Brazilian Journal of Biometrics, **41**, 110–121 (2023)
doi:10.28951/bjb.v41i2.601

BRAZILIAN JOURNAL OF BIOMETRICS
ISSN:2764-5290

ARTICLE

A Tarso model for studying the level of the Cuiabá river

Lucas Monteiro Bianchi,*[,1] Anderson Castro Soares de Oliveira,[1] Luciene Resende Gonçalves,[2] and Lia Hanna Martins Morita[1]

[1]Department of Statistics, Federal University of Mato Grosso, Cuiabá-MT, Brazil.
[2]Institute of Applied Social Sciences, Federal University of Alfenas, Varginha-MG, Brazil.
*Corresponding author. Email: estatistico.bianchi@gmail.com

(Received: January 24, 2022; Revised: June 9, 2022; Accepted: August 26, 2022; Published: June, 2023)

Abstract

The Cuiabá River registers cyclical occurrences of floods and droughts over the years, according to the hydrological periods, thus characterizing a non-linear behavior. The prediction of the level of the Cuiabá river is important to help institutions such as the Civil Defense of the state of Mato Grosso and many other institutions that are concerned with the prevention and mitigation of natural disasters. Thus, this study considered the nonlinear Threshold Autoregressive Self-Excking Open-loop (TARSO) model with 2 regimes, with a Bayesian approach. We tested models to which values of the linimetric quota (river water level in millimeters) with and without rainfall (mm) were associated. All models were compared using the lowest DIC, MAPE and MSE criterion, and the TARSO $(2; 1, 0, 3, 1, 1)$ model performed best according to these criteria. Finally, the selected model was shown to produce reliable predictions.
Keywords: Times Series, Bayesian analysis, nonlinear models, Cuiabá River, Floodings

1. Introduction

Natural disasters have been studying more frequently in the international science community due to the conflicting relationship between society and nature worldwide, even as the consequences of catastrophic events for the population of affected regions. Brazil still lacks papers that purport to present preventive decision-making tools to reduce the outcomes caused by natural phenomena.

These catastrophes are frequently triggered by heavy rainfalls. When precipitation is intense but cannot soak into the soil fast, a large part of its water volume flows into the drainage system, surpassing its natural drainage capacity. The surplus water volume not drained by the soil fills the floodplain, submerging according to the topography of the areas close to the rivers (Tucci, 2004). Generally, these phenomena are enhanced by human changes in the environment, such for example, the waterproofing of the soil and the rectification of watercourses due to urban interventions (Goerl & Kobiyama, 2005).

Floodings in fluvial systems are natural phenomena that occur anywhere worldwide. These events come about when there is an increase in the average level in a water system and does not necessarily cause harm to the population. However, they often gain sufficient proportions to extrapolate the smaller bed of the watercourse, affecting large cities, mainly where urban development did not occur in a planned way, severely impacting the economy and society in general (Paes, 2011).

According to IBGE (2014), the Midwest region had the lowest concentration of municipalities affected by flooding (19%). However, Cuiaba, the capital of Mato Grosso, has records of floods that devastated neighborhoods in the metropolitan area. In 1974, for example, the so-called "great flood" took place, where the waters of the Cuiaba river basin reached 10.87 meters. The Manso plant, located 280 km from Cuiaba, was then planned in the late 70s. It has the aim of attenuating the floods that used to compromise this region. Furthermore, other floods fell out in 1942, 1959, 1995, 2001 (Zamparoni, 2012)

Concerning the high frequency of these phenomena and their impacts on the local populations, the Superintendence of Civil Defense of State of Mato Grosso (SUDEC/MT) specified planimetric reference levels for several rivers in locations with potential floods (Paes, 2011). SUDEC/MT considered the experiences of floodable areas upstream of the Pantanal and adopted the alert (8.50m), emergency (9.50m), and calamity (11.00m) levels for the Cuiaba River.

The Cuiaba river basin is located in the state of Mato Grosso, with a total area of approximately 29,000 km^2, covering 13 municipalities. Understanding the hydrological dynamics of this basin is of fundamental importance for plans that minimize the impacts caused by floods and overflows. Therefore, the modeling of hydrometeorological variables provides support for the knowledge of the future behavior of the basin, having fundamental importance in the planning of flood forecasting systems. The region stands out by the presence of perennial rivers and a good water production capacity. Furthermore, the hydrological regime presents a severe drought period between June and September, and a lot of rain with peak flows from December up to February (SEDEC, 2013).

The dynamics of the Cuiaba river basin does not assume a linear behavior, then it is necessary to develop non-linear models, which is challenging, as the relationship between the parameters and the observed values are only hypotheses, without any general law leading to this relationship (Chakraborty *et al.,* 1992).

