INTEGER-VALUED SELF-EXCITING PERIODIC THRESHOLD AUTOREGRESSIVE PROCESSES

ISABEL PEREIRA MANUEL SCOTTO RAQUEL NICOLETTE*

Abstract. In this paper, the periodic self-exciting threshold integer-valued autoregressive model of order one with period T driven by a periodic sequence of independent Poisson-distributed random variables is introduced and analyzed in detail. Basic probabilistic and statistical properties of the model are discussed as well as parameter estimation and forecasting.

1. Introduction

Modeling the temporal dependence and evolution of integer-valued (and in particular low counts) time series is an area of research which is gaining importance in time series analysis. The problem of developing models for integer-valued time series is, indeed, very challenging because traditional approaches based on Gaussian autoregressive-moving average processes, are of little use to accurately describe time series defined over finite range of counts or exhibiting features such low counts, over dispersion, asymmetric marginal distributions, or excess of zeros. The need to analyze such data adequately led to a multiplicity of approaches and a diversification of models that explicitly account for such features.

Recently, models for dealing with integer-valued time series exhibiting the so-called piecewise phenomenon have been proposed in the literature. The fundamental reason for introducing such class of models is the need to model random cyclic behavior that exists in many time series. In the continuous-valued case,

2010 *Mathematics Subject Classification.* 62M10; 91B70; 60G10 .

Key words and phrases. Count processes; Binomial thinning; Threshold models.

This work was supported by Portuguese funds through the CIDMA - Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology (FCT Fundação para a Ciência e a Tecnologia), within project UID/MAT/04106/2013.

threshold models are typically characterized by having a linear (ARMA) structure in each regime; see, e.g., Turkman *et al.* (2014) for details. However, in the field of integer-valued time series modelling little research has been done so far to develop models to cope with time series of counts exhibiting piecewise-type patterns. For this purpose, Monteiro *et al.* (2012) introduced a class of self-exciting threshold integer-valued autoregressive (SETINAR, in short) models of order one and two regimes, defined by the recursive equation

$$X_t \;=\; \begin{cases} \alpha_1 \circ X_{t-1} + Z_t, \;\; X_{t-1} \leq R \\ \alpha_2 \circ X_{t-1} + Z_t, \;\; X_{t-1} > R \end{cases} .$$

Here, (Z_t) constitutes a sequence of integer-valued random variables (r.v's), and R represents the threshold level. The "$\alpha_i \circ$ " is the binomial thinning operator of Steutel and van Harn (1979). It is defined as $\alpha \circ X := \sum_{i=1}^{X} Y_i$, for X with range $\mathbb{N}_0 = \{0, 1, \ldots\}$, where the Y_i's are independent and identically distributed (i.i.d.) Bernoulli variables with probability $\alpha \in (0; 1)$. The authors discussed probabilistic and statistical properties related with this class of models. Note that the SETINAR models fall within the state-dependent thinning class.

In this article, we introduce the periodic self-exciting threshold integer-valued autoregressive model of order one with two regimes (hereafter referred to as $\texttt{PSETINAR}(2; 1, 1)_T$) which generalizes the SETINAR model by considering periodically varying threshold levels. For this class of models, we investigate some basic probabilistic and statistical properties. Furthermore, parameter estimation and forecasting are also addressed. Finally some concluding remarks are given.

The $\texttt{PSETINAR}(2; 1, 1)_T$ model is defined through the recursive equation

$$X_t = \begin{cases} \alpha_j^{(1)} \circ X_{t-1} + Z_t^{(1)}, \;\; X_{t-d} \leq R_t \\ \alpha_j^{(2)} \circ X_{t-1} + Z_t^{(2)}, \;\; X_{t-d} > R_t \end{cases}, \quad t \in \mathbb{N}_0 \tag{1.1}$$

with $R_t = r_j$, for $t = j + sT$, $j = 1, \ldots, T$ and $s \in \mathbb{N}_0$. Note that for the jth-period we have

$$X_{j+sT} = (\alpha_j^{(1)} \circ X_{j+sT-1} + Z_{j+sT}^{(1)}) I_j^{(1)} + (\alpha_j^{(2)} \circ X_{j+sT-1} + Z_{j+sT}^{(2)}) I_j^{(2)} \tag{1.2}$$

with $I_j^{(k)}$, for $k = \{1, 2\}$, defined as

$$I_j^{(1)} := \begin{cases} 1, \;\; X_{j+sT-d} \leq r_j \\ 0, \;\; X_{j+sT-d} > r_j \end{cases}, \quad I_j^{(2)} = 1 - I_j^{(1)}. \tag{1.3}$$

The threshold parameter R_t (which is assumed to be known) represents the level of the process and the regime switch is triggered by the lag-d value of the series.

