

Bayesian Analysis of Multiplicative Seasonal Threshold Autoregressive Processes

Análisis Bayesiano de procesos autorregresivos de umbrales estacionales multiplicativos

JOAQUÍN GONZÁLEZ^{1,a}, FABIO H. NIETO^{2,b}

¹DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD DEL
TOLIMA, IBAGUÉ, COLOMBIA

²DEPARTAMENTO DE ESTADÍSTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD NACIONAL DE
COLOMBIA, BOGOTÁ, COLOMBIA

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.

Key words: Bayesian analysis; Exogenous variable; Multiplicative model; Nonlinearity; Seasonality; Threshold autoregressive models.

Resumen

Las fluctuaciones estacionales son frecuentes en series de tiempo. En adición, la no linealidad y la relación con otras series de tiempo son comportamientos prominentes de muchas series. En este artículo, se considera 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

^aPh.D. E-mail: jgonzalezb@ut.edu.co

^bPh.D. E-mail: fnietos@unal.edu.co

de umbrales. La estacionalidad es estocástica y dependiente del régimen. Se desarrolla un procedimiento basado en métodos Bayesianos para identificar el modelo, estimar sus parámetros, validarlo y calcular pronósticos. En la etapa de identificación del modelo, se presenta una prueba estadística de estacionalidad multiplicativa por regímenes. La metodología propuesta es ilustrada con un ejemplo simulado y aplicada a datos empíricos económicos.

Palabras clave: Análisis bayesiano; Estacionalidad; Modelos autorregresivos de umbrales; Modelo multiplicativo; No linealidad; Variable exógena.

1. Introduction

The threshold autoregressive models (TAR) proposed by Tong & Chen (1978) and Tong & Lim (1980) have been widely applied in several areas, including economics, finance, hydrology, meteorology and biology. The TAR models assume that the values of a stochastic process $\{Z_t\}$ (threshold process) determines both the values of the stochastic process $\{X_t\}$ (interest process) and its dynamics. When the threshold process is the same process of interest but lagged, the model is known as SETAR (Self-Exciting TAR), Tong (1990). The TAR models can be easily extended, including some exogenous time series and called TARX. The use of Bayesian methods applied to the TAR and SETAR models, can be found in Chen & Lee (1995), So & Chen (2003), Nieto (2005), Nieto (2008), Nieto, Zhang & Li (2013), Zhang & Nieto (2015) and Calderón & Nieto (2017). While to the TARX models, can be found in Chen (1998), So, Chen & Liu (2006), Chen & So (2006), Chen, Gerlach & Lin (2010), among other works. Additionally, Hansen (2011) and Chen, Liu & So (2011) give a review of the theory and applications of the TAR and TARX models in economics and finance, respectively.

TAR models may be useful for nonseasonal or seasonally adjusted data that display nonlinearity by means of thresholds. However, diverse time series present some seasonal patterns. In addition, the nonlinearity of thresholds and the relationship with other time series are typical characteristics of many such time series in practice. Tong (2015) gives a compendium of what has been developed with respect to the TAR models, and possible issues to be investigated, including the study of seasonality. In many cases, it is convenient to study the seasonal patterns present in time series, since they can carry relevant information regarding the dynamic behavior of the series.

Thus, using a frequentist approach and data without seasonal adjustment, Franses, de Bruin & van Dijk (2000) and Franses & Van Dijk (2005) examine the performance of various models of linear and smooth transition autoregressive (STAR) time series, considering deterministic seasonality of series of quarterly industrial production data. Crespo (2001) considers seasonality as deterministic and dependent of the regime in the SETAR models and uses his model in time series of quarterly unemployment rates. De Gooijer & Vidiella-i Anguera (2003) define a seasonal SETAR model, where each regime has the form of a multiplicative seasonal autoregressive model. They carry out the study of time series of monthly inflation rates in several countries, finding more realistic dynamic characteristics

than those explained by the usual SETAR or linear seasonal time series models. In this paper, we propose a statistical model called the multiplicative seasonal threshold autoregressive model with exogenous input (TSARX), which explicitly and simultaneously incorporates multiplicative seasonality by regimes that arises from periodic autocorrelation structures and threshold nonlinearity. The proposed model arises from imposing certain restrictions on the parameters of a TARX model. Bayesian methods are used to determine the presence of multiplicative seasonality by regimes and to obtain joint estimates of parameters, generalized residuals and forecasts. The developed methodology is illustrated with a simulated TSARX model and with an empirical application in the Colombian economy. The variables to be used are the growth rates of total monthly unemployment and a monthly coincident index that traces the Colombian economy (ISE for its Spanish acronym of índice de seguimiento económico).

The paper is organized as follows: In Section 2, the TSARX model and the proposed methodology are presented. Section 3 shows a simulated example. The real-data application is presented in Section 4 and, finally, Section 5 concludes.

2. The TSARX Model

2.1. Model Specification

Let $\{X_t\}$ and $\{Z_t\}$ be stochastic processes related by means of the equation

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

if $r_{j-1} < Z_{t-d} \leq r_j$, for all $t \in \mathbb{Z}$ (\mathbb{Z} is the set of integers numbers) and some $j = 1, \dots, l$. The model assumes l regimes defined by the threshold values r_j that satisfy $-\infty = r_0 < r_1 < \dots < r_{l-1} < r_l = \infty$ and the threshold variable lagged d periods Z_{t-d} . In addition, the threshold variable is considered an exogenous input, in the sense that there is no feedback from $\{X_t\}$ towards $\{Z_t\}$. The parameters of the model can be divided into two groups (Nieto 2005):

- Structural parameters: the number of regimes l , the $l - 1$ thresholds $\mathbf{r} = (r_1, \dots, r_{l-1})'$, the delay parameter d , the autoregressive orders of the l regimes $k_1, \dots, k_l, K_1, \dots, K_l, q_1, \dots, q_l$ and seasonal period s .
- Nonstructural parameters: the autoregressive coefficients $a_i^{(j)}, b_u^{(j)}, c_v^{(j)}$, $i = 0, 1, \dots, k_j, u = 1, \dots, K_j, v = 1, \dots, q_j, j = 1, \dots, l$ and the variance weights $h^{(1)}, \dots, h^{(l)}$.

We assume that the seasonal period of model s is known, where $s = 4$ is usually taken for quarterly time series and $s = 12$ for monthly time series. Moreover, model

(1) could be used for time series that exhibit another type of seasonal patterns, as, for example, either daily (hour data) or weekly (daily observations). We consider $k_j < s$ for all $j(j = 1, \dots, l)$.

Also, $\{\epsilon_t\}$ is a zero mean Gaussian white noise process with variance 1 ($GWN(0, 1)$) such that $E(X_w \epsilon_t) = 0$ for $w < t$, Z_w and the set $\{X_t, X_{t-1}, \dots\}$ are mutually independent for all $w > t$ and $\{Z_t\}$ and $\{\epsilon_t\}$ are mutually independent. To describe the stochastic behavior of $\{Z_t\}$, we additionally assume that $\{Z_t\}$ is a homogeneous Markov chain of order $m \geq 1$, with initial density function $f(\cdot)$ and kernel density function $f_m(\cdot | \cdot)$, in the Lebesgue - measure sense. Additionally, the Markov chain converges weakly to a distribution F . For more details about Markov chains with general state space, see Meyn & Tweedie (2009). Equation (1) describes a dynamic system without feedback, with observable input $\{Z_t\}$ and observable output $\{X_t\}$, which is presented in Tong (1990) but without considering the modeling of multiplicative seasonality by regimes.

The symbol $TSARX(l; k_1, \dots, k_l; K_1, \dots, K_l; q_1, \dots, q_l)_s$, is used to denote model (1) and it is said that $\{X_t\}$ is a threshold multiplicative seasonal autoregressive process with exogenous input, with $\{Z_t\}$ as its threshold process and s as its seasonal period. The vector of nonstructural parameters of the TSARX model is defined as $\theta_{ns} = (A'_1, \dots, A'_l, B'_1, \dots, B'_l, C'_1, \dots, C'_l, \mathbf{h}')'$, where $A_j = (a_0^{(j)}, a_1^{(j)}, \dots, a_{k_j}^{(j)})$, $B_j = (b_1^{(j)}, \dots, b_{K_j}^{(j)})$, $C_j = (c_1^{(j)}, \dots, c_{q_j}^{(j)})$, $j = 1, \dots, l$ and $\mathbf{h} = ((h^{(1)})^2, \dots, (h^{(l)})^2)$ and $\theta_s = (k_1, \dots, k_l, K_1, \dots, K_l, q_1, \dots, q_l, \mathbf{r}', d, s, l)'$ is the vector of structural parameters of the model. $\theta = (\theta_x, \theta_z)$ is denoted as the complete parameter vector, where $\theta_x = (\theta'_{ns}, \theta'_s)$ is the parameter vector in the TSARX model and θ_z is the parameter vector of the Markov chain.

Following the theoretical results regarding the TAR models, given in Nieto & Moreno (2016), and making an extension to the TSARX models, we have the following considerations:

- (i) Let $K = \max\{k_j + sK_j, q_j, j = 1, \dots, l\}$, the distribution of X_t for all t , conditional on $\mathbf{x}_{1,t-1} = (x_1, \dots, x_{t-1})$, $\mathbf{z}_{1,t-1} = (z_1, \dots, z_{t-1})$ and $z_{t-d} \in R_j = (r_{j-1}, r_j]$, for some $j = 1, \dots, l$, $t > K$, which is denoted as $F_{t-1,j}(x_t | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, z_{t-d} \in R_j)$ is a normal distribution with mean

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

and variance $(h^{(j)})^2$, for each $j = 1, \dots, l$.

- (ii) The conditional cumulative distribution function of X_t given $\mathbf{x}_{1,t-1}$, $\mathbf{z}_{1,t-1}$, $t > K$, is given by

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

where $p_{t,j} = Pr(z_{t-d} \in R_j)$, $\sum_{j=1}^l p_{t,j} = 1$ and $F_{t-1,j}(\cdot | \cdot)$ ($j = 1, \dots, l$) is given in (i). Now, being that $\{Z_t\}$ converges weakly to F , $p_{t,j} \rightarrow p_j =$

$F(r_j) - F(r_{j-1})$ when $t \rightarrow \infty$ for all $j = 1, 2, \dots, l$. Of course, if $\{Z_t\}$ is a strictly stationary process, then the cumulative distribution function of Z_t is F for all t and, thus, $p_{t,j} = p_j$ for all t and for all $j = 1, \dots, l$.

The theoretical considerations (i) and (ii) above are taken into account for the calculation of the generalized residuals of the TSARX model.

The model (1) can be considered as a special case of a TARX model, as described below. A TARX($l; k_1 + sK_1, \dots, k_l + sK_l; q_1, \dots, q_l$) model is of the form

$$X_t = a_0^{(j)} + \sum_{i=1}^{k_j+sK_j} a_i^{(j)} X_{t-i} + \sum_{v=1}^{q_j} c_v^{(j)} Z_{t-v} + h^{(j)} \epsilon_t, \tag{2}$$

if $r_{j-1} < Z_{t-d} \leq r_j$ for each $t \in \mathbb{Z}$, some $j = 1, \dots, l, k_j < s$ for all j , and $\{\epsilon_t\} \sim GWN(0, 1)$.

