

*Modelamiento de Procesos Autorregresivos de  
Umbrales Estacionales*

JOAQUÍN GONZÁLEZ BORJA

CÓDIGO: 1832586



UNIVERSIDAD NACIONAL DE COLOMBIA  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE ESTADÍSTICA  
BOGOTÁ, D.C.  
2019

*Modelamiento de Procesos Autorregresivos de  
Umbrales Estacionales*

JOAQUÍN GONZÁLEZ BORJA

CÓDIGO: 1832586

DISERTACIÓN PRESENTADA PARA OPTAR AL TÍTULO DE  
DOCTOR EN ESTADÍSTICA

DIRECTOR  
FABIO HUMBERTO NIETO SÁNCHEZ  
DOCTOR EN ESTADÍSTICA

LÍNEA DE INVESTIGACIÓN  
SERIES DE TIEMPO

GRUPO DE INVESTIGACIÓN  
SERIES DE TIEMPO



UNIVERSIDAD NACIONAL DE COLOMBIA  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE ESTADÍSTICA  
BOGOTÁ, D.C.  
2019

**Título en español**

Modelamiento de Procesos Autorregresivos de Umbrales Estacionales.

**Title in English**

Modeling of Seasonal Threshold Autoregressive Processes.

**Resumen:** Fluctuaciones estacionales frecuentemente se hallan en muchas series de tiempo. En adición, la no linealidad y la relación con otras series de tiempo son comportamientos prominentes de muchas de tales series. En este artículo, consideramos el modelamiento de procesos autorregresivos de umbrales estacionales multiplicativos con entrada exógena (TSARX), los cuales incorporan en forma explícita y simultánea estacionalidad multiplicativa y no linealidad de umbrales. La estacionalidad es modelada a ser estocástica y dependiente del régimen. El modelo propuesto es un caso especial de un proceso autorregresivo de umbrales con entrada exógena (TARX). Desarrollamos un procedimiento basado en métodos Bayesiano para identificar el modelo, estimar parámetros, validar el modelo y calcular pronósticos. En la etapa de identificación del modelo, presentamos una prueba estadística de estacionalidad multiplicativa por regímenes. La metodología propuesta es ilustrada con un ejemplo simulado y aplicada a datos empíricos económicos.

**Abstract:** Seasonal fluctuations are often found in many time series. In addition, non-linearity and the relationship with other time series are prominent behaviors of several, of such series. In this paper, we consider the modeling of multiplicative seasonal threshold autoregressive processes with exogenous input (TSARX), which explicitly and simultaneously incorporate multiplicative seasonality and threshold nonlinearity. Seasonality is modeled to be stochastic and regime dependent. The proposed model is a special case of a threshold autoregressive process with exogenous input (TARX). We develop a procedure based on Bayesian methods to identify the model, estimate parameters, validate the model and calculate forecasts. In the identification stage of the model, we present a statistical test of regime dependent multiplicative seasonality. The proposed methodology is illustrated with a simulated example and applied to economic empirical data.

**Palabras clave:** Análisis de series de tiempo, entrada exógena, estacionalidad multiplicativa, estadística Bayesiana, modelos TSARX, no linealidad de umbrales.

**Keywords:** Bayesian statistics, exogenous input, multiplicative seasonality, threshold nonlinearity, time series analysis, TSARX models.

# Nota de aceptación

Trabajo de tesis

Aprobado

“Mención Meritoria o Laureada”

---

Jurado

---

Jurado

---

Jurado

---

Director  
Fabio Humberto Nieto Sánchez

Bogotá, D.C., 2019

---

---

# Índice general

---

---

Índice de tablas	II
Índice de figuras	IV
Introducción	VII
<b>1. El modelo TSARX</b>	<b>1</b>
1.1. La especificación del modelo. . . . .	1
1.2. La función de verosimilitud condicional. . . . .	3
1.3. Estimación de parámetros de un modelo TSARX. . . . .	5
1.4. La prueba Bayesiana de estacionalidad multiplicativa por regímenes. . . . .	11
1.5. Estimación del número de regímenes. . . . .	15
1.6. Diagnósticos con los residuales del modelo. . . . .	17
1.7. El cálculo de pronósticos. . . . .	18
1.8. La metodología propuesta. . . . .	19
<b>2. Un estudio de simulación</b>	<b>21</b>
2.1. Modelo 1. . . . .	21
2.1.1. Análisis exploratorio de los datos. . . . .	21
2.1.2. Resultados de la estimación. . . . .	24
2.1.3. Diagnósticos del modelo. . . . .	28
2.1.4. Cálculo y evaluación de pronósticos. . . . .	30
2.2. Modelo 2. . . . .	32
2.2.1. Análisis exploratorio de los datos. . . . .	32
2.2.2. Resultados de la estimación. . . . .	36

---

2.2.3. Diagnósticos del modelo. . . . .	40
2.2.4. Cálculo y evaluación de pronósticos. . . . .	41
<b>3. Una aplicación en economía</b>	<b>44</b>
3.1. Análisis exploratorio de los datos. . . . .	44
3.2. Resultados de la estimación. . . . .	50
3.3. Diagnósticos del modelo. . . . .	54
3.4. Cálculo y evaluación de pronósticos. . . . .	56
<b>Conclusiones</b>	<b>59</b>
<b>Bibliografía</b>	<b>60</b>

---

---

# Índice de tablas

---

---

2.1. Pruebas de raíces unitarias y raíces unitarias estacionales para el modelo 1. . . . .	22
2.2. Probabilidades medias a posteriori y DIC promedio para selección del número de regímenes para el modelo 1 (100 repeticiones). . . . .	25
2.3. Los mejores tres modelos identificados usando el método GVS, para el modelo 1 de 100 repeticiones. . . . .	25
2.4. Estimación de parámetros para el modelo 1 de 100 repeticiones. . . . .	26
2.5. Porcentajes de 100 repeticiones, para el cual el modelo ajustado es hallado ser adecuado para el modelo 1, usando residuales generalizados. . . . .	30
2.6. Pronósticos para las variables $X_t$ y $Z_t$ del modelo 1 con 100 repeticiones. . . . .	31
2.7. Pruebas de raíces unitarias y raíces unitarias estacionales para el modelo 2. . . . .	33
2.8. Probabilidades medias a posteriori y DIC promedio para selección del número de regímenes para el modelo 2 (100 repeticiones). . . . .	36
2.9. Los mejores tres modelos identificados usando el método GVS, para el modelo 2 de 100 repeticiones. . . . .	37
2.10. Estimación de parámetros para el modelo 2 de 100 repeticiones. . . . .	38
2.11. Porcentajes de 100 repeticiones, para el cual el modelo ajustado es hallado ser adecuado para el modelo 2, usando residuales generalizados. . . . .	40
2.12. Pronósticos para las variables $X_t$ y $Z_t$ del modelo 2 con 100 repeticiones. . . . .	42
3.1. Pruebas de raíces unitarias y raíces unitarias estacionales para los datos empíricos. . . . .	46
3.2. Ajuste de un modelo AR(26) multiplicativo para la variable de umbrales. . . . .	49
3.3. Prueba de no linealidad de umbrales (Tsay, 1998) para los datos empíricos, considerando diferentes órdenes autorregresivos y valores d. . . . .	50

---

3.4. Prueba de estacionalidad multiplicativa por regímenes para los datos empíricos. . . . .	51
3.5. Probabilidades a posteriori y DIC para la selección del número de regímenes de un modelo TSARX para los datos empíricos. . . . .	51
3.6. Los mejores tres modelos identificados usando el método GVS, para los datos empíricos. . . . .	52
3.7. Estimación de parámetros para los datos empíricos. . . . .	53
3.8. Pruebas de Ljung y Box para diferentes rezagos (LB) y de Jarque y Bera (JB) aplicada a los residuales generalizados del modelo TSARX ajustado a los datos empíricos. El valor $p$ de la prueba está en paréntesis. . . . .	54
3.9. Pronósticos para las variables de interés y de umbrales de los datos empíricos, mediante un TSARX. . . . .	56

---



---

# Índice de figuras

---



---

2.1.	(a) $X_t$ variable de interés, (b) $Z_t$ variable de umbrales, modelo 1. . . .	22
2.2.	Para $X_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 1. . . . .	23
2.3.	Para $Z_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 1. . . . .	23
2.4.	Para $X_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por meses, modelo 1. . . . .	24
2.5.	Gráficos cumuplot para: (a) $a_1^{(1)}$ , (b) $b_1^{(1)}$ , (c) $b_2^{(1)}$ , (d) $c_1^{(1)}$ (e) $a_0^{(2)}$ , (f) $b_1^{(2)}$ , (g) $(h^{(1)})^2$ , (h) $(h^{(2)})^2$ , (i) r, modelo 1. . . . .	27
2.6.	Gráficos $Z$ -score de Geweke: (a) $a_1^{(1)}$ , (b) $b_1^{(1)}$ , (c) $b_2^{(1)}$ , (d) $c_1^{(1)}$ (e) $a_0^{(2)}$ , (f) $b_1^{(2)}$ , (g) $(h^{(1)})^2$ , (h) $(h^{(2)})^2$ , (i) r, modelo 1. . . . .	28
2.7.	(a) FAC muestral, (b) FACP muestral, residuales generalizados del modelo 1. . . . .	29
2.8.	(a) Gráfico CUSUM, (b) gráfico CUSUMSQ, residuales generalizados del modelo 1. . . . .	29
2.9.	Distribuciones predictivas para la variable $X_t$ modelo 1, (a) h=1, (b) h=2, (c) h=3, (d) h=11, (e) h=12, (f) h=13, pasos adelante. . . . .	31
2.10.	Distribuciones predictivas para la variable $Z_t$ modelo 1, (a) h=1, (b) h=2, (c) h=3, (d) h=11, (e) h=12, (f) h=13, pasos adelante. . . . .	32
2.11.	(a) $X_t$ variable de interés, (b) $Z_t$ variable de umbrales, modelo 2. . . .	33
2.12.	Para $X_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 2. . . . .	34
2.13.	Para $Z_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 2. . . . .	34
2.14.	Para $X_t$ , (a) AR multiplicativo lineal con BIC mínimo, (b) diagramas de caja por trimestres, modelo 2. . . . .	35

2.15. Para $Z_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por trimestres, modelo 2. . . . .	35
2.16. Gráficos cumuplot para: (a) $b_1^{(1)}$ , (b) $c_1^{(1)}$ , (c) $a_2^{(2)}$ , (d) $b_1^{(2)}$ (e) $b_1^{(3)}$ , (f) $b_2^{(3)}$ , (g) $(h^{(2)})^2$ , (h) $(h^{(3)})^2$ , (i) $r_1$ , modelo 2. . . . .	39
2.17. Gráficos $Z$ -score de Geweke: (a) $b_1^{(1)}$ , (b) $c_1^{(1)}$ , (c) $a_2^{(2)}$ , (d) $b_1^{(2)}$ (e) $b_1^{(3)}$ , (f) $b_2^{(3)}$ , (g) $(h^{(2)})^2$ , (h) $(h^{(3)})^2$ , (i) $r_1$ , modelo 2. . . . .	39
2.18. (a) FAC muestral, (b) FACP muestral, residuales generalizados del modelo 2. . . . .	40
2.19. (a) Gráfico CUSUM, (b) gráfico CUSUMSQ, residuales generalizados del modelo 2. . . . .	41
2.20. Distribuciones predictivas para la variable $X_t$ modelo 2, (a) $h=1$ , (b) $h=2$ , (c) $h=3$ , (d) $h=4$ , (e) $h=5$ , (f) $h=6$ , pasos adelante. . . . .	42
2.21. Distribuciones predictivas para la variable $Z_t$ modelo 2, (a) $h=1$ , (b) $h=2$ , (c) $h=3$ , (d) $h=4$ , (e) $h=5$ , (f) $h=6$ , pasos adelante. . . . .	43
3.1. (a) $TD_t$ , tasa de desempleo total mensual Colombiana, (b) $ISE_t$ , índice de seguimiento económico mensual Colombiano, con base 2005. . . . .	45
3.2. (a) $X_t$ , crecimiento de la tasa de desempleo total mensual Colombiana, (b) $Z_t$ , crecimiento mensual de la transformación logaritmo del $ISE_t$ . . . . .	46
3.3. Para $X_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, datos empíricos. . . . .	47
3.4. Para $Z_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, datos empíricos. . . . .	47
3.5. Para $X_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por meses, datos empíricos. . . . .	48
3.6. Para $Z_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por meses, datos empíricos. . . . .	48
3.7. Residuales del modelo AR(26) multiplicativo para $Z_t$ , (a) series de tiempo, (b) FAC muestral, (c) QQ-norm y (d) estadísticas de Ljung y Box con diferentes rezagos. . . . .	49
3.8. Gráficos cumuplot para: (a) $a_0^{(1)}$ , (b) $a_1^{(1)}$ , (c) $b_1^{(1)}$ , (d) $a_1^{(2)}$ , (e) $b_1^{(2)}$ , (f) $b_2^{(2)}$ , (g) $(h^{(1)})^2$ , (h) $(h^{(2)})^2$ , (i) $r$ , datos empíricos. . . . .	53
3.9. Gráficos $Z$ -score de Geweke: (a) $a_0^{(1)}$ , (b) $a_1^{(1)}$ , (c) $b_1^{(1)}$ , (d) $a_1^{(2)}$ , (e) $b_1^{(2)}$ , (f) $b_2^{(2)}$ , (g) $(h^{(1)})^2$ , (h) $(h^{(2)})^2$ , (i) $r$ , datos empíricos. . . . .	54
3.10. (a) FAC muestral, (b) FACP muestral, residuales generalizados de los datos empíricos. . . . .	55
3.11. (a) CUSUM, (b) CUSUMSQ residuales generalizados de los datos empíricos. . . . .	55

- 
- 3.12. Distribuciones predictivas para la variable  $X_t$  datos empíricos, (a) h=1, (b) h=2, (c) h=3, (d) h=11, (e) h=12, (f) h=13, pasos adelante. 57
- 3.13. Distribuciones predictivas para la variable  $Z_t$  datos empíricos, (a) h=1, (b) h=2, (c) h=3, (d) h=11, (e) h=12, (f) h=13, pasos adelante. 57
- 3.14. Gráficas de pronósticos, (a) para  $X_t$ , (b) para  $Z_t$ , datos empíricos. . . 58

---

---

# Introducción

---

---

Los modelos autorregresivos de umbrales (TAR) propuestos por Tong (1978) y Tong y Lim (1980) han tenido una amplia aplicación en diversas áreas, incluyendo economía, finanzas, hidrología, meteorología y biología. Los modelos TAR asumen que los valores de un proceso estocástico  $\{Z_t\}$  (proceso de umbrales) determina tanto los valores del proceso estocástico  $\{X_t\}$  (proceso de interés) como su dinámica. Si el proceso de umbrales es el mismo proceso de interés pero rezagado, el modelo recibe el nombre de SETAR (*Self-Exciting TAR*, Tong, 1990).

El uso de métodos Bayesianos y técnicas MCMC aplicados a los modelos TAR y SETAR, pueden hallarse entre otros trabajos en: Chen y Lee (1995) para realizar inferencia en modelos SETAR, Chen (1998) para modelos SETAR con entradas exógenas (SETARX), So, Chen y Liu (2006) para la selección de modelos TARX con errores heterocedásticos, Nieto (2005) para la estimación de parámetros y datos faltantes en forma conjunta en modelos TAR con ruido Gaussiano, Nieto (2008) y Vargas (2012) en el cálculo de pronósticos en modelos TAR. Mientras que, Nieto, Zhang y Li (2013) proponen un procedimiento para identificar y estimar modelos TAR siguiendo la técnica MCMC de saltos reversibles, Zhang y Nieto (2015) para la estimación de parámetros, datos faltantes y el cálculo de pronósticos en modelos TAR con ruido t-Student. Además, Nieto y Moreno (2016) presentan nuevas características de los modelos TAR en términos de tres tipos de distribuciones marginales condicionales, Calderón y Nieto (2017) proponen una metodología para ajustar un modelo TARX multivariado a series de tiempo con datos faltantes, Hoyos (2006) y Vivas (2011) implementan la metodología propuesta en Nieto (2005) ajustando modelos TAR a series de tiempo económicas Colombianas. Adicionalmente, Hansen (2011) y Chen, So y Liu (2011) dan una revisión de teoría y aplicaciones de los modelos TAR en economía y finanzas, respectivamente.

Los modelos TAR han resultado útiles para datos no estacionales o ajustados estacionalmente que exhiban no linealidad de umbrales. Sin embargo, por ejemplo las tasas de crecimiento de muchas series de tiempo económicas trimestrales o mensuales presentan cierto comportamiento estacional. Además, la no linealidad de umbrales y la relación con otras series de tiempo son características típicas de

muchas de tales series.

Tong (2015) da un compendio de lo desarrollado respecto a los modelos TAR, desde su aparición y posibles temas a investigar entre los que se considera el estudio de la estacionalidad. En muchas ocasiones, resulta conveniente estudiar la dinámica estacional presente en las series de tiempo, ya que pueden llevar información relevante sobre el comportamiento de agentes económicos. Diversos estudios empíricos muestran que muchos métodos de ajuste estacional, llevan a distorsionar los datos en forma severa, las fases del ciclo económico y no linealidad son afectados por dicho ajuste. Estas discusiones aparecen en Hylleberg, Jørgensen y Sørensen (1993), Ghysels (1994), Franses (1996), Franses y Paap (1999), entre muchos otros. En el caso de usar datos sin ajuste estacional, existe la posibilidad que las fluctuaciones estacionales cambien sobre el tiempo y quizás cambien de acuerdo a las fases del ciclo económico como lo sugieren Canova y Ghysels (1994). Por consiguiente, para un mayor conocimiento de ciertas variables económicas se requiere cuidadosamente el modelamiento de su comportamiento estacional.

Una amplia variedad de modelos de series de tiempo lineales para capturar la dinámica estacional se encuentran en la literatura. Entre los modelos más populares de estacionalidad son los modelos con un patrón estacional determinístico usando variables dummies, con raíces unitarias estacionales, con diferenciación estacional y autorregresivos periódicos, véase Ghysels y Osborn (2001) para una revisión de los modelos en mención. En el modelamiento de la estacionalidad también se puede direccionar por los modelos ARIMA estacionales propuestos por Box y Jenkins (1970), donde la estacionalidad es caracterizada por coeficientes autorregresivos y promedios móviles, ambos estacionales.

Usando un enfoque frecuentista y datos sin ajuste estacional, Franses, De Bruin y Van Dijk (2000), Franses y Van Dijk (2005) examinan el funcionamiento de diversos modelos de series de tiempo lineales y autorregresivos de transición suave (STAR), considerando estacionalidad determinística en series de datos de producción industrial trimestral. Crespo (2001) contempla la estacionalidad como determinística y dependiente del régimen en modelos SETAR y usa su modelo en series de tiempo de tasas de desempleo trimestral. Mientras que, De Gooijer y Vidiella-i-Anguera (2003) definen un modelo SETAR estacional, donde cada régimen tiene la forma de un modelo autorregresivo estacional multiplicativo. Ellos realizan el estudio de series de tiempo de tasa de inflación mensual de varios países encontrando características dinámicas más realistas que al aplicar un SETAR usual o un SAR (procesos autorregresivos estacionales).

En este trabajo, proponemos un modelo estadístico denominado modelo autorregresivo estacional multiplicativo con entrada exógena de umbrales (TSARX) el cual incorpora explícitamente y simultáneamente estacionalidad multiplicativa por regímenes, que surgen de estructuras de autocorrelación periódicas y no linealidad

---

de umbrales. El modelo propuesto, surge de imponer ciertas restricciones sobre los parámetros de un modelo TARX. Se usan métodos Bayesianos y técnicas MCMC para determinar la presencia de estacionalidad multiplicativa por regímenes, para obtener la estimación conjunta de parámetros, cálculo de residuales generalizados y el cálculo de pronósticos. La metodología desarrollada es evaluada e ilustrada con diversos modelos TSARX simulados y con una aplicación empírica analizando las tasas de crecimiento del desempleo total mensual y del índice de seguimiento económico mensual en Colombia.

Este trabajo está organizado como sigue: en el capítulo 1, se presenta el modelo TSARX y la metodología propuesta. En el capítulo 2 se muestra el estudio de simulación. Una ilustración de la metodología propuesta se da en el capítulo 3 con datos de series de tiempo económicas. Finalmente, se dan las conclusiones en el capítulo 4.

# CAPÍTULO 1

---

---

## El modelo TSARX

---

---

### 1.1. La especificación del modelo.

Sean  $\{X_t\}$  y  $\{Z_t\}$  procesos estocásticos relacionados por la ecuación

$$X_t = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i} + \sum_{u=1}^{K_j} b_u^{(j)} X_{t-su} - \sum_{i=1}^{k_j} \sum_{u=1}^{K_j} a_i^{(j)} b_u^{(j)} X_{t-i-su} + \sum_{v=1}^{q_j} c_v^{(j)} Z_{t-v} + h^{(j)} \epsilon_t, \quad (1.1)$$

si  $r_{j-1} < Z_{t-d} \leq r_j$ , para todo  $t \in \mathbb{Z}$  ( $\mathbb{Z}$  es el conjunto de los números enteros) y algún  $j = 1, \dots, l$ ,  $d$  es un entero no negativo. El modelo asume  $l$  regímenes definidos por los valores umbrales  $r_j$  que satisfacen  $-\infty = r_0 < r_1 < \dots < r_{l-1} < r_l = \infty$  y la variable de umbrales rezagada  $d$  periodos  $Z_{t-d}$ . Además, se considera la variable de umbrales como una entrada exógena, en el sentido de que no hay retroalimentación de  $\{X_t\}$  hacia  $\{Z_t\}$ . Los números reales  $a_i^{(j)}$ ,  $b_u^{(j)}$ ,  $c_v^{(j)}$  y  $h^{(j)}$ ,  $i = 0, 1, \dots, k_j$ ,  $u = 1, \dots, K_j$ ,  $v = 1, \dots, q_j$ ; denotan, respectivamente los coeficientes autorregresivos no estacionales, autorregresivos estacionales, autorregresivos de la entrada exógena y las ponderaciones de las varianzas para cada régimen  $j = 1, \dots, l$ ; y reciben el nombre de parámetros no estructurales del modelo. Los números enteros no negativos  $k_1, \dots, k_l$  son los órdenes autorregresivos no estacionales,  $K_1, \dots, K_l$  son los órdenes autorregresivos estacionales; mientras que,  $q_1, \dots, q_l$  representan los órdenes autorregresivos de la entrada exógena. El número de regímenes junto a los órdenes autorregresivos, los valores umbrales  $\mathbf{r} = (r_1, \dots, r_{l-1})'$  y  $d$  llamado parámetro de rezago, conforman el denominado conjunto de parámetros estructurales del modelo. Se tiene que  $s$  es el periodo estacional del modelo que se asume conocido, donde en forma usual se toma  $s = 4$  para series de tiempo trimestrales y  $s = 12$  para series de tiempo mensuales. Además, el modelo (1) podría ser usado para series de tiempo que exhiben comportamientos estacionales diarios, semanales, anuales, entre otros. Se consideran  $k_j < s$  para todo  $j$  ( $j = 1, \dots, l$ ).

Adicionalmente,  $\{\epsilon_t\}$  es un proceso de ruido blanco Gaussiano con media ce-

ro y varianza uno (RBG(0,1)).  $\{\epsilon_t\}$  y  $\{Z_t\}$  son mutuamente independientes,  $E(X_w \epsilon_t) = 0$  para  $w < t$  y  $Z_w$  con el conjunto  $\{X_t, X_{t-1}, \dots\}$  son mutuamente independientes para todo  $w > t$ . Para describir el comportamiento estocástico de  $\{Z_t\}$ , en forma adicional se asume que  $\{Z_t\}$  es una cadena de Markov homogénea de orden  $m \geq 1$ , con función de densidad inicial  $f(\cdot)$  y función de densidad kernel  $f_m(\cdot|\cdot)$ , con respecto a una medida de Lebesgue. La cadena de Markov  $\{Z_t\}$  converge débilmente a una distribución  $F$ . Para mayores detalles sobre cadenas de Markov con espacio de estado general, se puede consultar Meyn y Tweedie (2009). La ecuación (1.1) describe un sistema dinámico sin retroalimentación, con entrada observable  $\{Z_t\}$  y salida observable  $\{X_t\}$ , que es presentado en Tong (1990) pero sin considerar el modelamiento de la estacionalidad multiplicativa por regímenes.

Se usa el símbolo  $\text{TSARX}(l; k_1, \dots, k_l; K_1, \dots, K_l; q_1, \dots, q_l)_s$ , para denotar (1.1) y se dice que  $\{X_t\}$  es un proceso autorregresivo estacional multiplicativo con entrada exógena de umbrales, con  $\{Z_t\}$  como su proceso de umbrales y  $s$  su periodo estacional. Se define el vector de parámetros no estructurales del modelo TSARX como  $\boldsymbol{\theta}_{ne} = (A'_1, \dots, A'_l, B'_1, \dots, B'_l, C'_1, \dots, C'_l, \mathbf{h}')'$ , donde  $A_j = (a_0^{(j)}, a_1^{(j)}, \dots, a_{k_j}^{(j)})$ ,  $B_j = (b_1^{(j)}, \dots, b_{K_j}^{(j)})$ ,  $C_j = (c_1^{(j)}, \dots, c_{q_j}^{(j)})$ ,  $j = 1, \dots, l$  y  $\mathbf{h} = ((h^{(1)})^2, \dots, (h^{(l)})^2)$  y  $\boldsymbol{\theta}_e = (k_1, \dots, k_l, K_1, \dots, K_l, q_1, \dots, q_l, \mathbf{r}', d, l)'$  es el vector de parámetros estructurales del modelo. Se denota  $\boldsymbol{\theta} = (\boldsymbol{\theta}_x, \boldsymbol{\theta}_z)$  como el vector de parámetros completo, donde  $\boldsymbol{\theta}_x = (\boldsymbol{\theta}'_{ne}, \boldsymbol{\theta}'_e)$  es el vector de parámetros del modelo TSARX y  $\boldsymbol{\theta}_z$  el vector de parámetros de la cadena de Markov.

Siguiendo resultados teóricos sobre los modelos TAR, dados en Nieto y Moreno (2016), y haciendo una extensión a los modelos TSARX, se tienen las siguientes consideraciones:

- (i) La función de distribución acumulativa condicional de  $X_t$  dados  $\mathbf{x}_{1,t-1} = (x_1, \dots, x_{t-1})$ ,  $\mathbf{z}_{1,t-1} = (z_1, \dots, z_{t-1})$  y  $z_{t-d} \in R_j = (r_{j-1}, r_j]$ , para algún  $j = 1, \dots, l$ ,  $t > K = \max\{k_j + sK_j, q_j, j = 1, \dots, l\}$ ,  $F_{t-1,j}(x_t | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, z_{t-d} \in R_j)$  es una normal con media

$$a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i} + \sum_{u=1}^{K_j} b_u^{(j)} x_{t-su} - \sum_{i=1}^{k_j} \sum_{u=1}^{K_j} a_i^{(j)} b_u^{(j)} x_{t-i-su} + \sum_{v=1}^{q_j} c_v^{(j)} z_{t-v}$$

y varianza  $(h^{(j)})^2$ , para cada  $j = 1, \dots, l$ .

- (ii) La función de distribución acumulativa condicional de  $X_t$  dados  $\mathbf{x}_{1,t-1}$ ,  $\mathbf{z}_{1,t-1}$ ,  $t > K$ , está dada por

$$F_t(x_t | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}) = \sum_{j=1}^l p_{t,j} F_{t-1,j}(x_t | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, z_{t-d} \in R_j)$$

donde  $p_{t,j} = Pr(z_{t-d} \in R_j)$ ,  $\sum_{j=1}^l p_{t,j} = 1$  y  $F_{t-1,j}(\cdot | \cdot)$  ( $j = 1, \dots, l$ ) es dada en (i). Ahora, ya que  $\{Z_t\}$  converge débilmente a  $F$ ,  $p_{t,j} \rightarrow p_j =$

$F(r_j) - F(r_{j-1})$  cuando  $t \rightarrow \infty$  para todo  $j = 1, \dots, l$ . Por supuesto, si  $\{Z_t\}$  es un proceso estrictamente estacionario, entonces la función de distribución acumulada de  $Z_t$  es  $F$  para todo  $t$  y, así,  $p_{t,j} = p_j$  para todo  $t$  y para todo  $j = 1, \dots, l$ .

Las consideraciones teóricas (i) y (ii) anteriores, se tendrán en cuenta para el cálculo de los residuales generalizados del modelo TSARX.