Note that the model in (1.1) can be represented as

$$X_t = \phi_t \circ X_{t-1} + Z_t, \tag{1.4}$$

where $Z_t = Z_j^{(1)} I_j^{(1)} + Z_j^{(2)} I_j^{(2)}$, $\phi_t = \alpha_j \equiv \alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)}$, such as $\alpha_j \in (0,1)$, $t = j + sT$, $j = 1, \dots, T$, and $s \in \mathbb{N}$. Furthermore, the thinning operator $\circ$ is defined as

$$\phi_t \circ X_{t-1} \overset{d}{=} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_j^{(1)}) I_j^{(1)} + \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_j^{(2)}) I_j^{(2)}$$

with $(U_{i,t}(\alpha_j^{(1)}))$ and $(U_{i,t}(\alpha_j^{(2)}))$, $i \in \mathbb{N}$, being periodic sequences of i.i.d. Bernoulli r.v's with success probabilities $P(U_{i,t}(\alpha_j^{(k)}) = 1) = \alpha_j^{(k)}$, for $k \in \{1, 2\}$. Moreover, the innovation process (Z_t) forms a periodic sequence of independent Poisson-distributed r.v's with mean v_t, $Z_t \sim Po(v_t)$, where $v_t = \lambda_j$, $t = j + sT$, $j = 1, \dots, T$, $s \in \mathbb{N}_0$. It is assumed that Z_t is independent of X_{t-1} and $\alpha_t \circ X_{t-1}$, for every t.

The $\texttt{PSETINAR}(2;1,1)_T$ process in (1.1) can be embedded in the following vectorial form

$$\mathbf{Y}_t = \mathbf{A} \circ \mathbf{Y}_{t-1} + \mathbf{M}_t, \tag{1.5}$$

being $\mathbf{Y}_t = [X_{1+tT}\, X_{2+tT}\, \cdots\, X_{T+tT}]'$, where $'$ denotes matrix transpose

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & \dots & \alpha_1 \\ 0 & 0 & \dots & \alpha_1 \alpha_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \prod_{j=0}^{T-1} \alpha_{T-j} \end{pmatrix}$$

and

$$\mathbf{M}_t := \mathbf{B} \circ \mathbf{Z}_t = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ \alpha_2 & 1 & \dots & 0 & 0 \\ \alpha_3\alpha_2 & \alpha_3 & \dots & 0 & 0 \\ \vdots & \ddots & \dots & \dots & \vdots \\ \prod_{i=0}^{T-1-2} \alpha_{T-1-i} & \prod_{i=0}^{T-1-3} \alpha_{T-1-i} & \dots & 1 & 0 \\ \prod_{i=0}^{T-1-1} \alpha_{T-i} & \prod_{i=0}^{T-1-2} \alpha_{T-i} & \dots & \alpha_T & 1 \end{pmatrix} \circ \begin{pmatrix} Z_{1+tT} \\ Z_{2+tT} \\ \vdots \\ Z_{T+tT} \end{pmatrix}.$$

Note that, $\mathbf{A} \circ \mathbf{Y}$ is a T-dimensional random vector with i-th component

$$[\mathbf{A} \circ \mathbf{Y}]_i = \sum_{j=1}^{T-1} 0 \circ X_{j+tT} + \left(\prod_{j=1}^{i} \alpha_j \right) \circ X_{T+tT},$$

for $i = 1, \dots, T$. The components of $\mathbf{B} \circ \mathbf{Z}$ can be defined similarly.

The rest of the paper is organized as follows: In Section 2, we demonstrate the existence of a strictly ciclostationary $\texttt{PSETINAR}(2;1,1)_T$ process satisfying (1.5). Furthermore, the expression for the periodic mean is given. Parameter estimation is covered in Section 3. Finally, forecasting is discussed in Section 4.

2. Some properties of the PSETINAR model with two regimes

Let (X_t) be the $\texttt{PSETINAR}(2;1,1)_T$ process defined in (1.2). We first prove that there exists a strictly ciclostationary $\texttt{PSETINAR}(2;1,1)_T$ process satisfying (1.2). Note that, since $\alpha_j^{(k)} \in (0,1)$ for $j = 1, \dots, T$ and $k = 1, 2$, and that $P(Z_{j+sT}^{(k)} = 0) \in (0,1)$, for $k = 1, 2$, it follows by Lemma 3 in Franke and Subba Rao (1995) that any solution $(\mathbf{Y}_t)$ of (1.5) is an irreducible and aperiodic Markov chain on $\mathbb{N}_0$. Thus, the existence of a ciclostationary solution of (1.5) relies upon the largest eigenvalue of the $\mathbf{A}$ matrix in (1.5). The result is quoted below.

