

Received January 2, 2022, accepted January 16, 2022, date of publication January 26, 2022, date of current version February 8, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3146724

Bayesian Estimation of Multivariate Pure Moving Average Processes

MOHAMMED ALBASSAM¹, EMAD E. A. SOLIMAN, AND SHERIF S. ALI

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Mohammed Albassam (malbassam@kau.edu.sa)

This work was supported by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant 71-130-1437.

ABSTRACT The multivariate estimation problems arise if the observations are available for several related variables of interest. The multivariate time series may be found in many fields of application such as economics, meteorology and utilities. The current study has three main objectives. The first one is to develop an approximate convenient Bayesian methodology to estimate the parameters of multivariate moving average processes. The second objective is to investigate the numerical efficiency of the proposed technique in solving the estimation problems by conducting a wide simulation study. The last objective is to study the sensitivity of the numerical efficiency with respect to the parameters values and time series length. The results show that the proposed technique succeeded in estimating the parameters of the multivariate moving average processes. The results are not sensitive to the changes in parameter values or in time series length.

INDEX TERMS Multivariate time series, multivariate (vector) moving average processes, Bayesian estimation, matrix normal-Wishart prior, Jeffreys' prior.

I. INTRODUCTION

The multivariate (vector) time series may be found in many fields of application such as economics, business, meteorology, hydrology and utilities. In economics, for example, one may record quarterly money supply $y(t,1)$, real interest rate $y(t,2)$ and gross national product $y(t,3)$. These variables are modeled, estimated, and forecasted jointly using a multivariate model in order to have an insight into the dynamic interrelationship among them and increase the precision of the estimates and forecasts (See [1]).

Usually the Bayesian and non-Bayesian approaches of univariate and multivariate time series are based on a class of parametric models such as moving average (MA) models. Model estimation and forecasting are two main phases in time series analysis. Although they are closely connected, they are usually treated as two separate steps. The non-Bayesian literature on univariate and multivariate of time series is vast and can be found in [1]–[6].

One may trace three different Bayesian approaches to analyze univariate time series processes. The first approach is to use the numerical integration. [7] presented this approach to implement the identification, estimation, and forecasting phases of autoregressive moving average (ARMA) models

with low orders. However, the use of numerical integration is difficult and time consuming especially in the multi-parameter's cases. The second approach is the use of analytical approximations in order to have standard posterior distributions. Several approximations have been developed to solve the estimation and forecasting problems such as [8]–[10]. The last approximation has been extended to the case of seasonal models by [11], [12]. Finally, the third approach is to use sampling based methods which include: the Gibbs sampler technique ([13]), [14]–[17], data augmentation algorithm ([18], [19]) and the importance sampling algorithm ([20]).

Regarding the multivariate version, the estimation and forecasting, from non-Bayesian viewpoint, of vector moving average processes, denoted by VMA, have been extensively studied and investigated. [21] gave methods to estimate the parameters of pure autoregressive and pure moving average processes. [22] presented a practical iterative procedure to estimate the parameters of mixed VARMA processes. [23] extended the Box and Jenkins method [24] to derive the exact likelihood function of pure vector moving average processes. [25] proposed three different methods to compute the exact likelihood function of vector moving average processes. However, it seems that the non-Bayesian literature on the estimation problems of multivariate processes traditionally focused on maximum likelihood methods because of

The associate editor coordinating the review of this manuscript and approving it for publication was Alessandra Bertoldo.

their desirable properties. For more about these methods, the reader is referred to [1], [26]–[29]. However, there have been extensive investigations in order to ease the maximum likelihood routine and make it faster. For instance, [30] developed an efficient numerical expression for the likelihood function of stationary and partially non-stationary autoregressive moving average processes. Another development was done by [31] who proposed a new procedure for the exact maximum likelihood estimation of mixed VARMA models. [32] proposed a consistent and fast iterative procedure to estimate the parameters of VARMA processes. Many other investigations have been done to simplify the computation process of VARMA models such as [33], [34].

On the other hand, the Bayesian analysis of multivariate time series is being developed and the Bayesian literature devoted to the analysis of VARMA models is sparse. [35] has introduced approximate techniques to estimate the parameters of VARMA process and predict their future values. [36] used sampling-based inference to analyze VARMA models. [37] developed a Gibbs sampler for the basic VARMA models. However, for well understood reasons, one may say that an analytic Bayesian procedure to estimate the parameters of vector moving average process and predict their future values have not been explored yet.

The main objective of this study is to develop an approximate convenient Bayesian methodology to estimate the parameters of multivariate moving average, denoted by $VMA_k(q)$, processes. Using n vectors of observations, the joint posterior probability density of the model parameters is developed in an approximate convenient form using a matrix normal-Wishart prior density (or Jeffreys' vague prior). Then, one may use this approximate joint posterior probability density to develop point estimates or construct HPD (Highest Posterior Density) regions for the parameters. A wide simulation study is conducted, using the modern specialized SCA package, in order to examine the numerical efficiency of the proposed Bayesian estimation procedure.

The study is designed as follows: Section 2 defines the multivariate (vector) pure moving average ($VMA_k(q)$) processes in scalar and matrix notations. Section 3 constructs a convenient approximation for the likelihood function of the $VMA_k(q)$ processes. In addition, it proposes appropriate informative and non-informative priors. Section 4 derives an approximate joint posterior probability density of the model parameters in a standard form, namely matrix t distribution. In addition, it explains how to construct point and interval estimates for the parameters. Finally, section 5 is devoted to examine and assess the numerical effectiveness of the proposed Bayesian estimation procedure in solving the estimation problems of some bivariate pure moving average processes.

