otros 2 métodos convencionales que se han empleado para la estimación de las señales vibratorias en las redes de distribución de agua, uno es el método de la transformada wavelet discreta y el otro es la suma acumulativa (CUSUM). El primer método utiliza la wavelet Haar y la desviación estándar de los coeficientes de detalle como detector en cada uno de los niveles de descomposición. Luego, se estima el tiempo que se demora en detectar la señal vibratoria usando los coeficientes de detalles. El segundo método elimina el ruido de la señal vibratoria por medio de filtros de mínimos cuadrado recursivos (RLS). Por consiguiente, se realiza la prueba CUSUM que cumple la función de determinar si ha ocurrido un cambio en la señal vibratoria. Los resultados obtenidos dan cuenta que el algoritmo (WANE) fue el mejor evaluado, pero posee ciertas falencias como los hiperparámetros, la compleja implementación computacional, entre otros puntos. Sin embargo, es más precisa en contraparte de las metodologías convencionales.

En la siguiente sección se presenta el planteamiento del problema que motiva el desarrollo de esta investigación.

## 1.7. Planteamiento del problema

A través de los años se han presentado distintos métodos para detectar puntos de cambio en las series temporales, muchos de ellos enfocan sus esfuerzos en la detección de cambios de nivel en media sin tomar en cuenta las zonas de alta variabilidad donde puede estar la información interesante de analizar, por lo tanto, el análisis wavelet puede ser el método apropiado para resolver esa desventaja.

Dado que los métodos tradicionales de detección de puntos de cambio tienen ciertas desventajas tales como, un pobre desempeño a la hora de detectar puntos de cambio en varianza, tanto cambios pequeños como grandes, las preguntas que guiaron esta investigación son:

1. ¿El análisis wavelet es capaz de detectar puntos de cambio en media o varianza en una serie temporal?
2. ¿El análisis wavelet tiene un mejor desempeño detectando puntos de cambio en media o varianza que los métodos tradicionales?

## 1.8. Objetivos

A continuación se presentan los objetivos de la investigación.

### 1.8.1. Objetivo general

- Evaluar el desempeño de la transformada wavelet continua (TWC) para la detección de puntos de cambio en series de tiempo.

### 1.8.2. Objetivos específicos

- Estudiar el uso de la transformada wavelet continua en series de tiempo.
- Investigar metodologías clásicas de detección de puntos de cambio en series de tiempo.
- Describir el uso del análisis wavelet en la detección de puntos de cambio en series de tiempo.
- Examinar el desempeño del análisis wavelet en la detección de puntos de cambio en la media y varianza introducidos en modelos AR, MA, ARMA, ARIMA y GARCH a través de simulaciones y compararlos con los tests tradicionales: MPQ, Chow, CUSUM y MOSUM.
- Implementar la detección de puntos de cambio en datos reales con métodos tradicionales y análisis wavelet.

## 1.9. Hipótesis

Las hipótesis que dan respuesta a las preguntas de investigación son las siguientes:

- La transformada wavelet continua (TWC) es el mejor método para detectar puntos de cambio en varianza.
- El análisis wavelet detecta puntos de cambio en media y varianza en una serie temporal.

Esta investigación consta de 6 capítulos, el primer capítulo se compone de los estudios realizados para la detección de puntos de cambio tanto en series temporales como en el análisis wavelet, el capítulo 2 presenta el contenido relacionado a la teoría de series temporales, el capítulo 3 presenta las bases teóricas del análisis wavelet, en el capítulo 4 se realiza la metodología aplicando la teoría wavelet y de series temporales para medir su desempeño en la detección de puntos de cambio, el capítulo 5 presenta la aplicación de los tests tradicionales y wavelet en los conjuntos de datos a evaluar el desempeño de cada uno de los tests. Por último, el capítulo 6 presenta el análisis de los resultados de los tests en las simulaciones y en las aplicaciones a los conjuntos de datos para determinar si se cumplen con las expectativas de esta investigación.

# Capítulo 2

## Marco teórico

En este capítulo se presentan los aportes científicos de [Valera \(2020\)](#) en las secciones 2.1 a 2.5 y los conocimientos teóricos de [Tsay \(2001\)](#) en las secciones 2.5 y 2.6.

### 2.1. Series de tiempo

El análisis de series de tiempo se fue desarrollando rápidamente gracias a las bases teóricas de Box-Jenkins y su influencia ha contribuido en los estudios científicos de mayor dificultad como el análisis espacial ([Valera, 2020](#)). En otras áreas como la economía, las empresas necesitan los mejores pronósticos para llevar a cabo la toma de decisiones sólidas. Los mejores pronósticos son aquellos que se ajusten bien a un modelo ([Singh, 2015](#)). Consecuentemente, se establece una relación en la cual se plantean estrategias para la toma de decisiones en base a las predicciones en las series de tiempo ([Valera, 2020](#)).

### 2.2. Componentes de las series de tiempo

El análisis clásico de las series temporales entrega indicios acerca de los valores que la variable puede llegar a tomar, con esa información es posible desarrollar modelos que brinden las características sobre los datos observados, de esta manera se puede establecer una relación entre la teoría y la muestra ([Valera, 2020](#)). Es posible detectar 4 componentes que son parte de las series temporales:

- (a). **Tendencia:** Representa el cambio a largo plazo de la serie de tiempo, se pueden definir como el cambio que pueda tener en la media y se puede observar los movimientos suaves o abruptos de la serie en un periodo de tiempo suficientemente extenso.
- (b). **Estacionalidad:** En la mayoría de las series temporales se pueden presentar variaciones recurrentes en ciertos períodos de tiempo. Este fenómeno no es complicado entender y de ser necesario se puede separar de la serie, a este suceso se le conoce como desestacionalización de la serie.
- (c). **Periodicidad:** Son variaciones que con frecuencia se manifiestan con una larga duración alrededor de la tendencia a través de los años o meses.
- (d). **Aleatoriedad:** Se define como los cambios aleatorios en los movimientos de los datos los cuales no tienen un patrón lógico reconocible.

## 2.3. Descomposición clásica de series de tiempo

Para obtener un adecuado análisis de series de tiempo, es necesario construir modelos capaces de otorgar una mirada simplificada de la evolución de las series en el transcurso del tiempo. En relación a ello, se afirma que el conjunto de variables  $X_t$  es una función conformada por los componentes de Tendencia  $T_t$ , estacionalidad  $S_t$  y el error  $\varepsilon_t$  (Valera, 2020). Estas componentes se caracterizan por:

- $T_t$ : Corresponde a la Tendencia y realiza la modelización de la serie a largo plazo.
- $S_t$ : Representa la estacionalidad o el comportamiento iterativo de la serie en distintos periodos de la serie de tiempo.
- $\varepsilon_t$ : Es la manifestación de distintos factores y por lo general son desconocidos. Este efecto es la representación de las fluctuaciones de la serie  $X_t$  entorno al promedio  $\mu_t$ .

Los modelos que se construyen para el análisis de series de tiempo a partir de estos componentes son:

- Modelo aditivo: Es ideal utilizarlo cuando la magnitud de las variaciones de las frecuencias de tipo estacional no poseen cambios bruscos. Además, se puede usar cuando la estacionalidad no tiene dependencia atribuida por otro componente. Es decir, son de utilidad en series que posean variación estacional constante, de no ser así, el modelo debe ser multiplicativo. Es posible construir un modelo aditivo mediante operadores.

$$X_t = T_t + S_t + \varepsilon_t .$$

- Modelo Multiplicativo: Se desempeñan mejor cuando la magnitud de las variaciones de las frecuencias estacionales de la serie aumentan y disminuyen gradualmente con las variaciones de la tendencia y se representa por la siguiente ecuación:

$$X_t = T_t \times S_t \times \varepsilon_t$$

## 2.4. Procesos Estocásticos

Un proceso estocástico se define como un conjunto de variables aleatorias,  $\{X_t\}_{t \in T}$  que van evolucionando en el tiempo,  $t \in T$  con su respectiva distribución de probabilidad, donde  $T$  puede ser un subconjunto de los números enteros o un intervalo o todos los reales (Valera, 2020).

## 2.5. Procesos estacionarios

Según Tsay (2001) las series de tiempo pueden presentar uno de los 2 tipos de estacionariedad. En el primer caso se considera un proceso estocástico **fuertemente estacionario o estrictamente estacionario** si la distribución conjunta  $(X_{t_1}, \dots, X_{t_k})$  es igual en distribución a  $(X_{t_1+t}, \dots, X_{t_k+t})$  para todo  $t$ , donde  $k$  es un entero positivo y  $(t_1, \dots, t_k)$  es un conjunto de  $k$  enteros positivos. Es decir, el supuesto de estacionariedad estricta se cumple si y sólo si la distribución conjunta no varía en el tiempo. Sin embargo, la condición de estacionariedad fuerte es muy difícil de demostrar empíricamente. Por lo tanto, se asume un nivel más débil de estacionariedad. Para que un proceso estocástico se considere **estacionario débil o débilmente estacionario** este debe tener una media y varianza constante en el tiempo y la covarianza entre 2 periodos depende exclusivamente de la distancia entre estos periodos de tiempo y no del tiempo en que se calculó la covarianza.

Las propiedades que conforman un proceso débilmente estacionario son:

- $E(X_t) = \mu$ , para todo  $t \in T$ .
- $V(X_t) < \infty$ , para todo  $t \in T$ .
- $Cov(X_t, X_{t+z}) = \gamma_z$ , depende sólo de  $z$ , (donde  $\gamma_z$  se define en la definición 2.5.1)

La última condición se puede analizar de la siguiente manera:

**Definición 2.5.1** Sea  $X_t$  un proceso débilmente estacionario. Entonces

- $\gamma_X(u, w) = \gamma_X(u - w, 0)$  para todo  $u, w \in T$ .
- $\gamma_z = \gamma_X(z, 0) = Cov(X_{t+z}, X_t)$ , para todo  $t \in T$  y  $z \in \mathbb{Z}$ .

de este modo, la función de autocovarianza  $\gamma_X$  se define como la función que solo depende de las diferencias entre  $(u - w) = z$ , el cual es denominado **retardo de z** (Valera, 2020). Por lo tanto, se desprenden 3 propiedades inherentes de la covarianza.

**Definición 2.5.2** Sea  $\gamma$  la función de autocovarianza de un proceso estacionario  $\{X_t, t \in T\}$ . Entonces:

- $\gamma_0(X_t) = V(X_t)$ .
- $|\gamma_z| \leq \gamma_0$ .
- $\gamma_z = \gamma_{-z}$ .

En el área de las finanzas, es común afirmar que una serie de rendimientos de activos sea un proceso estocástico estacionario débil. Para comprobar empíricamente este supuesto se necesita de un número indicado de rendimientos históricos.

La serie  $X_t$  es ruido blanco si su conjunto de variables aleatorias son independientes e idénticamente distribuidas con media 0 y varianza finita. También es conocido como ruido blanco Gaussiano ( $N(0, \sigma^2)$ ).

La **función de autocorrelación** se utiliza para una serie de tiempo débilmente estacionaria con dependencia lineal entre 2 rezagos en estudio

$$\rho_z = \frac{Cov(X_{t+z}, X_t)}{\sqrt{V(X_t)V(X_{t+z})}} = \frac{Cov(X_t, X_{t+z})}{V(X_t)} = \frac{\gamma_z}{\gamma_0}$$

Utilizando la propiedad  $V(x_{t+z}) = V(x_t)$  desde una serie débilmente estacionaria se puede utilizar la propiedad,  $\rho_0 = 1$  y  $-1 \leq \rho_z \leq 1$  y  $\rho_z = \rho_{-z}$ . Adicionalmente, la serie  $X_t$  es débilmente estacionaria y es no correlacionada si y sólo si  $\rho_z = 0$  para todo  $z > 0$ .

La **función de autocorrelación parcial** calcula la correlación entre 2 variables aleatorias que se encuentran separadas  $z$  períodos de tiempo y sin tener en cuenta la dependencia impuesta por los retardos que se encuentran en medio de esas 2 variables. La definición formal es:

$$\varphi_{zz} = Corr(X_t, X_{t-z}/X_{t-1}, X_{t-2}, \dots, X_{t-z+1})$$

En las siguientes subsecciones se van a describir diferentes tipos de procesos estacionarios. Tales como: Modelo AR(p), MA(q), ARMA (p,q), ARIMA (p,i,q), ARCH(p) y GARCH(p,q). Todos ellos son considerados como procesos estacionarios débiles.

### 2.5.1. Modelos autoregresivos (AR) de orden p

Uno de los modelos más frecuentes en el análisis de series temporales es el modelo autoregresivo. Un modelo autoregresivo de orden  $p$  se define como:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t, \quad (2.1)$$

A partir de la ecuación (2.1) es posible determinar su media :

$$E(X_t) = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p},$$

donde  $p$  es un entero no negativo,  $a_t$  es ruido blanco  $N(0, \sigma_a^2)$ , y  $\phi_i$  son los parámetros del modelo para  $i = 0, \dots, p$ . Este modelo de serie temporal comparte una estructura semejante a los modelos de regresión lineal simple, donde  $X_t$  es la variable respuesta o dependiente y  $X_{t-1}, \dots, X_{t-k}$  son las variables regresoras o independientes. Un modelo autoregresivo de orden 1 está definido por un valor de  $p = 1$  y se escribe de la siguiente manera:

$$X_t = \phi_0 + \phi_1 X_{t-1} + a_t. \quad (2.2)$$

Tsay (2001) postula que, bajo el supuesto de estacionariedad débil de un modelo AR(1) se tiene  $E(X_t) = \mu$ ,  $V(X_t) = \gamma_0$ ,  $E(a_t) = 0$  y  $Cov(X_t, X_{t-j}) = \gamma_j$ , donde  $\mu$  y  $\gamma_0$  son constantes. Además,  $\gamma_j$  está en función de  $j$  y no de  $t$ .

$$E(X_t | X_{t-1}) = \phi_0 + \phi_1 X_{t-1},$$

$$E(X_t) = \phi_0 + \phi_1 E(X_{t-1}),$$

$$\mu = \phi_0 + \phi_1 \mu,$$

$$E(X_t) = \mu = \frac{\phi_0}{1 - \phi_1}. \quad (2.3)$$

A raíz de este resultado se tienen las siguientes implicaciones en  $X_t$ . En primer lugar, la media de  $X_t$  existe si  $\phi_1 \neq 1$ . En segundo lugar, la media  $X_t$  es 0 si y solo si  $\phi_0 = 0$ .

Usando  $\phi_0 = (1 - \phi_1)\mu$  se tiene el siguiente modelo AR(1):

$$X_t - \mu = \phi_1 (X_{t-1} - \mu) + a_t. \quad (2.4)$$

Sustituyendo la expresión (2.4) de forma generalizada, se tiene la siguiente formula:

$$X_t - \mu = a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \dots + \phi_1^n a_{t-k} \quad ; n \geq 0 \quad ; k = 0, 1, \dots, T$$

$$X_t - \mu = \sum_{i=0}^{\infty} \phi_1^i a_{t-i}. \quad (2.5)$$

Entonces,  $X_t - \mu$  se identifica como una función lineal de  $a_{t-i}$  para  $i \geq 0$ . Por ello, es posible utilizar esta propiedad además de la propiedad de independencia en la serie ruido blanco  $\{a_t\}$ ,

$$E[(X_t - \mu)a_{t+1}] = 0.$$

En relación con este resultado se puede entender el hecho de que  $X_{t-1}$  está sucediendo en el tiempo  $t$  y no se encuentra sujeta a la información rezagada.

Utilizando la ecuación (2.2) se puede determinar su varianza:

$$V(X_t) = \phi_1^2 V(X_{t-1}) + \sigma_a^2, \quad (2.6)$$

donde  $\sigma_a^2$  está relacionada a la varianza de la serie  $a_t$  y teniendo en cuenta que la covarianza entre  $X_{t-1}$  y  $a_t$  es 0. Además, bajo el supuesto de estacionariedad se obtiene el siguiente resultado a partir de la ecuación (2.3):

$$V(X_t) = \frac{\sigma_a^2}{1 - \phi_1^2}.$$

La varianza de  $X_t$  es válida si y solo si  $\phi_1^2 < 1$ . Esta condición es justificada, dado que la varianza de una variable aleatoria  $\{X_t\}$  es acotada y no negativa. A partir de la idea anterior, bajo el supuesto de estacionariedad débil de un modelo AR(1) se tiene que  $-1 < \phi_1 < 1$ , es decir, la condición que es suficiente para que un modelo AR(1) a partir de la ecuación (2.3) sea un proceso débilmente estacionario es  $|\phi_1| < 1$ .

### 2.5.2. Modelo de medias móviles (MA) de orden p

Para algunos especialistas los modelos de Medias móviles (MA) son considerados como una extensión de una serie de ruido blanco o también como un modelo AR de orden infinito con ciertas limitaciones en sus parámetros. Para este caso es recomendable utilizar la segunda propuesta, por ende, la ecuación correspondiente a un proceso MA(q) es el siguiente:

$$X_t = c_0 + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (2.7)$$

donde  $c_0$  es una constante,  $a_t$  es una serie ruido blanco y  $\theta_1, \dots, \theta_q$  son constantes reales del modelo con  $q > 0$ .

