CÁLCULO DE LA DISTRIBUCIÓN PREDICTIVA EN UN MODELO TAR

Luz Elena Vargas Ortíz

Director:

Dr. Fabio Nieto Sánchez

Universidad Nacional de Colombia

Departamento de Estadística

2012

CÁLCULO DE LA DISTRIBUCIÓN PREDICTIVA EN UN MODELO TAR

Luz Elena Vargas Ortíz

Proyecto de Tesis para optar al titulo de maestro en
Ciencias Estadísticas

Director:

Dr. Fabio Nieto Sánchez

Universidad Nacional de Colombia

Departamento de Estadística

2012

A Dios por darme sabiduría.

A la memoria de mis padres María y Antonio porque siempre creyeron en mí.

A mi hijo Santiago por ser mi más grande motivación.

Luz Elena.

Tabla de Contenido

RESUMEN.

ABSTRAC.

INTRODUCCIÓN.

1. ESTADÍSTICA BAYESIANA. 1

1.1. Generalidades. 1

1.1.1. Tipos de Distribuciones a priori. 2

1.2. Métodos de simulación MCMC. 3

1.2.1. Algoritmo de Metropolis-Hastings. 4

1.2.2. Muestreador de Gibbs. 6

1.3. Predicción Bayesiana. 6

2. MODELOS TAR. 8

2.1. Especificación del Modelo TAR. 8

2.2. Algunas Propiedades del Modelo TAR. 9

2.3. Estimación de los Parámetros en un Modelo TAR. 11

2.3.1. Distribuciones a priori informativas 11

2.3.2. Distribuciones a priori no informativas. 12

3. CÁLCULO DE PRONÓSTICOS CON MODELOS TAR. 15

3.1. Algunos Resultados. 15

3.2. Predicción Bayesiana. 17

4. APLICACIONES. 20

4.1. Aplicación en Datos Simulados. 20

4.1.1. TAR(2;0,0). 21

4.1.2. TAR(2;0,1). 24

4.1.3. TAR(2;1,1). 26

4.1.4. TAR(2;1,2). 28

4.1.5. TAR(3;1,2,2). 30

4.2. Aplicación con Datos Reales. 32

CONCLUSIONES.

ANEXOS.

BIBLIOGRAFÍA.

RESUMEN.

Este trabajo propone un método para calcular la distribución predictiva Bayesiana en modelos TAR, utilizando distribuciones a priori informativas y no informativas. De acuerdo con los resultados obtenidos, bajo la función la Perdida Cuadrática Bayesiana y de los intervalos de credibilidad calculados. Los resultados indican que las distribuciones informativas proporcionan mejores predicciones que las distribuciones no informativas y que el método propuesto por Nieto (2008).

Palabras claves: Modelos TAR, Pronósticos, Predicción Bayesiana, Series de tiempo no lineales.

ABSTRAC.

This paper proposes a method to calculate the Bayesian predictive distribution in TAR models using informative priors and not informative. According to the results obtained under the quadratic loss function and Bayesian credibility intervals calculated. The results indicate that informative distributions provide better predictions than not informative distributions and that the method proposed by Nieto (2008).

Keywords: TAR Models, Forecasts, Bayesian prediction, nonlinear time series.

INTRODUCCIÓN.

Los modelos no lineales fueron mencionados a principios del siglo XX. Estos han adquirido importancia a través de los años, ya que muchos fenómenos de la naturaleza pueden explicarse por medio de ellos; como en el caso la hidrología y meteorología (ver por ejemplo Nieto (2005) y Briñez & Nieto (2005)) o Económicas (Hoyos & Nieto (2006)y Salamanca & Monrroy (2008)).

El modelo TAR (Threshold Auto-regressive) es un modelo no lineal utilizado en el análisis de series de tiempo, el cual se basa en una variable exógena denominada variable de umbrales. Nieto (2005) desarrolló una metodología Bayesiana basada en el procedimiento MCMC para modelar procesos bivariados, en presencia de datos faltantes, por medio de modelos TAR. Posterior a esto, Nieto (2008) obtuvo la distribución predictiva de una variable que sigue un modelo TAR. Sin embargo, esta distribución ignora el hecho de que los parámetros involucrados en el modelo pueden comportarse como variables aleatorias (como lo ilustran Tong (1990) y Nieto (2005)). Algunos autores tales como Ibañez & Simó (2005), Rojo & Sans (2010) y Bolívar & Cartaya (2009) han calculado las distribuciones predictivas en modelos de series de tiempo, pero existen pocas aplicaciones de estos en modelos no lineales.

El interés de este trabajo se centra en calcular la distribución predictiva de una variable que sigue un modelo TAR a partir de los métodos Bayesianos formales y

posterior a esto, realizar una comparación con la metodología de Nieto (2008) usando datos hidrologicos/meteorológicos. El trabajo se organiza de la siguiente manera:

En el capitulo 1 se consideran generalidades de la estadística Bayesiana, en el capitulo 2 se presentan prerrequisitos teóricos de los modelos TAR y algunas de sus propiedades, así como la estimación Bayesiana de sus parámetros, el capitulo 3 propone una expresión para la distribución predictiva de una variable X_T, que sigue un modelo TAR y en el capitulo 4 se incluyen ejemplos del uso de la distribución predictiva bayesiana propuesta y una aplicación en datos hidrológicos/meteorológicos.

Capítulo 1

ESTADÍSTICA BAYESIANA.

1.1. Generalidades.

En la inferencia Bayesiana el interés se centra en un vector de parámetros $\boldsymbol{\theta}$. De acuerdo con el teorema de Bayes se tiene que la función de distribución a posteriori de $\boldsymbol{\theta}$ dado los datos, es dada por

$$\pi(\boldsymbol{\theta}|\mathbf{x}) = \frac{f(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\Theta} f(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}, \theta \in \Theta, \tag{1.1}$$

donde: $\mathbf{x}$ es el vector de datos, $\boldsymbol{\theta}$ es el vector de parámetros, $f(\mathbf{x}|\boldsymbol{\theta})$ es la función de verosimilitud de $\mathbf{x}$ dado $\boldsymbol{\theta}$, $\pi(\boldsymbol{\theta})$ es la distribución a priori de $\boldsymbol{\theta}$ y Θ es el espacio de parámetros del vector $\boldsymbol{\theta}$.

(1.1) implica que

$$\pi(\boldsymbol{\theta}|\mathbf{x}) \propto f(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \tag{1.2}$$

donde la constante de proporcionalidad es $\left(\int_{\Theta} f(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}\right)^{-1}$.
Proponer una distribución a priori para $\boldsymbol{\theta}$ no es sencillo, debido a la incertidumbre que muchas veces se tiene acerca de $\pi(\boldsymbol{\theta})$.

Bajo el enfoque Bayesiano, se puede obtener:

- Un estimador para $\boldsymbol{\theta}$ vía $E(\boldsymbol{\theta}|\mathbf{x}) = \mu_\theta = \int_\Theta \boldsymbol{\theta}\pi(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}$ y una medida de variabilidad dada por:

$$E[(\boldsymbol{\theta} - \mu_\theta)^2|\mathbf{x}] = \int_\Theta (\boldsymbol{\theta} - \mu_\theta)^2)\pi(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}.$$

- Si $(\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_p)$ es el vector de parámetros de interés, se puede obtener la distribución marginal a posteriori de $\boldsymbol{\theta}_i$ como sigue:

$$\pi(\boldsymbol{\theta}_i|\mathbf{x}) = \int \pi(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}_{-i}, i : 1, \cdots, p.$$

- Pronósticos, ya sea de observaciones futuras o de datos faltantes en el vector de datos $\mathbf{x}$.

1.1.1. Tipos de Distribuciones a priori.

Las distribuciones a priori se pueden clasificar en informativas y no informativa, como se describe brevemente a continuación.

- **Distribución a priori informativa:** En este caso existe algún tipo de información que permite asociar una función de distribución o de masa a los parámetros.

- **Distribución a priori no informativa:** Se utiliza cuando no existe información alguna con respecto a los parámetro y toda la información a posteriori es dada por los datos. Entre estas se encuentra la distribución a priori de Laplace, la cual asume $\pi(\boldsymbol{\theta}) \propto 1$ y la distribución a priori de Jeffreys, la cual se define a continuación.

***Definición* 1.1.** *Sea $f(x|\boldsymbol{\theta})$ la densidad de x dado el vector $\boldsymbol{\theta}$. Si $\boldsymbol{\theta}$ tiene p componentes, la matriz de información de fisher es definida como*

$$\mathfrak{I}(\boldsymbol{\theta}) = E\left[\frac{\partial^2 log(f(x|\boldsymbol{\theta}))}{\partial\theta_i \partial\theta_j}\right]_{p\times p}$$

y la distribución a priori de Jeffreys se define

$$\pi(\boldsymbol{\theta}) \propto |\ \mathfrak{I}(\boldsymbol{\theta})\ |^{1/2}$$

El método de Jeffreys se utiliza para calcular distribuciones apriori no informativas para $\boldsymbol{\theta}$ y está basado en la matriz de información de Fisher. Fue Jeffreys quien propuso este método para encontrar distribuciones no informativas.

1.2. Métodos de simulación MCMC.

Los métodos MCMC (Markov Chain Monte Carlo) son muy utilizados en la estadística Bayesiana y se basan en las propiedades de las cadenas de Markov. La idea de los métodos MCMC es construir una cadena de Markov cuya distribución estacionaria sea la distribución a posteriori $f(\boldsymbol{\theta}|\mathbf{x})$ de la que se desea obtener una muestra. Para definir la convergencia de estos métodos de simulación se deben cumplir algunas propiedades en la cadena de Markov, son condiciones suficientes la irreducibilidad, recurrente positiva y aperiódica (Casella, 1992).

Se puede considerar el periodo de una cadena como el tiempo necesario para repetir un ciclo de valores de la cadena. Es conveniente, así, en el caso de MCMC que las cadenas sean no periódicas. Si una cadena es irreducible, recurrente positiva y no periódica entonces se dice que es ergódica. Esto equivale a decir que la cadena ya no abandona la distribución de equilibrio y que olvida cuál fue su primer estado inicial.

Existe un teorema de ergodicidad que es el equivalente a la ley fuerte de los grandes números pero con cadenas de Markov. Establece que cualquier función de la distribución se puede estimar mediante una muestra de la cadena de Markov obtenidas a partir de un estado ergódico. Así, tras alcanzar el estado ergódico, se obtiene una muestra válida de la distribución a posteriori y, así, se puede utilizar esos datos para estimar medias, varianzas o cuantiles. En estadística el problema es el inverso: cómo construir la cadena de Markov para obtener una muestra de una distribución dada. Existen dos técnicas básicas: El algoritmo de *Metropolis-Hastings* y el muestreador de *Gibbs*, que es un caso particular del primero.

1.2.1. Algoritmo de Metropolis-Hastings.

Este es el algoritmo más general. Se trata de simular muestras la densidad aposteriori $f(\boldsymbol{\theta}|\mathbf{x})$ salvo, quizás, la constante de integración. En este caso, se busca una densidad condicional $g(.|\boldsymbol{\theta})$ o distribución propuesta que sea fácil de muestrear. Después, se generan observaciones de esta distribución propuesta para decidir si pertenecen a la distribución de $\boldsymbol{\theta}|\mathbf{x}$. El algoritmo *Metropolis-Hastings* se puede resumir de la forma siguiente:

i) Definir un valor inicial $\boldsymbol{\theta}^{(0)}$.
ii) Para $t > 0$, generar $\phi \sim g\left(.|\boldsymbol{\theta}^{(t)}\right)$.
iii) Definir $\alpha = \text{mín}\left\{1, \dfrac{\pi(\phi) g\left(\boldsymbol{\theta}^{(t)}|\phi\right)}{\pi\left(\boldsymbol{\theta}^{(t)}\right) g\left(\phi|\boldsymbol{\theta}^{(t)}\right)}\right\}$
iv) Tomar:

$$\boldsymbol{\theta}^{(t+1)} = \begin{cases} \phi, \text{con probabilidad } \alpha \\ \boldsymbol{\theta}^{(t)}, \text{En caso contrario} \end{cases}$$

v) $t = t + 1$. Ir a *ii*).
Se puede demostrar que mediante este algoritmo se simula una muestra de una cadena

de Markov cuya distribución estacionaria es $f(\boldsymbol{\theta}|\mathbf{x})$. En teoría, se puede utilizar cualquier distribución $g(\phi|\boldsymbol{\theta})$. Lo más importante es que sea fácil de muestrear, aunque la convergencia del algoritmo será más rápida si $g \approx f$.

Comprobación de la convergencia.

i) Comparación de los resultados con valores iniciales diferentes.
ii) Gráficos de los valores $\boldsymbol{\theta}^{(t)}$.
iii) Gráficos de la media estimada de $\boldsymbol{\theta}$.
iv) Diagnósticos formales de convergencia; aunque los métodos garantizan la convergencia, se pueden verificar la autocorrelación de la muestra, ya que correlaciones altas conllevan a una convergencia lenta.