El modelo (1.1) puede considerarse como un caso especial de un modelo TARX, como se describe a continuación. Sea un TARX( $l; k_1 + sK_1, \dots, k_l + sK_l; q_1, \dots, q_l$ ) de la forma

$$X_t = a_0^{(j)} + \sum_{i=1}^{k_j + sK_j} a_i^{(j)} X_{t-i} + \sum_{v=1}^{q_j} c_v^{(j)} Z_{t-v} + h^{(j)} \epsilon_t, \quad (1.2)$$

si  $r_{j-1} < Z_{t-d} \leq r_j$  para todo  $t \in \mathbb{Z}$ , algún  $j = 1, \dots, l$ ,  $k_j < s$  para todo  $j$ , y  $\{\epsilon_t\} \sim \text{RBG}(0,1)$ .

Se define en forma conveniente el vector de parámetros completo del modelo TARX,  $\boldsymbol{\theta}_x = (\boldsymbol{\theta}'_{ne}, \boldsymbol{\theta}'_e)$  siendo  $\boldsymbol{\theta}_{ne} = (A'_{11}, \dots, A'_{l1}, A'_{12}, \dots, A'_{l2}, C'_1, \dots, C'_l, \mathbf{h}')'$  (vector de parámetros no estructurales) donde  $A_{j1} = (a_0^{(j)}, a_1^{(j)}, \dots, a_{k_j}^{(j)}, a_s^{(j)}, \dots, a_{sK_j}^{(j)})$ ,  $A_{j2} = (a_{1+k_j}^{(j)}, \dots, a_{s-1}^{(j)}, a_{1+s}^{(j)}, \dots, a_{k_j+s}^{(j)}, \dots, a_{(1+k_j)+s(K_j-1)}^{(j)}, \dots, a_{(s-1)+s(K_j-1)}^{(j)}, a_{1+sK_j}^{(j)}, \dots, a_{k_j+sK_j}^{(j)})$ ,  $C_j = (c_1^{(j)}, \dots, c_{q_j}^{(j)})$ ,  $j = 1, \dots, l$ ,  $\mathbf{h} = ((h^{(1)})^2, \dots, (h^{(l)})^2)$  y  $\boldsymbol{\theta}_e = (k_1 + sK_1, \dots, k_l + sK_l, q_1, \dots, q_l, \mathbf{r}', d, l)'$  (vector de parámetros estructurales) donde  $\mathbf{r} = (r_1, \dots, r_{l-1})'$ .

Si  $A_{j2} = (0, \dots, 0, -a_1^{(j)} a_s^{(j)}, \dots, -a_{k_j}^{(j)} a_s^{(j)}, \dots, 0, \dots, 0, -a_1^{(j)} a_{sK_j}^{(j)}, \dots, -a_{k_j}^{(j)} a_{sK_j}^{(j)})$ , para cada  $j = 1, \dots, l$ , el modelo TARX dado en (1.2) se convierte en el modelo TSARX dado en (1.1). Relacionado a las restricciones dadas sobre los coeficientes de  $A_{j2}$  de un modelo TARX, en una sección posterior se presentará una prueba estadística para determinar estacionalidad multiplicativa por regímenes.

## 1.2. La función de verosimilitud condicional.

Sean  $\mathbf{x}_{1,T} = (x_1, \dots, x_T)$  y  $\mathbf{z}_{1,T} = (z_1, \dots, z_T)$  vectores de datos observados para los procesos  $\{X_t\}$  y  $\{Z_t\}$  respectivamente, donde  $T$  es la longitud del periodo muestral. Se asume que el mecanismo probabilístico que genera  $\mathbf{z}_{1,T}$  no depende de  $\boldsymbol{\theta}_x$  y la distribución conjunta de  $\mathbf{x}_{1,T}$  condicional sobre  $\mathbf{z}_{1,T}$  y  $\boldsymbol{\theta}$ , no depende de  $\boldsymbol{\theta}_z$ . Con estas condiciones se puede ver, que es posible estimar inicialmente los parámetros de  $\{Z_t\}$ , y luego condicional sobre estos parámetros, se procede a estimar los parámetros del modelo TSARX considerando la función de verosimilitud

$$f(\mathbf{x}_{1,T}, \mathbf{z}_{1,T} | \boldsymbol{\theta}) = f(\mathbf{z}_{1,T} | \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) f(\mathbf{x}_{1,T} | \mathbf{z}_{1,T}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = f(\mathbf{z}_{1,T} | \boldsymbol{\theta}_z) f(\mathbf{x}_{1,T} | \mathbf{z}_{1,T}, \boldsymbol{\theta}_x).$$

Investigaciones futuras, pueden direccionarse a obtener estimaciones conjuntas de los parámetros  $\boldsymbol{\theta}_x$  y  $\boldsymbol{\theta}_z$ . Nuestro interés, es calcular la función de verosimilitud condicional de un modelo TSARX( $l; k_1, \dots, k_l; K_1, \dots, K_l; q_1, \dots, q_l$ )<sub>s</sub> dado en (1.1); es decir,  $f(\mathbf{x}_{1,T} | \mathbf{z}_{1,T}, \boldsymbol{\theta}_x)$ . Adicionalmente se asume que los primeros  $K$  valores de  $\mathbf{x}_{1,T}$ ; es decir,  $\mathbf{x}_{1,K} = (x_1, \dots, x_K)$  son fijos y conocidos con  $K = \text{máximo} \{k_j + sK_j, q_j, j = 1, \dots, l\}$ ; por consiguiente,

$$f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) = f(x_{K+1} | x_K, \dots, x_1, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) \cdots f(x_T | x_{T-1}, \dots, x_1, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x).$$

Ahora, como  $\{\epsilon_t\} \sim RBG(0, 1)$  se tiene que para  $t = (K + 1), \dots, T$ ,  $f(x_t | x_{t-1}, \dots, x_1, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x)$  se distribuye como una normal con media

$$\mu_t = a_0^{(j_t)} + \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i} + \sum_{u=1}^{K_{j_t}} b_u^{(j_t)} x_{t-su} - \sum_{i=1}^{k_{j_t}} \sum_{u=1}^{K_{j_t}} a_i^{(j_t)} b_u^{(j_t)} x_{t-i-su} + \sum_{v=1}^{q_{j_t}} c_v^{(j_t)} z_{t-v}$$

y varianza  $(h^{(j_t)})^2$ . La sucesión  $\{j_t\}$  es la serie de tiempo observada para el proceso estocástico  $\{J_t\}$ , siendo una sucesión de variables indicadoras definidas como  $J_t = j$  si  $r_{j-1} < z_{t-d} \leq r_j$ , para algún  $j = 1, \dots, l$ ; así,

$$\begin{aligned} f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) &= (2\pi)^{-\frac{(T-K)}{2}} \prod_{t=K+1}^T (h^{(j_t)})^{-1} \exp \left\{ \frac{1}{2} \sum_{t=K+1}^T \left( \frac{x_t - \mu_t}{h^{(j_t)}} \right)^2 \right\} \\ &= (2\pi)^{-\frac{(T-K)}{2}} \prod_{j=1}^l (h^{(j)})^{-n_j} \exp \left\{ -\frac{1}{2} \sum_{j=1}^l (\mathbb{X}_j - A_j^* B_j^*)' ((h^{(j)})^{-2} I_{n_j}) (\mathbb{X}_j - A_j^* B_j^*) \right\}, \end{aligned} \quad (1.3)$$

donde

$$\begin{aligned} a_t^{*'} &= (1, x_{t-1}, \dots, x_{t-k_j}, 0, \dots, 0, x_{t-s}, x_{t-(1+s)}, \dots, x_{t-(k_j+s)}, \dots, \\ &\quad 0, \dots, 0, x_{t-sK_j}, x_{t-(1+sK_j)}, \dots, x_{t-(k_j+sK_j)}, z_{t-1}, \dots, z_{t-q_j}), \\ A_j^* &= (a_{t_1, j}^*, \dots, a_{t_{n_j}, j}^*), \\ B_j^* &= (a_0^{(j)}, a_1^{(j)}, \dots, a_{k_j}^{(j)}, 0, \dots, 0, b_1^{(j)}, -a_1^{(j)} b_1^{(j)}, \dots, -a_{k_j}^{(j)} b_1^{(j)}, \dots, \\ &\quad 0, \dots, 0, b_{K_j}^{(j)}, -a_1^{(j)} b_{K_j}^{(j)}, \dots, -a_{k_j}^{(j)} b_{K_j}^{(j)}, c_1^{(j)}, \dots, c_{q_j}^{(j)})', \end{aligned}$$

$\mathbb{X}_j = (x_{t_1, j}, \dots, x_{t_{n_j}, j})'$ ,  $\{t_{1, j}, \dots, t_{n_j, j}\}$  son los índices en el tiempo, donde  $r_{j-1} < z_{t-d} \leq r_j$  y  $n_j$  es el número de observaciones en el régimen  $j$ ,  $j = 1, \dots, l$ .

Para el caso de un modelo TARX( $l; k_1 + sK_1, \dots, k_l + sK_l; q_1, \dots, q_l$ ) dado en (1.2), se tiene que

$$f(x_t | x_{t-1}, \dots, x_1, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) \sim N \left( a_0^{(j_t)} + \sum_{i=1}^{k_{j_t} + sK_{j_t}} a_i^{(j_t)} x_{t-i} + \sum_{v=1}^{q_{j_t}} c_v^{(j_t)} z_{t-v}, (h^{(j_t)})^2 \right),$$

con la sucesión  $\{j_t\}$  definida previamente; luego, la función de verosimilitud condicional de un modelo TARX( $l; k_1 + sK_1, \dots, k_l + sK_l, q_1, \dots, q_l$ ) es dada por

$$f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) = (2\pi)^{-\frac{(T-K)}{2}} \prod_{j=1}^l (h^j)^{-n_j} \exp \left\{ -\frac{1}{2} \sum_{j=1}^l (\mathbb{X}_j - A_j^* B_j^*)' ((h^{(j)})^{-2} I_{n_j}) (\mathbb{X}_j - A_j^* B_j^*) \right\}, \quad (1.4)$$

donde

$$\begin{aligned} \mathbf{a}_t^{*'} &= (1, x_{t-1}, \dots, x_{t-k_j}, x_{t-(1+k_j)}, \dots, x_{t-(s-1)}, x_{t-s}, x_{t-(1+s)}, \dots, x_{t-(k_j+s)}, \dots, \\ &\quad x_{t-[(1+k_j)+s(K_j-1)]}, \dots, x_{t-[(s-1)+s(K_j-1)]}, x_{t-sK_j}, x_{t-(1+sK_j)}, \dots, x_{t-(k_j+sK_j)}, \\ &\quad z_{t-1}, \dots, z_{t-q_j}), \\ A_j^* &= (a_{t_1,j}^*, \dots, a_{t_{n_j},j}^*), \\ B_j^* &= (a_0^{(j)}, a_1^{(j)}, \dots, a_{k_j}^{(j)}, a_{1+k_j}^{(j)}, \dots, a_{s-1}^{(j)}, a_s^{(j)}, a_{1+s}^{(j)}, \dots, a_{k_j+s}^{(j)}, \dots, a_{[(1+k_j)+s(K_j-1)]}^{(j)}, \dots, \\ &\quad a_{[(s-1)+s(K_j-1)]}^{(j)}, a_{sK_j}^{(j)}, a_{1+sK_j}^{(j)}, \dots, a_{k_j+sK_j}^{(j)}, c_1^{(j)}, \dots, c_{q_j}^{(j)})', \\ \mathbb{X}_j &= (x_{t_1,j}, \dots, x_{t_{n_j},j})' \text{ con } \{t_{1,j}, \dots, t_{n_j,j}\} \text{ y } n_j, j = 1, \dots, l, \text{ definidos anteriormente.} \end{aligned}$$

### 1.3. Estimación de parámetros de un modelo TSARX.

Se usan ideas de búsqueda estocástica, en particular el método Gibbs de selección de variables dado en Dellaportas, Forster y Ntzoufras (2002) y denotado por GVS para identificar los órdenes autorregresivos en un modelo TSARX con  $l$  regímenes. Con este método se logra la estimación de los órdenes autorregresivos en una sola etapa y se pueden involucrar estas estimaciones en el esquema muestral MCMC junto con el resto de estimaciones de los parámetros, asumiendo conocido el número de regímenes.

El método GVS es un híbrido entre el método para la selección de variables, propuesto por Kuo y Mallick (1998) y se denota por KM, y el de búsqueda estocástica para la selección de variables (SSVS) dado por George y McCulloch (1993). Estas y otras técnicas de selección Bayesiana de modelos pueden revisarse en O'hara y Sillanpää (2009).

So y Chen(2003) utilizan el método SSVS para determinar los órdenes autorregresivos y el resto de parámetros, conociendo el número de regímenes en los modelos SETAR. Mientras que, Calderon y Nieto (2017) emplean los métodos KM y GVS en modelos TARX multivariados.

El método GVS para un modelo TSARX con  $l$  regímenes, es descrito como sigue: Se definen las variables indicadoras latentes 0-1,  $s_{ij}$ ,  $m_{uj}$ ,  $n_{vj}$ ,  $i = 0, 1, \dots, k_j^*$ ,  $u = 1, \dots, K_j^*$ ,  $v = 1, \dots, q_j^*$ , con  $k_j^*$ ,  $K_j^*$  y  $q_j^*$ ,  $j = 1, \dots, l$  órdenes autorregresivos máximos considerados. Se tiene que si  $s_{ij} = 1$ , el parámetro asociado  $a_i^{(j)}$  se incluye en el modelo, y si  $s_{ij} = 0$ , el parámetro asociado  $a_i^{(j)}$  no se incluye. Si  $m_{uj} = 1$  el parámetro asociado  $b_u^{(j)}$  debe

incluirse en el modelo, y si  $m_{uj} = 0$  el parámetro asociado  $b_u^{(j)}$  no debe incluirse. Mientras, si  $n_{vj} = 1$  el parámetro asociado  $c_v^{(j)}$  se incluye en el modelo, y si  $n_{vj} = 0$  no se incluye dicho parámetro.

Sean  $S_j = (s_{0j}, s_{1j}, \dots, s_{k_j^*j})'$ ,  $M_j = (m_{1j}, \dots, m_{K_j^*j})'$  y  $N_j = (n_{1j}, \dots, n_{q_j^*j})'$ ,  $j = 1, \dots, l$ , vectores de indicadores latentes asociados a los coeficientes autorregresivos no estacionales, estacionales y de la entrada exógena, respectivamente.

Se reescribe el modelo (1.1) con los vectores  $S_j$ ,  $M_j$  y  $N_j$ ,  $j = 1, \dots, l$  como

$$X_t = a_0^{(j)} s_{0j} + \sum_{i=1}^{k_j^*} a_i^{(j)} s_{ij} X_{t-i} + \sum_{u=1}^{K_j^*} b_u^{(j)} m_{uj} X_{t-su} - \sum_{i=1}^{k_j^*} \sum_{u=1}^{K_j^*} a_i^{(j)} b_u^{(j)} s_{ij} m_{uj} X_{t-i-su} \\ + \sum_{v=1}^{q_j^*} c_v^{(j)} n_{vj} Z_{t-v} + h^{(j)} \epsilon_t, \quad \text{si } r_{j-1} < Z_{t-d} \leq r_j,$$

para algún  $j = 1, \dots, l$  y que se cumpla el resto de especificaciones dadas en la sección 1.1., excepto que el vector de parámetros estructurales del modelo, ahora es

$$\theta_e = (S_1, \dots, S_l, M_1, \dots, M_l, N_1, \dots, N_l, r', d, l)'$$

Para usar el método GVS, es necesario asumir que:

$$a_i^{(j)} | s_{ij} \sim (1 - s_{ij}) N(0, 1/\varphi_{ij}^2) + s_{ij} N(0, 1/\eta_{ij}^2 \varphi_{ij}^2), \\ b_u^{(j)} | m_{uj} \sim (1 - m_{uj}) N(0, 1/\chi_{uj}^2) + m_{uj} N(0, 1/\xi_{uj}^2 \chi_{uj}^2), \quad y \\ c_v^{(j)} | n_{vj} \sim (1 - n_{vj}) N(0, 1/\psi_{vj}^2) + n_{vj} N(0, 1/\lambda_{vj}^2 \psi_{vj}^2),$$

donde  $Pr(s_{ij} = 1) = \pi_{ij}$ ,  $Pr(m_{uj} = 1) = \lambda_{uj}$  y  $Pr(n_{vj} = 1) = \omega_{vj}$ ; además, se eligen valores específicos  $\eta_{ij} > 1$ ,  $\xi_{uj} > 1$ ,  $\lambda_{vj} > 1$ ,  $\varphi_{ij} > 0$ ,  $\chi_{uj} > 0$  y  $\psi_{vj} > 0$ ,  $i = 0, 1, \dots, k_j^*$ ,  $u = 1, \dots, K_j^*$ ,  $v = 1, \dots, q_j^*$ ,  $j = 1, \dots, l$ .

Las anteriores mezclas de distribuciones normales se pueden llevar a una estructura multivariada como sigue:

$$A_j | S_j \sim N_{k_j^*+1}(\mathbf{0}, (D_{S_j} V_{A_{j0}} D_{S_j})^{-1})', \\ B_j | M_j \sim N_{K_j^*}(\mathbf{0}, (D_{M_j} V_{B_{j0}} D_{M_j})^{-1})', \quad y \\ C_j | N_j \sim N_{q_j^*}(\mathbf{0}, (D_{N_j} V_{C_{j0}} D_{N_j})^{-1})',$$

donde  $V_{A_{j0}}$ ,  $V_{B_{j0}}$  y  $V_{C_{j0}}$  son matrices de covarianzas a priori,  $D_{S_j}$  es una matriz diagonal  $diag\{1/a_{0j}\varphi_{0j}, \dots, 1/a_{k_j^*j}\varphi_{k_j^*j}\}$  con  $a_{ij} = 1$  si  $s_{ij} = 0$  y  $a_{ij} = \eta_{ij}$  si  $s_{ij} = 1$ ,  $D_{M_j} = diag\{1/b_{1j}\chi_{1j}, \dots, 1/b_{K_j^*j}\chi_{K_j^*j}\}$  con  $b_{uj} = 1$  si  $m_{uj} = 0$  y  $b_{uj} = \xi_{uj}$  si  $m_{uj} = 1$ , y  $D_{N_j} = diag\{1/c_{1j}\psi_{1j}, \dots, 1/c_{q_j^*j}\psi_{q_j^*j}\}$  con  $c_{vj} = 1$  si  $n_{vj} = 0$  y  $c_{vj} = \lambda_{vj}$  si  $n_{vj} = 1$ ; lo anterior asumiendo independencia a priori de las  $s_{ij}$ ,  $m_{uj}$ ,  $n_{vj}$ ,  $i = 0, 1, \dots, k_j^*$ ,  $u = 1, \dots, K_j^*$ ,  $v = 1, \dots, q_j^*$ ,  $j = 1, \dots, l$ .

La función de verosimilitud condicional de un modelo TSARX con  $l$  regímenes considerando los vectores  $S_j$ ,  $M_j$  y  $N_j$ ,  $j = 1, \dots, l$  es la dada en (1.3), excepto que ahora,

$$B_j^* = (a_0^{(j)} s_{0j}, a_1^{(j)} s_{1j}, \dots, a_{k_j^*}^{(j)} s_{k_j^* j}, 0, \dots, 0, b_1^{(j)} m_{1j}, -a_1^{(j)} b_1^{(j)} s_{1j} m_{1j}, \dots, -a_{k_j^*}^{(j)} b_1^{(j)} s_{k_j^* j} m_{1j}, \dots, 0, \dots, 0, b_{K_j^*}^{(j)} m_{K_j^* j}, -a_1^{(j)} b_{K_j^*}^{(j)} s_{1j} m_{K_j^* j}, \dots, -a_{k_j^*}^{(j)} b_{K_j^*}^{(j)} s_{k_j^* j} m_{K_j^* j}, c_1^{(j)} n_{1j}, \dots, c_{q_j^*}^{(j)} n_{q_j^* j})', \text{ con } k_j^*, K_j^*, \text{ y } q_j^*, j = 1, \dots, l \text{ definidos anteriormente.}$$

Una vez se obtienen muestras MCMC a posteriori para  $S_j$ ,  $M_j$ ,  $N_j$ ,  $j = 1, \dots, l$ , se pueden identificar los órdenes autorregresivos del modelo considerando una combinación particular de  $S_j$ ,  $M_j$  y  $N_j$ ,  $j = 1, \dots, l$  la cual posea la probabilidad a posteriori más alta. Es decir, aquel evento con mayor frecuencia en las muestras MCMC.

Se asume independencia a priori para  $A_j$ ,  $B_j$  y  $C_j$  así que  $V_{A_{j0}} = \mathbb{I}_{k_j^*+1}$ ,  $V_{B_{j0}} = \mathbb{I}_{K_j^*}$  y  $V_{C_{j0}} = \mathbb{I}_{q_j^*}$ ,  $j = 1, \dots, l$  ( $\mathbb{I}$  es la matriz identidad). También, se asume que cada coeficiente autorregresivo tiene igual probabilidad de incluirse o no en el modelo; luego,  $\pi_{ij} = 1/2$ ,  $\lambda_{uj} = 1/2$  y  $\omega_{vj} = 1/2$ ,  $i = 0, 1, \dots, k_j^*$ ,  $u = 1, \dots, K_j^*$ ,  $v = 1, \dots, q_j^*$ ,  $j = 1, \dots, l$ .

Las distribuciones a priori para los grupos de parámetros  $A_j$ ,  $B_j$ ,  $C_j$ ,  $(h^{(j)})^2$ ,  $j = 1, \dots, l$ ,  $\mathbf{r}$  y  $d$  son:

- (i) Para  $A_j|S_j$  una distribución normal multivariada con media  $\mathbf{0}_{k_j^*+1}$  y covarianza  $(D_{S_j} D_{S_j})^{-1}$ ,  $j = 1, \dots, l$ .
- (ii) Para  $B_j|M_j$  una distribución normal multivariada con media  $\mathbf{0}_{K_j^*}$  y covarianza  $(D_{M_j} D_{M_j})^{-1}$ ,  $j = 1, \dots, l$ .
- (iii) Para  $C_j|N_j$  una distribución normal multivariada con media  $\mathbf{0}_{q_j^*}$  y covarianza  $(D_{N_j} D_{N_j})^{-1}$ ,  $j = 1, \dots, l$ .
- (iv) Para  $(h^{(j)})^2$  una distribución gamma invertida (IG) con parámetro de posición  $\nu_j/2$  y parámetro de escala  $\nu_j \lambda_j/2$ ,  $j = 1, \dots, l$ .
- (v)  $\mathbf{r} = (r_1, \dots, r_{l-1})'$  sigue la distribución a priori constante  $f(\mathbf{r}) = \frac{1}{\mathcal{C}} I(\mathbf{A})$  donde  $\mathcal{C} = \int \dots \int_{\mathbf{A}} dr_1 \dots dr_{l-1}$ , y  $\mathbf{A}$  es una región que satisface:

- (1)  $a \leq r_1 < \dots < r_{l-1} \leq b$ , con  $a$  y  $b$  cuantiles convenientes de  $Z_{t-d}$ , y
- (2) cada régimen contiene al menos un  $H\%$  de la muestra de  $Z_{t-d}$  (esto para asegurar inferencia válida). El valor de  $H$  es preestablecido.

$I(\mathbf{A})$  es la función indicadora de  $\mathbf{A}$  sobre  $\mathbb{R}^{l-1}$ . Para el caso particular de dos regímenes,  $f(\mathbf{r})$  es una uniforme continua definida en  $(a, b)$ .

- (vi)  $d$  sigue una distribución uniforme discreta, con función de masa de probabilidad  $Pr(d = i) = 1/(d_0 + 1)$ , para  $i = 0, 1, \dots, d_0$ , siendo  $d_0$  un rezago máximo dado.

Se asumen los hiperparámetros  $\{\varphi_{ij}, \eta_{ij}, \chi_{uj}, \xi_{uj}, \psi_{vj}, \lambda_{vj}, \nu_j, \lambda_j, i = 0, 1, \dots, k_j^*, u = 1, \dots, K_j^*, v = 1, \dots, q_j^*, j = 1, \dots, l, a, b, d_0\}$  conocidos. Combinando la función de verosimilitud condicional para un modelo TSARX, considerando los vectores  $S_j$ ,  $M_j$  y  $N_j$ ,

$j = 1, \dots, l$ , con cada una de las distribuciones a priori dadas en (i) a (vi), se obtienen las distribuciones a posteriori de los parámetros desconocidos del modelo. Muestras MCMC de  $A_j$ ,  $B_j$ ,  $C_j$ ,  $(h^{(j)})^2$ ,  $S_j$ ,  $M_j$ ,  $N_j$ ,  $j = 1, \dots, l$ ,  $\mathbf{r}$  y  $d$  son generados de distribuciones condicionales completas convenientes como sigue:

Algoritmo 1.

1. Inicializar:  $A_j^{(0)}$ ,  $B_j^{(0)}$ ,  $C_j^{(0)}$ ,  $[(h^{(j)})^2]^{(0)}$ ,  $S_j^{(0)}$ ,  $M_j^{(0)}$ ,  $N_j^{(0)}$ ,  $j = 1, \dots, l$ ,  $\mathbf{r}^{(0)}$ ,  $d^{(0)}$ .
2. Generar  $A_j$  (independiente de  $A_i$ ,  $i \neq j$ ) de

$$f(A_j | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-A_j}) \sim N_{k_j^*+1}(\boldsymbol{\theta}_{A_j}, V_{A_j}^{-1})$$

donde

$$V_{A_j} = [1/(h^{(j)})^2] S_j^{*'} S_j^* + D_{S_j} D_{S_j}$$

y

$$\boldsymbol{\theta}_{A_j} = V_{A_j}^{-1} [1/(h^{(j)})^2] S_j^{*'} (\mathbb{X}_j - M_j^* B_j - N_j^* C_j),$$

$$s_t^{*'} = (s_{0j}, s_{1j} x_{t-1} - \sum_{u=1}^{K_j^*} b_u^{(j)} s_{1j} m_{uj} x_{t-1-su}, \dots, s_{k_j^* j} x_{t-k_j^*} - \sum_{u=1}^{K_j^*} b_u^{(j)} s_{k_j^* j} m_{uj} x_{t-k_j^*-su}),$$

$$S_j^* = (s_{t_{1,j}}^*, \dots, s_{t_{n_j,j}}^*)', \quad m_t^{*'} = (m_{1j} x_{t-s}, m_{2j} x_{t-s_2}, \dots, m_{K_j^* j} x_{t-s_{K_j^*}}),$$

$$M_j^* = (m_{t_{1,j}}^*, \dots, m_{t_{n_j,j}}^*)', \quad n_t^{*'} = (n_{1j} z_{t-1}, \dots, n_{q_j^* j} z_{t-q_j^*}), \quad N_j^* = (n_{t_{1,j}}^*, \dots, n_{t_{n_j,j}}^*)',$$

$\mathbb{X}_j = (x_{t_{1,j}}, \dots, x_{t_{n_j,j}})'$  y  $n_j$  el número de observaciones en el régimen  $j$ ,  $j = 1, \dots, l$ .

Aquí  $\boldsymbol{\theta}_{x-A_j}$  es el vector de parámetros  $\boldsymbol{\theta}_x$  sin el subvector  $A_j$ .

3. Generar  $B_j$  (independiente de  $B_i$ ,  $i \neq j$ ) de

$$f(B_j | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-B_j}) \sim N_{K_j^*}(\boldsymbol{\theta}_{B_j}, V_{B_j}^{-1})$$

donde

$$V_{B_j} = [1/(h^{(j)})^2] T_j^{*'} T_j^* + D_{M_j} D_{M_j}$$

y

$$\boldsymbol{\theta}_{B_j} = V_{B_j}^{-1} [1/(h^{(j)})^2] T_j^{*'} (\mathbb{X}_j - U_j^* A_j - N_j^* C_j),$$

$$t_t^{*'} = (m_{1j} x_{t-s} - \sum_{i=1}^{k_j^*} a_i^{(j)} s_{ij} m_{1j} x_{t-i-s}, \dots, m_{K_j^* j} x_{t-s_{K_j^*}} - \sum_{i=1}^{k_j^*} a_i^{(j)} s_{ij} m_{K_j^* j} x_{t-i-s_{K_j^*}}),$$

$$T_j^* = (t_{t_{1,j}}^*, \dots, t_{t_{n_j,j}}^*)', \quad \mu_t^{*'} = (s_{0j}, s_{1j} x_{t-1}, \dots, s_{k_j^* j} x_{t-k_j^*}), \quad U_j^* = (\mu_{t_{1,j}}^*, \dots, \mu_{t_{n_j,j}}^*)',$$

$N_j^*$ ,  $\mathbb{X}_j$  y  $n_j$ ,  $j = 1, \dots, l$  son definidos anteriormente.