Los modelos MA(q) poseen la característica de ser débilmente estacionarios por naturaleza, esto se debe a que son combinaciones lineales finitas de una sucesión de ruido blanco.

Utilizando la ecuación (2.7) se puede determinar la media del modelo MA(q) que se define a continuación:

$$E(X_t) = c_0$$

Del mismo modo, a través de la ecuación (2.7) se establece la varianza de un modelo MA(q):

$$V(X_t) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\sigma_a^2$$

### 2.5.3. Modelos ARMA de orden (p,q)

Los modelos ARMA son considerados como una conjunción de propiedades de los modelos AR y MA en un modo compacto para que la cantidad de parámetros a usar sea la menor.

Un modelo ARMA(p,q) se define de la siguiente manera:

$$X_t = \phi_0 + \sum_{i=1}^p \phi_i X_{t-i} + a_t - \sum_{i=1}^q \theta_i a_{t-i}, \quad (2.8)$$

donde  $\{a_t\}$  es definida como una serie ruido blanco,  $(p, q)$  son enteros no negativos,  $(\phi_0, \dots, \phi_p), (\theta_0, \dots, \theta_q)$  son parámetros del modelo.

Otra manera de describir la ecuación (2.8) utilizando el operador de diferencias B se presenta a continuación:

$$(1 - \phi_0 B - \dots - \phi_p B^p)X_t = \phi_0 + (1 - \theta_0 B - \dots - \theta_q B^q)a_t. \quad (2.9)$$

El polinomio obtenido al lado izquierdo de la igualdad  $(1 - \phi_0 B - \dots - \phi_p B^p)$  corresponde a un polinomio de un modelo AR y, por otro lado, el polinomio  $(1 - \theta_0 B - \dots - \theta_q B^q)$  pertenece al polinomio modelo MA, donde  $BX_t = X_t - X_{t-1}$ . Para estos modelos se exige que no existan factores comunes entre los polinomios, ya que se puede reducir considerablemente el orden (p,q) del modelo. Bajo otro punto de vista, cuando un AR es puro, el polinomio AR involucra la ecuación característica de un modelo ARMA. Además, si todas las soluciones de una ecuación característica tienen valores menores a 1 en valor absoluto, entonces el modelo ARMA se puede considerar como débilmente estacionario. Siguiendo con la idea anterior, la media incondicional del modelo ARMA es la siguiente:

$$E(X_t) = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p}. \quad (2.10)$$

#### 2.5.4. Procesos de memoria larga ARIMA

Los modelos ARIMA poseen una fuerte memoria dado que los coeficientes en el modelo MA no decaen rápidamente en el tiempo a cero, esto conlleva a que el valor pasado en  $a_{i-1}$  del modelo no se desvanezca en la serie. Para contrarrestar esta situación de no estacionariedad por raíces unitarias es recomendable el uso de la diferenciación.

Una serie  $z_t$  se considera un proceso ARIMA (p,1,q) si el cambio en la serie  $c_t = z_t - z_{t-1} = (1 - B)z_t$  pertenece a un modelo ARMA(p,q) estacionario e invertible. En resumen, una serie no estacionaria se puede convertir en una serie estacionaria a través de  $c_t = z_t - z_{t-1}$  que se denomina como la primera diferencia de la serie  $z_t$ . Sin embargo, es posible que la serie  $z_t$  posea más de 1 raíz unitaria y, por lo tanto, se deben realizar múltiples diferenciaciones para transformar la serie a estacionaria.

Existen diferentes maneras de estacionarizar un modelo con raíz unitaria, algunos se complementan con otros para dar una mayor solidez al modelo de interés.

#### 2.5.5. Diferenciación en estacionalidad

El operador que diferencia en estacionalidad una serie de tiempo con estacionalidad  $z_t$  con  $s$  períodos es definido como:

$$\Delta_s z_t = z_t - z_{t-s} = (1 - B^s)z_t.$$

En algunas situaciones es posible utilizar la transformación logarítmica para que la serie sea estacionaria. Este operador es usado con mayor frecuencia en el análisis de series financieras y económicas. Por lo general, se utiliza la transformación logarítmica bajo 2 escenarios. El primero, se utiliza para mantener bajo control el crecimiento exponencial de la serie y el segundo, es utilizado para estabilizar la variabilidad de la serie.

#### 2.5.6. Modelo multiplicativo estacional

Un hecho importante en las series temporales es el de analizar su comportamiento estacional y su carácter estacionario, al detectarse un aumento exponencial desenfrenado y las raíces unitarias, se puede evidenciar que el modelo (ARIMA) posee correlaciones entre los pares de observaciones distanciadas por periodos estacionales bajo el supuesto de que la serie de ruido blanco  $\{a_t\}$  en la parte estacional esté correlacionado. A partir de lo anteriormente descrito se puede escribir su modelo como:

$$(1 - B^s)(1 - B)z_t = (1 - \theta B)(1 - \Theta B^s)a_t \quad (2.11)$$

donde  $s$  es la periodicidad de la serie,  $a_t$  es una serie con características de ruido blanco, con  $|\theta| < 1$  y  $|\Theta| < 1$ .

### 2.6. Modelos heterocedásticos condicionales

En econometría existen modelos que son especialmente utilizados para elaborar pronósticos de la volatilidad de las acciones o de activos financieros cuyo especial interés es el rendimiento o ganancia en el corto o largo plazo. Estos son los modelos heterocedásticos condicionales, en estas secciones se van a describir los modelos ARCH(p) y los modelos GARCH(p,q).

### 2.6.1. Modelo ARCH(p)

El modelo ARCH (*Autoregressive conditionally heterosedastic*) fue propuesto por Engle (1982) y es un modelo econométrico utilizado en el estudio de la evolución de la volatilidad ( $\sigma_t^2$ ) de activos financieros, cuyas principales características son:

- (a)  $g_t$  es el rendimiento de un activo que tiene media ajustada y no es correlacionada en serie, pero si es dependiente.
- (b) La dependencia de  $g_t$  se analiza en los valores de los rezagos de una función cuadrática simple.

Un modelo ARCH de orden p se define como:

$$g_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 g_{t-1}^2 + \dots + \alpha_p g_{t-p}^2, \quad (2.12)$$

donde  $\{\epsilon_t\}$  es un conjunto de variables aleatorias independientes e idénticamente distribuidas (iid) con media igual a 0 y varianza igual a 1,  $\alpha_0 > 0$ , y  $\alpha_i \geq 0$  para  $i > 0$ . Además, los coeficientes  $\alpha_i$  deben cumplir con ciertas condiciones de regularidad para que la varianza incondicional de  $g_t$  sea finita. Por lo general, se asume que  $\epsilon$  sigue una distribución t-Student estandarizada o en algunas ocasiones sigue una distribución normal estándar.

Supongamos que se tiene un modelo ARCH(1)

$$g_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 g_{t-1}^2,$$

donde  $\alpha_0 > 0$  y  $\alpha_1 \geq 0$ . La media no condicionada de  $g_t$  se mantiene en cero debido a

$$E(g_t) = E[E(g_t|F_{t-1})] = E[\sigma_t E(\epsilon_t)] = 0,$$

donde  $F_{t-1}$  denota la información obtenida del conjunto en el momento t-1. La varianza no condicionada de  $g_t$  se describe como

$$V(g_t) = E(g_t^2) = E[E(\sigma_t^2|F_{t-1})] = E(\alpha_0 + \alpha_1 g_{t-1}^2) = \alpha_0 + \alpha_1 E(g_{t-1}^2).$$

En síntesis,  $g_t$  es un proceso estacionario cuyo valor esperado es  $E(g_t) = 0$  y su varianza correspondiente es  $V(g_t) = V(g_{t-1}) = E(g_{t-1}^2)$ . En adición, la varianza de  $g_t$  posee la siguiente característica,  $V(g_t) = \alpha_0 + \alpha_1 V(g_t)$  y  $V(g_t) = \alpha_0 / (1 - \alpha_1)$ , ya que la varianza de  $g_t$  debe estar en los reales positivos  $0 \leq \alpha_1 < 1$ .

### 2.6.2. Modelo GARCH(p,q)

El modelo GARCH fue propuesto por Bollerslev(1986) y son una extensión generalizada de un modelo ARCH. Sea  $g_t = x_t - \mu_t$  una media ajustada de retornos logarítmicos. Por consiguiente,  $g_t$  se describe en el modelo GARCH de orden (p,q) como

$$g_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i g_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (2.13)$$

donde  $(\{\epsilon_t\}, \alpha_0, \alpha_i)$  posee las mismas características que en el modelo ARCH (p), ver en la ec(2.12),  $\beta_j \geq 0$  y para que el proceso sea estacionario debe cumplir con  $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$ .

En el siguiente capítulo se presentan las bases teóricas del análisis wavelet y el método que se utiliza para la detección de puntos de cambio.

# Capítulo 3

## Wavelet

Las wavelets forman un conjunto de herramientas de procesamiento de señales, cuya función es analizar en diferentes escalas de tiempo las propiedades locales de señales complejas donde no se cumpla el supuesto de estacionariedad. Son el pilar de nuevos métodos de análisis de señales y se pueden encontrar en distintos problemas como la compresión de datos y la eliminación de ruido (Misiti, Misiti, Oppenheim, y Poggi, 2007). El estudio del análisis wavelet ha surgido lentamente en el área de las matemáticas aplicadas (Daubechies, 1992). Son muchos los investigadores que han aportado al avance científico de los wavelet desde la teoría hasta la práctica. Los más destacados del siglo *XX* son Daubechies (1988), Donoho (1993), Chui(1992), Donoho y Johnstone(1994) y Silverman (1999) (Vega, 2003). Su aplicación se ha expandido a distintas áreas, tales como, astrofísica, geofísica, medicina, telecomunicaciones, entre otras (Misiti y cols., 2007).

El concepto de estudiar una señal como una superposición de funciones se remonta a principios del siglo *XIX*. Fue Joseph Fourier quien dedujo que al superponer una suma de funciones de frecuencias seno y coseno se podrían formar diversos tipos de funciones. Esta suma de funciones se conoce como desarrollo de Fourier el cual entrega una descomposición en componentes de frecuencia. Estas componentes añaden información de cada una de las frecuencias hacia la variabilidad de la función (Vega, 2003).

Si bien la teoría de Fourier ha logrado resolver problemas de diversa complejidad, su aplicación ha perdido efectividad a través del tiempo ya que las limitaciones de la teoría lo vuelven poco eficiente en la práctica. Por otro lado, el análisis wavelet es mucho más versátil y no es penalizado por las limitaciones que posee el desarrollo de Fourier (Vega, 2003). El análisis wavelet realiza una descomposición de componentes que entregan información acerca de los dominios de frecuencia y de tiempo. En resumen, el análisis wavelet es más confiable para describir los cambios de patrones en el dominio de frecuencias de las funciones a través del tiempo (Vega, 2003). Según Misiti y cols. (2007) se pueden distinguir 3 etapas propias de las características de las wavelets, estas se separan en:

- *Análisis*: Un wavelet es una función que va oscilando como una onda pero que velozmente se amortigua en 0 en distintos períodos de tiempo. Al estar bien posicionada en tiempo y frecuencia, se establece una familia de funciones de análisis por traslación en el tiempo y dilatación en escala. De esta manera, las wavelets conforman un “acercamiento” matemático para estudiar en sincronía las propiedades de una señal en distintas escalas de tiempo.

- *Eliminación de ruido:* Se refiere a la capacidad de reconstruir la señal útil al mismo tiempo que observamos una representación ruidosa de la señal. El método consiste en obtener la transformada wavelet de las observaciones de toda la señal, para después amoldar cuidadosamente los coeficientes usando al máximo las características locales e invertir la transformación.
- *Compresión:* La compresión se utiliza para disminuir la cantidad de volumen de información que se está codificando. Esta característica surge de la capacidad de las wavelets para compactar la energía de la señal en los coeficientes que son significativamente diferentes de 0.

### 3.1. Teoría Wavelet

Desde la sección 3.1 hasta la sección 3.3 se utilizarán las bases teóricas implementadas por [Misiti y cols. \(2007\)](#). Además, en la sección 3.4 se presenta la técnica para la detección de puntos de cambio por [Chen y Gupta \(2001\)](#).

Sea  $L^1(\mathbb{R})$  o  $L^1$  el espacio de funciones integrables sobre los reales y  $L^2(\mathbb{R})$  o  $L^2$  el espacio integrable de funciones cuadradas de energía finita sobre los reales.

La norma elevada al cuadrado  $\int_{\mathbb{R}} |s(t)|^2 dt$  se denomina energía de la señal  $s$ .

Sea  $\hat{\psi}(\omega)$  la transformada de Fourier:

$$\hat{\psi}(\omega) = \int_{\mathbb{R}} \psi(t) e^{-2i\pi\omega t} dt. \quad (3.1)$$

Sea  $\psi$  una función suficientemente regular y bien localizada. La función  $\psi \in L^1 \cap L^2$  se llama wavelet si cumple la condición de admisibilidad en el dominio de la frecuencia:

$$\int_{\mathbb{R}^+} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega = \int_{\mathbb{R}^-} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty, \quad (3.2)$$

donde  $\hat{\psi}$  conforma a la transformada de Fourier de  $\psi$ . Esto conlleva en particular a que la wavelet utilizada tenga integral igual a 0.

Para probar de manera simple la condición de admisibilidad se considera que, para una wavelet  $\psi$  perteneciente a los reales, se tiene que verificar que:

$$\psi \in L^1 \cap L^2, \quad t\psi \in L^1 \quad \text{y} \quad \int_{\mathbb{R}} \psi(t) dt = 0.$$

Para resumir los conceptos vistos hasta ahora, supongamos que por un período de tiempo la wavelet oscila particularmente de manera ondulatoria y es localizada por efecto de un amortiguamiento. En otras palabras, su oscilación está relacionada con la cantidad de momentos iguales a 0 y su localización está relacionada con el intervalo donde adquiere valores que son significativamente distintos de 0.

A partir de una función  $\psi$  utilizando los componentes de dilatación  $a$  y traslación  $b$  se puede construir una familia de funciones que se denominan átomos básicos:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a \in \mathbb{R}^+, b \in \mathbb{R}. \quad (3.3)$$

Para una función  $f$  de energía finita, se obtiene la transformada wavelet continua  $C_f$  de la función  $\psi_{a,b}$  como:

$$C_f(a, b) = \int_{\mathbb{R}} f(t) \overline{\psi_{a,b}(t)} dt \quad a \in \mathbb{R}^+ \quad y \quad b \in \mathbb{R}.$$

Realizar cálculos con la función  $C_f$  es lo mismo que hacer análisis en  $f$  con la transformada wavelet  $\psi$ . Por ende, la función  $f$  es posible describirla a través de los coeficientes de la transformada wavelet continua  $C_f(a, b)$ . Estos coeficientes son utilizados para medir las fluctuaciones de  $f$  en su escala  $a$ . Luego, la tendencia en la escala  $a$  posee un lento decaimiento y se eliminan en  $C_f(a, b)$ . Esto hace posible que el análisis wavelet realice un análisis local de la función  $f$ , además del comportamiento de los efectos de escala entre el  $C_f(a, b)$  y distintos valores de  $a$ . Si por ejemplo suponemos que la función wavelet  $\psi$  es cero en el exterior del intervalo  $[-Z, +Z]$ , entonces  $\psi_{(a,b)}$  va a ser cero en el exterior del intervalo  $[-aZ + b, +aZ + b]$ . Por consiguiente, el valor obtenido por  $C_f(a, b)$  va a depender del valor de  $f$  en la vecindad que está ubicado  $b$  con la longitud que posea  $a$ .

Con respecto a esto, las wavelet poseen ciertas diferencias con el análisis de Fourier, esto se debe a que el valor que posea la transformada de Fourier  $\hat{f}(\omega)$  de  $f$  en un punto  $\omega$  va a depender de los valores que tome  $f$  en todo el conjunto  $\mathbb{R}$ . Por contraparte, los valores grandes que tome  $C_f(a, b)$  van a entregar información en relación a la irregularidad local de  $f$  en el entorno de la posición  $b$  y en la escala de  $a$ . En síntesis, el análisis wavelet es considerado el análisis de las fluctuaciones que posea la función  $f$  en todas las escalas designadas.

Si solamente se requiere utilizar ciertas partes de la función, entonces el uso de un subconjunto discreto es lo indicado. Por ejemplo, si se tiene un subconjunto discreto de  $\mathbb{R}^+$ , y  $\mathbb{R}$ . Se dejan fijos a  $a_0 > 1, b_0 > 0$  y luego se toma a  $a \in \{a_0^h\}_{h \in \mathbb{Z}}$  y  $b \in \{qa_0^h b_0\}_{h, q \in \mathbb{Z}^2}$ .

