



A Novel Comparison between the Ordinary Estimators and Robust Estimators for the Parameters of Some Binary Mixed Models

Ahmad Khaldi^{1,*}

¹ Mutah University, Faculty of Science, Jordan

Email: khaldiahmad1221@gmail.com

Received: January 10, 2024 Accepted: July 04, 2024 ★ Corresponding author

ABSTRACT

A condensed study is conducted to compare ordinary estimators, particularly the maximum likelihood estimator, with a robust estimator for estimating the parameters of the first-order binary mixed model BARMA (1, 1). Simulation experiments are applied to several BARMA (1, 1) configurations using small, moderate, and large sample sizes, from which new results are obtained. The mean absolute percentage error (MAPE) is used as the statistical criterion for comparison.

Keywords: Binary system ▪ Mathematical model ▪ Estimator ▪ Numerical comparison

1. INTRODUCTION

One of the most widely used methods for estimating the parameters of time-series models is the least-squares method, denoted by LS when the distribution of errors is unknown. The maximum likelihood method, denoted by ML, is used when the distribution of errors is known. The method of moments, denoted by MOM, is also known as the Yule–Walker method.

The estimators produced by these methods may be efficient, consistent, and appropriate when the time-series conditions are satisfied. These conditions generally include normality, stationarity, and invertibility. When the conditions change because of an external or contingent factor, however, another estimation method is required. Even a small change in the data structure, usually caused by outliers appearing in the errors, directly affects the assumed distribution of those errors.

The aim of this research is to compare an ordinary method and a robust method for estimating the parameters of the lower-order binary mixed model BARMA (1,1). Both a

theoretical method (statistical theory) and an experimental method (simulation) are used to estimate the parameters of randomly generated time series, first without outliers and then after introducing outliers. The estimators are compared using the mean absolute percentage error (MAPE). This criterion is used because it is considered one of the most accurate criteria for comparing estimation methods in time-series applications. Mean squared error (MSE) is not used because squaring each error gives large errors greater weight than small errors, which can make practical comparisons between estimation methods less informative (Daniel, 2004).

2. MAIN DISCUSSION

2.1 First-Order Binary Mixed Autoregressive Moving-Average Model: BARMA (1, 1)

The BARMA (1, 1) model is written as

$$(I - \Phi_1 B)y_t = (I - \Theta_1 B)a_t. \quad (1)$$

First: the model is stable if the roots of $|I - \Phi_1 B| = 0$ lie outside the unit circle, or equivalently if the characteristic values of Φ_1 lie inside the unit circle.

Second: the model can be written in terms of random errors as

$$y_t = \sum_{s=0}^{\infty} \Psi_s a_{t-s}. \quad (2)$$

The weights Ψ_s are obtained by equating the coefficients of B^j in

$$(I - \Phi_1 B)(I + \Psi_1 B + \Psi_2 B^2 + \dots) = (I - \Theta_1 B).$$

Thus,

$$\Psi_j = \Phi_1 \Psi_{j-1} = \Phi_1^{j-1} (\Phi_1 - \Theta_1), \quad j \geq 1. \quad (3)$$

The model is invertible if the roots of $|I - \Theta_1 B| = 0$ lie outside the unit circle, or equivalently if the characteristic values of Θ_1 lie inside the unit circle.

Third: the covariance matrix can be derived from

$$E[y_t(y_t - y'_{t-1} \Phi'_1)] = E[y'_{t-k}(a'_t - a'_{t-1} \Theta'_1)].$$

We note that

$$\begin{aligned} E[y_t(a'_{t-1} \Theta'_1)] &= E[(\Phi_1 y_{t-1} + a_t - \Theta_1 a_{t-1})(a'_{t-1} \Theta'_1)] \\ &= \Phi_1 \Sigma \Theta'_1 - \Theta_1 \Sigma \Theta'_1. \end{aligned} \quad (4)$$