The complete parameter vector of the TARX model is conveniently defined as $\theta_x = (\theta'_{ns}, \theta'_s)$ being $\theta_{ns} = (A'_{11}, \dots, A'_{l1}, A'_{12}, \dots, A'_{l2}, C'_1, \dots, C'_l, \mathbf{h}')'$ (vector of nonstructural parameters) where $A_{j1} = (a_0^{(j)}, a_1^{(j)}, \dots, a_{k_j}^{(j)}, a_s^{(j)}, \dots, a_{sK_j}^{(j)})$, $A_{j2} = (a_{1+k_j}^{(j)}, \dots, a_{s-1}^{(j)}, a_{1+s}^{(j)}, \dots, a_{k_j+s}^{(j)}, \dots, a_{(1+k_j)+s(K_j-1)}^{(j)}, a_{(s-1)+s(K_j-1)}^{(j)}, a_{1+sK_j}^{(j)}, \dots, a_{k_j+sK_j}^{(j)})$, $C_j = (c_1^{(j)}, \dots, c_{q_j}^{(j)})$, $j = 1, \dots, l$, $\mathbf{h} = ((h^{(1)})^2, \dots, (h^{(l)})^2)$ and $\theta_s = (k_1 + sK_1, \dots, k_l + sK_l, q_1, \dots, q_l, \mathbf{r}', d, l)'$ (vector of structural parameters) where $\mathbf{r} = (r_1, \dots, r_{l-1})'$. If $A_{j2} = (0, \dots, 0, -a_1^{(j)} a_s^{(j)}, \dots, -a_{k_j}^{(j)} a_s^{(j)}, \dots, 0, \dots, 0, -a_1^{(j)} a_{sK_j}^{(j)}, \dots, -a_{k_j}^{(j)} a_{sK_j}^{(j)})$, for each $j = 1, \dots, l$, the TARX model given in (2) becomes the TSARX model given in (1). Related to the restrictions given regarding the coefficients of A_{j2} of a TARX model, in a later subsection a statistical test is presented to determine multiplicative seasonality by regimes

2.2. The Conditional Likelihood Function

Let $\mathbf{x}_{1,T} = (x_1, \dots, x_T)$ and $\mathbf{z}_{1,T} = (z_1, \dots, z_T)$ be observed data vectors for the processes $\{X_t\}$ and $\{Z_t\}$, respectively, where T is the length of the sample period. Following Nieto (2005), it is assumed that the probabilistic mechanism that generates $\mathbf{z}_{1,T}$ does not depend on θ_x and the joint distribution of $\mathbf{x}_{1,T}$ conditional on $\mathbf{z}_{1,T}$ and θ , does not depend on θ_z . With these conditions it can be seen that, initially, it is possible to estimate the parameters of $\{Z_t\}$ and then conditional on these parameters, we proceed to estimate the parameters of the TSARX model considering the likelihood function¹

$$\begin{aligned} f(\mathbf{x}_{1,T}, \mathbf{z}_{1,T} \mid \theta) &= f(\mathbf{z}_{1,T} \mid \theta_x, \theta_z) f(\mathbf{x}_{1,T} \mid \mathbf{z}_{1,T}, \theta_x, \theta_z) \\ &= f(\mathbf{z}_{1,T} \mid \theta_z) f(\mathbf{x}_{1,T} \mid \mathbf{z}_{1,T}, \theta_x). \end{aligned}$$

Our interest is to calculate the conditional likelihood function of a TSARX($l; k_1, \dots, k_l; K_1, \dots, K_l; q_1, \dots, q_l$)_s model given in (1); that is to say, $f(\mathbf{x}_{1,T} \mid \mathbf{z}_{1,T}, \theta_x)$. Additionally, it is assumed that the first K values of $\mathbf{x}_{1,T}$; that is,

¹Future research can be directed to obtain joint estimates of the parameters θ_x and θ_z .

$\mathbf{x}_{1,K} = (x_1, \dots, x_K)$ are fixed and known with $K = \max\{k_j + sK_j, q_j, j = 1, \dots, l\}$; consequently,

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

Now, since $\{\epsilon_t\} \sim \text{GWN}(0, 1)$ for $t = (K+1), \dots, T$, the variable $x_t | x_{t-1}, \dots, x_1, \mathbf{z}_{1,T}$ has normal distribution with mean

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

and variance $(h^{(j_t)})^2$. The sequence $\{j_t\}$ is the observed time series for the stochastic process $\{J_t\}$, where $J_t = j$ if $r_{j-1} < z_{t-d} \leq r_j$, for some $j = 1, \dots, l$. Then,

$$\begin{aligned} f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) &= (2\pi)^{-\frac{(T-K)}{2}} \prod_{t=K+1}^T (h^{(j_t)})^{-1} \exp \left\{ \frac{1}{2} \sum_{t=K+1}^T \left(\frac{x_t - \mu_t}{h^{(j_t)}} \right)^2 \right\} \\ &= (2\pi)^{-\frac{(T-K)}{2}} \prod_{j=1}^l (h^{(j)})^{-n_j} \exp \left\{ -\frac{1}{2} \sum_{j=1}^l (\mathbb{X}_j - A_j^* B_j^*)' ((h^{(j)})^{-2} I_{n_j}) (\mathbb{X}_j - A_j^* B_j^*) \right\}, \end{aligned} \quad (3)$$

where

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

$\mathbb{X}_j = (x_{t_1,j}, \dots, x_{t_{n_j,j}})^'$, $\{t_{1,j}, \dots, t_{n_j,j}\}$ are the indexes in time, where $r_{j-1} < z_{t-d} \leq r_j$ and n_j is the number of observations in the regime j , $j = 1, \dots, l$. Henceforth, the conditional likelihood function of a TSARX model is given by (3).

2.3. Estimation of Parameters of a TSARX Model

Stochastic search ideas are used to identify autoregressive orders in a TSARX model with l regimes; in particular, the Gibbs variable selection method given by Dellaportas, Forster & Ntzoufras (2002) and denoted by GVS. This method achieves the estimation of autoregressive orders in a single stage and can involve these estimates in the MCMC sampling scheme, along with the remaining parameter estimates, assuming that the number of regimes is known. The GVS method is a hybrid between the method for the selection of variables proposed by Kuo & Mallick (1998), denoted KM, and the stochastic search for the selection of

variables (SSVS) given by George & McCulloch (1993). These and other Bayesian model selection techniques can be reviewed in O’Hara & Sillanpää (2009).

So & Chen (2003) use the SSVS method to determine the autoregressive orders and the remaining parameters, when the number of regimes in the SETAR models is known. Calderón & Nieto (2017) use the KM and GVS methods in multivariate TARX models.

The GVS method for a TSARX model with l regimes, is described as follows: let k_j, K_j and $q_j, j = 1, \dots, l$, be the maximum autoregressive orders. We define the 0 – 1 latent indicator variables s_{ij}, m_{uj} and n_{vj} , for each $i = 0, 1, \dots, k_j, u = 1, \dots, K_j$ and $v = 1, \dots, q_j$ in the following way:

$$s_{ij} = \begin{cases} 1, & \text{then } a_i^{(j)} \text{ is included} \\ & \text{in the model,} \\ 0, & \text{otherwise,} \end{cases} \quad m_{uj} = \begin{cases} 1, & \text{then } b_u^{(j)} \text{ is included} \\ & \text{in the model,} \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{and } n_{vj} = \begin{cases} 1, & \text{then } c_v^{(j)} \text{ is included} \\ & \text{in the model,} \\ 0, & \text{otherwise.} \end{cases}$$

Let $S_j = (s_{0j}, s_{1j}, \dots, s_{k_jj})', M_j = (m_{1j}, \dots, m_{K_jj})'$ and $N_j = (n_{1j}, \dots, n_{q_jj})', j = 1, \dots, l$, latent indicator vectors associated with nonseasonal autoregressive, seasonal autoregressive and the exogenous input coefficients, respectively.

Model (1) is rewritten with the vectors S_j, M_j and $N_j, j = 1, \dots, l$ as

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

for some $j = 1, \dots, l$ and the other specifications given in subsection 2.1 are met, except that the vector of structural parameters of the model is now

$$\theta_s = (S_1, \dots, S_l, M_1, \dots, M_l, N_1, \dots, N_l, \mathbf{r}', d, s, l)'$$

To use the GVS method, it is necessary to assume that:

$$a_i^{(j)} | s_{ij} \sim (1 - s_{ij})N(0, 1/\varphi_{ij}^2) + s_{ij}N(0, 1/\eta_{ij}^2 \varphi_{ij}^2),$$

$$b_u^{(j)} | m_{uj} \sim (1 - m_{uj})N(0, 1/\chi_{uj}^2) + m_{uj}N(0, 1/\xi_{uj}^2 \chi_{uj}^2), \text{ and}$$

$$c_v^{(j)} | n_{vj} \sim (1 - n_{vj})N(0, 1/\psi_{vj}^2) + n_{vj}N(0, 1/\lambda_{vj}^2 \psi_{vj}^2),$$

where $Pr(s_{ij} = 1) = \pi_{ij}, Pr(m_{uj} = 1) = \lambda_{uj}$ and $Pr(n_{vj} = 1) = \omega_{vj}, i = 0, 1, \dots, k_j, u = 1, \dots, K_j, v = 1, \dots, q_j, j = 1, \dots, l$. According to George & McCulloch (1993), the interpretation of this formulation is as follows. First, we

set $1/\varphi_{ij}^2$ small so that if $s_{ij} = 0$, then $a_i^{(j)}$ would probably be so small that it could be safely estimated by 0. Second, we set $1/\eta_{ij}^2\varphi_{ij}^2$ large so that if $s_{ij} = 1$, then a non-0 estimate of $a_i^{(j)}$ should probably be included in the final model. A similar deduction can happen for the other mixtures of normal distributions. Additionally, those mixtures of normal distributions can be carried to a multivariate structure in the following way:

$$\begin{aligned} A_j | S_j &\sim N_{k_j+1}(\mathbf{0}, (D_{S_j} V_{A_{j0}} D_{S_j})^{-1})', \\ B_j | M_j &\sim N_{K_j}(\mathbf{0}, (D_{M_j} V_{B_{j0}} D_{M_j})^{-1})', \text{ and} \\ C_j | N_j &\sim N_{q_j}(\mathbf{0}, (D_{N_j} V_{C_{j0}} D_{N_j})^{-1})', \end{aligned}$$

where $V_{A_{j0}}$, $V_{B_{j0}}$ and $V_{C_{j0}}$ are matrices of prior covariances, D_{S_j} is a diagonal matrix $\text{diag}\{1/a_{0j}\varphi_{0j}, \dots, 1/a_{k_{ij}}\varphi_{k_{ij}}\}$ with $a_{ij} = 1$ if $s_{ij} = 0$ and $a_{ij} = \eta_{ij}$ if $s_{ij} = 1$, $D_{M_j} = \text{diag}\{1/b_{1j}\chi_{1j}, \dots, 1/b_{K_{jj}}\chi_{K_{jj}}\}$ with $b_{uj} = 1$ if $m_{uj} = 0$ and $b_{uj} = \xi_{uj}$ if $m_{uj} = 1$, and $D_{N_j} = \text{diag}\{1/c_{1j}\psi_{1j}, \dots, 1/c_{q_{jj}}\psi_{q_{jj}}\}$ with $c_{vj} = 1$ if $n_{vj} = 0$ and $c_{vj} = \lambda_{vj}$ if $n_{vj} = 1$. The above is valid, assuming priori independence of the s_{ij} , m_{uj} , n_{vj} , $i = 0, 1, \dots, k_j$, $u = 1, \dots, K_j$, $v = 1, \dots, q_j$, $j = 1, \dots, l$.

The conditional likelihood function of a TSARX model with l regimes considering the vectors S_j , M_j and N_j , $j = 1, \dots, l$ is given in (3), except now,

$$\begin{aligned} B_j^* &= (a_0^{(j)} s_{0j}, a_1^{(j)} s_{1j}, \dots, a_{k_j}^{(j)} s_{k_{jj}}, 0, \dots, 0, b_1^{(j)} m_{1j}, -a_1^{(j)} b_1^{(j)} s_{1j} m_{1j}, \dots, \\ &\quad - a_{k_j}^{(j)} b_1^{(j)} s_{k_{jj}} m_{1j}, \dots, 0, \dots, 0, b_{K_j}^{(j)} m_{K_{jj}}, -a_1^{(j)} b_{K_j}^{(j)} s_{1j} m_{K_{jj}}, \dots, \\ &\quad - a_{k_j}^{(j)} b_{K_j}^{(j)} s_{k_{jj}} m_{K_{jj}}, c_1^{(j)} n_{1j}, \dots, c_{q_j}^{(j)} n_{q_{jj}})', \end{aligned}$$

with, k_j , K_j , and q_j , $j = 1, \dots, l$, defined previously.

Once MCMC samples are obtained for S_j , M_j , N_j , $j = 1, \dots, l$, the autoregressive orders of the model can be identified considering a particular combination of S_j , M_j and N_j , $j = 1, \dots, l$ which possesses the highest posterior probability.

Prior independence is assumed for A_j , B_j and C_j so $V_{A_{j0}} = \mathbb{I}_{k_j+1}$, $V_{B_{j0}} = \mathbb{I}_{K_j}$ and $V_{C_{j0}} = \mathbb{I}_{q_j}$, $j = 1, \dots, l$ (\mathbb{I} is the identity matrix). Also, it is assumed that each autoregressive coefficient is equally likely to be included in the model; then, $\pi_{ij} = 1/2$, $\lambda_{uj} = 1/2$ and $\omega_{vj} = 1/2$, $i = 0, 1, \dots, k_j$, $u = 1, \dots, K_j$, $v = 1, \dots, q_j$, $j = 1, \dots, l$.