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a \in \mathbb{R}^+ \neq 0, b \in \mathbb{R}.$$

Para el caso de la transformada wavelet discreta se usará la familia wavelet indexada en  $\mathbb{Z}$ :

$$\psi_{h,q}(t) = a_0^{-\frac{h}{2}} \psi\left(a_0^{-h} t - qb_0\right) \quad a_0 > 1 \quad y \quad b_0 > 0 \quad \text{fijo} \quad y \quad h, q \in \mathbb{Z}^2.$$

Para  $f \in L^2$  se define una transformada wavelet discreta para una función de energía finita  $f$  como:

$$C_f(h, q) = \int_{\mathbb{R}} f(t) \overline{\psi_{h,q}(t)} dt = (f, \psi_{h,q})_{L^2} \quad (h, q) \in \mathbb{Z}^2.$$

Los valores de  $a$  y de  $b$  respectivamente se presentan como:

$$a = 2^j, b = r2^j \quad \text{o} \quad b = ra \quad \text{para} \quad (j, r) \in \mathbb{Z}^2.$$

Para el presente caso y para wavelets que corroboran propiedades más poderosas que la admisibilidad, como lo es el caso ortogonal, se determinan nuevas componentes para estudiar una señal a través de una función escala  $\varphi$  vinculada con  $\psi$  y la función  $\psi$  se dilata y traslada por toda la señal. En resumen, la función escala representada por  $\varphi$  es utilizada para establecer las aproximaciones locales y la función  $\psi$  se utiliza para denotar las fluctuaciones alrededor de las aproximaciones locales como se define a continuación:

$$\psi_{j,r}(y) = 2^{\frac{-j}{2}} \psi(2^{-j}y - r), \quad \text{para } (j,r) \in \mathbb{Z}^2, \quad (3.4)$$

$$\varphi_{j,r}(y) = 2^{\frac{-j}{2}} \varphi(2^{-j}y - r), \quad \text{para } (j,r) \in \mathbb{Z}^2. \quad (3.5)$$

Los coeficientes wavelet de una señal,  $s(t)$ , están definidos por

$$\alpha_{j,r} = \int_{\mathbb{R}} s(t) \psi_{j,r}(t) dt, \quad (3.6)$$

y, cumpliendo ciertas propiedades, los coeficientes obtenidos pueden reconstruir la señal de la siguiente manera:

$$s(t) = \sum_{j \in \mathbb{Z}} \sum_{r \in \mathbb{Z}} \alpha_{j,r} \psi_{j,r}(t). \quad (3.7)$$

dada la existencia de una función  $\psi$  tal que la familia de funciones  $\{\psi_{j,r}\}_{(j,r) \in \mathbb{Z}^2}$  se encuentra relacionada con las descripciones del análisis multiresolución (AM) que se describe en la siguiente sección.

## 3.2. Análisis Multiresolución

El análisis multiresolución (AM), se utiliza para descomponer una señal  $s(t)$  en un espacio vectorial de funciones como la suma de subespacios bajo ciertas propiedades, de tal forma que cualquier función se describe como una serie constituida por sus proyecciones en aquellos subespacios.

Un AM de  $L^2(\mathbb{R})$  se define como una familia de subespacios de señales de energía finita,  $\{V_j\}_{j \in \mathbb{Z}}$ , que son cerrados anidados, es decir:

- (a)  $V_j$  es un subespacio cerrado de  $L^2$ ,
- (b)  $V_j \subset V_{j+1}$ ,
- (c)  $\overline{\cup_{j \in \mathbb{Z}} V_j} = L^2$  y  $\cap_{j \in \mathbb{Z}} V_j = \{0\}$ .

Cada subespacio,  $V_j$  proviene del espacio de origen (o espacio central)  $V_0$  por contracción (para  $j < 0$ ) o por dilatación (para  $j > 0$ ) de sus funciones, en otras palabras, las funciones de  $V_j$  se obtienen por dilatación o contracción por diádicos ( $2^j$ ) de funciones que pertenecen al espacio único:

$$f(t) \in V_j \leftrightarrow f(2t) \in V_{j+1} \quad \forall j \in \mathbb{Z}$$

Los subespacios  $\{V_j\}_{j \in \mathbb{Z}}$  del AM forman los espacios de aproximación. Por otro lado, se utilizan los subespacios de detalle denominados como  $\{W_j\}_{j \in \mathbb{Z}}$ . Para  $j \in \mathbb{Z}$  fijo, se tiene que  $W_j$  es el complemento ortogonal del subespacio  $V_j$  en  $V_{j-1}$ :

$$V_{j+1} = V_j \oplus W_j \quad \text{para } j \in \mathbb{Z}$$

Las traslaciones enteras de  $\varphi$ ,  $\{\varphi_{j,r}\}_{r \in \mathbb{Z}}$  van creando los subespacios  $V_j$ . De manera similar, las traslaciones enteras de  $\psi$ ,  $\{\psi_{j,r}\}_{r \in \mathbb{Z}}$  son las responsables de crear a los  $W_j$ . Como  $L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$ , cualquier función de energía finita perteneciente a  $\mathbb{R}$  se describe como la suma de todos los detalles pertenecientes a la señal y  $\{\psi_{j,r}\}_{(j,r) \in \mathbb{Z}^2}$  conforman una base wavelet ortonormal de  $L^2(\mathbb{R})$ . En otras palabras, los coeficientes  $\alpha_{j,r}$  están relacionados a los  $\psi_{j,r}$  en las denominadas proyecciones ortogonales de la señal  $s(t)$  sobre  $W_j$ .

Recopilando los conceptos vistos anteriormente, se puede organizar las wavelets usando 2 parámetros:

- $r$ , el tiempo cuya función es trasladar los segmentos para un nivel en específico,
- $2^j$  la escala cuya función es cambiar de un nivel  $j$  a un nivel menor.

La función escala o padre wavelet ( $\varphi$ ) y la función wavelet o madre wavelet ( $\psi$ ) cumplen funciones importantes por separado en la descomposición de las señales. En primer lugar, las funciones de escala son usadas para determinar el valor de los coeficientes de aproximación  $\beta_{j,r} = \int_{\mathbb{R}} s(t) \varphi_{j,r}(t) dt$ , los cuales entregan los promedios locales de la señal  $s(t)$ . En consecuencia, las aproximaciones de la señal se describen como,  $A_j(t) = \sum_{r \in \mathbb{Z}} \beta_{j,r} \varphi_{j,r}(t)$ . En segundo lugar, las funciones wavelet son usadas para determinar el valor de los coeficientes wavelets  $\alpha_{j,r} = \int_{\mathbb{R}} s(t) \psi_{j,r}(t) dt$ , los cuales se calculan a partir de la diferencia de 2 promedios locales sucesivos. Estos son denominados como detalles de la señal  $s(t)$ ,  $D_j(t) = \sum_{r \in \mathbb{Z}} \alpha_{j,r} \psi_{j,r}(t)$ .

### 3.3. Familia wavelet

El análisis wavelet tiene muchas más libertades que el análisis de Fourier, ya que uno puede elegir sus átomos básicos. Más aún, bajo las características del procesamiento de wavelets, es mejor utilizar la transformada wavelet continua en lugar de la transformada wavelet discreta, si la redundancia es de utilidad para poder analizar la señal. Por el contrario, será mejor utilizar las transformadas wavelet discretas si el objetivo es la compresión de la señal. En el último caso, para la TWD, se utilizarán wavelets con filtros, mientras que en el primer caso, para la transformada wavelet continua la mayoría de las funciones que integren cero son óptimas ([Misiti y cols., 2007](#)).

Existen diversos tipos de familias wavelet los cuales poseen diferentes propiedades como:

- el soporte de las funciones madre wavelet  $\psi$ ,  $\hat{\psi}$  (y también en padre wavelet  $\varphi$ ,  $\hat{\varphi}$ ) o la velocidad en que converge a cero en las funciones  $\psi(t)$  y  $\hat{\psi}(\omega)$  cuando el tiempo  $t$  o la frecuencia  $\omega$  crece hacia infinito, lo que permite cuantificar la localización wavelet en tiempo y en frecuencia.
- La simetría, en el análisis de señales es muy útil para no considerar las perturbaciones.
- La cantidad de momentos en que se hace cero en la madre wavelet  $\psi$  o padre wavelet  $\varphi$ , son utilizados para entender su compresión.
- La regularidad, es de mucha utilidad para conseguir imágenes o señales que son reestructuradas uniformes y regulares.

En la literatura y en la práctica se pueden encontrar diversos tipos de señales que comparten características similares con las familias wavelets, las más utilizadas son:

- Wavelet Haar

La familia wavelet Haar es una de las wavelets más antiguas y simples, es utilizada frecuentemente en muchas investigaciones, es discontinua, tiene forma cuadrada, ver la madre wavelet Haar en la 3.1, y se define como:

$$\psi(x) = 1 \quad \text{si } x \in [0, 0.5[,$$

$$\psi(x) = -1 \quad \text{si } x \in [0.5, 1[,$$

$$\psi(x) = 0 \quad \text{e.o.c.}$$

Para la función escala  $\varphi$ (padre wavelet) se define de la siguiente manera:

$$\varphi(x) = 1 \quad \text{si } x \in [0, 1[,$$

$$\varphi(x) = 0 \quad \text{e.o.c.}$$

Una vez obtenido su escala y traslación se puede construir la wavelet Haar como se observa en la figura 3.1.

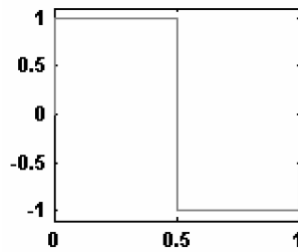


Figura 3.1: Ejemplo de una wavelet Haar.

**Fuente:** Wavelets and their Applications.

Otro tipo de wavelet muy utilizada es la familia wavelet de Daubechies donde la wavelet Haar también forma parte de esta familia.

- Wavelets Daubechies

La familia wavelet Daubechies de un solo parámetro son de soporte compacto, regularidad arbitraria y posee las siguientes características:

- La longitud de soporte de la wavelet madre  $\psi$  y wavelet padre  $\varphi$  es de  $2N - 1$ , donde  $N$  es la cantidad de momentos en que la wavelet madre  $\psi$  es cero.
- Las wavelets de Daubechies son asimétricas (Sobre todo para aquellos valores bajos de  $N$ ) pero la wavelet Haar es un caso especial, es la *db1*.

La madre wavelet Daubechies de orden  $N = 2$  hasta  $N = 10$  se observa en la figura 3.2.

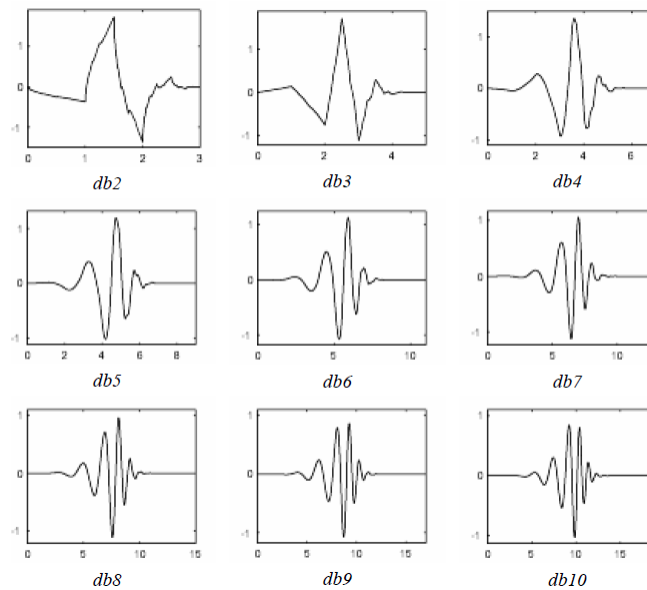


Figura 3.2: Se presenta una wavelet Daubechies para las familias db2-db10.

**Fuente:** Wavelets and their Applications.

En algunas situaciones la capacidad de análisis local y de precisión de las funciones en tiempo y en escala son confusas. Es por esto que se da un mayor enfoque en wavelets que cumplen con la condición de admisibilidad muy débil. Algunos de ellos son:

- Wavelet Gorro mexicano

La wavelet Gorro mexicano es una función continua que es proporcional a la segunda derivada de una función de densidad gaussiana, se define como:

$$\psi(x) = \left( \frac{2}{\sqrt{3}} \pi^{-1/4} \right) (1 - x^2) e^{-x^2/2}$$

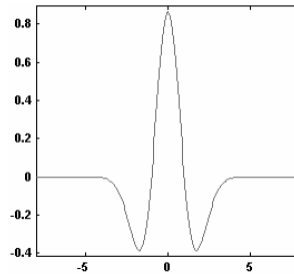


Figura 3.3: Ejemplo de una wavelet gorro mexicano

**Fuente:** Wavelets and their Applications

En la figura 3.3 se puede observar que la wavelet Gorro mexicano posee variaciones de corta duración con frecuencias de alta y baja amplitud.

- Wavelet Morlet

La wavelet Morlet cumple de manera parcial la condición de admisibilidad aunque sea una función que es estándar, se define de la siguiente manera:

$$\psi(x) = C e^{-x^2/2} \cos(5x),$$

donde  $C$  es una constante de normalización que es utilizada para hacer la reconstrucción de señales.

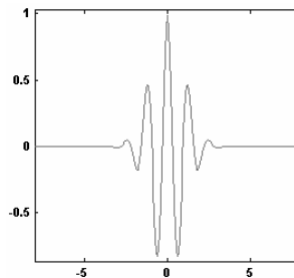


Figura 3.4: Ejemplo de una wavelet Morlet.

**Fuente:** Wavelets and their Applications.

Observando la figura 3.4 se puede apreciar que posee la wavelet de Morlet tiene una mayor cantidad de ondas en su frecuencia con una mayor amplitud que la wavelet Gorro mexicano.

Cada wavelet posee sus ventajas y desventajas en el estudio de señales. Sin embargo, la wavelet gorro mexicano posee una gran ventaja sobre otras wavelets, ya que muchos de los algoritmos utilizados para la detección de picos, poseen una alta tasa de detección de falsos positivos debido al ruido presente en los espectros de espectrometría de masas. Pero, con la transformada wavelet Gorro mexicano, la tasa de detección de falsos positivos disminuye considerablemente, dado su rango de soporte  $[-5, 5]$  y su valor de escala  $a = 1$  la vuelven ideal para detectar picos de baja y alta amplitud (Du, Kibbe, y Lin, 2006).

En la siguiente sección se presentan las bases teóricas para detectar puntos de cambio a través del análisis wavelet realizada por (Chen y Gupta, 2001).

### 3.4. Modelo de detección de puntos de cambio utilizando wavelets

La teoría desarrollada en wavelets ha facilitado la localización de componentes oscilantes a través de la descomposición de las funciones o señales. Por ende se construye un modelo de ruido blanco para combinar los modelos de series temporales y wavelets. El modelo de ruido blanco es el siguiente:

$$Y(dx) = f(x)dx + \tau W(dx) \quad x \in [0, 1], \quad (3.8)$$

donde  $W$  corresponde a un proceso Wiener,  $\tau$  es el parámetro en el nivel de ruido que idealmente es pequeño,  $f$  es una función desconocida con aumentos y caídas agudos.

Los cambios en la amplitud de las funciones se pueden localizar a través de la siguiente definición:

**Definición 3.4.1** *Dada una función  $f$  se dice que posee  $\alpha$ -cúspides en  $x_0$  si existe una constante positiva  $K$  tal que para  $h$  acercándose a 0 por la izquierda o la derecha,*

$$|f(x_0 + h) - f(x_0)| \geq K|h|^\alpha.$$

*Si  $\alpha = 0$ , la función desconocida  $f$  tendrá un salto en  $x_0$ . Para cúspides agudas el valor de  $\alpha$  está en,  $0 \leq \alpha < 1$ .*

La prueba a contrastar esta bajo las siguientes hipótesis:

$$H_0 : f \text{ es diferenciable,}$$

vs

$$H_1 : f \text{ tiene al menos 1 punto de cambio } \alpha - \text{cúspides,}$$

donde  $0 \leq \alpha < 1$ .

Según Chen y Gupta (2001) el modelo apropiado es desarrollado por Wang(1995), definiendo la transformada wavelet de la función desconocida  $f$  como:

$$(C_f)(s, x) = \int_{-\infty}^{\infty} \psi_s(x - u)f(u)du, \quad (3.9)$$

donde  $\psi$  es un wavelet de Daubechies, y

$$\psi_s(x) = \frac{1}{\sqrt{s}}\psi(x/s).$$

La  $(C_f)(s, x)$  es una función que posee una escala  $s$  y una localización espacial en el tiempo  $x$ .