The time series model is widely used to represent hydrological data. However, the models that belong to the well-known Box-Jenkins methodology may be inappropriate, as they assume linear relationships between the variables. Therefore, non-linear time series models can perform better in fitting hydrological data.

Among the various nonlinear models, the threshold models stand out (Tong & Lim, 1980). These models present as the main characteristic the presence of different states or regimes for the time series and allow the variables to have several dynamic behaviors depending on the regime pattern that occurs at each time slot. In the autoregressive models with a threshold (TAR), the threshold variable controls the regimes. The threshold autoregressive (TAR) and the self-excited threshold autoregressive (SETAR) models stand out in this class (de Almeida *et al.,* 2020).

Another process belonging to the threshold class is the threshold autoregressive self-exciting open-loop (TARSO) proposed by Tong (1990). This model emerged to time series modeling due to several factors, such as flow series, water level series, plant disease levels, weather variables, among other aspects.

Several works using threshold models for hydrological data were presented in the scientific community. Jian *et al.* (1998) used the TARSO model to describe the dynamics of groundwater flow in Shanxi province, in northwestern China, in which the threshold parameter enables different precipitation processes, and the lag parameter stamps the time interval between precipitation and spring flow increase. Vasas *et al.* (2007) used this model to describe the daily series of water discharges from the Tisza River in Hungary. de Almeida *et al.* (2020) modeled the average daily quota in the Cuiaba

River, Brazil using the SETAR model, with a threshold set by the median.

This work aims to analyze the behavior of the time series of the daily quota of the Cuiaba River and propose forecast models through the TARSO model.

2. Matherials and Methods

Study area

The Cuiaba watershed is located in the western part of central Brazil, in the state of Mato Grosso, with a total area of approximately 29,000 km^2 with 841 km in perimeter belonging to the Upper Paraguay River basin. The basin is located between the geographic coordinates 1418 and 1700S and 5440 and 5655W. The headwaters belong to the municipality of Rosário Oeste, on the riverside of Serra Azul, and their main sources are the Cuiaba da Larga and Cuiaba do Bonito rivers. After the confluence of these rivers, it changes its name to Cuiabazinho river and, only after the encounter with the Manso, it turns to the name of Cuiaba River (de Almeida *et al.,* 2020; MATO GROSSO, 2003).

The Cuiaba river basin consists of two large geological formations with well-defined biotic and abiotic characteristics: the Pantanal flatland and the surrounding highland and mountain ranges. These orographic characteristics enable us to distinguish three different regions in the Cuiaba river basin, namely: Upper Cuiaba, Middle Cuiaba, and Lower Cuiaba (de Almeida *et al.,* 2020)

Pluviometric anad Fluviometric data

The pluviometric and fluviometric daily data were obtained from January 1st, 2001, up to December 31st, 2012, resulting in a total number of 4384 observations for each data set. The fluviometric data consists of measurements from the limnimetric level (river water level in millimeters), and the pluviometric data consists of rainfall measurements (mm). The limnimetric quotas refer to the Rosário Oeste stations (code 66250001) obtained from the Hidroweb system of the National Water Agency (ANA). Furthermore, the rainfall data were collected through the TRMM satellite (The Tropical Rainfall Measuring Mission).

2.1 TARSO MODEL

In this work, it is considered the TARSO model with two regimes and threshold value r (Tsay, 1998). Consider the model

$$Y_t = \begin{cases} \phi_{10} + \sum_{i=1}^{p_1} \phi_{1i} Y_{t-i} + \sum_{j=1}^{q_1} \theta_{1i} X_t 1 - 1 + \varepsilon_t^{(1)} & \text{se} \quad Y_{t-d} \leq r \\ \phi_{20} + \sum_{i=1}^{p_2} \phi_{2i} Y_{t-i} + \sum_{j=1}^{q_2} \theta_{2i} X_t 1 - 1 + \varepsilon_t^{(2)} & \text{se} \quad Y_{t-d} > r \end{cases} \tag{1}$$

where

- Y_t is the linimetric level of the Cuiabá river at time t;
- X_t is the precipitation at time t
- r is the threshold parameter which breaks the series into two parts;
- $d > 0$ is the lag parameter;
- p_1, p_2, q_1, q_2 are orders of submodels in each regime;
- ϕ and θ are the autoregressive parameters;
- ε_t is the white noise, uncorrelated with zero mean and constant variance.