II. MULTIVARIATE MOVING AVERAGE PROCESSES

Let $\{t\}$ be a sequence of integers, $q \in \{1, 2, \dots\}$, $k \in \{2, 3, \dots\}$, θ_i ($i = 1, 2, \dots, q$) are $k \times k$ unknown matrices of real constants, $\{y(t)\}$ is a sequence of $k \times 1$ real observable random

vectors and $\{\varepsilon(t)\}$ is a sequence of $k \times 1$ independent and normally distributed unobservable random vectors with zero mean and a $k \times k$ unknown precision matrix. Then the Multivariate (vector) moving average process of orders q , denoted by $VMA_k(q)$, is defined as

$$y(t) = \Theta_q(B)\varepsilon(t) \tag{2.1}$$

where

$$y(t) = [y(t, 1) \quad y(t, 2) \quad \dots \quad y(t, k)]',$$

$$\theta_q(B) = I_k - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

And

$$\varepsilon(t) = [\varepsilon(t, 1) \quad \varepsilon(t, 2) \quad \dots \quad \varepsilon(t, k)]'$$

I_k is the identity matrix of order k and B is the usual backward shift operator. The $k \times k$ matrix polynomial $\theta_q(B)$, of degree q in the backshift operator B , is known as the moving average operator of order q . The process $y(t)$ is invertible if all the roots of determinantal equation $|\theta_q(B)| = 0$ lie outside the unite circle.

Consider the special case $VMA_2(1)$, which is called bivariate moving average process of order one, with moving average coefficients

$$\theta = \theta_1 = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}$$

Then the model (2.1) can be written as

$$y(t) = (I_2 - \theta B) \varepsilon(t) \tag{2.2}$$

where

$$y(t) = [y(t, 1) \quad y(t, 2)]',$$

$$\varepsilon(t) = [\varepsilon(t, 1) \quad \varepsilon(t, 2)]'$$

And

$$I_2 - \theta B = \begin{bmatrix} 1 - \theta_{11}B & -\theta_{12}B \\ -\theta_{21}B & 1 - \theta_{22}B \end{bmatrix}$$

Thus, one may write the observation $y(t)$ of the $VMA_2(1)$ process as

$$y(t, 1) = -\theta_{11}\varepsilon(t - 1, 1) - \theta_{12}\varepsilon(t - 1, 2) + \varepsilon(t, 1)$$

$$y(t, 2) = -\theta_{21}\varepsilon(t - 1, 1) - \theta_{22}\varepsilon(t - 1, 2) + \varepsilon(t, 2) \tag{2.3}$$

The model (2.2) can be written compactly for n observations as

$$y(t) = \varepsilon(t) - \theta\varepsilon(t - 1), \quad t = 1, 2, \dots, n \tag{2.4}$$

where $\varepsilon(t - 1) = \begin{bmatrix} \varepsilon(t - 1, 1) \\ \varepsilon(t - 1, 2) \end{bmatrix}$

Here we consider $y(t, 1)$ and $y(t, 2)$ as dependent variables, while $\varepsilon(t-1, 1)$ and $\varepsilon(t-1, 2)$ are considered as input or independent variables.

In general, one can write the $VMA_k(q)$ process as

$$y(t) = \varepsilon(t) - \theta_1\varepsilon(t - 1) - \theta_2\varepsilon(t - 2) - \dots - \theta_q\varepsilon(t - q),$$

$$t = 1, 2, \dots, n \tag{2.5}$$

The model (2.5) can be written in more compact expression as

$$Y = X\Gamma + U \tag{2.6}$$

where Y is a matrix of order $n \times k$ with ij^{th} element equals $y(i,j)$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$. That is

$$Y = Y_{n \times k} = [y(1) \quad y(2) \quad \dots \quad y(n)]'$$

The matrix X of order $n \times kq$ is defined by

$$X = \begin{bmatrix} -\varepsilon'(0) & -\varepsilon'(-1) & \dots & -\varepsilon'(1-q) \\ -\varepsilon'(1) & -\varepsilon'(0) & \dots & -\varepsilon'(2-q) \\ \vdots & \vdots & \vdots & \vdots \\ -\varepsilon'(n-1) & -\varepsilon'(n-2) & \dots & -\varepsilon'(n-q) \end{bmatrix}$$

Furthermore, Γ is the $kq \times k$ matrix of coefficients defined by

$$\Gamma = \begin{bmatrix} \theta'_1 \\ \theta'_2 \\ \dots \\ \theta'_q \end{bmatrix}$$

where

$$\theta_i = \begin{bmatrix} \theta_{i,11} & \theta_{i,12} & \dots & \theta_{i,1k} \\ \theta_{i,21} & \theta_{i,22} & \dots & \theta_{i,2k} \\ \vdots & \vdots & \vdots & \vdots \\ \theta_{i,k1} & \theta_{i,k2} & \dots & \theta_{i,kk} \end{bmatrix}, \quad i = 1, 2, \dots, q$$

III. AN APPROXIMATE LIKELIHOOD FUNCTION OF MULTIVARIATE MOVING AVERAGE PROCESSES

The class of models (2.6) represents the general class of vector moving average models of order q , denoted by $VMA_K(q)$. It is very useful in modeling time series data arise in many areas of application such as economics, medicine, ecology and chemistry. The likelihood function of the parameters Γ and T is

$$L(\Gamma, T | S_n) \propto |T|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left\{\sum_{t=1}^n \varepsilon(t) \varepsilon(t)'\right\}\right) \tag{3.1}$$

where $\Gamma \in R^{kq \times k}$, $T > 0$, and

$$\varepsilon'(t) = y'(t) - X'(t-1)\Gamma, \quad t = 1, 2, \dots, n \tag{3.2}$$

where

$$X'(t-1) = [-\varepsilon'(t-1) \quad -\varepsilon'(t-2) \quad \dots \quad -\varepsilon'(t-q)]$$