La transformada wavelet del ruido blanco  $W(dx)$  se define en la siguiente ecuación:

$$(C_W)(s, x) = \int_{-\infty}^{\infty} \psi_s(x - u)W du, \quad (3.10)$$

y la transformada wavelet de  $Y$  es,

$$(C_Y)(s, x) = \int_{-\infty}^{\infty} \psi_s(x - u)Y du = (C_f)(s, x) + (C_W)(s, x). \quad (3.11)$$

Además,  $(C_W)(s, x)$  posee las siguientes propiedades:

$(C_W)(s, x)$  es definido como un proceso gaussiano estacionario con media igual a 0 y función de covarianza definidas a continuación:

$$E\{(C_W)(s, x)(C_W)(s, y)\} = \int_{-\infty}^{\infty} \psi_s(x - u)\psi_s(y - u)du, \quad (3.12)$$

y

$$Var\{(C_f(s, x))\} = \int_{-\infty}^{\infty} \psi^2(u)du = 1. \quad (3.13)$$

**Definición 3.4.2** Sea  $s_\tau$  el valor de la escala para el cual, cuando  $f(x)$  es diferenciable, el orden de magnitud de  $C_f(s_\tau, x)$  y  $C_W(s_\tau, x)$  están balanceados. Esto ocurre cuando  $\tau$  tiende a 0,  $\eta > 1$  y  $s_\tau \approx (\tau^2 |\log \tau|^\eta)^{1/(2\alpha+1)}$ .

En la siguiente sección se presenta el estadístico de prueba para el test wavelet y, además, se incorpora un nuevo componente.

### 3.4.1. Identificación del valor crítico aproximado del test y la estimación de discontinuidades y cúspides

A partir de las hipótesis mencionadas en la sección (3.4), bajo  $H_0$  la función  $f$  es diferenciable y, en consecuencia,  $(C_Y)(s_\tau, x)$  tiene el orden de magnitud de  $(C_W)(s_\tau, x)$  y su valor está dominado por ella. Por otro lado, bajo  $H_1$  la función  $f$  posee discontinuidades y/o cúspides, para valores cercanos a  $x$ ,  $(C_Y)(s_\tau, x)$  está dominada por  $(C_f)(s_\tau, x)$ . Además, el valor máximo de la función  $|(C_Y)(s_\tau, x)|$  para  $x \in [0, 1]$  es de orden más grande bajo  $H_1$  que  $H_0$ , y a partir de esto se puede construir un estadístico de prueba definido en el teorema 1:

**Teorema 1** *Bajo  $H_0$ , para  $0 < \gamma < 1$ ,*

$$\lim_{\tau \rightarrow 0} P[\max_{0 \leq x \leq 1} |(C_Y)(s_\tau, x)| \geq C_{\tau, \gamma}] = \gamma,$$

*donde  $C_{\tau, \gamma}$  es un valor crítico aproximado obtenido de la siguiente manera:*

$$C_{\tau, \gamma} = \tau(2|\log s_\tau|)^{-1/2} \cdot [2|\log s_\tau| - \log\{\log \frac{1-\gamma}{2}\} + \log\{[\int_{-\infty}^{\infty} (\psi(u))^2 du]^{1/2}/(2\pi)\}].$$

En el siguiente capítulo se aplica la metodología teórica de los capítulos anteriores a través de los tests clásicos y wavelet para detectar puntos de cambio en los diferentes modelos de series temporales clásicos y de usos más específicos, como en la economía.

## Capítulo 4

# Simulaciones

En este capítulo se realiza una evaluación de la efectividad que poseen los distintos tests para detectar puntos de cambio. Se compararán los test clásicos, el test de múltiples puntos de quiebre (MPQ), el test de Chow, el test CUSUM y el test MOSUM con el test utilizando el análisis wavelet descrito en la sección 3.4 del capítulo 3.

Para evaluar el desempeño de cada test utilizado en esta investigación se desarrolla un estudio de simulación con el diseño factorial  $2^5$  con 2 niveles y 5 factores los cuales se definen en el siguiente cuadro:

Factores	Niveles
A: Puntos de cambio al principio o al final	(+) al principio ; (-) al final
B: Puntos de cambio juntos o separados	(+) juntos ; (-) separados
C: Cambios pequeños o grandes en magnitud	(+) grandes ; (-) pequeños
D: Cambios en la media o en la varianza	(+) media ; (-) varianza
E: Cantidad de puntos de cambio	(+) 4 puntos ; (-) 2 puntos

Cuadro 4.1: Diseño factorial del estudio de simulación para evaluar y comparar el desempeño de los tests para la detección de puntos de cambio.

**Fuente:** Elaboración propia.

Se realizan 32 iteraciones utilizando el software estadístico *Rstudio* combinando los 5 factores bajo diferentes niveles, los cuales empiezan todos en el nivel (-) y al avanzar en las iteraciones se va cambiando el nivel de uno de los factores, este proceso termina cuando todos los factores estén en el mismo nivel (+).

Factores						
Simulación	A	B	C	D	E	tipo de caso
1	-	-	-	-	-	(1)
2	+	-	-	-	-	a
3	-	+	-	-	-	b
4	+	+	-	-	-	ab
5	-	-	+	-	-	c
6	+	-	+	-	-	ac
7	-	+	+	-	-	bc
8	+	+	+	-	-	abc
9	-	-	-	+	-	d
10	+	-	-	+	-	ad
11	-	+	-	+	-	bd
12	+	+	-	+	-	abd
13	-	-	+	+	-	cd
14	+	-	+	+	-	acd
15	-	+	+	+	-	bcd
16	+	+	+	+	-	abcd
17	-	-	-	-	+	e
18	+	-	-	-	+	ae
19	-	+	-	-	+	be
20	+	+	-	-	+	abe
21	-	-	+	-	+	ce
22	+	-	+	-	+	ace
23	-	+	+	-	+	bce
24	+	+	+	-	+	abce
25	-	-	-	+	+	de
26	+	-	-	+	+	ade
27	-	+	-	+	+	bde
28	+	+	-	+	+	abde
29	-	-	+	+	+	cde
30	+	-	+	+	+	acde
31	-	+	+	+	+	bcde
32	+	+	+	+	+	abcde

Cuadro 4.2: Diseño factorial para las simulaciones.

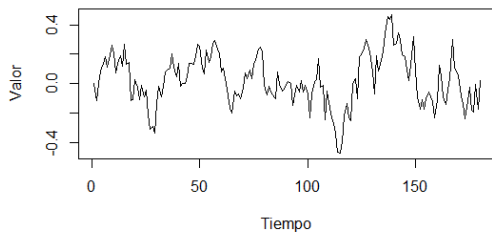
**Fuente:** Elaboración propia.

Este proceso se repite para los modelos que se presentan en la figura 4.1 y los tests deben detectar los puntos de cambio para todos los casos que se presenten.

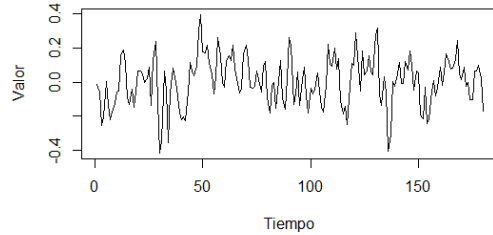
Los modelos clásicos utilizados en las simulaciones son los siguientes:

- Modelo AR(1),  $X_t = 0.8 X_{t-1} + u_t$ ,  $u_t$  es ruido blanco con  $\sigma=0.1$ ,
- Modelo MA(1),  $X_t = -0.8 u_{t-1} + u_t$ ,  $u_t$  es ruido blanco con  $\sigma=0.1$ ,
- Modelo ARMA(1,1),  $X_t = 0.9X_{t-1} + 0.8u_{t-1} + u_t$ ,  $u_t$  es ruido blanco con  $\sigma=0.1$ ,
- Modelo ARIMA (1,1,1),  $X_t - X_{t-1}$  es ARMA(1,1) con  $\phi_1 = 0.4$ ,  $\theta_1 = 0.8$  y con  $\sigma=0.1$ ,
- Modelo GARCH(1,1),  $\xi_t^2 = 0.1\pi_{t-1}^2 + 0.8\xi_{t-1}^2$

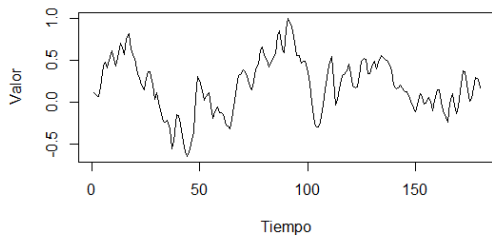
donde  $\xi_t^2$  es la volatilidad implícita lognormal de los rendimientos de la serie,  $\pi_t = \xi_t u_t$  es el rendimiento corregido de la media y  $u_t$  son (iid) con media 0 y varianza 1 (Tsay, 2001).



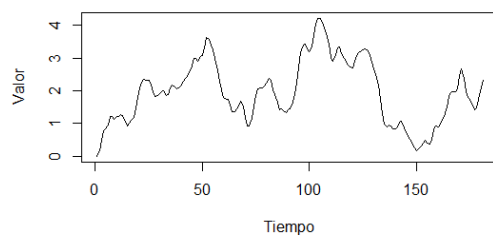
(a) Modelo AR(1)



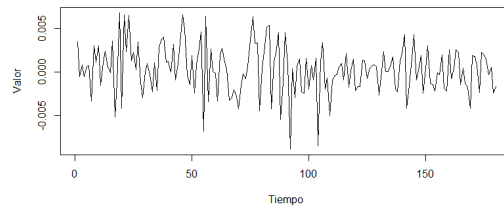
(b) Modelo MA(1)



(c) Modelo ARMA(1,1)



(d) Modelo ARIMA(1,1,1)



(e) Modelo GARCH(1,1).

Figura 4.1: Simulación de modelos clásicos de series temporales.

**Fuente:** Elaboración propia.

En la figura 4.1 se presentan los modelos. Estas simulaciones se modificaron según las características de cada combinación de factores en el diseño. Es decir, por cada modelo simulado se obtienen 32 series de tiempo, en total 5x32. Además, a cada modelo se les aplican los 5 tests (4 tests clásicos y 1 test wavelet).

## 4.1. Resultados de las simulaciones

En el cuadro 4.3 se presentan los resultados obtenidos con cada test para cada uno de los modelos de series temporales luego de considerar todas las combinaciones de los niveles de los factores. El desempeño de cada uno se evalúa mediante un porcentaje que representa los puntos de cambio ingresados y que al menos 1 de ellos fueron detectados. Este porcentaje se separa en dos medidas, el primer porcentaje corresponde a los puntos de cambio detectados por un cambio en la varianza, el segundo porcentaje corresponde los puntos de cambio que fueron detectados en un cambio en la media.

Tests de detección de puntos de cambio					
Modelo	MPQ	Chow	CUSUM	MOSUM	Wavelet
AR(1)	(6 %)(56 %)	(6 %)(37 %)	(0 %)(43 %)	(6 %)(37 %)	(43 %)(18 %)
MA(1)	(0 %)(81 %)	(0 %)(56 %)	(0 %)(56 %)	(6 %)(43 %)	(12 %)(31 %)
ARMA(1,1)	(12 %)(43 %)	(0 %)(12 %)	(6 %)(25 %)	(18 %)(31 %)	(18 %)(25 %)
ARIMA(1,1,1)	(31 %)(12 %)	(18 %)(6 %)	(18 %)(18 %)	(43 %)(25 %)	(37 %)(25 %)
GARCH(1,1)	(6 %)(50 %)	(6 %)(43 %)	(0 %)(37 %)	(6 %)(25 %)	(0 %)(12 %)

Cuadro 4.3: Porcentajes de puntos de cambio detectados en **media** o en **varianza** por los tests en los modelos de series temporales.

**Fuente:** Elaboración propia.

Para el modelo AR (1) los tests clásicos tuvieron un mejor rendimiento en la detección de puntos de cambios en la media, entre el 60 % y el 30 %. Sin embargo, la mayoría de los tests no superan el 6 % de la detección de puntos de cambio en la varianza para los casos cuando los puntos de cambio están al final de la serie y a una distancia cercana. Por otro lado, el test wavelet fue superior a los tests clásicos en la detección de puntos de cambio en la varianza.

Para el modelo MA (1), los tests clásicos tuvieron un exitoso rendimiento en la detección de cambios en media, sobre todo el test MPQ el cual detectó alrededor del 80 % de los puntos ingresados. Además, antes de ingresar los puntos de cambio, 3 de los 4 test tradicionales (MPQ, Chow y CUSUM) detectaron el mismo punto de cambio en la media. Si bien los tests clásicos tuvieron un mejor rendimiento en media, en la detección de puntos de cambio en varianza su rendimiento fue inferior o igual al 6 %, el test MOSUM solo logró detectar un punto de cambio. Por otro lado, el test wavelet tuvo un mejor rendimiento en la detección de puntos de cambio en varianza que los tests clásicos en presencia de varios puntos de cambio. Respecto a la detección de puntos de cambio en la media, la mayoría de los puntos detectados por el análisis wavelet fueron producto de fuertes cambios en magnitud.

Para el modelo ARMA (1,1) los tests tuvieron un rendimiento superior en la detección de puntos de cambio en media, incluso el test wavelet. Los tests con mejor rendimiento en media fueron el test MPQ y el test MOSUM. Respecto a los cambios en varianza, el test MPQ sólo logró detectar puntos de cambio al final de la serie con fuertes cambios en magnitud y donde los puntos estaban juntos. El test MOSUM tuvo un alto rendimiento en la detección de puntos de cambio en varianza al igual que el test wavelet, aunque la mayoría de los casos que logró detectar fueron cuando los puntos de cambio se encontraban cerca de los grandes cambios en magnitud. Por último, el test wavelet tuvo un mejor rendimiento en la detección de puntos de cambio en media que en varianza, esto producto a los fuertes cambios en magnitud al final de la serie temporal superando al test de Chow e igualando el desempeño del test CUSUM.

El modelo ARIMA (1,1,1) al ser una serie ARMA diferenciada posee menos cambios bruscos a través del tiempo. El test MOSUM se desenvuelve mejor en la detección de cambios en la varianza, pero el test wavelet supera o iguala a todos los otros tests en la detección de puntos de cambio en la media. Los tests clásicos, sorprendentemente tuvieron un mejor desempeño detectando más cambios en varianza que en media al igual que el test MOSUM. En el caso del test de Chow la precisión de la detección de los puntos de cambio aumenta en presencia de cambios fuertes en magnitud aunque su rendimiento se ve limitado por su capacidad de detectar 1 punto de cambio a la vez, por lo que su rendimiento es bajo en ciertas situaciones. Por último el test MPQ tuvo un mejor rendimiento detectando cambios leves en varianza cuando se presentaban varios puntos de cambio y en otras ocasiones con fuertes cambios en magnitud.

En el modelo GARCH (1,1) se obtuvieron resultados parecidos al modelo AR(1), algunos tests lograron detectar cambios en varianza con gran dificultad, ya que en el caso de los tests MOSUM y CUSUM el puente browniano apenas logró traspasar las bandas de confianza en presencia de cambios leves y fuertes en magnitud de la varianza. Por el contrario, en media, los tests MOSUM y CUSUM logran traspasar con facilidad las bandas de confianza, al menos en los cambios fuertes en magnitud. En los tests MPQ y Chow la precisión aumenta cuando se presentan cambios fuertes en magnitud en la media, en los casos de cambios leves su precisión disminuye considerablemente, aunque en los casos de cambios al principio se mantienen cerca del rango de cambio. En el caso del análisis wavelet no se obtuvo un buen rendimiento para detectar puntos de cambio en varianza, los puntos de cambio si fueron detectados en cierto modo, pero el test wavelet no es capaz de determinar los puntos de cambio óptimos. En media, se obtuvo un mejor resultado, si bien no detecta los puntos de cambios ingresados, se mantienen a una corta distancia de los puntos ingresados. En general, el test wavelet tuvo un desempeño igual o inferior a los tests clásicos en la detección de puntos de cambio en la media, salvo para el caso del modelo ARIMA donde el ocurre el efecto contrario.

# Capítulo 5

## Aplicaciones

En este capítulo se presenta la aplicación de los test clásicos y el análisis wavelet para la detección de puntos de cambio en las series temporales del río Nilo, la tasa de desempleo en Chile, el IPC general de Chile y los datos del Seeing del observatorio Paranal de Chile.

### 5.1. Flujo anual del caudal del río Nilo

El río Nilo es considerado uno de los ríos más grandes de África, la superficie de su cuenca cubre alrededor del 10% del continente, su amplio contorno abastece de agua a 10 países, nace del río Kagera hasta desembocar en el extremo suroeste del Mediterráneo y actualmente las fuentes hídricas que alimentan al caudal del río Nilo son el Nilo Azul y el Nilo Blanco ([GeoEnciclopedia, 2016](#)). La serie del flujo anual del caudal del río Nilo cuenta con 100 observaciones desde el año 1871 hasta el año 1970, ver Figura 5.1.