Habitualmente, se desechan los primeros datos generados por la cadena de Markov para que la cadena olvide su estado inicial, de modo que para la estimación solo se utiliza la última parte de los datos generados. Se estima (por ejemplo) la media aposteriori de $\boldsymbol{\theta}$ usando

$$E(\boldsymbol{\theta}|\mathbf{x}) \approx \frac{1}{N-M} \sum_{t=M+1}^{N} \boldsymbol{\theta}^{(t)}, \tag{1.3}$$

donde M es el número de iteraciones iniciales seleccionadas y N es el número total de iteraciones. La convergencia puede ser más rápida si $\boldsymbol{\theta}^{(0)}$ se aproxima a la moda de la distribución final.

Para estimar la forma de la distribución final, se pueden usar estimadores kernel (equivalen al suavizado del histograma).

1.2.2. Muestreador de Gibbs.

El muestreador de Gibbs es una versión del método de *Metropolis-Hastings* donde se toma como distribución de propuesta la distribución condicionada completa de la distribución original de donde se quería simular el parámetro $\boldsymbol{\theta}_i$, es decir:

$$g_i(.|\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}) = f(.|\boldsymbol{\theta}_{-i}, \mathbf{x}),$$

donde, $\boldsymbol{\theta}_{-i}$ representa el vector de parámetros $\boldsymbol{\theta}$ omitiendo el parámetro i . El muestreador de Gibbs es una técnica introducida por Geman & Geman (1984) y desarrollada posteriormente por Gelfand & Smith (1990). Se puede demostrar que la probabilidad de aceptar ϕ es siempre 1 y, así, los valores propuestos son siempre aceptados. Aunque, en este caso, las distribuciones condicionadas completas deben ser todas fáciles de simular.

El algoritmo para el muestreador de Gibbs es el siguiente:

i) Comenzar con valores iniciales arbitrarios $\boldsymbol{\theta}^{(0)}$
ii) Generar $\boldsymbol{\theta}_1^{(t+1)} \sim f\left(\boldsymbol{\theta}_1|\mathbf{x}, \boldsymbol{\theta}_2^{(t)}, \boldsymbol{\theta}_3^{(t)}, \cdots, \boldsymbol{\theta}_k^{(t)}\right)$
iii) Generar $\boldsymbol{\theta}_2^{(t+1)} \sim f\left(\boldsymbol{\theta}_2|\mathbf{x}, \boldsymbol{\theta}_1^{(t+1)}, \boldsymbol{\theta}_3^{(t)} \cdots, \boldsymbol{\theta}_k^{(t)}\right)$
$\vdots$
Generar $\boldsymbol{\theta}_k^{(t+1)} \sim f\left(\boldsymbol{\theta}_k|\mathbf{x}, \boldsymbol{\theta}_1^{(t+1)}, \boldsymbol{\theta}_2^{(t+1)}, \cdots, \boldsymbol{\theta}_{k-1}^{(t+1)}\right)$
iv) Ir a *ii*.

1.3. Predicción Bayesiana.

La estadística Bayesiana tiene un método general para encontrar la distribución condicional de una variable en un tiempo futuro, dada una muestra de valores observados. Esta es llamada la distribución predictiva.

El problema es cómo combinar la incertidumbre de la muestra dada con la incer-

tidumbre futura de la variable. En el método Bayesiano se llama marginación, lo cual implica encontrar la distribución aposteriori conjunta para la observación Y y el parámetro θ, dada la muestra aleatoria. La distribución marginal de la observación futura, dada la muestra aleatoria se encuentra por la integración del parámetro de la distribución aposteriori conjunta (Casella, 1985).

***Definición* 1.2.** *Sea Y la variable a predecir en el futuro. La distribución predictiva de Y, dado el vector de datos $\boldsymbol{x}$, es dada por*

$$p(y|\boldsymbol{x}) = \int_{\Theta} p(y|\theta, \boldsymbol{x})\pi(\theta|\boldsymbol{x})d\boldsymbol{\theta}, \tag{1.4}$$

donde, $p(y|\theta, \boldsymbol{x})$ es la distribución de Y dado el parámetro θ y el vector de datos $\boldsymbol{x}$.

El predictor para la observación futura Y y la perdida cuadrática Bayesiana (varianza de la distribución predictiva) son dados por $E(y|\boldsymbol{x}) = \int yp(y|\boldsymbol{x})dy$ y $V(y|\boldsymbol{x}) = E(y^2|\boldsymbol{x}) - E^2(y|\boldsymbol{x})$, respectivamente.

Capítulo 2

MODELOS TAR.

2.1. Especificación del Modelo TAR.

Un modelo TAR representa la relación, mediante un sistema dinámico, entre dos procesos estocásticos, uno de entrada y uno de salida, denotados por $\{Z_t\}$ y $\{X_t\}$ respectivamente. La expresión que explica la relación anterior es dada por

$$X_t = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i} + h^{(j)}\varepsilon_t, \;\; t \in \mathbb{Z}, \tag{2.1}$$

si Z_t pertenece al intervalo real $B_j = (r_{j-1}, r_j]$ para algún j; $j = 1, \cdots, l$, donde $r_0 = -\infty$, $r_l = \infty$ y $r \geq 2$. Los números reales $a_i^{(j)}$ y el número entero k_j denotan, respectivamente, los coeficientes y el orden autoregresivo de $\{X_t\}$ en el régimen j. Los números reales r_j; $j = 1, \cdots, l-1$; son denominados valores umbrales del proceso $\{Z_t\}$ y son ellos quienes definen los l regímenes para el modelo.

Además, $\{\varepsilon_t\}$ es un proceso de ruido blanco Gaussiano con media cero y varianza 1 independiente de $\{Z_t\}$. Se utiliza la notación TAR$(l, k_1, \cdots, k_l)$ para denotar un modelo TAR con l regímenes y $k_1, \cdots, k_l$ los órdenes de los modelos autorregresivos de los respectivos regímenes. Los parámetros l; $r_1, ..., r_{l-1}$ y $k_1, ..., k_l$ son denominados

parámetros estructurales del modelo TAR. Por otro lado, $\{Z_t\}$ es un proceso exógeno con las propiedades de una cadena de Markov homogénea de orden p, con distribución invariante $f_p(.)$ y kernel de transición $f(.|.)$.

Estos modelos fueron introducidos por Tong (1990) asumiendo que la variable exógena Z_t corresponde a la variable X_t rezagada d veces (con d entero positivo), es decir $Z_t = X_{t-d}$.

2.2. Algunas Propiedades del Modelo TAR.

De acuerdo con Nieto (2008), es posible calcular la distribución marginal de X_t bajo la expresión

$$F_t(x) = \sum_{j=1}^{l} p_j F_{t,j}(x), x \in \mathbb{R}, \tag{2.2}$$

donde $p_j = P(Z_t \in B_j), F_{t,j}(x) = P(X_t \leq x | Z_t \in B_j)$ y $\mathbb{R}$ denota el conjunto de números reales. Además, $p_j = F_z(r_j) - F_z(r_{j-1})$, donde F_z es la función de distribución invariante de la cadena Markov $\{Z_t\}$.

Bajo los supuestos del modelo (2.1), siempre que las raíces del polinomio autorregresivo $\phi_j(y) = a_0^{(j)} - \sum_{i=1}^{k_j} a_i^{(j)} y^i$, $j = 1, \ldots, l$, estén por fuera del circulo unitario, Moreno (2011) demuestra que $\{X_t\}$, es estacionaria en el sentido débil y que, en particular, la función de distribución acumulada condicional para X_t en el regimen j es dada por

$$F_{t,j}(x) = \mathbf{\Phi}_{0,1}\left(\frac{x - \psi_j(1)a_0^{(j)}}{h^{(j)}\sigma_j}\right), \tag{2.3}$$

donde $\psi_j(y) = \frac{1}{\phi_j(y)} = \sum_{i=0}^{\infty} \psi_i^{(j)} y^i$, $\sigma_j^2 = \sum_{i=0}^{\infty} \left(\psi_i^{(j)}\right)^2$ y $\mathbf{\Phi}_{0,1}(.)$ denota la función de distribución acumulada de una normal estandarizada. De lo anterior se tiene que

$$X_t | Z_t \in B_j \sim N\left(\psi_j(1) a_0^{(j)}, (h^{(j)}\sigma_{(j)})^2\right).$$

De (2.2) se pueden calcular las expresiones para el valor esperado y varianza del proceso X_t como sigue:

$$E(X_t) = \sum_{j=1}^{l} p_j \int x_t dF_{t,j}(x_t) = \sum_{j=1}^{l} p_j \mu_{t,j,1}, \tag{2.4}$$

donde $\mu_{t,j,1} = \int x_t dF_{t,j}(x_t)$, $j = 1, \ldots, l$, y

$$E(X_t^2) = \sum_{j=1}^{l} p_j \int x_t^2 dF_{t,j}(x_t) = \sum_{j=1}^{l} p_j \mu_{t,j,2} \tag{2.5}$$

con $\mu_{t,j,2} = \int x_t^2 dF_{t,j}(x_t)$, $j = 1, \ldots, l$. Los valores $\mu_{t,j,i}$, $i = 1, 2$; representan el primer y segundo momento de X_t en el regímen j respectivamente. En consecuencia:

$$VAR(X_t) = E(X_t^2) - [E(X_t)]^2; \text{para cada } t \in \mathbb{Z}.$$

Similarmente se puede obtener la distribución condicional de X_t dado $X_s = x_s$, $s = t-1, \ldots, 1$, como una mezcla de l distribuciones Gaussianas con medias $a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i}$ y varianzas $[h^{(j)}]^2$, $j = 1, \ldots, l$ (Moreno, 2011). Este resultado muestra que X_t es también una cadena de Markov homogénea, con kernel de transición $f_X(x_t | x_{t-1}, \ldots, x_{t-k+1})$, donde $k = \text{máx}\{k_j | j = 1, \ldots, l\}$ (Nieto, 2008).

Otro importante resultado en el modelo TAR, es su función de verosimilitud condicionada a $\mathbf{x}_k = (x_1, ..., x_k)$. Esta función puede expresarse como

$$f(\mathbf{x}, \mathbf{z} | \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = f(\mathbf{x} | \mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) f(\mathbf{z} | \boldsymbol{\theta}_x, \boldsymbol{\theta}_z), \tag{2.6}$$

donde $\boldsymbol{\theta}_x$ representa al vector de parámetros no estructurales en el modelo TAR. Siendo además:

$$f(\mathbf{z} | \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = f(\mathbf{z}_p | \boldsymbol{\theta}_z) f(z_{p+1} | \mathbf{z}_p, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) \ldots f(z_T | \mathbf{z}_{T-1}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) \tag{2.7}$$

y

$$f(\mathbf{x}|\boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = (2\pi)^{-(T-k)/2} \left[\prod_{t:j_t=j} \left\{ h^{(j_t)} \right\}^{-1} \right] \exp\left(-\frac{1}{2} \sum_{t:j_t=j} e_t^2 \right) \tag{2.8}$$

con

$$e_t = \frac{x_t - a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i}}{h^{(j)}},\ t \in \mathbb{Z}$$

2.3. Estimación de los Parámetros en un Modelo TAR.

La estimación de los parámetros no estructurales del modelo TAR se realizan mediante estimación bayesiana utilizando el muestreador de Gibbs, considerando distribuciones a priori informativas y no informativas, como se describe a continuación.

2.3.1. Distribuciones a priori informativas

Sea $\boldsymbol{\theta_x} = (\boldsymbol{\theta_1}, ..., \boldsymbol{\theta_l}, \boldsymbol{h})$, el vector de parámetros no estructurales para $\{X_t\}$, donde $\boldsymbol{\theta_j} = \left(a_0^{(j)}, ..., a_{k_j}^{(j)}\right)'$ y $\boldsymbol{h} = \left(h^{(1)}, ..., h^{(l)}\right)'$. Nieto (2005) consideró las siguientes distribuciones a priori informativas para $\boldsymbol{\theta_j}$ y $\boldsymbol{h}$.

$$\boldsymbol{\theta_j} \sim \boldsymbol{N}\left(\boldsymbol{\theta_0}, V_{0,j}^{-1}\right) \text{ y } \left[h^{(j)}\right]^{-2} \sim \boldsymbol{G}\left(\gamma_{0,j}, \beta_{0,j}\right); j = 1, ..., l$$

De acuerdo con esto, se obtienen las distribuciones a posteriori para $\boldsymbol{\theta_j}$; $j = 1, ..., l$; y para $\boldsymbol{h}$ como sigue.