4. Generar  $C_j$  (independiente de  $C_i$ ,  $i \neq j$ ) de

$$f(C_j | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-C_j}) \sim N_{q_j^*}(\boldsymbol{\theta}_{C_j}, V_{C_j}^{-1})$$

donde

$$V_{C_j} = [1/(h^{(j)})^2]N_j^{*'}N_j^* + D_{N_j}D_{N_j}$$

y

$$\begin{aligned}\boldsymbol{\theta}_{C_j} &= V_{C_j}^{-1}[1/(h^{(j)})^2]N_j^{*'}(\mathbb{X}_j - F_j^*E_j - U_j^*A_j - M_j^*B_j), \\ f_t^{*'} &= (0, \dots, 0, s_{1j}m_{1j}x_{t-1-s}, \dots, s_{k_j^*j}m_{1j}x_{t-k_j^*-s}, 0, \dots, 0, \\ &\quad s_{1j}m_{K_j^*j}x_{t-1-sK_j^*}, \dots, s_{k_j^*j}m_{K_j^*j}x_{t-k_j^*-sK_j^*}),\end{aligned}$$

$$\begin{aligned}F_j^* &= (f_{t_{1,j}}^*, \dots, f_{t_{n_j,j}}^*)', \\ E_j &= (a_{1+k_j^*}^{(j)}, \dots, a_{s-1}^{(j)}, -a_1^{(j)}b_1^{(j)}, \dots, -a_{k_j^*}^{(j)}b_1^{(j)}, \dots, a_{(1+k_j^*)+s(K_j^*-1)}^{(j)}, \\ &\quad \dots, a_{(s-1)+s(K_j^*-1)}^{(j)}, -a_1^{(j)}b_{K_j^*}^{(j)}, \dots, -a_{k_j^*}^{(j)}b_{K_j^*}^{(j)}),\end{aligned}$$

$N_j^*$ ,  $U_j^*$ ,  $M_j^*$ ,  $\mathbb{X}_j$  y  $n_j$ ,  $j = 1, \dots, l$  son definidos anteriormente.

5. Generar  $(h^{(j)})^2$  (independiente de  $(h^{(i)})^2$ ,  $i \neq j$ ) de

$$f((h^{(j)})^2 | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-(h^{(j)})^2}) \sim IG\left(\frac{\nu_j + n_j}{2}, \frac{\nu_j \lambda_j + n_j s_j^2}{2}\right),$$

donde  $n_j s_j^2 = (\mathbb{X}_j - A_j^* B_j^*)'(\mathbb{X}_j - A_j^* B_j^*)$ ,  $j = 1, \dots, l$ .

6. Generar  $\mathbf{r}$  de

$$f(\mathbf{r} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-\mathbf{r}}) \propto f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x)I(\mathbf{A})$$

donde  $f(\cdot | \cdot)$  es la función de verosimilitud condicional del modelo TSARX teniendo en cuenta  $S_j$ ,  $M_j$  y  $N_j$ ,  $j = 1, \dots, l$  variables indicadoras latentes y  $\mathbf{A}$  es la región especificada en la a priori para  $\mathbf{r}$ .

7. Generar  $d$  de la distribución multinomial, con función de masa de probabilidad

$$Pr(d = i | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-d}) = \frac{f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-d}, d = i)Pr(d = i)}{\sum_{m=0}^{d_0} f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-d}, d = m)Pr(d = m)},$$

$i = 0, 1, \dots, d_0$ .

8. Generar el vector  $S_j = (s_{0j}, s_{1j}, \dots, s_{k_j^*j})'$  (independiente de  $S_i$ ,  $M_i$  y  $N_i$ ,  $i \neq j$ ), obteniendo muestras de  $s_{ij}$ ,  $i = 0, 1, \dots, k_j^*$ ,  $j = 1, \dots, l$  individualmente de la distribución Bernoulli, con función de masa de probabilidad

$$Pr(s_{ij} = 1 | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-s_{ij}}) = \frac{\alpha_{ij}}{\alpha_{ij} + \beta_{ij}},$$

donde

$$\begin{aligned}\alpha_{ij} &= f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-s_{ij}}, s_{ij} = 1)f(A_j | S_{j-s_{ij}}, s_{ij} = 1)\pi_{ij} \text{ y} \\ \beta_{ij} &= f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-s_{ij}}, s_{ij} = 0)f(A_j | S_{j-s_{ij}}, s_{ij} = 0)(1 - \pi_{ij})\end{aligned}$$

9. Generar el vector  $M_j = (m_{1j}, m_{2j}, \dots, m_{K_j^*j})'$  (independiente de  $M_i, S_i$  y  $N_i, i \neq j$ ), obteniendo muestras de  $m_{uj}, u = 1, \dots, K_j^*, j = 1, \dots, l$  individualmente de la distribución Bernoulli, con función de masa de probabilidad

$$Pr(m_{uj} = 1 | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-m_{uj}}) = \frac{\gamma_{uj}}{\gamma_{uj} + \varrho_{uj}},$$

donde

$$\begin{aligned} \gamma_{uj} &= f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-m_{uj}}, m_{uj} = 1) f(B_j | M_{j-m_{uj}}, m_{uj} = 1) \lambda_{uj} \text{ y} \\ \varrho_{uj} &= f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-m_{uj}}, m_{uj} = 0) f(B_j | M_{j-m_{uj}}, m_{uj} = 0) (1 - \lambda_{uj}). \end{aligned}$$

10. Generar el vector  $N_j = (n_{1j}, n_{2j}, \dots, n_{q_j^*j})'$  (independiente de  $N_i, S_i$  y  $M_i, i \neq j$ ), obteniendo muestras de  $n_{vj}, v = 1, 2, \dots, q_j^*, j = 1, \dots, l$  individualmente de la distribución Bernoulli, con función de masa de probabilidad

$$Pr(n_{vj} = 1 | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-n_{vj}}) = \frac{\varsigma_{vj}}{\varsigma_{vj} + \xi_{vj}},$$

donde

$$\begin{aligned} \varsigma_{vj} &= f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-n_{vj}}, n_{vj} = 1) f(C_j | N_{j-n_{vj}}, n_{vj} = 1) \omega_{vj} \text{ y} \\ \xi_{vj} &= f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-n_{vj}}, n_{vj} = 0) f(C_j | N_{j-n_{vj}}, n_{vj} = 0) (1 - \omega_{vj}). \end{aligned}$$

Todas las distribuciones a posteriori del Algoritmo 1, excepto la de  $\mathbf{r}$  son estándar. Para simular valores de  $\mathbf{r}$ , se emplea el algoritmo de Metropolis-Hastings MH (Metropolis *et al*, 1953), en particular un MH de caminata aleatoria con distribución instrumental Gaussiana. Denotando  $f(\mathbf{r} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-\mathbf{r}})$  por  $f(\mathbf{r})$ , el algoritmo para generar muestras de  $\mathbf{r}$  es descrito como sigue.

Algoritmo 2.

1. En la iteración  $i$ , un candidato  $\mathbf{r}^*$  se extrae del kernel de la caminata aleatoria

$$\mathbf{r}^* = \mathbf{r}^{(i-1)} + \boldsymbol{\xi}, \{\boldsymbol{\xi}\} \sim N_{l-1}(\mathbf{0}, \Sigma).$$

donde  $\Sigma = \text{diag}(\sigma_{r_1}^2, \dots, \sigma_{r_{l-1}}^2)$  y  $\mathbf{r}^{(i-1)}$  es la  $(i-1)$ -ésima iteración de  $\mathbf{r}$ .

2. Genere  $u$  de una uniforme  $(0, 1)$ , si  $u \leq \text{mínimo}\{1, f(\mathbf{r})/f(\mathbf{r}^{(i-1)})\}$ ,  $\mathbf{r}^*$  es aceptado como el nuevo estado  $\mathbf{r}^{(i)}$ ; en otro caso,  $\mathbf{r}^{(i)} = \mathbf{r}^{(i-1)}$ .

Se fijan y monitorean los tamaños de paso  $\sigma_{r_1}, \dots, \sigma_{r_{l-1}}$  para producir una rápida convergencia de las muestras MCMC. Valores adecuados de  $\sigma_{r_1}, \dots, \sigma_{r_{l-1}}$  con buenas propiedades de convergencia se pueden lograr al tener una tasa de aceptación entre el 25% al 50% (Gelman, Roberts y Gilks, 1996). Se obtienen los resultados a posteriori del muestreador de Gibbs incorporando el algoritmo MH, usando los Algoritmos 1 y 2. Detalles del muestreador de Gibbs y del algoritmo MH se pueden hallar en Casella y George (1992) y Chib y Greenberg (1995), respectivamente. Se toma un tamaño de muestra MCMC suficientemente grande  $N$ , se descartan las iteraciones del llamado periodo de calentamiento  $M$ , luego se queda con  $G = N - M$  iteraciones para realizar

inferencia a posteriori.

Se calculan las medias y desviaciones estándar de las distribuciones a posteriori de  $A_j$ ,  $B_j$ ,  $C_j$ ,  $(h^{(j)})^2$ ,  $j = 1, \dots, l$  y  $\mathbf{r}$  para obtener estimaciones puntuales y una medida de incertidumbre de dichos estimaciones, respectivamente. Además, con las muestras generadas se obtienen intervalos creíbles  $100(1 - \alpha)\%$ , para algún  $\alpha$ ,  $0 < \alpha < 1$ , donde tales intervalos son dados por  $(q_{\alpha/2}, q_{1-\alpha/2})$ , con  $q_\alpha$  el  $\alpha$ -ésimo cuantil muestral, o regiones de densidades altas  $100(1 - \alpha)\%$ . Para el caso, de la estimación del parámetro de rezago  $d$  se toma es la moda a posteriori y para las estimaciones de los vectores de ceros y unos,  $S_j$ ,  $M_j$ ,  $N_j$ ,  $j = 1, \dots, l$  se toman las probabilidades de ocurrencia más altas.

Con el fin de examinar cada uno de los parámetros, si las muestras MCMC obtenidas convergen a la distribución a posteriori, se usa el gráfico *Z-score* de Geweke (Geweke, 1992) de la función `geweke.plot` de `coda` en R (Plummer, Best, Cowles y Vines, 2006). Para obtener el *Z-score* de Geweke se construye el equivalente a una prueba *t*, para verificar la igualdad de las medias de la primera y última parte de las muestras MCMC, y la gráfica muestra los valores de los *Z-score* donde se eliminan sucesivamente un mayor número (a lo sumo la mitad de la cadena) de iteraciones desde el inicio de la cadena. Además, se revisa la evolución de la estimación puntual (medias a posteriori) de cada uno de los parámetros, cuando el número de iteraciones se va incrementando, a través del gráfico `cumuplot` de `coda` en R. Este gráfico presenta los cuantiles muestrales 0.025, 0.50 y 0.975, como una función del número de iteraciones (promedios acumulados) de las cadenas de Markov.

## 1.4. La prueba Bayesiana de estacionalidad multiplicativa por regímenes.

En esta sección se presenta una prueba estadística Bayesiana para decidir entre un modelo TSARX y un TARX. De Gooijer y Vidiella-i-Anguera (2003), utilizando un enfoque frecuentista, presentan un modelo que captura estacionalidad multiplicativa por regímenes y no linealidad de umbrales en forma simultánea, el cual es un caso especial de un modelo SETAR usual, al imponer ciertas restricciones sobre los coeficientes autorregresivos en cada régimen del SETAR. De Gooijer y Vidiella-i-Anguera (2003), introducen dos pruebas estadística relacionadas a las restricciones sobre los coeficientes autorregresivos de un modelo SETAR. Las pruebas estadísticas son del tipo Wald con distribuciones asintóticas chi cuadrado, los cuales forman la base de una prueba de estacionalidad multiplicativa por regímenes.

En este trabajo, para probar la presencia de estacionalidad multiplicativa en cada uno de los regímenes de un modelo de la familia TARX; se propone una prueba estadística Bayesiana donde se emplea, una técnica aproximada para el cálculo del factor Bayes, en particular se hace uso de la razón de distribuciones de Savage-Dickey, dado en Dickey (1971).

La metodología consiste en comparar el modelo de umbrales con estacionalidad multiplicativa por regímenes (TSARX) con un modelo donde no se capture dicha estacionalidad

(TARX); es decir, que el vector de parámetros  $A_{j2}$  definido para el modelo TARX dado en (1.2), sea igual al vector

$$A_{j20} = (0, \dots, 0, -a_1^{(j)} a_s^{(j)}, \dots, -a_{k_j}^{(j)} a_s^{(j)}, \dots, 0, \dots, 0, -a_1^{(j)} a_{sK_j}^{(j)}, \dots, -a_{k_j}^{(j)} a_{sK_j}^{(j)}),$$

para cada  $j = 1, \dots, l$ , siguiendo las restricciones sobre los parámetros, para que un modelo TSARX sea un caso especial de un TARX, dados en la sección 1.1. Asumiendo conocido o estimado en forma preliminar el número de regímenes, se requieren estimaciones del resto de parámetros del modelo TARX, para implementar la prueba estadística.

Se desea probar la hipótesis  $H_{0j} : A_{j2} = A_{j20}$  versus  $H_{aj} : A_{j2} \neq A_{j20}$ ,  $j = 1, \dots, l$ . El factor de Bayes para  $A_{j2} = A_{j20}$  es dado por

$$FB_j = \frac{\int f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) f(\boldsymbol{\theta}_x) d\boldsymbol{\theta}_x}{\int f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x_0}) f(\boldsymbol{\theta}_{x_0}) d\boldsymbol{\theta}_{x_0}}, \quad (1.5)$$

donde  $f(\cdot)$  y  $f(\cdot|\cdot)$  denotan respectivamente, la distribución a priori y función de verosimilitud condicional del modelo TARX( $l; k_1 + sK_1, \dots, k_l + sK_l; q_1, \dots, q_l$ ),  $K = \max\{k_j + sK_j, q_j, j = 1, \dots, l\}$  y  $\boldsymbol{\theta}_{x_0}$  resume los parámetros del modelo TARX en el caso que  $A_{j2} = A_{j20}$  para  $j = 1, \dots, l$ .

Para calcular este factor de Bayes se usa la razón de distribuciones de Savage-Dickey, como una aproximación del mismo (véase, Verdinelli y Wasserman (1995)). El factor de Bayes (1.5) es igual a

$$FB_j = \frac{f(A_{j2} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}) |_{A_{j2}=A_{j20}}}{f(A_{j2}) |_{A_{j2}=A_{j20}}},$$

para  $j = 1, \dots, l$ ; es decir, la razón de los pesos de la distribución a posteriori condicional de  $A_{j2}$  y los pesos de la distribución a priori de  $A_{j2}$ , ambas evaluadas en  $A_{j20}$ ,  $j = 1, \dots, l$ . Koop y Potter (1999) utilizan la razón de distribuciones de Savage-Dickey para crear una prueba de no linealidad de umbrales; mientras que, Fok, Franses y Paap (2007) la emplean para proponer una prueba de estacionalidad determinística para modelos Bayesianos jerárquicos de dos niveles.

El peso de la distribución a posteriori marginal de  $A_{j2}$  en  $A_{j20}$  puede obtenerse con el cálculo del promedio de la distribución a posteriori condicional de  $A_{j2}$  sobre las muestras MCMC de los otros parámetros; así,

$$f(A_{j2} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}) |_{A_{j2}=A_{j20}} = \frac{1}{G} \sum_{i=1}^G f(A_{j2} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-A_{j2}}^{(i)}) |_{A_{j2}=A_{j20}},$$

para  $j = 1, \dots, l$ , si la distribución a posteriori condicional completa de  $A_{j2}$  tiene una distribución estándar conocida (véase, Gelfand y Smith, 1990) y  $\boldsymbol{\theta}_{x-A_{j2}}$  es el vector de parámetros  $\boldsymbol{\theta}_x$  sin el subvector  $A_{j2}$ .

Aquí es necesario obtener muestras MCMC de la distribución a posteriori de  $A_{j2}$ , metodologías que se presenta a continuación:

Dado el número de regímenes  $l$  y órdenes autorregresivos  $k_j + sK_j$ ,  $q_j$ ,  $j = 1, \dots, l$ , se estiman el resto de parámetros de un modelo TARX definido en (1.2). Se asume independencia a priori entre los grupos de parámetros y métodos Bayesianos estándar, en el que se incluye distribuciones a priori conjugadas.

Las distribuciones a priori para los bloques de parámetros son:

- (1)  $A_{j1} \sim N_{k_j+K_j+1}(\boldsymbol{\theta}_{j10}, V_{j10}^{-1})$ ,  $j = 1, \dots, l$ .
- (2)  $A_{j2} \sim N_{K_j(s-1)}(\boldsymbol{\theta}_{j20}, V_{j20}^{-1})$ ,  $j = 1, \dots, l$ .
- (3)  $C_j \sim N_{q_j}(\boldsymbol{\theta}_{j0}, V_{j0}^{-1})$ ,  $j = 1, \dots, l$ .
- (4)  $(h^{(j)})^2 \sim IG(\frac{\nu_j}{2}, \frac{\nu_j \lambda_j}{2})$ ,  $j = 1, \dots, l$ .
- (5)  $\mathbf{r} = (r_1, \dots, r_{l-1})' \sim f(\mathbf{r}) = \frac{1}{C} I(\mathbf{A})$ , con  $\mathbf{C}$  y  $\mathbf{A}$  definida en (v) de la sección 1.3. Se requieren cuantiles convenientes  $a$  y  $b$  de  $Z_{t-d}$ .
- (6)  $d \sim U_d\{0, 1, 2, \dots, d_0\}$ , una uniforme discreta, recalando que  $d_0$  es un rezago máximo dado.

Los hiperparámetros  $(\boldsymbol{\theta}_{j10}, V_{j10}, \boldsymbol{\theta}_{j20}, V_{j20}, \boldsymbol{\theta}_{j0}, V_{j0}, \nu_j, \lambda_j, a, b, d_0)$  se asumen conocidos. Teniendo en cuenta la función de verosimilitud condicional para un modelo TARX dado en (1.4) y cada una de las distribuciones a priori dadas en (1) a (6), se obtienen las distribuciones a posteriori de los parámetros  $A_{j1}$ ,  $A_{j2}$ ,  $C_j$ ,  $(h^{(j)})^2$ ,  $j = 1, \dots, l$ ,  $\mathbf{r}$  y  $d$ ; y se logran conseguir muestras de los parámetros en mención, a través del siguiente esquema:

Algoritmo 3.

1. Inicializar:  $A_{j1}^{(0)}, A_{j2}^{(0)}, C_j^{(0)}, [(h^{(j)})^2]^{(0)}, j = 1, \dots, l, \mathbf{r}^{(0)}, d^{(0)}$ .
2. Generar  $A_{j1}$  (independiente de  $A_{i1}$ ,  $i \neq j$ ) de

$$f(A_{j1} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-A_{j2}}) \sim N_{k_j+K_j+1}(\boldsymbol{\theta}_{j1}, V_{j1}^{-1})$$

donde

$$\begin{aligned}
V_{j1} &= [1/(h^{(j)})^2] \tilde{S}'_j \tilde{S}_j + V_{j10} \quad y \\
\boldsymbol{\theta}_{j1} &= V_{j1}^{-1} ([1/(h^{(j)})^2] \tilde{S}'_j (\mathbb{X}_j - \tilde{M}_j A_{j2} - \tilde{N}_j C_j) + V_{j10} \boldsymbol{\theta}_{j10}) \\
\tilde{s}'_t &= (1, x_{t-1}, \dots, x_{t-k_j}, x_{t-s}, \dots, x_{t-sK_j}), \\
\tilde{S}_j &= (\tilde{s}_{t_1, j}, \dots, \tilde{s}_{t_{n_j}, j})', \\
\tilde{m}'_t &= (x_{t-(1+k_j)}, \dots, x_{t-(s-1)}, x_{t-(1+s)}, \dots, x_{t-(k_j+s)}, \dots, \\
&\quad x_{t-[(s-1)+s(K_j-1)], x_{t-(1+sK_j)}, \dots, x_{t-(k_j+sK_j)}), \\
\tilde{M}_j &= (\tilde{m}_{t_1, j}, \dots, \tilde{m}_{t_{n_j}, j})', \\
\tilde{n}'_t &= (z_{t-1}, \dots, z_{t-q_j}), \\
\tilde{N}_j &= (\tilde{n}_{t_1, j}, \dots, \tilde{n}_{t_{n_j}, j})', \quad \mathbb{X}_j = (x_{t_1, j}, \dots, x_{t_{n_j}, j})', \quad y \quad n_j \text{ es el número de observa-} \\
&\quad \text{ciones en el régimen } j, \quad j = 1, \dots, l.
\end{aligned}$$

3. Generar  $A_{j2}$  (independiente de  $A_{i2}$ ,  $i \neq j$ ) de

$$f(A_{j2} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-A_{j2}}) \sim N_{K_j(s-1)}(\boldsymbol{\theta}_{j2}, V_{j2}^{-1})$$

donde

$$\begin{aligned}
V_{j2} &= [1/(h^{(j)})^2] \tilde{M}'_j \tilde{M}_j + V_{j20} \quad y \\
\boldsymbol{\theta}_{j2} &= V_{j2}^{-1} ([1/(h^{(j)})^2] \tilde{M}'_j (\mathbb{X}_j - \tilde{S}_j A_{j1} - \tilde{N}_j C_j) + V_{j20} \boldsymbol{\theta}_{j20}),
\end{aligned}$$

con  $\tilde{M}_j$ ,  $\tilde{S}_j$ ,  $\tilde{N}_j$  y  $\mathbb{X}_j$ ,  $j = 1, \dots, l$  definidos anteriormente.

4. Generar  $C_j$  (independiente de  $C_i$ ,  $i \neq j$ ) de

$$f(C_j | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-C_j}) \sim N_{q_j}(\boldsymbol{\theta}_j, V'_j)$$

donde

$$\begin{aligned}
V_j &= [1/(h^{(j)})^2] \tilde{N}'_j \tilde{N}_j + V_{j0} \quad y \\
\boldsymbol{\theta}_j &= V_j^{-1} ([1/(h^{(j)})^2] \tilde{N}'_j (\mathbb{X}_j - \tilde{S}_j A_{j1} - \tilde{M}_j A_{j2}) + V_{j0} \boldsymbol{\theta}_{j0}),
\end{aligned}$$

con  $\tilde{M}_j$ ,  $\tilde{S}_j$ ,  $\tilde{N}_j$  y  $\mathbb{X}_j$ ,  $j = 1, \dots, l$  definidos anteriormente.

5. Generar  $(h^{(j)})^2$  (independiente de  $(h^{(i)})^2$ ,  $i \neq j$ ) de

$$f((h^{(j)})^2 | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-(h^{(j)})^2}) \sim IG \left( \frac{\nu_j + n_j}{2}, \frac{\nu_j \lambda_j + n_j s_j^2}{2} \right)$$

donde  $n_j s_j^2 = (\mathbb{X}_j - A_j^* B_j^*)' (\mathbb{X}_j - A_j^* B_j^*)$ , con  $A_j^*$  y  $B_j^*$   $j = 1, \dots, l$ , definidos en la ecuación (1.4).

6. Generar  $\mathbf{r} = (r_1, \dots, r_{l-1})'$  de

$$f(\mathbf{r} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-\mathbf{r}}) \propto f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}) I(\mathbf{A}),$$

donde  $f(\cdot|\cdot)$  es la función de verosimilitud condicional dado en (1.4), que depende de los valores umbrales  $\mathbf{r}$ , a través del número de observaciones  $n_j$  de cada régimen,  $j = 1, \dots, l$  y  $\mathbf{A}$  es la región especificada en la distribución a priori para  $\mathbf{r}$ .

7. Generar  $d$  de una distribución multinomial, con función de masa de probabilidad

$$Pr(d = i | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-d}) = \frac{f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-d}, d = i) Pr(d = i)}{\sum_{m=0}^{d_0} f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-d}, d = m) Pr(d = m)},$$

$$i = 0, 1, \dots, d_0.$$

Se logran muestras de las distribuciones a posteriori dadas en el Algoritmo 3, usando un muestreador de Gibbs incorporando el algoritmo MH para  $\mathbf{r}$  (Algoritmo 2). Para monitorear la convergencia de las cadenas MCMC y obtener estimaciones de los parámetros desconocidos del modelo TARX, se sigue una metodología similar a la descrita en la sección 1.3.

Regresando al cálculo del factor Bayes, debido que

$$A_{j2}^{(i)} \sim N(\boldsymbol{\theta}_{j2}^{(i)}, (V_{j2}^{-1})^{(i)}), \quad i = 1, \dots, G, \quad j = 1, \dots, l,$$

se utilizan los parámetros  $\boldsymbol{\theta}_{j2}, V_{j2}$  de la a posteriori y  $\boldsymbol{\theta}_{j20}, V_{j20}$  de la a priori para  $A_{j2}$ ; para el cálculo del factor de Bayes por regímenes, dado en la ecuación (1.5). Por lo tanto,

$$FB_j = \frac{\frac{1}{G} \sum_{i=1}^G 1/\sqrt{|(V_{j2}^{-1})^{(i)}|} \exp\{-\frac{1}{2}(A_{j20} - \boldsymbol{\theta}_{j2}^{(i)})' V_{j2}^{(i)} (A_{j20} - \boldsymbol{\theta}_{j2}^{(i)})\}}{1/\sqrt{|(V_{j20}^{-1})|} \exp\{-\frac{1}{2}(A_{j20} - \boldsymbol{\theta}_{j20})' V_{j20} (A_{j20} - \boldsymbol{\theta}_{j20})\}},$$

$j = 1, \dots, l$ . No se rechaza  $H_{0j} : A_{j2} = A_{j20}$  si  $2 \ln(FB_j) > 6$ , (evidencia fuerte o muy fuerte a favor de  $H_{0j}$ ) para cada  $j = 1, \dots, l$ , según el criterio de decisión dado por Kass y Raftery (1995).

Se decide utilizar un modelo TSARX( $l; k_1, \dots, k_l; K_1, \dots, K_l; q_1, \dots, q_l$ )<sub>s</sub> para ajustar los datos, en lugar de un modelo TARX( $l; k_1 + sK_1, \dots, k_l + sK_l; q_1, \dots, q_l$ ), si en todos los regímenes  $j$ ,  $H_{0j}$  no se rechaza,  $j = 1, \dots, l$ .

## 1.5. Estimación del número de regímenes.

El identificar el número de regímenes  $l$  en un modelo TSARX, puede verse como un problema de selección Bayesiana de modelos vía técnicas MCMC. La decisión entre modelos  $M_j$ ,  $j = 1, \dots, l_0$ , siendo  $M_j$  un modelo TSARX con  $j$  regímenes y  $l_0$  un número máximo de regímenes considerado, es realizada por el cálculo de probabilidades a posteriori, donde se elige el modelo  $M_j$  si  $Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T})$  es un máximo sobre  $j = 1, \dots, l_0$ . Para estimar estas probabilidades para cada modelo se involucran integraciones múltiples sobre cada espacio de parámetros del modelo. Varias técnicas MCMC han sido propuestos para la selección de modelos, tales como el de Carlin y Chib (1995) aplicado para determinar el número de regímenes en modelos TAR en Nieto (2005), el muestreador de saltos reversibles RJMCMC propuesto por Green (1995)

y utilizado en Nieto, Zhang y Li (2013); mientras, los métodos de aproximación de la verosimilitud marginal dado por Chib (1995) y la versión metropolizada del algoritmo de Carlin y Chib mencionada en Dellaportas, Forster y Ntzoufras (2002), son aplicados en Calderón y Nieto (2017) para modelos TARX multivariados. Sin embargo, todas estas técnicas involucran extensiones MCMC, así que el tiempo computacional requerido es usualmente grande o se requiere la especificación de distribuciones a priori de enlace, cuya escogencia pueden afectar la selección del modelo final en forma significativa. En este trabajo, se emplea una técnica MCMC propuesto por Congdon (2006) el cual con algunas condiciones, no presenta las dificultades de las técnicas anteriormente mencionadas.

Congdon (2006) desarrolló el siguiente estimador de probabilidad a posteriori, que requiere muestras MCMC separadas independientes para cada uno de los modelos candidatos y denotado como  $\{\boldsymbol{\theta}^{(i)} = (\boldsymbol{\theta}_1^{(i)}, \dots, \boldsymbol{\theta}_{l_0}^{(i)}), i = 1, \dots, G\}$ . Una estimación Monte Carlo aproximada de  $Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T})$  es

$$Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T}) \approx \frac{1}{G} \sum_{i=1}^G Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_j^{(i)}) = \frac{1}{G} \sum_{i=1}^G \frac{\varphi_j^{(i)}}{\sum_{k=1}^{l_0} \varphi_k^{(i)}}, \quad (1.6)$$

donde

$$\varphi_k^{(i)} = f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_k^{(i)}, M_k) f(\boldsymbol{\theta}_k^{(i)} | M_k) Pr(M_k),$$

$\boldsymbol{\theta}_k$  son los parámetros del modelo  $k$ ,  $\boldsymbol{\theta}_k^{(i)}$  es la  $i$ -ésima iteración MCMC de la distribución a posteriori del modelo  $k$ ,  $Pr(M_k)$  se elige como  $Pr(M_k) = \frac{1}{l_0}$ , para reflejar ignorancia a priori en la selección del modelo,  $f(\boldsymbol{\theta}_k^{(i)} | M_k)$  es la distribución a priori para el modelo  $k$  y  $f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_k^{(i)})$  es la función de verosimilitud condicional para  $M_k$ , para algún  $k = 1, \dots, l_0$ .