Consequently,

$$\begin{aligned} \Gamma(0) - \Gamma'(1) \Sigma &= \Sigma - (\Phi_1 - \Theta_1) \Sigma \Theta_1, & k=0, \\ \Gamma(1) - \Gamma(0) \Phi'_1 &= -\Sigma \Theta'_1, & k=1, \\ \Gamma(k) - \Gamma'(k-1) \Phi'_1 &= 0, & k \geq 2. \end{aligned} \quad (5)$$

Equivalently,

$$\Gamma(k) = \begin{cases} \Gamma(1) \Sigma + \Sigma - (\Phi - \Theta) \Sigma \Theta, & k=0, \\ \Gamma(0) \Phi'_1 - \Sigma \Theta_1, & k=1, \\ \Gamma(k-1) \Phi'_1, & k \geq 2. \end{cases}$$

For a mixed two-variable BARMA (1,1) model,

$$(I - \Phi_1 B)y_t = (I - \Theta_1 B)\varepsilon_t,$$

where

$$\Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}, \quad \Theta_1 = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}.$$

The covariance matrix of the vector $[\varepsilon_{1t}, \varepsilon_{2t}]'$ is

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}.$$

Assuming that $[y_{1t}, y_{2t}]'$ follows a bivariate normal distribution, the log-likelihood for the sample $(y_{11}, y_{21}, \dots, y_{1n}, y_{2n})$ is

$$\ln L(\Phi_1, \Theta_1, \Sigma | y) = C - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^n \varepsilon_t^T \Sigma^{-1} \varepsilon_t. \quad (6)$$

It can also be written as

$$\ln L = C - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \text{tr}[\Sigma^{-1} S(\Phi_1, \Theta_1)], \quad (7)$$

where

$$S(\Phi_1, \Theta_1) = \sum_{t=1}^n \varepsilon_t \varepsilon_t^T, \quad \varepsilon_t = y_t - \Phi_1 y_{t-1} + \Theta_1 \varepsilon_{t-1}.$$

The maximum-likelihood estimates of Φ_1 , Θ_1 , and Σ can therefore be calculated by maximizing the log-likelihood function.

The two-variable autoregressive model BARMA (1,0) is a special case:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}. \quad (8)$$

Its log-likelihood is

$$\ln L(\Phi_1, \Sigma | y) = C - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^n \varepsilon_t^T \Sigma^{-1} \varepsilon_t. \quad (9)$$

Using matrix properties, this becomes

$$\ln L = C - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \text{tr}[\Sigma^{-1} S(\Phi_1)], \quad (10)$$

with

$$S(\Phi_1) = \sum_{t=1}^n \varepsilon_t \varepsilon_t^T, \quad \varepsilon_t = y_t - \Phi_1 y_{t-1}.$$

The mixed two-variable model BARMA (0,1) is another special case:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} - \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}. \quad (11)$$

After simplification, the log-likelihood for a sample of size n is

$$\ln L = C - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \text{tr}[\Sigma^{-1} S(\Theta_1)], \quad (12)$$

where

$$S(\Theta_1) = \sum_{t=2}^n \varepsilon_t \varepsilon_t^T, \quad \varepsilon_t = y_t + \Theta_1 \varepsilon_{t-1}.$$

2.2 Robust Estimation of Model Parameters

Outlying values in a time series can adversely affect both least-squares estimators and M-estimators of autoregressive parameters. The focus here is on obtaining robust estimators of first-order autoregressive coefficients. Let the observations be $y_t = x_t + v_t$. Two models are considered: an innovation-outlier (IO) model, in which $v_t = 0$ and x_t may follow a non-Gaussian distribution, and an additive-outlier (AO) model, in which v_t is nonzero with a small probability and may be large while x_t is normal.

This paper studies robust estimation for lower-order autoregressive models when the corresponding time series contains outliers. Proposing robust procedures requires characterizing contaminated time series with suitable probability models. Because complete probability models are difficult to formu-

late (Martin, 1980), it is useful to begin with simple outlier-generating models. In practice, outliers commonly follow one of three forms:

- a– The occurrence of an outlier is generally related to the remaining observations, except in the case of an initial abrupt change.
- b– A large-error outlier may occur for several reasons, including a recording error.
- c– An outlier may behave differently from the rest of the sample, possibly because the recording medium was used inadequately.