***Proposición* 2.1.** *Para cada $j = 1, \ldots, l$, sea $\{t : j_t = j\} = \{t_{1,j}, \ldots, t_{n_j,j}\}$ donde n_j denota el número de observaciones de $\boldsymbol{X}$ en el régimen j, la distribución condicional completa de $\boldsymbol{\theta}_j$, dado $\boldsymbol{\theta}_i$; $i \neq j$; $\boldsymbol{h}$; $\boldsymbol{\theta}_z$, $\boldsymbol{x}$ y $\boldsymbol{z}$ es una normal multivariada con media*

$$\boldsymbol{\theta}_j^* = V_j^{-1} \left[\left(1/\left\{ h^{(j)} \right\}^2 \right) \mathbf{W}_j' \mathbf{X}_j + V_{0,j} \boldsymbol{\theta}_{0,j} \right] \tag{2.9}$$

y matriz de varianza

$$V_j^{-1} = \left[\left(1/\left\{h^{(j)}\right\}^2\right)\mathbf{W}_j'\mathbf{W}_j + V_{0,j}\right]^{-1}, \tag{2.10}$$

donde,

$\mathbf{w}_t' = (1, x_{t-1}, \ldots, x_{t-k_j})$, $\mathbf{W}_j = (\mathbf{w}_{t_{1,j}}, \ldots, \mathbf{w}_{t_{n_j,j}})'$ *y* $\mathbf{X}_j = (x_{t_{1,j}}, \ldots, x_{t_{n_j,j}})'$.

Prueba. Ver Nieto (2005).

***Proposición* 2.2.** *La distribución condicional completa para* $\left[h^{(j)}\right]^{-2}$*, dado* $\boldsymbol{\theta}_j$*,* $\boldsymbol{\theta}_z$*,* $\boldsymbol{x}$ *y* $\boldsymbol{z}$*; j=1,. . . , l; corresponde a una distribución Gamma con parámetro de forma* $\gamma_{0,j} + n_j/2$ *y parámetro de escala* $\beta_{0,j} + (1/2)(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j)'(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j)$.

Prueba. Ver Nieto (2005).

2.3.2. Distribuciones a priori no informativas.

En este caso la distribución no informativa que se utiliza es la calculada mediante el método de Jeffreys mencionado en el Capítulo 1, de donde se obtiene:

$$\pi\left(\boldsymbol{\theta_j}\right) \propto 1 \text{ y } \pi\left(\left[h^{(j)}\right]^{-2}\right) \propto \left[h^{(j)}\right]^2.$$

A partir de estas distribuciones a priori, se obtienen distribuciones a posteriori para $\boldsymbol{\theta}_j$ y para $[h^{(j)}]^{-2}$; $j = 1, ..., l$; como sigue.

***Proposición* 2.3.** *Para cada* $j = 1, \ldots, l$*, sea* $\{t : j_t = j\} = \{t_{1,j}, \ldots, t_{n_j,j}\}$ *donde* n_j *denota el número de observaciones de* $\mathbf{X}$ *en el régimen* j*, asumiendo la distribución apriori de Jeffreys especificada para* $\boldsymbol{\theta}_j$*, la distribución condicional completa de* $\boldsymbol{\theta}_j$ *dado* $\boldsymbol{\theta}_i$*;* $i \neq j$*;* $\boldsymbol{h}$ *y* $\boldsymbol{\theta}_z$*, es una normal multivariada con media*

$$\boldsymbol{\theta}_0^{(j)} = \left\{\mathbf{W}_j'\mathbf{W}_j\right\}^{-1}\mathbf{W}_j'\mathbf{X}_j$$

y matriz de varianza

$$\left(V_0^{(j)}\right)^{-1} = \left\{h^{(j)}\right\}^2\left(\mathbf{W}_j'\mathbf{W}_j\right)^{-1},$$

donde $\mathbf{W}_j$ *y* $\mathbf{X}_j$ *son las establecidas en la proposición* (2.1).

Prueba. Ver Anexo B1.

Proposición 2.4. *La distribución a posteriori de* $[h^{(j)}]^{-2}$*, dado* $\boldsymbol{\theta}_j$*,* $\boldsymbol{\theta}_z$*,* $\boldsymbol{x}$ *y* $\boldsymbol{z}$*; j=1,. . . , l; con distribución a priori no informativa; corresponde a una distribución Gamma con parámetro de forma* $n_j/2$ *y parámetro de escala escala* $(1/2)(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j)'(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j)$.

Prueba. Ver Anexo B2.

Al comparar los resultados de las distribuciones a posteriori obtenidas en el presente trabajo con las propuestas por Nieto (2005), se observa que dichas expresiones son muy similares, salvo que para estas ultimas se involucran los parámetros de las distribuciones a priori informativas. Se puede notar que las distribuciones a posteriori obtenidas bajo distribuciones a priori no informativas son aproximaciones de las distribuciones a posteriori resultantes de distribuciones a priori informativas. En el caso de $\boldsymbol{\theta}_j$, se observa que las dos distribuciones condicionales propuestas se aproximan cuando $\boldsymbol{\theta}_{0,j} = \mathbf{0}$ y $V_{0,j}^{-1}$ es lo suficientemente cercana a la matriz de ceros de orden j; es decir, a más baja variabilidad para los $\boldsymbol{\theta}_j$, más cercanas van a ser las dos distribuciones. Mientras que para $[h^{(j)}]^{-2}$, las dos distribuciones a posteriori propuestas se aproximan cuando $\beta_{0,j} = 0$ y $\gamma_{0,j}$ es lo suficientemente cercana a 0, es decir cuando la distribución apriori informativa es una distribución exponencial con tasa de ocurrencia muy cercana a cero.

Se puede mostrar que cuando se tienen distribuciones apriori no informativas de laplace para los parámetros no estructurales, es decir

$$\pi\left(\boldsymbol{\theta_j}\right) \propto 1 \text{ y } \pi\left(\left[h^{(j)}\right]^{-2}\right) \propto 1.$$

Las distribuciones condicionales completas para $[h^{(j)}]^{-2}$ y para $\boldsymbol{\theta}_j$, son muy parecidas a las obtenidas en las proposiciones (2.3) y (2.4) con las a prioris de Jeffreys, con la diferencia de que el parámetro de escala para la la distribución condicional completa de $[h^{(j)}]^{-2}$ es $(n_j/2) + 1$.

Capítulo 3

CÁLCULO DE PRONÓSTICOS CON MODELOS TAR.

3.1. Algunos Resultados.

Argumentando que bajo el criterio del error cuadrático medio mínimo ($ECMM$), el mejor predictor para la observación futura X_{T+h} en un modelo TAR es $E(X_{T+h}|\mathbf{x}_T, \mathbf{z}_T)$, donde $\mathbf{x}_T = (x_1, \ldots, x_T)$, $\mathbf{z}_T = (z_1, \ldots, z_T)$ y h es el horizonte de predicción; Nieto (2008) propone la distribución predictiva para X_{T+h} como sigue.

***Proposición* 3.1.** *Bajo los supuestos del modelo TAR y asumiendo independencia entre Z_t y X_s para todo t y s, tal que $s \leq t$; la densidad predictiva de X_{T+h} dado $\boldsymbol{x}_T$ y $\boldsymbol{z}_T$, para cada $h > 1$, está dada por*

$$p(x_{T+h}|\boldsymbol{x}_T, \boldsymbol{z}_T) = \int p(x_{T+h}|z_{T+h}, \boldsymbol{x}_T, \boldsymbol{z}_T)p(z_{T+h}|\boldsymbol{z}_T)dz_{T+h},$$

donde

$$p(z_{T+h}|\boldsymbol{z}_T) = \int \cdots \int p(z_{T+h}|z_{T+h-1}, \cdots, z_{T+1}, \boldsymbol{z}_T)\times$$
$$p(z_{T+h-1}|z_{T+h-2}, \cdots, z_{T+1}, \boldsymbol{z}_T), \cdots, p(z_{T+1}|\boldsymbol{z}_T)$$

y

$$p(x_{T+h}|z_{T+h}, \boldsymbol{x}_T, \boldsymbol{z}_T) = \int \cdots \int p(x_{T+h}|\boldsymbol{x}_{T+h-1}, \boldsymbol{z}_{T+h}) \times$$
$$p(x_{T+h-1}|\boldsymbol{x}_{T+h-2}, \boldsymbol{z}_{T+h-1}) \times \cdots \times p(x_{T+1}|\boldsymbol{x}_T, \boldsymbol{z}_{T+1}) dx_{T+h-1} \cdots dx_{T+1}$$

***Prueba*:** Ver Nieto (2008).

Nieto (2008) encontró que para parámetros no estructurales conocidos, la función de distribución predictiva $F_{T+h}(x|\mathbf{x}_T, \mathbf{z}_T)$, $h \geq 1$, es una mezcla de las l funciones de distribución $F_{T+h,j}(x|z_{T+h} \in B_j, \mathbf{x}_T, \mathbf{z}_T)$ con pesos $p_j(h) = P(Z_{T+h} \in B_j|\mathbf{x}_T, \mathbf{z}_T)$, $j = 1, \ldots, l$. Además, a partir de la distribución predictiva se pueden realizar pronósticos para la variable X, en el horizonte h, dada la información observada hasta el tiempo T, mediante la siguiente expresión:

$$x_{T+h|T} = E(X_{T+h}|\mathbf{x}_T, \mathbf{z}_T) = \sum_{j=1}^{l} p_j(h) E(X_{T+h}|z_{T+h} \in B_j, \mathbf{x}_T, \mathbf{z}_T) \tag{3.1}$$

Se observa que este pronóstico se puede obtener mediante un promedio ponderado de las predicciones condicionadas a cada regímen.

De acuerdo con la metodología establecida por Nieto (2008), se puede calcular el error cuadrático medio para el predictor de X_{T+h}, como sigue:

$$E[(X_{T+h} - X_{T+h|T})^2|\mathbf{x}_T, \mathbf{z}_T] = \int (x_{T+h} - x_{T+h|T})^2 f(x_{T+h}|\mathbf{x}_T, \mathbf{z}_T) dx_{T+h} \tag{3.2}$$

$$= \sum_{j=1}^{l} p_j(h) MSE_j \left(X_{T+h|T}^{(j)} \right), \tag{3.3}$$

donde

$$MSE_j \left(X_{T+h|T}^{(j)} \right) = \int (x_{T+h} - x_{T+h|T})^2 dF_{T+h,j}(x_{T+h}|z_{T+h} \in B_j, \mathbf{x}_T, \mathbf{z}_T) \tag{3.4}$$

Luego, el error cuadrático medio es un promedio ponderado de los errores cuadráticos por regímen con pesos $p_j(h)$.

En los resultados obtenidos por Nieto (2008) para obtener pronósticos para X_{t+h}, se observa, que aún cuando las estimaciones de los parámetros no estructurales son Bayesianas, la densidad predictiva propuesta para X_{t+h} no considera la variabilidad inicial en estos parámetros, como se muestra a continuación.

3.2. Predicción Bayesiana.

Siguiendo el enfoque Bayesiano, la distribución predictiva exacta para X_{T+h} en un modelo TAR, se define a continuación.

***Proposición* 3.2.** *Bajo los supuestos del modelo TAR y siendo $\boldsymbol{\theta}_j$ el vector de parámetros no estructurales del modelo en el regímen j, la densidad predictiva Bayesiana de X_{T+h}, $h \geq 1$, dados los vectores de datos $\boldsymbol{x}_T$ y $\boldsymbol{z}_T$, es dada por*

$$f_{T+h}(x_{T+h}|\boldsymbol{x}_T, \boldsymbol{z}_T) = \sum_{j=1}^{l} p\left(x_{T+h}|B_j, \boldsymbol{x}_T, \boldsymbol{z}_T\right) p_j(h), \tag{3.5}$$

donde

$$p_j(h) = P(Z_{t+h} \in B_j|\boldsymbol{x}_T, \boldsymbol{x}_z), h = 1, 2, ...; j = 1, 2, ..., l.$$

y

$$p(x_{T+h}|B_j, \boldsymbol{x}_T, \boldsymbol{x}_z) = \int_{\boldsymbol{\Theta}_j} p(x_{T+h}, \boldsymbol{\theta}^{(j)}|B_j, \boldsymbol{x}_T, \boldsymbol{x}_z) d\boldsymbol{\theta}^{(j)} =$$

$$\int_{\boldsymbol{\Theta}_j} p(x_{T+h}|\boldsymbol{\theta}^{(j)}, B_j, \boldsymbol{x}_T, \boldsymbol{x}_z) \pi(\boldsymbol{\theta}^{(j)}|B_j, \boldsymbol{x}_T, \boldsymbol{x}_z) d\boldsymbol{\theta}^{(j)}, \tag{3.6}$$

con

$$p(x_{T+h}|\boldsymbol{\theta}^{(j)}, B_j, \boldsymbol{x}_T, \boldsymbol{x}_z) = \int ... \int p(x_{T+h}|\boldsymbol{\theta}^{(j)}, B_j, \boldsymbol{x}_{T+h-1}) \times$$

$$p(x_{T+h-1}|\boldsymbol{\theta}^{(j)}, B_j, \boldsymbol{x}_{T+h-2}) \times ... \times p(x_{T+1}|\boldsymbol{\theta}^{(j)}, B_j, \boldsymbol{x}_T) dx_{T+1}...dx_{T+h-1} \tag{3.7}$$

siendo $\mathbf{x}_{T+i} = (\mathbf{x}_T, x_{T+1}, ..., x_{T+i})$.

Prueba. Ver Anexo B3.

