MOAD'2024
The Sixth International Colloquium on
Methods and Tools for Decision Support
October 20th–22nd 2024, Tizi-Ouzou, Algeria

Bayesian Inference in Autoregressive Models with trend under additive outliers contaminations

Ait Mohammed Noura
Faculty of Sciences, Mouloud Mammeri University of Tizi-Ouzou. Algeria

Guerbyenne Hafida
Faculty of Mathematics, University of Science and Technology Houari Boumediene. Algeria

Abstract

In this paper we consider additive outliers (AO) which can occur in data. We have generalized this type of contamination to the multiple case for an autoregressive models of order p with a regression trend. We adopt the Bayesian approach combined with Gibbs sampling to jointly estimate the model parameters and the outliers on the first hand, and on the other hand we use a test based on p-values to detect the location and the magnitude of the of outliers. A simulation study is presented for illustrating the performance of the method relative to maximum likelihood estimation, mainly for small sample sizes.

Keywords and phrases Gibbs sampling; multiple additive outliers; autoregressive model; bayesian analysis; outliers detection.

Received before September 15, 2024 **Accepted** October 5, 2024

1 Introduction

Identification and estimation of outliers play an important role in statistical analysis. It is well known that statistical data, collected for analysis and interpretation in diverse fields, economics, industrial, finance, etc., most often contain one or more observations which do not look similar to the rest of the data. Such observations are called outliers. These outliers may have a significant impact on the results of standard methodology for time series. Fox [1] introduced additive outlier (AO) which affects a single observation and innovation outlier (IO) which affects a single innovation and their effects on an autoregressive model of order p ($AR(p)$).

Suppose that an AO occurs in time a series $\{y_t\}$ at time $t = k$ and let y_k be the affected observation. The contaminated observation will differ from the original observations according to the following rule (see, e.g., Tsay [4]):

$$y_t^* = \begin{cases} y_t & for \quad t \neq k \\ y_t + \Delta & for \quad t = k \end{cases}$$

That is, the shock caused by an AO affects the observation at time $t = k$ only, with magnitude of Δ, while the rest remains unaffected.

In this paper we generalize the contamination AO to the multiple case for an $AR(p)$ with a regression trend. We adopt the Bayesian approach to jointly estimate the model parameters and the outliers on the first hand, and on the other hand, we use an unconditional Bayesian test based on p-values, to detect the location and the magnitude of the multiple outliers AO.

The paper is organized as follow. Section 2 presents the generalization of the AO model, the notations considered in the paper and the Bayesian analysis of the model. In section 3, we perform a test based on p-values to detect location and magnitude of outliers. Section 4 contains some simulation results.

2 Multiple additive outliers contamination

2.1 The model and notations

In this section, we consider an $AR(p)$ model with regression trend which is contaminated by m multiple additive outliers

$$y_t = \underline{x}_t'\underline{\beta} + \sum_{l=1}^{m} \delta_{k_l}(t)\Delta_{k_l} + \mu_t, \qquad t = 1, ..., n. \tag{1a}$$

and $\{\mu_t\}$ is a causal solution of the following $AR(p)$ model,

$$\mu_t = \sum_{i=1}^{p} \phi_i \mu_{t-i} + \varepsilon_t \tag{1b}$$

So $(1a) - (1b)$ can be rewritten as follows:

$$y_t - \underline{x}_t'\underline{\beta} - \sum_{l=1}^{m} \delta_{k_l}(t)\Delta_{k_l} = \sum_{i=1}^{p} \phi_i (y_{t-i} - \underline{x}_{t-i}'\underline{\beta} - \sum_{l=1}^{m} \delta_{k_l}(t-i)\Delta_{k_l}) + \varepsilon_t \tag{1c}$$

where $m = 1, 2, ..., n-1$ is the number of contaminants which occur at time $k_1, k_2, ..., k_m$, where $1 < k_l < n$, $\delta_{k_l}(.)$ is the usual Kronecker function, $l = 1, ..., m$ and $k_1 < k_2 < ... < k_m$.

We suppose that we dispose of the first p initial values of y and x. Otherwise, we fix $y_1, ..., y_p$ and $\underline{x}_1, ..., \underline{x}_p$. In our calculation, we suppose that $\{\varepsilon_t, t = 1, ..., n\}$ are independent random innovations with mean zero and variance σ^2.