The expression (3.2) is a recurrence relation for the residuals and the m^{th} component of the residual $\varepsilon(t)$ can be written as

$$\varepsilon(t, m) = y(t, m) - \sum_{i=1}^q \sum_{j=1}^k \theta_{i,mj} \varepsilon(t-i, j), \quad t = 1, 2, \dots, n; \quad m = 1, 2, \dots, k \tag{3.3}$$

The recurrence relation (3.3) causes the main problem in developing the exact analysis of multivariate $VMA_K(q)$ processes. However, this recurrence may be used to evaluate the

residuals recursively if θ_i and the initial values of the residuals are known. The proposed Bayesian approximation is based on replacing the exact residuals by their least squares estimates and assuming that the initial residuals equal their means, namely zero. Thus, we estimate the residuals recursively by

$$\hat{\varepsilon}(t, m) = y(t, m) + \sum_{i=1}^q \sum_{j=1}^k \hat{\theta}_{i,mj} \hat{\varepsilon}(t-i, j), \quad t = 1, 2, \dots, n; \quad m = 1, 2, \dots, k \tag{3.4}$$

where $\hat{\theta}_{i,mj}$ are the nonlinear least squares estimates of parameters $\theta_{i,mj}$. Using the estimates of the residuals, it is possible to write the likelihood function approximately as

$$L^*(\Gamma, T | S_n) \propto |T|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left\{\sum_{t=1}^n [y(t) - \Gamma' \hat{X}(t-1)] [y(t) - \Gamma' \hat{X}(t-1)]' T\right\}\right) \tag{3.5}$$

where $\hat{X}(t-1)$ is the same as $X(t-1)$ but using the estimated residuals instead of the exact ones. A convenient choice of the joint prior density of the parameters Γ and T is the following matrix normal-Wishart distribution

$$\xi(\Gamma, T) = \xi_1(\Gamma | T) \xi_2(T) \tag{3.6}$$

where

$$\xi_1(\Gamma | T) \propto |T|^{\frac{kq}{2}} \exp\left(-\frac{1}{2} \text{tr}[\Gamma - D]' W [\Gamma - D] T\right)$$

and

$$\xi_2(T) \propto |T|^{\frac{a-(k+1)}{2}} \exp\left(-\frac{1}{2} \text{tr} \Psi T\right)$$

The hyper-parameters a is a scalar, $D \in R^{kq \times k}$, W is a $kq \times kq$ positive definite matrix and Ψ is a $k \times k$ positive definite matrix. If there is little information about the parameters, a priori, it is possible to use Jeffreys' vague prior

$$\xi(\Gamma, T) \propto |T|^{-\frac{(k+1)}{2}}, \quad \Gamma \in R^{kq \times k}, T > 0 \tag{3.7}$$

IV. THE POSTERIOR ANALYSIS OF THE MULTIVARIATE MOVING AVERAGE PROCESSES

In developing the proposed Bayesian estimation technique for $MA_K(q)$ models, we will assume that the order q is known. The posterior density of the model parameters Γ and T is the Bayesian tool to estimate the unknown parameters. The posterior density $\xi(\Gamma, T | S_n)$ is the conditional density of the parameters given the time series data

Theorem 1: If the approximate likelihood function (3.5) is combined with the joint prior density (3.6), the marginal posterior distribution of Γ is a matrix t with parameters:

$(A^{-1}B, A^{-1}, C-BA^{-1}B, \nu)$ where

$$A = W + \sum_{t=1}^n \hat{X}(t-1) \hat{X}'(t-1),$$

$$B = WD + \sum_{t=1}^n \hat{X}(t-1) y'(t),$$

$$C = D'WD + \Psi + \sum_{t=1}^n y(t)y'(t)$$

and $v = n + a - k + 1$

In addition, the marginal posterior distribution of T is a Wishart with parameters $(n + a, C-BA^{-1}B)$.

Corollary 2: If the approximate conditional likelihood function (3.5) is combined with the Jeffreys' vague prior (3.7), the marginal posterior distribution of Γ is a matrix t with parameters:

$(A^{-1}B, A^{-1}, C-B'A^{-1}B, v)$. However, the quantities A, B, C and v will be modified by letting $W \rightarrow 0$ ($kq \times kq$), $a \rightarrow -kq$ and $\Psi \rightarrow 0$ ($k \times k$).

Theorem (1) solves the estimation problem of vector moving average processes. In order to see this let γ_i be the i^{th} column of the coefficients matrix Γ and write $\mathbf{\Gamma} = [\gamma_1' \gamma_2' \dots \gamma_{kq}']$; i.e. we write $\mathbf{\Gamma}$ as a row vector of dimension k_2q . Partition $A^{-1}B$ similar to $\mathbf{\Gamma}$, then the posterior expectation and variance of the parameters matrix $\mathbf{\Gamma}$ are

$$E(\mathbf{\Gamma}|S_n) = A^{-1}B$$

$$COV(\mathbf{\Gamma}|S_n) = [(v - 2)A]^{-1} \otimes [C - B'A^{-1}B]$$

where \otimes denotes the usual Kronecker direct product. The variance covariance matrix is of dimension k^2q . The elements of the principle diagonal represent the variances of the elements of the coefficients matrix $\mathbf{\Gamma}$ as partitioned above. The matrix t approximation can be analytically used to make exact statistical inference for a parameter point $\mathbf{\Gamma}_0$ for the complete set Γ using F distribution if $k= 2$, see [38] (p. 451). For $k \geq 3$ approximate statistical inferences can be done using χ^2 distribution, see [38] (p. 452).