Ahora bien, $p\left(x_{T+h}|B_j, \mathbf{x}_T, \mathbf{x}_z\right)$ es el promedio de $p(x_{T+h}|\boldsymbol{\theta}^{(j)}, B_j, \mathbf{x}_T, \mathbf{x}_z)$ con respecto a la distribución a posteriori de $\boldsymbol{\theta}^{(j)}$ en el regímen j, donde el cómputo de $p(x_{T+h}|\boldsymbol{\theta}^{(j)}, B_j, \mathbf{x}_T, \mathbf{z}_T)$ se tiene generando primero a x_{T+1} de $p(x_{T+1}|\boldsymbol{\theta}^{(j)}, B_j, \mathbf{x}_T)$, y así sucesivamente hasta generar a x_{T+h} de $p(x_{T+h}|\boldsymbol{\theta}^{(j)}, B_j, x_{T+1}, ..., x_{T+h-1}, \mathbf{x}_T)$. Además, para $i = 1, 2, ..., h$, se tiene que $(x_{T+i}|\boldsymbol{\theta}^{(j)}, B_j, \mathbf{x}_{T+i-1})$ posee distribución normal con media $a_0^{(j)} + \sum_{m=1}^{k_j} a_m^{(j)} x_{T+i-m}$ y varianza $\left[h^{(j)}\right]^2$. Este proceso se realiza para cada uno de los l regímenes, para finalmente generar x_{T+i} de $f(x_{T+i}|\mathbf{x}_T, \mathbf{z}_T)$ mediante el promedio ponderado de los valores obtenidos a partir de $p(x_{T+i}|B_j, \mathbf{x}_T, \mathbf{x}_z)$ y los pesos $p_j(h)$.

De la distribución predictiva Bayesiana se pueden obtener pronósticos para la variable X, en el horizonte h, dados el vector de parámetros no estructurales en el regímen j y la información observada hasta el tiempo T, mediante la siguiente expresión.

$$x_{T+h|T} = E(X_{T+h}|\mathbf{x}_T, \mathbf{z}_T) = \sum_{j=1}^{l} p_j(h) E(X_{T+h}|B_j, \mathbf{x}_T, \mathbf{z}_T) = \sum_{j=1}^{l} p_j(h) x_{T+h|B_j,T} \quad (3.8)$$

Se observa que tanto la densidad predictiva como el pronostico para X_t involucran la incertidumbre de los parámetros no estructurales del modelo TAR, lo cual no ocurre con la distribución propuesta por Nieto (2008).

A partir de esta distribución predictiva, se puede calcular la varianza de la distribución predictiva Bayesiana (pérdida cuadrática Bayesiana) mediante expresiones similares a las dadas en (3.3) y (3.4), como sigue.

$$V(X_{T+h}|\mathbf{x}_T, \mathbf{z}_T) = \int (x_{T+h} - x_{T+h|T})^2 f(x_{T+h}|\mathbf{x}_T, \mathbf{z}_T) dx_{T+h} = \sum_{j=1}^{l} p_j(h) V_{T+h|B_j,T}, \quad (3.9)$$

donde

$$V_{T+h|B_j,T} = E\left[(X_{T+h} - X_{T+h|B_j,T})^2 | B_j, \mathbf{x}_T, \mathbf{z}_T\right].$$

Capítulo 4

APLICACIONES.

4.1. Aplicación en Datos Simulados.

Inicialmente se aplica el método de predicción bayesiana propuesto en modelos TAR con dos regímenes, variando los respectivos ordenes autorregresivos y posteriormente se realiza el análisis aumentando el número de regímenes. Adicionalmente, se realiza una comparación con la propuesta hecha por Nieto (2008) en el modelo TAR simulado con tres regímenes. En cada uno de los modelos, se simulan series de longitud 1000 para los procesos $\{X_t\}$ y $\{Z_t\}$. Para la distribución de los parámetros se toman 30000 muestras y para las predicciones se obtienen 20000 muestras. Se eliminaron las 1000 primeras muestras, tanto para los parámetros, como para las predicciones.

Los resultados de la simulación de las series, estimación de los parámetros y distribuciones predictivas se realizaron a través del programa R-2.12. Para obtener las muestras a partir de la distribución de los parámetros se utilizó el muestreador de *Gibbs* y las predicciones se calcularon utilizando dichas muestras y a partir de la Proposición 3.2.

Para determinar la convergencia del muestreador de *Gibbs* se realiza el test KPSS (Kwiatkowski-Phillips-Schmidt-Shin), el cual se obtiene mediante la función kpss.test del paquete tseries en el lenguaje R, en esta se observa que el valor p es 0.1 para cada una de las muestras simuladas , lo cual indica que la hipótesis nula de estacionariedad no se rechaza en ninguna de las muestras simuladas para los parámetros no estructurales de los distintos modelos TAR considerados; también se realizan las pruebas de PP(Phillips-Perron) y ADF(Dickey Fuller Aumentado), para las cuales el valor p es 0,01, lo cual corrobora lo encontrado con la prueba KPSS, ya que se rechaza la hipótesis nula a favor de la hipótesis alternativa de estacionariedad.

En las aplicaciones para modelos TAR de dos regímenes, $\{Z_t\}$ es un proceso $AR(1)$ dado por $Z_t = 1.8 + 0.6Z_{t-1} + a_t$ y $r_1 = 4$, siendo $\{a_t\}$ un proceso estocástico de ruido blanco Gaussiano con media 0 y varianza 1, independiente de $\{\varepsilon_t\}$, de aquí que $p_1 = 0.34$ y $p_2 = 0.66$. Evidentemente $\{Z_t\}$ tiene una distribución invariante, la cual es $N(4.5,1.56)$ y la densidad del kernel de transición $f_1(z_t|z_{t-1})$ es la de una $N(1.8+0.6z_{t-1}, 1)$.

4.1.1. TAR(2;0,0).

En esta aplicación se simula un modelo TAR(2;0,0) dado por:

$$X_t = \begin{cases} 0.7 + 1.4\varepsilon_t, \text{ si } Z_t < 4 \\ 0.9 + 1.8\varepsilon_t, \text{ si } Z_t \geq 4. \end{cases}$$

En la Figura 1 se encuentran los gráficos de las dos series involucradas en este modelo. Se tiene que las medias de X_t en los dos regímenes están dadas por 0.7 y 0.9 respectivamente y las varianzas son 1.96 y 3.24, respectivamente. Utilizando las formulas (2.4) y (2.5), se tiene que $E(X_t)$=0.83 y $V(X_t)$=2.81, para todo t.

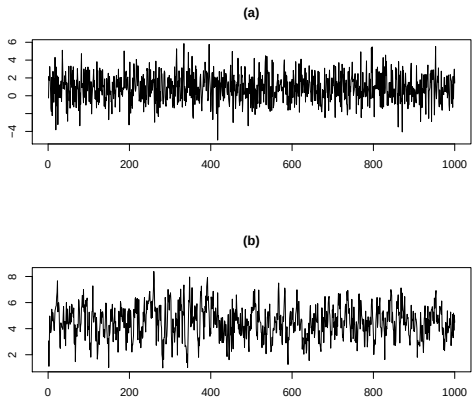

Figura 1. Simulación de un modelo TAR(2;0,0): (a) Variable Z_t (b) Variable X_t.

La Tabla 1 muestra las estimaciones Bayesianas de los respectivos parámetros no estructurales de dicho modelo.

Tabla 1. Estimación de los parámetros en la simulación de un TAR(2;0,0).

$\boldsymbol{\theta}$	$\widehat{\boldsymbol{\theta}}^{(a)}$	$\widehat{\boldsymbol{\theta}}^{(b)}$	$\boldsymbol{\theta}^{(1)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$
$a_0^{(1)}$	0.73	0.71	0.7	[0.65,0.81]	[0.66,0.76]	[0.66,0.81]	[0.64,0.78]
$[h^{(1)}]^2$	2.02	2.05	1.96	[1.79,2.29]	[1.79,2,25]	[1.77,2.27]	[1.79,2.25]
$a_0^{(2)}$	0.93	0.91	0.9	[0.89,0.97]	[0.89,0.96]	[0.85,0.77]	[0.89,0.96]
$[h^{(2)}]^2$	2.98	2.99	3.24	[2.72 3.26]	[2.73,3.27]	[2.71,3.25]	[2.73,3.26]

$^{(1)}$:Valor simulado $^{(a)}$:A priori no informativa $^{(b)}$: A priori informativa

Se observa que las estimaciones de los parámetros, cuando se tienen las diferentes distribuciones a priori, son muy parecidas entre si y al valor simulado del mismo. Por otro lado, existe mucha similitud entre los intervalos de credibilidad y regiones de alta densidad (**IC** y **RAD**, respectivamente), lo cual indica que no hay un comportamiento multimodal en las distribuciones muestreadas para los diferentes parámetros involucrados en el modelo. Lo anterior se puede corroborar en la Figura 2, donde se observan la distribuciónes a posteriori y regiónes de alta densidad para el parámetro $a_0^{(1)}$, con distribución a priori no informativa e informativa. Este mismo comportamiento es dado para los diferentes parámetros.

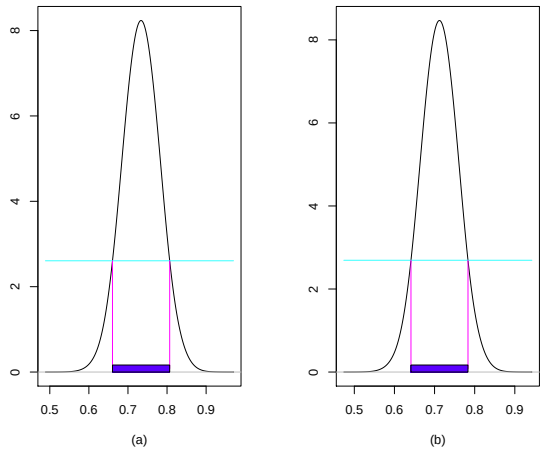

Figura 2. Distribución a posteriori para el parámetro $a_0^{(1)}$, con apriori no informativa (a) e informativa (b).

Aunque los valores en las estimaciones de los parámetros no estructurales no difieren con las distintas distribuciones a priori, se puede observar que los intervalos de credibilidad y las regiones de alta densidad, son de mayor amplitud cuando se consideran distribuciones a priori no informativas.

Tabla 2. Predicciones en el ejemplo teórico TAR(2;0,0).

t	$x_t{}^{(1)}$	$x_{t\|990}^{(a)}$	$x_{t\|990}^{(b)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$	$\mathbf{RPC}^{(a)}$	$\mathbf{RPC}^{(b)}$
991	1.25	2.16	2.02	[-1.13,2.76]	[-1.01,2.72]	[-1.10,2.78]	[-1.08,2.65]	1.45	1.42
992	0.33	0.79	0.81	[-1.06,2.61]	[-1.21,2.62]	[-1.08,2.59]	[-1.24,2.62]	1.45	1.46
993	-1.09	0.88	0.87	[-1.05,2.61]	[-1.11,2.63]	[-1.04,2.62]	[-1.13,2.66]	1.46	1.46
994	0.32	0.90	0.88	[-1.33,3.03]	[-1.22,2.99]	[-1.32,3.02]	[-1.24,3.05]	1.49	1.48
995	1.93	0.86	0.88	[-1.13,2.74]	[-1.04,2.82]	[-1.13,2.74]	[-1.09,2.74]	1.51	1.50
996	3.63	0.83	0.87	[-1.13,2.76]	[-1.22,2.83]	[-1.13,2.76]	[-1.14,2.79]	1.51	1.50
997	-0.57	0.87	0.85	[-1.39,3.15]	[-1.42,3.22]	[-1.43,3.11]	[-1.43,3.23]	1.56	1.55
998	1.18	0.88	0.89	[-1.06,2.62]	[-0.98,2.59]	[-1.06,2.62]	[-0.95,2.61]	1.60	1.61
999	-0.06	0.82	0.83	[-1.17,2.85]	[-1.15,2.80]	[-1.18,2.84]	[-1.18,2.76]	1.65	1.63
1000	0.31	0.83	0.85	[-1.19,2.85]	[-1.24,2.92]	[-1.19,2.85]	[-1.21,2.94]	1.78	1.77

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa.

En la Tabla 2 se encuentran las predicciones estimadas, donde el periodo de la muestra fue 1-990 y el horizonte de predicción fue 10, así como los intervalos de credibilidad

del 90 % (**IC**), regiones de alta densidad del 90 % (**RAD**) y raíz de la varianza predictiva (**RPC**). Todas las predicciones se realizaron considerando distribuciones a priori informativas (b) y no informativas (a).

En este caso se observa, que las predicciones considerando las distintas distribuciones a priori son similares, al considerar distribuciones no informativas e informativas; al igual que los intervalos de credibilidad, regiones de alta densidad y raíces de las varianzas predictivas. También es de resaltar que las raíces de las varianzas predictivas muestran una tendencia creciente a medida que aumenta el horizonte de predicción, esto indica que aumenta la variabilidad a medida que aumenta la incertidumbre, como es de esperar.

Los verdaderos valores de las observaciones se encuentren incluidos en los intervalos creíbles. La similitud entre los intervalos de credibilidad y las regiones de alta densidad, indican que la distribución de las 10 predicciones muestreadas no poseen un comportamiento multimodal.