The equation 1 assigns that Y_t belongs to two different processes with different orders $(k, p_1, p_2, m_1, m_2, q_1, q$ where k is the number of regimes that depends on the threshold variable X_{td}. It is denoted by

$$TARSO(k = 2; p_1, p_2, q_1, q_2, d)$$

Assuming that the orders (p_i, q_i), $i = 1, 2$ are known, the model parameters are (γ_i, τ_i) where

$$\gamma_i = \{\phi_{i0}, \phi_{i1}, ..., \phi_{ip_i}, \theta_{i0}, \theta_{i1}, ..., \theta_{iq_i}\}, \quad i = 1, 2$$

Let $Z_{it} = (1, Y_{t-1}, Y_{t-2}, ..., Y_{t-p_i}, X_{t-1}, X_{t-2}, ...X_{t-p_i})$ with $i = 1, 2$ the model 1 can be matrix rewritten as

$$Y_t = \begin{cases} \gamma_1 Z_{1t} + \varepsilon_t^{(1)} & \text{se} \quad Y_{t-d} \leq r \\ \gamma_2 Z_{2t} + \varepsilon_t^{(2)} & \text{se} \quad Y_{t-d} > r \end{cases} \tag{2}$$

Let $D = Y_t, X_t, t = p + 1, ..., n$ the entire sequence, $p = \max(p_1, p_2, q_1, q_2)$ and d a fixed value, in which the model is conditioned and n_1, n_2 the number of observations in each regime. Conditioning in the p-first observations then $\varepsilon_t^{(1)} \sim N(0, \tau_i)$, the conditional likelihood function can be approximated by

$$L(\Phi|d, r, D) \approx \tau_1^{\frac{n_1}{2}} \tau_2^{\frac{n_2}{2}} exp \left\{ -\frac{\tau_1}{2} \sum_{t=p+1}^{n} (Y_t - \gamma_1 Z_{1t})^2 - \frac{\tau_2}{2} \sum_{t=p+1}^{n} (Y_t - \gamma_2 Z_{2t})^2 \right\}$$

where

- $\sum_{t=p+1}^{n} (Y_t - \gamma_1 Z_{1t})^2$ is the sum in t under the first regime, that is, $X_{t-d} \leq r$
- $\sum_{t=p+1}^{n} (Y_t - \gamma_2 Z_{2t})^2$ is the sum em t under the second regime, that is, $X_{t-d} > r$
- $\Phi = \{\gamma_1, \gamma_2, \tau_1, \tau_2\}$

2.2 Statistical Analysis

All analyzes were performed in the R Core Team (2020) software. Initially, descriptive analyzes of the linimetric quota series and rainfall were performed using the graphs of the time series, as well as the application of the Cox-Stuart and g Fisher tests (Morettin & Toloi, 2006) to verify, respectively, the existence of trend and seasonality.

2.2.1 Bayesian Analysis

The Bayesian analysis was performed considering the article by Sáfaddi & Morettin (2001) assuming a previous distribution by Jeffreys:

$$P(\Phi) \approx \tau_1^{-1} \tau_2^{-1}$$

The parameters r and d are known, their values were specified as $r = 119$ (median of the quota), and $d = 1$ in the same way as used by de Almeida *et al.* (2020).

Thus, the posterior distribution is gamma-normal and the complete conditional distributions of γ_i and τ_i are given by

$$\begin{aligned} \gamma_i|\tau_i, d, r, D &\sim N\left(\mathbf{A}_i^{-1}\mathbf{B}_i, \tau_i \mathbf{A}_i^{-1}\right) \\ \tau_i|\gamma_i, d, r, D &\sim Gama\left(\frac{n_i - (p_i + m_i + 1)}{1}, \frac{\mu}{2}\right) \end{aligned}$$

in which

$$
\begin{aligned}
\mu &= \mathbf{C}_i - \mathbf{B}_i'\mathbf{A}_i^{-1}\mathbf{B}_i + \left(\gamma_i - \mathbf{A}_i^{-1}\mathbf{B}_i\right)'\mathbf{A}_i\left(\gamma_i - \mathbf{A}_i^{-1}\mathbf{B}_i\right)\\
\mathbf{A}_i &= \begin{pmatrix} \sum_{t=p+1}^{n_i} Y_{t-1}Y_{t-1} & \sum_{t=p+1}^{n_i} Y_{t-1}Z_{t-1} \\ \sum_{t=p+1}^{n_i} Y_{t-1}Z_{t-1} & \sum_{t=p+1}^{n_i} Z_{t-1}Z_{t-1} \end{pmatrix} \qquad \mathbf{B}_i = \begin{pmatrix} \sum_{t=p+1}^{n_i} Y_{t-1}Y_t \\ \sum_{t=p+1}^{n_i} Z_{t-1}Y_t \end{pmatrix}\\
\mathbf{C}_i &= \left(\sum_{t=p+1}^{n_i} Y_tY_t\right)
\end{aligned}
$$

The posterior distribution estimates for γ_i and τ_i were obtained using the MCMC method and the Gibbs sampling algorithm Geman & Geman, 1984.