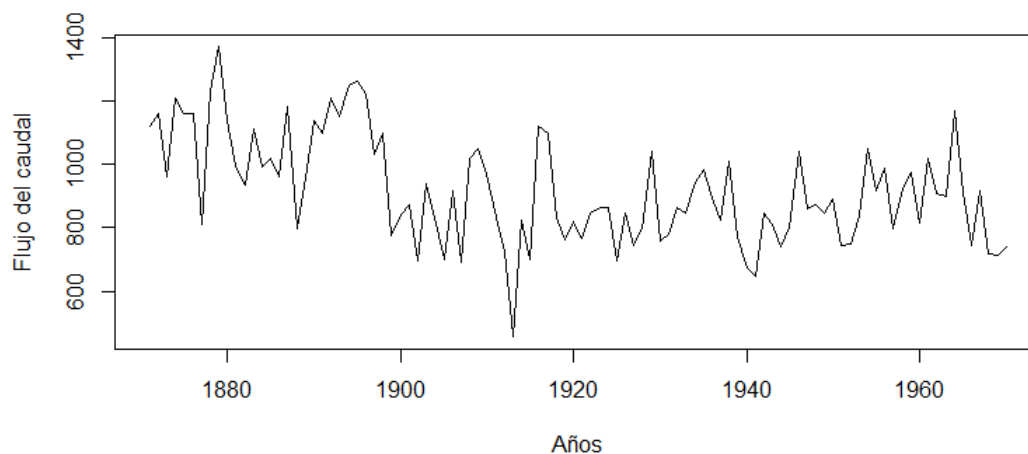


Figura 5.1: Flujo anual del caudal del río Nilo 1871-1970.

**Fuente:** Inference for Change Point and Post Change Means After a CUSUM Test, pág. 27.

Se puede observar en la Figura 5.1 que la serie temporal posee una leve tendencia a la baja con cambios bruscos en las etapas iniciales desde 1871 hasta 1880 y luego decae hasta llegar a un punto mínimo en 1913. Posteriormente se estabiliza alrededor de la media de  $919 \text{ m}^3$ . La estabilización del flujo del caudal está justificado por obras de construcción realizadas durante los años 1899 y 1902, ya que en el río Asuán se impulsó un proyecto de construcción llamada la presa baja para disminuir los daños provocados por las inundaciones (Fredes y Hidalgo, 2019).

Se establecen como hipótesis a contrastar:

$H_0$ : El flujo anual del caudal del río Nilo es estable a lo largo del tiempo.

vs

$H_1$ : El flujo anual del caudal del río Nilo no es estable a lo largo del tiempo.

El resultado de los tests de hipótesis aplicados para este conjunto de datos fueron los siguientes:

Test de hipótesis	Estadístico de prueba	P-valor
Test de Chow	75,93	2,22e-16
Test CUSUM por MCO	2,95	5,409e-08
Test MOSUM por MCO	1,53	0,01

Cuadro 5.1: Test de hipótesis del flujo anual del caudal del río Nilo.

**Fuente:** Elaboración propia.

Del cuadro 5.1 se puede observar que los p-valores son inferiores a  $\alpha = 0,05$ , por lo tanto, se tiene evidencia suficiente para rechazar la hipótesis nula y afirmar que el flujo anual del caudal del río Nilo no es estable a lo largo del tiempo.

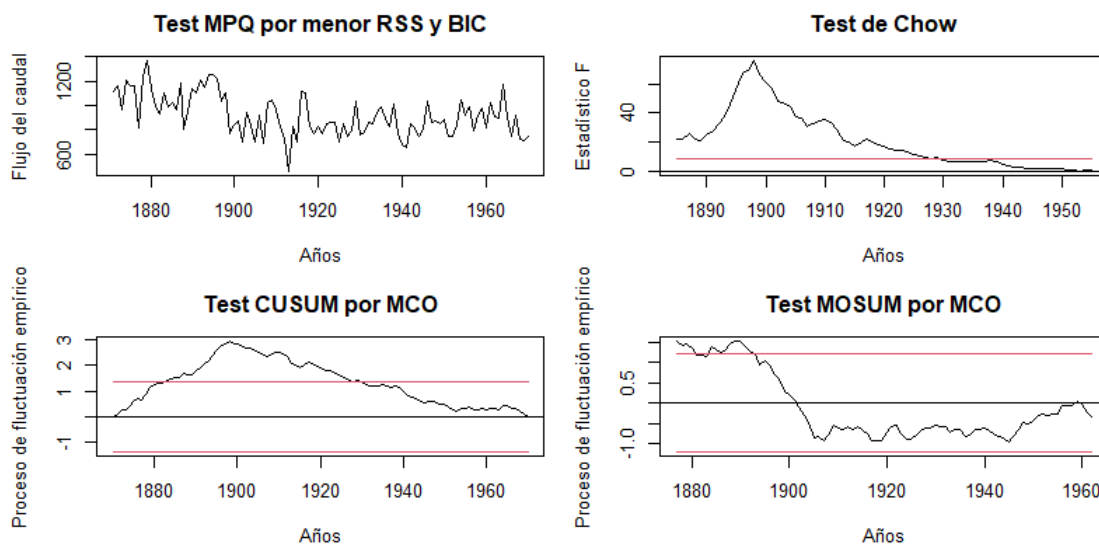


Figura 5.2: Puntos de cambio del flujo anual del caudal del río Nilo.

Fuente: Elaboración propia.

En la Figura 5.2 se observa que en el gráfico correspondiente a la estimación de puntos de cambio bajo el test de múltiples puntos de quiebre (MPQ) hay 2 segmentos que separan la serie temporal cuyo punto de cambio fue estimado con el menor BIC y menor SSR o suma de cuadrados residuales los cuales fueron (1270,1597457). Por ende, la cantidad de puntos de cambio para esta serie temporal fue de 1 correspondiente al año 1898.

En el lado superior derecho de la Figura 5.2 se presenta el estadístico del test de Chow. El punto de cambio obtenido corresponde al año 1898 con un estadístico observado de  $Sup.F = 75,93$ . El año está relacionado a las primeras etapas de construcción de la presa baja de Asuán.

De manera similar, se puede observar en el panel inferior izquierdo de la Figura 5.2, bajo la estimación de mínimos cuadrados ordinarios el test CUSUM, que el punto de cambio se produjo en el año 1898 con un estadístico de prueba  $S_0 = 2,95$ , es decir, hasta ahora el test de múltiples puntos de quiebre, el test de Chow y el test de CUSUM han detectado el punto de cambio en la misma fecha designada. Sin embargo, el test de CUSUM posee un decaimiento más lento cuando cruza la banda superior de confianza, lo cual dificulta la identificación del punto de cambio, ver Figura 5.2.

Finalmente, en el panel inferior derecho de la Figura 5.2 se puede observar, bajo la estimación de los mínimos cuadrados ordinarios del test de MOSUM, que el punto de cambio se produjo en el año 1889. Esto evidencia que 3 de los 4 test lograron detectar el punto de cambio en la misma fecha bajo distintos métodos.

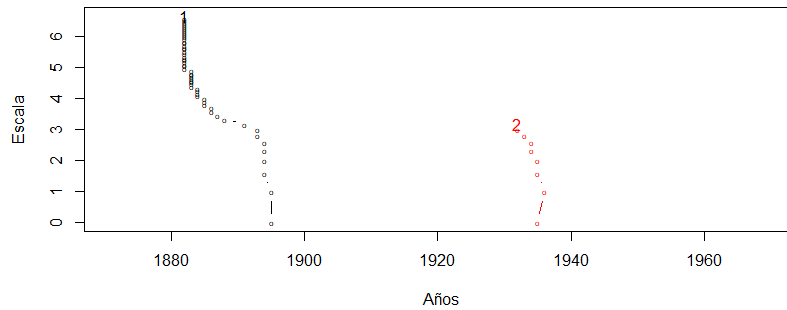


Figura 5.3: Ramificación de candidatos a ser puntos de cambio para la variable del flujo anual del caudal del río Nilo.

**Fuente:** Elaboración propia.

En la Figura 5.3 utilizando el análisis wavelet se pueden observar 2 ramificaciones que marcan los máximos locales o cúspides de la transformada wavelet Gorro mexicano que son candidatos a ser puntos de cambio, la ramificación 1 corresponde al año 1895 y la ramificación 2 al año 1935, la ramificación perteneciente al año 1895 tiene mayor posibilidad de ser punto de cambio ya que una de las condiciones para ser considerada punto de cambio es que su rama sea superior o igual a 9 núcleos.

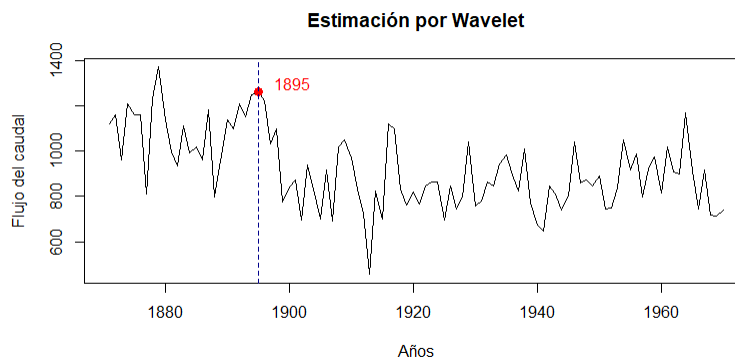


Figura 5.4: Puntos de cambio con wavelet para la variable flujo anual del caudal del río Nilo.

**Fuente:** Elaboración propia.

Se puede observar en la Figura 5.4 que, de las 2 ramificaciones en la Figura 5.3, candidatos a ser puntos de cambio, la ramificación del año 1895 fue seleccionada como el punto de cambio óptimo.

Tests de puntos de cambio	Puntos detectados
Test MPQ	1898
Test Chow	1898
Test CUSUM por MCO	1898
Test MOSUM por MCO	1889
Análisis wavelet	1895

Cuadro 5.2: Detección de puntos de cambio del flujo anual del caudal del río Nilo bajo diferentes métodos y tests.

**Fuente:** Elaboración propia.

En el cuadro 5.2 se presentan los puntos detectados por cada uno de los métodos utilizados. Todos son cercanos en el tiempo entre si, 3 de los métodos coincidieron en el año 1898 mientras el test MOSUM detectó una fecha más temprana (9 años antes) y el test wavelet casi coincidió con los resultados de los 3 primeros tests del cuadro (3 años antes).

En la siguiente sección se va a describir la segunda aplicación de los tests en los datos de la tasa de desempleo en Chile.

## 5.2. Tasa de desempleo en Chile

El desempleo, desocupación o cesantía, se refiere a la ocupación del ciudadano que no se encuentra ejerciendo una actividad laboral para una entidad o empresa. Por consiguiente, pertenece a la proporción de personas que en edad, condiciones y disposición de trabajar que no han obtenido un puesto de trabajo ([Instituto Nacional de Estadísticas, 2020](#)).

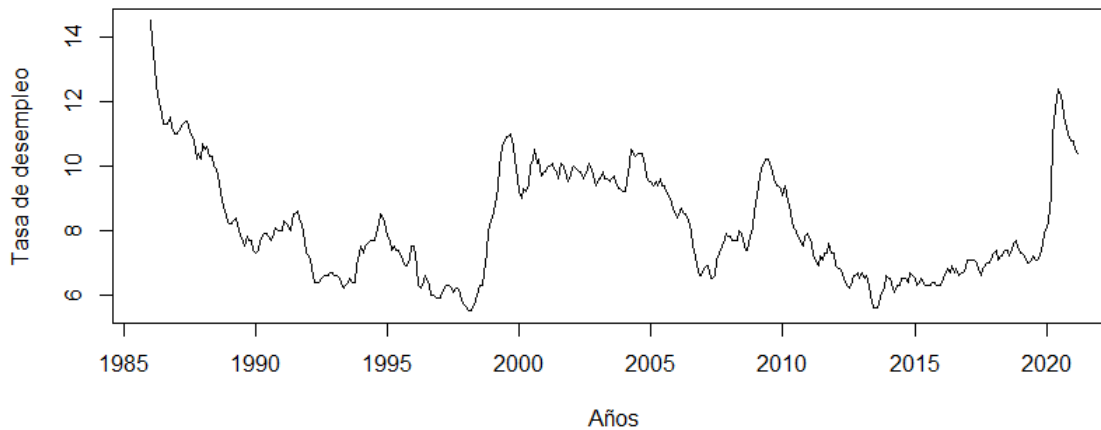


Figura 5.5: Tasa de desempleo en Chile 1986-2021.

**Fuente:** Federal Reserve economic data.

La base de datos de la tasa de desempleo en Chile considera como desempleados a personas sobre los 15 años de edad. Los datos se obtienen de la página web de FRED (Federal Reserve Economic Data) y abarca un rango de tiempo desde enero 1986 hasta marzo 2021.

A primera vista se pueden distinguir 6 posibles puntos de cambio en la Figura 5.5. El primer punto de cambio se produce alrededor de los años 1990 y 1992 después de una rápida caída de la tasa del desempleo en el año 1986, desde un 14,5 % a un 8,5 % aprox. El segundo punto de cambio se presenta alrededor de los años 1994 y 1995 con el mismo valor porcentual de los años 1990 y 1992. El tercer punto de cambio se presenta alrededor de los años 1999 y 2000 con un aumento en la tasa de desempleo desde un 5,5 % a principios de 1998 hasta un 11 % en el tercer trimestre del año 1999. Después del año 2000 hasta el año 2005 la tasa de desempleo se mantiene estable variando entre 9,5 % y 10,5 % hasta llegar al cuarto punto de cambio entre el segundo y tercer trimestre del año 2004 con un valor porcentual de 10,4 %. El quinto punto de cambio se observa alrededor de los años 2009 y 2010 después de un lento descenso hasta un 6.5 % en la tasa de desempleo, luego se recupera rápidamente hasta alcanzar valores cercanos entre los 9,5 y 10 %. Finalmente, el último punto de cambio se observa al final de la serie temporal entre los años 2020 y 2021 con un valor porcentual de un 12,5 % aprox., donde la tasa de desempleo mantuvo un lento crecimiento desde el año 2013 hasta el año 2019, entre el 6 % hasta un 8 %, luego tiene un rápido y vertiginoso aumento a comienzos del año 2020.

Se establecen como hipótesis a contrastar:

$H_0$ : La tasa de desempleo en Chile es estable a lo largo del tiempo.

vs

$H_1$ : La tasa de desempleo en Chile no es estable a lo largo del tiempo.

El resultado de los tests de hipótesis aplicados para este conjunto de datos se presenta en el siguiente cuadro:

Test de hipótesis	Estadístico de prueba	P-valor
Test de Chow	64,889	4,855e-14
test CUSUM por MCO	3,1963	2,675e-09
test MOSUM por MCO	3,1172	0,01

Cuadro 5.3: Test de hipótesis de la tasa de desempleo en Chile.

**Fuente:** Elaboración propia.

Observando el 5.3 los p-valores obtenidos para cada tests son inferiores al nivel de significación  $\alpha = 0,05$ . Por lo tanto, en todos los casos se rechaza la hipótesis nula y se concluye que la tasa de desempleo en Chile no es estable a lo largo del tiempo.

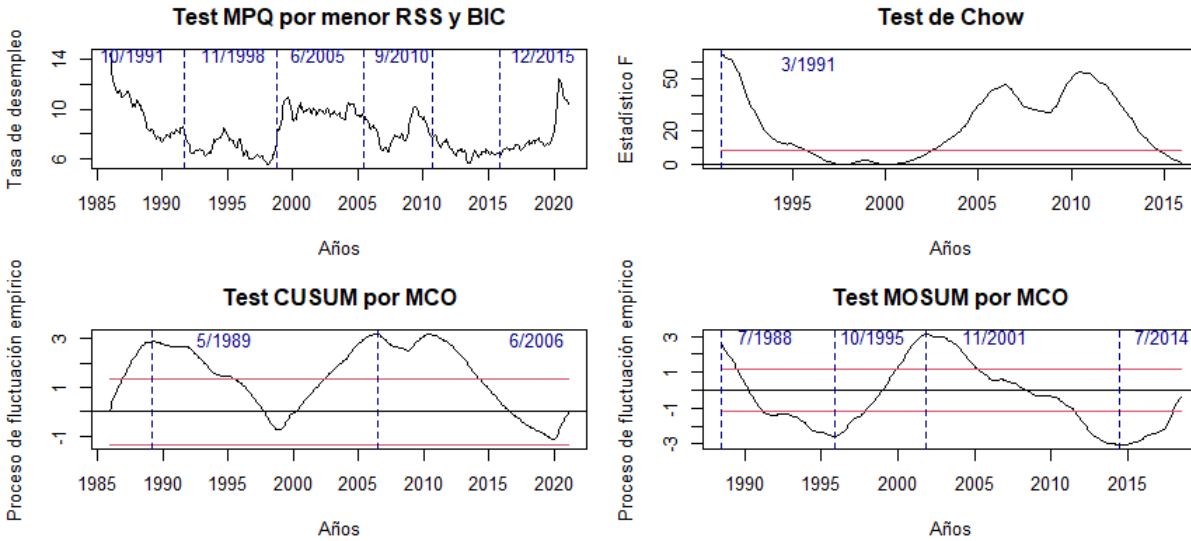


Figura 5.6: Puntos de cambio de la tasa de desempleo en Chile.