4.1.2. TAR(2;0,1).

En este caso se simula un modelo TAR(2;0,1) como sigue:

$$X_t = \begin{cases} 0.7 + 1.4\varepsilon_t, \text{ si } Z_t < 4 \\ 0.9 - 0.5X_{t-1} + 1.8\varepsilon_t, \text{ si } Z_t \geq 4 \end{cases}$$

La Figura 3 muestra el comportamiento de las series X_t y Z_t. El valor esperado y varianza para este modelo son dados por 0.61 y 4.01 respectivamente. Esta variabilidad es aportada más que todo por el segundo regímen, ya que en este se observa una mayor varianza (5.06) y se sabe que p_2 =0.66.

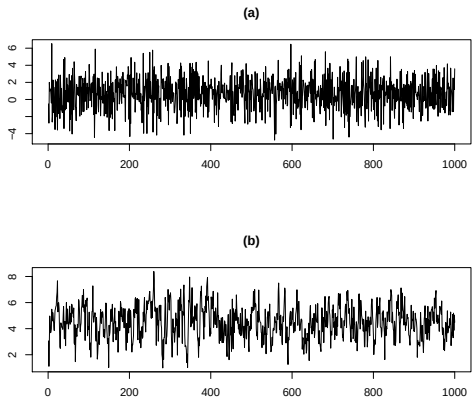

Figura 3. Simulación de un modelo TAR(2;0,1): (a) Variable Z_t (b) Variable X_t.

Los valores de los parámetros para este modelo (Tabla 3) son similares a los dados en la Tabla 1, con excepción del coeficiente autorregresivo en el segundo regímen, el cual no se considera en el modelo simulado anteriormente. Se puede ver que los valores de las estimaciones no difieren a los obtenidos en la Tabla 1, así como tampoco los intervalos de credibilidad y alta densidad, esto puede indicar que el orden autoregresivo puede no afectar la estimación de los parámetros.

Tabla 3. Estimación de los parámetros de la simulación de un TAR(2;0,1).

$\boldsymbol{\theta}$	$\widehat{\boldsymbol{\theta}}^{(a)}$	$\widehat{\boldsymbol{\theta}}^{(b)}$	$\boldsymbol{\theta}^{(1)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$
$a_0^{(1)}$	0.71	0.702	0.7	[0.65,0.77]	[0.68,0.77]	[0.65,0.77]	[0.68,0.75]
$[h^{(1)}]^2$	1.95	1.95	1.96	[1.73,2.19]	[1.75,2.17]	[1.72,2.17]	[1.72,2.17]
$a_0^{(2)}$	0.88	0.88	0.9	[0.84,0.92]	[0.84,0.92]	[0.84,0.92]	[0.84,0.92]
$a_1^{(2)}$	-0.59	-0.61	-0.6	[-0.62,-0.58]	[-0.61,-0.58]	[-0.62,-0.58]	[-0.61,-0.58]
$[h^{(2)}]^2$	3.31	3.32	3.24	[2.96,3.58]	[2.97,3.42]	[2.94,3.56]	[2.97,3.46]

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

Para este modelo, igual que para el modelo anterior, las predicciones difieren un poco a medida que aumenta el horizonte de predicción (ver Tabla 4), también se puede ver que los intervalos de credibilidad y regiones de alta densidad son de mayor longitud y las raíces de las varianzas son más grandes en las predicciones realizadas con base

en las distribuciones a priori no informativas. Se puede ver que los valores observados en la serie se encuentran incluidos en el intervalo de credibilidad (**IC**).

Tabla 4. Predicciones en el ejemplo teórico TAR(2;0,1).

t	$x_t{}^{(1)}$	$x^{(a)}_{t\|990}$	$x^{(b)}_{t\|990}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$	$\mathbf{RPC}^{(a)}$	$\mathbf{RPC}^{(b)}$
991	1.33	0.64	0.62	[-1.03,2.66]	[-0.88,2.09]	[-1.00,2.68]	[-0.88,2.11]	1.23	0.90
992	2.32	0.66	0.63	[-1.07,2.61]	[-0.90,2.10]	[-1.07,2.60]	[-0.91,2.12]	1.25	0.91
993	2.15	0.66	0.64	[-1.15,2.60]	[-0.92,2.16]	[-1.13,2.52]	[-0.92,2.120]	1.67	1.08
994	1.38	0.67	0.67	[-1.22,2.57]	[-0.95,2.16]	[-1.22,2.58]	[-0.94,2.15]	1.89	1.12
995	0.51	0.66	0.68	[-1.24,2.58]	[-0.96,2.19]	[-1.23,2.59]	[-0.96,2.17]	1.96	1.21
996	-0.42	0.67	0.67	[-1.55,2.78]	[-0.98,2.34]	[-1.54,2.80]	[-1.02,2.15]	1.86	1.23
997	0.29	0.67	0.65	[-1.82,3.02]	[-0.99,3.22]	[-1.81,3.03]	[-1.09,3.20]	1.95	1.24
998	-0.89	0.64	0.65	[-1.72,2.66]	[-1.08,2.38]	[-1.72,2.67]	[-1.20,2.41]	1.98	1.45
999	-0.50	0.61	0.63	[-1.76,2.73]	[-1.12,2.75]	[-1.75,2.74]	[-1.13,2.75]	2.06	1.57
1000	-0.66	0.61	0.62	[-2.15,3.10]	[-1.18,2.93]	[-2.13,3.12]	[-2.15,3.01]	2.11	1.76

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

4.1.3. TAR(2;1,1).

Se considera un modelo TAR con dos regímenes de orden uno, el modelo para X_t se define a continuación:

$$X_t = \begin{cases} 0.7 + 0.2X_{t-1} + 1.4\varepsilon_t, \text{ si } Z_t < 4 \\ 0.9 - 0.5X_{t-1} + 1.8\varepsilon_t, \text{ si } Z_t \geq 4 \end{cases}$$

La Figura 4 muestra el comportamiento de las dos series. El valor esperado y la varianza para este modelo están dados por 0.67 y 4.04 respectivamente. Los valores de las estimaciones, los intervalos de credibilidad y regiones de alta densidad, no difieren de los valores calculados en los modelos anteriores. Lo anterior ilustra que el orden autoregresivo de los regímenes puede no estar influyendo en la estimación de los parámetros (ver Tabla 5).

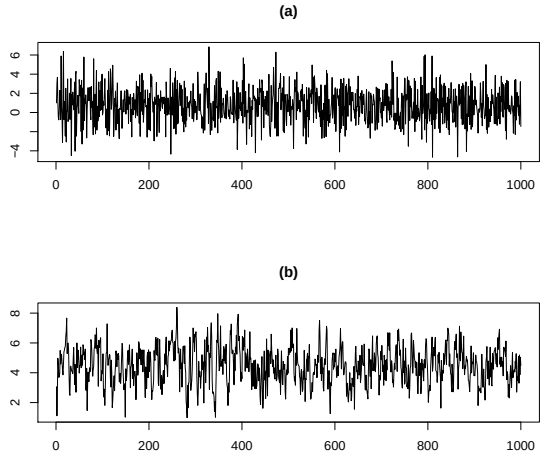

Figura 4. Simulación de un modelo TAR(2,1,1): (a) Variable Z_t (b) Variable X_t.

Tabla 5. Estimación de los parámetros para la simulación TAR(2;1,1).

$\boldsymbol{\theta}$	$\widehat{\boldsymbol{\theta}}^{(a)}$	$\widehat{\boldsymbol{\theta}}^{(b)}$	$\boldsymbol{\theta}^{(1)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$
$a_0^{(1)}$	0.72	0.74	0.7	[0.66,0.82]	[0.66,0.81]	[0.66,0.82]	[0.66,0.81]
$a_1^{(1)}$	0.23	0.22	0.2	[0.23,0.32]	[0.23,0.32]	[0.23,0.32]	[0.24,0.31]
$[h^{(1)}]^2$	1.94	1.92	1.96	[1.62,2.08]	[1.62,2.00]	[1.61,2.06]	[1.63,2.01]
$a_0^{(2)}$	0.91	0.91	0.9	[0.87,0.94]	[0.88,0.95]	[0.87,0.94]	[0.88,0.96]
$a_1^{(2)}$	-0.64	-0.63	-0.6	[-0.61,-0.58]	[-0.61,-0.58]	[-0.61,-0.58]	[-0.61,-0.58]
$[h^{(2)}]^2$	3.42	3.44	3.24	[3.13,3.75]	[3.13,3.71]	[3.11,3.73]	[3.12,3.71]

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

Las predicciones para este modelo (Tabla 6) muestran que a medida que aumentan los ordenes autoregresivos en los regímenes, aumentan los intervalos de credibilidad (**IC**), regiones de alta densidad (**RAD**) y la raíz de las varianzas predictivas (**RPC**), así, aunque el número de parámetros no influye en las estimaciones de los mismos, si afecta en las predicciones. Por otro lado se observa más marcada el aumento de longitud en los intervalos y los cuadrados medios cuando las estimaciones se realizan vía distribuciones a priori no informativas, con respecto a las informativas, aún cuando el valor las predicciones no difieren.

Tabla 6. Predicciones en el ejemplo teórico TAR(2;1,1).

t	$x_t{}^{(1)}$	$x^{(a)}_{t\|990}$	$x^{(b)}_{t\|990}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$	$\mathbf{RPC}^{(a)}$	$\mathbf{RPC}^{(b)}$
991	2.21	0.64	0.60	[-1.13,2.76]	[-0.91,2.80]	[-1.10,2.78]	[-0.92,2.81]	1.20	1.09
992	-2.19	0.66	0.62	[-1.18,2.50]	[-0.95,2.50]	[-1.19,2.50]	[-0.96,2.51]	1.20	1.12
993	0.05	0.66	0.67	[-1.15,2.66]	[-0.95,2.63]	[-1.16,2.66]	[-0.95,2.54]	1.22	1.15
994	1.92	0.67	0.65	[-1.18,2.69]	[-0.97,2.64]	[-1.18,2.69]	[-0.97,2.61]	1.28	1.15
995	-0.59	0.66	0.62	[-1.19,2.56]	[-0.98,2.51]	[-1.19,2.56]	[-0.98,2.52]	1.30	1.18
996	0.65	0.67	0.69	[-1.20,2.50]	[-0.98,2.47]	[-1.20,2.50]	[-0.98,2.48]	1.30	1.21
997	0.29	0.67	0.63	[-1.29,2.51]	[-0.99,2.51]	[-1.27,2.52]	[-1.99,2.51]	1.32	1.22
998	-0.43	0.64	0.64	[-1.33,2.50]	[-1.14,2.51]	[-1.36,2.53]	[-1.13,2.51]	1.33	1.24
999	-0.86	0.61	0.60	[-1.99,3.21]	[-1.17,3.22]	[-1.99,3.21]	[-1.17,3.20]	1.40	1.28
1000	1.11	0.61	0.59	[-2.10,3.30]	[-1.21,3.29]	[-2.12,3.29]	[-1.20,3.26]	1.45	1.32

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

4.1.4. TAR(2;1,2).

Finalmente, se observan las estimaciones para un TAR(2,1,2), definido por el siguiente modelo (ver Figura 5):

$$X_t = \begin{cases} 0.7 + 0.2X_{t-1} + 1.4\varepsilon_t, \text{ si } Z_t < 4 \\ 0.9 - 0.5X_{t-1} - 0.8X_{t-2} + 1.8\varepsilon_t, \text{ si } Z_t \geq 4 \end{cases}$$

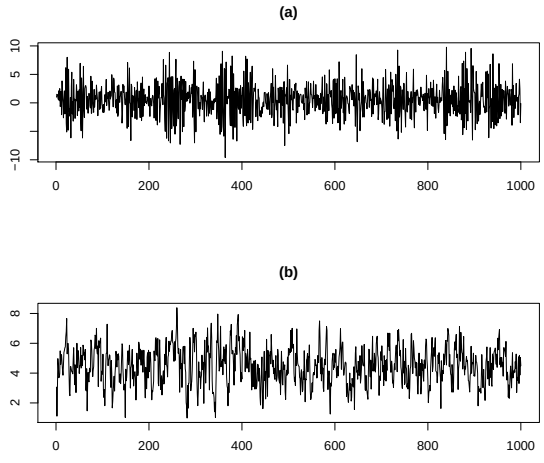

Figura 5. Simulación de un modelo TAR(2,1,2): (a) Variable Z_t (b) Variable X_t.

El valor esperado y la varianza para este modelo son 0.56 y 14.95 respectivamente, donde la mayor parte de la variabilidad es aportada por el regímen dos. En la Tabla 7 se observan las estimaciones para los parámetros no estructurales asociados a este modelo.

Tabla 7. Estimación de los parámetros para la simulación TAR(2;1,2).