The first behavior can be represented by an innovation-outlier model if the observed values equal the underlying process values (Zch, 1979):

$$y_i = \hat{x}_i. \quad (13)$$

If the IO error distribution is symmetric, the innovation model may use a t distribution or a contaminated normal distribution G :

$$G(p, \sigma_1, \sigma_2) = (1 - p)N(0, \sigma_1^2) + pN(0, \sigma_2^2), \quad \sigma_2^2 > \sigma_1^2,$$

where p is usually small. Thus, if the white-noise values ε_t are independent and identically distributed according to a symmetric distribution G with mean zero and scale parameter σ , they are called innovations.

The second and third behaviors are represented by the additive-outlier model

$$Y_t = X_t + V_t, \quad (14)$$

where V_t is independent of X_t and, for a relatively small contamination probability, has the marginal distribution

$$P(V_t = 0) = 1 - \gamma.$$

Time-series experiments suggest that γ commonly lies between 0.01 and 0.25 (Stochinger and Duter, 1987). A contaminated normal distribution for V_t is

$$\text{CND}(p, \sigma_3) = (1 - \gamma)\delta_0 + \gamma N(p, \sigma_3^2), \quad (15)$$

where δ_0 denotes a degenerate distribution whose mass is concentrated at zero.

Fox (1973) distinguished between outliers that affect only the observation at which they occur (type I, or innovation outliers) and outliers that affect subsequent observations as well (type II, or additive outliers). He also proposed two approaches for identifying the outlier type: select the model in which the outlying observation is more extreme, or select the model according to the observed propagation of the outlier's effect into later observations.

3. THE EXPERIMENTAL SIDE

3.1 Simulation Model Formulation

Simulation imitates a real system or model without requiring the system itself. It is particularly useful when statistical problems and theories are difficult to prove mathematically,

because they can be translated into experimental populations from which random samples are drawn.

To achieve the main goal of this research, a simulation model was formulated to compare the ordinary and robust methods for estimating the parameters of the mixed two-variable BARMA (1, 1) model.

A: Data generation. For the BARMA (1, 1) model,

$$\begin{aligned} \Phi(B)y_t &= \Theta(B)\varepsilon_t, \\ \varepsilon_t &= \Theta^{-1}(B)\Phi(B)y_t, \\ \varepsilon'_t \varepsilon_t &= y'_t \Phi'(B) \Theta^{-1}(B) \Theta^{-1}(B) \Phi(B) y_t. \end{aligned}$$

The last expression is differentiated with respect to the mixed-model parameters (Φ, Θ) to obtain their estimates. The vector ε_t follows a bivariate normal distribution:

$$\varepsilon_t \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right).$$

B: Model dimension. The number of variables in the model is set to $m = 2$.

C: Parameter configurations. The following 13 parameter configurations satisfy the stability and invertibility requirements:

$$\begin{aligned} 1 \text{ ;; } \Phi &= \begin{bmatrix} 0.8 & -0.2 \\ 0.2 & -0.6 \end{bmatrix}, & \Theta &= \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & -0.6 \end{bmatrix}, \\ 2 \text{ ;; } \Phi &= \begin{bmatrix} 0.8 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.6 & 0.3 \\ -0.3 & -0.8 \end{bmatrix}, \\ 3 \text{ ;; } \Phi &= \begin{bmatrix} 0.8 & -0.1 \\ -0.2 & 0.2 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.2 & 0.6 \\ -0.6 & 0.2 \end{bmatrix}, \\ 4 \text{ ;; } \Phi &= \begin{bmatrix} 0.8 & -0.6 \\ 0.5 & 0.6 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.2 & 0.1 \\ -0.5 & 0.2 \end{bmatrix}, \\ 5 \text{ ;; } \Phi &= \begin{bmatrix} 0.8 & -0.6 \\ 0.2 & 0.8 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.8 & 0.4 \\ 0.3 & 0.8 \end{bmatrix}, \\ 6 \text{ ;; } \Phi &= \begin{bmatrix} 0.8 & 0.2 \\ -0.2 & 0.6 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.6 & 0.2 \\ 0.2 & 0.6 \end{bmatrix}, \\ 7 \text{ ;; } \Phi &= \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.2 \end{bmatrix}, & \Theta &= \begin{bmatrix} 0.3 & 0.1 \\ 0.2 & 0.5 \end{bmatrix}, \\ 8 \text{ ;; } \Phi &= \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.7 \end{bmatrix}, & \Theta &= \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 0.7 \end{bmatrix}, \\ 9 \text{ ;; } \Phi &= \begin{bmatrix} 0.6 & -0.4 \\ -0.3 & 0.2 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.3 & 0.1 \\ 0.2 & 0.5 \end{bmatrix}, \\ 10 \text{ ;; } \Phi &= \begin{bmatrix} -0.8 & -0.6 \\ -0.5 & 0.4 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.2 & 0.1 \\ -0.5 & 0.3 \end{bmatrix}, \\ 11 \text{ ;; } \Phi &= \begin{bmatrix} -0.8 & 0.6 \\ 0.6 & 0.8 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}, \\ 12 \text{ ;; } \Phi &= \begin{bmatrix} -0.8 & 0.2 \\ 0.2 & -0.8 \end{bmatrix}, & \Theta &= \begin{bmatrix} -0.8 & 0.2 \\ 0.2 & -0.8 \end{bmatrix}, \end{aligned}$$

$$13 :: \Phi = \begin{bmatrix} -0.9 & 0.1 \\ 0.2 & 0.7 \end{bmatrix}, \quad \Theta = \begin{bmatrix} 0.8 & 0.4 \\ 0.6 & -0.5 \end{bmatrix}.$$

D: Sample size and contamination. Three sample sizes are considered: $n = 25, 50$, and 100 . The parameters of the mixed two-variable model are estimated by the ordinary and robust methods. The data are then contaminated with 10% outliers. MAPE is used for all comparisons. The ordinary method performs better before contamination, whereas the robust method is generally superior after 10% contamination. Table 1 gives the results for $n = 25$; results for $n = 50$ and $n = 100$ are reported by Safawi (2005).

4. CONCLUSION

- I– For a small sample, maximum likelihood is preferred in 87.5% of the uncontaminated cases according to MAPE, whereas the robust method is superior in 95% of the contaminated cases.
- II– For a moderate sample, maximum likelihood is preferred in 95% of the uncontaminated cases, whereas the robust method is superior in 93% of the contaminated cases.
- III– For a large sample, maximum likelihood is preferred in 97% of the uncontaminated cases, whereas the robust method is superior in 97% of the contaminated cases.

Table 1. The mean absolute percentage errors of the BARMA (1, 1) parameter estimates for sample size $n = 25$.