With regard to a point estimate of $\mathbf{\Gamma}$, one may use the posterior expectation $A^{-1}B$. In addition, the matrix t approximation can be used to construct an exact H.P.D. region for a specific row or a specific column using the usual multivariate t distribution. Furthermore, an exact HPD interval for any element of the matrix $\mathbf{\Gamma}$ can be constructed using univariate student t distribution. Finally, we can make approximate statistical inferences about parameters belonging to a certain block sub matrix of $\mathbf{\Gamma}$ using χ^2 distribution, see Box and Tiao (1973, page 453). To estimate the precision matrix T , one may use the posterior precision $(n + a)(C - B'A^{-1}B)^{-1}$.

Moreover, to estimate the covariance matrix $V = T^{-1}$, one may use the inverse of the posterior precision, i.e. $\frac{(C-B'A^{-1}B)}{(n+a)}$.

V. RESULTS

One of the main objectives of this study is to investigate the effectiveness of the proposed Bayesian estimation methodology. In order to achieve this objective, three simulation studies have been conducted. The proposed methodology is employed to estimate the parameters of $MA_2(1)$ processes with various parameters values. All computations were performed on a PC using SCA package.

The simulation process has the following general design: first, a time series is generated from a given bivariate moving average model of order one with certain parameters. Second,

the generated data are used to evaluate the posterior densities of the coefficients and the covariance matrix of the model. Third, performance criteria are calculated for the posterior density. Fourth, 500 replications of the above three steps are done. Finally, the output is summarized in tables.

Generally, the simulation process begins by generating 500 data sets of bivariate normal variates, each of size 500, to represent the noise $\varepsilon(t)$, which is assumed to follow $N_2(0, V)$. These data sets are then used recursively to generate 500 realizations, each of size 500, from a certain bivariate moving average process of order one with certain parameters. The initial values of the errors $\varepsilon(t) = [\varepsilon(t, 1)\varepsilon(t, 2)]$ are considered to equal their unconditional mean, namely zero. The first 200 pairs of observations are deleted in order to remove the initialization effect and hence we get 500 bivariate time series each of length 300. From the 300 observations, a bivariate time series of the desired length is used to estimate the posterior densities of the coefficients' matrix θ and the covariance matrix of the noise term using the proposed Bayesian methodology. In our simulation studies, the time series lengths are taken to be 30, 50, 100, 150, 200, and 300. Each simulation study corresponds to specific coefficients and for all of them the variance-covariance matrix of the noise is fixed at

$$V(\varepsilon) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \tag{5.1}$$

The values of the coefficients are selected to represent different positions in the invertibility domain of the $MA_2(1)$ model. It might be important here to emphasis that the Jefferys' non-informative prior is used to conduct all the simulation studies.

Our main concern is to study the numerical efficiency of the proposed Bayesian estimation methodology by calculating three goodness criteria, namely P^* , MAD and MAPE. The measure P^* checks the goodness of interval estimates calculated from a specified posterior density of the model's coefficients. Defining 95% highest posterior density (HPD) region as the shortest interval having probability 0.95 centered at the mean of the posterior density, the percentage P_{lm}^* of time series for which the HPD region of the posterior density contains the true value of the coefficient is defined as

$$P_{lm}^* = (n_{lm}^*/500) * 100, \quad l, m = 1, 2 \tag{5.2}$$

where n_{lm}^* is the number of time series where the HPD region includes the true value of θ_{lm} . P_{lm}^* is evaluated such that the higher the value P_{lm}^* , the better the performance of the posterior density in estimation. It should be noted that according to P_{lm}^* a certain coefficient may be ruled as belonging to the HPD region or not. However, P^* does not account for the distance of this coefficient from the center of the region or its boundaries. The MAD stands for the mean absolute deviation of the true coefficient from the mean of its marginal posterior density and is defined by

$$MAD_{lm} = \sum_{j=1}^{500} |\theta_{lm} - E(\theta_{j,lm})|/500, \quad l, m = 1, 2 \tag{5.3}$$

TABLE 1. Results for θ_{11} , true value is 0.9.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	0.9195	0.2862	0.0833	0.7258	0.9109	1.1027	1.9918	0.2252	0.2503	95.4
50	0.9184	0.2110	0.2686	0.7829	0.9130	1.0553	1.5944	0.1656	0.1840	95.0
100	0.9121	0.1470	0.4722	0.8083	0.9164	1.0096	1.3300	0.1179	0.1310	94.6
150	0.9077	0.1219	0.5804	0.8224	0.9105	0.9903	1.2985	0.0980	0.1089	94.2
200	0.9038	0.1058	0.6184	0.8265	0.9040	0.9841	1.2214	0.0863	0.0958	95.6
300	0.9007	0.0826	0.6491	0.8489	0.9005	0.9559	1.1535	0.0652	0.0724	94.4

TABLE 2. Results for θ_{12} , true value is -0.2.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	-0.1937	0.4143	-1.5918	-0.4425	-0.1899	0.0742	1.0485	0.3239	1.6195	95.0
50	-0.2009	0.3035	-1.1634	-0.3972	-0.1952	-0.0127	0.6725	0.2399	1.1994	94.4
100	-0.2094	0.1996	-0.8190	-0.3489	-0.2029	-0.0766	0.4247	0.1593	0.7967	96.0
150	-0.2039	0.1750	-0.8312	-0.3206	-0.1977	-0.0825	0.2264	0.1412	0.7062	93.6
200	-0.2020	0.1503	-0.7175	-0.3110	-0.2019	-0.1012	0.2456	0.1215	0.6076	94.8
300	-0.1973	0.1194	-0.5985	-0.2771	-0.1954	-0.1123	0.1673	0.0957	0.4784	95.6