$\boldsymbol{\theta}$	$\widehat{\boldsymbol{\theta}}^{(a)}$	$\widehat{\boldsymbol{\theta}}^{(b)}$	$\boldsymbol{\theta}^{(1)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$
$a_0^{(1)}$	0.69	0.710	0.7	[0.61,0.77]	[0.63,0.77]	[0.61,0.77]	[0.62,0.76]
$a_1^{(1)}$	0.30	0.30	0.3	[0.26,0.34]	[0.27,0.34]	[0.26,0.34]	[0.27,0.31]
$[h^{(1)}]^2$	2.12	2.19	2.24	[1.86,2.42]	[1.88,2.42]	[1.84,2.40]	[1.88,2.43]
$a_0^{(2)}$	0.92	0.90	0.9	[0.89,0.96]	[0.89,0.91]	[0.89,0.96]	[0.892,0.92]
$a_1^{(2)}$	-0.60	-0.61	-0.6	[-0.61,-0.59]	[-0.61,-0.60]	[-0.61,-0.60]	[-0.61,-0.60]
$a_2^{(2)}$	-0.81	-0.81	-0.8	[-0.82,-0.80]	[-0.82,-0.80]	[-0.82,-0.80]	[-0.82,-0.79]
$[h^{(2)}]^2$	3.30	3.25	3.24	[3.01,3.60]	[3.02,3.50]	[3.00,3.58]	[3.02,3.51]

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

En las predicciones para este modelo (Tabla 8.) son coherentes con lo encontrado en las simulaciones anteriores, en el sentido de que se observa mayor variabilidad cuando se utiliza distribuciones a priori no informativas.

Se observa que los valores difieren más entre los distintos horizontes de predicción, así como los intervalos son de mayor longitud, esto puede estar asociado al hecho de que la varianza de X_t es superior en este modelo con respecto a los anteriores.

Tabla 8. Predicciones en el ejemplo teórico TAR(2;1,2).

$\boldsymbol{t}$	$\boldsymbol{x_t}^{(1)}$	$\boldsymbol{x}_{\boldsymbol{t}\mid\mathbf{990}}^{(a)}$	$\boldsymbol{x}_{\boldsymbol{t}\mid\mathbf{990}}^{(b)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$	$\mathbf{RPC}^{(a)}$	$\mathbf{RPC}^{(b)}$
991	-1.00	1.65	1.65	[-1.60,2.77]	[-1.65,2.80]	[-1.60,2.70]	[-1.67,2.70]	1.32	1.23
992	-0.27	1.45	1.45	[0.33,3.993]	[0.24,3.99]	[0.33,3.99]	[0.23,3.97]	1.35	1.26
993	1.42	1.52	1.52	[-1.63,2.71]	[-1.73,2.71]	[-1.61,2.72]	[-1.72,2.72]	1.76	1.50
994	-8.77	2.18	2.18	[-3.49,1.85]	[-2.54,1.85]	[-3.48,1.86]	[-2.53,1.86]	1.85	1.62
995	1.28	1.95	1.95	[-2.92,1.60]	[-2.97,1.60]	[-2.93,1.58]	[-2.97,1.60]	1.98	1.85
996	6.73	2.23	2.22	[-4.47,1.16]	[-4.65,1.16]	[-4.46,1.17]	[-4.64,1.16]	2.32	1.88
998	-4.95	1.96	1.96	[-7.79,1.80]	[-7.99,1.80]	[-1.79,1.80]	[-1.94,1.81]	2.51	1.96
999	-0.23	-0.02	-0.01	[-7.72,1.46]	[-7.85,1.46]	[-7.71,1.47]	[-7.82,1.46]	2.60	2.21
1000	6.16	2.20	2.19	[-7.93,2.51]	[-8.12,2.51]	[-7.92,2.51]	[-7.99,2.51]	2.72	2.37

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

4.1.5. TAR(3;1,2,2).

De acuerdo con lo especificado, en este caso se asume un modelo TAR con tres regímenes, en el cual los valores simulados de los coeficientes autoregresivos en los diferentes regímenes, son distintos a los dados en las simulaciones para modelos con dos regímenes. La aplicación es hecha bajo el siguiente modelo TAR:

$$X_t = \begin{cases} 0.9 + 0.2X_{t-1} + 1.5\varepsilon_t, \text{ si } Z_t < 2 \\ 0.8X_{t-1} - 0.4X_{t-2} + 1.8\varepsilon_t, \text{ si } 2 \leq Z_t < 5.2 \\ -0.6X_{t-2} + 1.2\varepsilon_t, \text{ si } Z_t \geq 5.2 \end{cases}$$

La variable Z_t es un proceso $AR(1)$, donde $Z_t = 2 + 0.8Z_{t-1} + a_t$, con $a_t \sim N(0,1)$ independiente de ε_t.Ademas, $Z_t \sim N(2.5, 2.78)$ y $Z_t|Z_{t-1} \sim N(2 + 0.8Z_{t-1}, 1)$. De acuerdo con la definición de $\{Z_t\}$, se tiene que $p_1 = 0.5$, $p_2 = 0.24$ y $p_3 = 0.26$, mientras que los valores esperados para los tres regímenes son 1.125, 0 y 0, respectivamente y las varianzas 2.34, 16.2 y 2.25, respectivamente por lo cual, $E(X_t) = 0.56$ y $V(X_t) = 5.64$.

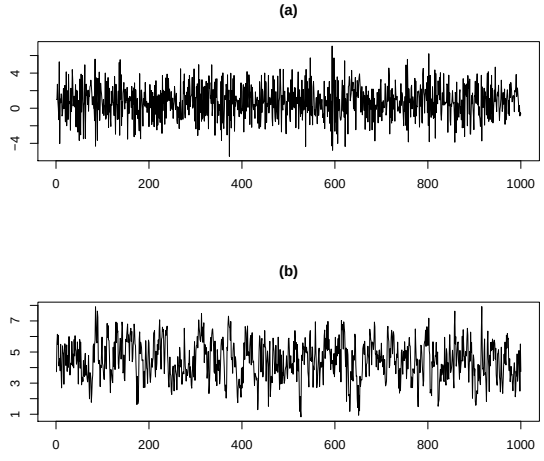

Figura 6. Simulación de un modelo TAR(3;1,2,2): (a) Variable Z_t (b) Variable X_t.

En la Figura 6 se encuentran las gráficas de los procesos $\{X_t\}$ y $\{Z_t\}$. Las estimaciones de los parámetros del modelo anterior son dadas en la Tabla 9. Los intervalos creíbles y las regiones de alta densidad dados en distribuciones a priori informativas y no informativas son muy similares, siendo estos de menor longitud al considerar distribuciones a priori informativas.

Tabla 9. Estimación de los parámetros en el ejemplo teórico TAR(3;1,2,2).

$\boldsymbol{\theta}$	$\widehat{\boldsymbol{\theta}}^{(a)}$	$\widehat{\boldsymbol{\theta}}^{(b)}$	$\boldsymbol{\theta}^{(1)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$
$a_0^{(1)}$	0.89	0.91	0.9	[0.82,0.97]	[0.81,0.95]	[0.78,0.96]	[0.83,0.98]
$a_1^{(1)}$	0.21	0.203	0.2	[0.17,0.25]	[0.18,0.23]	[0.16,0.27]	[0.12,0.21]
$[h^{(1)}]^2$	2.1	2.1	2.25	[1.87,2.36]	[1.90,2.30]	[1.90,2.27]	[2.05,2.69]
$a_1^{(2)}$	0.82	0.81	0.8	[0.79,0,83]	[0.79,0,82]	[0.79, 0.83]	[0.80,0.85]
$a_2^{(2)}$	-0.42	-0.42	-0.4	[-0.37,-0.44]	[-0.38,-0.43]	[-0.37,-0.43]	[-0.45,-0.40]
$[h^{(2)}]^2$	3.35	3.35	3.24	[3.13,3.96]	[3.10,3.56]	[3.11,3.93]	[2,94,3.74]
$a_2^{(3)}$	-0.58	-0.602	-0.6	[-0.66,0.53]	[-0.65,0.50]	[-0.61,-0.44]	[-0.61,-0.42]
$[h^{(3)}]^2$	1.42	1.43	1.44	[1.25,1.62]	[1.23,1.59]	[1.21,1.45]	[1.24,1.51]

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

Las varianzas de las predicciones son superiores a las encontradas en las predicciones hechas para modelos de dos regímenes, comportamiento que se había observado en los demás modelos, donde se observaba que a medida que aumentaban los ordenes autoregresivos en los umbrales, también aumenta la variabilidad en las predicciones, al parecer el aumento de los regímenes puede influenciar en la variabilidad de las estimaciones, también se puede ver que las varianzas predictivas para este modelo, son superiores a las calculadas en los modelos de dos regímenes.

Al comparar los resultados anteriores con la metodología expuesta por en Nieto (2008) (ver Anexo A1), se observa que las predicciones dadas en la Tabla 10 presentan intervalos de credibilidad y los cuadrados medios presentan una menor longitud (tanto para distribuciones a priori informativas como no informativas). Además, se puede ver que los intervalos creíbles obtenidos en la presente propuesta están contenidos en los calculados en [9]. Lo anterior muestra que la de predicción Bayesiana propuesta para

este modelo simulado arroja mejores resultados que la utilizada en Nieto (2008) con cualquiera de las dos distribuciones a priori consideradas, siendo las distribuciones a priori informativas las que proporcionan predicciones Bayesianas más favorables.

Tabla 10. Predicciones en el ejemplo teórico TAR(3;1,2,2).

t	$x_t{}^{(1)}$	$x^{(a)}_{t\|990}$	$x^{(b)}_{t\|990}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$	$\mathbf{RPC}^{(a)}$	$\mathbf{RPC}^{(b)}$
991	1.19	1.05	0.93	[0.59,3.14]	[0.58,3.17]	[0.56,3.12]	[0.55,3.08]	1.91	1.85
992	2.05	1.93	1.96	[0.68,3.21]	[0.67,3.25]	[0.64,2.85]	[0.64,2.84]	1.93	1.93
993	0.76	0.13	0.18	[0.05,3.31]	[0.08,3.39]	[-0.01,2.89]	[-0.04,2.91]	2.05	1.98
994	0.90	-0.11	-0.15	[-0.48,2.93]	[-0.51,2.95]	[-1.42,2.95]	[-1.41,2.97]	2.11	2.08
995	-0.75	0.12	0.25	[-1.31,2.78]	[-1.35,2.75]	[-1.21,2.70]	[-1.22,2.69]	2.45	2.40
996	1.08	1.02	1.12	[-1.62,2.82]	[-0.61,2.85]	[-1.60,2.81]	[-0.62,2.86]	2.51	2.39
997	0.84	-0.13	0-0.20	[-1.67,2.83]	[-1.67,1.82]	[-1.40,2.81]	[-1.41,1.78]	2.89	2.51
998	0.69	-0.30	0.38	[-1.03,3.15]	[-1.07,3.22]	[-1.01,2.98]	[-1.03,2.97]	3.08	2.75
999	0.86	0.81	0.78	[-1.31,3.35]	[-1.37,3.31]	[-1.21,3.13]	[-1.23,3.12]	3.11	2.82
1000	0.91	0.56	0.49	[-1.60,3.40]	[-1.57,3.40]	[-1.50,3.21]	[-1.48,3.24]	3.41	3.19

$^{(1)}$: Valor simulado $^{(a)}$: A priori no informativa $^{(b)}$: A priori informativa

4.2. Aplicación con Datos Reales.

Para la aplicación en datos reales se toman los datos utilizados por Nieto (2008), los cuales corresponden a las variables precipitación diaria (en mm), considerada la variable de umbrales (Z_t) y como la variable respuesta (X_t) se tiene el caudal de un río (en m^2/s).

Las variables fueron medidas en una determinada región geográfica Colombiana. La precipitación se mide en la estación meteorológica de la Laguna de San Rafael, con una altitud de 3420 metros y coordenadas geográficas 2.23° latitud norte y 76.27° al oeste. El caudal se mide en el río Bedón en la estación hidrológica de la Laguna de San Rafael con una altitud de 3300 metros y coordenadas de 2.19° al norte y 76.15° al oeste. Estas estaciones están situadas cerca de la linea del ecuador de la tierra, en zonas muy secas.

Las dos series temporales consideradas, tienen una longitud de 3152 datos observados a partir del 1 de enero de 1992, hasta 30 de noviembre de 2000. Ambas series poseían datos faltantes, para lo cual Nieto (2005), implementó una metodología para estimación de datos faltantes y finalmente determina que estas series temporales pueden modelarse mediante un TAR(2;1,1). En esta aplicación se asumen estos parámetros estructurales y mediante el muestreador de *Gibbs* se estiman los parámetros no estructurales del modelo, tanto para distribuciones a prioris informativas (las consideradas por Nieto (2005)) como para a priori no informativas.

Para distribuciones a priori no informativas de Jeffreys tiene que

$$X_t = \begin{cases} 1.31 + 0.62X_{t-1} + 1.37\varepsilon_t, \text{ si } Z_t < 6, \\ 2.61 + 0.39X_{t-1} + 2.44\varepsilon_t, \text{ si } Z_t \geq 6, \end{cases}$$

y para distribuciones a priori informativas,

$$X_t = \begin{cases} 1.32 + 0.61X_{t-1} + 1.34\varepsilon_t, \text{ si } Z_t < 6, \\ 2.58 + 0.38X_{t-1} + 2.41\varepsilon_t, \text{ si } Z_t \geq 6. \end{cases}$$

Se observa que los dos modelos son muy similares, siendo el segundo modelo más parecido al encontrado en [9], lo cual es de esperarse, ya que en este ultimo las estimaciones también fueron basadas en distribuciones a priori informativas.