Lineup	Default parameter	Before polluting			After polluting		
		ML	Robust	Best	ML	Robust	Best
1	$\phi_{11} = 0.8$	0.00191551	0.0794192	ML	0.00050000	0.000306387	RO
	$\phi_{12} = -0.2$	0.00245530	0.0020000	RO	0.00200000	0.000000200	RO
	$\phi_{21} = 0.2$	0.00165431	0.0020000	ML	0.00200000	0.000000200	RO
	$\phi_{22} = -0.6$	0.00221439	0.0020000	RO	0.00533333	0.000000200	RO
	$\theta_{11} = 0.6$	0.00174262	7.7386600	ML	0.00173600	0.000773866	RO
	$\theta_{12} = 0.4$	0.00276808	1.8020000	ML	0.00269300	0.000180200	RO
	$\theta_{21} = 0.5$	0.00237627	1.4420000	ML	0.00231680	0.000144200	RO
	$\theta_{22} = -0.6$	0.00218199	1.5308200	ML	0.0024620	0.000153082	RO
2	$\phi_{11} = 0.8$	0.001942910	0.0816621	ML	0.0005000	0.00027991	RO
	$\phi_{12} = 0.1$	0.002479630	0.0020000	RO	0.0020000	0.00000020	RO
	$\phi_{21} = 0.1$	0.001736420	0.0020000	RO	0.0020000	0.00000020	RO
	$\phi_{22} = 0.2$	0.002167800	0.0020000	RO	0.0053333	0.00000020	RO
	$\theta_{11} = -0.6$	0.001761250	98.3582000	ML	0.0017360	0.00974395	ML
	$\theta_{12} = 0.3$	0.002716140	1.8020000	ML	0.0026930	0.00018770	RO
	$\theta_{21} = -0.3$	0.002328190	1.8420000	ML	0.0023168	0.00015020	RO
	$\theta_{22} = -0.8$	0.002132190	0.3077610	ML	0.0024620	0.000033175	RO
3	$\phi_{11} = 0.8$	0.00191844	0.8367640	ML	0.00050000	0.000074563	RO
	$\phi_{12} = -0.1$	0.00249507	0.0020000	RO	0.00200000	0.000000200	RO
	$\phi_{21} = -0.2$	0.00179517	0.0020000	ML	0.00200000	0.000000200	RO
	$\phi_{22} = 0.2$	0.00212376	0.0020000	RO	0.00533333	0.000000200	RO
	$\theta_{11} = -0.2$	0.00178199	81.9702000	ML	0.00570565	0.008203560	ML
	$\theta_{12} = 0.6$	0.00273501	1.802000	ML	0.00921596	0.000187700	RO
	$\theta_{21} = -0.6$	0.00238797	1.442000	ML	0.00244678	0.000150200	RO
	$\theta_{22} = 0.2$	0.00215055	8.945280	ML	0.00547731	0.000904132	RO
4	$\phi_{11} = 0.8$	0.00195391	0.0823780	ML	0.00166667	0.000235597	RO
	$\phi_{12} = -0.6$	0.00273353	0.0020000	RO	0.10666770	0.000000200	RO
	$\phi_{21} = 0.5$	0.00172039	0.0020000	ML	0.00666770	0.000000200	RO
	$\phi_{22} = 0.6$	0.00224935	0.0020000	RO	0.10666770	0.000000200	RO
	$\theta_{11} = -0.2$	0.00177402	15.4597000	ML	0.00200000	0.001541920	RO
	$\theta_{12} = 0.1$	0.00272224	1.8020000	ML	0.00533333	0.000187700	RO
	$\theta_{21} = -0.5$	0.00235209	1.4420000	ML	0.00269300	0.000150200	RO
	$\theta_{22} = 0.2$	0.00215510	1.2216700	ML	0.00246200	0.000133549	RO
5	$\phi_{11} = 0.8$	0.00195735	0.0811064	ML	0.00050000	0.000310016	RO
	$\phi_{12} = -0.6$	0.00236050	0.0020000	RO	0.00020000	0.000000200	RO
	$\phi_{21} = 0.2$	0.00179132	0.0020000	ML	0.00020000	0.000000200	RO
	$\phi_{22} = -0.8$	0.00219714	0.0020000	RO	0.00533330	0.000000200	RO
	$\theta_{11} = -0.8$	0.00172990	49.6713000	ML	0.00213891	0.004970960	ML
	$\theta_{12} = 0.4$	0.00271187	1.802000	ML	0.00239570	0.000187700	RO
	$\theta_{21} = 0.3$	0.00232818	1.442000	ML	0.00183331	0.000150200	RO
	$\theta_{22} = 0.8$	0.00210534	0.236148	ML	0.00226380	0.000018466	RO
6	$\phi_{11} = 0.8$	0.00194614	0.0818381	ML	0.0005000	0.000391503	RO
	$\phi_{12} = 0.2$	0.00253378	0.0020000	RO	0.0020000	0.000000200	RO
	$\phi_{21} = -0.2$	0.00152255	0.0020000	ML	0.0020000	0.000000200	RO
	$\phi_{22} = 0.6$	0.00222693	0.0020000	RO	0.0005333	0.000000200	RO
	$\theta_{11} = -0.6$	0.00180898	1.5479300	ML	0.0017360	0.000155020	RO
	$\theta_{12} = 0.2$	0.00268366	1.8020000	ML	0.0026930	0.000187700	RO
	$\theta_{21} = 0.2$	0.00239610	1.4420000	ML	0.0023168	0.000150200	RO
	$\theta_{22} = 0.6$	0.00216408	0.4944350	ML	0.0024620	0.000057406	RO