Proposition 2.1. *Let $(\mathbf{Y}_t)$ be the $\texttt{PSETINAR}(2;1,1)_T$ process defined in (1.5). If $E||\mathbf{M}_t|| < +\infty$ and if the largest eigenvalue, say η, of $\mathbf{A}$ is less than one, then there exists a strictly ciclostationary $\texttt{PSETINAR}(2;1,1)_T$ process satisfying (1.5).*

Proof. Proposition B in Dion *et al.* (1995, p. 126) allows to concluded that

$$\alpha_j^{(1)} I_j^{(1)} + \alpha_j^{(2)} I_j^{(2)} < 1, j = 1, \dots, T \Leftrightarrow \eta < 1. \tag{2.6}$$

Conditions in (2.6) imply that all roots of the characteristic polynomial of $\mathbf{A}$ lie inside the unit circle. Furthermore, if $E||\mathbf{Z}_t|| < \infty$ it follows by Theorem 1 in Franke and Subba Rao (1995) that there exists a strictly ciclostationary $\texttt{PSETINAR}(2;1,1)_T$ process satisfying (1.5). $\square$

Without employing any distributional assumption on the periodic sequences $Z_t^{(1)}$ and $Z_t^{(2)}$, the periodic mean of the process is given in the next result. For simplicity in notation we define

$$u_{1,2,\ldots,j}^{(k_1,k_2,\ldots,k_j)} := E\left[X_{sT}|X_{1+sT-d} \in r_1^{(k_1)}, X_{2+sT-d} \in r_2^{(k_2)}, \ldots, X_{j+sT-d} \in r_j^{(k_j)}\right],$$

$$p_{1,2,\ldots,j}^{(k_1,k_2,\ldots,k_j)} := P\left[X_{1+sT-d} \in r_1^{(k_1)}, X_{2+sT-d} \in r_2^{(k_2)}, \ldots, X_{j+sT-d} \in r_j^{(k_j)}\right],$$

for $k_1, k_2, \ldots, k_j = \{1, 2\}$, where r_j denotes the regime corresponding to the period j, i.e.,

$$r_j = \begin{cases} \textit{regime } r_j^{(1)}, & X_{j+sT-d} \leq r_j \\ \textit{regime } r_j^{(2)}, & X_{j+sT-d} > r_j \end{cases}, \quad t \in \mathbb{N}_0. \tag{2.7}$$

Lemma 2.2. *Let (X_t) be the $\texttt{PSETINAR}(2;1,1)_T$ process in (1.5). The mean of the process takes the form*

$$\begin{aligned} E[X_t] &= \sum_{k_1=1}^{2}\sum_{k_2=1}^{2}\cdots\sum_{k_j=1}^{2}\left[\alpha_1^{(k_1)}\alpha_2^{(k_2)}\cdots\alpha_j^{(k_j)} \times u_{1,2,\ldots,j}^{(k_1,k_2,\ldots,k_j)} \times p_{1,2,\ldots,j}^{(k_1,k_2,\ldots,k_j)}\right] \\ &+ \sum_{l=1}^{j}\lambda_l\left(\sum_{k_{l+1}=1}^{2}\cdots\sum_{k_j=1}^{2}\alpha_{l+1}^{(k_{l+1})}\cdots\alpha_j^{(k_j)} \times p_{l+1,\ldots,j}^{k_{l+1},\ldots,k_j}\right), \end{aligned}$$

for $t = j + sT, j = 1, \ldots, T$ and $s \in \mathbb{N}_0$.

3. Parameters estimation

In this section, we consider the parameter estimation of the $\texttt{PSETINAR}(2;1,1)_T$ process. In particular, the conditional least squares (CLS) and conditional maximum likelihood methods are adopted. For this purpose, let $(X_1, \ldots, X_n)$ be a sequence of r.v's satisfying (1.4) and denote by

$$\boldsymbol{\theta} := (\alpha_1^{(1)}, \alpha_1^{(2)}, \lambda_1, \ldots, \alpha_T^{(1)}, \alpha_T^{(2)}, \lambda_T),$$

the vector of unkown parameters. Recall that R_t is assumed to be known.

3.1. Conditional least squares estimators.

The CLS-estimators

$$\hat{\boldsymbol{\theta}}_{CLS} := (\hat{\alpha}_{1,CLS}^{(1)}, \hat{\alpha}_{1,CLS}^{(2)}, \hat{\lambda}_{1,CLS}, \ldots, \hat{\alpha}_{T,CLS}^{(1)}, \hat{\alpha}_{T,CLS}^{(2)}, \hat{\lambda}_{T,CLS}),$$

are obtained by minimizing the expression

$$Q(\boldsymbol{\theta}) := \sum_{s=0}^{N-1}\sum_{j=1}^{T}\Big(X_{j+sT} - g_j(\boldsymbol{\theta}_j, X_{j+sT-1})\Big)^2$$

with N and T denoting the number of complete cycles and number of periods, respectively. Moreover, $\boldsymbol{\theta}_j := \left(\alpha_j^{(1)}, \alpha_j^{(2)}, \lambda_j\right)$ and the function g_j takes the form

$$g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) = \alpha_j^{(1)} X_{j+sT-1} I_j^{(1)} + \alpha_j^{(2)} X_{j+sT-1} I_j^{(2)} + \lambda_j.$$