Para eficiencia numérica, en la práctica el cálculo de (1.6) es basado sobre versiones escaladas de las logverosimilitudes. La escala emplea el máximo  $L_{\max}^{(i)}$  de las logverosimilitudes en cada iteración y sustrae este de cada modelo de logverosimilitud, es decir

$$\varphi_k^{(i)} = \exp(L_k^{(i)} - L_{\max}^{(i)}) f(\boldsymbol{\theta}_k^{(i)} | M_k) Pr(M_k), \quad i = 1, \dots, G, \text{ para algún } k = 1, \dots, l_0.$$

Algunos resultados soportan que el método es bastante preciso y potente para selección de modelos. Por ejemplo, Chen, Gerlach y Lin (2010) comparan modelos TARX con errores GARCH con diferentes regímenes; mientras, Gerlach y Chen (2008), lo usan en modelos STAR con errores GARCH.

Se examina la efectividad de esta aproximación, comparándolo con el criterio de información de la devianza (DIC), el cual se describe a continuación. Spiegelhalter *et al.* (2002) proponen un criterio de comparación Bayesiano de modelos definido como:

$$DIC = 2E_{\boldsymbol{\theta}_j | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}}(D(\boldsymbol{\theta}_j)) - D(E_{\boldsymbol{\theta}_j | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}}(\boldsymbol{\theta}_j)), \quad (1.7)$$

donde

$$D(\tilde{\boldsymbol{\theta}}_j) = -2 \ln f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \tilde{\boldsymbol{\theta}}_j),$$

con  $f(\cdot|\cdot)$  la función de verosimilitud condicional del modelo TSARX con  $j$  regímenes,  $j = 1, \dots, l_0$ .

Si  $\boldsymbol{\theta}_j^{(1)}, \dots, \boldsymbol{\theta}_j^{(G)}$  son muestras MCMC de la distribución  $f(\boldsymbol{\theta}_j|\mathbf{x}_{1,T}, \mathbf{z}_{1,T})$ , entonces el DIC puede aproximarse vía Monte Carlo por

$$\text{DIC} = \frac{2}{G} \sum_{i=1}^G D(\boldsymbol{\theta}_j^{(i)}) - D\left(\frac{1}{G} \sum_{i=1}^G \boldsymbol{\theta}_j^{(i)}\right),$$

donde

$$D(\tilde{\boldsymbol{\theta}}_j^{(i)}) = -2 \ln f(\mathbf{x}_{K+1,T}|\mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \tilde{\boldsymbol{\theta}}_j^{(i)}),$$

$i = 1, \dots, G, j = 1, \dots, l_0$ . Se elige el modelo  $M_j, j = 1, \dots, l_0$  cuyo correspondiente DIC sea el mínimo entre los modelos candidatos.

## 1.6. Diagnósticos con los residuales del modelo.

Probar la adecuación de un modelo es una parte importante de cualquier análisis de una serie de tiempo. Usualmente, se utilizan los residuales estandarizados, dados por  $e_t = (X_t - X_{t|t-1})/h^{(j)}$ , si  $r_{j-1} < Z_{t-d} \leq r_j$  para algún  $j = 1, \dots, l$ , donde

$$X_{t|t-1} = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i} + \sum_{u=1}^{K_j} b_u^{(j)} X_{t-su} - \sum_{i=1}^{k_j} \sum_{u=1}^{K_j} a_i^{(j)} b_u^{(j)} X_{t-i-su} + \sum_{v=1}^{q_j} c_v^{(j)} Z_{t-v}$$

es el predictor un paso adelante de  $X_t$  para un modelo TSARX, donde  $t = (K+1), \dots, T$  con  $K = \text{máximo}\{k_j + sK_j, q_j, j = 1, \dots, l\}$  y  $T$  la longitud del periodo muestral.  $\{e_t\}$  debe aproximarse al comportamiento dinámico estocástico de un ruido blanco Gaussiano, si el modelo considerado es el correcto. Nieto (2005) considera el gráfico CUSUMSQ para chequear heterocedasticidad marginal en los residuales y el gráfico CUSUM para determinar la especificación de un modelo TAR. Ahora, basado sobre técnicas MCMC e inferencia Bayesiana, es natural considerar estas técnicas que logre incorporar la incertidumbre de la estimación de los parámetros del modelo en el cálculo de los residuales. En este trabajo, se emplea el método de Gerlach, Carter y Kohn (1999), el cual es basado sobre técnicas MCMC, muestreo por importancia y la serie de tiempo  $u_t = F_t(x_t|\mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1})$ ,  $t = (K+1), \dots, T$ , donde  $F_t(\cdot|\cdot)$  es la función de distribución condicional acumulada dada en la sección 1.1, y  $K = \text{máximo}\{k_j + sK_j, q_j, j = 1, \dots, l\}$ . Gerlach, Carter y Kohn (1999) muestran que para  $k \geq t$ , el estimador

$$\hat{u}_t = \frac{\sum_{i=1}^G F_t(x_t|\mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, \boldsymbol{\theta}_k^{(i)})/f(x_{t,k}|\mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, \boldsymbol{\theta}_k^{(i)})}{\sum_{i=1}^G 1/f(x_t|\mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, \boldsymbol{\theta}_k^{(i)})} \quad (1.8)$$

converge a  $u_t$  cuando  $G \rightarrow \infty$ . Aquí  $\boldsymbol{\theta}_k^{(i)}$  es la  $i$ -ésima iteración muestral MCMC de la distribución a posteriori  $f(\boldsymbol{\theta}_k|\mathbf{x}_{1,k}, \mathbf{z}_{1,k})$  para un  $k$  dado y  $f(\cdot|\cdot)$  es la función de verosimilitud condicional para el modelo TSARX con  $l$  regímenes. Debido que la varianza de  $\hat{u}_t$  en la ecuación (1.8) se incrementa con  $(k-t)$ , se requiere calcular  $\hat{u}_t$  con  $t$  razonablemente cercano a  $k$ . Esto puede ser alcanzado incrementando  $k$  en forma

secuencial, digamos  $100, 200, \dots, T$  y evaluar  $\hat{u}_t$  usando la ecuación (1.8) con  $(k - t)$  no más grande que los incrementos. Basado sobre las propiedades de la convergencia de  $\hat{u}_t$ , los residuales generalizados  $\hat{v}_t = \Phi^{-1}(\hat{u}_t)$  ( $\Phi$  es la función de distribución normal estándar acumulada) son aproximadamente independientes e idénticamente distribuidos  $N(0, 1)$ , bajo el modelo correcto. Pruebas diagnósticas estándar pueden aplicarse a  $\hat{v}_t$  tal como la prueba de Ljung y Box para determinar no autocorrelación, la prueba estadística de Jarque y Bera para verificar normalidad, los gráficos de sumas acumuladas (CUSUM) para evidenciar la especificación correcta del modelo y de sumas acumuladas al cuadrado (CUSUMSQ) para indicar presencia o no de heterocedasticidad marginal en  $\{X_t\}$ , entre otros.

Chen y So (2006) y Chen, Gerlach y Lin (2010) conducen el chequeo de diagnósticos en modelos TARX con errores GARCH, usando residuales generalizados siendo muy efectivos en determinar la adecuación del modelo.

## 1.7. El cálculo de pronósticos.

En esta sección se desarrolla la etapa de pronósticos con los modelos TSARX con  $l$  regímenes. Para ello, se requiere hallar  $\hat{X}_{T+h} = E(X_{T+h} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T})$ , donde  $h \geq 1$  es el horizonte de predicción,  $\mathbf{x}_{1,T} = (x_1, \dots, x_T)$ ,  $\mathbf{z}_{1,T} = (z_1, \dots, z_T)$  y  $T$  es la longitud del periodo muestral.  $\hat{X}_{T+h}$  es la mejor predicción en el sentido del mínimo error cuadrático medio, cuya expresión analítica es difícil o imposible de obtener en el contexto de las series de tiempo no lineales (véase, Franses y Van Dijk, 2000). Por tal razón, se aborda el problema de calcular  $\hat{X}_{T+h}$  a través del enfoque Bayesiano mediante el cálculo de las distribuciones predictivas  $f(x_{T+h} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T})$  para  $h \geq 1$ . Las distribuciones predictivas de interés pueden obtenerse en forma marginal de la distribución predictiva conjunta

$$f(x_{T+1}, \dots, x_{T+h}, z_{T+1}, \dots, z_{T+h}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}), \quad (1.9)$$

con  $\boldsymbol{\theta}_x$  el vector de parámetros del modelo TSARX y  $\boldsymbol{\theta}_z$  el vector de parámetros de la variable  $Z_t$ . Trabajos previos en esta dirección son dados en Nieto (2008) y Vargas (2012), entre otros.

Se supondrá en lo que sigue, que las especificaciones dadas en secciones anteriores, respecto al modelo elegido se cumplen. Se requiere obtener muestras de (1.9), basado sobre la descomposición de la distribución predictiva conjunta:

$$\begin{aligned} & f(x_{T+1}, \dots, x_{T+h}, z_{T+1}, \dots, z_{T+h}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}) = \\ & f(x_{T+1}, \dots, x_{T+h}, z_{T+1}, \dots, z_{T+h}, \boldsymbol{\theta}_z | \boldsymbol{\theta}_x, \mathbf{x}_{1,T}, \mathbf{z}_{1,T}) f(\boldsymbol{\theta}_x | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}) \\ & = \left\{ \prod_{i=1}^h f(z_{T+i} | \mathbf{z}_{1,T+i-1}, \boldsymbol{\theta}_z) f(x_{T+i} | z_{T+i}, \mathbf{x}_{1,T+i-1}, \mathbf{z}_{1,T+i-1}, \boldsymbol{\theta}_x) \right\} f(\boldsymbol{\theta}_x | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}), \end{aligned}$$

donde

- (i)  $f(z_{T+i} | \mathbf{z}_{1,T+i-1}, \boldsymbol{\theta}_z)$  es la función de densidad kernel de la cadena de Markov  $\{Z_t\}$ .

(ii)  $f(x_{T+i}|z_{T+i}, \mathbf{x}_{1,T+i-1}, \mathbf{z}_{1,T+i-1}, \boldsymbol{\theta}_x)$  es una distribución normal con media

$$a_0^{(j)} + \sum_{m=1}^{k_j} a_m^{(j)} x_{T+i-m} + \sum_{u=1}^{K_j} b_u^{(j)} x_{T+i-su} - \sum_{m=1}^{k_j} \sum_{u=1}^{K_j} a_m^{(j)} b_u^{(j)} x_{T+i-m-su} + \sum_{v=1}^{q_j} c_v^{(j)} z_{T+i-v},$$

y varianza  $(h^{(j)})^2$ , si  $r_{j-1} < z_{T+i-d} \leq r_j$ , para algún  $j = 1, \dots, l$ .

(iii)  $f(\boldsymbol{\theta}_x|\mathbf{x}_{1,T}, \mathbf{z}_{1,T})$  es la distribución a posteriori de los parámetros del modelo TSARX;

además,  $\mathbf{x}_{1,T+i-1} = (x_1, \dots, x_{T+i-1})$  y  $\mathbf{z}_{1,T+i-1} = (z_1, \dots, z_{T+i-1})$ . Así, de la distribución predictiva conjunta mediante simulación, se puede extraer  $z_{T+1}, x_{T+1}, z_{T+2}, x_{T+2}, \dots, z_{T+h}, x_{T+h}$  (Congdon, 2001) en forma secuencial como sigue:

Algoritmo 4.

1. Generar una muestra MCMC de  $\boldsymbol{\theta}_x^{(k)}$  de la distribución a posteriori  $f(\boldsymbol{\theta}_x|\mathbf{x}_{1,T}, \mathbf{z}_{1,T})$ .
2. Generar  $z_{T+1}^{(k)}$  de  $f(z_{T+1}|\mathbf{z}_{1,T}, \boldsymbol{\theta}_z)$ .
3. Generar  $x_{T+1}^{(k)}$  de  $f(x_{T+1}|z_{T+1}^{(k)}, \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x^{(k)})$ .
4. Generar  $z_{T+i}^{(k)}$  de  $f(z_{T+i}|\mathbf{z}_{1,T}, z_{T+1}^{(k)}, \dots, z_{T+i-1}^{(k)}, \boldsymbol{\theta}_z)$ .
5. Generar  $x_{T+i}^{(k)}$  de  $f(x_{T+i}|z_{T+i}^{(k)}, \mathbf{x}_{1,T}, x_{T+1}^{(k)}, \dots, x_{T+i-1}^{(k)}, \mathbf{z}_{1,T}, z_{T+1}^{(k)}, \dots, z_{T+i-1}^{(k)}, \boldsymbol{\theta}_x^{(k)})$ .
6. Repetir los pasos 4 y 5 para  $i = 2, 3, \dots, h$ .

Con las muestras  $\{z_{T+h}^{(k)}, x_{T+h}^{(k)}\}$ ,  $k = 1, \dots, G$ , con  $G$  la longitud de la cadena y  $h \geq 1$ , se calcula las medias (pronósticos puntuales), desviaciones estándar (medida de incertidumbre), y cotas inferior y superior de intervalos a posteriori Bayes (intervalos creíbles), o regiones de credibilidad del  $100(1 - \alpha)\%$  ( $0 < \alpha < 1$ ), tanto para la variable de umbrales, como para la variable de interés.

El procedimiento descrito en esta sección, incluye la incertidumbre de la estimación de los parámetros del modelo en el cálculo de pronósticos como lo trata Vargas (2012) en modelos TAR.

## 1.8. La metodología propuesta.

La metodología propuesta para ajustar y pronósticar con un modelo TSARX dada en (1.1), es dado por los siguientes pasos:

- Paso 1. Se realiza un análisis exploratorio de las variables de interés y de umbrales bajo objeto de estudio, para determinar en forma preliminar la presencia o no de estacionalidad multiplicativa. Para ello, se aplican pruebas de raíces unitarias y de raíces unitarias estacionales a las series de tiempo, el cálculo de raíces de

---

los polinomios característicos por regímenes y se usan gráficas de las series de tiempo, correlogramas, gráficas de las series de tiempo por meses o trimestres según sea el caso, entre otras.

- Paso 2. Se fijan órdenes autorregresivos máximos, los cuales se obtienen ajustando modelos candidatos ARX a los datos y eligiendo aquel modelo con DIC mínimo. Se procede a realizar la prueba de no linealidad de umbrales propuesta por Tsay (1988) y se prueba la presencia de estacionalidad multiplicativa por regímenes, usando la prueba estadística propuesta en la sección 1.4. Dado el número de regímenes y órdenes autorregresivos máximos, son necesarios valores intrínsecos del resto de parámetros de un modelo TARX dado en (1.2), para implementar la prueba estadística.
- Paso 3. Según los resultados obtenidos en el paso 2, se decide ajustar los datos a un TARX o un TSARX. Se procede a estimar el número de regímenes, órdenes autorregresivos y el resto de parámetros del modelo TSARX siguiendo lo descrito en las secciones 1.3 (método GVS) y 1.5 (método de Congdon (2006)). Es de anotar, que los métodos de Congdon (2006) y GVS se pueden implementar en modelos TARX, siguiendo una metodología similar a la presentada para los TSARX.
- Paso 4. Se valida el modelo seleccionado, mediante la aplicación de pruebas diagnósticas a los residuales generalizados, descrito en la sección 1.6.
- Paso 5. Se hace uso del modelo seleccionado, con el cálculo de pronósticos  $h$  pasos adelante, de las variables de interés y de umbrales, aplicando la metodología presentada en la sección 1.7.

# CAPÍTULO 2

---

---

## Un estudio de simulación

---

---

En este capítulo se evalúa el funcionamiento de la metodología propuesta, dada en la sección 1.8. Se conduce un estudio de simulación bajo las especificaciones de dos modelos TSARX, que se definen a continuación.

### 2.1. Modelo 1.

Se asume que  $\{X_t\}$  sigue el modelo  $TSARX(2; 1, 1; 2, 1; 1, 3)_{12}$  dado por

$$X_t = \begin{cases} 2.34 + 0.50X_{t-1} + 0.20X_{t-12} - 0.10X_{t-13} + 0.10X_{t-24} \\ -0.05X_{t-25} + 1.23Z_{t-1} + \epsilon_t, & \text{si } Z_{t-2} \leq r, \\ -4.50 + 0.60X_{t-1} + 0.10X_{t-12} - 0.06X_{t-13} - 1.15Z_{t-1} \\ + 3.30Z_{t-2} - 1.92Z_{t-3} + 4\epsilon_t, & \text{si } Z_{t-2} > r, \end{cases} \quad (2.1)$$

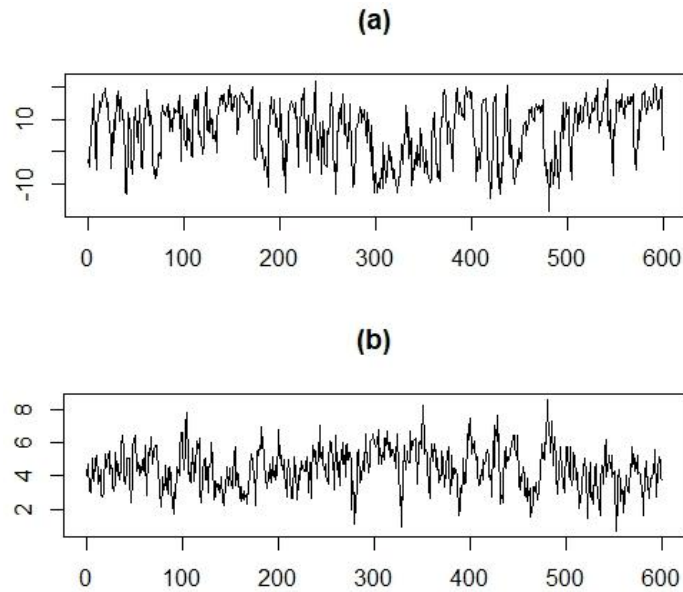
donde  $\{Z_t\}$  obedece un modelo AR(1),  $Z_t = 1.80 + 0.60Z_{t-1} + a_t$ , con  $\{a_t\} \sim RBG(0, 1)$ , mutuamente independientes de  $\{\epsilon_t\}$  y  $r = 4.46$ , siendo  $r$  la mediana de  $Z_{t-2}$ . Note que  $\{Z_t\}$  es una cadena de Markov homogénea de primer orden, y un proceso estocástico estrictamente estacionario.

#### 2.1.1. Análisis exploratorio de los datos.

En el paso uno, se realiza el análisis exploratorio de las series de tiempo involucradas en la ecuación (2.1). Se hace uso de pruebas de raíces unitarias y raíces unitarias estacionales, y de gráficos de las series de tiempo, autocorrelación muestral (FAC), autocorrelación parcial muestral (FACP) y el índice de información mutua promedio (IMP), para determinar en forma preliminar la presencia o no de estacionalidad multiplicativa, en las series de tiempo en mención. El IMP mide la dependencia mutua entre las variables  $W_i$  y  $W_{i-\tau}$ ,  $\tau = 0, 1, \dots, \tau^*$ , siendo  $\tau^*$  un número de rezagos máximo especificado. Si  $W_i$  y  $W_{i-\tau}$  son independientes, entonces conocer  $W_i$  no da información sobre  $W_{i-\tau}$  y viceversa,

por lo que su información mutua es cero. Se usa la función mutua de `tseriesChaos` en R (Hegger, Kantz y Schreiber, 1999) para estimar el IMP.

Se calcula la prueba de raíces unitarias de Dickey y Fuller aumentada (DFA) dado en Dickey y Fuller (1979). La prueba rechaza la hipótesis nula de raíz unitaria al nivel del 5% para  $X_t$  y  $Z_t$ ; además, se calcula la prueba de raíces unitarias estacionales de Osborn, Chui, Smith y Birchenhall (1988) denotada por OCSB. La hipótesis nula de raíz unitaria estacional, se rechaza con un valor crítico al nivel del 5%, para las dos series de tiempo consideradas. En la Tabla 2.1, se reportan las estadísticas, valores  $p$  y valores críticos de las pruebas de raíces unitarias en mención.

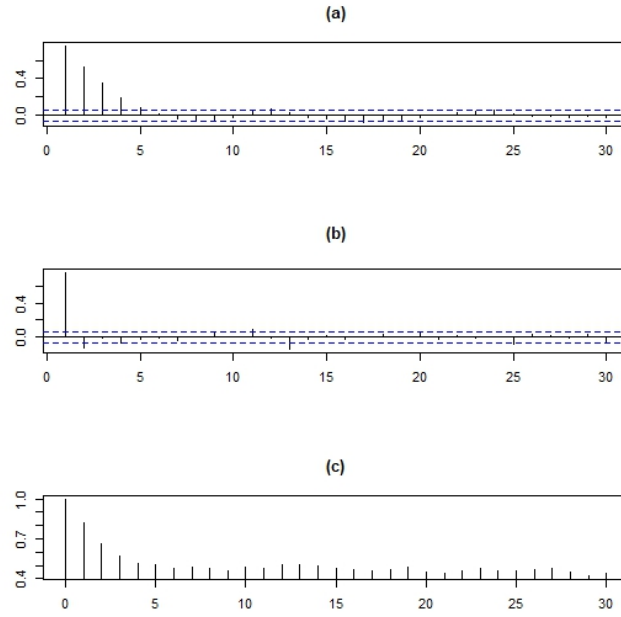


**Figura 2.1.** (a)  $X_t$  variable de interés, (b)  $Z_t$  variable de umbrales, modelo 1.

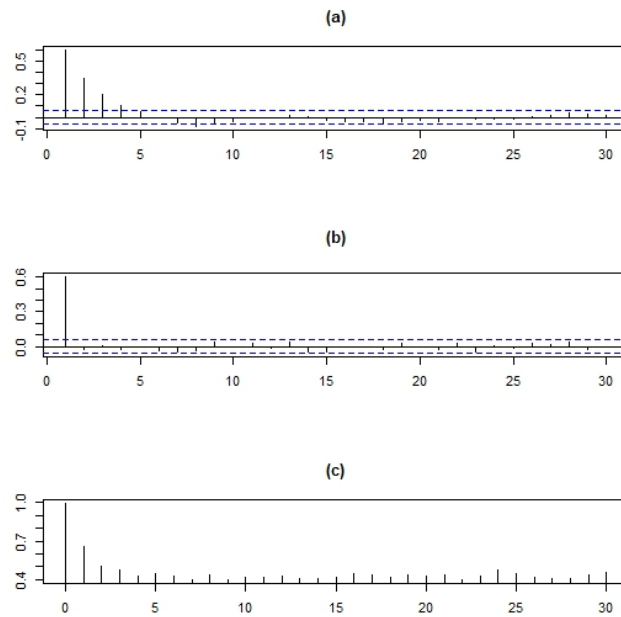
En la Figura 2.1 se muestran las dos series de tiempo simuladas dadas en (2.1), con un tamaño muestral de  $T = 600$  observaciones. Se puede ver, cierto comportamiento estable en ambas series y no evidencia a simple vista un comportamiento estacional. Para  $\{X_t\}$  (véase Figura 2.2) se observa que se logra identificar autocorrelaciones significativas en los primeros rezagos y el rezago 12 esencialmente en el FACP; ya que, el IMP solo muestra los primeros tres rezagos ser significativos. Mientras, para  $\{Z_t\}$  (véase Figura 2.3) solo es significativo el primer rezago, según muestran las gráficas FAC, FACP e IMP. Comportamientos esperados en dichas series simuladas.

**Tabla 2.1.** Pruebas de raíces unitarias y raíces unitarias estacionales para el modelo 1.

	$X_t$	$Z_t$
Estadística DFA	-4.716	-6.185
Valor $p$	0.010	0.010
Estadística OCSB	-18.991	-22.325
Valor crítico del 5%	-1.803	-1.803

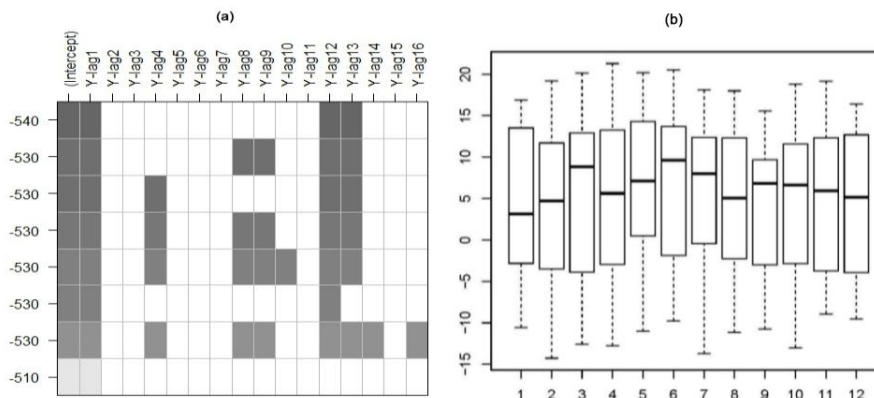


**Figura 2.2.** Para  $X_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 1.



**Figura 2.3.** Para  $Z_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 1.

Siguiendo con el análisis gráfico exploratorio, para  $\{X_t\}$  (Figura 2.4) se tiene que el mejor modelo que se ajusta a los datos es un posible AR(13) multiplicativo, según el criterio BIC mínimo; además, los diagramas de caja por meses indican que la función de medias de la serie no es constante y la variabilidad difiere entre meses.



**Figura 2.4.** Para  $X_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por meses, modelo 1.

Ahora correspondiente al modelo dado en la ecuación (2.1), algunas de las raíces de los polinomios característicos  $\phi_1(w) = 1 - 0.50w - 0.20w^{12} + 0.10w^{13} - 0.10w^{24} + 0.05w^{25}$  del primer régimen, y  $\phi_2(w) = 1 - 0.60w - 0.10w^2 + 0.06w^{13}$  ( $w \in \mathbb{C}$ , donde  $\mathbb{C}$  es el conjunto de los números complejos) del segundo régimen, le corresponden periodos de longitud 6, 12, 24 y 6, 12, respectivamente. Lo que lleva a suponer, presencia de ciclos estocásticos estacionales en la serie de tiempo  $\{X_t\}$ .

### 2.1.2. Resultados de la estimación.

En el paso dos y tres para ilustrar la metodología propuesta, se generan 100 conjuntos de datos separados de tamaño muestral  $T=600$ , del modelo (2.1). Para el conjunto de datos replicado, el muestreador MCMC se corre para 12.000 iteraciones y las primeras 6.000 iteraciones son descartadas como el periodo de calentamiento. La inferencia a posteriori es basada sobre las restantes 6.000 iteraciones. Se procede a evaluar el funcionamiento en la identificación de la presencia de estacionalidad multiplicativa por regímenes, identificación del número de regímenes, de los órdenes autorregresivos y estimación del resto de parámetros de un modelo TSARX. En la prueba Bayesiana de estacionalidad multiplicativa por regímenes se requiere estimar parámetros de un modelo TARX con dos regímenes, con  $k_1^* = k_2^* = 2$ ,  $K_1^* = K_2^* = 2$ ,  $q_1^* = q_2^* = 3$ , órdenes autorregresivos máximos considerados y  $s = 12$  el periodo estacional.

Se tomó valores para los hiperparámetros  $\theta_{j10} = \mathbf{0}_{k_j^* + K_j^* + 1}$ ,  $V_{j10} = \text{diag}\{0.10\}_{k_j^* + K_j^* + 1}$ ,  $\theta_{j20} = \mathbf{0}_{K_j^*(s-1)}$ ,  $V_{j20} = \text{diag}\{0.10\}_{K_j^*(s-1)}$ ,  $\theta_{j0} = \mathbf{0}_{q_j^*}$ ,  $V_{j0} = \text{diag}\{0.10\}_{q_j^*}$ ,  $j = 1, 2$ ; además,  $\nu_j = 3$ ,  $\lambda_j = \frac{\tilde{\sigma}^2}{3}$ , donde  $\tilde{\sigma}^2$  es la varianza muestral de  $\{X_t\}$ ,  $j = 1, 2$ ;  $a = \text{mínimo}\{Z_{t-d}\}$ ,  $b = \text{máximo}\{Z_{t-d}\}$ ,  $H = 10$  (asegura al menos 60 observaciones en cada régimen) y  $d_0 = 5$ . Las distribuciones a priori son seleccionadas a ser razonablemente

**Tabla 2.2.** Probabilidades medias a posteriori y DIC promedio para selección del número de regímenes para el modelo 1 (100 repeticiones).