Set $\underline{\phi} := (\phi_1, ..., \phi_p)'$, $\underline{\Delta} := (\Delta_{k_1}, ..., \Delta_{k_m})'$.

The parameters $\underline{\beta}$, $\underline{\phi}$, $\underline{\Delta}$, σ of model (1c) are assumed to be unknown and m is known. $\underline{\beta} \in \mathbb{R}^K$, $\underline{\Delta} \in \mathbb{R}^m$, $\underline{\phi} \in \mathbb{R}^p$, $X = (x_{t,j}), t = 1, ..., n$ and $j = 1, ..., K$ is a full rank $(n \times K)$-matrix with fixed elements.

We assume that the autoregressive parameters correspond to stationary and causal processes $\{\mu_t, t \in \mathbb{Z}\}$ in the sense that the parameter vector $\underline{\phi}$ lies in the region $\Phi^{(p)} = \{z \in \mathbb{C} \quad \text{such that} \quad 1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p = 0 \quad \text{implies} \quad |z| > 1\}$.

So, $\underline{\theta} = (\underline{\beta}', \underline{\phi}', \underline{\Delta}', \sigma)' \in \mathbb{R}^{(K+p+m)} \times \mathbb{R}_+^*$ is the vector parameter ($\mathbb{R}_+^* = (0, +\infty)$).

2.2 A Bayesian analysis

We use Gaussian quasi likelihood (GQL) in a Bayesian framework to estimate our parameters.
We assume that our prior knowledge about the parameter σ is vague or diffuse given by $\pi(\sigma) \propto \frac{1}{\sigma}$ and the priors of the parameters $\underline{\beta}$, $\underline{\phi}$ and $\underline{\Delta}$ are given by

$$\pi(\underline{\beta}) = \prod_{i=1}^{K} \pi(\beta_i) \propto c, \quad \pi(\underline{\phi}) = \prod_{i=1}^{p} \pi(\phi_i) \propto c \quad \text{and} \quad \pi(\Delta_{k_l}) \propto c, \quad l = 1, ..., m,$$

where c stands for a generic positive constant. The parameters $\underline{\beta}$, $\underline{\phi}$, $\underline{\Delta}$ and σ are assumed independent. Therefore, the prior of $\underline{\theta}$ is $\pi(\underline{\theta}) \propto \frac{1}{\sigma}$.
With the prior given above and the likelihood function, application of Bayes' Theorem yields to the

following joint posterior distribution

$$
\begin{aligned}
\pi(\underline{\theta}|X,\underline{y}) \quad \propto \quad & \frac{1}{\sigma^{n+1}}\exp\Bigg\{-\frac{1}{2\sigma^2}\Bigg[\sum_{t=1}^{k_1-1}(y_t-\underline{x}_t'\underline{\beta}-\sum_{i=1}^{p}\phi_i(y_{t-i}-\underline{x}_{t-i}'\underline{\beta}))^2 \\
+ \quad & \sum_{l=1}^{m}(y_{k_l}-\underline{x}_{k_l}'\underline{\beta}-\Delta_{k_l}-\sum_{i=1}^{p}\phi_i(y_{k_l-i}-\underline{x}_{k_l-i}'\underline{\beta}-\sum_{\tau=1}^{l-1}\delta_{k_l-k_\tau}(i)\Delta_{k_\tau}))^2 \\
+ \quad & \sum_{l=1}^{m}\sum_{j=1}^{k_{l+1}-k_l-1}(y_{k_l+j}-\underline{x}_{k_l+j}'\underline{\beta}-\sum_{i=1}^{p}\phi_i(y_{k_l+j-i}-\underline{x}_{k_l+j-i}'\underline{\beta}-\sum_{\tau=1}^{l-1}\delta_{k_l-k_\tau+j}(i)\Delta_{k_\tau} \\
- \quad & \delta_j(i)\Delta_{k_l}))^2+\sum_{t=k_m+1}^{n}(y_t-\underline{x}_t'\underline{\beta}-\sum_{i=1}^{p}\phi_i(y_{t-i}-\underline{x}_{t-i}'\underline{\beta}-\sum_{l=1}^{m}\delta_{t-k_l}(i)\Delta_{k_l}))^2\Bigg]\Bigg\}.
\end{aligned}
\tag{2}
$$

Our aim is to show how robust is GQL when the underlying errors are Student-t-distributed or are distributed according to a mixture distribution.
The conditional posterior distribution of the vector regression model $\underline{\beta}$, the conditional posterior distribution of the vector autoregressive parameter $\underline{\phi}$ and the conditional posterior distribution of Δ_{k_l}, for $l=1,...,m$ are given in Theorem 2.1 (see Ait Mohammed and Guerbyenne [3]).