Fuente: Elaboración propia.

En la Figura 5.6 se puede observar que en la esquina superior izquierda se encuentra el test de múltiples puntos de quiebre (MPQ), el cual separa la serie de la tasa de desempleo en Chile en 6 segmentos con el menor BIC y RSS cuyos valores fueron (541.9, 1377.8). Por ende, la cantidad de puntos de cambio para esta serie temporal fue de 5 (octubre del año 1991, noviembre del año 1998, mayo del año 2005, septiembre del año 2010 y diciembre del año 2015), la mayoría de los puntos de cambio que detectó este test estuvo cerca de los puntos mencionados en la Figura 5.5. En la esquina superior derecha de la Figura 5.6 se encuentra el test de Chow, el cual detectó un punto de cambio en marzo del año 1991 el cual estuvo a unos meses de diferencia del punto de cambio que detectó el test de múltiples puntos de quiebres (MPQ). En la esquina inferior izquierda de la Figura 5.6 se encuentra el test CUSUM que detectó 2 puntos de cambio en mayo del año 1989 y en junio del año 2006, estos puntos de cambio son cercanos a los puntos de cambio detectados por el test de múltiples puntos de quiebre con un par de meses de diferencia. En la esquina inferior derecha de la Figura 5.6 se encuentra el test MOSUM que detectó 4 puntos de cambio en julio del año 1988, en octubre del año 1995, en noviembre del año 2001 y en julio del año 2014. Este test y el test de múltiples puntos de quiebre fueron los que más puntos de cambio detectaron, solo que el test MOSUM tiene hasta 4 años de diferencia en comparación con el test de múltiples puntos de quiebre.

Entre los años 1998 y 2002 hubo una caída en la generación de nuevos empleos que afectaron a las personas jóvenes (entre 18 y 25 años) de mayor nivel educacional y baja experiencia laboral, los sectores más afectados fueron en la construcción por el aumento del salario mínimo y en la industria (Cowan, Micco, Mizala, Pagés, y Romaguera, 2003). Entre los años 1990 y 1992 Chile pasaba por una etapa de transición y, además, en esos años Chile se posicionó como una de las mejores economías de Latinoamérica (Fredes y Hidalgo, 2019). Entre los años 1995 y 1999 se produce la crisis asiática provocando una caída en las exportaciones y del PIB (Fredes y Hidalgo, 2019).

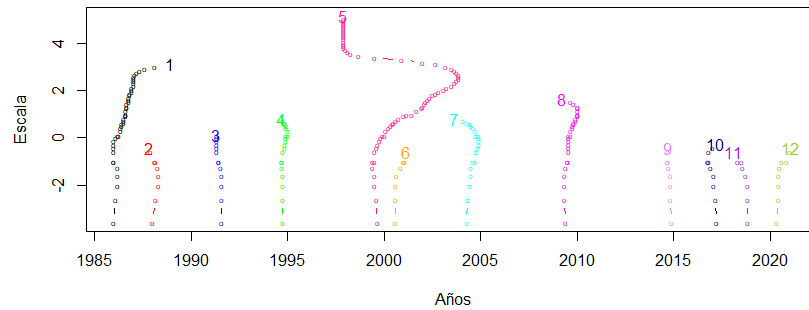


Figura 5.7: Ramificación de candidatos a puntos de cambio en los datos de la tasa de desempleo en Chile.

**Fuente:** Elaboración propia.

En la Figura 5.7 se observa que hay muchos candidatos para ser puntos de cambio que cumplen el criterio de poseer 9 o más núcleos. Las ramificaciones que más se destacan son las ramas 1 del año 1986, la rama 4 del año 1995, la rama 5 del año 1999, la rama 7 del año 2004 y la rama 8 del año 2009, por lo que es muy posible que esas ramas sean puntos de cambio óptimos.

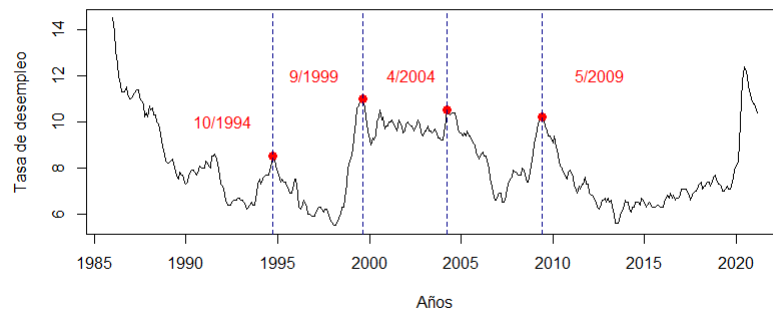


Figura 5.8: Puntos de cambio con wavelet en la tasa de desempleo en Chile.

**Fuente:** Elaboración propia.

En la Figura 5.8 se observa que, de todas las ramificaciones candidatas a ser puntos de cambio solo 4 fueron elegidas, estas marcan la mayoría de los puntos de interés mencionados al inicio de la sección de los datos, lo que lo convierte en un buen método para detectar cambios en las variaciones.

A continuación en el cuadro 5.4 se presenta un tabla resumen de los punto de cambio detectados por los tests y el análisis wavelet.

Tests de puntos de cambio	Puntos detectados
Test MPQ	Octubre 1991 Noviembre 1998 Junio 2005 Septiembre 2010 Diciembre 2015
Test Chow	Marzo 1991
Test CUSUM por MCO	Mayo 1989 Junio 2006
Test MOSUM por MCO	Julio 1988 Octubre 1995 Noviembre 2001 Julio 2014
Análisis wavelet	Octubre 1994 Septiembre 1999 Abril 2004 Mayo 2009

Cuadro 5.4: Detección de puntos de cambio en la tasa de desempleo en Chile bajo diferentes métodos y tests.

**Fuente:** Elaboración propia.

En el cuadro 5.4 se puede observar que los tests de múltiples puntos de quiebre y Chow detectaron, a lo más, 1 punto de cambio en la misma fecha, con 8 meses de diferencia. Por otro lado, el test CUSUM detectó puntos de cambios muy cerca de los puntos que detectó el test de múltiples puntos de quiebre con una diferencia de 1 a 2 años aproximadamente, por lo que puede ser una zona de punto de cambio entre los años 1990 y 2005. En el caso del test de MOSUM detectó casi la misma cantidad de puntos de cambio que detectó el test de múltiples puntos de quiebre y muy cerca de las fechas que ese mismo test detectó, con una diferencia de hasta 4 años. Por último, el análisis wavelet tuvo el mismo rendimiento que el test MOSUM y también los puntos de cambio que detectó estuvieron a una distancia de a lo más 3 años.

### 5.3. IPC en Chile

El índice de precios al consumidor (IPC) mide la variación de los precios de una canasta de bienes y servicios de consumo en los hogares, se construye en base a la encuesta de presupuestos familiares y se realiza mes a mes por el Instituto Nacional de Estadísticas. En la Figura 5.9 se presenta la variación mensual del IPC en Chile desde enero del año 2004 hasta febrero del año 2021.

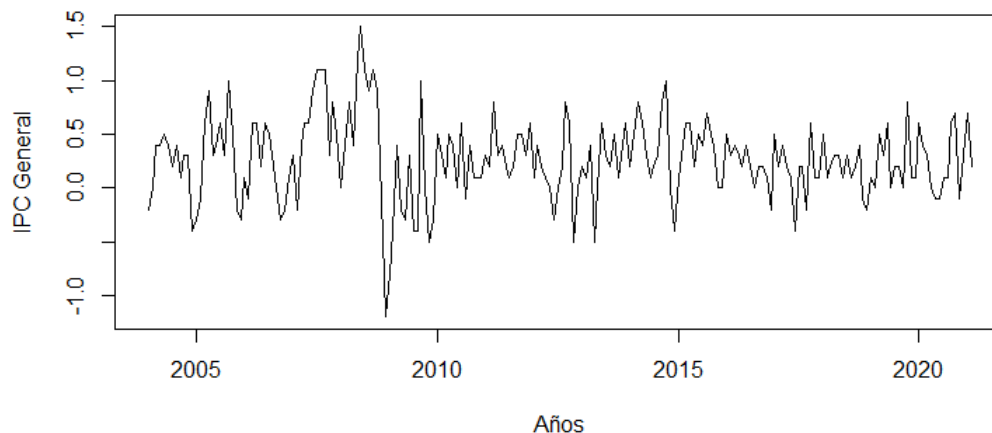


Figura 5.9: IPC General en Chile 2004-2021.

**Fuente:** Banco Central de Chile.

En la Figura 5.9 se puede observar que en el año 2008 hubo una caída estrepitosa desde un 1,5% a un -1,2%, las causas se deben a muchos acontecimientos internacionales. Uno de ellos fue la denominada crisis Subprime o crisis hipotecaria producto de los préstamos hipotecarios a instituciones de inversiones cuyas acciones desestabilizaron la economía mundial con pocos retornos para sus accionistas y aumentaron la incertidumbre de otros inversionistas. El sector inmobiliario también recibió un efecto negativo con el aumento del valor de las propiedades (Laval, 2009).

Se establecen como hipótesis a contrastar:

$H_0$ : El IPC General en Chile es estable a lo largo del tiempo.

vs

$H_1$ : EL IPC General en Chile no es estable a lo largo del tiempo.

Test de hipótesis	Estadístico de prueba	P-valor
Test de Chow	13,27	0,005785
Test CUSUM por MCO	1,59	0,01261
Test MOSUM por MCO	1,63	0,01

Cuadro 5.5: Test de hipótesis del IPC General en Chile.

**Fuente:** Elaboración propia.

A través del cuadro 5.5 se pueden observar los resultados de cada test los cuales obtuvieron un p-valor inferior a  $\alpha = 0,05$ . Por lo tanto, ese tiene evidencia suficiente para afirmar que el IPC General en Chile no es estable a lo largo del tiempo.

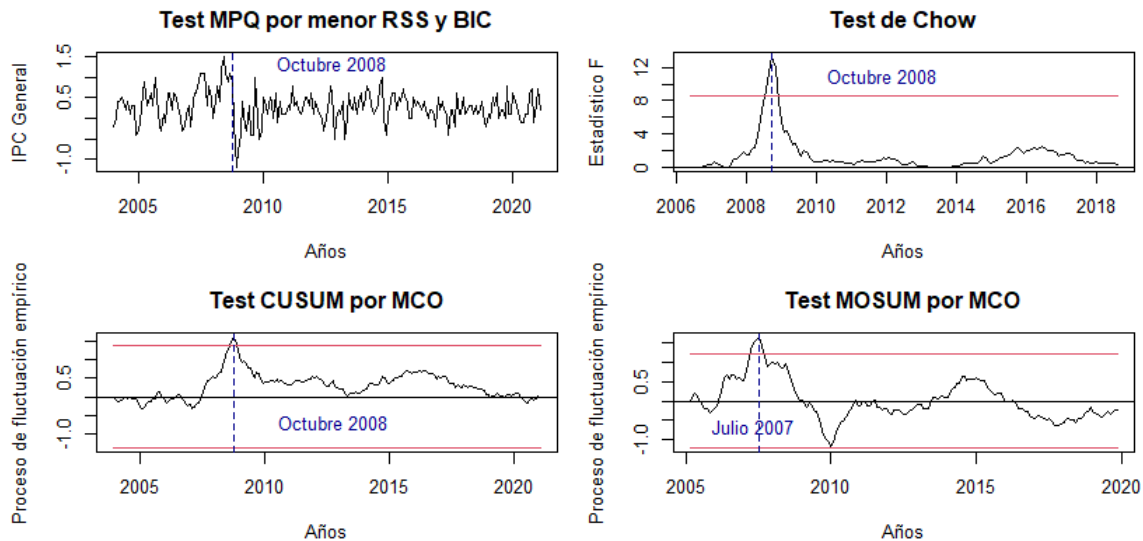


Figura 5.10: Puntos de cambio en el IPC General en Chile.

**Fuente:** Elaboración propia.

En la esquina superior izquierda de la Figura 5.10 se observa que el test de múltiples puntos de quiebre (MPQ) divide en 2 segmentos la serie temporal, cuyo punto de cambio fue estimado con el menor BIC y menor SSR, cuyos valores fueron (192.97, 27.75) respectivamente. Por ende, la cantidad de puntos de cambio para esta serie temporal fue de 1, correspondiente a octubre del año 2008. En la esquina superior derecha de la Figura 5.10 se encuentra el test de Chow que detectó un punto de cambio en octubre del año 2008 con una variación porcentual del IPC General del 0,9%, en la cual se produjo un aumento del precio del petróleo (Laval, 2007). También la crisis Subprime ya alcanzaba su punto máximo provocando estragos en las grandes agencias de inversiones que posteriormente llevaría a la quiebra a muchas de ellas para el año 2009 (Laval, 2009). En la esquina inferior izquierda de la Figura 5.10 se encuentra el test CUSUM que muestra el mismo resultado que en los tests de MPQ y Chow, un punto de cambio en octubre del año 2008. En la esquina inferior derecha de la Figura 5.10 se encuentra el test MOSUM que detectó un punto de cambio en julio del año 2007 correspondiente a una variación del IPC General del 1,1%. En esta fecha comenzaban los efectos de la crisis Subprime con el aumento de las tasas de interés interbancarias en Estados Unidos (Laval, 2007).

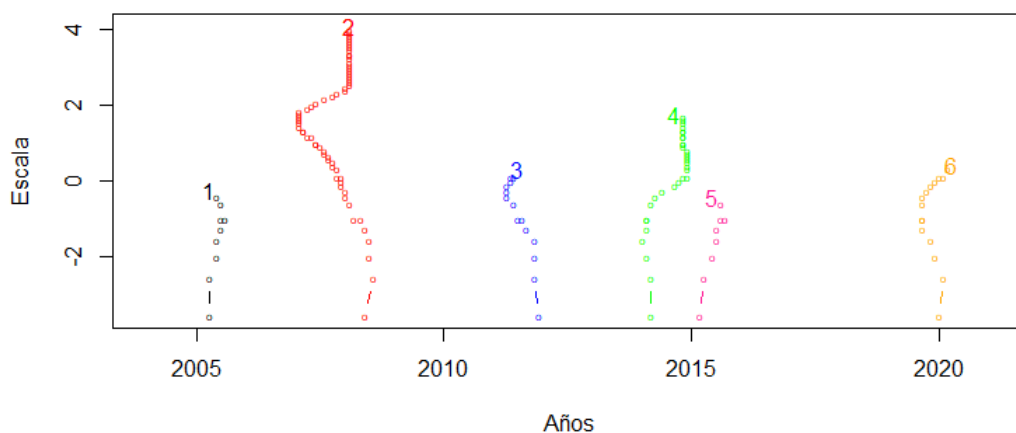


Figura 5.11: Ramificación de candidatos a puntos de cambio del IPC General en Chile.

**Fuente:** Elaboración propia.

En la Figura 5.11 se puede observar que de todos los candidatos a ser punto de cambio la rama 2 es la más indicada para ser elegida ya que es la más larga, aunque otras ramas como la 4 y la 6 también pueden ser consideradas como puntos de cambio.

El análisis wavelet no pudo determinar un punto de cambio para el IPC General de Chile debido a problemas en el software. En el cuadro 5.6 se presentan los puntos de cambio detectados por los distintos tests.

Tests de puntos de cambio	Puntos detectados
Test MPQ	Octubre 2008
Test Chow por MCO	Octubre 2008
Test CUSUM por MCO	Octubre 2008
Test MOSUM por MCO	Julio 2007

Cuadro 5.6: Detección de puntos de cambio del IPC General en Chile bajo diferentes tests.

**Fuente:** Elaboración propia.

En éste se observa que los tests de múltiples puntos de quiebre (MPQ), Chow y CUSUM obtuvieron los mismos resultados. Por otro lado, el test de MOSUM detectó un cambio cerca de donde se encuentra el punto de cambio detectado por los otros tests, por lo que es posible que ese punto detectado se encuentre en una zona de posibles puntos de cambio.

En la siguiente sección se presenta la serie temporal de los datos del *Seeing*. Además, se describe brevemente su uso en la astronomía y la importancia del observatorio donde se obtuvo la información para este estudio.

## 5.4. Seeing

En Chile existe un gran interés por la comunidad científica en el estudio de la astronomía y de sus derivados dada las condiciones favorables de visibilidad atmosférica que posee, particularmente, en la zona norte del país en la región de Antofagasta se encuentra uno de los observatorios internacionales más importantes, el observatorio Paranal que cuenta con el *Very Large Telescope (VTL)*.