TABLE 3. Results for θ_{21} , true value is 1.1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	1.1249	0.2115	0.4223	0.9812	1.1273	1.2503	1.8928	0.1686	0.1533	93.6
50	1.1114	0.1552	0.6249	1.0083	1.1083	1.2153	1.5913	0.1233	0.1121	94.4
100	1.1084	0.1049	0.8121	1.0382	1.1124	1.1772	1.4507	0.0844	0.0767	94.2
150	1.1042	0.0875	0.7970	1.0453	1.1064	1.1614	1.3735	0.0699	0.0636	93.2
200	1.1009	0.0747	0.8679	1.0504	1.0986	1.1512	1.2995	0.0597	0.0543	94.4
300	1.1002	0.0589	0.9303	1.0635	1.0955	1.1367	1.2778	0.0464	0.0421	94.6

TABLE 4. Results for θ_{22} , true value is -0.9.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	-0.9158	0.3021	-2.0609	-1.1181	-0.9212	-0.7258	-0.0709	0.2363	0.2626	93.8
50	-0.9060	0.2240	-1.5938	-1.0445	-0.9001	-0.7671	-0.1510	0.1762	0.1957	93.8
100	-0.9121	0.1499	-1.4233	-1.0114	-0.9220	-0.8109	-0.4123	0.1209	0.1344	94.6
150	-0.9064	0.1249	-1.2389	-0.9845	-0.9028	-0.8213	-0.5383	0.0985	0.1094	94.4
200	-0.9022	0.1064	-1.1620	-0.9774	-0.9031	-0.8312	-0.5914	0.0858	0.0953	95.0
300	-0.8996	0.0864	-1.1876	-0.9543	-0.9011	-0.8456	-0.6286	0.0673	0.0747	94.6

where θ_{lm} and $E(\theta_{j,lm})$ are the lm^{th} component of the coefficients matrix and the mean of its marginal posterior density of the j^{th} simulated series, respectively. The MAPE stands for the mean absolute percentage deviation of the true coefficient from the mean of its marginal posterior density and is defined as

$$MAPE_{lm} = \left(\sum_{j=1}^{500} \left| \frac{\theta_{lm} - E(\theta_{j,lm})}{\theta_{lm}} \right| / 500 \right), \quad l, m = 1, 2 \tag{5.4}$$

where θ_{lm} and $E(\theta_{j,lm})$ are defined as above.

The numerical efficiency of the proposed estimation procedure will be examined with respect to the time series length (n) as well as the parameters of the selected model.

In order to study the numerical efficiency of the estimation process of the covariance matrix (5.1), the MAD and the MAPE will be computed for the elements of the covariance matrix using the forms

$$MAD_{lm} = \sum_{j=1}^{500} |V_{lm} - E(V_{j,lm})| / 500, \quad l, m = 1, 2 \tag{5.5}$$

$$MAPE_{lm} = \left(\sum_{j=1}^{500} \left| \frac{V_{lm} - E(V_{j,lm})}{V_{lm}} \right| / 500 \right), \quad l, m = 1, 2 \tag{5.6}$$

where V_{lm} is the lm^{th} component of the covariance matrix and $E(V_{j,lm})$ is the lm^{th} component of the

TABLE 5. Results for V_{11} , true value is 2.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	2.0268	0.5441	0.8477	1.6136	1.9756	2.3609	4.0065	0.4322	0.2161
50	2.0037	0.4225	0.9727	1.7271	1.9607	2.2774	3.3701	0.3346	0.1673
100	2.0102	0.2860	1.2995	1.8124	1.9968	2.1910	2.8843	0.2268	0.1134
150	2.0071	0.2295	1.3423	1.8414	2.0144	2.1547	2.7487	0.1850	0.0925
200	2.0028	0.1940	1.4546	1.8694	1.9995	2.1242	2.5945	0.1524	0.0762
300	2.0030	0.1578	1.5069	1.8954	2.0067	2.1096	2.6396	0.1258	0.0629

TABLE 6. Results for V_{12} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	1.0242	0.3408	0.2454	0.7771	1.0052	1.2536	2.2891	0.2740	0.2740
50	1.0071	0.2644	0.3388	0.8224	0.9873	1.1775	1.8868	0.2105	0.2105
100	1.0083	0.1750	0.4582	0.8960	0.9991	1.1281	1.5934	0.1380	0.1380
150	1.0058	0.1375	0.5613	0.9090	1.0105	1.0933	1.4504	0.1075	0.1075
200	1.0033	0.1188	0.6081	0.9249	1.0020	1.0772	1.4160	0.0939	0.0939
300	1.0032	0.0953	0.6594	0.9414	1.0028	1.0609	1.2713	0.0754	0.0754

TABLE 7. Results for V_{21} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	1.0242	0.3408	0.2454	0.7771	1.0052	1.2536	2.2891	0.2740	0.2740
50	1.0071	0.2644	0.3388	0.8224	0.9873	1.1775	1.8868	0.2105	0.2105
100	1.0083	0.1750	0.4582	0.8960	0.9991	1.1281	1.5934	0.1380	0.1380
150	1.0058	0.1375	0.5613	0.9090	1.0105	1.0933	1.4504	0.1075	0.1075
200	1.0033	0.1188	0.6081	0.9249	1.0020	1.0772	1.4160	0.0939	0.0939
300	1.0032	0.0953	0.6594	0.9414	1.0028	1.0609	1.2713	0.0754	0.0754

TABLE 8. Results for V_{22} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	1.0256	0.2922	0.3558	0.8178	1.0141	1.2000	2.3188	0.2299	0.2299
50	1.0090	0.2159	0.5277	0.8642	0.9830	1.1315	1.7757	0.1705	0.1705
100	1.0060	0.1432	0.5755	0.8997	1.0057	1.1055	1.5509	0.1144	0.1144
150	1.0039	0.1126	0.6386	0.9301	1.0038	1.0738	1.4220	0.0885	0.0885
200	1.0027	0.0974	0.7037	0.9367	1.0034	1.0680	1.3350	0.0768	0.0768
300	1.0020	0.0804	0.7684	0.9482	1.0017	1.0495	1.2552	0.0628	0.0628

estimated covariance matrix of the j^{th} time series defined above.