	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
Prob. media a posteriori	0.00026	0.95415	0.03290	0.00429	0.00840
DIC promedio	3569.24	2603.58	2608.90	2700.77	2704.05

no informativas, así que la verosimilitud domine la inferencia.

Los valores iniciales para cada uno de los parámetros fueron:  $A_{j1}^{(0)} = 0.05_{k_j^*+K_j^*+1}$ ,  $A_{j2}^{(0)} = 0.05_{K_j^*(s-1)}$ ,  $C_j^{(0)} = 0.05_{q_j^*}$ ,  $[(h^{(j)})^2]^{(0)} = 0.20$ ,  $j = 1, 2$ ;  $d^{(0)} = 3$ , el tamaño de paso  $\sigma_r = 0.025$  para ejecutar el Algoritmo MH (Algoritmo 2), para generar muestras de  $r$  y  $r^{(0)}$  el cuantil 0.50 de  $Z_{t-d}$ . Al realizar la prueba Bayesiana descrita en la sección 1.4, para los 100 conjuntos de datos simulados que siguen la ecuación (2.1), se tiene que el número de aciertos de la prueba en detectar la presencia de estacionariedad multiplicativa en cada uno de los dos regímenes es dada en todas las 100 repeticiones.

**Tabla 2.3.** Los mejores tres modelos identificados usando el método GVS, para el modelo 1 de 100 repeticiones.

Mejor	$\left( \begin{array}{ccc cc ccc} 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \end{array} \right)$ (0.1801)
Segundo mejor	$\left( \begin{array}{ccc cc ccc} 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \end{array} \right)$ (0.1379)
Tercer mejor	$\left( \begin{array}{ccc cc ccc} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \end{array} \right)$ (0.1156)

Los paréntesis muestran las correspondientes probabilidades a posteriori. Los tres modelos están asociados con  $d = 2$ .

En la identificación del número de regímenes en modelos TSARX siguiendo la metodología dada en la sección 1.5, se requieren muestras MCMC separadas e independientes para los modelos candidatos TSARX con  $l = 1, 2, 3, 4, 5$  regímenes, siendo  $l_0 = 5$  el número de regímenes máximo considerado. Se consideran valores para los hiperparámetros: aparte de los ya dados en la sección 1.3, para  $A_j$ ,  $B_j$ ,  $C_j$ , se tiene que  $\varphi_{ij} = 25$ ,  $\eta_{ij} = 1.50$  para todo  $i, j$ ,  $\chi_{uj} = 25$ ,  $\xi_{uj} = 1.50$  para todo  $u, j$ ,  $\psi_{vj} = 25$ ,  $\lambda_{vj} = 1.50$  para todo  $v, j$ ,  $i = 0, 1, 2$ ,  $u = 1, 2$ ,  $v = 1, 2, 3$ , y  $j = 1, 2, 3, 4, 5$ . Estas constantes son las mismas que eligieron Calderón y Nieto (2017), para implementar el método GVS en modelos TARX multivariados.

Adicionalmente,  $\nu_j$  y  $\lambda_j$ ,  $j = 1, 2, 3, 4, 5$ ;  $a$ ,  $b$ ,  $H$  y  $d_0$  son los mismos hiperparámetros que se asumieron para el modelo TARX usado en la prueba de estacionalidad. Ahora, los valores iniciales para cada uno de los parámetros fueron:  $A_j^{(0)} = 0.05k_j^*+1$ ,  $B_j^{(0)} = 0.05K_j^*$ ,  $C_j^{(0)} = 0.05q_j^*$ ,  $S_j^{(0)} = \mathbb{1}_{k_j^*+1}$ ,  $M_j^{(0)} = \mathbb{1}_{K_j^*}$ ,  $N_j^{(0)} = \mathbb{1}_{q_j^*}$  ( $\mathbb{1}$  es el vector de unos),  $[(h^{(j)})^2]^{(0)} = 0.20$ ,  $j = 1, 2, 3, 4, 5$ ; además,  $d^{(0)} = 3$ ,  $r^{(0)} = q_{0.50}\{Z_{t-d}\}$  para dos regímenes,  $\mathbf{r}^{(0)} = (r_1, r_2)' = (q_{0.25}\{Z_{t-d}\}, q_{0.75}\{Z_{t-d}\})'$  para tres regímenes,  $\mathbf{r}^{(0)} = (r_1, r_2, r_3)' = (q_{0.25}\{Z_{t-d}\}, q_{0.50}\{Z_{t-d}\}, q_{0.75}\{Z_{t-d}\})'$  para cuatro regímenes y  $\mathbf{r}^{(0)} = (r_1, r_2, r_3, r_4)' = (q_{0.20}\{Z_{t-d}\}, q_{0.40}\{Z_{t-d}\}, q_{0.60}\{Z_{t-d}\}, q_{0.80}\{Z_{t-d}\})'$  para cinco regímenes, siendo  $q_\alpha\{Z_{t-d}\}$  el cuantil  $\alpha$  de  $Z_{t-d}$ . Los tamaños de paso  $\sigma_{r1} = \sigma_{r2} = \sigma_{r3} = \sigma_{r4} = 0.025$  del Algoritmo 2 de la sección 1.3, para generar muestras de  $\mathbf{r}$ .

**Tabla 2.4.** Estimación de parámetros para el modelo 1 de 100 repeticiones.

Parámetro	Verdadero	Estimación	D.E.	I.C. 95 %	Cobertura
$a_0^{(1)}$	2.34	2.66	0.19	(2.29, 3.01)	96
$a_1^{(1)}$	0.50	0.50	0.01	(0.49, 0.51)	93
$b_1^{(1)}$	0.20	0.19	0.01	(0.17, 0.21)	97
$b_2^{(1)}$	0.10	0.10	0.01	(0.08, 0.11)	96
$c_1^{(1)}$	1.23	1.15	0.04	(1.06, 1.24)	95
$a_0^{(2)}$	-4.50	-4.32	1.40	(-7.08, -1.45)	98
$a_1^{(2)}$	0.60	0.57	0.02	( 0.52, 0.61)	94
$b_1^{(2)}$	0.10	0.09	0.03	( 0.03, 0.15)	97
$c_1^{(2)}$	-1.15	-1.15	0.18	(-1.49, -0.81)	97
$c_2^{(2)}$	3.30	3.23	0.27	( 2.72, 3.76)	95
$c_3^{(2)}$	-1.92	-1.92	0.20	(-2.32, -1.53)	94
$(h^{(1)})^2$	1	1.08	0.07	( 0.96, 1.24)	97
$(h^{(2)})^2$	16	16.14	1.01	(14.28, 18.22)	96
$r$	4.46	4.46	0.01	( 4.30, 4.72)	94
$d$	2	2(96)*			

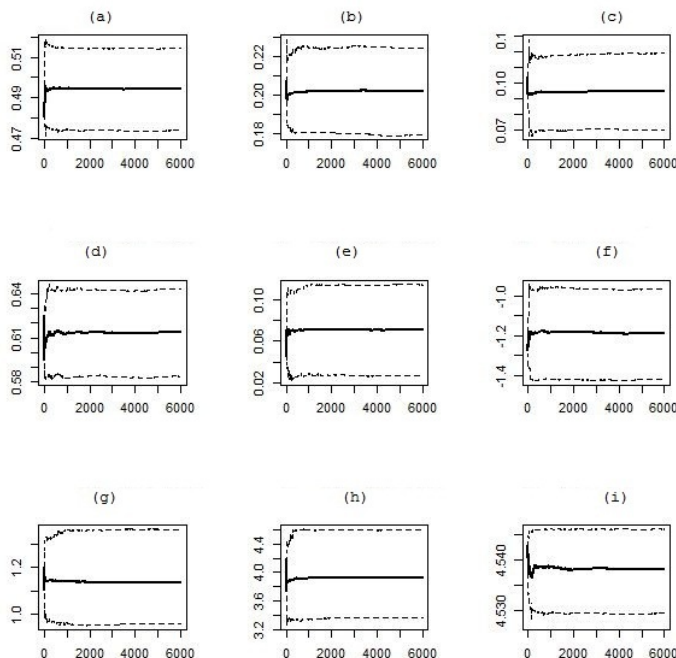
\* Número de veces (de un total de 100) donde el rezago 2 es elegido.

Los resultados de la estimación del número de regímenes son mostrados en la Tabla 2.2. Los números muestran la media de las probabilidades a posteriori dadas en (1.6), sobre las 100 repeticiones del modelo (2.1), y como otra referencia también, se calculó el promedio de las medidas DIC estándar dadas en (1.7). El modelo verdadero, un TSARX con dos regímenes obtuvo la probabilidad a posteriori promedio más alta y el criterio DIC promedio más pequeño. Igualmente, se calculó el número de veces que cada uno de los modelos TSARX con  $j = 1, 2, 3, 4, 5$  fue seleccionado de los 100 conjuntos de datos simulados. El modelo verdadero fue seleccionado 98 de las 100 repeticiones consideradas, usando el método de Congdon (2006) y 85 de las 100 repeticiones con el criterio DIC. Ahora, para estimar los órdenes autorregresivos del modelo, se emplea

el método presentado en la sección 1.3. Considerando  $l = 2$  y órdenes autorregresivos máximos  $k_1^* = k_2^* = 2$ ,  $K_1^* = K_2^* = 2$  y  $q_1^* = q_2^* = 3$  fijos, el modelo verdadero puede ser expresado como

$$\left( \begin{array}{c|c|c} S_1 & M_1 & N_1 \\ \hline S_2 & M_2 & N_2 \end{array} \right) = \left( \begin{array}{c|c|c} 1 & 1 & 0 \\ \hline 1 & 1 & 0 \end{array} \middle| \begin{array}{c|c} 1 & 1 \\ \hline 1 & 0 \end{array} \middle| \begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 1 & 1 & 1 \end{array} \right),$$

donde los unos significan que los coeficientes asociados a dichas variables latentes están presentes en el modelo. Se reporta en la Tabla 2.3, los tres mejores modelos seleccionados, basados sobre las probabilidades a posteriori de ocurrencia del evento más alta, y que son definidas en la sección 1.3. Note que el mejor modelo seleccionado usando el método GVS coincide con el modelo verdadero con una probabilidad a posteriori de 0.1801 y el segundo y tercer mejor modelo difieren al mejor modelo, solo en la presencia o no de algunos coeficientes de la entrada exógena en el primer régimen. Los tres modelos seleccionados están asociados con  $d = 2$ . En forma adicional se registra el número de veces de las 100 repeticiones que el modelo verdadero es declarado como el mejor modelo o el mejor o segundo mejor modelo, llegando al 88 % y el 95 % de los ensayos, respectivamente.

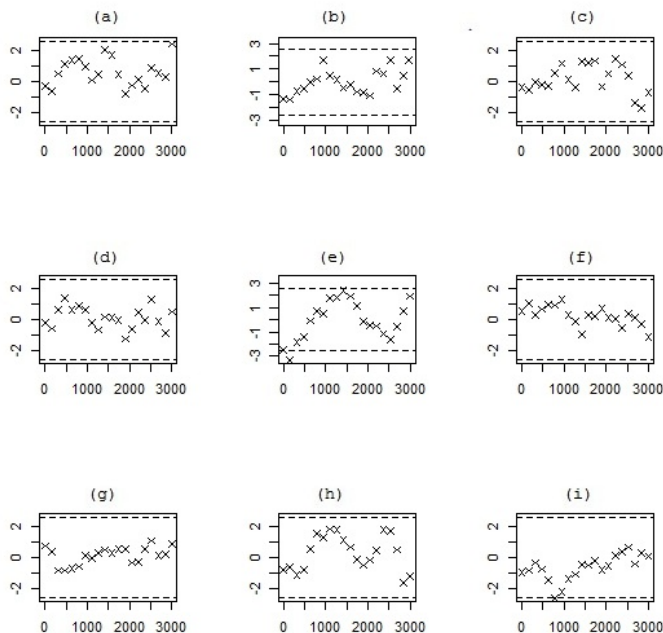


**Figura 2.5.** Gráficos cumuplot para: (a)  $a_1^{(1)}$ , (b)  $b_1^{(1)}$ , (c)  $b_2^{(1)}$ , (d)  $c_1^{(1)}$ , (e)  $a_0^{(2)}$ , (f)  $b_1^{(2)}$ , (g)  $(h^{(1)})^2$ , (h)  $(h^{(2)})^2$ , (i)  $r$ , modelo 1.

Se procede ahora, a estimar el resto de parámetros del modelo TSARX. La estimación de los parámetros autorregresivos en cada régimen, las ponderaciones de las varianzas de cada uno de los regímenes, el valor umbral y rezago de las 100 repeticiones son resumidas en la Tabla 2.4, mostrando cada uno de los valores verdaderos, más el promedio,

desviación estándar, los cuantiles 0.025 y 0.975 de las 100 medias a posteriori para cada uno de los parámetros. Además, se da el porcentaje de cobertura, que es el número de veces que el verdadero valor del parámetro cae en el intervalo creíble. Para  $d$ , la moda de las 100 modas a posteriori es dada.

Todos los promedios de las medias a posteriori son cercanas a los valores verdaderos y los



**Figura 2.6.** Gráficos *Z-score* de Geweke: (a)  $a_1^{(1)}$ , (b)  $b_1^{(1)}$ , (c)  $b_2^{(1)}$ , (d)  $c_1^{(1)}$ , (e)  $a_0^{(2)}$ , (f)  $b_1^{(2)}$ , (g)  $(h^{(1)})^2$ , (h)  $(h^{(2)})^2$ , (i)  $r$ , modelo 1.

intervalos de confianza del 95 % de las medias a posteriori cubren los valores verdaderos. Los porcentajes de cobertura son razonables y muy cercanas al 95 %. También, el rezago  $d = 2$  es correctamente identificado con el esquema muestral considerado.

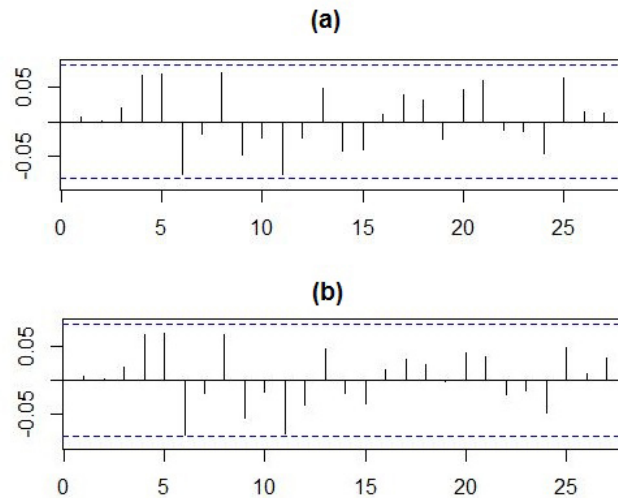
En orden de asegurar que para cada parámetro, las muestras MCMC convergen a la distribución a posteriori, se usan los gráficos cumuplot y *Z-score* de Geweke en una de las cien series de tiempo simuladas. Debido que existe una gran cantidad de parámetros en el modelo, solo se muestran algunas de ellas (Figuras 2.5 y 2.6), donde se observan que la convergencia es alcanzada en cada caso.

### 2.1.3. Diagnósticos del modelo.

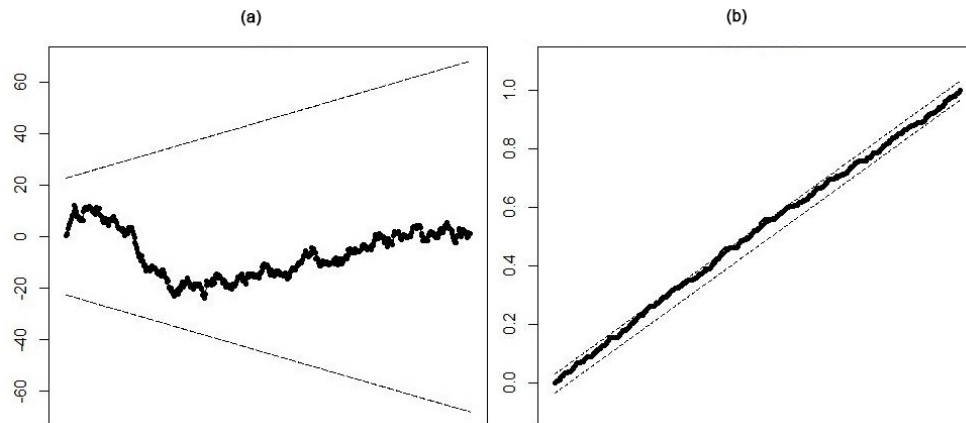
En el paso cuatro de la metodología propuesta, se conducen pruebas de residuales de adecuación del modelo. Para ello, en cada una de las 100 series de tiempo simuladas que siguen el modelo dado en (2.1), se estiman  $u_t$ ,  $t = 101, \dots, 600$ , (ver la ecuación (1.8)). Esto se hace estableciendo el valor de  $k$  en forma secuencial en 150, 200, 250,  $\dots$ , 600 y estimando  $u_t$  con  $(k - t) < 50$  para cada  $k$ ; luego, se calcula  $\hat{v}_t$  los cuales son los residuales

generalizados del modelo.

En la Tabla 2.5, se listan las probabilidades empíricas de no rechazo de las pruebas de Ljung y Box para los rezagos 1, 12 y 24 y de Jarque y Bera, obtenidas de las 100 series simulados de la ecuación (2.1). Se hace el cálculo de los residuales generalizados y se comprueba si ellos son independientes e idénticamente distribuidos como una normal y así determinar si el modelo considerado es el apropiado. Las probabilidades empíricas están entre 0.92 y 0.96, valores cercanos al valor nominal de 0.95 (nivel de significancia de 0.05).



**Figura 2.7.** (a) *FAC muestral*, (b) *FACP muestral*, residuales generalizados del modelo 1.



**Figura 2.8.** (a) *Gráfico CUSUM*, (b) *gráfico CUSUMSQ*, residuales generalizados del modelo 1.

**Tabla 2.5.** *Porcentajes de 100 repeticiones, para el cual el modelo ajustado es hallado ser adecuado para el modelo 1, usando residuales generalizados.*

$LB_1$	$LB_{12}$	$LB_{24}$	$JB$
0.96	0.96	0.92	0.94

*El nivel de significancia es 5 %.*

Seguido, tomando una de las cien series de tiempo simuladas, se realizan gráficos de autocorrelación muestral y autocorrelación parcial muestral de los residuales generalizados (véase Figura 2.7), donde se verifica no autocorrelación serial en dichos residuales. Además, se presentan gráficos CUSUM y CUSUMSQ, dados en la Figura 2.8, indicando que no hay evidencia de una especificación incorrecta del modelo y de presencia de heterocedasticidad marginal en  $\{X_t\}$ , respectivamente.

#### 2.1.4. Cálculo y evaluación de pronósticos.

Finalmente en el paso cinco, se ilustra el procedimiento de pronósticos desarrollado en la sección 1.7. El periodo muestral considerado fue 1-587, y el horizonte de pronóstico fue 588-600. En la Tabla 2.6, se muestran los valores simulados, los pronósticos puntuales, desviaciones estándar e intervalos creíbles del 95 % promedios para  $X_t$  y  $Z_t$ , de las 100 series de tiempo simuladas del modelo (2.1), con  $h = 1, 2, 3, 11, 12$  y  $13$ ; en la última columna de la tabla, aparece el porcentaje de cobertura.

Para las variables  $X_t$  y  $Z_t$ , los valores de los pronósticos son muy cercanos a los valores simulados; también, puede verse que todos los valores simulados caen en los correspondientes intervalos creíbles del 95 %, las desviaciones estándar se incrementan a medida que aumenta el horizonte de pronóstico, y el porcentaje de cobertura son cercanos al 95 %.

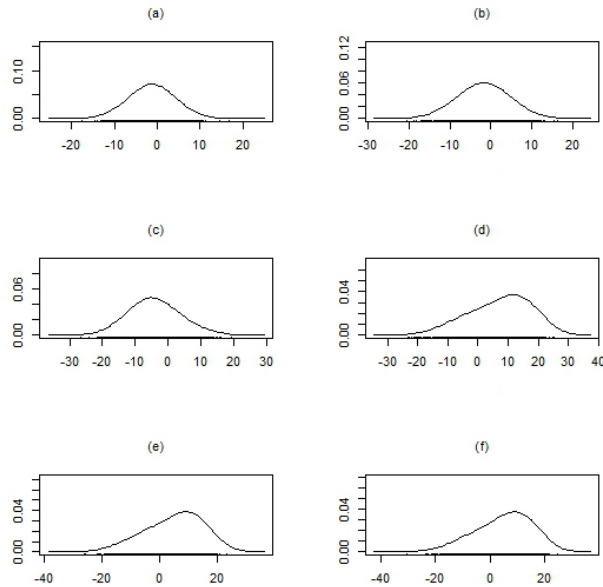
Además, se muestran las distribuciones predictivas para  $X_t$  y  $Z_t$ , (véase Figuras 2.9 y 2.10) estimadas usando métodos kernel no paramétricos estándar para una de las cien series simuladas descritas anteriormente. Las distribuciones predictivas de las variables de interés y de umbrales, se caracterizan por ser simétricas y unimodales, excepto para  $h = 11, 12$  y  $13$  de  $X_t$ , que muestran asimetrías en forma negativa.

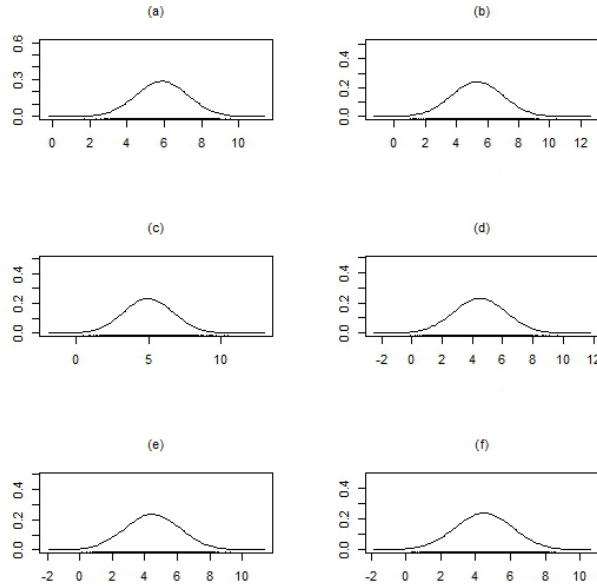
**Tabla 2.6.** Pronósticos para las variables  $X_t$  y  $Z_t$  del modelo 1 con 100 repeticiones.

Horizonte (h)	$X_{T+h}$	$\widehat{X}_{T+h}$	D.E.	I.C. 95 %	Cobertura
1	6.49	5.80	2.88	( 0.10, 11.30)	94
2	5.76	5.74	3.48	( -1.12, 12.51)	96
3	7.09	6.36	5.39	( -4.62, 15.19)	95
11	6.72	6.35	8.54	(-11.96, 17.80)	100
12	7.48	6.41	8.56	(-11.89, 17.89)	100
13	6.12	6.34	8.60	(-12.05, 17.97)	100

Horizonte (h)	$Z_{T+h}$	$\widehat{Z}_{T+h}$	D.E.	I.C. 95 %	Cobertura
1	4.25	4.36	0.99	( 2.39, 6.32)	93
2	4.39	4.41	1.16	( 2.13, 6.69)	95
3	4.14	4.44	1.21	( 2.06, 6.83)	99
11	4.59	4.49	1.24	( 2.05, 6.94)	98
12	4.72	4.49	1.25	( 2.05, 6.93)	100
13	4.56	4.50	1.25	( 2.04, 6.95)	100

**Figura 2.9.** Distribuciones predictivas para la variable  $X_t$  modelo 1, (a)  $h=1$ , (b)  $h=2$ , (c)  $h=3$ , (d)  $h=11$ , (e)  $h=12$ , (f)  $h=13$ , pasos adelante.



**Figura 2.10.** *Distribuciones predictivas para la variable  $Z_t$  modelo 1, (a)  $h=1$ , (b)  $h=2$ , (c)  $h=3$ , (d)  $h=11$ , (e)  $h=12$ , (f)  $h=13$ , pasos adelante.*

## 2.2. Modelo 2.

Ahora se asume que,  $\{X_t\}$  sigue el modelo TSARX(3; 1, 2, 1; 1, 1, 2; 2, 1, 0)<sub>4</sub> dado por

$$X_t = \begin{cases} 1.32 - 0.20X_{t-1} + 0.60X_{t-4} + 0.12X_{t-5} + 2.32Z_{t-1} \\ -2.00Z_{t-2} + 3\epsilon_t, & \text{si } Z_{t-1} \leq r_1, \\ 1.92 + 0.20X_{t-1} + 0.30X_{t-2} + 0.50X_{t-4} - 0.10X_{t-5} \\ -0.15X_{t-6} - 1.50Z_{t-1} + \epsilon_t, & \text{si } r_1 < Z_{t-1} \leq r_2, \\ -2.34 + 0.50X_{t-1} + 0.20X_{t-4} - 0.10X_{t-5} + 0.10X_{t-8} \\ -0.05X_{t-9} + 2\epsilon_t, & \text{si } Z_{t-1} > r_2, \end{cases} \quad (2.2)$$

donde  $\{Z_t\}$  es un modelo AR(5) multiplicativo,  $Z_t = 1.80 + 0.60Z_{t-1} + 0.50Z_{t-4} - 0.30Z_{t-5} + 2a_t$ , con  $\{a_t\} \sim \text{RBG}(0,1)$  mutuamente independientes de  $\{\epsilon_t\}$ ,  $r_1 = 8.22$ ,  $r_2 = 10.77$  son respectivamente, el primer y tercer cuartil de  $Z_{t-1}$ . Es bien conocido que  $\{Z_t\}$  es una cadena de Markov homogénea de orden cinco.

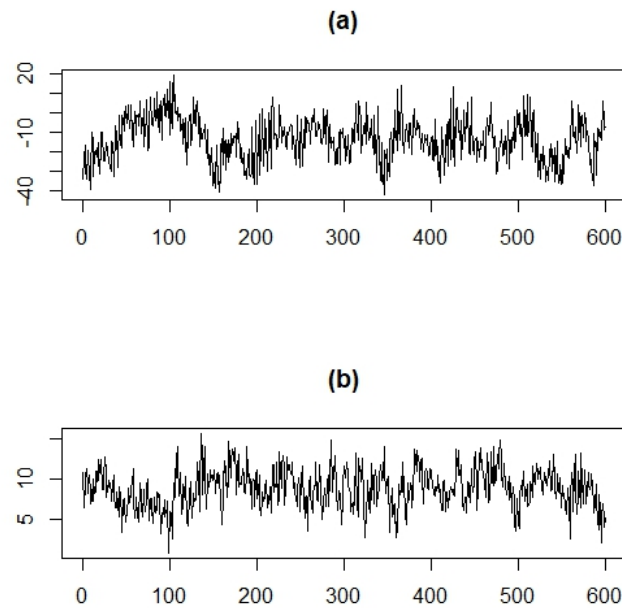
### 2.2.1. Análisis exploratorio de los datos.

En el paso uno, el análisis exploratorio con la aplicación de raíces unitarias y gráficos a las series de tiempo  $\{X_t\}$  y  $\{Z_t\}$  dadas en (2.2), son realizadas. En la Tabla 2.7, se

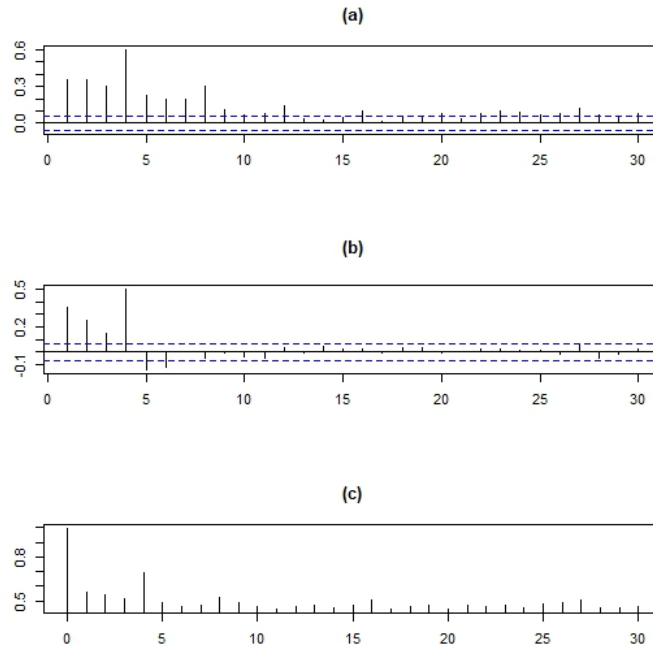
presentan las pruebas DFA y OCSB para  $X_t$  y  $Z_t$  del modelo 2. En los dos casos, no hay evidencia de presencia de raíces unitarias y raíces unitarias estacionales en las series de tiempo simuladas, considerando un nivel de significancia del 5%. En la Figura 2.11 se exhiben las dos series simuladas, con 600 observaciones. En las gráficas de las series se detalla cierto comportamiento estable y no hay evidencia de un comportamiento estacional, por lo cual se requieren de otras herramientas para asegurar o descartar la presencia de estacionalidad en las dos series. Para ello, en la Figura 2.12, para  $\{X_t\}$  se logra identificar autocorrelaciones significativas en los primeros dos rezagos, seguido de los rezagos cuatro y ocho; mientras, para  $\{Z_t\}$  (véase Figura 2.13) las autocorrelaciones son significativas hasta el rezago cinco. El IMP para las dos series muestran mayor significancia en las dependencias mutuas entre  $W_i$  y  $W_{i-4}$ , que lo presentado en el modelo dado en (2.1), para el rezago 12. Los anteriores hallazgos son los esperados para  $X_t$  y  $Z_t$ .