The prior distributions for parameter groups A_j , B_j , C_j , $(h^{(j)})^2$, $j = 1, \dots, l$, \mathbf{r} and d are:

- (i) For $A_j | S_j$ a multivariate normal distribution with mean $\mathbf{0}_{k_j+1}$ and covariance $(D_{S_j} D_{S_j})^{-1}$, $j = 1, \dots, l$.
- (ii) For $B_j | M_j$ a multivariate normal distribution with mean $\mathbf{0}_{K_j}$ and covariance $(D_{M_j} D_{M_j})^{-1}$, $j = 1, \dots, l$.
- (iii) For $C_j | N_j$ a multivariate normal distribution with mean $\mathbf{0}_{q_j}$ and covariance $(D_{N_j} D_{N_j})^{-1}$, $j = 1, \dots, l$.

- (iv) For $(h^{(j)})^2$ an inverse Gamma distribution (IG) with position parameter $\nu_j/2$ and scale parameter $\nu_j\lambda_j/2$, $j = 1, \dots, l$.
- (v) $\mathbf{r} = (r_1, \dots, r_{l-1})'$ follows the constant prior distribution $f(\mathbf{r}) = \frac{1}{C}I(\mathbf{A})$ where $C = \int \dots \int_{\mathbf{A}} dr_1 \dots dr_{l-1}$, and \mathbf{A} is a region that satisfies:
 - (1) $a \leq r_1 < \dots < r_{l-1} \leq b$, with a and b convenient quantiles of Z_{t-d} , and
 - (2) each regime contains at least one $H\%$ from the sample of Z_{t-d} (this to ensure valid inference). The value of H is preset.

$I(\mathbf{A})$ is the indicator function of \mathbf{A} on \mathbb{R}^{l-1} . For the particular case of two regimes, $f(r)$ is a uniform distribution defined in (a, b) .

- (vi) d follows a discrete uniform distribution, with probability mass function $Pr(d = i) = 1/(d_0 + 1)$, for $i = 0, 1, \dots, d_0$; d_0 being a maximum delay given.

The hyperparameters $\{\varphi_{ij}, \eta_{ij}, \chi_{uj}, \xi_{uj}, \psi_{vj}, \lambda_{vj}, \nu_j, \lambda_j, (i = 0, 1, \dots, k_j, u = 1, \dots, K_j, v = 1, \dots, q_j, j = 1, \dots, l), a, b, d_0\}$ are assumed known. Combining the conditional likelihood function for a TSARX model, considering the vectors S_j , M_j and N_j , $j = 1, \dots, l$, with each of the prior distributions given in (i) to (vi), the posterior distributions of the unknown parameters of the model are obtained. MCMC samples of A_j , B_j , C_j , $(h^{(j)})^2$, S_j , M_j , N_j , $j = 1, \dots, l$, \mathbf{r} and d are generated from convenient complete conditional distributions as follows.

Algorithm 1.

1. Initialize: $A_j^{(0)}$, $B_j^{(0)}$, $C_j^{(0)}$, $[(h^{(j)})^2]^{(0)}$, $S_j^{(0)}$, $M_j^{(0)}$, $N_j^{(0)}$, $j = 1, \dots, l$, $\mathbf{r}^{(0)}$, $d^{(0)}$.
2. Generate A_j (independent of A_i , $i \neq j$) from

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

where $V_{A_j} = [1/(h^{(j)})^2]S_j^* S_j^* + D_{S_j} D_{S_j}$ and $\boldsymbol{\theta}_{A_j} = V_{A_j}^{-1}[1/(h^{(j)})^2]S_j^* (\mathbb{X}_j - M_j^* B_j - N_j^* C_j)$, $s_t^* = (s_{0j}, s_{1j}x_{t-1} - \sum_{u=1}^{K_j} b_u^{(j)} s_{1j} m_{uj} x_{t-1-su}, \dots, s_{k_j j} x_{t-k_j - \sum_{u=1}^{K_j} b_u^{(j)} s_{k_j j} m_{uj} x_{t-k_j-su})$, $S_j^* = (s_{t_1, j}^*, \dots, s_{t_{n_j}, j}^*)'$, $m_t^* = (m_{1j} x_{t-s}, m_{2j} x_{t-s-2}, \dots, m_{K_j j} x_{t-sK_j})$, $M_j^* = (m_{t_1, j}^*, \dots, m_{t_{n_j}, j}^*)'$, $n_t^* = (n_{1j} z_{t-1}, \dots, n_{q_j j} z_{t-q_j})$, $N_j^* = (n_{t_1, j}^*, \dots, n_{t_{n_j}, j}^*)'$, $\mathbb{X}_j = (x_{t_1, j}, \dots, x_{t_{n_j}, j})'$ and n_j the number of observations in the regime j , $j = 1, \dots, l$.

Here $\boldsymbol{\theta}_{x-A_j}$ is the parameter vector $\boldsymbol{\theta}_x$ without the subvector A_j .

3. Generate B_j (independent of B_i , $i \neq j$) from

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

where $V_{B_j} = [1/(h^{(j)})^2]T_j^* T_j^* + D_{M_j} D_{M_j}$ and $\boldsymbol{\theta}_{B_j} = V_{B_j}^{-1}[1/(h^{(j)})^2]T_j^* (\mathbb{X}_j - U_j^* A_j - N_j^* C_j)$, $t_t^* = (m_{1j} x_{t-s} - \sum_{i=1}^{k_j} a_i^{(j)} s_{ij} m_{1j} x_{t-i-s}, \dots, m_{K_j j} x_{t-sK_j} -$

$\sum_{i=1}^{k_j} a_i^{(j)} s_{ij} m_{K_{ij}} x_{t-i-s_{K_j}}$, $T_j^* = (t_{t_1,j}^*, \dots, t_{t_{n_j},j}^*)'$, $\mu_t^* = (s_{0j}, s_{1j}x_{t-1}, \dots, s_{k_j,j}x_{t-k_j})$, $U_j^* = (\mu_{t_1,j}^*, \dots, \mu_{t_{n_j},j}^*)'$, N_j^*, \mathbb{X}_j and $n_j, j = 1, \dots, l$ are defined above.

4. Generate C_j (independent of $C_i, i \neq j$) from

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

where $V_{C_j} = [1/(h^{(j)})^2]N_j^* N_j^* + D_{N_j} D_{N_j}$ and $\boldsymbol{\theta}_{C_j} = V_{C_j}^{-1}[1/(h^{(j)})^2]N_j^* (\mathbb{X}_j - F_j^* E_j - U_j^* A_j - M_j^* B_j)$, $f_t^* = (0, \dots, 0, s_{1j}m_{1j}x_{t-1-s}, \dots, s_{k_j,j}m_{1j}x_{t-k_j-s}, 0, \dots, 0, s_{1j}m_{K_{ij}}x_{t-1-s_{K_j}}, \dots, s_{k_j,j}m_{K_{ij}}x_{t-k_j-s_{K_j}})$, $F_j^* = (f_{t_1,j}^*, \dots, f_{t_{n_j},j}^*)'$, $E_j = (a_{1+k_j}^{(j)}, \dots, a_{s-1}^{(j)}, -a_1^{(j)}b_1^{(j)}, \dots, -a_{k_j}^{(j)}b_1^{(j)}, \dots, a_{(1+k_j)+s(K_j-1)}^{(j)}, \dots, a_{(s-1)+s(K_j-1)}^{(j)}, -a_1^{(j)}b_{K_j}^{(j)}, \dots, -a_{k_j}^{(j)}b_{K_j}^{(j)})$, $N_j^*, U_j^*, M_j^*, \mathbb{X}_j$ y $n_j, j = 1, \dots, l$ are defined above.

5. Generate $(h^{(j)})^2$ (independent of $(h^{(i)})^2, i \neq j$) from

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

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

6. Generate \mathbf{r} from

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

where $f(\cdot \mid \cdot)$ is the conditional likelihood function of the TSARX model taking into account, that S_j, M_j and $N_j, j = 1, \dots, l$ are latent indicator variables and \mathbf{A} is the region specified in the prior distribution for \mathbf{r} .

7. Generate d from the multinomial distribution, with probability mass function

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

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

8. Generate vector $S_j = (s_{0j}, s_{1j}, \dots, s_{k_j,j})'$ (independent from S_i, M_i and $N_i, i \neq j$), obtaining samples from $s_{ij}, i = 0, 1, \dots, k_j, j = 1, \dots, l$ individually based on the Bernoulli distribution, with probability mass function

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

where

$$\alpha_{ij} = f(\mathbf{x}_{K+1,T} \mid \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-s_{ij}}, s_{ij} = 1) f(A_j \mid S_{j-s_{ij}}, s_{ij} = 1) \pi_{ij}$$

and

$$\beta_{ij} = f(\mathbf{x}_{K+1,T} \mid \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-s_{ij}}, s_{ij} = 0) f(A_j \mid S_{j-s_{ij}}, s_{ij} = 0) (1 - \pi_{ij}).$$

9. Generate vector $M_j = (m_{1j}, m_{2j}, \dots, m_{K_jj})'$ (independent from M_i , S_i and N_i , $i \neq j$), obtaining samples from m_{uj} , $u = 1, \dots, K_j$, $j = 1, \dots, l$ individually based on the Bernoulli distribution, with probability mass function

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

where

$$\gamma_{uj} = f(\mathbf{x}_{K+1,T} \mid \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-m_{uj}}, m_{uj} = 1) f(B_j \mid M_{j-m_{uj}}, m_{uj} = 1) \lambda_{uj}$$

and

$$\varrho_{uj} = f(\mathbf{x}_{K+1,T} \mid \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-m_{uj}}, m_{uj} = 0) f(B_j \mid M_{j-m_{uj}}, m_{uj} = 0) (1 - \lambda_{uj}).$$

10. Generate vector $N_j = (n_{1j}, n_{2j}, \dots, n_{q_jj})'$ (independent from N_i , S_i and M_i , $i \neq j$), obtaining samples from n_{vj} , $v = 1, 2, \dots, q_j$, $j = 1, \dots, l$ individually based on the Bernoulli distribution, with probability mass function

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

where

$$\varsigma_{vj} = f(\mathbf{x}_{K+1,T} \mid \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-n_{vj}}, n_{vj} = 1) f(C_j \mid N_{j-n_{vj}}, n_{vj} = 1) \omega_{vj}$$

and

$$\xi_{vj} = f(\mathbf{x}_{K+1,T} \mid \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x-n_{vj}}, n_{vj} = 0) f(C_j \mid N_{j-n_{vj}}, n_{vj} = 0) (1 - \omega_{vj}).$$

All the posterior distributions of Algorithm 1, except that of \mathbf{r} , are standard. To simulate values of \mathbf{r} , a Gaussian random walk Metropolis algorithm with step sizes $\sigma_{r_1}, \dots, \sigma_{r_{l-1}}$ is used (Metropolis, Rosenbluth, Rosenbluth & Teller 1953). We set the MCMC sample size N sufficiently large, discarding the burn-in iterates, and keep the last $G = N - M$ iterates for posterior analysis.

The means and standard deviations of the posterior distributions of A_j , B_j , C_j , $(h^{(j)})^2$, $j = 1, \dots, l$ and \mathbf{r} are calculated to obtain point estimates and a measure of uncertainty for such estimates, respectively. Also, with the samples generated, $100(1 - \alpha)\%$ credible intervals or $100(1 - \alpha)\%$ Highest-Density Regions (HDR), for α , $0 < \alpha < 1$, can be obtained. The estimation of the delay parameter d is taken as the posterior mode and for the estimates of the vectors of zeros and ones, S_j , M_j , N_j , $j = 1, \dots, l$ the highest occurrence probabilities are taken.

In order to examine each of the parameters, if the MCMC samples obtained converge to the posterior distribution, the Geweke's Z-score plot (Geweke 1992)

of the *geweke.plot* function of *coda* in R software (Plummer, Best, Cowles & Vines 2006) is used. Also, we review the evolution of the point estimate (posterior averages) of each of the parameters, when the number of iterations is increased, by means of the *cumuplot* function of *coda* in R. This plot presents the sample quantiles 0.025, 0.50 and 0.975, as a function of the number of iterations (accumulated averages) of the Markov chains.

2.4. A Bayesian Test of Multiplicative Seasonality by Regimes

De Gooijer & Vidiella-i Anguera (2003) using a frequentist approach, present a model that captures multiplicative seasonality by regimes and nonlinearity of thresholds simultaneously in SETAR models.