The models were fit to datasets considering the different combinations of

- $p_1 = 1, ..., 5$
- $p_2 = 1, ..., 5$
- $m_1 = 0, ..., 5$
- $m_2 = 0, ..., 5$
- with and without the parameters $\phi_{10}, \beta_{10}, \phi_{20}, \beta_{20}$

3. Results and Discussion

Figure 1 exhibits the time series of the daily average linimetric quotas at the Rosário Oeste Station, in which a seasonal behavior is observed. The Fisher g test shows the existence of seasonality (p-value <0.0001) with a period of oscillation of the series around 366 days, which was determined by the periodogram (Morettin & Toloi, 2006), with the maximum in the rainy periods and minimal in the dry period. The Cox-Stuart test had no trend (p-value=0.8639), indicating that there was no gradual increase or decrease in the series over time.

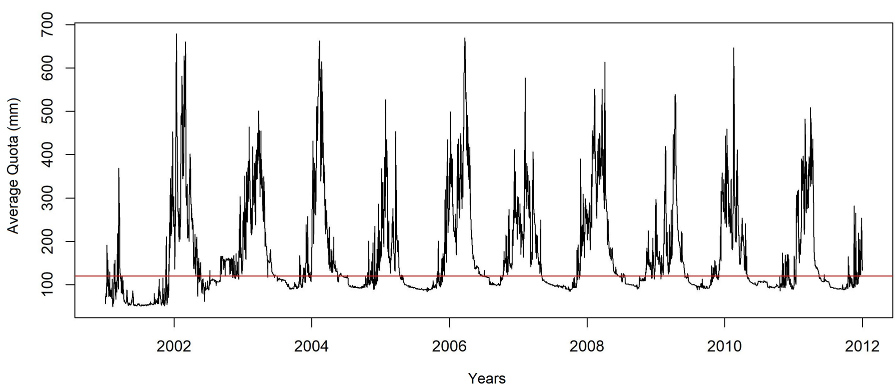

Figure 1. Series of daily average linimetric quotas, for Rosário Oeste station, from January 1st, 2001 up to December 31, 2012.

Figure 2 shows the precipitation series, which has similar behavior to the quota, with seasonality (p-value <0.0001), and an oscillation period of 367 days, and without a trend (p-value p=0.0985).

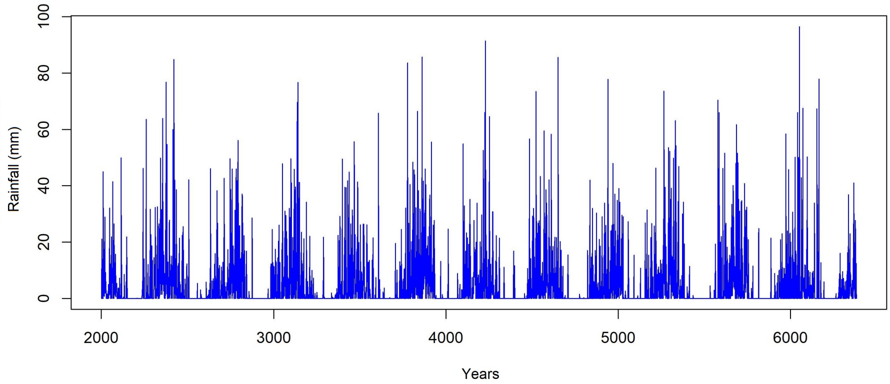

Figure 2. Daily rainfall series for Rosário Oeste station, from January 1st, 2001 up to December 31, 2012.

The cross-correlation function between the mean elevation and precipitation is shown in figure 3. In positive and negative lags they are significant up to around 30 days, indicating that precipitation positively influences the average quota and the opposite also occurs, but it is also observed that positive lags have a greater correlation, indicating a greater force of precipitation in influencing the quota.

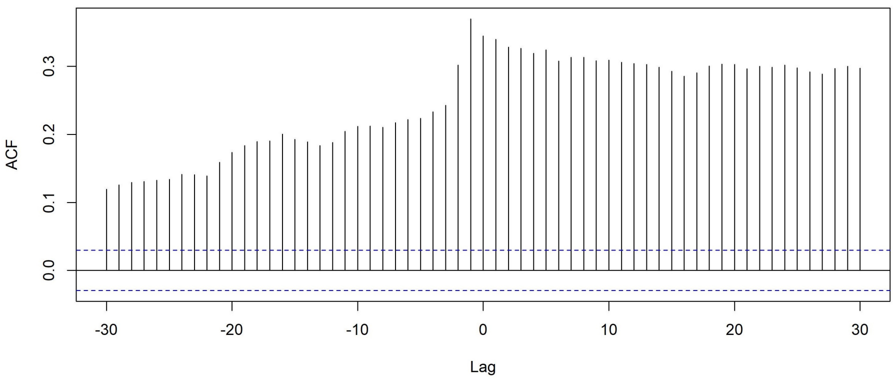

Figure 3. Cross-correlation function between the average quota and precipitation.