En la Tabla 11 se encuentran las estimaciones de los distintos parámetros, así como los intervalos de credibilidad y regiones de alta densidad.

En la Figura 7 se observa la gráfica de las dos variables.

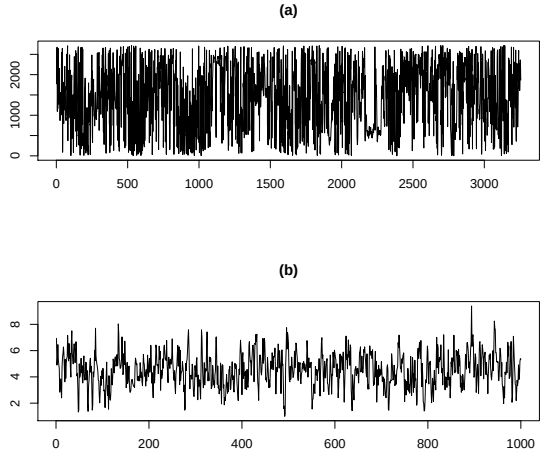

Figura 7. Series de tiempo para el ejemplo con datos reales: (a) Caudal (b) Precipitación.

Tabla 11. Estimación de los Parámetros para los datos reales.

$\boldsymbol{\theta}$	$\widehat{\boldsymbol{\theta}}^{(a)}$	$\widehat{\boldsymbol{\theta}}^{(b)}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$
$a_0^{(1)}$	1.31	1.32	[1.25,1.37]	[1.26,1.35]	[1.24,1.37]	[1.26,1.36]
$a_1^{(1)}$	0.62	0.61	[0.60,0.64]	[0.61,0.64]	[0.60,0.64]	[0.61,0.64]
$[h^{(1)}]^2$	1.89	1.8	[1.79,1.99]	[1.80,1.91]	[1.79,1.99]	[1.81,1.92]
$a_0^{(2)}$	2.61	2.58	[2.57,2.64]	[2.59,2.62]	[2.57,2.64]	[2.59,2.62]
$a_1^{(2)}$	0.39	0.38	[0.38,0.39]	[0.38,0.40]	[0.379,0.39]	[0.38,0.40]
$[h^{(2)}]^2$	5.95	5.8	[5.59,6.33]	[5.61,6.34]	[5.577,6.32]	[5.61,6.30]

$^{(a)}$:A priori no informativa $^{(b)}$: A priori informativa

Tomando p_1=0.6 y p_2=0.4, como se calculan en Nieto (2008), se obtienen las predicciones para la variable caudal. Para efectos de comparar con valores observados, se toman las predicciones en las observaciones que van de $t = 3191$ hasta $t = 3196$, dado $t = 3190$ (Tabla 12). Estas predicciones son coherentes con los valores reales, ya que las estimaciones son cercanas y los intervalos de credibilidad contienen a los valores observados. También se observa que la varianza de la predictiva es creciente, conforme

crece el horizonte de predicción. Estas predicciones son inferiores a las obtenidas por la propuesta de Nieto (2008) (ver Anexo A2), mientras que los intervalos y la raíces de las varianzas predictivas son de menor longitud a los encontrados en Anexo A2, y están contenidos en los mismos (Tanto para a priori informativas y no informativas), este mismo comportamiento se observa al comparar los resultados de las predicciones para el caudal (Tabla 13.) y las predicciones hechas mediante la metodología de Nieto (2008), situadas en el Anexo A3. Lo anterior indica que lo propuesto en el presente trabajo es una buena estrategia para pronósticos en modelos TAR. En este caso las predicciones más adecuadas, como ha sido en todos los resultados, en a priori informativas. Lo encontrado en esta aplicación coincide con lo obtenido en la simulación del modelo TAR con tres regímenes (Tabla 10), donde también se muestra que las predicciones Bayesianas realizadas a partir de distribuciones a priori informativas presentan mejores resultados.

Tabla 12. Predicciones intra-muestra para el caudal.

t	$x_t{}^{(1)}$	$x^{(a)}_{t\|990}$	$x^{(b)}_{t\|990}$	$\mathbf{IC}^{(a)}$	$\mathbf{IC}^{(b)}$	$\mathbf{RAD}^{(a)}$	$\mathbf{RAD}^{(b)}$	$\mathbf{RPC}^{(a)}$	$\mathbf{RPC}^{(b)}$
3191	2.22	5.72	3.72	[0.63,6.84]	[0.83,5.42]	[0.61,7.81]	[0.83,5.32]	1.72	1.61
3192	3.00	4.62	2.83	[0.54,6.72]	[0.84,5.52]	[0.53,6.74]	[0.84,5.50]	1.71	1.61
3193	1.30	4.63	2.63	[0.67,5.79]	[0.87,5.52]	[0.25,6.43]	[0.85,5.50]	1.74	1.62
3194	1.23	3.34	2.44	[0.25,6.47]	[0.75,5.34]	[0.24,6.46]	[0.73,5.35]	1.93	1.73
3195	4.91	3.33	2.52	[0.23,6.44]	[0.73,5.37]	[0.24,6.44]	[0.72,5.35]	2.02	1.76
3196	3.80	3.92	2.97	[0.18,6.45]	[0.78,5.42]	[0.19,6.46]	[0.80,5.43]	2.09	1.78

$^{(1)}$: Valor simulado $^{(a)}$:A priori no informativa $^{(b)}$: A priori informativa

La Figura 8 muestra las distribuciones de las predicciones obtenidas en la Tabla 12. En esta se corrobora lo enunciado en las aplicaciones, en ella se observa que las distribuciones aposteriori de las distintas predicciones presentan mayor variabilidad cuando se utilizan distribuciones apriori no informativas, este comportamiento es más evidente a medida que aumenta el horizonte de predicción. También se observa que en general las predicciones realizadas con base en distribuciones a priori no informativas son superiores a las obtenidas a partir de a priori informativas.

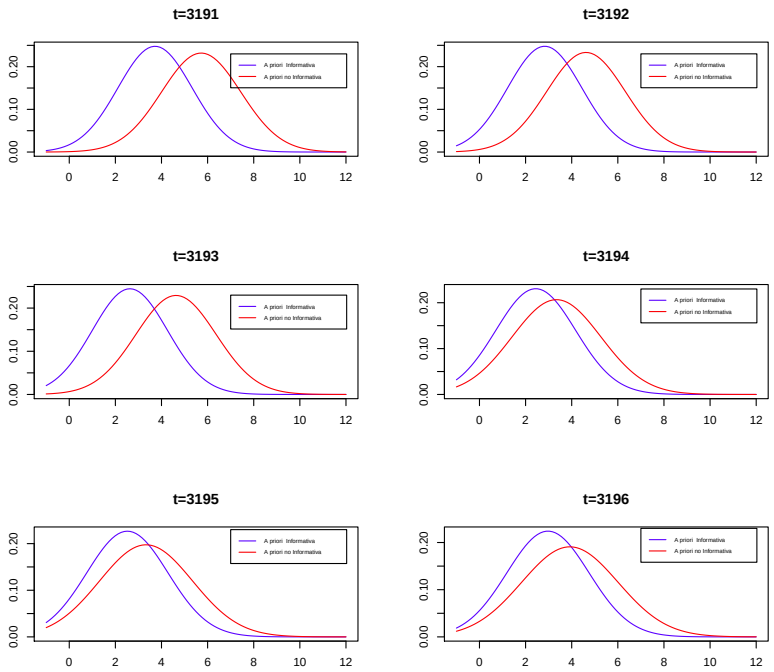

Figura 8. Distribución a posteriori de las predicciones Bayesianas para caudal del río.

En la Tabla 13 se encuentran las predicciones para X_{3256+h}, $h = 1, ..., 5$. en dicha serie.

Tabla 13. Predicciones para el caudal.

t	$x_{t\|3266}^{(a)}$	$x_{t\|3256}^{(b)}$	IC$^{(a)}$	IC$^{(b)}$	RAD$^{(a)}$	RAD$^{(b)}$	RPC$^{(a)}$	RPC$^{(b)}$
3257	3.26	2.81	[0.70,6.92]	[0.61,5.82]	[0.70,6.93]	[0.61,6.93]	1.87	1.73
3258	3.07	2.42	[0.32,6.51]	[0.30,5.67]	[0.34,6.33]	[0.39,5.34]	1.87	1.74
3259	3.05	2.35	[0.25,6.43]	[0.28,6.31]	[0.25,6.43]	[0.30,6.32]	1.88	1.81
3260	3.12	2.34	[0.26,6.44]	[0.35,6.31]	[0.24,6.42]	[0.30,6.31]	1.90	1.84
3261	3.27	2.02	[0.23,6.83]	[0.38,6.23]	[0.23,6.83]	[0.37,6.16]	2.08	1.89

$^{(a)}$:A priori no informativa $^{(b)}$: A priori informativa.

En la Figura se encuentran las distribuciones a posteriori informativa y no informativa para los pronósticos de la variable caudal del río en los primeros 5 horizontes de predicción. En ella se observa un comportamiento similar en las distribuciones de los cinco pronósticos a favor de la distribución a priori informativa, como se ha observado

en todas las aplicaciones del método, tanto en aplicaciones con modelos simulados como en las aplicaciones con datos reales.

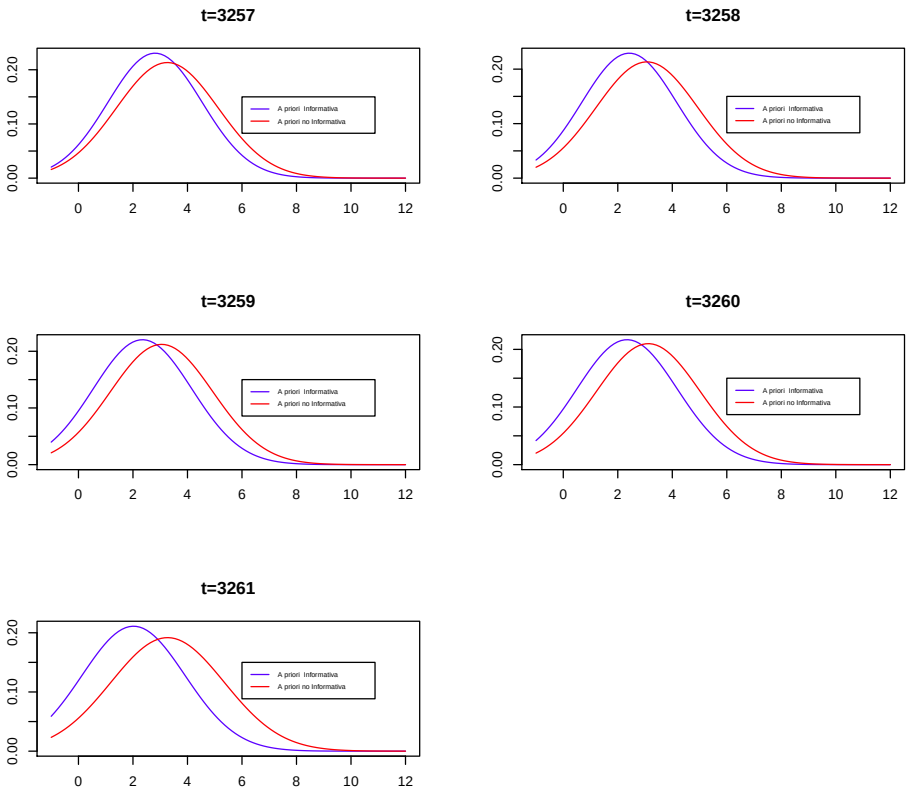

Figura 9. Distribución a posteriori para el pronóstico de caudal del río.

CONCLUSIONES.

Se obtuvo la expresión para el cálculo de la distribución predictiva Bayesiana en modelos no lineales TAR. La función de probabilidad de la predictiva funciona como una combinación entre las funciones de probabilidad predictivas bayesianas por regímen y los pesos $p_j(h)$, permitiendo así generar predicciones Bayesianas ponderadas por regímen.

Los resultados muestran que esta metodología es mejor que la utilizada por Nieto (2008), en el sentido de que proporciona predicciones, intervalos creíbles y varianzas predictivas de menor longitud y posee una expresión más practica de cómputo.

Al utilizarse distribuciones a priori informativas y no informativas, se observó que estas ultimas poseen intervalos de credibilidad y varianzas predictivas de mayor longitud, por lo cual el presente trabajo sugiere obtener predicciones Bayesianas para los modelos TAR utilizando distribuciones a priori informativas.

ANEXOS.

Anexo A. Predicciones realizadas mediante la metodología propuesta por Nieto (2008).