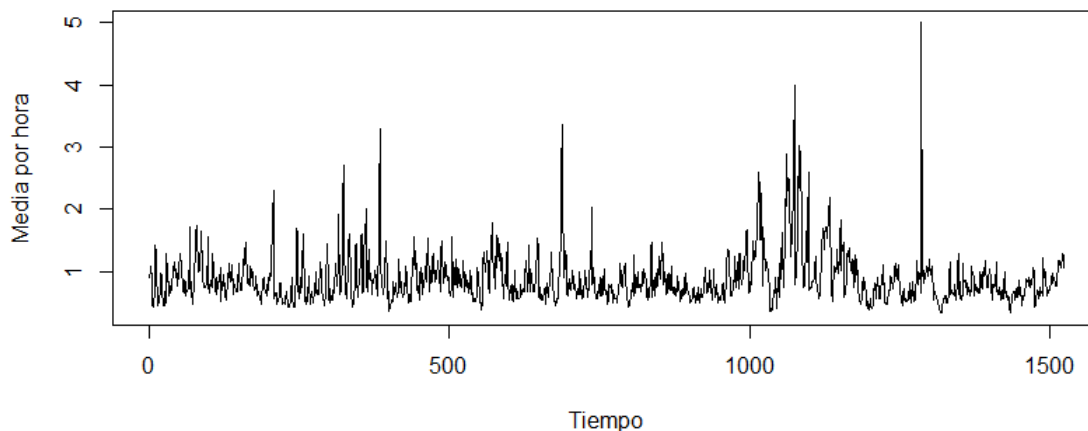


Figura 5.12: Media por hora del Seeing 1° semestre 2021.

**Fuente:** Observatorio Paranal de Chile.

Los datos bajo análisis corresponden a la variable *Seeing*, medida utilizada para determinar el grado de distorsión de la atmósfera. Los datos obtenidos por el observatorio están registrados por segundos en horarios de las 22 PM a las 10 AM. En este estudio se considera la media por hora como se observa en la figura 5.12. El motivo por el que se decide utilizar las medias por hora en vez de las medias por segundos es debido a que las diferencias entre los valores registrados por segundos eran demasiado pequeña por lo que no se podría observar ni detectar un cambio lo suficientemente pequeño para ser considerado como punto de cambio. El *Seeing* o grado de turbulencia atmosférica, es producido por las corrientes de aire con valores que se encuentran entre los 0.4 grados de arco en zonas de gran altitud y sobre 1 grado de arco en zonas de baja altitud. Su escala de medición está separada en 2 categorías, la escala *Antoniadi* y la escala *Pickering*, en esta investigación se utiliza la escala *Pickering* que va desde 1(excelente) a 10(pésimo) (Bravo, 2019).

Se establecen como hipótesis a contrastar:

$H_0$ : La media por hora del Seeing es estable a lo largo del tiempo.

vs

$H_1$ : La media por hora del Seeing no es estable a lo largo del tiempo.

Test de hipótesis	Estadístico de prueba	P-valor
Test de Chow	47,072	3,571e-10
Test CUSUM por MCO	2,8265	2,301e-07
Test MOSUM por MCO	4,8639	0,01

Cuadro 5.7: Test de hipótesis de la media por hora del Seeing.

**Fuente:** Elaboración propia.

A través del cuadro 5.7 se pueden observar los resultados de cada test los cuales obtuvieron un p-valor inferior a  $\alpha = 0,05$ . Por lo tanto, ese tiene evidencia suficiente para afirmar que la media por hora del Seeing no es estable a lo largo del tiempo.

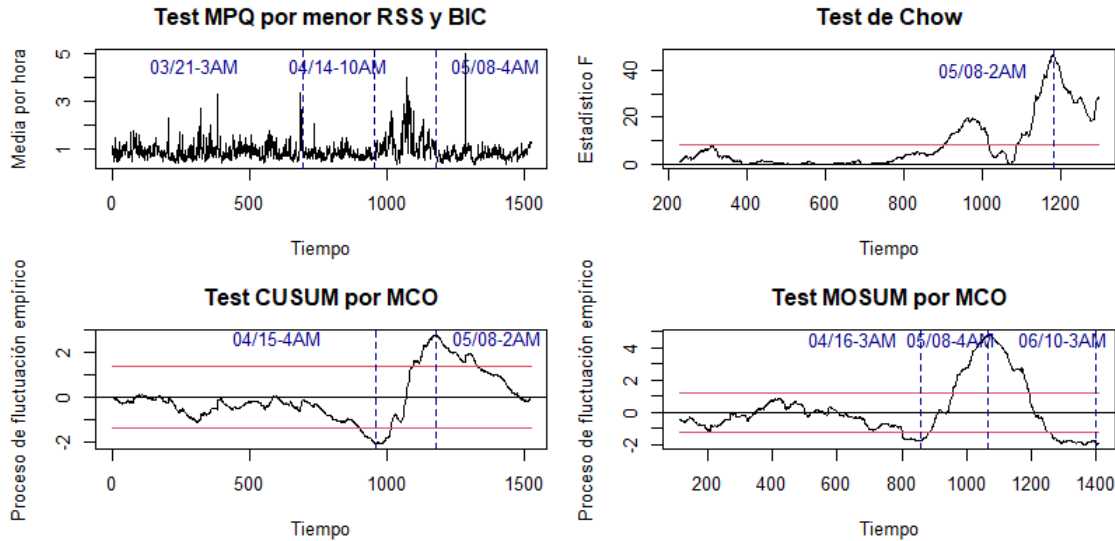


Figura 5.13: Puntos de cambios en la media por hora del Seeing.

Fuente: Elaboración propia.

En la esquina superior izquierda de la Figura 5.13 se puede observar que el test de múltiples puntos de quiebre (MPQ) divide a la serie temporal en 4 segmentos y 3 puntos de cambio utilizando el menor BIC y RSS, los valores correspondientes a estos criterios fueron (1385.6, 213.1) respectivamente. El primero, el 03/21-03AM, se detectó cuando la media disminuyó levemente a 0.5 puntos del arco de distorsión aprox. hasta llegar al segundo punto de cambio, el 04/14-10AM, con un aumento fuerte de la media a 1.3 puntos aprox. Finalmente, el tercer punto, el 05/08-04AM, con una caída fuerte en la media de 0.5 puntos aprox. En la esquina superior derecha de la Figura 5.13 se encuentra el test de Chow el cual detectó un punto de cambio correspondiente al 05/08-02AM con un valor de 1.025 de *Seeing*, lo que gráficamente coincide con el cambio en la media que fue detectado por el test de múltiples puntos de quiebre sólo que un poco antes. En la esquina inferior izquierda de la Figura 5.13 se encuentra el test CUSUM el cual detectó 2 puntos de cambio, el 04/15-04AM, que marca un cambio fuerte en la media que está por delante del segundo punto de cambio que detectó el test de múltiples puntos de quiebre, y el 05/08-02AM, que estuvo cerca del último punto de cambio que detectó el test de múltiples puntos de quiebre. En síntesis, tanto el test CUSUM como el test de múltiples puntos de quiebre tienen estimaciones muy similares. En la esquina inferior derecha de la Figura 5.13 se encuentra el test MOSUM el cual detectó 3 puntos de cambios, el 04/16-03AM con un *Seeing* por hora de 0.67, que marca un cambio fuerte en la media que está por delante del primer punto de cambio que detectó el test CUSUM y un poco más lejos, con 2 días de diferencia, del segundo punto de cambio detectado por el test de múltiples puntos de quiebre (MPQ), el 05/08-04AM, coincidió con el tercer punto de cambio que fue detectado por el test de múltiples puntos de quiebre. Por último, se detectó el tercer punto de cambio al final de la serie, el 06/10-03AM, con un leve aumento en la media de 1.2 grados del arco de distorsión.

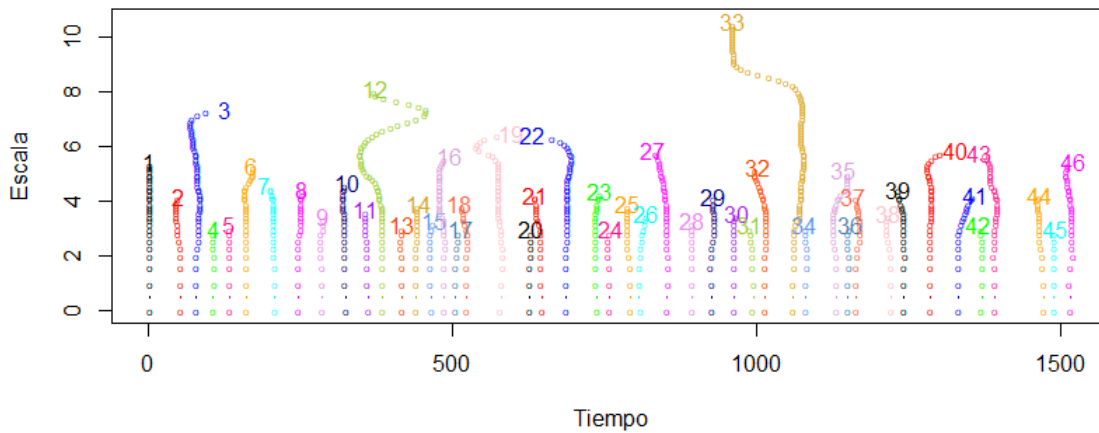


Figura 5.14: Ramificación de candidatos a puntos de cambio en la media por hora del Seeing.

**Fuente:** Elaboración propia.

En la Figura 5.14 se puede observar que hay muchas ramificaciones candidatas a ser puntos de cambio, pero si recordamos 1 de los criterios de selección el cual es que posea sobre 10 núcleos la cantidad de candidatos se reduce significativamente, en lo que respecta a las posibles ramas a ser puntos de cambio se puede observar que la rama 3, 12, 22, 33, 40, 42 y 46 tienen grandes posibilidades de ser seleccionadas.

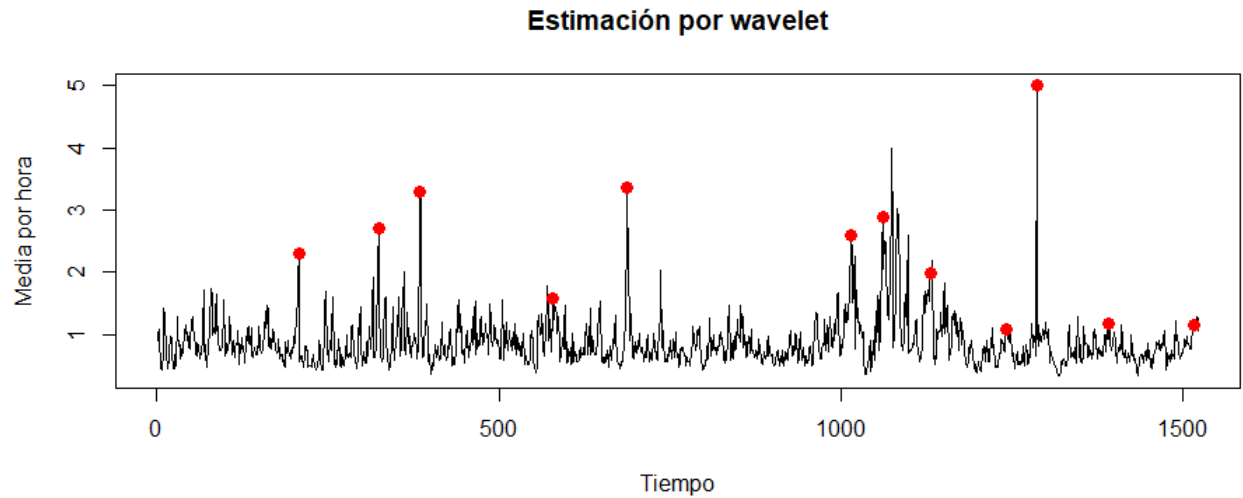


Figura 5.15: Puntos de cambio con wavelet en la media por hora del Seeing.

**Fuente:** Elaboración propia

En la Figura 5.15 se observa que el análisis wavelet seleccionó un total de 12 puntos de cambio, los primeros 5 puntos fueron detectados antes del primer punto de cambio detectado por el test de múltiples puntos de quiebre, de los cuales 4 marcan cambios fuertes en varianza entre valores de los 2 a 4 grados del arco de distorsión. Además, el quinto punto, 03/20-04AM, se detectó cerca del primer punto de cambio del test de múltiples puntos de quiebre. Luego, se detectaron 3 puntos de cambio fuertes en varianza sobre los 1.5 y por debajo los 3 grados del arco de distorsión. Finalmente, de los últimos 4 puntos de cambio que fueron detectados solo 3 de ellos alcanzaron valores por debajo de 1.5 grados de arco de distorsión, el valor más grande, 05/20-05AM, fue de 5 grados del arco de distorsión.

En el cuadro resumen 5.8 se muestran los resultados obtenidos para los distintos tests y del análisis wavelet en la detección de puntos de cambio de todas las aplicaciones a los distintos conjuntos de datos.

Test de puntos de cambio					
Bases de datos	MPQ	Chow	CUSUM	MOSUM	Wavelet
Río Nilo	1898	1898	1898	1889	1895
IPC	10/2008	10/2008	10/2008	07/2007	N/D
Desempleo (Mes/Año)	10/1991 11/1998 06/2005 09/2010 12/2015	03/1991	05/1989 06/2006	07/1988 10/1995 11/2001 07/2014	10/1994 09/1999 04/2004 05/2009
<i>Seeing</i> (Mes-Día-Hora)	03-21-03AM 04-14-10AM 05-08-04AM	05-08-02AM	04-15-04AM 05-08-02AM	04-16-03AM 05-08-04AM 06-10-03AM	01-30-09AM 02-11-07AM 02-18-07AM 03-10-03AM 03- 20-04AM 04-20-05AM 04-24-04AM 05-02-07AM 05- 20-05AM 05-31-06AM 06-10-05AM

Cuadro 5.8: Detección de puntos de cambio en los datos del flujo del caudal del río Nilo, la tasa de desempleo en Chile, el IPC General en Chile y la media por hora del Seeing.

**Fuente:** Elaboración propia.

En el cuadro 5.8 se observa que tanto los tests de múltiples puntos de quiebre (MPQ), Chow y CUSUM obtuvieron los mismos resultados en los datos del flujo anual del caudal del río Nilo y del IPC General en Chile. Por otro lado, en los datos de la tasa de desempleo en Chile los puntos de cambio obtenido por el test CUSUM mantuvieron una corta distancia de los puntos de cambio obtenidos por los tests de múltiples puntos de quiebre y Chow, lo cual posiblemente se deba a una diferencia en la metodología del test. En cambio el test de MOSUM fue capaz de detectar puntos de cambios más alejados que los otros tests clásicos, aunque eso no significa que su capacidad de detección sea mala, sino que es capaz de detectar puntos de cambio en zonas de poca variabilidad. Por último, el análisis wavelet logró detectar puntos de cambio a una corta distancia de los puntos que detectaron los tests clásicos en los datos del río Nilo, el IPC General en Chile y en la tasa de desempleo en Chile con una diferencia de hasta 5 años, para los datos del IPC General en Chile hubo problemas para que el análisis wavelet determinara los puntos de cambios óptimos, pero si el análisis wavelet hubiese tenido éxito, en los candidatos a puntos de cambio la ramificación 2 en el año 2008 hubiese sido el punto de cambio óptimo.

En la serie del Seeing hubo una detección de múltiples puntos de cambio, con excepción del test de Chow, aunque la detección de un punto a otro sea diferente en términos de hora, lograron detectarlos en los mismos días. Por ejemplo, los tests de múltiples puntos de quiebre, CUSUM y MOSUM lograron detectar puntos de cambio con una diferencia al menos de 2 horas (representados con el color magenta) en el 5.8 y de 2 días (representados con el color azul) en ese mismo cuadro. Por otro lado, el análisis wavelet fue el que se destacó por sobre los otros tests ya que fue capaz de detectar más puntos de cambio y varios de ellos están relacionados a las fechas de detección de los tests clásicos. Pese a que la mayoría de los puntos detectados por el análisis wavelet fueron en variaciones altas, hay algunos puntos detectados en variación pequeñas que mantienen una corta distancia con los diferentes puntos de cambio de los tests clásicos, posiblemente posea una mayor sensibilidad que los otros tests al momento de detectar puntos de cambio.

En el siguiente capítulo se van a exponer las conclusiones de esta investigación con las observaciones de los anteriores capítulos.

## Capítulo 6

# Conclusiones

Durante el transcurso de este trabajo de titulación se investigaron las metodologías clásicas de detección de puntos de cambio en series de tiempo, incluyendo el test de múltiples puntos de quiebre, el test de Chow, el test CUSUM y test MOSUM ampliamente utilizados en la literatura para detectar puntos de cambio en series de tiempo. Además se logró implementar cada uno de ellos en el software R.

Se estudió la teoría de análisis wavelet, los fundamentos, algunas de las principales familias wavelet, como por ejemplo, wavelet Haar, wavelet Daubichies, wavelet Morlet y wavelet Sombrero mexicano. Se profundizó en la transformada wavelet continua y discreta, y en el uso del escalograma y su implementación en el software R.

Se investigó el uso del análisis wavelet en series de tiempo. Se utilizó el análisis wavelet para la detección de puntos de cambio. Se hizo una revisión bibliográfica del estado del arte en la literatura para detectar puntos de cambio en series de tiempo. En este trabajo se evaluó el desempeño de la transformada wavelet continua (TWC) para la detección de puntos de cambio en series de tiempo a través de simulaciones.