3 A Bayesian significance test

The conditional posterior distribution of Δ_{k_l}, $(l=1,...,m)$ is a Student distribution with parameter location $\tilde{\Delta}_{k_l}$, scale parameter S_{k_l} and $(n-1)$ d.o.f.

$$
\pi(\Delta_{k_l}|\underline{\beta},\underline{\phi},\underline{\Delta}^{(-k_l)},X,\underline{y}) \propto \left\{1+\frac{(\Delta_{k_l}-\tilde{\Delta}_{k_l})^2}{(n-1)S_{k_l}^2}\right\}^{-\frac{n}{2}} \tag{3}
$$

Equivalently, the quantity,

$$
t_l(\Delta_{k_l})=\frac{(\Delta_{k_l}-\tilde{\Delta}_{k_l})}{S_{k_l}}, \qquad l=1,...,m
$$

is distributed a posteriori as a conditional standard Student-t distribution with $(n-1)$ d.o.f given $\underline{\beta}$, $\underline{\phi}$ and $\underline{\Delta}^{(-k_l)}$.
The null hypothesis H_0, that there is no contaminant in the model (1) is

$$H_0:\Delta_{k_l}=0, \quad \forall \quad l=1,...,m$$

against

$$H_1:\Delta_{k_l}\neq 0, \quad \text{for some} \quad l=1,...,m$$

The null hypothesis H_0 can be divided into m sub-nulls

$$H_{0l}:\Delta_{k_l}=0 \quad \text{for} \quad l=1,...,m$$

and H_0 could be rejected if either of these m sub-nulls is rejected.
The proposed test is based on the posterior distribution of Δ_{k_l} (see, Kim [2]).
The unconditional posterior distribution of $t_l(\Delta_{k_l})$, $\quad (l=1,...,m)$ is

$$
\begin{aligned}
\pi(t_l(\Delta_{k_l})|X,\underline{y}) \quad = \quad & \int_{\Phi^p}\int_{\mathbb{R}^K}\int_{\mathbb{R}^{m-1}}\pi(t_l(\Delta_{k_l})|\underline{\beta},\underline{\phi},\underline{\Delta}^{(-k_l)},X,\underline{y})\pi(\underline{\beta}|\underline{\phi},\underline{\Delta}^{(-k_l)},X,\underline{y}) \\
\times \quad & \pi(\underline{\phi}|\underline{\Delta}^{(-k_l)},X,\underline{y})\pi(\underline{\Delta}^{(-k_l)}|X,\underline{y})d\underline{\Delta}^{(-k_l)}d\underline{\beta}d\underline{\phi}.
\end{aligned}
\tag{4}
$$

One defines the highest posterior density credible sets of $t_l(\Delta_{k_l})$. The credible set will be used to define the unconditional p-value and therefore an unconditional test.

For $l = 1, ..., m$, given $\underline{\beta}$, $\underline{\phi}$ and $\underline{\Delta}^{(-k_l)}$, the $(1-\alpha)-$ credible set $(0 < \alpha < 1)$ for $t_l(\Delta_{k_l})$ is defined as

$$C_l = \left\{ t_l(\Delta_{k_l}) : |t_l(\Delta_{k_l})| < t_{\frac{\alpha}{2}}(n-1) \right\}$$

where $t_{\frac{\alpha}{2}}(n-1)$ is the $(1-\frac{\alpha}{2})$-th quantile of a t-distribution with $(n-1)$ d.o.f. Hence, given $\underline{\beta}$, $\underline{\phi}$ and $\underline{\Delta}^{(-k_l)}$, the decision rule for H_{0l} is to reject if $t_l(0) \in \bar{C}_l$, where $\bar{C}_l$ is the complement of C_l.