3.1 Models

Among the 14,400 fitted models, only 38 had all significant parameters, and their configuration of these models is presented in table 1. Of these 38 models, we have to:

- In the first regime for the quota autoregressive order 31 models with $p_1 = 1$ and 7 models with $p_1 = 2$, all models presented $m_1 = 0$, indicating that when the quota is lower than its median, its value depends on 1 or two previous ones and has no influence of precipitation.
- In the second regime for the autoregressive order of the quota, 22 models with $p_2 = 1$, 1 model with $p_2 = 2$ and 15 models $p_2 = 3$, for rainfall variable 19 models presented $m_2 = 0$ and $m_2 = 1$, indicating that when the quota is higher than its median, its value depends on the previous 1 to 3 and is influenced by the previous day's rainfall.

From Tabel 1 it is also possible to observe that the models 26 and 29 presented better performance since they have the lowest DIC value (9503,4).

Table 1. TARSO models fitted with (p_1, m_1, p_2, m_2) and with and without ϕ_{10}, β_{10}, ϕ_{20}, β_{20}

model	p_1	m_1	p_2	m_2	ϕ_{10}	β_{10}	ϕ_{20}	β_{20}	DIC
1	1	0	1	0	without	without	without	without	9520.8
2	1	0	1	0	with	without	without	without	9517.6
3	1	0	1	0	without	without	with	without	9520.6
4	1	0	1	0	with	without	with	without	9517.7
5	1	0	1	0	without	without	without	with	9532.1
6	1	0	1	0	with	without	without	with	9529.2
7	1	0	1	0	without	without	with	with	9531.4
8	1	0	1	0	with	without	with	with	9528.4
9	1	0	1	1	without	without	without	without	9525.2
10	1	0	1	1	with	without	without	without	9510.5
11	1	0	1	1	without	without	with	without	9525.0
12	1	0	1	1	with	without	with	without	9510.0
13	1	0	1	1	without	without	without	with	9531.7
14	1	0	1	1	with	without	without	with	9528.8
15	1	0	1	1	without	with	without	with	9530.1
16	1	0	1	1	without	without	with	with	9531.4
17	1	0	1	1	with	without	with	with	9528.3
18	1	0	2	1	with	without	with	with	9524.9
19	1	0	3	0	without	without	without	without	9508.9
20	1	0	3	0	with	without	without	without	9506.0
21	1	0	3	0	without	with	without	without	9507.0
22	1	0	3	0	without	without	with	without	9508.6
23	1	0	3	0	with	without	with	without	9506.0
24	1	0	3	0	with	without	with	with	9514.9
25	1	0	3	1	without	without	without	without	9506.8
26	1	0	3	1	with	without	without	without	9503.4
27	1	0	3	1	without	with	without	without	9509.5
28	1	0	3	1	without	without	with	without	9506.2
29	1	0	3	1	with	without	with	without	9503.4
30	1	0	3	1	without	without	with	with	9511.4
31	1	0	3	1	with	without	with	with	9508.6
32	2	0	1	0	without	without	without	without	152623.7
33	2	0	1	0	without	with	with	without	151641.7
34	2	0	1	0	without	with	without	with	151069.7
35	2	0	1	0	with	without	with	with	144412.2
36	2	0	1	1	without	without	without	without	152959.8
37	2	0	3	0	without	with	without	without	10022.3
38	2	0	3	1	without	without	with	with	10041.0

Figure 4 shows the final densities and chains of the parameters for model 26 – TARSO (2,1,0,3,1,1) with intercept ϕ_{10}.

The estimates for model 26 – TARSO(2,1,0,3,1,1) with intercept ϕ_{10} – are presented in table ??,