Table 1 continued

Lineup	Default parameter	Before polluting			After polluting		
		ML	Robust	Best	ML	Robust	Best
7	$\phi_{11} = 0.6$	0.00182011	0.0277795	ML	0.00022222	0.000000670	RO
	$\phi_{12} = 0.4$	0.00175529	18.2857000	ML	0.00200000	0.001691860	RO
	$\phi_{21} = 0.3$	0.00158952	9.33511000	ML	0.00200000	0.000947070	RO
	$\phi_{22} = 0.2$	0.00156523	0.18340000	ML	0.00085714	0.000054450	RO
	$\theta_{11} = 0.3$	0.00185794	19.47450000	ML	0.00168320	0.002040110	ML
	$\theta_{12} = 0.1$	0.00418934	48.74230000	ML	0.00338600	0.000488670	RO
	$\theta_{21} = 0.2$	0.00276094	48.74230000	ML	0.00279200	0.000488670	RO
	$\theta_{22} = 0.5$	0.00196439	0.246852000	ML	0.00160400	0.000018926	RO
8	$\phi_{11} = 0.9$	0.001969780	0.0277795	ML	0.00022222	0.000000067	RO
	$\phi_{12} = 0.1$	0.000704745	18.2857000	ML	0.00200000	0.001691860	RO
	$\phi_{21} = 0.2$	0.001634340	9.3351100	ML	0.00200000	0.000947570	RO
	$\phi_{22} = 0.7$	0.001809310	0.1834000	ML	0.00085714	0.000054450	RO
	$\theta_{11} = 0.5$	0.001753060	19.4745000	ML	0.00168320	0.002040110	ML
	$\theta_{12} = 0.2$	0.003484420	48.7423000	ML	0.00338600	0.000488670	RO
	$\theta_{21} = 0.2$	0.003125470	48.7423000	ML	0.00279200	0.000488670	RO
	$\theta_{22} = 0.7$	0.001840750	0.2468520	ML	0.00160400	0.000018926	RO
9	$\phi_{11} = 0.6$	0.00183650	0.07494880	ML	0.00133333	0.00026086	RO
	$\phi_{12} = 0.4$	0.00226427	12.14130010	ML	0.00200000	0.00193030	RO
	$\phi_{21} = -0.3$	0.00226835	18.20250000	ML	0.00200000	0.00178275	RO
	$\phi_{22} = 0.2$	0.00153823	2.67203000	ML	0.00800000	0.000133702	RO
	$\theta_{11} = -0.3$	0.00218242	1.66585000	ML	0.00165187	0.000166585	RO
	$\theta_{12} = 0.1$	0.00440359	8.36564000	ML	0.00484763	0.000836564	RO
	$\theta_{21} = 0.2$	0.00271048	4.18382000	ML	0.00147781	0.000418380	RO
	$\theta_{22} = 0.5$	0.00198422	34.47560000	ML	0.00249695	0.000344756	RO
10	$\phi_{11} = -0.8$	0.00216153	0.4701300	ML	0.00450000	0.000572389	RO
	$\phi_{12} = -0.6$	0.00220332	1.9932900	ML	0.00200000	0.000209329	RO
	$\phi_{21} = -0.5$	0.00218370	5.2971500	ML	0.00200000	0.000679715	RO
	$\phi_{22} = 0.4$	0.00180604	1.0726100	ML	0.00300000	0.000266666	RO
	$\theta_{11} = -0.2$	0.00223749	8.7182500	ML	0.01199520	0.000870834	RO
	$\theta_{12} = 0.1$	0.00441582	1.4358000	ML	0.01001520	0.000161169	RO
	$\theta_{21} = -0.5$	0.00168816	0.2895600	ML	0.00759807	0.000304720	RO
	$\theta_{22} = 0.3$	0.00203284	39.608400	ML	0.00600506	0.004069800	RO
11	$\phi_{11} = -0.8$	0.00200798	0.275581	ML	0.00450000	0.00095766	RO
	$\phi_{12} = 0.6$	0.00188508	0.950183	ML	0.00200000	0.00007830	RO
	$\phi_{21} = 0.6$	0.00190857	1.336290	ML	0.00200000	0.000143004	RO
	$\phi_{22} = -0.8$	0.00209497	6.274660	ML	0.00450000	0.000329157	RO
	$\theta_{11} = -0.8$	0.00217061	0.205902	ML	0.0021980	0.000021845	RO
	$\theta_{12} = 0.6$	0.00249157	1.372900	ML	0.0024620	0.000141910	RO
	$\theta_{21} = 0.6$	0.00225630	1.372900	ML	0.0022640	0.000144191	RO
	$\theta_{22} = -0.8$	0.00209207	8.237870	ML	0.002346	0.000844106	RO
12	$\phi_{11} = -0.8$	0.00206821	0.0507938	ML	0.00450000	0.000300210	RO
	$\phi_{12} = 0.2$	0.00162002	12.0031000	ML	0.00200000	0.001050310	RO
	$\phi_{21} = 0.2$	0.00164455	7.4499100	ML	0.00200000	0.000735616	RO
	$\phi_{22} = -0.8$	0.00211576	0.3514030	ML	0.00450000	0.000017863	RO
	$\theta_{11} = -0.8$	0.00220751	4.7931400	ML	0.0021980	0.000478334	RO
	$\theta_{12} = 0.2$	0.00360058	14.4777000	ML	0.0033860	0.001067310	RO
	$\theta_{21} = 0.2$	0.00293582	14.4777000	ML	0.0027920	0.001607310	RO
	$\theta_{22} = -0.8$	0.00209540	0.3059830	ML	0.0023465	0.000067775	RO
13	$\phi_{11} = -0.9$	0.00207900	0.0503912	ML	0.004222220	0.000179170	RO
	$\phi_{12} = 0.1$	0.00103921	6.3309000	ML	0.002000000	0.001099760	RO
	$\phi_{21} = 0.2$	0.00203168	2.7594100	ML	0.002000000	0.000444691	RO
	$\phi_{22} = 0.7$	0.00187348	0.3486920	ML	0.000857143	0.000020147	RO
	$\theta_{11} = 0.8$	0.00179222	0.6476790	ML	0.001802000	0.000070287	RO
	$\theta_{12} = 0.4$	0.00280789	1.7994800	ML	0.002693000	0.000184471	RO
	$\theta_{21} = 0.6$	0.00231321	1.2003200	ML	0.002264000	0.000123048	RO
	$\theta_{22} = -0.5$	0.00219796	4.7930100	ML	0.002554400	0.000482457	RO