The unconditional p-value of the sub hypothesis H_{0l} is calculated from (4) to yield

$$\begin{aligned} P_{\Delta_{k_l}=0|\underline{y}} &= 2\int_{\Phi^p}\int_{\mathbb{R}^K}\int_{\mathbb{R}^{m-1}} \{1-\tau_{n-1}(|t_l(0)|)\}\pi(\underline{\beta}, \underline{\phi}, \underline{\Delta}^{(-k_l)}|X, \underline{y}) d\underline{\Delta}^{(-k_l)} d\underline{\beta} d\underline{\phi} \\ &= 2E_{\underline{\phi}}E_{\underline{\beta}}E_{\underline{\Delta}^{(-k_l)}}\{1-\tau_{n-1}(|t_l(0)|)\} \end{aligned} \tag{5}$$

where τ_{n-1} is the cumulative density function of the standard Student-t distribution with $(n-1)$ d.o.f. The expectation is taken with respect to $\underline{\phi}$, $\underline{\beta}$ and $\underline{\Delta}^{(-k_l)}$.

- If the locations of the contaminants are known, then for $l = 1, ..., m$ the hypothesis H_{0l} is rejected unconditionally at α significance level if $P_{\Delta_{k_l}=0|\underline{y}} < \alpha$.
  Therefore, H_0 is rejected unconditionally at α significance level if

$$P_{\underline{\Delta}=0|\underline{y}} := \min_{1\le l\le m}\{P_{\Delta_{k_l}=0|\underline{y}}\} < \alpha \tag{6}$$

- If the locations of the contaminants are unknown then, we propose the following unconditional p-value based test

$$P_{\Delta_{k_l}=0|\underline{y}} = 2\sum_{\substack{j=1\\ j\neq l}}^{m}\sum_{k_j=k_{j-1}+1}^{n-1}\frac{1}{\prod_{\substack{j=1\\ j\neq l}}^{m}(n-k_j-1)}E_{\underline{\phi}}E_{\underline{\beta}}E_{\Delta^{(-k_l)}}\{1-\tau_{n-1}(|t_l(0)|)\}, \quad k_0 = 1, l = 1, ..., m \tag{7}$$

4 A simulation study

- The Gibbs sampler algorithm is applied with $N = 1000$ repetitions. We approximate the posterior mean using a squared error loss function by $\hat{\theta} = \frac{1}{N}\sum_{i=1}^{N}\theta_i^{(m)}$, with a burn-in period of $m-1$ observations to avoid dependence on initial conditions, for N Gibbs sequences of length m. We give the standard deviation (std), the root mean square error (rmse) and the credible interval (CI) of the parameters β, ϕ, Δ_{k_1} and Δ_{k_2}. The parameters estimations are obtained for the true values $(T.V.)$: $\Delta_{k_1} = 2$, $\Delta_{k_2} = 6$, $\phi = 0.5$ and $\beta = 4$ with different sample sizes $n = 50$, $n = 200$ and $n = 400$.

Since in practice, we cannot guarantee that the white noise distribution is standard Gaussian (symmetric, centred and normalized), we examine the accuracy and robustness of the Gibbs estimates when the underlying distribution of the white noise process is standard Gaussian $\mathbb{N}(0,1)$ as a reference against a Student with 5 d.o.f. $(t5)$ and a mixture $(0.8\times\mathbb{N}(0,1)+0.2\times\mathbb{N}(0,2^2))$. For comparison, we give maximum likelihood estimates (MLE) for all parameters, with the same underlying distributions of the white noise process. The results are given in Tables 1-3.