Solving the systems of the form

$$\left\{\begin{array}{l} \frac{\partial Q}{\partial \alpha_j^{(1)}} = 0 \\ \frac{\partial Q}{\partial \alpha_j^{(2)}} = 0, \ j = 1, \ldots, T, \\ \frac{\partial Q}{\partial \lambda_j} = 0 \end{array}\right. \tag{3.1}$$

we obtain the following set of CLS-estimators

$$\left\{\begin{array}{l} \hat{\alpha}_{j,MQC}^{(1)} = \dfrac{N\sum\limits_{s=0}^{N-1} X_{j+sT}X_{j+sT-1}I_j^{(1)} - \sum\limits_{s=0}^{N-1} X_{j+sT}\sum\limits_{s=0}^{N-1} X_{j+sT-1}I_j^{(1)}}{N\sum\limits_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(1)} - \left(\sum\limits_{s=0}^{N-1} X_{j+sT-1}I_j^{(1)}\right)^2} \\ \\ \hat{\alpha}_{j,MQC}^{(2)} = \dfrac{N\sum\limits_{s=0}^{N-1} X_{j+sT}X_{j+sT-1}I_j^{(2)} - \sum\limits_{s=0}^{N-1} X_{j+sT}\sum\limits_{s=0}^{N-1} X_{j+sT-1}I_j^{(2)}}{(N\sum\limits_{s=0}^{N-1} X_{j+sT-1}^2 I_j^{(2)} - \left(\sum\limits_{s=0}^{N-1} X_{j+sT-1}I_j^{(2)}\right)^2}, \\ \\ \hat{\lambda}_{j,MQC} = N^{-1}\left(\sum\limits_{s=0}^{N-1} X_{j+sT} - \alpha_j^{(1)}\sum\limits_{s=0}^{N-1} X_{j+sT-1}I_j^{(1)} - \alpha_j^{(2)}\sum\limits_{s=0}^{N-1} X_{j+sT-1}I_j^{(2)}\right) \end{array}\right.$$

for $j = 1, \ldots, T$. The consistent and asymptotic distribution of the CLS-estimators is established in the result given below.

Theorem 3.1. *The CLS-estimators are strongly consistent and asymptotically normal, i.e.,*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}^{-1}\mathbf{W}\mathbf{V}^{-1}), \tag{3.2}$$

where $\mathbf{V}$ *and* $\mathbf{W}$ *are square matrices of order* $3T$ *defined by blocks of* 3×3 *given by*

$$\mathbf{V} = \begin{bmatrix} \mathbf{\Psi}_{:1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Psi}_{:2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{\Psi}_{:T} \end{bmatrix} \quad \text{and} \quad \mathbf{W} = \begin{bmatrix} \mathbf{\Omega}_{:1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega}_{:2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{\Omega}_{:T} \end{bmatrix},$$

where

$$\Psi_{(k,l):j} := E\left[\frac{\partial}{\partial \theta_{k:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) \frac{\partial}{\partial \theta_{l:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1})\right],$$

$$\Omega_{(k,l):j} := E\left[U_{j+sT}^2 \frac{\partial}{\partial \theta_{a:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1}) \frac{\partial}{\partial \theta_{l:j}} g_j(\boldsymbol{\theta}_j, X_{j+sT-1})\right],$$

are the elements of the matrices $\mathbf{\Psi}_{:j}$ *and* $\mathbf{\Omega}_{:j}, j = 1, \dots, T$*, respectively, with* $k, l \in \{1, 2, 3\}$*;* $j = 1, \dots, T$ *and* $\boldsymbol{\theta}_j := (\theta_{1:j}, \theta_{2:j}, \theta_{3:j}) \equiv \left(\alpha_j^{(1)}, \alpha_j^{(2)}, \lambda_j\right)$*, are the parameters associated to* j*-th period.*