Anexo A1. Predicciones en el ejemplo teórico TAR(3;1,2,2).

t	$x_t{}^{(1)}$	$x_{t\|990}^{(b)}$	IC$^{(b)}$	RAD$^{(b)}$	RPC$^{(b)}$
991	1.19	1.30	[-1.55,3.58]	[-0.55,3.09]	2.01
992	2.05	3.06	[-1.81,3.56]	[-1.84,3.54]	2.16
993	0.76	1.60	[-1.31,3.51]	[-1.34,3.91]	2.56
994	0.90	-0.78	[-1.41,3.15]	[-1.47,3.17]	2.49
995	-0.75	1.01	[-1.61,3.56]	[-1.62,3.59]	2.86
996	1.08	1.98	[-1.67,4.19]	[-0.72,4.21]	2.99
997	0.84	-0.67	[-1.41,4.67]	[-1.47,4.70]	3.89
998	0.69	0,89	[-1.21,4.62]	[-1.31,4.62]	4.99
999	0.86	1.01	[-1.48,4.32]	[-1.46,4.39]	4.99
1000	0.91	1.25	[-1.14,5.74]	[-1.17,5.79]	4.99

$^{(1)}$: Valor simulado $^{(a)}$: Apriori no informativa $^{(b)}$: Apriori informativa.

Anexo A2. Predicciones intra-muestra del caudal.

t	$x_t{}^{(1)}$	$x_{t\|3190}$	IC	RAD	RPC
3191	2.22	6.58	[3.14,10.08]	[3.03,9.80]	2.32
3192	3.00	5.26	[1.37,9.29]	[1.27,9.17]	2.59
3193	1.30	4.47	[0.33,8.68]	[0.18,8.44]	2.67
3194	1.23	3.98	[-0.18,8.12]	[-0.23,9.19]	2.65
3195	4.91	3.72	[-0.60,8.10]	[-0.66,7.87]	2.74
3196	3.80	3.55	[-0.68,7.79]	[-0.59,7.78]	2.73
3197	3.16	3.46	[-0.85,7.74]	[-0.71,7.65]	2.76

$^{(1)}$: Valor real.

Anexo A3. Predicciones para el caudal.

t	$x_{t\|3256}$	IC	RPC
3257	4.08	[0.59,7.70]	2.46
3258	3.74	[-0.27,7.85]	2.65
3259	3.58	[-0.54,7.89]	2.68
3260	3.50	[-0.71,7.66]	2.68
3261	3.36	[-0.83,7.56]	2.73

Anexo B. Demostración de las Proposiciones.

Anexo B1. Demostración de la proposición 2.3.

Se tiene que la función (2.6) puede expresarse como el producto de (2.7) y (2.8), dado que la ultima expresión no depende de los parámetros no estructurales (los de interés a estimar), se puede expresar:

$$f(\mathbf{x}, \mathbf{z}|\boldsymbol{\theta}_x, \boldsymbol{\theta}_z) \propto f(\mathbf{x}|\mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_y)$$

El exponente en (2.8) se puede escribir:

$$-\frac{1}{2\left\{h^{(j)}\right\}^2}\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j\right)'\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j\right) \tag{4.1}$$

Mediante algunos cálculo en (4.1) se llega a la siguiente forma:

$$-\frac{1}{2\left\{h^{(j)}\right\}^2}\left[\mathbf{X}'\mathbf{X} - \mathbf{X}'\mathbf{W}\left(\mathbf{W}'\mathbf{W}\right)^{-1}\mathbf{W}'\mathbf{X}\right] + \left(\boldsymbol{\theta}_j - \boldsymbol{\theta}_0^{(j)}\right)' - \frac{1}{2\left\{h^{(j)}\right\}^2}\left(\mathbf{W}'\mathbf{W}\right)\left(\boldsymbol{\theta}^{(j)} - \boldsymbol{\theta}_0^{(j)}\right)$$

de lo cual:

$$p\left(\boldsymbol{\theta}^{(j)}|\mathbf{h}, \boldsymbol{\theta}^{(i)}; i \neq j; \mathbf{h}, \boldsymbol{\theta}_z, \mathbf{x}, \mathbf{z}\right) \propto \exp\left[-\frac{1}{2}\left(\boldsymbol{\theta}^{(j)} - \boldsymbol{\theta}_0^{(j)}\right)' V_0^{(j)}\left(\boldsymbol{\theta}^{(j)} - \boldsymbol{\theta}_0^{(j)}\right)\right]\pi\left(\boldsymbol{\theta}^{(j)}\right)$$

donde:

$\boldsymbol{\theta}_0^{(j)} = \left\{\mathbf{W}_j'\mathbf{W}_j\right\}^{-1}\mathbf{W}_j'\mathbf{X}_j$

y

$$V_0^{(j)} = \frac{1}{\left\{h^{(j)}\right\}^2}\mathbf{W}_j'\mathbf{W}$$

pero, $\pi\left(\boldsymbol{\theta}^{(j)}\right) \propto 1$, así

$$p\left(\boldsymbol{\theta}^{(j)}|\mathbf{h}, \boldsymbol{\theta}^{(i)}; i \neq j; \mathbf{h}, \boldsymbol{\theta}_z, \mathbf{x}, \mathbf{z}\right) \propto \exp\left[-\frac{1}{2}\left(\boldsymbol{\theta}^{(j)} - \boldsymbol{\theta}_0^{(j)}\right)' V_0^{(j)}\left(\boldsymbol{\theta}^{(j)} - \boldsymbol{\theta}_0^{(j)}\right)\right]$$

luego $p\left(\boldsymbol{\theta}^{(j)}|\left\{h^{(j)}\right\}^2, \boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}_z, \mathbf{x}, \mathbf{z}\right)$ es directamente proporcional al kernel de una distribución normal multivariante con media $\boldsymbol{\theta}_0^{(j)}$ y varianza $\left(V_0^{(j)}\right)^{-1}$.

Anexo B2. Demostración de la proposición 2.4.

Como se ilustró en la proposición anterior, para la función de verisimilitud en un modelo TAR se tiene:

$$f(\mathbf{x}, \mathbf{z}|\boldsymbol{\theta}_x, \boldsymbol{\theta}_z) \propto f(\mathbf{x}|\mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_y)$$

con

$$f(\mathbf{x}|\boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = (2\pi)^{-(T-k)/2}\left[\prod_{t:j_t=j}\left\{h^{(j_t)}\right\}^{-1}\right]\exp\left(-\frac{1}{2}\sum_{t:j_t=j}e_t^2\right)$$

así

$$f(\mathbf{x}|\boldsymbol{\theta}_x, \boldsymbol{\theta}_z) \propto \left[\left\{h^{\{(j)\}}\right\}^{-2}\right]^{\frac{n_j}{2}}\exp\left(-\frac{1}{2}\sum_{t:j_t=j}e_t^2,\right)$$

siendo

$$\exp\left(-\frac{1}{2}\sum_{t:j_t=j}e_t^2\right) = \exp\left(-\frac{1}{2\left\{h^{(j)}\right\}^2}\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j\right)'\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j\right)\right)$$

luego

$$f(\mathbf{x}|\boldsymbol{\theta}_x, \boldsymbol{\theta}_z) \propto \left[\left\{h^{\{(j)\}}\right\}^{-2}\right]^{\frac{n_j}{2}}\exp\left(-\frac{1}{2\left\{h^{(j)}\right\}^2}\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j\right)'\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}_j\right)\right)$$

de aquí

$$p\left(\left\{h^{(j)}\right\}^{-2}|\boldsymbol{\theta}_{-j,h}, \mathbf{x}, \mathbf{z}\right) \propto \left[\left\{h^{(j)}\right\}^{-2}\right]^{\frac{n_j}{2}}\exp\left(-\frac{1}{2\left\{h^{(j)}\right\}^2}\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}^{(j)}\right)'\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}^{(j)}\right)\right)\pi\left(\left\{h^{(j)}\right\}^{-2}\right)$$

pero, $\pi\left(\left\{h^{(j)}\right\}^{-2}\right) \propto \left\{h^{(j)}\right\}^{2}$, así

$$p\left(\left\{h^{(j)}\right\}^{-2}|\boldsymbol{\theta}_{-j,h}, \mathbf{x}, \mathbf{z}\right) \propto \left[\left\{h^{(j)}\right\}^{-2}\right]^{\frac{n_j}{2}-1} \exp\left(-\frac{1}{2\left\{h^{(j)}\right\}^{2}}\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}^{(j)}\right)'\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}^{(j)}\right)\right)$$

luego, $p\left(\left\{h^{(j)}\right\}^{-2}|\boldsymbol{\theta}_{-j,h}, \mathbf{x}, \mathbf{z}\right)$ es directamente proporcional a una distribución Gamma con parámetro de forma $n_j/2$ y escala $(1/2)\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}^{(j)}\right)'\left(\mathbf{X}_j - \mathbf{W}_j\boldsymbol{\theta}^{(j)}\right)$.

Anexo B3. Demostración de la Proposición 3.2

Se tiene que:

$$F_{T+h}(x_{T+h}|\mathbf{x}_T, \mathbf{z}_T) = p\left(X_{T+h} \leq x_{T+h}|\mathbf{x}_T, \mathbf{z}_T\right) = p\left(X_{T+h} \leq x_{T+h}, \Omega|\mathbf{x}_T, \mathbf{z}_T\right)$$

$$= p\left(\left[X_{T+h} \leq x_{T+h}\right] \cap \bigcup_{j=1}^{l} B_j|\mathbf{x}_T, \mathbf{z}_T\right) = p\left(\bigcup_{j=1}^{l}(X_{T+h} \leq x_{T+h}, B_j)|\mathbf{x}_T, \mathbf{z}_T\right)$$

$$= \sum_{j=1}^{l} p\left(X_{T+h} \leq x_{T+h}, B_j|\mathbf{x}_T, \mathbf{z}_T\right) = \sum_{j=1}^{l} p\left(X_{T+h} \leq x_{T+h}|B_j, \mathbf{x}_T, \mathbf{z}_T\right) p(B_j|\mathbf{x}_T, \mathbf{z}_T)$$

así

$$f_{T+h}(x_{T+h}|\mathbf{x}_T, \mathbf{z}_T) = \sum_{j=1}^{l} p\left(x_{T+h}|B_j, \mathbf{x}_T, \mathbf{z}_T\right) p_j(h),$$

se tiene que $p\left(x_{T+h}|B_j, \mathbf{x}_T, \mathbf{z}_T\right)$ es la distribución predictiva de x_{T+h} dados $\mathbf{x}_T$ y $\mathbf{z}_T$, siempre que $z_{T+h} \in B_j$, así esto se reduce a calcular la distribución predictiva de x_{T+h} en el regímen j, donde, finalmente se llega a la expresión (3.6) mediante la definición (1.2).

Bibliografía

[1] Bolivar, W. & Cartaya, V (2009) Determinación de ciclos y tendencias en series de tiempo macroeconómicas mediante un enfoque bayesiano. Banco Central de Venezuela, 109, 1-35.

[2] Briñez, A. & Nieto, F. (2005) Ajuste de un modelo no lineal a la variable precipitación en una estación hidro-metereologica de Colombia, revista colombiana de estadística, 28, 113-124.

[3] Casella, G. (1985) An Introduction to Empirical Bayes Data Analysis. The American Statistician, Vol. 39, No. 2, pp. 83-87,

[4] Gelfand, A. & Smith, A. F. M. (1990), Sampling based approaches to calculating marginal densities, J. Amer. Stat. Assoc 85, 398-409.

[5] Geman, S. & Geman, D. (1984), Stochastic relaxation, Gibbs distributions and the bayesian restoration of images, IEEE Trans. on Pattern Analysis and Machine Intelligence 6, 721-741.

[6] Hoyos, M. & Nieto, F. (2006) Aplicación del modelo no lineal TAR para analizar la relación entre las expectativas de actividad económica y la estructura a plazo: los casos de Colombia y Estados Unidos. Reporte Interno No. 8, Unidad de Investigación, Departamento de Estadística, Universidad Nacional de Colombia.

[7] Ibáñes, M. & Simó A. (2003) Modelos Estadísticos Espacio temporales en perimetría. Tesis Doctoral, Universidad Jaume , 241-271.

[8] Moreno, E. (2011) Modelos TAR en series de tiempo financieras. Tesis de Maestría, Departamento de Estadística, Universidad Nacional de Colombia.

[9] Nieto, F. (2008) Forecasting with univariate TAR models. Reporte interno, Unidad de Investigación, Departamento de Estadística, Universidad Nacional de Colombia.

[10] Nieto, F. (2005) Modeling bivariate threshold autoregressive processes in the presence of missing data. Communications in Statistics: Theory and Methods, 34, 905-930.

[11] Rojo, J. & Sans, J. (2010). Descomposición de series de tiempo cortas: un enfoque bayesiano. Estadística Española, 52, 127 - 153.

[12] Salamanca, A & Monroy, V (2008) Deuda externa pública e inversión en Colombia 1994- 2007: Evidencia de un modelo no-lineal TAR. Borradores de Economía.

[13] Tong, H. (1990) Nonlinear Time Series, A Dynamical System Approach, Oxford Uni- versity Press, Oxford.