		$n = 50$		$n = 200$		$n = 400$	
parameter	T.V.	Gibbs	MLE	Gibbs	MLE	Gibbs	MLE
$\beta^{(std)}_{(rmse)}$	4	$4.0184^{(0.3259)}_{(0.3377)}$ [3.4552,4.6058]	$4.0108^{(0.3947)}_{(0.4347)}$ [3.2779,4.7944]	$4.0328^{(0.1857)}_{(0.1908)}$ [3.6658,4.3855]	$3.9953^{(0.1756)}_{(0.1756)}$ [3.6543,4.5233]	$3.9992^{(0.1242)}_{(0.1243)}$ [3.6915,4.2353]	$3.9983^{(0.1209)}_{(0.1208)}$ [3.7653,4.2393]
$\phi^{(std)}_{(rmse)}$	0.5	$0.4975^{(0.1279)}_{(0.1280)}$ [0.2451,0.6655]	$0.4856^{(0.1962)}_{(0.2069)}$ [0.1725,0.8196]	$0.4983^{(0.0653)}_{(0.0765)}$ [0.3547,0.6220]	$0.4979^{(0.0623)}_{(0.0623)}$ [0.3698,0.6093]	$0.5040^{(0.0402)}_{(0.0415)}$ [0.4296,0.5732]	$0.4997^{(0.0431)}_{(0.0431)}$ [0.4123,0.5813]
$\Delta_{k_1}{}^{(std)}_{(rmse)}$	2	$2.0129^{(0.8146)}_{(0.8268)}$ [0.2697,3.4152]	$1.9885^{(0.8909)}_{(0.9508)}$ [0.1314,3.8899]	$2.0491^{(0.8413)}_{(0.8430)}$ [0.3010,3.2097]	$1.9645^{(0.8678)}_{(0.8684)}$ [0.2354,3.6658]	$1.9766^{(0.8308)}_{(0.8410)}$ [0.3413,3.5012]	$2.0048^{(0.8945)}_{(0.8940)}$ [0.3894,3.6011]
$\Delta_{k_2}{}^{(std)}_{(rmse)}$	6	$6.0230^{(0.8741)}_{(0.8809)}$ [4.691,7.8392]	$5.9736^{(0.9796)}_{(0.9998)}$ [4.3047,7.7796]	$6.0692^{(0.8638)}_{(0.8608)}$ [4.7513,7.7898]	$5.9565^{(0.8826)}_{(0.8828)}$ [4.5907,7.5337]	$5.9867^{(0.8214)}_{(0.8222)}$ [4.8205,7.6824]	$5.9914^{(0.9159)}_{(0.9154)}$ [4.6063,7.4163]

Table 1 The estimation results when the white noise is Gaussian $N(0,1)$.

		$n=50$		$n=200$		$n=400$	
parameter	T.V.	Gibbs	MLE	Gibbs	MLE	Gibbs	MLE
$\beta^{(std)}_{(rmse)}$	4	$3.9867^{(0.4039)}_{(0.3411)}$ [3.1950,4.9167]	$3.9883^{(0.4941)}_{(0.5040)}$ [2.9976,4.9226]	$4.0191^{(0.2156)}_{(0.2401)}$ [3.6193,4.5085]	$3.9990^{(0.2243)}_{(0.2242)}$ [3.5499,4.4714]	$4.0173^{(0.1573)}_{(0.1989)}$ [3.7291,4.2705]	$4.0075^{(0.1627)}_{(0.1628)}$ [3.6914,4.3388]
$\phi^{(std)}_{(rmse)}$	0.5	$0.4834^{(0.1203)}_{(0.1355)}$ [0.2096,0.6948]	$0.4849^{(0.1990)}_{(0.2498)}$ [0.1729,0.7114]	$0.4960^{(0.0623)}_{(0.0780)}$ [0.3940,0.6140]	$0.4950^{(0.0622)}_{(0.0624)}$ [0.3660,0.6114]	$0.4908^{(0.0417)}_{(0.0562)}$ [0.4223,0.5726]	$0.4972^{(0.0440)}_{(0.0441)}$ [0.4026,0.5857]
$\Delta_{k_1}{}^{(std)}_{(rmse)}$	2	$2.0052^{(1.0512)}_{(0.9401)}$ [−0.1220,4.3084]	$2.0637^{(1.1992)}_{(1.2391)}$ [−0.1048,4.6009]	$2.0967^{(1.1000)}_{(0.9311)}$ [0.0215,4.2002]	$1.9854^{(1.1421)}_{(1.1460)}$ [0.1442,4.4105]	$2.0469^{(1.0971)}_{(0.9280)}$ [0.2480,4.1131]	$2.0126^{(1.1126)}_{(1.1022)}$ [0.2186,4.2008]
$\Delta_{k_2}{}^{(std)}_{(rmse)}$	6	$5.9908^{(1.1227)}_{(1.0601)}$ [3.7981,7.9838]	$6.0335^{(1.1831)}_{(1.1743)}$ [3.6190,8.2662]	$6.0916^{(1.0933)}_{(1.0100)}$ [3.9765,7.2901]	$5.9732^{(1.0772)}_{(1.0770)}$ [3.9188,7.3634]	$6.0119^{(0.9267)}_{(0.9209)}$ [4.1707,7.2235]	$5.9862^{(0.9044)}_{(0.9161)}$ [3.9878,7.2831]

Table 2 The estimation results when the white noise is Student $t5$.