In the second experiment, when the sample size is small ($n = 25$), maximum likelihood outperforms the robust method for 87.5% of the parameter values before contamination. After contamination, the robust method outperforms maximum likelihood in 95% of the cases. MAPE is used as the comparison criterion throughout.

REFERENCES

- [1] S. Y. Safawi, "Comparison of normal and robust estimators of mixed binary time series models of lower ranks," Ph.D. dissertation, Faculty of Computer Science and Mathematics, University of Mosul, 2005.
- [2] M. D. Abdul-Ahad, "Hippocampal estimation in a first-order self-regression model," unpublished master's thesis, Faculty of Computer Science and Mathematics, University of Mosul, 2004.
- [3] R. D. Martin, "Robust estimation of autoregressive models," in *Directions in Time Series*, D. R. Brillinger and G. C. Tiao, Eds. Hayward, CA, USA: Institute of Mathematical Statistics, 1980.
- [4] A. J. Fox, "Outliers in time series," *Journal of the Royal Statistical Society: Series B*, vol. 34, pp. 350–363, 1973.
- [5] N. Stochinger and R. Duter, "Robust time series analysis: A survey," supplement to *Kybernetika*, 1987.
- [6] J. E. Zch, "Efficiency robustness of generalized M-estimates for autoregression and their use in determining outliers," Ph.D. dissertation, University of Washington, Seattle, WA, USA, 1979.