In this Subsection, to test the presence of multiplicative seasonality in each of the regimes of the TARX model, a Bayesian statistical test is proposed where an approximate technique is used for the calculation of the Bayes factor. In particular, we use the Savage-Dickey distributions ratio given in Dickey (1971) paper, which is an approximation of the Bayes factor (see, Verdinelli & Wasserman, 1995).

The methodology consists in comparing the TSARX model with a model where such seasonality is not captured (a TARX model). That is, that the parameter vector A_{j2} defined for the TARX model given in (2), equals the vector

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

for each $j = 1, \dots, l$, so that a TSARX model is a special case of a TARX model.

Formally, we want to test the hypotheses $H_{0j} : A_{j2} = A_{j20}$ versus $H_{aj} : A_{j2} \neq A_{j20}$, for each $j = 1, \dots, l$. The Bayes factor is given by

$$FB_j = \frac{\int f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x) f(\boldsymbol{\theta}_x) d\boldsymbol{\theta}_x}{\int f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_{x_0}) f(\boldsymbol{\theta}_{x_0}) d\boldsymbol{\theta}_{x_0}}, \quad (4)$$

where, respectively, $f(\cdot)$ and $f(\cdot | \cdot)$ denote the prior distribution and the conditional likelihood function of the TARX and $\boldsymbol{\theta}_{x_0}$ summarizes the parameters of the TARX model in the case $A_{j2} = A_{j20}$, for $j = 1, \dots, l$.

The Bayes factor (4) is equal to

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

for $j = 1, \dots, l$; that is, the ratio of the weights of the conditional posterior distribution of A_{j2} and the weights of the prior distribution of A_{j2} , both evaluated in A_{j20} , $j = 1, \dots, l$.

The weight of the marginal posterior distribution of A_{j2} in A_{j20} can be obtained with the calculation of the average of the conditional posterior distribution of A_{j2} over the MCMC samples of the other parameters; thus,

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

for $j = 1, \dots, l$, if the complete conditional posterior distribution of A_{j2} has a known standard distribution (see, Gelfand & Smith, 1990) and $\theta_{x-A_{j2}}$ is the parameter vector θ_x without subvector A_{j2} . Here it is necessary to obtain MCMC samples of the posterior distributions of A_{j2} . Given the number of regimes l and autorregressive orders $k_j + sK_j, q_j, j = 1, \dots, l$, the remaining parameters of the TARX model defined in (2), are estimated using a Gibbs sampler.

Returning to the Bayes factor calculation, because

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

the parameters θ_{j2} and V_{j2} are used from the posterior distribution and θ_{j20} and V_{j20} from the prior distribution for A_{j2} , for the calculation of the Bayes factor by regimes as indicated in equation (4). Thus,

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

$j = 1, \dots, l$. $H_{0j} : A_{j2} = A_{j20}$ is not rejected if $2 \ln(FB_j) > 6$, (strong or very strong evidence in favor of H_{0j}) for each $j = 1, \dots, l$, according to the decision criterion given by Kass & Raftery (1995).

We decide to use a TSARX($l; k_1, \dots, k_l; K_1, \dots, K_l; q_1, \dots, q_l$)_s model to adjust the data, instead of a TARX($l; k_1 + sK_1, \dots, k_l + sK_l; q_1, \dots, q_l$) model, if in all the regimes j , H_{0j} is not rejected, $j = 1, \dots, l$.

2.5. Estimation of the Number of Regimes

The identification of l , the number of regimes, can be seen as a problem of Bayesian model choice. The decision among models $M_j, j = 1, \dots, l_0$, M_j being a TSARX model with j regimes and l_0 being a maximum number of regimes, is performed by the calculation of posterior probabilities, where the model M_j is chosen if $Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T})$ is the highest one over $j = 1, \dots, l_0$. Congdon (2006) developed the following posterior probability estimator, which requires separate independent MCMC samples for each of the candidate models and is denoted as $\{\theta^{(i)} = (\theta_1^{(i)}, \dots, \theta_{l_0}^{(i)}), i = 1, \dots, G\}$. An approximate Monte Carlo estimate of $Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T})$ is then

$$Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T}) \approx \frac{1}{G} \sum_{i=1}^G Pr(M_j | \mathbf{x}_{K+1,T}, \mathbf{z}_{1,T}, \theta_j^{(i)}) = \frac{1}{G} \sum_{i=1}^G \frac{\varphi_j^{(i)}}{\sum_{k=1}^{l_0} \varphi_k^{(i)}}, \quad (5)$$

where $\varphi_k^{(i)} = f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \theta_k^{(i)}, M_k) f(\theta_k^{(i)} | M_k) Pr(M_k)$, θ_k is the parameter vector of model k , $\theta_k^{(i)}$ is the MCMC i -th iteration of the posterior distribution of model k , $Pr(M_k)$ is chosen as $Pr(M_k) = \frac{1}{l_0}$, $f(\theta_k^{(i)} | M_k)$ is the prior distribution for model k and $f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \theta_k^{(i)})$ is the conditional

likelihood function for M_k , for each $k = 1, \dots, l_0$. For numerical efficiency, the calculation of (5) is based on scaled versions of the log-likelihoods. The scale uses the maximum $L_{\max}^{(i)}$ of the log-likelihoods in each iteration and subtracts this from each log-likelihood model, that is

$$\varphi_k^{(i)} = \exp(L_k^{(i)} - L_{\max}^{(i)})f(\theta_k^{(i)} | M_k)Pr(M_k), \quad i = 1, \dots, G, \text{ for some } k = 1, \dots, l_0.$$

Some results support that the method is quite accurate and powerful for model selection. For example, Chen et al. (2010) compare TARX models with GARCH errors for different regimes, while Gerlach & Chen (2008) use it for STAR models with GARCH errors. The effectiveness of this approach is examined, comparing it with the deviance information criterion (DIC), which is described below. Spiegelhalter, Best, Carlin & Van Der Linde (2002) propose a Bayesian model comparison criterion defined as:

$$DIC = 2E_{\theta_j | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}}(D(\theta_j)) - D(E_{\theta_j | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}}(\theta_j)), \quad (6)$$

where $D(\tilde{\mathbf{A}}_j) = -2 \ln f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \tilde{\mathbf{A}}_j)$, with $f(\cdot | \cdot)$ the conditional likelihood function of the TSARX model with j regimes, $j = 1, \dots, l_0$. If $\theta_j^{(1)}, \dots, \theta_j^{(G)}$ are MCMC samples from the distribution $f(\theta_j | \mathbf{x}_{1,T}, \mathbf{z}_{1,T})$, then the DIC can be approached via Monte Carlo by

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

where $D(\tilde{\mathbf{A}}_j^{(i)}) = -2 \ln f(\mathbf{x}_{K+1,T} | \mathbf{x}_{1,K}, \mathbf{z}_{1,T}, \tilde{\mathbf{A}}_j^{(i)})$, $i = 1, \dots, G$, $j = 1, \dots, l_0$. The model M_j , $j = 1, \dots, l_0$ is chosen if its DIC is the smallest among the candidate models.

2.6. Diagnostic Checking

Testing for model adequacy is an important part of any time series analysis. Based on MCMC techniques and Bayesian inference, it is natural to consider techniques that take into account the uncertainty of the parameter estimators in the calculation of the residuals. In this paper, the method of Gerlach, Carter & Kohn (1999) is used, which is based on MCMC techniques, importance sampling and the time series $u_t = F_t(x_t | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1})$, $t = (K + 1), \dots, T$, where $F_t(\cdot | \cdot)$ is the cumulative conditional distribution function given in Subsection 2.1, and $K = \max\{k_j + sK_j, q_j, j = 1, \dots, l\}$. Gerlach et al. (1999) show that for $k \geq t$,

$$\hat{u}_t = \frac{\sum_{i=1}^G F_t(x_t | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, \theta_k^{(i)})/f(x_{t,k} | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, \theta_k^{(i)})}{\sum_{i=1}^G 1/f(x_t | \mathbf{x}_{1,t-1}, \mathbf{z}_{1,t-1}, \theta_k^{(i)})} \quad (7)$$

converges in distribution to u_t when $G \rightarrow \infty$. Here $\theta_k^{(i)}$ is the i -th MCMC iterated sampled from the posterior distribution $f(\theta_k | \mathbf{x}_{1,k}, \mathbf{z}_{1,k})$ for a given k and

$f(\cdot | \cdot)$ is the conditional likelihood function for the TSARX model with l regimes. Because the variance of \hat{u}_t in equation (7) increases with $(k - t)$, it is necessary to calculate \hat{u}_t with t reasonably close to k . This can be achieved by increasing k sequentially, say to 100, 200, \dots , T and evaluate \hat{u}_t using the equation (7) with $(k - t)$ not greater than the increments. Based on the convergence properties of \hat{u}_t , the generalized residuals $\hat{v}_t = \Phi^{-1}(\hat{u}_t)$ (Φ is the standard cumulative normal distribution function) are approximately independent and identically distributed $N(0, 1)$, under the correct model. Standard diagnostic tests can be applied to \hat{v}_t such as the Ljung-Box test to determine autocorrelation, the Jarque-Bera test to verify normality, the CUSUM plot to determine the correct specification of model and the CUSUMSQ plot to check marginal heteroscedasticity, among others. Chen & So (2006) and Chen et al. (2010) conduct the diagnostic check on TARX models with GARCH errors using generalized residuals and their results were very satisfactory for determining the adequacy of the model.

2.7. Forecasting

Here, the idea is to find $\hat{X}_{T+h} = E(X_{T+h} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T})$, where $h \geq 1$ is the forecast horizon, $\mathbf{x}_{1,T} = (x_1, \dots, x_T)$, $\mathbf{z}_{1,T} = (z_1, \dots, z_T)$, and T is the length of the sample period. \hat{X}_{T+h} is the best prediction in the sense of the minimum value that minimizes the quadratic lose function and its analytical expression is difficult or impossible to obtain in the context of nonlinear time series (see, Franses & Van Dijk 2000). For this reason, we focus on the so-called predictive distribution of X_{T+h} , that is, $f(x_{T+h} | \mathbf{x}_{1,T}, \mathbf{z}_{1,T})$. The predictive distributions of interest can be obtained marginally from the joint predictive distribution

$$f(x_{T+1}, \dots, x_{T+h}, z_{T+1}, \dots, z_{T+h}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z | \mathbf{x}_{1,T}, \mathbf{z}_{1,T}), \tag{8}$$

with $\boldsymbol{\theta}_x$ the parameters vector of the TSARX model and $\boldsymbol{\theta}_z$ the parameters vector of the variable Z_t . Previous work in this direction is presented in Nieto (2008) and Vargas (2012), among others. It will be assumed in what follows, that the specifications given in previous subsections, with respect to the chosen model, are fulfilled. To obtain samples from the predictive distribution (8), we take into account that the joint predictive distribution can be expressed as

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

where

- (i) $f(z_{T+i} | \mathbf{z}_{1,T+i-1}, \boldsymbol{\theta}_z)$ is the kernel density function of the Markov chain $\{Z_t\}$,

(ii) $f(x_{T+i} \mid z_{T+i}, \mathbf{x}_{1,T+i-1}, \mathbf{z}_{1,T+i-1}, \boldsymbol{\theta}_x)$ is a normal distribution with mean

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

and variance $(h^{(j)})^2$ if $r_{j-1} < z_{T+i-d} \leq r_j$, for some $j = 1, \dots, l$,

(iii) $f(\boldsymbol{\theta}_x \mid \mathbf{x}_{1,T}, \mathbf{z}_{1,T})$ is the posterior distribution of the parameters of the TSARX model;

Also, $\mathbf{x}_{1,T+i-1} = (x_1, \dots, x_{T+i-1})$ and $\mathbf{z}_{1,T+i-1} = (z_1, \dots, z_{T+i-1})$. Thus, from the joint predictive distribution, $z_{T+1}, x_{T+1}, z_{T+2}, x_{T+2}, \dots, z_{T+h}, x_{T+h}$ can be extracted sequentially as follows (Congdon 2007):

Algorithm 2.