		$n=50$		$n=200$		$n=400$	
parameter	T.V.	Gibbs	MLE	Gibbs	MLE	Gibbs	MLE
$\beta^{(std)}_{(rmse)}$	4	$3.9931^{(0.3533)}_{(0.3731)}$ [3.1374,4.9836]	$4.0223^{(0.4814)}_{(0.4917)}$ [3.0884,4.9939]	$4.0022^{(0.2242)}_{(0.2346)}$ [3.6319,4.4939]	$4.0093^{(0.2190)}_{(0.2193)}$ [3.6043,4.4625]	$4.0030^{(0.1567)}_{(0.1662)}$ [3.6606,4.3379]	$4.0053^{(0.1511)}_{(0.1511)}$ [3.7112,4.3039]
$\phi^{(std)}_{(rmse)}$	0.5	$0.4879^{(0.1063)}_{(0.1066)}$ [0.1975,0.7004]	$0.4809^{(0.2090)}_{(0.2673)}$ [0.1762,0.7127]	$0.4950^{(0.0741)}_{(0.0766)}$ [0.3293,0.6089]	$0.4939^{(0.0645)}_{(0.0647)}$ [0.3649,0.6108]	$0.4976^{(0.0445)}_{(0.0453)}$ [0.4177,0.5803]	$0.4989^{(0.0432)}_{(0.0432)}$ [0.4118,0.5808]
$\Delta_{k_1}{}^{(std)}_{(rmse)}$	2	$1.9976^{(1.1247)}_{(1.1303)}$ [−0.1031,4.0756]	$2.0407^{(1.2863)}_{(1.2968)}$ [−0.2793,4.2814]	$1.998^{(1.0928)}_{(1.1064)}$ [0.1520,4.0054]	$2.0323^{(1.1597)}_{(1.1598)}$ [0.0199,4.1473]	$1.9990^{(1.0203)}_{(1.0246)}$ [0.2202,3.9783]	$1.9907^{(1.1035)}_{(1.1116)}$ [0.3157,3.8826]
$\Delta_{k_2}{}^{(std)}_{(rmse)}$	6	$6.0088^{(1.0088)}_{(1.0198)}$ [3.8687,8.6828]	$5.9494^{(1.1898)}_{(1.1799)}$ [3.5398,8.3619]	$6.0017^{(1.0001)}_{(1.0091)}$ [3.8149,7.7579]	$6.0060^{(1.0960)}_{(1.0996)}$ [3.7132,8.3250]	$6.0071^{(0.9704)}_{(0.9806)}$ [3.9456,7.7356]	$5.9852^{(0.9870)}_{(0.9876)}$ [3.9215,8.3027]

Table 3 The estimation results when the white noise is Gaussian mixture $0.8 \times \mathbb{N}(0,1) + 0.2 \times \mathbb{N}(0,2^2)$.

We notice that, for the three cases of the white noise, the two estimation methods, Gibbs sampler and MLE, give good results; when the sample size increases, the $rmse$ and the std of the estimations decrease. Comparing the two methods, we notice that, when the sample size is small, the Bayesian method is better than the classical method. When the sample size is large the two methods are almost similar.

- The Gibbs sampler algorithm is applied with 1000 repetitions for $n = 200$ and $\phi = 0.5$ to approximate the unconditional p-value test for the hypothesis $H_0 : \Delta_{k_l} = 0, \quad l \in \{1, 2\}$ for different values of σ and $(\Delta k_1, \Delta k_2)$ (small, moderate and great values). The results are given in Tables 4-6.

$\sigma = 0.1$	k_1-2	k_1-1	k_1	k_1+1	k_1+2		k_2-2	k_2-1	k_2	k_2+1	k_2+2
$\Delta_{k_1}=0.5$	0.2408	0.2084	0.0335	0.6389	0.4151	$\Delta_{k_2}=1$	0.3301	0.1086	0.00009	0.7551	0.3891
$\Delta_{k_1}=2$	0.56797	0.0220	0.00009	0.0404	0.6611	$\Delta_{k_2}=4$	0.6871	0.0155	0.00009	0.0293	0.6085
$\Delta_{k_1}=8$	0.8787	0.0062	0.00009	0.0097	0.8864	$\Delta_{k_2}=15$	0.9805	0.0004	0.00009	0.0045	0.8541

Table 4 The unconditional p-value test of H_0, $P_{\Delta_{k_l}=0}|\underline{y}$ for $\sigma = 0.1$.