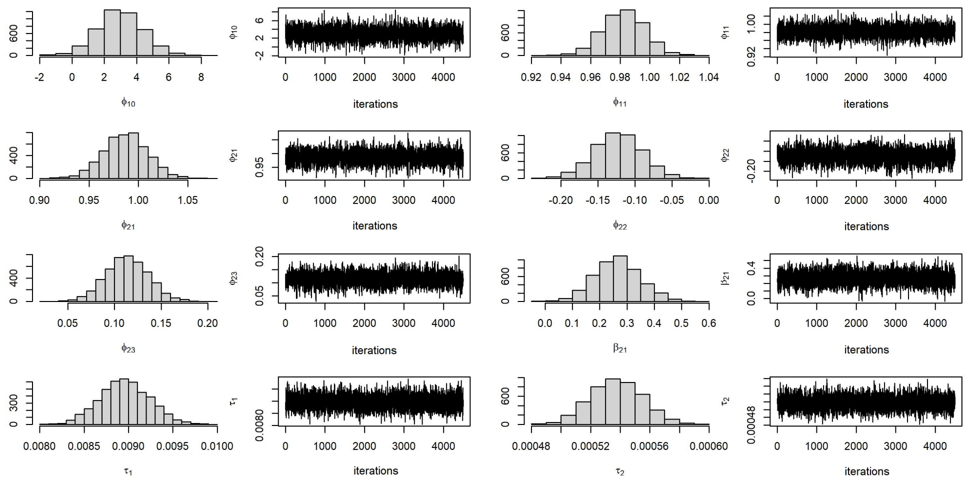

Figure 4. Densities and final strings of parameters for the TARSO model (2,1,0,3,1,1) with ϕ_{10} intercept, for daily average linimetric quotas, for Rosário Oeste station, from January 1st, 2001 up to December 31, 2012.

and from the 95% credible intervals, all parameters are found to be meaningful. From the estimates, it is possible to observe that:

- In the first regime, the quota value revolves around a constant 2.97 and is positively affected by the value of the previous day, it does not influence precipitation;
- In the second regime, the quota value is influenced by up to three past values, with the first and third day having a positive influence, and the second day negatively. The quota is also positively influenced by the past day's rainfall.

Table 2. TARSO (2,1,0,3,1,1) model estimates with intercept ϕ_{10} e ϕ_{20}

Parameters	a Posteriori	95% credible intervals	
	Mode	lower	upper
ϕ_{10}	2.97980	0.34485	5.67386
ϕ_{11}	0.98320	0.95477	1.00923
ϕ_{21}	0.98661	0.94151	1.02818
ϕ_{22}	-0.12656	-0.18791	-0.06094
ϕ_{23}	0.11584	0.06729	0.15763
β_{21}	0.25900	0.10282	0.42901
τ_1	0.00894	0.00837	0.00946
τ_2	0.00054	0.00050	0.00057

Figure 5 shows the densities and final chains of the parameters for model 29 - TARSO(2,1,0,3,1,1) with intercept ϕ_{10} and ϕ_{20}.

The table 3 shows the estimates for model 29 - TARSO(2,1,0,3,1,1) with intercept ϕ_{10} and ϕ_{20} - where all parameters are found to be significant. From the estimates, it is possible to observe that:

- In the first regime, the quota value revolves around a constant 3.00 and is positively affected by the value of the previous day and it does not influence precipitation;

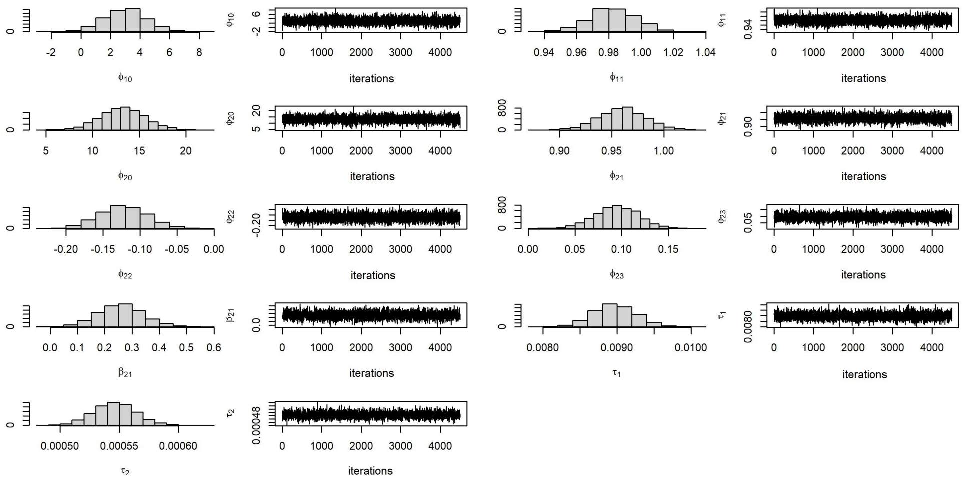

Figure 5. Densities and final strings of the parameters for the TARSO model(2,1,0,3,1,1) with ϕ_{10} and ϕ_{20} intercept, for daily average linimetric quotas, for Rosário Oeste station, from January 1st, 2001 to December 31, 2012.

- In the second regime, the quota value revolves around a constant 13.24 and it is influenced by its values of up to 3 previous ones, with the first and third day having a positive influence, and the second day negatively. The quota is also positively influenced by the previous day's rainfall.