1. Generate a MCMC sample of $\boldsymbol{\theta}_x^{(k)}$ of the posterior distribution $f(\boldsymbol{\theta}_x \mid \mathbf{x}_{1,T}, \mathbf{z}_{1,T})$.
2. Generate $z_{T+1}^{(k)}$ from $f(z_{T+1} \mid \mathbf{z}_{1,T}, \boldsymbol{\theta}_z)$.
3. Generate $x_{T+1}^{(k)}$ from $f(x_{T+1} \mid z_{T+1}^{(k)}, \mathbf{x}_{1,T}, \mathbf{z}_{1,T}, \boldsymbol{\theta}_x^{(k)})$.
4. Generate $z_{T+i}^{(k)}$ from $f(z_{T+i} \mid \mathbf{z}_{1,T}, z_{T+1}^{(k)}, \dots, z_{T+i-1}^{(k)}, \boldsymbol{\theta}_z)$.
5. Generate $x_{T+i}^{(k)}$ from $f(x_{T+i} \mid z_{T+i}^{(k)}, \mathbf{x}_{1,T}, x_{T+1}^{(k)}, \dots, x_{T+i-1}^{(k)}, \mathbf{z}_{1,T}, z_{T+1}^{(k)}, \dots, z_{T+i-1}^{(k)}, \boldsymbol{\theta}_x^{(k)})$.
6. Repeat steps 4 and 5 for $i = 2, 3, \dots, h$.

With the samples $\{z_{T+h}^{(k)}, x_{T+h}^{(k)}\}$, $k = 1, \dots, G$, with G being the length of the chain and $h \geq 1$, we calculate the means (point forecasts), the standard deviations, and credible intervals or HDR, for both variables.

The procedure described above includes the uncertainty of the estimation of the model parameters in the calculation of the forecasts, as is discussed by Vargas (2012) in TAR models.

In summary, our proposed methodology for fitting a TSARX model to a time series of interest that exhibits seasonality, consists of the following steps:

- Step 1. An exploratory analysis is made for the threshold and interest variables, to detect the presence of seasonality. For this, plots of the time series, correlograms, and plots of the time series for seasonal periods are used.

- Step 2. Maximum autoregressive orders are fixed, which are obtained by fitting candidate ARX models to the data and choosing that order with minimum DIC. The threshold nonlinearity test proposed by Tsay (1998) is carried out and the presence of multiplicative seasonality is tested, using the statistical test proposed in Subsection 2.4.
- Step 3. We proceed to estimate the number of regimes, autoregressive orders and the remaining parameters of the TSARX model, following the procedures described in Subsections 2.3 and 2.5.
- Step 4. The estimated model is validated, using the generalized residuals as described in Subsection 2.6.
- Step 5. Finally, the fitted model is used to obtain forecasts of the threshold and interest variables, as indicated in Subsection 2.7.

All procedures were implemented in R package (Team 2016).

3. A Simulated Example

In this section we present a simulation example in order to illustrate the performance of the proposed methodology².

3.1. The model

We assume that $\{X_t\}$ follows the TSARX(2; 1, 1; 2, 1; 1, 3)₁₂ model given by

$$X_t = \begin{cases} 2.34 + 0.50X_{t-1} + 0.20X_{t-12} - 0.10X_{t-13} + 0.10X_{t-24} \\ -0.05X_{t-25} + 1.23Z_{t-1} + \epsilon_t, & \text{if } Z_{t-2} \leq 4.46, \\ -4.50 + 0.60X_{t-1} + 0.10X_{t-12} - 0.06X_{t-13} - 1.15Z_{t-1} \\ + 3.30Z_{t-2} - 1.92Z_{t-3} + 4\epsilon_t, & \text{if } Z_{t-2} > 4.46, \end{cases} \quad (9)$$

where $\{Z_t\}$ obeys the AR(1) model $Z_t = 1.80 + 0.60Z_{t-1} + a_t$, with $\{a_t\}$ a zero-mean Gaussian white noise process with variance 1, which is independent of $\{\epsilon_t\}$ and $r = 4.46$ is the median of Z_{t-2} . Notice that $\{Z_t\}$ is a homogeneous first-order Markov chain, with kernel density function $f_1(z_t | \mathbf{z}_{t-1}, \theta_z)$ and corresponds to a $N(1.80 + 0.60z_{t-1}, 1)$.

3.2. Exploratory Analysis of the Data

Time series plots, the sample autocorrelation function (ACF), the sample partial autocorrelation function (PACF), and the average mutual information index (AMI) were used to preliminarily determine the presence of seasonality.

²Other simulated examples are given in González (2019).

The AMI measures the mutual dependence between variables W_i and $W_{i-\tau}$, $\tau = 0, 1, \dots, \tau^*$, where τ^* is a specified maximum number of lags. If W_i and $W_{i-\tau}$ are independent, then knowing W_i does not give information about $W_{i-\tau}$ and viceversa, so their mutual information is zero. The *mutual* function of the package *tseriesChaos* in R (Di Narzo & Di Narzo 2013) is used to estimate the AMI.

In Figure 1, the two simulated time series are plotted, with a sample size $T = 600$ observations. Seasonal behavior is not detected at first sight. For $\{X_t\}$ (see Figure 2) it is observed that significant autocorrelations can be identified in the first lags and lag 12, essentially in the PACF. The AMI only shows the first three lags to be significant. For $\{Z_t\}$ (see Figure 2), only the first lag is significant, as indicated by the ACF, PACF and AMI plots. Now, looking at Figure 3(a), (the *armasubsets* function of *TSA* in R) each row in this plot represents an AR model; the shaded rectangles in the columns indicate the variables included in the given model. The numbers on the left margin are the values of Bayesian Information Criterion (BIC). We have that an appropriate model for $\{X_t\}$ is a multiplicative AR(13), according to the minimum BIC. Additionally, at Figure 3(b), the box plots by months indicate that the function of the means of the series is not constant, which leads to suggest the presence of a seasonal cycle in the process $\{X_t\}$.

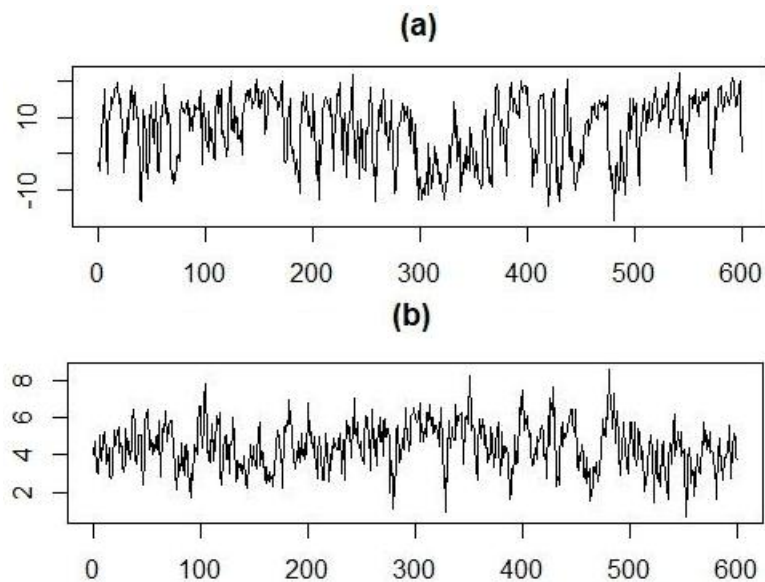


FIGURE 1: (a) X_t variable, (b) Z_t variable.

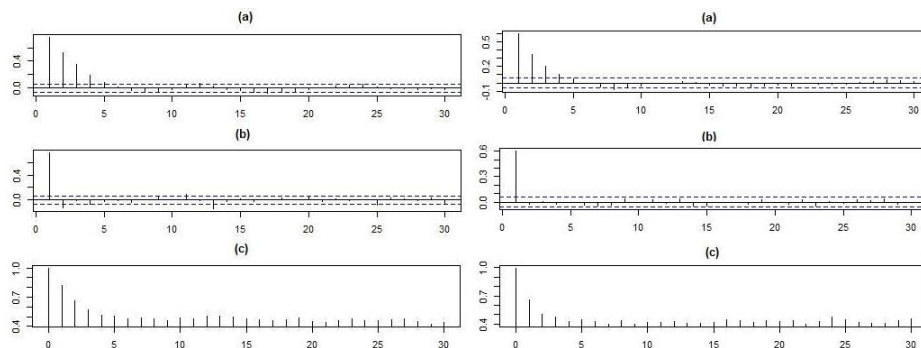


FIGURE 2: For X_t (left) and Z_t (right), (a) sample ACF, (b) sample PACF, (c) AMI sample.

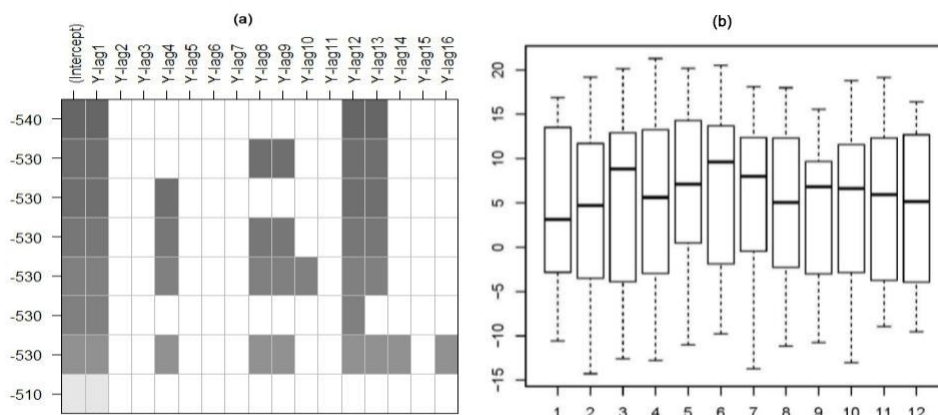


FIGURE 3: For X_t , (a) multiplicative AR with minimum BIC, (b) box plots by months.

3.3. Estimation of the TSARX Model Parameters

100 simulated time series of sample size $T = 600$ are generated from model (9). For each simulated time series, the MCMC sampler is run for 12.000 iterations and the first 6.000 iterations are taken as the burn-in period. The posterior inference is based on the remaining 6.000 iterations. The prespecified maximum autoregressive orders k_1 , K_1 , and q_1 , are chosen from an $ARX(k_1 + sK_1; q_1)$ model that best fits the simulated data, according to the minimum DIC. We take candidate ARX models, with $k_1 = 0, 1, 2, 3, 4, 5$, $K_1 = 0, 1, 2, 3$, and $q_1 = 0, 1, 2, 3, \dots, 15$ and $s = 12$. The values $k_1 = 2$, $K_1 = 2$, and $q_1 = 3$ were obtained. The Tsay threshold nonlinearity test (Tsay 1998) is applied for delay $d = 0, 1, 2, 3$. The null hypothesis of a linear $ARX(24; 3)$ model is rejected for $d = 2$ at the 1% significance level. In the Bayesian test of multiplicative seasonality by regimes, it is required to estimate parameters of a TARX model with two regimes, with $k_1 = k_2 = 2$, $K_1 = K_2 = 2$, $q_1 = q_2 = 3$, and $s = 12$.

TABLE 1: Averages of posterior probabilities and DICs for selection of the number of regimes (100 replications).