**Tabla 2.7.** Pruebas de raíces unitarias y raíces unitarias estacionales para el modelo 2.

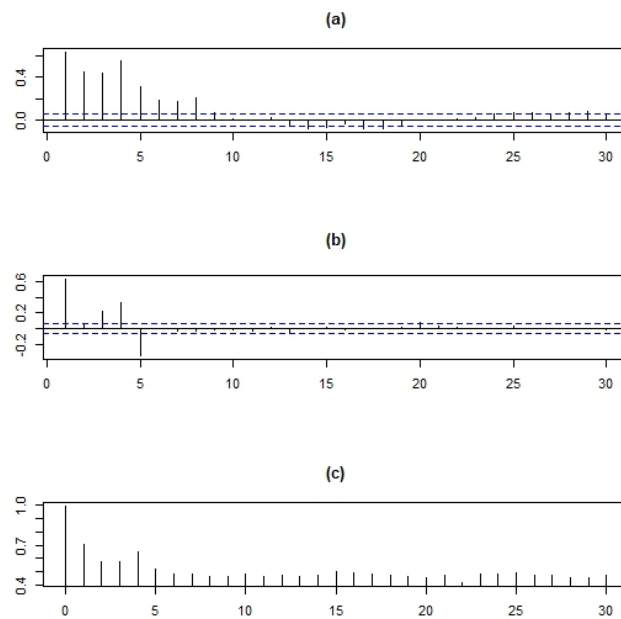
	$X_t$	$Z_t$
Estadística DFA	-3.824	-4.385
Valor $p$	0.017	0.010
Estadística OCSB	-9.926	-13.150
Valor crítico del 5%	-1.893	-1.893



**Figura 2.11.** (a)  $X_t$  variable de interés, (b)  $Z_t$  variable de umbrales, modelo 2.



**Figura 2.12.** Para  $X_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 2.



**Figura 2.13.** Para  $Z_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, modelo 2.

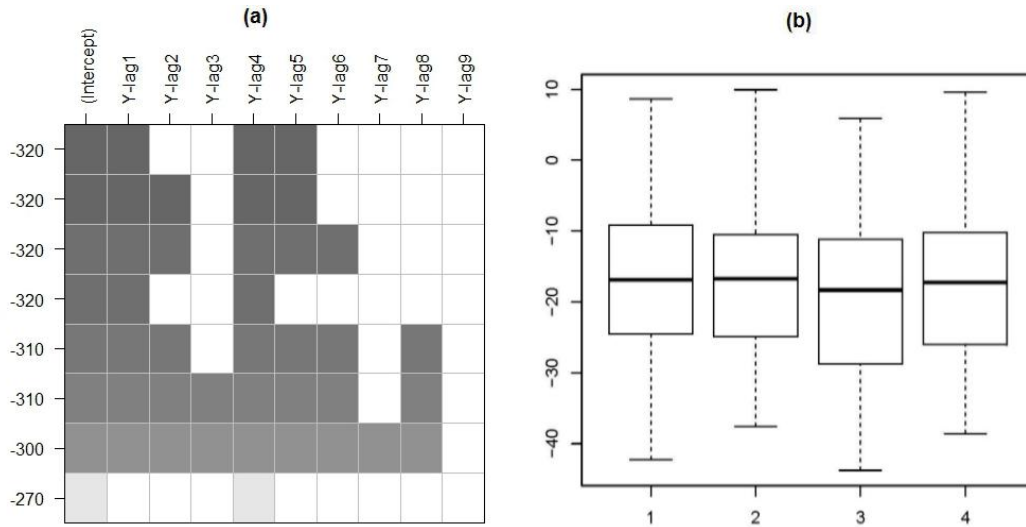


Figura 2.14. Para  $X_t$ , (a) AR multiplicativo lineal con BIC mínimo, (b) diagramas de caja por trimestres, modelo 2.

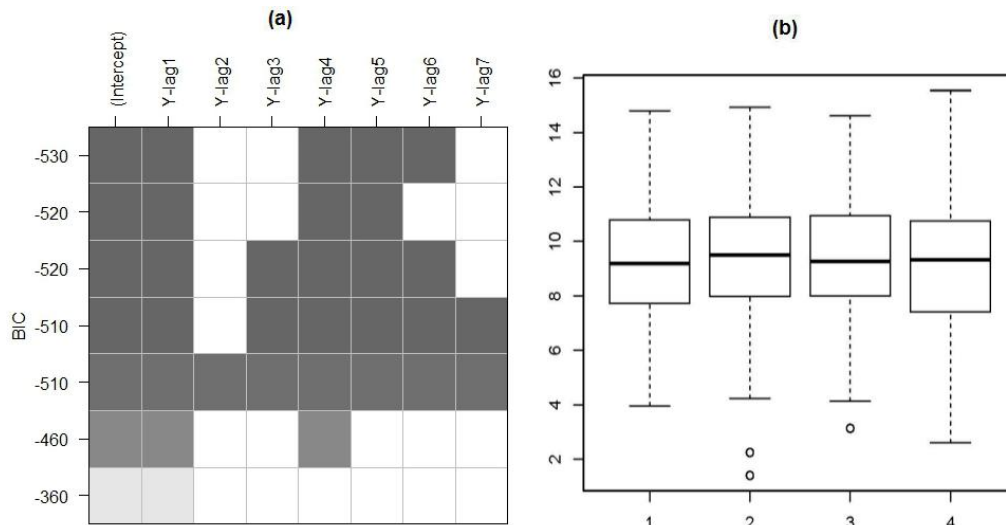


Figura 2.15. Para  $Z_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por trimestres, modelo 2.

Un análisis complementario se da en la Figura 2.14 para  $\{X_t\}$ , y la Figura 2.15 para  $\{Z_t\}$ . En las dos series se pueden apreciar un ajuste de un modelo de series de tiempo lineal mediante un posible SAR(5) multiplicativo usando el BIC mínimo; además, en las gráficas de diagramas de caja presentan casi medias iguales entre trimestres, pero distinta variabilidad.

También se puede ver, correspondiente al modelo dado en la ecuación (2.2), que algunas raíces de los polinomios característicos  $\phi_1(w) = 1 + 0.20w - 0.60w^4 - 0.12w^5$  para el régimen uno,  $\phi_2(w) = 1 - 0.20w - 0.30w^2 - 0.50w^4 + 0.10w^5 + 0.15w^6$  para el régimen

dos, y  $\phi_3(w) = 1 - 0.50w - 0.20w^4 + 0.10w^5 - 0.10w^8 + 0.05w^9$  para el régimen tres, le corresponden un periodo de longitud 2 y 4, lo cual indica presencia de ciclos estacionales en  $\{X_t\}$ .

### 2.2.2. Resultados de la estimación.

En el paso dos y tres de la metodología propuesta, se usan 100 repeticiones con tamaño de muestra  $T = 600$  del modelo (2.2). Para cada conjunto de datos, se corre 12.000 iteraciones MCMC y un periodo de calentamiento de 6.000 iteraciones. Aquí se evalúa el funcionamiento de la prueba Bayesiana de estacionalidad propuesta en este trabajo, y también, la identificación del modelo y estimación de parámetros del modelo dado en (2.2). En la prueba Bayesiana de estacionalidad multiplicativa por regímenes se requiere estimar parámetros de un modelo TARX con tres regímenes, con  $k_1^* = k_2^* = k_3^* = 2$ ,  $K_1^* = K_2^* = K_3^* = 2$ ,  $q_1^* = q_2^* = q_3^* = 2$  órdenes autorregresivos máximos y  $s = 4$ .

Los hiperparámetros y valores iniciales para cada uno de los parámetros, son los mismos que se consideraron para el modelo 1 de la sección 2.1, salvo que se extienden a tres regímenes. Tal es el caso, para  $\mathbf{r}^{(0)} = (r_1, r_2)' = (q_{0.25}\{Z_{t-d}\}, q_{0.75}\{Z_{t-d}\})'$  y los tamaños de paso  $\sigma_{r1} = \sigma_{r2} = 0.025$  del Algoritmo 2, para generar muestras de  $\mathbf{r}$ . El número de aciertos de la prueba estadística Bayesiana es como mínimo 94 de las 100 repeticiones, en los tres regímenes. Si  $2 \ln(FB_j) > 6$ ,  $j = 1, 2, 3$  se considera un acierto de la prueba en cada régimen (véase sección 1.4). Para la identificación del número de regímenes, se consideran modelos candidatos TSARX con  $l = 1, 2, 3, 4, 5$ . Se tienen en cuenta los hiperparámetros y valores iniciales para cada uno de los modelos candidatos dados en la sección 2.1.

**Tabla 2.8.** Probabilidades medias a posteriori y DIC promedio para selección del número de regímenes para el modelo 2 (100 repeticiones).

	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
Prob. media a posteriori	0.00000	0.00037	0.983	0.0161	0.00053
DIC promedio	4105.37	3430.56	2225.89	2576.93	2636.60

En la Tabla 2.8, se muestran las probabilidades a posteriori promedio y DIC promedio de los TSARX candidatos. Se puede observar que el número de regímenes asociado con la probabilidad a posteriori más grande es tres, coincidiendo dicha elección con el valor DIC promedio más pequeño. En forma adicional, se calculó el número de veces que cada modelo candidato fue elegido de las 100 repeticiones. El modelo verdadero fue seleccionado 93 de las 100 series simuladas, usando el método de Congdon (2006) y 82 de las 100 series con el DIC. Lo cual indica que el método de Congdon es eficiente en la identificación del número de regímenes en los dos ejemplos simulados bajo objeto de estudio.

Seguido, se estimó los órdenes autorregresivos del modelo TSARX empleando el método GVS descrito en la sección 1.3. Se consideró  $l = 3$  y  $k_1^* = k_2^* = k_3^* = 2$ ,  $K_1^* = K_2^* = K_3^* = 2$  y  $q_1^* = q_2^* = q_3^* = 2$  fijos, luego el modelo verdadero se puede expresar en forma matricial

**Tabla 2.9.** *Los mejores tres modelos identificados usando el método GVS, para el modelo 2 de 100 repeticiones.*

---

Mejor	$\left( \begin{array}{ccc ccc} 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right)$
	(0.1456)
Segundo mejor	$\left( \begin{array}{ccc ccc} 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right)$
	(0.0973)
Tercer mejor	$\left( \begin{array}{ccc ccc} 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right)$
	(0.0651)

---

*Los paréntesis muestran las correspondientes probabilidades a posteriori. Los tres modelos están asociados con  $d = 1$ .*

como

$$\left( \begin{array}{c|c|c} S_1 & M_1 & N_1 \\ S_2 & M_2 & N_2 \\ S_3 & M_3 & N_3 \end{array} \right) = \left( \begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right).$$

Se presentan en la Tabla 2.9, los tres mejores modelos seleccionados, según las probabilidades a posteriori de ocurrencia más alta. Se tiene que el segundo modelo elegido por el método GVS es el modelo verdadero, con una probabilidad a posteriori de 0.0973 y el primer y tercer mejor modelo difieren del modelo verdadero, solo en los coeficientes de los interceptos del primer y segundo régimen del modelo. Los tres modelos seleccionados están asociados con  $d = 1$ . También, se llega a que 80 de las 100 repeticiones, el modelo verdadero es declarado como el mejor modelo y 96 de las 100 repeticiones, el modelo verdadero se elige como el mejor o segundo mejor modelo. El método GVS posee buenos resultados, en las dos series de tiempo simuladas dadas.

A continuación, se estiman el resto de parámetros del modelo TSARX simulado. Estimaciones Bayesianas son presentadas en la Tabla 2.10, la cual contiene los promedios de los valores verdaderos de los parámetros, medias a posteriori, errores estándar, intervalos creíbles del 95 % junto con el porcentaje de cobertura, de las 100 repeticiones del modelo dado en (2.2). Las estimaciones puntuales no son significativamente diferentes de sus

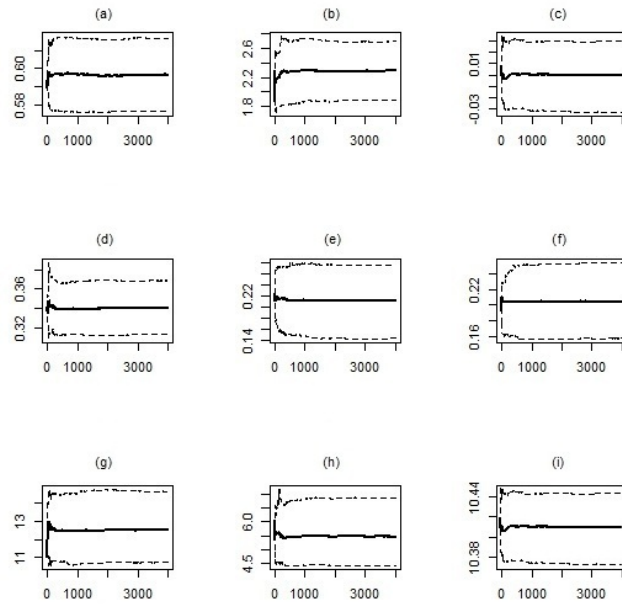
**Tabla 2.10.** *Estimación de parámetros para el modelo 2 de 100 repeticiones.*

Parámetro	Verdadero	Estimación	D.E.	I.C. 95 %	Cobertura
$a_0^{(1)}$	1.32	1.25	1.08	(-0.87, 3.36)	95
$a_1^{(1)}$	-0.20	-0.20	0.02	(-0.24, -0.15)	96
$b_1^{(1)}$	0.60	0.60	0.01	( 0.57, 0.62)	98
$c_1^{(1)}$	2.32	2.31	0.17	( 1.97, 2.64)	95
$c_2^{(1)}$	-2.00	-1.98	0.12	(-2.22, -1.74)	99
$a_0^{(2)}$	1.92	1.85	0.72	( 0.42, 3.28)	96
$a_1^{(2)}$	0.20	0.20	0.01	( 0.18, 0.22)	99
$a_2^{(2)}$	0.30	0.30	0.01	( 0.28, 0.32)	98
$b_1^{(2)}$	0.50	0.50	0.01	( 0.48, 0.52)	99
$c_1^{(2)}$	-1.50	-1.49	0.08	(-1.65, -1.33)	98
$a_0^{(3)}$	-2.34	-1.73	1.59	(-4.80, 1.40)	97
$a_1^{(3)}$	0.50	0.50	0.02	( 0.47, 0.54)	98
$b_1^{(3)}$	0.20	0.20	0.02	( 0.16, 0.25)	96
$b_2^{(3)}$	0.10	0.10	0.02	( 0.06, 0.14)	97
$(h^{(1)})^2$	9	9.41	0.51	(8.08, 10.01)	98
$(h^{(2)})^2$	1	1.22	0.11	(0.21, 1.90)	95
$(h^{(3)})^2$	4	3.95	1.36	(3.32, 4.75)	96
$r_1$	8.22	7.53	0.01	( 7.50, 8.94)	98
$r_2$	10.77	10.46	0.01	(10.23, 11.45)	97
$d$	1	1(100)*			

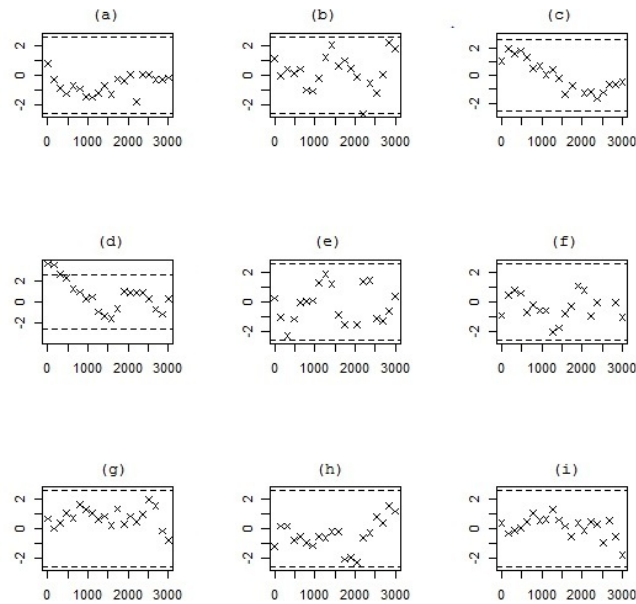
\* Número de veces (de un total de 100) donde el rezago 1 es elegido.

valores verdaderos, cada intervalo creíble del 95% contiene el valor verdadero y los porcentajes de cobertura parecen bastante razonables, en su mayoría cerca al 95%. También, el rezago  $d = 1$  es correctamente identificado de todas las 100 repeticiones del modelo. En resumen, las estimaciones de los parámetros parecen trabajar bien para los modelos TSARX usando la metodología propuesta.

En seguida, se presentan gráficos cumuplot y  $Z$ -score de Geweke, de algunos de los parámetros del modelo dado en (2.2), (véase las Figuras 2.16 y 2.17), con el fin de verificar convergencia a las respectivas distribuciones a posteriori. Se observa que el esquema muestral propuesto genera iteraciones MCMC que convergen rápidamente a las distribuciones a posteriori deseadas.



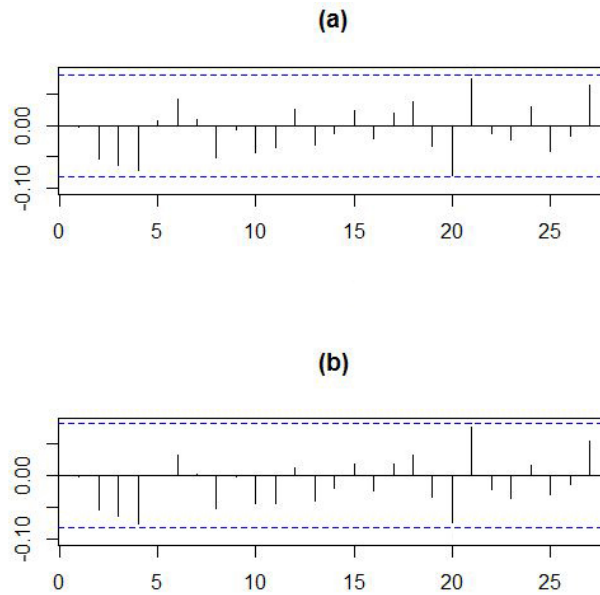
**Figura 2.16.** Gráficos cumuplot para: (a)  $b_1^{(1)}$ , (b)  $c_1^{(1)}$ , (c)  $a_2^{(2)}$ , (d)  $b_1^{(2)}$ , (e)  $b_1^{(3)}$ , (f)  $b_2^{(3)}$ , (g)  $(h^{(2)})^2$ , (h)  $(h^{(3)})^2$ , (i)  $r_1$ , modelo 2.



**Figura 2.17.** Gráficos Z-score de Geweke: (a)  $b_1^{(1)}$ , (b)  $c_1^{(1)}$ , (c)  $a_2^{(2)}$ , (d)  $b_1^{(2)}$ , (e)  $b_1^{(3)}$ , (f)  $b_2^{(3)}$ , (g)  $(h^{(2)})^2$ , (h)  $(h^{(3)})^2$ , (i)  $r_1$ , modelo 2.

### 2.2.3. Diagnósticos del modelo.

En el paso cuatro, se verifica la adecuación del modelo a través de los residuales generalizados. Para cada una de las 100 series de tiempo simuladas del modelo dado en la ecuación (2.2), se estiman  $v_t = \Phi^{-1}(u_t)$ ,  $t = 101, \dots, 600$  y siguiendo la metodología dada en la sección 1.6. En la Tabla 2.11, se dan las probabilidades empíricas de no rechazar normalidad (prueba de Jarque y Bera) y no autocorrelación (prueba de Ljung y Box) para los rezagos 1, 4 y 8, logradas de las 100 repeticiones. Las probabilidades empíricas mencionadas están entre 0.93 y 0.97, las cuales son valores cercanos al valor nominal de 0.95, indicando que las pruebas estadísticas aplicadas a los residuales generalizados son útiles en la validación de los modelos TSARX.



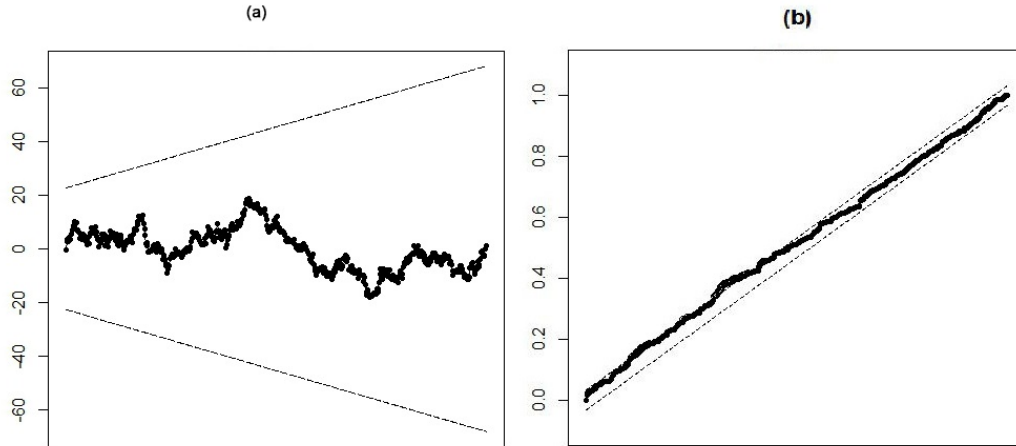
**Figura 2.18.** (a) *FAC muestral*, (b) *FACP muestral*, residuales generalizados del modelo 2.

Ahora, usando solo una de las series de tiempo simuladas del modelo dado en (2.2), se procede a determinar independencia, especificación correcta del modelo y no presencia de heterocedásticidad marginal en el modelo, mediante el análisis de las Figuras 2.18 y 2.19, las cuales corroboran un adecuado ajuste del modelo.

**Tabla 2.11.** *Porcentajes de 100 repeticiones, para el cual el modelo ajustado es hallado ser adecuado para el modelo 2, usando residuales generalizados.*

$LB_1$	$LB_4$	$LB_8$	$JB$
0.95	0.96	0.97	0.93

*El nivel de significancia es 5%.*



**Figura 2.19.** (a) *Gráfico CUSUM*, (b) *gráfico CUSUMSQ*,  
residuales generalizados del modelo 2.

#### 2.2.4. Cálculo y evaluación de pronósticos.

Por último en el paso cinco, se usa el método de predicción Bayesiana dado en la sección 1.7, para generar pronósticos de  $X_t$  y  $Z_t$  de las distribuciones predictivas, para unos valores  $h$  determinados. El periodo muestral fue 1 – 594, y el horizonte de pronóstico fue 595 – 600. En la Tabla 2.10, se listan los promedios de los valores simulados, pronósticos puntuales, desviaciones estándar, intervalos creíbles del 95% y porcentajes de cobertura de las 100 repeticiones de las series de tiempo  $\{X_t\}$  y  $\{Z_t\}$  dadas en (2.2), con  $h = 1, 2, 3, 4, 5, 6$ . Se puede concluir que, los pronósticos puntuales son muy cercanos a los valores simulados, las desviaciones estándar se incrementan a medida que aumentan el valor de  $h$  y los porcentajes de cobertura son cercanos al 95%.

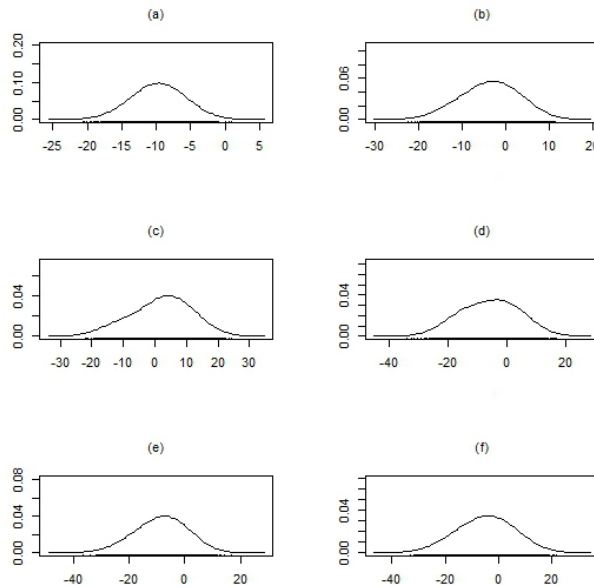
Adicionalmente, se calcula las distribuciones predictivas de  $X_t$  y  $Z_t$ , de una serie del conjunto de cien series simuladas, que se presentan en las Figuras 2.20 y 2.21, respectivamente. Se puede notar que, las distribuciones predictivas en ambas series, son simétricas y unimodales para todos los  $h$  tomados. Lo anterior, nos lleva a decir, que la metodología propuesta posee buenos resultados para realizar pronósticos en los modelos TSARX.

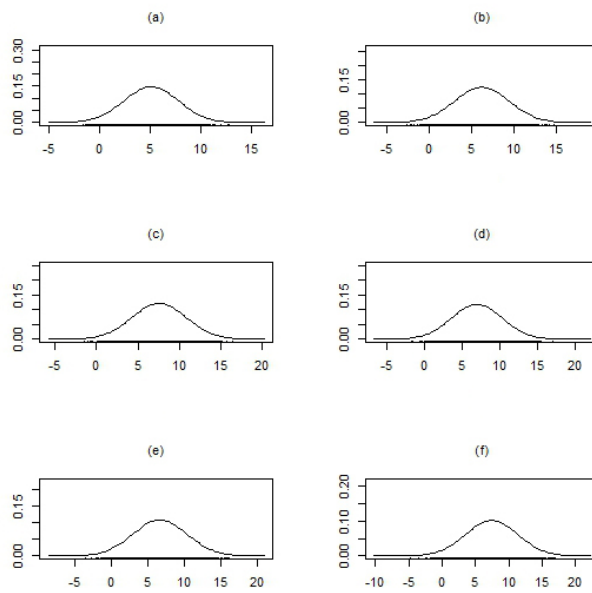
**Tabla 2.12.** *Pronósticos para las variables  $X_t$  y  $Z_t$  del modelo 2 con 100 repeticiones.*

Horizonte (h)	$X_{T+h}$	$\widehat{X}_{T+h}$	D.E.	I.C. 95 %	Cobertura
1	-10.08	-9.98	2.74	(-15.34, -4.47)	98
2	-10.49	-11.58	6.80	(-22.96, 0.12)	94
3	-12.02	-11.83	7.64	(-24.68, 1.42)	95
4	-10.27	-11.03	8.08	(-25.59, 3.08)	98
5	-10.50	-10.83	8.49	(-26.56, 3.97)	93
6	-11.23	-11.95	9.04	(-30.09, 4.25)	100

Horizonte (h)	$Z_{T+h}$	$\widehat{Z}_{T+h}$	D.E.	I.C. 95 %	Cobertura
1	9.03	9.29	2.00	( 5.37, 13.21)	93
2	8.56	8.95	2.33	( 4.38, 13.52)	95
3	8.34	8.88	2.43	( 4.09, 13.65)	99
4	8.66	9.13	2.47	( 4.27, 13.97)	98
5	8.81	9.17	2.77	( 3.72, 14.60)	100
6	8.61	8.99	2.87	( 3.35, 14.63)	100

**Figura 2.20.** *Distribuciones predictivas para la variable  $X_t$  modelo 2, (a)  $h=1$ , (b)  $h=2$ , (c)  $h=3$ , (d)  $h=4$ , (e)  $h=5$ , (f)  $h=6$ , pasos adelante.*



**Figura 2.21.** *Distribuciones predictivas para la variable  $Z_t$  modelo 2, (a)  $h=1$ , (b)  $h=2$ , (c)  $h=3$ , (d)  $h=4$ , (e)  $h=5$ , (f)  $h=6$ , pasos adelante.*

# CAPÍTULO 3

---

---

## Una aplicación en economía

---

---

En este capítulo se aplica la metodología propuesta a series de tiempo económicas mensuales.