Proof. Consistency and asymptotic normality can be proved using the results of Klimko and Nelson (1978, sec.3). It is easily checked that all regularity conditions by Klimko and Nelson (1978, p. 634) are satisfied, and thus, by their Theorem 3.1 it follows that the CLS-estimators are strongly consistent. Furthermore, in proving asymptotic normality we have to check first conditions (A)-(C) in Monteiro *et al.* (2012, p. 2725). To this extent, note first that conditions (A) and (B) follow easily by the arguments given in Monteiro *et al.* (2012). Finally, note that each block $\mathbf{\Psi}_{:j}$ of matrix V in (3.3) is defined as

$$\mathbf{\Psi}_{:j} = \begin{bmatrix} q_j^{(1)} m_{j,2}^{(1)} & 0 & q_j^{(1)} u_j^{(1)} \\ 0 & q_j^{(2)} m_{j,2}^{(2)} & q_j^{(2)} u_j^{(2)} \\ q_j^{(1)} u_j^{(1)} & q_j^{(2)} u_j^{(2)} & 1 \end{bmatrix},$$

where

$$u_j^{(1)} := E\left[X_{j-1+sT} | X_{j-1+sT-d} \le r_j\right] \ u_j^{(2)} := E\left[X_{j-1+sT} | X_{j-1+sT-d} > r_j\right]$$
$$m_{j,i}^{(1)} := E\left[X_{j-1+sT}^i | X_{j-1+sT-d} \le r_j\right] \ m_{j,i}^{(2)} := E\left[X_{j-1+sT}^i | X_{j-1+sT-d} > r_j\right]$$
$$q_j^{(1)} := P\left[X_{j+sT-d} \le r_j\right] \ q_j^{(1)} := P\left[X_{j+sT-d} > r_j\right].$$

Noticing that $det|\mathbf{\Psi}_{:j}| \neq 0$, $\forall j = 1, \dots, T$ lead us to conclude that V is invertible and condition (C) is thus fulfilled. Thus, Theorem 3.2 of Klimko and Nelson (1978) is thereby satisfied implying that the CLS-estimators are asymptotically normal. This concludes the proof. □

3.2. **Conditional maximum likelihood estimators.** For a fixed value of x_0, the conditional likelihood function for the $\texttt{PSETINAR}(2;1,1)_T$ takes the form

$$\begin{aligned} L(\boldsymbol{\theta}) &:= \prod_{s=0}^{N-1}\prod_{j=1}^{T} P_j\left(X_{j+sT}=x_{j+sT}|X_{j-1+sT}=x_{j-1+sT}\right) \\ &= \prod_{s=0}^{N-1}\prod_{j=1}^{T} p_j\left(x_{j-1+sT},x_{j+sT},\alpha_j^{(1)}I_j^{(1)}+\alpha_j^{(2)}I_j^{(2)},\lambda_j\right) \end{aligned}$$

with

$$p_j\left(x_{j-1+sT},x_{j+sT},\alpha_j^{(1)}I_j^{(1)}+\alpha_j^{(2)}I_j^{(2)},\lambda_j\right)=$$

$$\begin{aligned} &= e^{-\lambda_j}\sum_{m=0}^{M^*}\sum_{k=1}^{2} C_m^{x_{j-1+sT}}\alpha_j^{(k)^m}\left(1-\alpha_j^{(k)}\right)^{(x_{j-1+sT}-m)}\frac{\lambda^{x_{j+sT}-m}}{(x_{j+sT}-m)!}I_j^{(k)} \\ &\equiv p_j\left(x_{j-1+sT},x_{j+sT},\alpha_j^{(1)}I_j^{(1)},\lambda_j\right)+p_j\left(x_{j-1+sT},x_{j+sT},\alpha_j^{(2)}I_j^{(2)},\lambda_j\right) \\ &= p_j\left(x_{j-1+sT},x_{j+sT}\right)^{(1)}+p_j\left(x_{j-1+sT},x_{j+sT}\right)^{(2)} \end{aligned}$$

and $M^* := min(x_{j-1+sT},x_{j+sT})$.

The CML-estimators

$$\hat{\boldsymbol{\theta}}_{CML} := (\hat{\alpha}_{1,CML}^{(1)},\hat{\alpha}_{1,CML}^{(2)},\hat{\lambda}_{1,CML},\ldots,\hat{\alpha}_{T,CML}^{(1)},\hat{\alpha}_{T,CML}^{(2)},\hat{\lambda}_{T,CML}),$$

are obtained by maximizing the conditional log-likelihood function

$$l(\boldsymbol{\theta}) = \sum_{s=0}^{N-1}\sum_{j=1}^{T}\ln\ p_j\left(x_{j-1+sT},x_{j+sT},\alpha_j^{(1)}I_j^{(1)}+\alpha_j^{(2)}I_j^{(2)},\lambda_j\right) .$$