Table 3. TARSO (2,1,0,3,1,1) model estimates with intercept ϕ_{10} e ϕ_{20}

Parameters	a Posteriori	95% credible intervals	
	Mode	lower	upper
ϕ_{10}	3.00497	0.44726	5.84880
ϕ_{11}	0.98194	0.95433	1.00954
ϕ_{20}	13.24514	8.68038	17.71779
ϕ_{21}	0.96031	0.91703	1.00350
ϕ_{22}	-0.12385	-0.18383	-0.06123
ϕ_{23}	0.09665	0.04868	0.13786
β_{21}	0.25160	0.09830	0.41759
τ_1	0.00895	0.00844	0.00955
τ_2	0.00055	0.00051	0.00058

Figure 6 shows the daily average linimetric quotas observed and estimated by the TARSO model(2,1,0,3,1,1) with ϕ_{10} intercept for the Rosário Oeste station. It is noted observed that the values presented by the model were similar to the observed values.

The daily average linimetric quotas observed and estimated by the TARSO model(2,1,0,3,1,1) with ϕ_{10} and ϕ_{20} intercept for Rosário Oeste station are presented in Figure 7, in which similarity is observed between the results obtained by the model and the observed values.

The table 4 presents the MAPE and MSE of the TARSO models (2,1,0,3,1,1), in which there is a good predictive capacity of the two models, but the model with intercept shows a better performance considering the first regime ϕ_{10}.

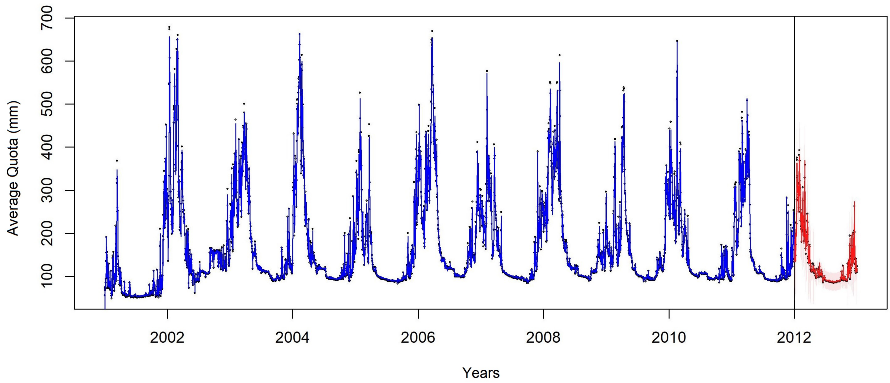

Figure 6. Fit of the TARSO model (2,1,0,3,1,1) with intercept ϕ_{10}, for daily linimetric quotas, for Rosário Oeste station, from January 1st, 2001 up to December 31, 2012.

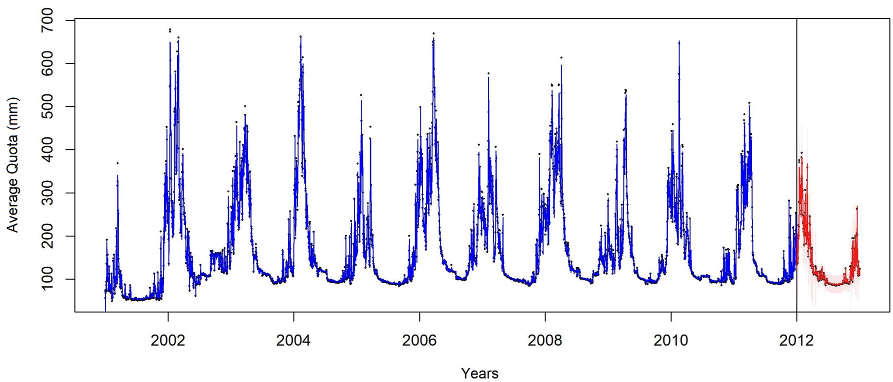

Figure 7. Fit of the TARSO model (2,1,0,3,1,1) with ϕ_{10} and ϕ_{20} intercept, for daily average linimetric quotas, for Rosário Oeste station, from January 1st, 2001 up to December 31, 2012.

Table 4. Forecast quality indicators for TARSO models obtained for daily average linimetric quotas, for Rosário Oeste station, from January 1st, 2001 up to December 31, 2012

Model	MAPE	MSE
TARSO(2,1,0,3,1,1) with intercept ϕ_{10}	7.23	543.25
TARSO(2,1,0,3,1,1) with intercept ϕ_{10} e ϕ_{20}	7.48	552.42

4. Conclusions

Forecasting the stage of the Cuiabá river basin is crucial to help the Civil Defense of Mato Grosso and many other authorities concerned with the anticipation and mitigation of natural disasters. Due to the history of flooding in the Cuiabá river basin and its complex dynamics, linear time series models do not present good fits. Consequently, they are inappropriate for fitting seasonal time series. Accordingly, this paper considered the *Threshold Autoregressive Self-exciting Open-loop* (TARSO) non-linear with two regimes to study the daily quota of the Cuiabá river.