Simulation I, as an illustration, begins with the generation of 500 data sets of bivariate normal variates, each with 500 observations to represent $\varepsilon(t,1)$ and $\varepsilon(t,2)$ respectively. These data sets are then used to generate pairs of 500 realizations, each of size 300, from $MA_2(1)$ process with

$$\Theta = \begin{bmatrix} 0.9 & -0.2 \\ 1.1 & -0.9 \end{bmatrix}$$

Assuming the starting values of the errors are zeros and Jefferys' prior is employed, the second step is to carry out all computations required to estimate the posterior density of

the coefficients matrix Θ of each of the 500 realizations and compute the P^* , MAD and MAPE values. Such computations are done for a specific time series length using the first n observations of each generated set. This second step is repeated for each chosen time series length. The results of Simulation I are summarized in tables (1) up to (8). Table (1) is devoted to the results of the coefficient θ_{11} . It consists of six rows corresponding to the selected time series lengths and ten columns corresponding to the computed measures. Tables (2) up to (4) display the results of θ_{12} , θ_{21} and θ_{22} respectively and are designed similarly. Moreover, the results of the estimation of the covariance matrix are displayed similarly in tables (5) up to (8).

TABLE 9. Results for θ_{11} , true value is 0.5.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	0.5224	0.2919	-0.1759	0.3242	0.5140	0.6935	1.7690	0.2285	0.4570	94.0
50	0.5128	0.2137	-0.0857	0.3749	0.5088	0.6568	1.0917	0.1672	0.3344	94.2
100	0.5099	0.1456	0.0806	0.4143	0.5094	0.6071	0.8972	0.1166	0.2331	95.4
150	0.5066	0.1213	0.1578	0.4247	0.5080	0.5857	0.8737	0.0975	0.1950	94.6
200	0.5028	0.1056	0.1984	0.4263	0.5060	0.5792	0.8078	0.0858	0.1716	95.6
300	0.5000	0.0826	0.2448	0.4494	0.5017	0.5556	0.7393	0.0650	0.1301	93.8

TABLE 10. Results for θ_{12} , true value is -0.4.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	-0.4189	0.4077	-1.8978	-0.6995	-0.3958	-0.1819	0.7807	0.3186	0.7966	94.2
50	-0.4070	0.3002	-1.2214	-0.6204	-0.4036	-0.2069	0.6052	0.2400	0.5999	94.8
100	-0.4120	0.1987	-0.9884	-0.5456	-0.4094	-0.2696	0.0921	0.1591	0.3977	96.8
150	-0.4038	0.1737	-0.9840	-0.5225	-0.4040	-0.2815	0.0557	0.1393	0.3482	94.6
200	-0.4004	0.1507	-0.8902	-0.5073	-0.3998	-0.2921	0.0374	0.1214	0.3034	94.8
300	-0.3965	0.1199	-0.7668	-0.4830	-0.4015	-0.3159	-0.0357	0.0961	0.2404	95.0

TABLE 11. Results for θ_{21} , true values is -0.3.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	-0.3051	0.2140	-0.8253	-0.4463	-0.3060	-0.1806	0.3623	0.1677	0.5589	91.8
50	-0.3078	0.1548	-0.8201	-0.4073	-0.3091	-0.2013	0.1986	0.1229	0.4095	94.4
100	-0.2993	0.1018	-0.5839	-0.3642	-0.2964	-0.2355	-0.0040	0.0806	0.2686	95.2
150	-0.2993	0.0865	-0.6097	-0.3530	-0.2972	-0.2412	-0.0799	0.0690	0.2299	92.4
200	-0.3014	0.0732	-0.5438	-0.3485	-0.3032	-0.2493	-0.1088	0.0585	0.1950	94.8
300	-0.3014	0.0576	-0.4734	-0.3360	-0.3056	-0.2651	-0.1308	0.0453	0.1509	94.8

TABLE 12. Results for θ_{22} , true value is 0.2.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	0.2074	0.3057	-0.7272	0.0108	0.1871	0.3994	1.1991	0.2416	1.2082	93.4
50	0.2080	0.2273	-0.4576	0.0472	0.2039	0.3546	1.1291	0.1806	0.9029	92.2
100	0.1929	0.1491	-0.2326	0.0941	0.1895	0.2916	0.6574	0.1211	0.6053	94.8
150	0.1956	0.1248	-0.1980	0.1166	0.1942	0.2794	0.5325	0.0993	0.4967	93.6
200	0.1995	0.1064	-0.1440	0.1293	0.1983	0.2752	0.5112	0.0861	0.4304	94.8
300	0.2012	0.0856	-0.1057	0.1493	0.2010	0.2568	0.4451	0.0672	0.3358	94.4

TABLE 13. Results for V_{11} , true value is 2.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	2.0072	0.5398	0.8345	1.6403	1.9413	2.3528	4.2102	0.4277	0.2139
50	2.0017	0.4159	0.8881	1.7135	1.9760	2.2715	3.2704	0.3303	0.1651
100	2.0077	0.2807	1.2831	1.8147	1.9834	2.1849	2.8589	0.2205	0.1102
150	2.0033	0.2302	1.3251	1.8372	2.0084	2.1558	2.7406	0.1857	0.0929
200	1.9991	0.1943	1.4452	1.8634	1.9969	2.1231	2.6113	0.1533	0.0766
300	2.0007	0.1582	1.4946	1.8943	2.0061	2.1080	2.6454	0.1252	0.0626