Los resultados de las simulaciones fueron importantes para destacar la función que cumple cada uno de los test aplicados en esta investigación, sobre todo para el caso del análisis wavelet que posee un mayor alcance en la detección de cambios en la varianza. Se comparó el desempeño de los tests clásicos con el análisis wavelet en la detección de puntos de cambio en series de tiempo con los siguientes modelos, AR(1), MA(1), ARMA(1,1), ARIMA( 1,1,1) y GARCH(1,1). En general los tests clásicos tuvieron un mejor desempeño que el test basado en el análisis wavelet detectando cambios en la media salvo en el caso del modelo AR(1) donde el análisis wavelet muestra un desempeño superior. Respecto de los cambios en varianza, los tests clásicos no lograron superar el desempeño que tuvo el análisis wavelet, salvo en el modelo ARIMA(1,1,1) y el modelo GARCH(1,1).

Finalmente, se aplicó la metodología de detección de puntos de cambio en 4 conjuntos de datos reales utilizando tests clásicos y el test basado en el análisis wavelet.

Se detectó un cambio en la media del flujo anual del caudal del río Nilo, de  $1100 m^3$  a  $925 m^3$ , debido a la construcción de la presa de agua para disminuir los daños producto de las inundaciones en temporadas de alta precipitación. En la serie de la tasa de desempleo en Chile, se encontraron muchas fechas en donde la tasa era alta con valores de hasta un 12,5%. En otros casos disminuyó hasta un 5,5% producto de diversos acontecimientos en esas fechas. En los datos del IPC se observa un cambio brusco y de gran magnitud alrededor del año 2008 (tests MPQ, Chow y CUSUM) en la misma fecha, marcando un cambio en la media que va desde el 1.1% hasta el 0.8% producto de un aumento del precio del petróleo y de la crisis hipotecaria. Lamentablemente, el test wavelet no pudo determinar los candidatos óptimos para la serie temporal del IPC General en Chile. Por último, los datos del Seeing poseen gran variabilidad y tienen muchos datos faltantes porque las mediciones se obtienen durante la noche y algunos días no tienen mediciones. Los datos fueron registrados por segundo, lo que hace imposible determinar un cambio evidente entre los registros debido a su variabilidad. Por ende, se decidió tomar promedio por hora para realizar el análisis. Entre los resultados obtenidos, los tests clásicos lograron detectar cambios en la media de los datos, algunos leves y otros más fuertes. Es más, algunos de los tests consiguieron detectar el mismo punto de cambio con tan solo horas de diferencia. En cambio el test wavelet se destacó detectando más cambios en la variabilidad con márgenes que superaban los 3 puntos hasta los 5 puntos del arco de distorsión.

Como conclusión, hubo una estrecha cercanía entre los tests MPQ, Chow y CUSUM en la detección de puntos de cambio en las aplicaciones. Si bien hubieron algunas diferencias entre los valores obtenidos por cada uno de los tests, es posible que el punto que detecte se deba a una región o zona de análisis en donde se encuentren los distintos candidatos a ser puntos de cambio. El test MOSUM y el test Wavelet coincidieron en las fechas detectadas en los datos de la tasa de desempleo y los datos del Seeing. El análisis wavelet se desempeñó mejor detectando cambios en varianza. En algunos casos detectó varios puntos de cambio relativamente cerca a los encontrados por otros tests.

# Apéndice A

## Anexos

Versión del software utilizado

- Rstudio version 4.0.4

Librerías tests clásicos

- library(strucchange)

### 1. Test de múltiples puntos de quiebre

- se va a tomar de ejemplo los datos del flujo anual del caudal del río Nilo dentro del software Rstudio.

```
data("Nile")
```

- se construye un modelo de regresión con la constante uno como variable regresora.

```
bp.nile <- breakpoints(Nile ~ 1)
```

- a través de un resumen se obtienen los candidatos a ser puntos de cambio, a medida que aumentan los candidatos aumentan las particiones dentro de la serie temporal. Además se presetan los criterios minimizando los *RSS* (*residual sum of squares*) y *BIC* (*Bayesian information criterion*) utilizados para establecer la cantidad óptima de puntos de cambio mediante los valores más pequeños de ambos criterios. Sin embargo el criterio dominante es el mínimo valor que posea el BIC.

```
summary(bp.nile)
```

- el siguiente comando designa automáticamente la cantidad de puntos de cambio o segmentos de cambios de nivel siguiendo las normas de menor valor de los criterios RSS y BIC.

```
breakpoints(bp.nile)
```

- se ajusta un modelo de regresión con un punto de cambio como variable regresora, dependiendo de los resultados de `breakpoints(bp.nile)` la cantidad de puntos de cambio es diferente.

```
fm1 <- lm(Nile ~ breakfactor(bp.nile, breaks = 1))
```

- se gráfica la serie temporal utilizada como ejemplo.

```
plot(Nile, ylab="Flujo del río Nilo", xlab="Años")
```

- utilizando el comando `lines` se pueden observar el recorrido del modelo de regresión ajustado transformado a serie temporal donde se denota el cambio del valor de la media .

```
lines(ts(fitted(fm1), start = 1871), col = 4)
```

- a diferencia del comando anterior aca se utiliza la cantidad de puntos de cambio óptimos designados por el comando `breakpoint` marcando las fechas dentro de la serie temporal con una línea vertical segmentada.

```
lines(bp.nile,col="darkblue")
```

## 2. Test de Chow

- se construye un estadístico de prueba F el cual por defecto utiliza desde el 15 % de los datos en adelante .

```
fs.nile <- Fstats(Nile ~ 1)
```

- se presenta el test de hipótesis para el test de Chow, se rechaza la hipótesis nula cuando el supremo del estadístico F atraviese la banda de confianza con una probabilidad igual a un  $\alpha = 0,05$ .

```
sctest(fs.nile)
```

- El comando `breakpoint` es compatible con el algoritmo del test de Chow, por lo que se puede determinar la fecha en que se produce el punto de cambio.

```
breakpoints(fs.nile)
```

- se grafican los valores del estadístico F en función del test de hipótesis, la cúspide que se forma marca el único punto de cambio de la serie temporal.

```
plot(fs.nile,main=" Test de Chow",ylab=" EstadísticoF",xlab=" Años")
```

- al igual que en el test de Bai-Perron se utiliza el comando `lines` para marcar con una línea vertical segmentada la fecha en que se detectan los puntos de cambio.

```
lines(breakpoints(fs.nile),col="darkgreen")
```

### 3. Test CUSUM

- en la construcción del modelo de regresión se pueden utilizar 2 tipos de métodos, en este estudio solo se utilizara el método *OLS* (*Ordinary least square*) .

```
ocus.nile <- efp(Nile ~ 1, type = "OLS - CUSUM")
```

- el test de hipótesis consiste en que los residuos de un proceso browniano estándar puedan atravesar ambas bandas de confianzas por el lado positivo y negativo paralelas a la media de los residuos con un probabilidad de  $\alpha = 0,05$ . Además, se obtiene el máximo valor en valor absoluto del estadístico de prueba.

```
sctest(ocus.nile)
```

- se grafican los valores de los residuos del modelo de regresión en función del test de hipótesis. Se pueden detectar más de 1 punto de cambio atravesando las bandas de confianza por ambos lados formando cúspides, las cúspides más grandes marcan el punto de cambio anunciando la fecha de detección.

```
plot(ocus.nile, main="Test CUSUM por método OLS", ylab="Proceso de fluctuación empírico", xlab="Años")
```

### 3. Test MOSUM

- al igual que en el algoritmo del test CUSUM se construye el modelo de regresión para el test MOSUM. Sin embargo, el test MOSUM utiliza por defecto un componente adicional, una banda de análisis  $h = 0,15$  para calcular los residuos del movimiento browniano, esta banda de análisis comienza desde el 15 % de los datos de la serie temporal.

```
omus.nile <- efp(Nile ~ 1, type = "OLS - MOSUM")
```

- se presenta el test de hipótesis del test MOSUM que posee las mismas condiciones que en el test CUSUM.

```
sctest(omus.nile)
```

- se grafican los residuos del proceso browniano en función del test de hipótesis, donde los residuos atraviesan las bandas de confianza por ambos lados formando cúspides los cuales marcan el punto de cambio y la fecha.

```
plot(omus.nile, main="Test MOSUM por método OLS", ylab="Proceso de fluctuación empírico", xlab="Años")
```

#### 4. Wavelet

Las librerías para el análisis wavelet se deben instalar en el siguiente orden

- (1) `install.packages("wmtsa", repos="http://R-Forge.R-project.org")`
- (2) `install.packages("ifultools")`
- (3) `install.packages("splus2R")`

Luego se leen las librerías en el siguiente orden

- A. `library(splus2R)`
- B. `library(ifultools)`
- C. `library(wmtsa)`

- se construye una transformada wavelet continua *mexican hat* a partir de la serie temporal

```
wav.Nile <- wavCWT(Nile)
```

- se identifican los valores máximos dentro de la transformada wavelet formando un subconjunto de ramificaciones, estas ramificaciones pueden ser largas o cortas.

```
wavtree.Nile <- wavCWTTree(wav.Nile)
```

- se utiliza el comando *summary* para analizar los candidatos a ser puntos de cambio, entre los datos más relevantes se encuentra el *End time* (o la fecha), el *Length* y los *Octaves*.

```
summary(wavtree.Nile)
```

- se grafican los resultados de *summary(wavtree.Nile)* a través de  $\log_2(\text{scale})$  en el eje Y y las fechas en el eje X. El valor  $\log_2(\text{scale})$  está relacionado al valor de *Octaves* y el largo de las ramas está relacionado con el *length*.

```
plot(wavtree.Nile)
```

- se determinan los puntos de cambios de las ramas cuyo valor *length* sea mayor o igual a 10 y el mínimo ruido local sea 3.

```
Peak.Nile <- wavCWTPeaks(wavtree.Nile)
```

- se gráfica la serie de tiempo para luego determinar el punto de cambio.

```
plot(Nile, main="Estimación por Wavelet Peaks", ylab="Flujo", xlab="Años")
```

- El punto de cambio se muestra a través del comando *points* con la fecha en que se detecta el punto de cambio.

```
points(Peak.Nile, pch=16, col="red", cex=1.2)
```

# Referencias

- Alarcon-Aquino, V., y Barria, J. (2009). Change detection in time series using the maximal overlap discrete wavelet transform. *Latin American Applied Research*, 39, 145-152. Descargado de [https://www.researchgate.net/publication/230710713.Change\\_detection\\_in\\_time\\_series\\_using\\_the\\_maximal\\_overlap\\_discrete\\_wavelet\\_transform](https://www.researchgate.net/publication/230710713.Change_detection_in_time_series_using_the_maximal_overlap_discrete_wavelet_transform)
- Aminikhanghahi, S., y Cook, D. (2016). A survey of methods for time series change point detection. *Knowledge and Information Systems*, 51(2), 339-367. doi: 10.1007/s10115-016-0987-z
- Bai, J., y Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66(1), 47. doi: 10.2307/2998540
- Banda-Ortiz, H., y González, A. (2020). El impacto de las crisis financieras en los cambios estructurales de los factores de fecundidad. *PANORAMA ECONÓMICO*, 15(30), 105-123.
- Bravo, R. (2019, 06). *Calidad del cielo: Seeing y transparencia*. Descargado de <https://astroaficion.com/2019/08/19/calidad-cielo-seeing-transparencia/>
- Brown, R. L., Durbin, J., y Evans, J. M. (1975). Techniques for Testing the Constancy of Regression Relationships Over Time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(2), 149-163. doi: 10.1111/j.2517-6161.1975.tb01532.x
- Chan, T., Chin, C., , Li, Y., Shafiee, E., y Sela, L. (2021). A better estimation of wave arrival time in water distribution networks using wavelet knee (wane). *Advanced Engineering Informatics*, 48, 101-111. doi: 10.1016/j.aei.2021.101287
- Chen, J., y Gupta, A. (2001). On change point detection and estimation. *Communications in Statistics - Simulation and Computation*, 30(3), 665-697. doi: 10.1081/SAC-100105085
- Chu, C.-S. J., Hornik, K., y Kuan, C.-M. (1995). MOSUM tests for parameter constancy. *Biometrika*, 82(3), 603-617. doi: 10.1093/biomet/82.3.603
- Cowan, K., Micco, A., Mizala, A., Pagés, C., y Romaguera, P. (2003). *Un Diagnóstico del Desempleo en Chile* (Inf. Téc.). Descargado de [https://www.dipres.gob.cl/598/articles-21661\\_doc\\_pdf.pdf](https://www.dipres.gob.cl/598/articles-21661_doc_pdf.pdf)
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. Philadelphia, USA: Society for Industrial and Applied Mathematics (SIAM).
- Du, P., Kibbe, W. A., y Lin, S. M. (2006). Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics*, 22(17), 2059-2065. doi: 10.1093/bioinformatics/btl355
- Fredes, C., y Hidalgo, C. (2019). *Detección de inestabilidad de los parámetros en modelos lineales* [Trabajo de grado]. Pontificia Universidad católica de Valparaíso, Chile: Valparaíso.
- GeoEnciclopedia. (2016, 05). *Río Nilo - Información y Características*. Descargado de <https://www.geoenciclopedia.com/rio-nilo/>
- Instituto Nacional de Estadísticas. (2020). *Documento Metodológico Encuesta Nacional de Empleo (ENE)* (Inf. Téc.). Descargado de <https://www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion>

- Keković, G., y Sekulić, S. (2019). Detection of change points in time series with moving average filters and wavelet transform: Application to eeg signals. *Neurophysiology*, 51(1), 2–8. doi: 10.1007/s11062-019-09783-y
- Laval, J. (2007). *Informe de política monetaria septiembre 2007* (Inf. Téc.). Banco central de Chile.
- Laval, J. (2009). *Informe de estabilidad financiera segundo semestre 2008* (Inf. Téc.). Banco central de Chile.
- Misiti, M., Misiti, Y., Oppenheim, G., y Poggi, J. (2007). *Wavelets and their Applications*. Hoboken, NJ, Estados Unidos: Wiley.
- Morales, N. (2019). *Identificación de cambios estructurales en los principales índices bursátiles de los países que conforman el mila entre los años 2008 y 2018* [Trabajo de grado]. Universidad de la Salle, Colombia: Bogotá.
- Olsen, L., Chaudhuri, P., y Godtlielsen, F. (2008). Multiscale spectral analysis for detecting short and long range change points in time series. *Computational Statistics Data Analysis*, 52(7), 3310–3330. doi: 10.1016/j.csda.2007.10.027
- Ploberger, W., y Krämer, W. (1992). The Cusum Test with Ols Residuals. *Econometrica*, 60(2), 271. doi: 10.2307/2951597
- Raimondo, M., y Tajvidi, N. (2004). A peaks over threshold model for change-point detection by wavelets. *Statistica Sinica*, 14, 395–412.
- Sastri, T., Flores, B., y Valdés, J. (1989). Detecting points of change in time series. *Computers Operations Research*, 16(3), 271-293. doi: 10.1016/0305-0548(89)90027-0
- Singh, P. (2015). *Applications of Soft Computing in Time Series Forecasting* (Vol. 330). doi: 10.1007/978-3-319-26293-2
- Su, W., Zhu, Y., Liu, F., y Hu, K. (2013). On-line outlier and change point detection for time series. *Journal of Central South University*, 20(1), 114–122. doi: 10.1007/s11771-013-1466-2
- Tsay, R. (2001). *Analysis of Financial Time Series: Financial Econometrics* (1.<sup>a</sup> ed.). Danvers, United States: Wiley-Interscience.
- Valera, M. (2020). *Detección de un punto de cambio en la media de un modelo de series de tiempo* [Trabajo de grado]. Pontificia Universidad Javeriana Cali. Colombia: Santiago de Cali.
- Vega, C. (2003). *Aplicación de técnicas wavelet al análisis de series de temporales* (Tesis Doctoral no publicada). Universidad de Granada, Departamento de Estadística e investigación Operativa.
- Xie, P., Gu, H., Sang, Y., Wu, Z., y Singh, V. (2019). Comparison of different methods for detecting change points in hydroclimatic time series. *Journal of Hydrology*, 577, 123973. doi: 10.1016/j.jhydrol.2019.123973
- Yazgan, M., y Özkan, H. (2015). Detecting structural changes using wavelets. *Finance Research Letters*, 12, 23–37. doi: 10.1016/j.frl.2014.12.003
- Zeileis, A., Kleiber, C., Krämer, W., y Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics Data Analysis*, 44(1-2), 109–123. doi: 10.1016/s0167-9473(03)00030-6
- Zeileis, A., Leisch, F., Hornik, K., y Kleiber, C. (s.f.). *strucchange: An R Package for Testing for Structural Change in Linear Regression Models*. Descargado de <https://cran.r-project.org/web/packages/strucchange/vignettes/strucchange-intro.pdf>