	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
Aver. post. prob.	0.00026	0.95415	0.03290	0.00429	0.00840
Average DIC	3569.24	2603.58	2608.90	2700.77	2704.05

For the hyperparameters the following values were taken: $\theta_{j10} = \mathbf{0}_{k_j+K_j+1}$, $V_{j10} = \text{diag}\{0.10\}_{k_j+K_j+1}$, $\theta_{j20} = \mathbf{0}_{K_j(s-1)}$, $V_{j20} = \text{diag}\{0.10\}_{K_j(s-1)}$, $\theta_{j0} = \mathbf{0}_{q_j}$, $V_{j0} = \text{diag}\{0.10\}_{q_j}$, $j = 1, 2$; also $\nu_j = 3$, $\lambda_j = \frac{\tilde{\sigma}^2}{3}$, where $\tilde{\sigma}^2$ is the sample variance of $\{X_t\}$, $j = 1, 2$; $a = \min\{Z_{t-d}\}$, $b = \max\{Z_{t-d}\}$, $H = 10$ (assures at least 60 observations in each regime) and $d_0 = 5$. The prior distributions are noninformative, thus the likelihood dominates the inference. The initial values for the parameters were: $A_{j1}^{(0)} = 0.05_{k_j+K_j+1}$, $A_{j2}^{(0)} = 0.05_{K_j(s-1)}$, $C_j^{(0)} = 0.05_{q_j}$, $[(h^{(j)})^2]^{(0)} = 0.20$, $j = 1, 2$; $d^{(0)} = 3$; the step size to execute the Metropolis algorithm for drawing samples of r was $\sigma_r = 0.025$, and $r^{(0)}$ the 0.50 quantile of Z_{t-d} . When we run the Bayesian test for the 100 simulated series, the presence of multiplicative seasonality it is detected in both regimes, with average values of $2\ln(FB_1)=3725.30$ and $2\ln(FB_2) = 15072.52$. In the identification of the number of regimes, separate and independent MCMC samples are used for the TSARX candidate models, with $l = 1, 2, 3, 4, 5$ regimes ($l_0 = 5$). Additionally, we set $\varphi_{ij} = 25$, $\eta_{ij} = 1.50$ for all i, j , $\chi_{uj} = 25$, $\xi_{uj} = 1.50$ for all u, j , $\psi_{vj} = 25$, $\lambda_{vj} = 1.50$ for all v, j , $i = 0, 1, 2$, $u = 1, 2$, $v = 1, 2, 3$, and $j = 1, 2, 3, 4, 5$. These values are the same as those chosen by Calderón & Nieto (2017). Furthermore, ν_j and λ_j , $j = 1, 2, 3, 4, 5$; a , b , H and d_0 are the same hyperparameters that were assumed for the TARX model used in the seasonality test. Now, the initial values for the parameters under the null hypothesis were $A_j^{(0)} = 0.05_{k_j+1}$, $B_j^{(0)} = 0.05_{K_j}$, $C_j^{(0)} = 0.05_{q_j}$, $S_j^{(0)} = \mathbf{1}_{k_j+1}$, $M_j^{(0)} = \mathbf{1}_{K_j}$, $N_j^{(0)} = \mathbf{1}_{q_j}$ ($\mathbf{1}$ is the vector of ones), $[(h^{(j)})^2]^{(0)} = 0.20$, $j = 1, 2, 3, 4, 5$; also, $d^{(0)} = 3$, $r^{(0)} = q_{0.50}\{Z_{t-d}\}$ for two regimes, $\mathbf{r}^{(0)} = (r_1, r_2)' = (q_{0.25}\{Z_{t-d}\}, q_{0.75}\{Z_{t-d}\})'$ for three regimes, $\mathbf{r}^{(0)} = (r_1, r_2, r_3)' = (q_{0.25}\{Z_{t-d}\}, q_{0.50}\{Z_{t-d}\}, q_{0.75}\{Z_{t-d}\})'$ for four regimes and $\mathbf{r}^{(0)} = (r_1, r_2, r_3, r_4)' = (q_{0.20}\{Z_{t-d}\}, q_{0.40}\{Z_{t-d}\}, q_{0.60}\{Z_{t-d}\}, q_{0.80}\{Z_{t-d}\})'$ for five regimes, $q_\alpha\{Z_{t-d}\}$ being the quantile α of Z_{t-d} . With the step sizes $\sigma_{r1} = \sigma_{r2} = \sigma_{r3} = \sigma_{r4} = 0.025$ of the Metropolis algorithm we draw from the distribution of \mathbf{r} .

The results of the estimation of the number of regimes are shown in Table 1. Here we see the averages of the posterior probabilities given in (5), which are computed using each one of the 100 time series, and the average of the standard DICs given in (6), as an alternative to doing this task. We find that the true model is selected.

Considering $l = 2$ and maximum autoregressive orders $k_1 = k_2 = 2$, $K_1 = K_2 = 2$ and $q_1 = q_2 = 3$, the true model can be expressed as:

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

TABLE 2: The best three models identified using the GVS method.

Best	$\left(\begin{array}{ccc cc} 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{array} \middle \begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 1 \end{array} \right)$
	(0.1801)*
Second best	$\left(\begin{array}{ccc cc} 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{array} \middle \begin{array}{ccc} 1 & 0 & 1 \\ 1 & 1 & 1 \end{array} \right)$
	(0.1379)
Third best	$\left(\begin{array}{ccc cc} 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{array} \middle \begin{array}{ccc} 1 & 1 & 0 \\ 1 & 1 & 1 \end{array} \right)$
	(0.1156)

*posterior probability.

where the ones indicate that the coefficients associated with these latent variables are present in the model. The three best selected models, are reported in Table 2. Using the GVS method, the best model that is selected matches the true model with a posterior probability of 0.1801 and the second and third best models differ from the true model only in the presence or not of some coefficients of the exogenous input in the first regime. The three selected models are associated with $d = 2$. Also, the proportion of times that the model is declared as either the best or the second best are 88% and 95%, respectively.

TABLE 3: Estimation of parameters for model 1 from 100 replications.

Parameter	True	Estimate	S.D.	C.I. 95%	Coverage
$a_0^{(1)}$	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)*			

*Number of times (out of a total of 100) where delay 2 is chosen.

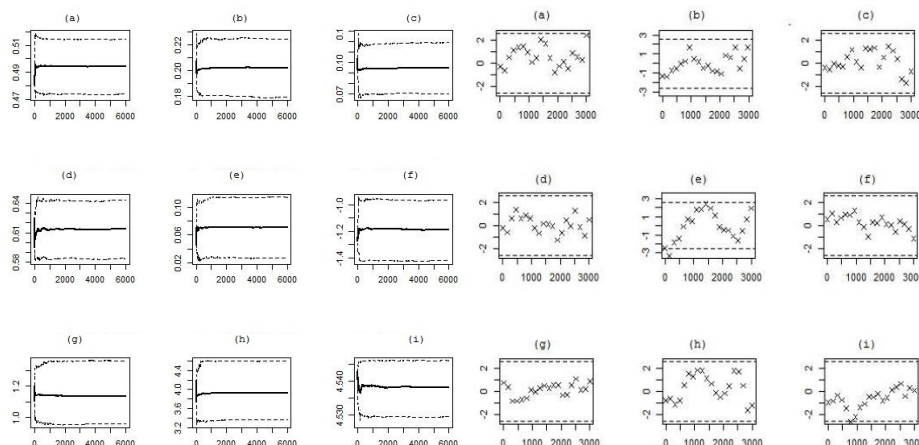


FIGURE 4: Cumulplot function (left) and Geweke's *Z-score* plot (right) for: (a) $a_1^{(1)}$, (b) $b_1^{(1)}$, (c) $b_2^{(1)}$, (d) $c_1^{(1)}$, (e) $a_0^{(2)}$, (f) $b_1^{(2)}$, (g) $(h^{(1)})^2$, (h) $(h^{(2)})^2$, (i) r , simulated model.

We proceed now, to estimate the other parameters of the TSARX model. The estimation of the autoregressive parameters in each regime, the variance weights by regimes, the threshold and delay values of the 100 replications are summarized in Table 3, showing each of the true values, plus the average, the standard deviation, the 0.025 and 0.975 quantiles of the 100 posterior means for each of the parameters. In addition, the percentage of coverage is given, which is the number of times that the true value of the parameter fall within the credible interval. For d , the average of the 100 posterior mode is given. All the averages of the posterior means are close to the true values and the 95% confidence intervals of the posterior means contain the true values. The percentages of coverage are reasonable and very close to 95%. Also the delay $d = 2$ is correctly identified with the sampling scheme considered.

In order to ensure that for each parameter, the MCMC samples converge to the posterior distribution, cumulplots function and Geweke's Z-score plots are used in one of the one hundred simulated time series. Because there are several parameters in the model, only some of them are shown (Figure 4), where we observe that convergence is reached in each case.

3.4. Model Diagnostics

For each one of the 100 simulated time series we estimate u_t , $t = 101, \dots, 600$ (see equation 7). This is done by setting the value of k sequentially at 150, 200, 250, $\dots, 600$, and for each t such that $k - t < 50$; then, \hat{v}_t is calculated, which is the generalized residual at time t . The Ljung-Box tests for lags 1, 12 and 24 and the Jarque-Bera test are used to check whether \hat{v}_t violates the hypothesis that it is independent and identically distributed $N(0,1)$. We reports the number of non-rejections out of 100 replications (not shown in table). The model (9) is correctly

fitted, it is good to see that the empirical sizes of the Ljung-Box tests are 96%, 96% and 92% and the Jarque-Bera test is 94%, respectively, match the nominal values of 95% very closely. Using one of the 100 simulated time series, CUSUM and CUSUMSQ plots of the generalized residuals are presented in Figure 5, indicating that there is no evidence of an incorrect specification of the model nor the presence of marginal heteroscedasticity in the generalized residuals, respectively.

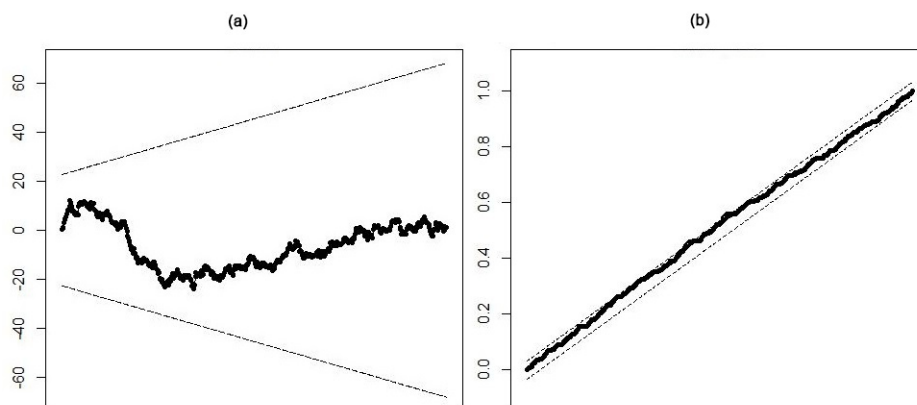


FIGURE 5: (a) CUSUM plot, (b) CUSUMSQ plot, of generalized residuals in the simulated model.

3.5. Forecasts Computation

Now, the forecast procedure developed in Subsection 2.7 is illustrated. In order to check the forecasting performance of the model, we considered as sample period 1-587 and the forecast horizon was 588-600. In Table 4, the simulated values, point forecasts, standard deviations and 95% credible interval averages for the variables X and Z are presented. It is worth noticing that these estimates are obtained via the averages of the corresponding statistics in each one of the 100 simulated time series. We put $h = 1, 2, 3, 11, 12, 13$. Additionally, in the last column of the table, the coverage percentage is presented.

We find that the forecasts are very close to the simulated values, all the simulated values fall within the corresponding 95% credible intervals, the standard deviations increase as the forecast horizon increases, and the coverage percentage is close to the 95% confidence level. The posterior distributions of the predictive distributions present a high variability. In this way, We assume that it is due to the use of priori noninformative distributions. Similar results to those found in Table 4, are given in Nieto (2008) and Vargas (2012), for TAR models. The behavior of the uncertainty measure of the long-term forecasts for TSARX models, maybe investigated, since it escapes the purpose of this paper.

TABLE 4: Forecasts for variables X and Z with 100 replications.

Horizon (h)	X_{T+h}	\widehat{X}_{T+h}	S.D.	C.I. 95%	Coverage
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
Horizon (h)	Z_{T+h}	\widehat{Z}_{T+h}	S.D.	C.I. 95%	Coverage
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

4. An Empirical Application

In this section, we apply the proposed methodology to Colombian monthly economic time series. The variables are the total monthly unemployment rate, TD_t , and a monthly coincident index that traces the Colombian economy, ISE_t . In our analysis, we focus on the growth of the total monthly unemployment rate, $X_t = TD_t - TD_{t-1}$, as the interest variable, and the monthly growth of the logarithmic transformation of ISE_t , $Z_t = (\ln(ISE_t) - \ln(ISE_{t-1})) \times 100$ (in percent) as the threshold variable. The sample period is from February, 2000 to December, 2016 (203 observations). The data is provided by the DANE, the Colombian official agency for the study and diffusion of official statistics.

4.1. Exploratory Analysis of the Data

The time series $\{X_t\}$ and $\{Z_t\}$ are given in Figure 6(a) and (b), respectively. $\{X_t\}$ presents annual positive peaks (Januaries), while $\{Z_t\}$ presents very pronounced annual negative peaks (Januaries). This indicates some seasonal behavior in both series, inherited from the original time series. It can also be observed that the series have no trends, but present seasonal cycles.