$\sigma = 0.5$	k_1-2	k_1-1	k_1	k_1+1	k_1+2		k_2-2	k_2-1	k_2	k_2+1	k_2+2
$\Delta_{k_1}=0.5$	0.9449	0.7812	0.4033	0.5938	0.9360	$\Delta_{k_2}=1$	0.4402	0.0810	0.00009	0.0212	0.6255
$\Delta_{k_1}=2$	0.1690	0.0808	0.00009	0.0449	0.2317	$\Delta_{k_2}=4$	0.7245	0.0339	0.00009	0.0479	0.1941
$\Delta_{k_1}=8$	0.8632	0.0098	0.00009	0.0211	0.4190	$\Delta_{k_2}=15$	0.9508	0.0010	0.00009	0.0025	0.7241

Table 5 The unconditional p-value test of H_0, $P_{\Delta_{k_l}=0}|\underline{y}$ for $\sigma = 0.5$.

$\sigma = 1$	k_1-2	k_1-1	k_1	k_1+1	k_1+2		k_2-2	k_2-1	k_2	k_2+1	k_2+2
$\Delta_{k_1}=0.5$	0.1759	0.4780	0.2042	0.3646	0.3909	$\Delta_{k_2}=1$	0.1231	0.7433	0.1134	0.2127	0.8529
$\Delta_{k_1}=2$	0.7493	0.2928	0.0020	0.1245	0.7861	$\Delta_{k_2}=4$	0.4361	0.0625	0.0002	0.0471	0.7212
$\Delta_{k_1}=8$	0.8880	0.0495	0.00009	0.0297	0.7625	$\Delta_{k_2}=15$	0.2665	0.0034	0.00009	0.0015	0.8541

Table 6 The unconditional p-value test of H_0, $P_{\Delta_{k_l}=0}|\underline{y}$ for $\sigma = 1$.

One notices, when the variance is smaller, the p-value test can detect even small amplitudes of contamination. We notice also that, the p-value test detects the locations of contamination k_1 and k_2; it also detects the locations k_1-1, k_1+1, k_2-1 and k_2+1. This is due to the fact that $\tilde{\Delta}_{k_l}$ in the test statistic at time k_l, depends only on the observations y_{k_l}, y_{k_l-1} and y_{k_l+1}, $l = 1, ..., m$.

5 Conclusion

In this paper, a Bayesian inference has been conducted for an autoregressive time series model of order p with a regression trend when it is contaminated by multiple additive outliers. Using a quasi maximum likelihood combined with non informative or diffuse priors, posterior distributions of the models parameters are obtained. We used the Gaussian quasi likelihood (GQL) in a Bayesian framework to estimate our parameters. Our aim is to show how robust is GQL when the underlying errors are not Gaussian.

A simulation study was conducted in which we examined the accuracy and robustness of the Gibbs estimates when the underlying distribution of the error process is standard Gaussian as a reference against, Student with 5 d.o.f. and a Gaussian mixture. It shows that model parameters are well estimated for AO contamination. A comparison with the MLE approach shows that the Bayesian method is better than the classical one when the sample size is small.

Unconditional p-value based tests are proposed for detecting the locations and the magnitudes of contaminations for multiple AO outliers. The p-value based test is suggested, when the number of contaminants is unknown.

References

1 AJ. Fox. Outliers in time series. *Journal of the Royal Statistical Society*, 1972;B:350–363, 34.

2 D. Kim. A bayesian significance test of the stationarity of regression parameters. *Biometrika*, 78(3):667–675, 1991.

3 N. Ait Mohammed and H. Guerbyenne. Bayesian inference in a multiple contaminated autoregressive model with trend. *Journal of statistical computation and simulation*, 93(12):2067–2109, 2023.

4 RS. Tsay. Time series model specification in the presence of outliers. *Journal of the American Statistical Association*, 81:132–141, 1986.