From the partial derivatives of first-order we obtain the set of systems

$$\begin{cases} \dfrac{I_j^{(k)}}{\alpha_j^{(k)}(1-\alpha_j^{(k)})}\displaystyle\sum_{s=0}^{N-1}(x_{j+sT}-x_{j-1+sT}\alpha_j^{(k)})- \\ -\lambda_j\dfrac{p_j\left(x_{j-1+sT},x_{j+sT}-1,\alpha_j^{(1)}I_j^{(1)}+\alpha_j^{(2)}I_j^{(2)},\lambda_j\right)}{p_j\left(x_{j-1+sT},x_{j+sT},\alpha_j^{(1)}I_j^{(1)}+\alpha_j^{(2)}I_j^{(2)},\lambda_j\right)}=0 \quad , \\ \displaystyle\sum_{s=0}^{N-1}\frac{p_j\left(x_{j-1+sT},x_{j+sT}-1,\alpha_j^{(1)}I_j^{(1)}+\alpha_j^{(2)}I_j^{(2)},\lambda_j\right)}{p_j\left(x_{j-1+sT},x_{j+sT},\alpha_j^{(1)}I_j^{(1)}+\alpha_j^{(2)}I_j^{(2)},\lambda_j\right)}=N \end{cases}$$

for $k = 1, 2$ and $j = 1, \ldots, T$. In order to solve those systems numerical procedures have to be employed. Note, however, that the CML-estimates for the λ's, are readily available from those for the α's through the following expression

$$\hat{\lambda}_{j,CML} = \frac{1}{N}\sum_{s=0}^{N-1}\left(x_{j+sT} - \hat{\alpha}^{(k)}_{j,CML}x_{j-1+sT}\right), \; j = 1, \ldots, T.$$

The following result establishes the consistency and the asymptotic distribution of the CML-estimators.

Theorem 3.2. *Let (X_t) be the* PSETINAR$(2;1,1)_T$ *model in (1.1). The CML-estimators are asymptotically normal, i.e,*

$$\sqrt{n}\left|\begin{array}{c} \hat{\alpha}_1^{(1)} - \alpha_1^{(1)} \\ \hat{\alpha}_1^{(2)} - \alpha_1^{(2)} \\ \hat{\lambda}_1 - \lambda_1 \\ \vdots \\ \hat{\alpha}_T^{(1)} - \alpha_T^{(1)} \\ \hat{\alpha}_T^{(2)} - \alpha_T^{(2)} \\ \hat{\lambda}_T - \lambda_T \end{array}\right| \xrightarrow{d} N(\mathbf{0}, \mathbf{I}^{-1}), \; where \quad \mathbf{I} = \left[\begin{array}{cccc} \mathbf{M}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{M}_T \end{array}\right],$$

is the Fisher information matrix with

$$\mathbf{M}_j = \left[\begin{array}{ccc} -E\left[\dfrac{\partial^2 l(\theta)}{\partial(\alpha_j^{(1)})^2}\right] & -E\left[\dfrac{\partial^2 l(\theta)}{\partial\alpha_j^{(1)}\partial\alpha_j^{(2)}}\right] & -E\left[\dfrac{\partial^2 l(\theta)}{\partial\alpha_j^{(1)}\partial\lambda_j}\right] \\ -E\left[\dfrac{\partial^2 l(\theta)}{\partial\alpha_j^{(1)}\partial\alpha_j^{(2)}}\right] & -E\left[\dfrac{\partial^2 l(\theta)}{\partial(\alpha_j^{(2)})^2}\right] & -E\left[\dfrac{\partial^2 l(\theta)}{\partial\alpha_j^{(2)}\partial\lambda_j}\right] \\ -E\left[\dfrac{\partial^2 l(\theta)}{\partial\alpha_j^{(1)}\partial\lambda_j}\right] & -E\left[\dfrac{\partial^2 l(\theta)}{\partial\alpha_j^{(2)}\partial\lambda_j}\right] & -E\left[\dfrac{\partial^2 l(\theta)}{\partial\lambda_j^2}\right] \end{array}\right],$$

for $j = 1, \ldots, T$, *and*

$$E\left[\frac{\partial^2 l(\theta)}{\partial(\alpha_j^{(k)})^2}\right] = N\sum_a^{+\infty}\sum_b^{+\infty} P\left(X_{j+sT} = b\right)\frac{I_j^{(k)}}{\alpha_j^{(k)}(1-\alpha_j^{(k)})}\Big\{\left[(2\alpha_j^{(k)}-1)b - a\alpha_j^{(k)^2}\right]\times$$
$$\times p_j(b|a) + 2(1-\alpha_j^{(k)})\lambda_j p_j(b-1|a)^{(k)} + \lambda_j^2 p_j(b-2|a)^{(k)} +$$
$$+2(1-\alpha_j^{(k)})\lambda_j p_j(b-1|a)^{(k)} + \lambda_j^2 p_j(b-2|a)^{(k)} - \lambda_j^2\frac{p_j^2(b-1|a)^{(k)}}{p_j(b|a)^{(k)}}\Big\};$$

$$E\left[\frac{\partial^2 l(\theta)}{\partial\alpha_j^{(k)}\partial\lambda_j}\right] = \frac{N}{\alpha_j^{(k)}(1-\alpha_j^{(k)})}\sum_a^{+\infty}\sum_b^{+\infty} P\left(X_{j+sT} = b\right)\Big\{-p_j(b-1|a)^{(k)} -$$
$$-\lambda_j p_j(b-2|a)^{(k)} + \lambda_j\frac{p^2(b-1|a)^{(k)}}{p(b|a)^{(k)}}\Big\};$$

$$E\left[\frac{\partial^2 l(\theta)}{\partial\lambda_j^2}\right] = N\sum_a^{+\infty}\sum_b^{+\infty} P\left(X_{j+sT} = b\right)\left\{\lambda_j p_j(b-2|a)^{(k)} - \frac{p_j^2(b-1|a)^{(k)}}{p_j(b|a)^{(k)}}\right\},$$

for $k = 1, 2$.