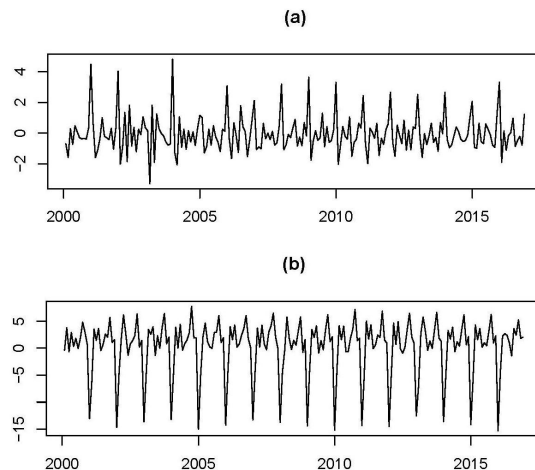


FIGURE 6: (a) X_t , growth of the Colombian total monthly unemployment rate, (b) Z_t , monthly growth of the logarithmic transformation of ISE_t .

In Figure 7, autocorrelograms and the AMI for X_t and Z_t are shown. In both cases, the samples ACF show seasonal lags that do not decrease and the PACF and AMI identify significant correlations in the first lags and lags 12, 24 and 36. Figure 8 suggests fitting a multiplicative AR(26) model to the two time series, following the minimum BIC. Looking at the box plots, X_t presents a distribution with higher values in the months of January, and a distribution with lower values in the months of March. On the other hand, for Z_t there is a distribution with lower values in the months of January, and a distribution with higher values in the months of October. In this way, there is some empirical evidence about the presence of seasonality in the time series.

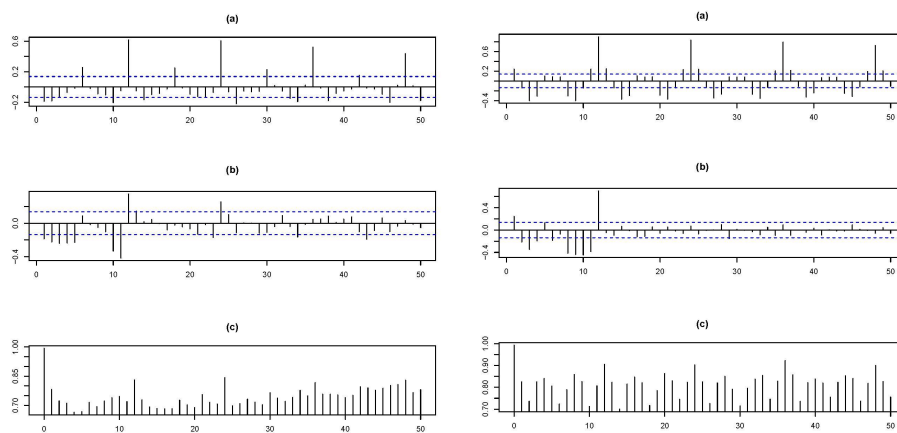


FIGURE 7: For X_t (left) and Z_t (right), (a) sample ACF, (b) sample PACF, (c) sample AMI, for the empirical data.

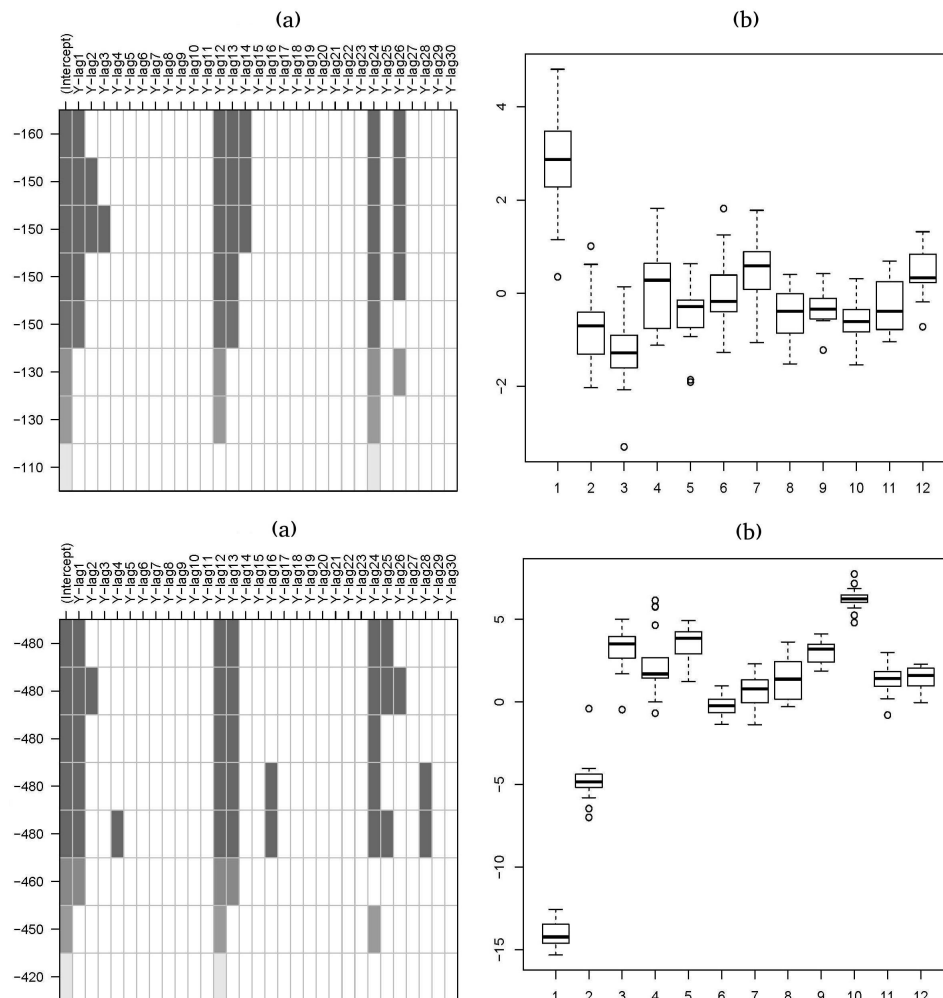


FIGURE 8: For X_t (top) and Z_t (bottom), (a) multiplicative AR with minimum BIC, (b) box plots by months, empirical data.

An adequate linear AR model fitted to the threshold process is

$$Z_t = -0.53Z_{t-1} - 0.21Z_{t-2} + 0.60Z_{t-12} + 0.32Z_{t-13} + 0.13Z_{t-14} + 0.38Z_{t-24} + 0.20Z_{t-25} + 0.08Z_{t-26} + \sqrt{0.57}a_t. \quad (10)$$

4.2. Estimation of the TSARX Model Parameters

To obtain the maximum autoregressive orders, the same candidate ARX models of Subsection 3.3 are considered. The values $k_1 = 3$, $K_1 = 2$ and $q_1 = 2$ were achieved. These values are fixed to apply the Tsay threshold nonlinearity

test (Tsay 1998), the proposed Bayesian seasonality statistical test, the Congdon method and the GVS technique.

Then, the threshold nonlinearity test is applied for delay $d = 0, 1, \dots, 10$, autoregressive orders $k_1 + sK_1 = 27$ and $q_1 = 0, 2$ and 15. The null hypothesis of a linear $ARX(k_1 + sK_1; q_1)$ model is rejected for different lags at the 1% significance level; in particular, for $d = 0, k_1 + sK_1 = 27, q_1 = 0$ and for $d = 3, k_1 + sK_1 = 27, q_1 = 0$ and 2. These results are shown in Table 5.

TABLE 5: Threshold nonlinearity test for the empirical data.

$k_1 + sK_1 = 27$ and $q_1 = 0$											
d	0	1	2	3	4	5	6	7	8	9	10
F statistics	2.54	1.37	0.96	1.58	1.01	1.51	1.67	1.06	1.18	1.19	1.52
p value	0.00	0.13	0.53	0.05	0.46	0.07	0.03	0.40	0.26	0.25	0.07
$k_1 + sK_1 = 27$ and $q_1 = 2$											
d	0	1	2	3	4	5	6	7	8	9	10
F statistics	2.47	1.39	0.87	1.86	0.86	1.52	1.32	0.93	1.45	1.51	1.85
p value	0.00	1.39	0.67	0.001	0.68	0.06	0.14	0.58	0.08	0.06	0.01
$k_1 + sK_1 = 27$ and $q_1 = 15$											
d	0	1	2	3	4	5	6	7	8	9	10
F statistics	1.32	1.12	1.13	1.08	0.64	1.13	0.99	1.12	0.84	0.72	1.53
p value	0.14	0.32	0.31	0.37	0.95	0.31	0.50	0.32	0.74	0.88	0.05

The above results allows us to consider a TARX model with a minimum of two regimes, which is specified in the seasonality test. For the Bayesian test of multiplicative seasonality by regimes, we considered $l = 2$ and $l = 3$ regimes in a TARX model. It is intrinsically required to estimate parameters of a $TARX(2; 3 + 12(2), 3 + 12(2); 2, 2)$ model and of a $TARX(3; 3 + 12(2), 3 + 12(2), 3 + 12(2); 2, 2, 2)$ model.

The MCMC sampler for posterior analysis is again based on 6.000 iterations after a burn-in period of 6.000 iterations. The hyperparameters and initial values for each of the parameters of the TARX models, are the same as those in Subsection 3.3. In Table 6, the results of the seasonality test are listed. A multiplicative TSARX model with two regimes is selected. For the case of $l = 3$ regimes, the second regime does not show the presence of multiplicative seasonality.

TABLE 6: Multiplicative seasonality test for the empirical data.

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

Now, we estimate the number of regimes and the autoregressive orders of a TSARX model. The hyperparameters and initial values of the parameters for each possible TSARX model are the same, as those described in Subsection 3.3. The

posterior probability of the number of regimes of a TSARX with $j = 1, 2, 3, 4, 5$ regimes, is given in Table 7, where it can be seen that the number of regimes associated with the highest posterior probability is two. This result is confirmed with the smallest DIC value.

TABLE 7: Posterior probabilities and DIC for choosing the number of regimes in the empirical data.

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

In Table 8, the three best models are observed based on the highest posterior probabilities, for the latent variable matrices S_j , M_j and N_j , $j = 1, 2$. The probabilities of the best three TSARX models with two regimes are 0.2122, 0.1536 and 0.1078. The three models are associated with $d = 3$ and they look very similar. Indeed, they have the same structure in the second regime and differ mainly in the entries of coefficients $a_3^{(1)}$ and $b_2^{(1)}$. The best model is chosen and the remaining parameters are estimated, which are shown in Table 9. Following the obtained results, the TSARX model is explicitly given by

$$X_t = \begin{cases} -0.22 - 0.47X_{t-1} + 0.53X_{t-12} + 0.25X_{t-13} + \sqrt{0.36}\epsilon_t, & \text{if } Z_{t-3} \leq 1.80, \\ -0.70X_{t-1} - 0.58X_{t-2} - 0.47X_{t-3} + 0.40X_{t-12} + \\ 0.28X_{t-13} + 0.23X_{t-14} + 0.19X_{t-15} + 0.47X_{t-24} + \\ 0.33X_{t-25} + 0.27X_{t-26} + 0.22X_{t-27} + \sqrt{0.46}\epsilon_t, & \text{if } Z_{t-3} > 1.80, \end{cases}$$

where the model for Z_t is given in model (10).

TABLE 8: The best three models identified using the GVS method, for the empirical data.

Best	$\left(\begin{array}{cccc cc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{array} \middle \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right)$ (0.2122)*
Second best	$\left(\begin{array}{cccc cc} 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{array} \middle \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right)$ (0.1536)
Third best	$\left(\begin{array}{cccc cc} 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{array} \middle \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right)$ (0.1078)

*posterior probability.

An interpretation of the fitted TSARX model is the following: first of all, only two regimes for the growth of the Colombian total monthly unemployment rate are detected, which could be termed as regimes of the monthly growth of the logarithmic transformation of the ISE_t , low and high. In the low regime, the most

pronounced negative peaks are also included. In addition, it is interesting to note that for the first regime, the intercept is negative and affects the epochs of recession of the economy by 22%, this reflects the economic impact that the growth of the unemployment rate has in that period; on the contrary, in the periods of expansion of the economy that are modeled by the second regime, the intercept is zero, giving evidence of a rapid economic recovery. Another interesting characteristics are the following: (1) the multiplicative AR model of the second regime has autoregressive orders higher than those of the first regime, (2) the coefficients of the nonseasonal component have negative signs, and the coefficients of the seasonal component have positive signs of similar magnitude, in the two regimes of the model, (3) the variances of the regimes are similar, and (4) the monthly growth of the logarithmic transformation of the index, affects the growth of the unemployment rate, which took place three months earlier.

TABLE 9: Estimation of parameters for empirical data.