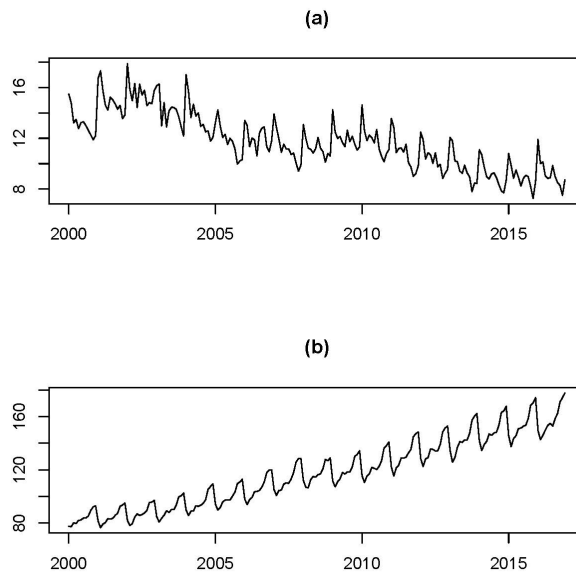
### 3.1. Análisis exploratorio de los datos.

Las series de tiempo iniciales son la tasa de desempleo total mensual Colombiana,  $TD_t$ , y el índice de seguimiento económico mensual Colombiano con base 2005,  $ISE_t$ . La tasa de desempleo es la relación porcentual entre el número de personas que están buscando trabajo y el número de personas que integran la población económicamente activa (DIMPE, 2016); además, la tasa de desempleo es un indicador muy importante para el análisis macroeconómico, ya que tiene la característica de indicar las principales señales del estado de la economía (véase, Leamer, 2009).

Una característica principal de la tasa de desempleo es la dinámica contracíclica asimétrica en las fases del ciclo económico; es decir, el comportamiento dinámico de la tasa de desempleo es generalmente caracterizado, por etapas de incremento durante recesiones y disminuciones lentas, durante las expansiones (véase, Skalin y Teräsvirta, 2001). De otro lado, el  $ISE_t$  es un índice sintético cuyo fin es proporcionar una medida de la evolución de la actividad real de la economía Colombiana en el corto plazo, el cual se ajusta a la metodología utilizada en las cuentas nacionales trimestrales, compuesto por un conjunto heterogéneo de indicadores mensuales representativos de las actividades económicas (DSCN, 2016). En la Figura 3.1 (a) y (b), se dan las gráficas de las series de tiempo  $TD_t$  e  $ISE_t$ , respectivamente; donde se puede ver, que la relación dinámica inversa entre las dos variables es clara y se evidencia cierto comportamiento estacional en las dos series de tiempo.

Múltiples trabajos en estudios empíricos se han orientado al análisis de la dinámica contracíclica asimétrica y no linealidad de umbrales de la tasa de desempleo; entre los que se cuentan, Rothman (1998) quien comparó el desempeño en pronósticos de

algunos modelos de series de tiempo lineales con modelos SETAR y STAR, ajustadas a la tasa de desempleo de Estados Unidos, superando en precisión los modelos no lineales a los lineales. Koop y Potter (1999) usaron un modelo SETAR con dos regímenes para ajustar la tasa de desempleo de Estados Unidos, a partir del cual encontraron mejores resultados en pronósticos en comparación con un modelo ARMA. Skalin y Teräsvirta (2001) modelaron la asimetría observada en la tasa de desempleo de varios países Europeos mediante un STAR; mientras que, Vivas (2011) estudia la dinámica del crecimiento anual de la tasa de desempleo trimestral Colombiano y el índice de confianza industrial trimestral para Colombia, en el periodo de 1985 hasta 2011 con un modelo TAR, hallando que el crecimiento anual de la tasa de desempleo trimestral se caracteriza por seguir una dinámica inversa a la evolución de la confianza industrial.



**Figura 3.1.** (a)  $TD_t$ , tasa de desempleo total mensual Colombiana, (b)  $ISE_t$ , índice de seguimiento económico mensual Colombiano, con base 2005.

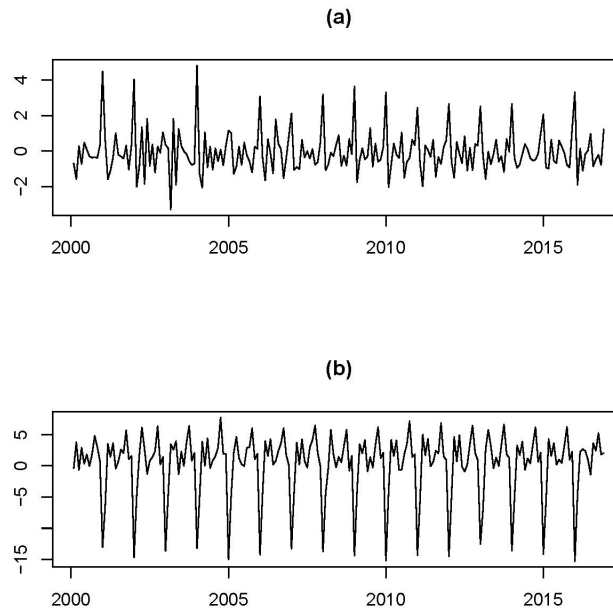
En nuestro análisis, nos enfocamos en el crecimiento de la tasa de desempleo mensual Colombiana,  $X_t = TD_t - TD_{t-1}$  como variable de interés y el crecimiento mensual de la transformación logarítmica del índice de seguimiento económico mensual Colombiano,  $Z_t = (\ln(ISE_t) - \ln(ISE_{t-1})) \times 100$  (en porcentaje) como variable de umbrales. El periodo muestral es desde Febrero 2001 hasta Diciembre 2016 (191 observaciones). Los datos son suministrados por el DANE, entidad encargada del estudio y difusión de las estadísticas oficiales de Colombia.

En la Tabla 3.1, se muestran resultados de las pruebas de raíces unitarias DFA y raíces unitarias estacionales OCSB, para las variables de interés y de umbrales de los datos empíricos. Solo la hipótesis nula de raíz unitaria estacional no se rechaza para  $Z_t$  con un valor crítico al nivel del 5%. Las series de tiempo  $\{X_t\}$  y  $\{Z_t\}$  son dadas en la Figura 3.2 (a) y (b), respectivamente.  $\{X_t\}$  presenta picos positivos anuales (eneros); mientras,

**Tabla 3.1.** Pruebas de raíces unitarias y raíces unitarias estacionales para los datos empíricos.

	$X_t$	$Z_t$
Estadística DFA	-8.202	-8.000
Valor $p$	0.010	0.010
Estadística OCSB	-3.562	-1.449
Valor crítico del 5 %	-1.803	-1.803

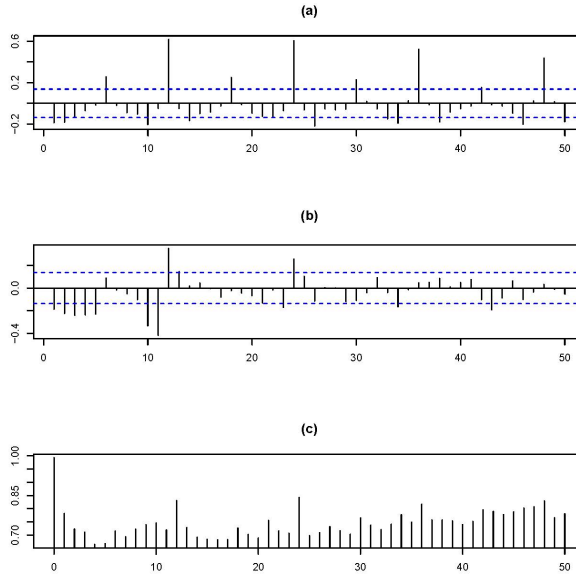
$\{Z_t\}$  presenta picos negativos muy pronunciados anuales (eneros). Esto indica cierto comportamiento estacional en ambas series, heredadas de las series de tiempo originales. También puede observarse que las series no tienen tendencias, pero sí ciclos estacionales. Por lo tanto, las series no son estacionarias.



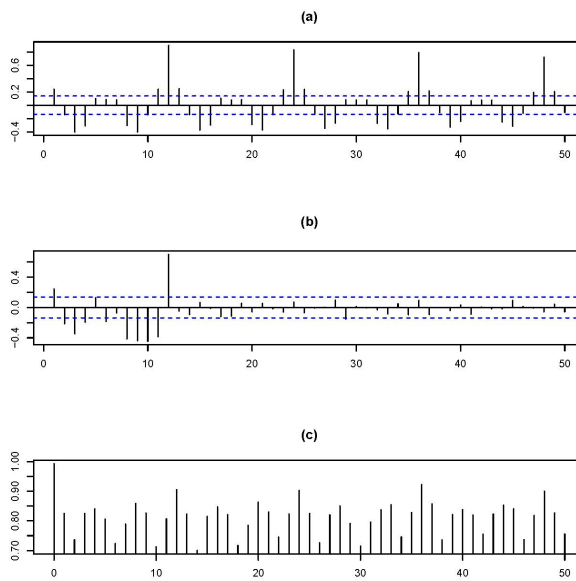
**Figura 3.2.** (a)  $X_t$ , crecimiento de la tasa de desempleo total mensual Colombiana, (b)  $Z_t$ , crecimiento mensual de la transformación logaritmo del  $ISE_t$ .

En las Figuras 3.3 y 3.4, se muestran autocorrelogramas y el IMP para  $\{X_t\}$  y  $\{Z_t\}$ , respectivamente. En los dos casos, las FAC muestral, muestran rezagos estacionales que no terminan de descender y las FACP e IMP se identifican correlaciones significativas en los primeros rezagos y en los rezagos 12, 24 ó 36. Seguido, en la Figura 3.5 para  $\{X_t\}$  y Figura 3.6 para  $\{Z_t\}$ , se observan un ajuste de un posible modelo AR(26) multiplicativo, en las dos series de tiempo, siguiendo el criterio BIC mínimo. En cuanto a los diagramas de caja, para  $\{X_t\}$  presenta una distribución con valores más altos en los meses de enero, y una distribución con valores más bajos en los meses de marzo; en cambio, para  $\{Z_t\}$  se observa una distribución con valores más bajos en los meses de enero, y una distribución

con valores más altos en los meses de octubre. En resumen, el análisis exploratorio realizado a las series de tiempo, llevan a deducir que  $\{X_t\}$  requiere un modelo que capture las estacionalidad detectada.

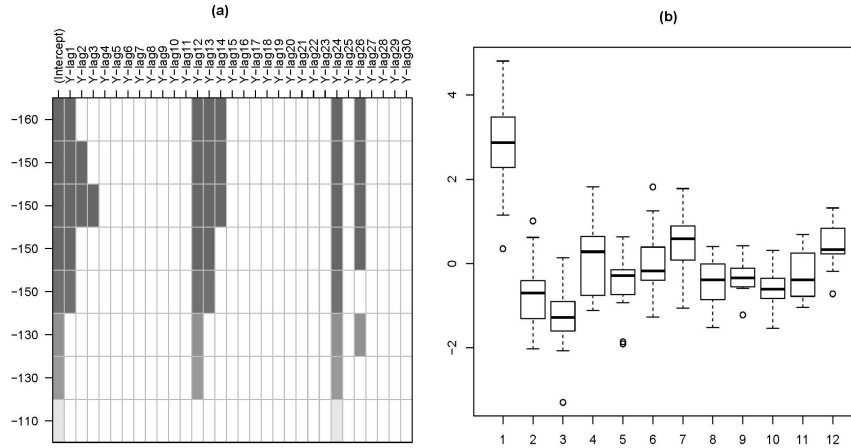


**Figura 3.3.** Para  $X_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, datos empíricos.

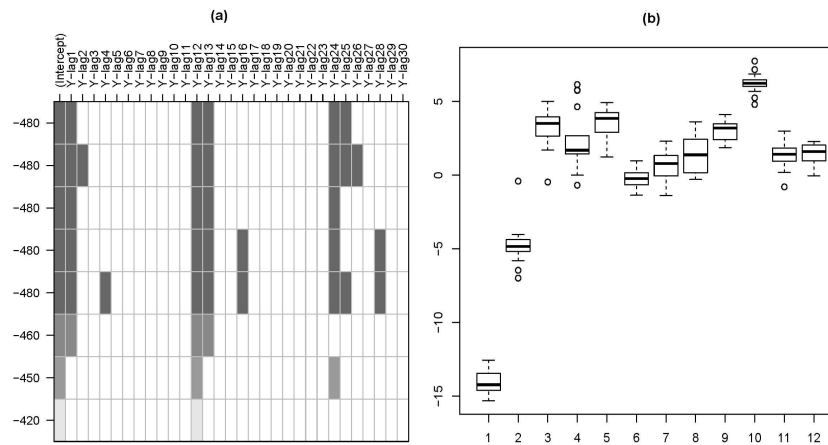


**Figura 3.4.** Para  $Z_t$ , (a) FAC muestral, (b) FACP muestral, (c) IMP muestral, datos empíricos.

Siguiendo las evidencias de ciclos estacionales que muestran las Figuras 3.4 y 3.6 para la variable de umbrales, se ajusta un modelo AR(26) multiplicativo a  $Z_t$ , según el criterio BIC mínimo.



**Figura 3.5.** Para  $X_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por meses, datos empíricos.



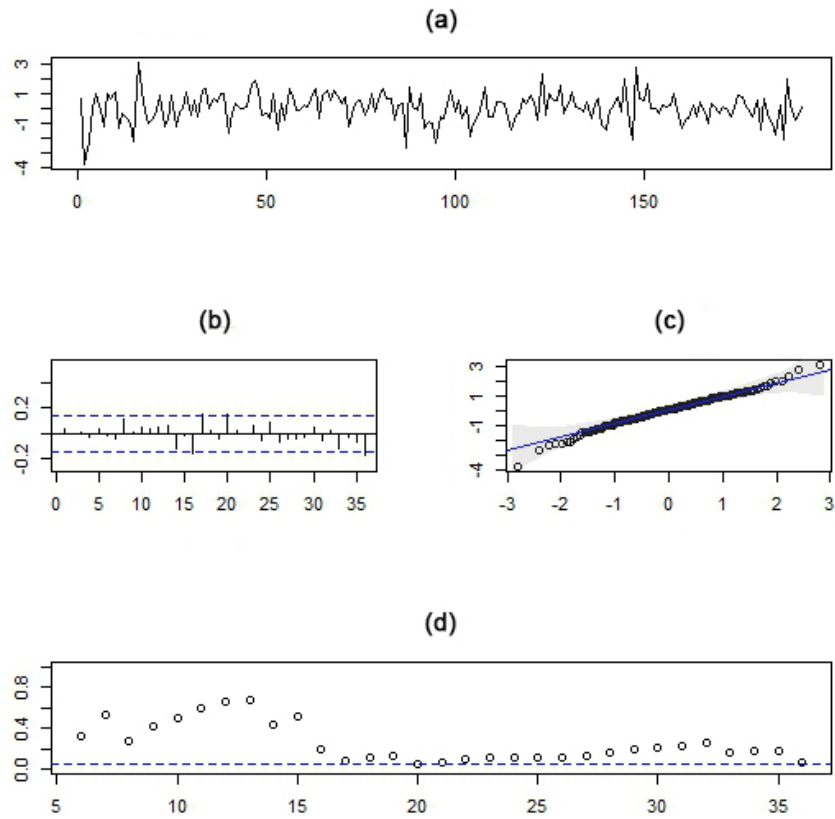
**Figura 3.6.** Para  $Z_t$ , (a) AR multiplicativo con BIC mínimo, (b) diagramas de caja por meses, datos empíricos.

La Tabla 3.2, exhibe las estimaciones de los parámetros del modelo AR en mención. Se puede notar, en los parámetros estimados, que el crecimiento mensual de la transformación logarítmica del  $ISE_t$ , se encuentra afectado en forma inversa por los dos anteriores meses y en la componente estacional por uno y dos años anteriores, afectado en forma directa. Los residuales del modelo ajustado poseen un buen comportamiento, ya que se asemejan a un proceso de ruido blanco Gaussiano (véase Figura 3.7).

**Tabla 3.2.** Ajuste de un modelo AR(26) multiplicativo para la variable de umbrales.

Parámetro	Estimación	E.E.	Valor $t$	valor $p$
$a_0$	0.20	0.99	0.21	0.83
$a_1$	-0.53	0.07	-7.17	0.00
$a_2$	-0.21	0.06	-2.87	0.01
$b_1$	0.60	0.07	8.51	0.00
$b_2$	0.38	0.07	5.39	0.00
$h^2$	1.565			
AIC	720.03			
*AIC si	718.07			

\*Criterio de información de Akaike del modelo AR(26) multiplicativo *sín* intercepto.



**Figura 3.7.** Residuales del modelo AR(26) multiplicativo para  $Z_t$ ,  
 (a) series de tiempo, (b) FAC muestral, (c) QQ-normal y  
 (d) estadísticas de Ljung y Box con diferentes rezagos.

### 3.2. Resultados de la estimación.

Los órdenes autorregresivos máximos preespecificados  $k^*$ ,  $K^*$ , y  $q^*$ , son elegidos de un modelo  $ARX(k + sK, q)$  que mejor se ajuste a los datos empíricos, de acuerdo al criterio DIC mínimo. Se toman modelos ARX candidatos, con  $k = 0, 1, 2, 3, 4, 5$ ,  $K = 0, 1, 2, 3$ , y  $q = 0, 1, 2, 3, \dots, 15$  y  $s = 12$ . Se obtuvo los valores  $k^* = 3$ ,  $K^* = 2$  y  $q^* = 2$ . Estos valores se fijan para aplicar la prueba de no linealidad de umbrales de Tsay (1998), la prueba estadística Bayesiana de estacionalidad propuesta, el método de Congdon y la técnica GVS.

A continuación, se procede a aplicar la prueba de no linealidad de umbrales para rezagos  $d = 0, 1, \dots, 10$ , órdenes autorregresivos  $k + sK = 27$  y  $q = 0, 2$  y 15. La hipótesis nula de linealidad de umbrales  $ARX(k + sK, q)$  es rechazada para diferentes rezagos al 1%; en particular, para  $d = 0$ ,  $k + sK = 27$ ,  $q = 0$  y para  $d = 3$ ,  $k + sK = 27$ ,  $q = 0$  y 2. Estos resultados se muestran en la Tabla 3.3.

**Tabla 3.3.** Prueba de no linealidad de umbrales (Tsay, 1998) para los datos empíricos, considerando diferentes órdenes autorregresivos y valores  $d$ .

$k + sK = 27$ y $q = 0$											
$d$	0	1	2	3	4	5	6	7	8	9	10
Estadística F	2.54	1.37	0.96	1.58	1.01	1.51	1.67	1.06	1.18	1.19	1.52
Valor $p$	0.00	0.13	0.53	0.05	0.46	0.07	0.03	0.40	0.26	0.25	0.07

$k + sK = 27$ y $q = 2$											
$d$	0	1	2	3	4	5	6	7	8	9	10
Estadística F	2.47	1.39	0.87	1.86	0.86	1.52	1.32	0.93	1.45	1.51	1.85
Valor $p$	0.00	1.39	0.67	0.001	0.68	0.06	0.14	0.58	0.08	0.06	0.01

$k + sK = 27$ y $q = 15$											
$d$	0	1	2	3	4	5	6	7	8	9	10
Estadística F	1.32	1.12	1.13	1.08	0.64	1.13	0.99	1.12	0.84	0.72	1.53
Valor $p$	0.14	0.32	0.31	0.37	0.95	0.31	0.50	0.32	0.74	0.88	0.05

Lo anterior, permite considerar un modelo TARX dado en (1.2) con mínimo dos regímenes, que se precisa en la prueba de estacionalidad dada en la sección 1.4. Para la prueba Bayesiana de estacionalidad multiplicativa por regímenes, se consideró  $l = 2$  y 3 regímenes en un modelo TARX. Se requiere en forma intrínseca estimar parámetros de un modelo TARX(2; 3, 3; 2, 2; 2, 2) maximal y de un modelo TARX(3; 3, 3, 3; 2, 2, 2; 2, 2, 2) maximal, para nuestros datos empíricos.

Las salidas MCMC para el análisis a posteriori son de nuevo basadas sobre 6.000 iteraciones después de un periodo de calentamiento de 6.000. Los hiperparámetros y valores iniciales para cada uno de los parámetros de los modelos TARX, son los mismos que se usaron para los modelos 1 y 2 simulados, del capítulo 2. En la Tabla 3.4, se lista

los resultados de la prueba de estacionalidad. Un modelo TSARX multiplicativo con dos regímenes es el opcionado, de acuerdo a la prueba de estacionalidad. Para el caso de  $l = 3$  regímenes, el segundo régimen no evidencia presencia de estacionalidad multiplicativa dado en (1.2).

**Tabla 3.4.** *Prueba de estacionalidad multiplicativa por regímenes para los datos empíricos.*

$l = 2$	$l = 3$
$2 \ln(FB_1) = 321.20$	$2 \ln(FB_1) = 9.464 \times 10^{32}$
$2 \ln(FB_2) = 240611.40$	$2 \ln(FB_2) = 1.698 \times 10^{-12}$
	$2 \ln(FB_3) = 1.649 \times 10^{28}$

Se prosigue a estimar el número de regímenes y órdenes autorregresivos de un modelo TSARX. Valores de los hiperparámetros y valores iniciales de los parámetros para cada modelo TSARX candidatos, son iguales a los descritos en la sección 2.1. La probabilidad a posteriori del número de regímenes de un TSARX con  $j = 1, 2, 3, 4, 5$  regímenes ajustados a los datos empíricos, es dada en la Tabla 3.5, donde se puede ver que el número de regímenes asociado con la probabilidad a posteriori más grande es dos. Este resultado se corrobora con el valor DIC más pequeño y con lo obtenido en la prueba de estacionalidad.

**Tabla 3.5.** *Probabilidades a posteriori y DIC para la selección del número de regímenes de un modelo TSARX para los datos empíricos.*

	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
Prob. a posteriori	0.00021	0.91054	0.00138	0.00168	0.08619
DIC	2088.29	1304.98	1830.86	1838.24	1312.73

En la Tabla 3.6, se observan los tres mejores modelos basado sobre las probabilidades a posteriori para la matriz de variables latentes formada por  $S_j$ ,  $M_j$  y  $N_j$ ,  $j = 1, 2$ . Las probabilidades de los mejores tres modelos TSARX con dos regímenes para los datos empíricos son 0.2122, 0.1536 y 0.1078. Los tres modelos están asociados con  $d = 3$ . Los tres modelos, se ven muy similares. En efecto, ellos tienen la misma estructura en el segundo régimen y difieren principalmente en las entradas de los coeficientes  $a_3^{(1)}$  y  $b_2^{(1)}$ . Se elige el mejor modelo y se estiman el resto de parámetros (estimación puntual, desviaciones estándar e intervalos creíbles del 95 %), los cuales se presentan en la Tabla 3.7.

Siguiendo los resultados de las Tablas 3.2 y 3.7, el modelo TSARX multiplicativo es dado en forma explícita por

$$X_t = \begin{cases} -0.22 - 0.47X_{t-1} + 0.53X_{t-2} + 0.25X_{t-3} + \sqrt{0.36}\epsilon_t, & \text{si } Z_{t-3} \leq 1.80, \\ -0.70X_{t-1} - 0.58X_{t-2} - 0.47X_{t-3} + 0.40X_{t-2} + 0.28X_{t-3} \\ + 0.23X_{t-4} + 0.19X_{t-5} + 0.47X_{t-24} + 0.33X_{t-25} + 0.27X_{t-26} \\ + 0.22X_{t-27} + \sqrt{0.46}\epsilon_t, & \text{si } Z_{t-3} > 1.80, \end{cases}$$

siendo  $Z_t = -0.53Z_{t-1} - 0.21Z_{t-2} + 0.60Z_{t-12} + 0.32Z_{t-13} + 0.13Z_{t-14} + 0.38Z_{t-24} + 0.20Z_{t-25} + 0.08Z_{t-26} + \sqrt{1.57}a_t$ . Algunas raíces de los correspondientes polinomios característicos de los regímenes del modelo TSARX hallado, poseen un periodo de longitud 3, 6, 12 ó 24, indicando ciclos estacionales.

Una interpretación del modelo TSARX ajustado a los datos empíricos es el siguiente: ante todo, solamente dos regímenes para el crecimiento de la tasa de desempleo total mensual Colombiana son detectadas, lo cual podría denominarse como regímenes del crecimiento mensual de la transformación logaritmo del ISE<sub>t</sub>, bajo y alto. En el régimen bajo, también se incluye los picos negativos más pronunciados. Además, es interesante notar que (1) para el primer régimen, el intercepto es negativo y afecta las épocas de recesión de la economía en un 22%, esto refleja el impacto económico que posee el crecimiento de la tasa de desempleo en dicho periodo; por el contrario, en las épocas de expansión de la economía que es modelado por el segundo régimen, el intercepto es nulo evidencia de una rápida recuperación económica, (2) el modelo SAR multiplicativo del segundo régimen es de órdenes autorregresivos más altos que las del primer régimen, (3) los coeficientes de la componente no estacional son de signos negativos, y los coeficientes de la componente estacional son de signos positivos y de magnitud similar, en los dos regímenes del modelo.

Esto indica que, existe una relación inversa del crecimiento de la tasa de desempleo total

**Tabla 3.6.** *Los mejores tres modelos identificados usando el método GVS, para los datos empíricos.*

Mejor	$\left( \begin{array}{cccc cc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{array} \middle  \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right)$ <p style="text-align: center;">(0.2122)</p>
Segundo mejor	$\left( \begin{array}{cccc cc} 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{array} \middle  \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right)$ <p style="text-align: center;">(0.1536)</p>
Tercer mejor	$\left( \begin{array}{cccc cc} 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{array} \middle  \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right)$ <p style="text-align: center;">(0.1078)</p>

*Los paréntesis muestran las correspondientes probabilidades a posteriori. Los tres modelos están asociados con  $d = 3$ .*

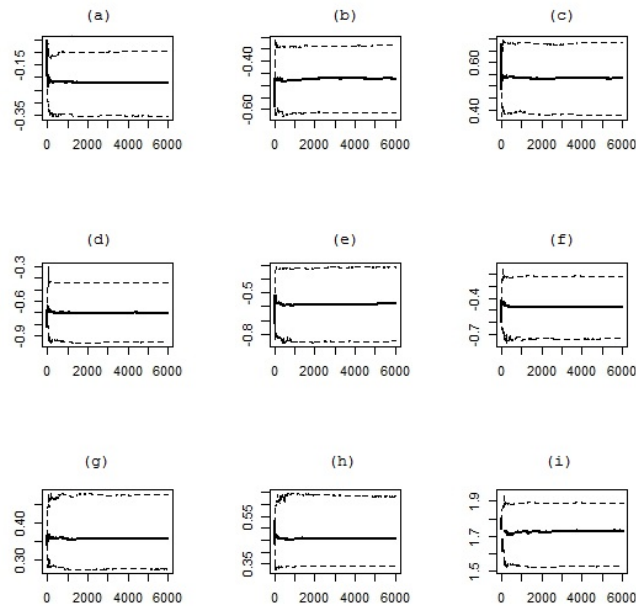
mensual, con los meses inmediatamente anteriores, máximo de tres meses y una relación directa con uno o dos años atrás, en ambos regímenes. (4) las varianzas de los regímenes, son similares, y (5) el crecimiento mensual de la transformación logarítmica del índice de seguimiento económico, afecta el crecimiento de la tasa de desempleo, lo que le suceda tres meses antes.

En las Figuras 3.8 y 3.9, se muestran los gráficos cumuplot y  $Z$ -score de Geweke, respectivamente, para algunas de las cadenas MCMC de los parámetros estimados. La convergencia, a las distribuciones a posteriori, es casi inmediata y claramente bien alcanzadas, después del periodo de calentamiento en cada caso considerado.