TABLE 14. Results for V_{12} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	0.9988	0.3251	0.2218	0.7654	0.9707	1.1936	2.2858	0.2604	0.2604
50	0.9986	0.2502	0.3774	0.8248	0.9766	1.1696	1.9541	0.2021	0.2021
100	1.0043	0.1670	0.4387	0.8997	1.0026	1.1052	1.5199	0.1297	0.1297
150	1.0019	0.1363	0.5926	0.9138	1.0020	1.0885	1.4077	0.1071	0.1071
200	1.0000	0.1180	0.6042	0.9209	0.9997	1.0791	1.3395	0.0932	0.0932
300	1.0006	0.0951	0.6403	0.9399	1.0018	1.0604	1.2668	0.0752	0.0752

TABLE 15. Results for V_{21} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	0.9988	0.3251	0.2218	0.7654	0.9707	1.1936	2.2858	0.2604	0.2604
50	0.9986	0.2502	0.3774	0.8248	0.9766	1.1696	1.9541	0.2021	0.2021
100	1.0043	0.1670	0.4387	0.8997	1.0026	1.1052	1.5199	0.1297	0.1297
150	1.0019	0.1363	0.5926	0.9138	1.0020	1.0885	1.4077	0.1071	0.1071
200	1.0000	0.1180	0.6042	0.9209	0.9997	1.0791	1.3395	0.0932	0.0932
300	1.0006	0.0951	0.6403	0.9399	1.0018	1.0604	1.2668	0.0752	0.0752

TABLE 16. Results for V_{22} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	1.0000	0.2660	0.4198	0.8078	0.9888	1.1705	2.3394	0.2083	0.2083
50	0.9986	0.1993	0.5527	0.8560	0.9934	1.1286	1.8763	0.1577	0.1577
100	1.0017	0.1374	0.6164	0.9091	1.0012	1.0878	1.6206	0.1075	0.1075
150	1.0000	0.1110	0.6517	0.9281	0.9966	1.0718	1.4147	0.0878	0.0878
200	0.9998	0.0969	0.6988	0.9329	1.0024	1.0639	1.3225	0.0762	0.0762
300	0.9997	0.0803	0.7322	0.9491	1.0010	1.0504	1.2607	0.0626	0.0626

TABLE 17. Results for θ_{11} , true value is 0.2.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	0.2325	0.2835	-0.5033	0.0558	0.2161	0.4049	1.4889	0.2207	1.1034	94.4
50	0.2167	0.2085	-0.4238	0.0724	0.2111	0.3503	0.8540	0.1636	0.8179	96.0
100	0.2126	0.1426	-0.1840	0.1183	0.2110	0.3073	0.6051	0.1145	0.5725	95.6
150	0.2088	0.1191	-0.1109	0.1289	0.2092	0.2886	0.5938	0.0956	0.4778	94.6
200	0.2049	0.1044	-0.0959	0.1293	0.2054	0.2829	0.5202	0.0849	0.4246	95.6
300	0.2016	0.0817	-0.0566	0.1522	0.2034	0.2569	0.4436	0.0643	0.3216	94.0

TABLE 18. Results for θ_{12} , true value is 0.3.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	0.3120	0.4062	-1.1629	0.0603	0.3010	0.5749	1.6102	0.3201	1.0671	94.8
50	0.3078	0.3022	-0.5011	0.0905	0.3118	0.4991	1.3206	0.2393	0.7977	94.2
100	0.2916	0.1970	-0.2689	0.1575	0.2932	0.4387	0.8482	0.1584	0.5282	96.4
150	0.2968	0.1726	-0.3225	0.1836	0.2981	0.4147	0.7880	0.1380	0.4600	94.6
200	0.2990	0.1497	-0.2270	0.1959	0.2989	0.3986	0.7310	0.1206	0.4019	94.4
300	0.3120	0.4062	-1.1629	0.0603	0.3010	0.5749	1.6102	0.3201	1.0671	94.8

TABLE 19. Results for θ_{21} , true value is -0.6 .

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	-0.5790	0.2191	-1.1503	-0.7254	-0.5723	-0.4456	0.0418	0.1745	0.2908	92.2
50	-0.5878	0.1604	-1.0937	-0.6863	-0.5870	-0.4820	-0.0787	0.1271	0.2119	94.4
100	-0.5877	0.1035	-0.8903	-0.6540	-0.5849	-0.5227	-0.2821	0.0828	0.1380	95.0
150	-0.5911	0.0867	-0.9103	-0.6499	-0.5904	-0.5277	-0.3851	0.0697	0.1161	93.4
200	-0.5947	0.0734	-0.8402	-0.6414	-0.5980	-0.5405	-0.3869	0.0585	0.0975	95.6
300	-0.5966	0.0572	-0.7734	-0.6323	-0.6006	-0.5610	-0.4286	0.0448	0.0746	94.2

TABLE 20. Results for θ_{22} , true value is 1.1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE	P*
30	1.0875	0.3192	0.1393	0.8708	1.0766	1.3172	2.1670	0.2550	0.2318	90.6
50	1.0863	0.2384	0.4802	0.9178	1.0774	1.2477	2.0122	0.1912	0.1738	92.2
100	1.0774	0.1522	0.6065	0.9712	1.0738	1.1808	1.5696	0.1249	0.1135	94.6
150	1.0834	0.1277	0.6605	0.9972	1.0862	1.1667	1.4655	0.1031	0.0937	92.8
200	1.0890	0.1081	0.7311	1.0144	1.0929	1.1696	1.4128	0.0875	0.0795	93.6
300	1.0936	0.0865	0.8052	1.0376	1.0955	1.1501	1.3400	0.0688	0.0625	94.2