Parameter	Estimate	S.D.	C.I. 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%)*		

* Percentage frequency of estimated delay.

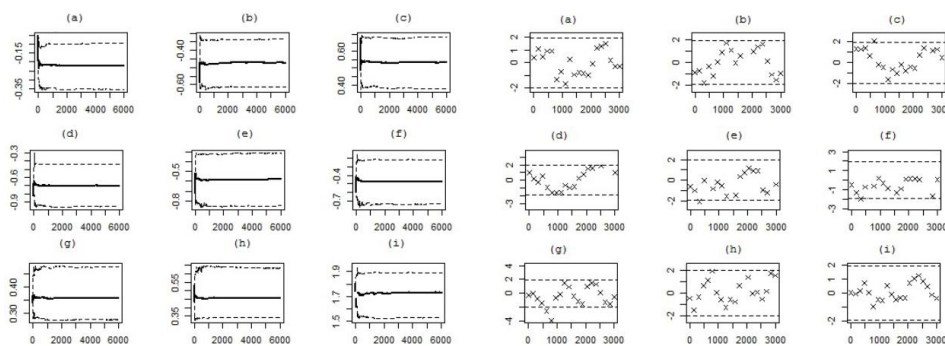


FIGURE 9: Cumuplot function (left) and Geweke's Z -score plot (right) for: (a) $a_0^{(1)}$, (b) $a_1^{(1)}$, (c) $b_1^{(1)}$, (d) $a_1^{(2)}$, (e) $b_1^{(2)}$, (f) $b_2^{(2)}$, (g) $(h^{(1)})^2$, (h) $(h^{(2)})^2$, (i) r , empirical data.

In Figure 9 the cumuplots function and Geweke's Z-score plots are shown, respectively, for some of the MCMC chains of the estimated parameters. The other chains can be provided by the authors upon request. The convergence, to the posterior distributions, is almost immediate and clearly reached, after the burn-in period in each case considered.

4.3. Model Diagnostics

In Table 10 we present the results of applying the Ljung-Box test for lags 1, 2, 3, 12 and 24 and the Jarque-Bera statistical test to the generalized residuals. In all cases, the null hypotheses of zero autocorrelations and normality are not rejected, under the 5% significance level. These results suggest that the TSARX model with two regimes fit reasonably well the empirical data. Figure 10 shows the CUSUM and CUSUMSQ plots of the generalized residuals, which complements the good fit of the model.

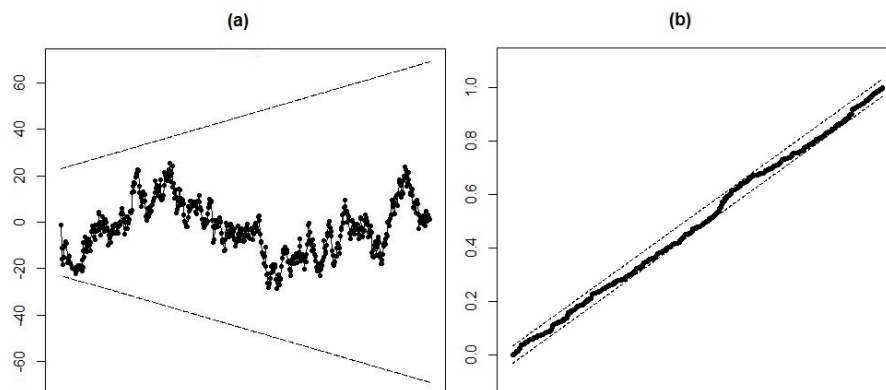


FIGURE 10: (a) CUSUM plot, (b) CUSUMSQ plot, of the generalized residuals in the real data.

TABLE 10: Ljung-Box test for different lags (LB) and Jarque-Bera test (JB) in the real data. p -values in parentheses.

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)

4.4. Forecasts of the Variables

As in the simulation example, we compute now forecasts for the variables X and Z . Our goal is double: first, to compute point predictions and second, to check the model performance for forecasting. We use 6.000 MCMC iterations in obtaining samples from the corresponding predictive distributions. We present the results in

Table 11, beginning in January, 2017 ($h = 1$) until January, 2018 ($h = 13$), period for which we include additional observations. The table includes the observed values, point forecasts, standard deviations of the predictive distributions and 95% confidence-level credible intervals. It can be noted that in both variables, (1) the point forecasts are very close to the observed values, (2) the standard deviations increase with larger values of h , and (3) all the observed values fall within the credible intervals.

TABLE 11: Forecasts for threshold and interest variables in the empirical application.

Horizon (h)	X_{T+h}	\hat{X}_{T+h}	S.D.	C.I. 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)
Horizon (h)	Z_{T+h}	\hat{Z}_{T+h}	S.D.	C.I. 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)

In future research, we will compare forecasts from the TSARX models with those from linear seasonal models and other nonlinear seasonal models, following the ideas given in Franses & Van Dijk (2005) and Vaca (2018).

5. Conclusions

In this paper, an open-loop threshold autoregressive model is specified for explaining the presence of seasonality in a nonlinear time series. The autoregressive model is, essentially, multiplicative but depending on the regimes. A Bayesian test is proposed for detecting this kind of seasonality. To fit the model to a time series, a full Bayesian methodology is developed that consists in the phases of (i) identifying the model, (ii) estimating the unknown parameters, (iii) checking the diagnostic via generalized residuals, and (iv) using the model to compute forecasts of the variables. The proposed methodology was illustrated with a simulated example and a real-data application. In future research, a more extensive characterization of the model may be conducted, in order to detect in the data more closely this kind of seasonality. Also, an extension to nonlinear multivariate time series may be investigated.

Acknowledgements

We gratefully acknowledge the very useful comments and suggestions of an anonymous referee, which helped to improve substantially the paper.

[Recibido: julio de 2019 — Aceptado: enero de 2020]

References

- Calderón, S. A. & Nieto, F. H. (2017), ‘Bayesian analysis of multivariate threshold autoregressive models with missing data’, *Communications in Statistics-Theory and Methods* **46**(1), 296–318.
- Chen, C. W. (1998), ‘A bayesian analysis of generalized threshold autoregressive models’, *Statistics & Probability Letters* **40**(1), 15–22.
- Chen, C. W., Gerlach, R. H. & Lin, A. M. (2010), ‘Falling and explosive, dormant, and rising markets via multiple-regime financial time series models’, *Applied Stochastic Models in Business and Industry* **26**(1), 28–49.
- Chen, C. W. & Lee, J. C. (1995), ‘Bayesian inference of threshold autoregressive models’, *Journal of Time Series Analysis* **16**(5), 483–492.
- Chen, C. W., Liu, F. C. & So, M. K. (2011), ‘A review of threshold time series models in finance’, *Statistics and its Interface* **4**(2), 167–181.
- Chen, C. W. & So, M. K. (2006), ‘On a threshold heteroscedastic model’, *International Journal of Forecasting* **22**(1), 73–89.
- Congdon, P. (2006), ‘Bayesian model choice based on monte carlo estimates of posterior model probabilities’, *Computational Statistics & Data Analysis* **50**(2), 346–357.
- Congdon, P. (2007), *Bayesian statistical modelling*, Vol. 704, John Wiley & Sons, Southern Gate, Chichester.
- Crespo, J. (2001), Modelling seasonal asymmetries using seasonal setar models. Working Paper.
- De Gooijer, J. G. & Vidiella-i Anguera, A. (2003), ‘Nonlinear stochastic inflation modelling using seasetars’, *Insurance: Mathematics and Economics* **32**(1), 3–18.
- Dellaportas, P., Forster, J. J. & Ntzoufras, I. (2002), ‘On bayesian model and variable selection using mcmc’, *Statistics and Computing* **12**(1), 27–36.
- Di Narzo, A. F. & Di Narzo, F. (2013), ‘Tserieschaos: Analysis of nonlinear time series. R package version 0.1-13’.

- Dickey, J. M. (1971), 'The weighted likelihood ratio, linear hypotheses on normal location parameters', *The Annals of Mathematical Statistics* **42**(1), 204–223.
- Franses, P. H., de Bruin, P. & van Dijk, D. (2000), Seasonal smooth transition autoregression, Technical report, Erasmus University, Amsterdam.
- Franses, P. H. & Van Dijk, D. (2000), *Non-linear time series models in empirical finance*, Cambridge University Press, Cambridge.
- Franses, P. H. & Van Dijk, D. (2005), 'The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production', *International Journal of Forecasting* **21**(1), 87–102.
- Gelfand, A. E. & Smith, A. F. (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association* **85**(410), 398–409.
- George, E. I. & McCulloch, R. E. (1993), 'Variable selection via gibbs sampling', *Journal of the American Statistical Association* **88**(423), 881–889.
- Gerlach, R., Carter, C. & Kohn, R. (1999), 'Diagnostics for time series analysis', *Journal of Time Series Analysis* **20**(3), 309–330.
- Gerlach, R. & Chen, C. W. (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.
- González, J. (2019), Modelamiento de procesos autorregresivos de umbrales estacionales, PhD thesis, Disertación doctoral en ciencias estadísticas, Universidad Nacional de Colombia, Bogotá.
- Hansen, B. E. (2011), 'Threshold autoregression in economics', *Statistics and its Interface* **4**(2), 123–127.
- Kass, R. E. & Raftery, A. E. (1995), 'Bayes factors', *Journal of the American Statistical Association* **90**(430), 773–795.
- Kuo, L. & Mallick, B. (1998), 'Variable selection for regression models', *Sankhyā: The Indian Journal of Statistics, Series B* **60**(1), 65–81.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. & Teller, A. H. (1953), 'Equation of state calculations by fast computing machines', *The Journal of Chemical Physics* **21**(6), 1087–1092.
- Meyn, S. P. & Tweedie, R. L. (2009), *Markov chains and stochastic stability*, 2 edn, Springer Science & Business Media, New York.

- Nieto, F. H. (2005), 'Modeling bivariate threshold autoregressive processes in the presence of missing data', *Communications in Statistics. Theory and Methods* **34**, 905–930.
- Nieto, F. H. (2008), 'Forecasting with univariate tar models', *Statistical Methodology* **5**(3), 263–276.
- Nieto, F. H. & Moreno, E. C. (2016), 'Univariate conditional distributions of an open-loop tar stochastic process', *Revista Colombiana de Estadística* **39**(2), 149–165.
- Nieto, F. H., Zhang, H. & Li, W. (2013), 'Using the reversible jump mcmc procedure for identifying and estimating univariate tar models', *Communications in Statistics-Simulation and Computation* **42**(4), 814–840.
- O'Hara, R. B. & Sillanpää, M. J. (2009), 'A review of bayesian variable selection methods: what, how and which', *Bayesian Analysis* **4**(1), 85–117.
- Plummer, M., Best, N., Cowles, K. & Vines, K. (2006), 'CODA: convergence diagnosis and output analysis for MCMC', *R news* **6**(1), 7–11.
- So, M. K. & Chen, C. W. (2003), 'Subset threshold autoregression', *Journal of Forecasting* **22**(1), 49–66.
- So, M. K., Chen, C. W. & 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 (Applied Statistics)* **55**(2), 201–224.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van Der Linde, A. (2002), 'Bayesian measures of model complexity and fit', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(4), 583–640.
- Team, R. D. C. (2016), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.
*<https://www.R-project.org/>
- Tong, H. (1990), *Non-linear 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.
- Tong, H. & Chen, C. H. (1978), *Pattern recognition and signal processing*, Springer Netherlands, chapter On a Threshold Model.
- Tong, H. L. & Lim, K. (1980), 'Threshold autoregression, limit cycles, and cyclical data', *Journal of the Royal Statistical Society, Series B* **42**(3), 245–292.
- Tsay, R. S. (1998), 'Testing and modeling multivariate threshold models', *Journal of the American Statistical Association* **93**(443), 1188–1202.

- Vaca, P. A. (2018), Analysis of the forecasting performance of the threshold autoregressive model, Master's thesis, Universidad Nacional de Colombia, Bogotá.
- Vargas, L. (2012), Cálculo de la distribución predictiva en un modelo TAR, Master's thesis, Universidad Nacional de Colombia.
- Verdinelli, I. & Wasserman, L. (1995), 'Computing bayes factors using a generalization of the savage-dickey density ratio', *Journal of the American Statistical Association* **90**(430), 614–618.
- Zhang, H. & Nieto, F. H. (2015), 'Tar modeling with missing data when the white noise process follows a student's t-distribution', *Revista Colombiana de Estadística* **38**(1), 239–266.