The Bayesian framework was implemented through the MCMC methods and Gibbs sampling algorithm, which is available in R software, thus the parameter estimates of the TARSO model were obtained. Among the total of 14,400 models considered in this analysis, only 38 of them present significance in all the parameters. According to selection model criteria, the TARSO model (2,1,0,3,1,1) presented the best performance, having the intercept coefficient in the first regime. This model holds a good predictive capacity for estimating the Cuiabá river basin level, being suitable for a system of flood alarm.

de Almeida *et al.* (2020) used the SETAR model to forecast the daily quota of the Cuiabá river. The authors supposed that the river flow dynamics depend only on its past values. The TARSO model, however, may take into account external variables. In this work, incorporating the daily precipitation in the TARSO model improved its predictive capacity, mainly in estimating high quota values.

The TARSO model presented satisfactory results in capturing the non-linear dynamics of the daily quota in the Cuiabá river. This threshold model is suitable in capturing seasonality and non-stationarity characteristics of the Cuiabá river. Furthermore, it captures the relationship between river height and precipitation. It is worth mentioning that the rainfall-runoff process presents asymmetric effects on river flow and it is adequately captured by the TARSO model with two regimes.

In TARSO model, d parameter expresses the threshold, and r parameter is the one that splits the series into two parts. In this work, these quantities are fixed and known. Future work may consider these parameters as unknown, affording more flexibility to the inference methodology.

Acknowledgments

The authors acknowledge support from Fundação de Amparo à Pesquisa do Estado de Mato Grosso (FAPEMAT).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Chakraborty, K., Mehrotra, K., Mohan, C. K. & Ranka, S. Forecasting the behavior of multivariate time series using neural networks. *Neural networks* **5,** 961–970 (1992).
2. De Almeida, E. M., de Oliveira, A. C. S., Sanches, L. & de Pascoa, M. A. R. Comportamento e previsão do nível do rio Cuiabá por meio de modelos SETAR. *Sigmae* **2,** 1–13 (2020).
3. Geman, S. & Geman, D. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Recognition* **6,** 721–741 (1984).
4. Goerl, R. F. & Kobiyama, M. Considerações sobre as inundações no Brasil. *XVI Simpósio Brasileiro de Recursos Hídricos* (2005).
5. IBGE. *Perfil dos estados e dos municípios brasileiros, 2013* (Coordenação de População e Indicadores Sociais - Rio de Janeiro, Rio de Janeiro, 2014).
6. Jian, W.-B., Yao, H., Wen, X.-H. & Chen, B.-R. A Nonlinear Time Series Model for Spring Flow: An Example from Shanxi Province, China. *Groundwater* **36,** 147–150 (1998).
7. MATO GROSSO. Diagnostico da bacia hidrográfica do Rio Cuiabá. Relatório final de diagnostico ambiental realizado pela empresa Ecoplan (2003).
8. Morettin, P. & Toloi, C. M. C. *Análise de séries temporais* (Edgard Blucher, 2006).
9. Paes, R. P. d. *Análise da translação da onda de cheia efluente do reservatório da UHE Manso na bacia hidrográfica do rio Cuiabá, Mato Grosso* PhD thesis (Universidade de São Paulo, 2011).
10. R Core Team. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria, 2020). https://www.R-project.org/.

11. Sáfaddi, T. & Morettin, P. A. A. AnáLise Bayesiana do modelo Open Loop Threshold Autoregressive. *Revista Brasileira de Estatística* **62,** 91–105 (2001).
12. SEDEC. *Anuário brasileiro de desastres naturais: 2013* (Ministério da Integração Nacional. Secretaria Nacional de Proteção e Defesa Civil. Centro Nacional de Gerenciamento de Riscos e Desastres - CENAD, Brasíia, 2013).
13. Tong, H. & Lim, K. S. Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society: Series B (Methodological)* **42,** 245–268 (1980).
14. Tong, H. *Non-linear time series: a dynamical system approach* (Clarendon Press, Oxford, 1990).
15. Tsay, R. S. Testing and Modeling Multivariate Threshold Models. *Journal of the American Statistical Association* **93,** 1188–1202 (1998).
16. Tucci, C. E. Gerenciamento integrado das inundações urbanas no Brasil. *Rega/Global Water Partnership South America* **1,** 59–73 (2004).
17. Vasas, K., Elek, P. & Márkus, L. A two-state regime switching autoregressive model with an application to river flow analysis. *Journal of Statistical Planning and Inference* **137.** Special Issue: Bayesian Inference for Stochastic Processes, 3113–3126 (2007).
18. Zamparoni, C. G. Riscos e desastres naturais em ambiente urbano: o exemplo de Cuiabá/MT. *Revista Brasileira de Climatologia* **10** (2012).