PROOF. In order to derive the large sample distribution of the CML-estimators, we use the same arguments as in Franke and Seligmann (1993, pp. 324–325). Note that the consistency and asymptotic distribution of the CML-estimators for the INAR(1) process can be obtained by means of Theorems 2.1 and 2.2 in Billingsley (1961, pp. 10–13). Since the innovation process is Poisson-distributed the arguments used by Monteiro *et al.* (2010, 2012) for the periodic INAR of order one and period T, PINAR$(1)_T$, and the SETINAR$(2;1,1)$ process can be easily generalized for the SETINAR$(2;1,1)$ with periodic structure. We omit the details. □

4. Point Prediction in the PSETINAR model

In this section we consider the forecasting of future values X_{i+NT+h}, with $h = j + lT$, given past observations up through time $i + NT$, i.e,

$$(X_1, \ldots, X_T, \ldots, X_{i+NT}, \ldots, +X_{i+NT}).$$

First note that by iterating equation (1.1) it follows that X_t can be expressed as

$$X_t \stackrel{d}{=} \left(\prod_{j=0}^{n-1}\phi_{t-j}\right)\circ X_{t-n} + \sum_{i=1}^{n-1}\left(\prod_{j=0}^{i-1}\phi_{t-j}\right)\circ Z_{t-i} + Z_t$$
$$\equiv \beta_{t,n}\circ X_{t-n} + \sum_{i=0}^{n-1}\beta_{t,i}\circ Z_{t-i},$$

where, for $t > i$

$$\beta_{t,i} := \begin{cases} \prod_{j=0}^{i-1} \phi_{t-j}, \ i > 0 \\ 1, \ i = 0 \end{cases} = \begin{cases} \beta_{t,j}\beta_{T,T}^{l}, \ i = j + lT, j = 1, \ldots, T \\ 1, \ i = 0 \end{cases},$$

leading to obtain

$$X_{i+NT+h} \stackrel{d}{=} \beta_{i+NT+h,h} \circ X_{i+NT} + \sum_{m=0}^{h-1} \beta_{i+NT+h,m} \circ Z_{i+NT+h-m}. \tag{4.1}$$

Since $h = j + lT$, it follows that

$$X_{i+NT+h} \stackrel{d}{=} \beta_{i+j+(N+l)T,j+lT} \circ X_{i+NT} + \sum_{m=0}^{j+lT-1} \beta_{i+j+(N+l)T,m} \circ Z_{i+j+(N+l)T-m}.$$

Due to the periodicity of β's, $\beta_{i+j+(N+l)T,j+lT} = \beta_{i+j,j+lT} = \beta_{i+j,j}\beta_{T,T}^{l}$, and considering the relation

$$\sum_{m=0}^{lT-1} \beta_{j+lT,m} \circ Z_{j+lT-m} \stackrel{d}{=} \sum_{w=0}^{l-1} \sum_{m=0}^{T-1} \beta_{j+lT,m}\beta_{T,T}^{l} \circ Z_{j+(l-w)T-m},$$

the expression in (4.1) takes the form

$$X_{i+NT+h} \stackrel{d}{=} \left(\beta_{i+j,j}\beta_{T,T}^{l}\right) \circ X_{i+NT} + V_{i+j+lT} \tag{4.2}$$

with

$$V_{i+j+lT} := \sum_{m=0}^{i-1} \beta_{i+j,m} \circ Z_{i+j-m+NT} + \sum_{w=0}^{l-1} \sum_{m=0}^{T-1} \beta_{i+j+(N+l)T,m+j+Tw} \circ Z_{i+(N+l-w)T-m}.$$