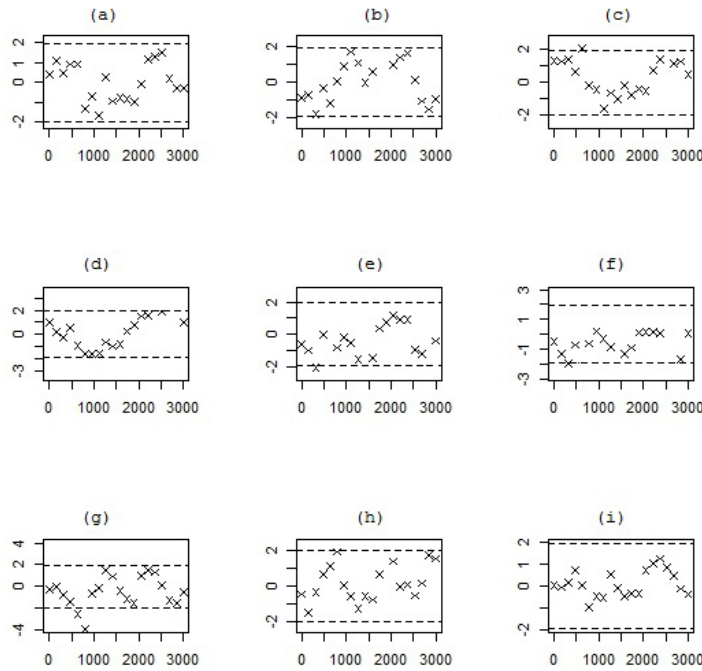
**Tabla 3.7.** *Estimación de parámetros para los datos empíricos.*

Parámetro	Estimación	D.E.	I.C. 95 %
$a_0^{(1)}$	-0.22	0.06	(-0.35, -0.09)
$a_1^{(1)}$	-0.47	0.07	(-0.61, -0.33)
$b_1^{(1)}$	0.53	0.08	( 0.37, 0.68)
$a_1^{(2)}$	-0.70	0.13	(-0.94, -0.44)
$a_2^{(2)}$	-0.58	0.14	(-0.84, -0.31)
$a_3^{(2)}$	-0.47	0.13	(-0.73, -0.21)
$b_1^{(2)}$	0.40	0.09	( 0.22, 0.58)
$b_2^{(2)}$	0.47	0.09	( 0.30, 0.66)
$(h^{(1)})^2$	0.36	0.05	( 0.27, 0.48)
$(h^{(2)})^2$	0.46	0.07	( 0.34, 0.63)
$r$	1.80	0.09	( 1.56, 1.94)
$d$	3(96 %)*		

\* *Frecuencia porcentual del rezago estimado.*



**Figura 3.8.** *Gráficos cumuplot para: (a)  $a_0^{(1)}$ , (b)  $a_1^{(1)}$ , (c)  $b_1^{(1)}$ , (d)  $a_1^{(2)}$ , (e)  $b_1^{(2)}$ , (f)  $b_2^{(2)}$ , (g)  $(h^{(1)})^2$ , (h)  $(h^{(2)})^2$ , (i)  $r$ , datos empíricos.*



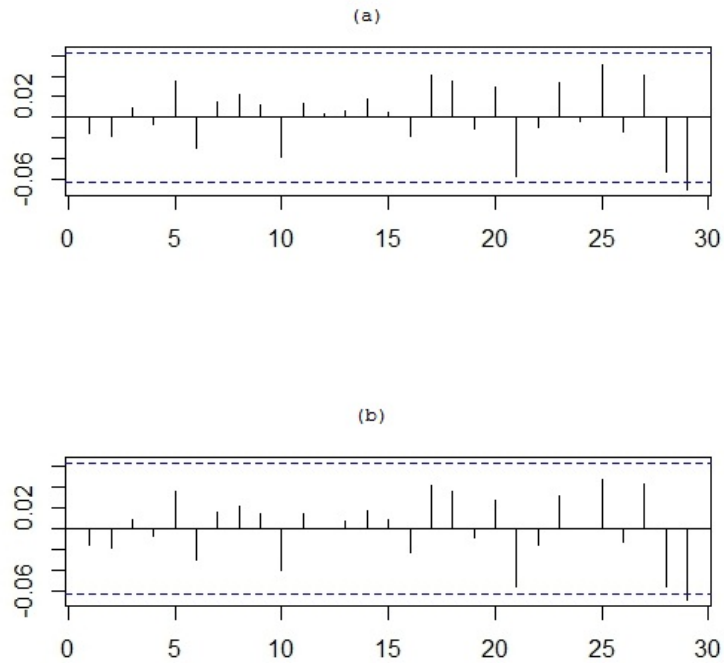
**Figura 3.9.** Gráficos Z-score de Geweke: (a)  $a_0^{(1)}$ , (b)  $a_1^{(1)}$ , (c)  $b_1^{(1)}$ , (d)  $a_1^{(2)}$ , (e)  $b_1^{(2)}$ , (f)  $b_2^{(2)}$ , (g)  $(h^{(1)})^2$ , (h)  $(h^{(2)})^2$ , (i)  $r$ , datos empíricos.

### 3.3. Diagnósticos del modelo.

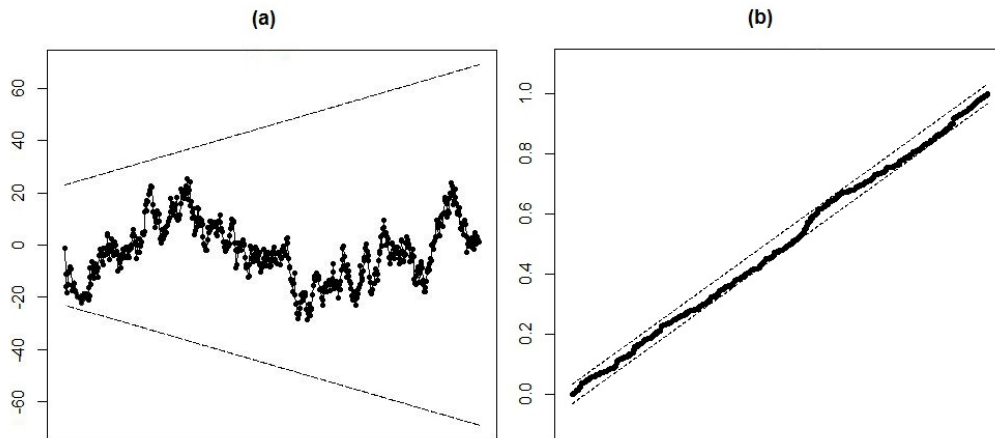
Para conducir el chequeo de diagnósticos para el modelo que se ajustó a los datos empíricos, se produce los residuales generalizados  $\hat{v}_t = \Phi^{-1}(\hat{u}_t)$ ,  $t = 51, \dots, 203$ . Para aplicar la prueba de Ljung y Box para rezagos 1, 2, 3, 12 y 24 y la prueba de Jarque y Bera, 153  $\hat{v}_t$  fueron obtenidos. Las estadísticas de prueba y el valor  $p$  correspondientes son dadas en la Tabla 3.8. En todos los casos, no se rechaza las hipótesis nulas de no autocorrelación cero y normalidad, tomando un nivel de significancia del 5%. Estos resultados sugieren que el modelo TSARX con dos regímenes elegido ajusta los datos empíricos bajo estudio bastante bien. Las Figuras 3.10 y 3.11, muestran los autocorrelogramas, gráficos CUSUM y CUSUMSQ, los cuales corroboran un adecuado ajuste del modelo.

**Tabla 3.8.** Pruebas de Ljung y Box para diferentes rezagos (LB) y de Jarque y Bera (JB) aplicada a los residuales generalizados del modelo TSARX ajustado a los datos empíricos. El valor  $p$  de la prueba está en paréntesis.

$LB_1$	$LB_2$	$LB_3$	$LB_{12}$	$LB_{24}$	$JB$
1.01313	1.06389	1.33908	8.35656	16.56653	38.55404
(0.31415)	(0.58746)	(0.71987)	(0.75668)	(0.86655)	(0.25323)



**Figura 3.10.** (a) *FAC muestral*, (b) *FACP muestral*, *residuales generalizados de los datos empíricos*.



**Figura 3.11.** (a) *CUSUM*, (b) *CUSUMSQ* *residuales generalizados de los datos empíricos*.

### 3.4. Cálculo y evaluación de pronósticos.

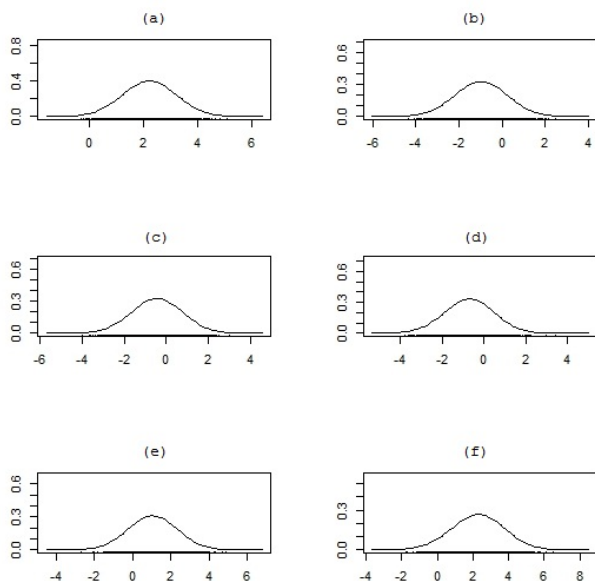
Finalmente, se procede a obtener algunos pronósticos para las variables  $X_t$  y  $Z_t$ . Hay un total de 6.000 iteraciones MCMC de  $Z_{T+h}$  y  $X_{T+h}$ ,  $h = 1, 2, 3, 11, 12$ , y 13 simuladas de las distribuciones predictivas, que surgen de usar el Algoritmo 4 de la sección 1.7.

En la Tabla 3.9, se presentan los pronósticos en mención, iniciando en enero de 2017 ( $h = 1$ ) hasta enero de 2018 ( $h = 13$ ). En la Tabla se incluyen los valores verdaderos, estimaciones puntuales, desviaciones estándar y estimaciones por intervalo del 95 %, de los diferentes  $h$ . Se puede notar que en ambas variables (1) las estimaciones puntuales de los pronósticos son muy cercanas a los valores observados, (2) las desviaciones estándar aumentan con valores de  $h$  más grande, y (3) todos los valores observados caen en los intervalos creíbles.

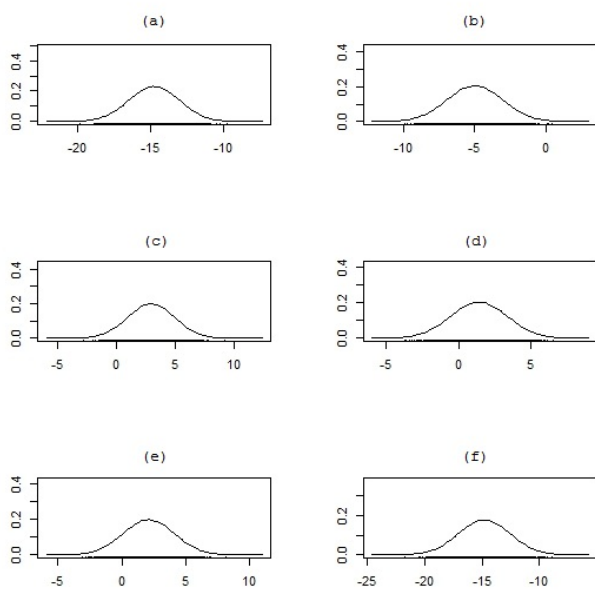
**Tabla 3.9.** *Pronósticos para las variables de interés y de umbrales de los datos empíricos, mediante un TSARX.*

Horizonte (h)	$X_{T+h}$	$\widehat{X}_{T+h}$	D.E.	I.C. 95 %
1	2.99	2.16	0.76	( 0.67, 3.66)
2	-1.23	-0.99	0.80	( -2.64, 0.81)
3	-0.80	-0.36	0.80	( -2.10, 1.31)
11	-0.19	-0.74	0.85	( -2.47, 0.91)
12	0.26	0.96	0.90	( -0.76, 2.79)
13	3.13	2.42	1.08	( 0.28, 4.56)
Horizonte (h)	$Z_{T+h}$	$\widehat{Z}_{T+h}$	D.E.	I.C. 95 %
1	-15.64	-14.79	1.25	(-17.23, -12.32)
2	-5.64	-4.95	1.40	( -7.74, -2.18)
3	3.46	2.97	1.41	( 0.19, 5.75)
11	2.17	1.46	1.43	( -1.32, 4.25)
12	2.26	2.10	1.45	( -0.67, 4.89)
13	-14.78	-14.77	1.61	(-17.93, -11.61)

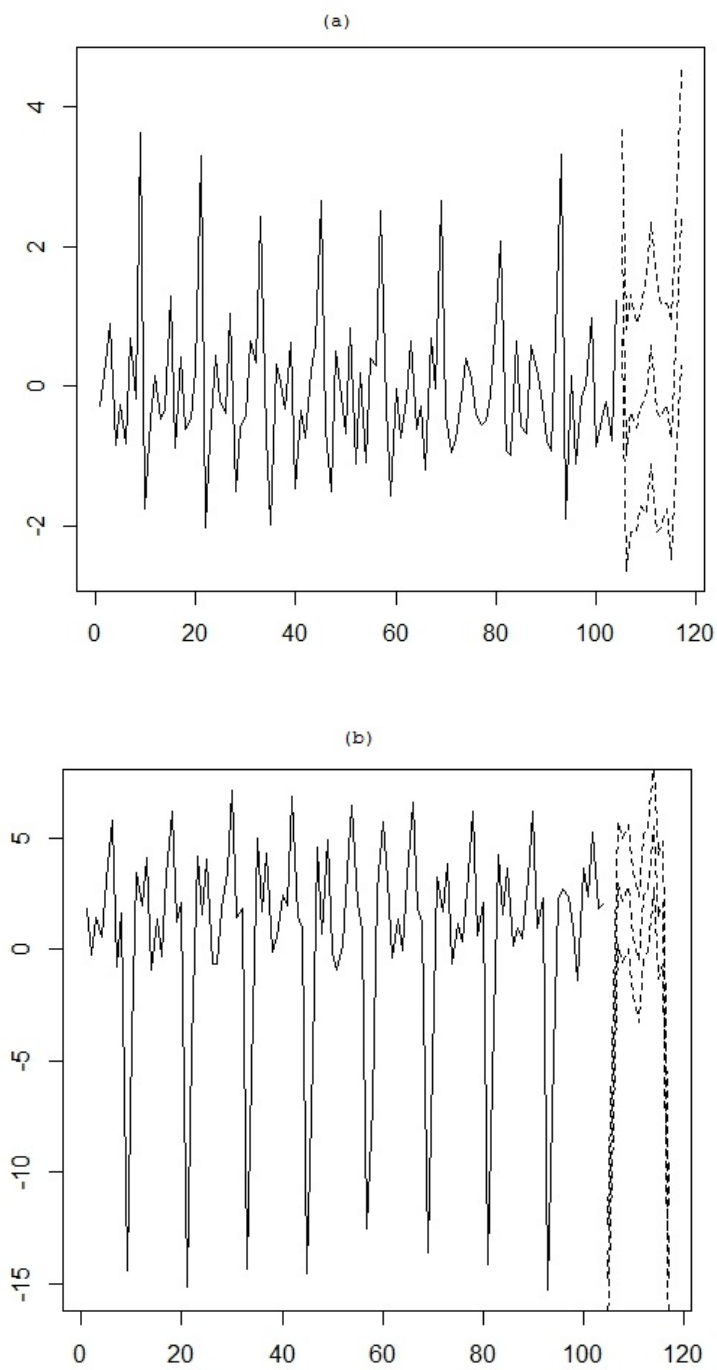
Adicionalmente, de las distribuciones predictivas para el horizonte de pronóstico  $h$  tomado para las dos variables  $X_t$  y  $Z_t$ , se pueden concluir que estas son unimodales y simétricas (véase Figuras 3.12 y 3.13). También se puede decir que, las estimaciones puntuales y por intervalo siguen un patrón estacional que se repite cada doce meses, capturando la dinámica de los datos empíricos, característica que se puede observar en la Figura 3.14. En resumen, los pronósticos obtenidos para la variable de interés y de umbrales son eficientes.



**Figura 3.12.** *Distribuciones predictivas para la variable  $X_t$  datos empíricos, (a)  $h=1$ , (b)  $h=2$ , (c)  $h=3$ , (d)  $h=11$ , (e)  $h=12$ , (f)  $h=13$ , pasos adelante.*



**Figura 3.13.** *Distribuciones predictivas para la variable  $Z_t$  datos empíricos, (a)  $h=1$ , (b)  $h=2$ , (c)  $h=3$ , (d)  $h=11$ , (e)  $h=12$ , (f)  $h=13$ , pasos adelante.*



**Figura 3.14.** Gráficas de pronósticos, (a) para  $X_t$ , (b) para  $Z_t$ , datos empíricos.

---

---

## Conclusiones

---

---

En este trabajo se propone un modelo autorregresivo de umbrales, denominado TSARX, el cual puede usarse para capturar estacionalidad multiplicativa dependiente del régimen. El modelo propuesto es un caso especial de un modelo TARX al imponer ciertas restricciones sobre los coeficientes autorregresivos del modelo.

La metodología propuesta para ajustar un modelo TSARX a series de tiempo, está basada sobre métodos Bayesianos y técnicas MCMC y consiste en las fases de (i) explorar la presencia de estacionalidad multiplicativa por regímenes y no linealidad de umbrales en los datos, a través de gráficos y pruebas estadísticas. Una prueba de estacionalidad multiplicativa por regímenes es propuesta. (ii) Estimar el número de regímenes, órdenes autorregresivos y el resto de parámetros del modelo, en forma simultánea, (iii) validar el modelo, examinando los residuales generalizados y (iv) hacer uso del modelo, mediante el cálculo de pronósticos.

El funcionamiento de la metodología desarrollada es satisfactoria en datos simulados, donde la variable de umbrales también puede presentar un comportamiento estacional o de ciclos estocásticos no necesariamente estacionales. Los datos empíricos utilizados fueron, el crecimiento de la tasa de desempleo total mensual Colombiana como variable de interés y la diferencia logarítmica del índice de seguimiento a la economía mensual para Colombia como variable de umbrales, en el periodo muestral desde Febrero, 2001 hasta Diciembre, 2016. Un modelo TSARX con dos regímenes captura muy bien las características intramuestra y en pronósticos de las series de tiempo económicas en mención.

En futuras investigaciones, se podría considerar la extensión al caso multivariado del modelo TSARX propuesto; además, se puede investigar la caracterización y modelamiento, de distintas formas de estacionalidad presentes en series de tiempo, en el contexto de la familia de los modelos TAR. También, sería interesante la comparación de pronósticos usando otros modelos de series de tiempo no lineales en presencia de estacionalidad, estudiar la potencia de la prueba de no linealidad de umbrales de Tsay ante presencia de estacionalidad en los datos, considerar herramientas diagnósticos adicionales para validar el modelo e investigar modelos TSARX cuando el proceso de ruido no sigue una distribución normal.

---

---

## Bibliografía

---

---

- Box, G.E.P., and Jenkins, G.M. (1970), *Time series analysis, forecasting and control*, Holden-Day, San Francisco.
- Calderón, S., and Nieto, F. (2017), Bayesian analysis of multivariate threshold autoregressive models with missing data, *Communications in Statistics, Theory and Methods*, **46**, 296-318.
- Canova, F., and Ghysels, E. (1994), Changes in seasonal patterns are they cyclical?, *Journal of Economic Dynamics and Control*, **18**, 1143-1171.
- Casella, G., and George, E.I. (1992), “Explaining the Gibbs Sampler”, *The American Statistician*, **46**, 167-174.
- Chen, C.W.S., and Lee, J.C. (1995), Bayesian inference of threshold autoregressive models, *Journal of Time Series Analysis*, **16**, 483-492.
- Chen, C.W.S. (1998), A Bayesian analysis of generalized threshold autoregressive models, *Statistics & Probability Letters*, **40**, 15-22.
- Chen, C.W.S., and So, M.K.P. (2006), On a threshold heteroscedastic model, *International Journal of Forecasting*, **22**, 73-89.
- Chen, C.W.S., Gerlach, R., and Lin, A.M.H. (2010), Falling and explosive, dormant and rising markets via multiple - regime financial time series models, *Applied Stochastic Models in Business and Industry*, **26**, 28-49.
- Chen, C.W.S., So, M.K.P., and Lui, F.C. (2011), A review of threshold time series models in finance, *Statistics and Its Interface*, **4**, 167-181. Chib, S. (1995), Marginal likelihood from the Gibbs output, *Journal of the American Statistical Association*, **90**, 1313-1321.
- Chib, S., and Greenberg, E. (1995), Understanding the Metropolis-Hastings algorithm, *The American Statistician*, **49**, 327-335.
- Congdon, P. (2001), *Bayesian statistical modeling*, John Wiley & Sons, New York.
- Congdon, P. (2006), Bayesian model choice based on Monte Carlo estimates of posterior model probabilities, *Computational Statistics and Data Analysis*, **50**, 346-357.
- Crespo, J. (2001), *Modelling seasonal asymmetries using seasonal SETAR models*, Working Paper, Department of Economics, University of Vienna.
- De Gooijer, J. G., and Vidiella-i-Anguera, A. (2003), Nonlinear stochastic inflation modeling using SEASETARs, *Insurance: Mathematics and Economics*, **32**, 3-18.

- Dellaportas, P., Forster, J., and Ntzoufras, I. (2002), On Bayesian model and variable selection using MCMC, *Statistics and Computing*, **12**, 27-36.
- Dickey, J. (1971), The weighted likelihood ratio, linear hypothesis on normal location parameters, *Annals of Mathematical Statistics*, **42**, 204-223.
- Dickey, D.A., and Fuller, W.A. (1979), Distribution of the estimators for autoregressive times series with a unit root, *Journal of the American Statistical Association*, **74**, 427-431.
- Dirección de Síntesis y Cuentas Nacionales DSCN (2016), *Metodología general, indicador de seguimiento a la economía ISE*, Departamento Administrativo Nacional de Estadística DANE, Bogotá.
- Dirección de Metodología y Producción Estadística DIMPE (2016), *Metodología general, gran encuesta integrada de hogares GEIH*, Departamento Administrativo Nacional de Estadística DANE, Bogotá.
- Franses, P.H. (1996), *Periodicity and stochastic trends in economic time series*, Oxford University Press, Oxford.
- Franses, P.H., and Paap, R. (1999), Does seasonality influence the dating of business cycle turning points?, *Journal of Macroeconomics*, **21**, 79-92.
- Franses, P.H., De Bruin, P., and Van Dijk, D. (2000), *Seasonal smooth transition autoregression*, Econometric Institute Report 2000-06/A, Econometric Institute, Erasmus University, Amsterdam.
- Franses, P.H., and Van Dijk, D. (2000), *Nonlinear time series models in empirical finance*, Cambridge University Press, Cambridge.
- Franses, P.H., and Van Dijk, D. (2005), The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production, *International Journal of Forecasting*, **21**, 87-102.
- Fok, D., Franses, P.H., and Paap, R. (2007), Seasonality and no-linear price effects in scanner-data-based market - response models, *Journal of Econometrics*, **138**, 231-251.
- Fuller, W.A. (1996), *Introduction to statistical time series, second edition.*, John Wiley & Sons, New York.
- Gelfand, A.E., and Smith, A.F.M. (1990), Sampling -based approaches to calculating marginal densities, *Journal of the American Statistical Association*, **85**, 398-409.
- Gelman, A., Roberts, G. O., and Gilks, W.R. (1996), *Efficient Metropolis jumping rules*, In J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith (Eds.), *Bayesian Statistics 5* (pp. 599-607), Oxford University Press, Oxford.
- George, E., and McCulloch, R. (1993), Variable selection via Gibbs sampling, *Journal of the American Statistical Association*, **8**, 881-889.
- Gerlach, R., Carter, C.K., and Kohn, R. (1999), Diagnostics for time series analysis, *Journal of Time Series Analysis*, **20**, 309-330.
- Gerlach, R., and Chen, C.W.S. (2008), Bayesian inference and model comparison for asymmetric smooth transition heteroskedastic models, *Statistics and Computing*, **18**, 391-408.
- Geweke, J. (1992), *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, In *Bayesian Statistics 4*, Oxford University Press, Oxford.

- Ghysels, E. (1994), *On the economics and econometrics of seasonality in economic time series*, In C. Sims (ED.), *Advances in econometrics*, 6th world congress of the econometric society, Cambridge University Press, Cambridge.
- Ghysels, E., and Osborn, D.R. (2001), *The econometric analysis of seasonal time series*, Cambridge University Press, Cambridge.
- Hansen, B.E. (2011), Threshold autoregression in economics, *Statistics and Its Interface*, **4**, 123-127.
- Hegger, R., Kantz, H., and Schreiber, T. (1999), Practical implementation of nonlinear time series methods: the TISEAN package, *CHAOS*, **9**, 413-435.
- Hoyos, M. (2006), *Una aplicación del modelo no lineal TAR en economía*, Tesis de grado maestría en ciencias estadística, Universidad Nacional de Colombia, Bogotá.
- Hylleberg, S., Jørgensen, C., and Sørensen, N.K. (1993), Seasonality in macroeconomic time series, *Empirical Economics*, **18**, 321-325.
- Kass, R.E., and Raftery, A.E. (1995), Bayes factors, *Journal of the American Statistical Association*, **90**, 773-795.
- Koop, G., and Potter, S. (1999), Dynamic asymmetries in U.S. unemployment, *Journal of Business & Economic Statistics*, **17**, 298-312.
- Kuo, L., and Mallick, B. (1998), Variable selection for regression models, *Sankhyā: The Indian Journal of Statistical Series B*, **60**, 65-81.
- Leamer, E.E. (2009), *Macroeconomic patterns and stories*, Springer-Verlag, Berlín.
- Metropolis, Rosenbluth, A.W., Rosenbluth, M.N., and Teller, A.H. (1953), Equations of state calculations by fast computing machines, *Journal of Chemical Physics*, **21**, 1087-1091.
- Meyn, S., and Tweedie, R.L. (2009), *Markov chains and stochastic stability, second edition*, Cambridge University Press, New York.
- Nieto, F. (2005), Modeling bivariate threshold autoregressive processes in the presence of missing data, *Communications in Statistics, Theory and Methods*, **34**, 905-930.
- Nieto, F. (2008), Forecasting with univariate TAR models, *Statistical Methodology*, **5**, 263-276.
- Nieto, F., Zhang, H., and Li, W. (2013), Using the reversible jump MCMC procedure for identifying and estimating univariate TAR models, *Communications in Statistics, Simulation and Computation*, **42**, 814-840.
- Nieto, F., and Moreno, E. (2016), Univariate conditional distributions of an open-loop TAR stochastic process, *Revista Colombiana de Estadística*, **39**, 149-165.
- O'hara, R.B., and Sillanpää, M.J. (2009), A review of Bayesian variable selection methods: What, how and which, *Bayesian Analysis*, **4**, 85-118.
- Osborn, D.R., Chui, A.P.L., Smith, J., and Birchenhall, C.R. (1988), Seasonality and the order of integration for consumption, *Oxford Bulletin of Economics and Statistics*, **50**, 361-377.
- Plummer, M., Best, N., Cowles, K., and Vines, K. (2006), Coda: Convergence diagnosis and output analysis for MCMC, *R News*, **6**, 7-11.

- Rothman, P. (1998), Forecasting asymmetric unemployment rates, *Review of Economics and statistics*, **80**, 164-168.
- Skalin, J., and Teräsvirta, T. (2001), Modelling asymmetries and moving equilibria in unemployment rates, *Macroeconomic Dynamics*, **6**, 202-241.
- So, M., and Chen, C.W.S. (2003), Subset threshold autoregression, *Journal of Forecasting*, **22**, 49-66.
- So, M.K.P., Chen, C.W.S., and Liu, F.C. (2006), Best subset selection of autoregressive models with exogenous variables and generalized autoregressive conditional heteroscedasticity errors, *Journal of the Royal Statistical society series C*, **55**, 201-224.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Van der Linde, A. (2002), Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society Series B* **64**, 583-640.
- Tong, H. (1978), *On a threshold model*, in C.H. Chen, ed., Pattern recognition and signal processing, Amsterdam: Sijhoff & Noordhoff.
- Tong, H., and Lim, K.S. (1980), Threshold autoregression, limit cycles, and cyclical data, *Journal of the Royal Statistical Society series B*, **42**, 245-292.
- Tong, H. (1990), *Nonlinear time series: A dynamical system approach*, Oxford University Press, Oxford.
- Tong, H. (2015), Threshold models in time series analysis-some reflections, *Journal of Econometrics*, **189**, 485-491.
- Tsay, R. (1998), Testing and modeling multivariate threshold models, *Journal of the American Statistical Association*, **93**, 1188-1202.
- Vargas, L. (2012), *Cálculo de la distribución predictiva en un modelo TAR*, Tesis de grado maestría en ciencias estadística, Universidad Nacional de Colombia, Bogotá.
- Verdinelli, I., and Wasserman, L. (1995), Computing Bayes Factors using a generalization of the Savage Dickey density ratio, *Journal of the American Statistical Association*, **90**, 614-618.
- Vivas, K.L. (2011), *Modelación no lineal de la tasa de desempleo urbano para Colombia*, Tesis de grado maestría en ciencias económicas, Universidad Nacional de Colombia, Bogotá.
- Zhang, H., and Nieto, F. (2015), TAR modelling with missing data when white noise process follows a Student's t-distribution, *Revista Colombiana de Estadística*, **38**, 239-266.