In order to generate the h-step ahead prediction the mean, median or mode of the predictive distribution of $X_{i+NT+h}|X_{i+NT}$ can be employed as a point forecast. Note that the median and mode are considered as coherent (i.e., integer-valued) predictions, whereas the mean is not. In order to evaluate the prediction performance given by the mean, median or mode of the predictive distribution we can use the square root of the mean squared error (RMSE), the mean absolute error (MAE) or the loss function everything or nothing (LFEN), respectively. Note that the h-step-ahead point predictor that minimizes the mean square error (MSE) is given by

$$
\begin{aligned}
\hat{X}_{i+NT+h} &= E\left[X_{i+NT+h}|X_{i+NT}\right] \\
&= E\left[\left(\beta_{i+j,j}\beta^{l}_{T,T}\right)\circ X_{i+NT}|X_{i+NT}\right] \\
&+ \sum_{m=0}^{j-1}\beta_{i+j,m}\left(\lambda^{(1)}_{i+j-m}p^{(1)}_{i+j-m}+\lambda^{(2)}_{i+j-m}p^{(2)}_{i+j-m}\right) \\
&+ \sum_{w=0}^{l-1}\sum_{m=0}^{T-1}\beta_{i+j,m+j+Tw}\left(\lambda^{(1)}_{i-m}p^{(1)}_{i-m}+\lambda^{(2)}_{i-m}p^{(2)}_{i-m}\right)
\end{aligned}
$$

with

$$
\begin{aligned}
&p^{(1)}_{i+j-m} := P\left(X_{i+j-m+NT-1}\le r_{i+j-m}\right); \\
&p^{(2)}_{i+j-m} = 1-p^{(1)}_{i+j-m}, i+j>m; \\
&p^{(1)}_{i-m} := P\left(X_{i+(N+l-w)T-m-1}\le r_{i-m}\right); \\
&p^{(2)}_{i-m} = 1-p^{(1)}_{i-m}, i>m.
\end{aligned}
$$

Turning now to the particular case $h=1$, the one-step-ahead predictive function is given by

$$
\begin{aligned}
&P(X_{j+NT+1}=y|X_{j+NT}=x) \\
&\quad=\sum_{m=0}^{min(x,y)}\sum_{k=1}^{2}C^{x}_{m}\left(\alpha^{(k)}_{j+1}\right)^{m}\left(1-\alpha^{(k)}_{j+1}\right)^{x-m}e^{-\lambda^{(k)}_{j+1}}\frac{\left(\lambda^{(k)}_{j+1}\right)^{(y-m)}}{(y-m)!}I^{(k)}_{j+1}
\end{aligned}
$$

with $\lambda^{(k)}_{j+1}=E[Z^{(k)}_{j+NT+1}]$, $k\in\{1,2\}$. Finally, from (4.2), the most commonly used one-step-ahead predictor of X_{j+NT+1}, takes the form

$$
\begin{aligned}
\hat{X}_{j+NT+1} &= \left(\alpha^{(1)}_{j+1}X_{j+NT}+\lambda^{(1)}_{j+1}\right)P(X_{i+NT}\le r_{i+1}) \\
&+ \left(\alpha^{(2)}_{j+1}X_{j+NT}+\lambda^{(2)}_{j+1}\right)P(X_{i+NT}> r_{i+1}).
\end{aligned}
$$

5. Concluding Remarks

This paper has presented the periodic self-exciting threshold integer-valued autoregressive model of order one with period T, driven by a periodic sequence of independent Poisson-distributed random variables. Basic probabilistic and statistical properties of the model are established as well as parameter estimation and forecasting.

We would like to stress here that an important issue when fitting `PINAR` models lies with parsimony. Even every simple `PINAR` model can have an inordinately large number of parameters. This is also true when dealing with `PSETINAR` models. Therefore, the development of procedures for dimensionality reduction is an impeding problem. This remains a topic of future research.

References

[1] Billingsley, P. (1961). *Statistical Inference for Markov Processes.* University of Chicago Press, Chicago.

[2] Dion, J-P., Gauthier, G. & Latour, A. (1995). Branching processes with immigration and integer-valued time series. *Serdica Math. J.* **21**, 123–136.

[3] Franke, J. & Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modeling epileptic seizure counts. *Time Series Analysis* (T. Subba Rao, ed.). London, Chapman and Hall, pp. 310-330.

[4] Franke, J. & Subba Rao, T. (1995). Multivariate first order integer valued autoregressions. *Technical Report*, Universitat Kaiserslautern.

[5] Klimko, L.A. & Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* **6**, 629-642.

[6] Monteiro, M., Scotto, M.G. & Pereira, I. (2010). Integer-valued autoregressive processes with periodic structure. *J. Statist. Plann. Inference* **140**, 1529–1541.

[7] Monteiro, M, Scotto, M.G. & Pereira, I. (2012). Integer-valued self-exciting threshold autoregressive processes. *Comm. Statist. Theory Methods* **41**, 2717–2737.

[8] Steutel, F.W. & van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Ann. Probab.* **7**, 893–899.

[9] Turkman, K.F., Scotto, M.G. & de Zea Bermudez, P. (2014). *Non-Linear Time Series: Extreme Events and Integer Value Problems.* Springer-Verlag, Switzerland.

Departamento de Matemática and CIDMA, Universidade de Aveiro, Portugal

*CIDMA

E-mail address: isabel.pereira@ua.pt mscotto@ua.pt raquel.nicolette@gmail.com