TABLE 21. Results for V_{11} , true value is 2.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	1.9840	0.5401	0.8400	1.5909	1.9423	2.3220	4.0371	0.4331	0.2165
50	1.9905	0.4163	1.0535	1.6959	1.9709	2.2473	3.2727	0.3337	0.1668
100	2.0049	0.2835	1.2959	1.8056	1.9840	2.1794	2.8615	0.2241	0.1120
150	2.0023	0.2322	1.3225	1.8355	2.0053	2.1665	2.7110	0.1877	0.0939
200	1.9978	0.1955	1.4597	1.8581	2.0008	2.1275	2.5830	0.1539	0.0769
300	1.9999	0.1587	1.5087	1.8897	2.0022	2.1038	2.6386	0.1262	0.0631

TABLE 22. Results for V_{12} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	0.9966	0.3375	0.2273	0.7609	0.9659	1.2089	2.4223	0.2657	0.2657
50	0.9966	0.2573	0.3816	0.8194	0.9725	1.1696	2.0912	0.2077	0.2077
100	1.0064	0.1733	0.4612	0.8835	1.0070	1.1137	1.5388	0.1358	0.1358
150	1.0040	0.1389	0.6142	0.9155	1.0044	1.0906	1.3977	0.1088	0.1088
200	1.0015	0.1193	0.6091	0.9223	1.0037	1.0777	1.3165	0.0944	0.0944
300	1.0018	0.0960	0.6662	0.9416	1.0015	1.0638	1.2540	0.0757	0.0757

Simulations II and III are designed in a similar manner but using different coefficients. They are

$$\Theta = \begin{bmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{bmatrix}$$

and

$$\Theta = \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix}$$

respectively. Their results are reported in tables (9) up to (24) respectively.

A. RESULTS OF SIMULATION I

The results of this simulation study are shown in tables (1) up to (8). The first four tables present the results of the estimation

of the elements of the coefficients matrix. Whereas, the last four tables present the results of the estimation of the elements of the covariance matrix.

Regarding table (1) one may notice the following comments concerning the coefficient θ_{11} : First, the values of the percentage P* fluctuate around 95% (its theoretical value). This means that the marginal posterior distribution of θ_{11} provide a good interval estimation. Second, the values of both the MAD and the MAPE are small at all time series lengths and tend to decrease as the time series length increases. This indicates that the posterior means of θ_{11} get closer to it as the time series length increases. Moreover, both the mean and the median of the posterior means tend to concentrate around the true value of the coefficient and get closer to it as the time series length increases. Finally, the standard

TABLE 23. Results for V_{21} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	0.9966	0.3375	0.2273	0.7609	0.9659	1.2089	2.4223	0.2657	0.2657
50	0.9966	0.2573	0.3816	0.8194	0.9725	1.1696	2.0912	0.2077	0.2077
100	1.0064	0.1733	0.4612	0.8835	1.0070	1.1137	1.5388	0.1358	0.1358
150	1.0040	0.1389	0.6142	0.9155	1.0044	1.0906	1.3977	0.1088	0.1088
200	1.0015	0.1193	0.6091	0.9223	1.0037	1.0777	1.3165	0.0944	0.0944
300	1.0018	0.0960	0.6662	0.9416	1.0015	1.0638	1.2540	0.0757	0.0757

TABLE 24. Results for V_{22} , true value is 1.

LENGTH	MEAN	STDEV	MIN	Q1	MEDIAN	Q2	MAX	MAD	MAPE
30	1.0375	0.3010	0.3753	0.8215	1.0158	1.2073	2.5118	0.2368	0.2368
50	1.0212	0.2157	0.4923	0.8691	1.0101	1.1434	1.9768	0.1696	0.1696
100	1.0160	0.1467	0.6131	0.9171	1.0114	1.1080	1.6668	0.1164	0.1164
150	1.0106	0.1162	0.6525	0.9322	1.0099	1.0804	1.4886	0.0912	0.0912
200	1.0080	0.0995	0.7037	0.9441	1.0049	1.0708	1.3974	0.0776	0.0776
300	1.0056	0.0823	0.7740	0.9516	1.0042	1.0580	1.2896	0.0643	0.0643

deviation of the posterior means tend to decrease as the time series length increases.

The conclusions obtained from tables (2) to (4) are similar.

The following four tables display the results of the estimation of the covariance matrix. The conclusions obtained from tables (5) to (8) are similar to those obtained from table (1).

Regarding tables (5) to (8), one may notice the following comments concerning the elements of the covariance matrix: First, the values of both the MAD and the MAPE for all elements are small for all time series lengths and tend to decrease as the time series length increases. This indicates that the posterior mean of the covariance matrix gets closer to the true one as time series length increases. Second, the means and the medians of the posterior means tend to concentrate around the true values of the elements and get closer to them as the time series length increases. Finally, the standard deviations of the posterior means tend to decrease as the time series length increases.

B. RESULTS OF SIMULATION II

The results of this simulation study are shown in tables (9) up to (16). The tables are designed the same way as the tables in the previous subsection and the conclusions obtained are similar.

C. RESULTS OF SIMULATION III

The results of this simulation study are shown in tables (17) up to (24). The tables are designed the same way as the tables in the previous subsection and the conclusions obtained are similar.

VI. CONCLUSION

The results of the three simulation studies displayed in the previous three sections show that the suggested estimation technique succeeded in estimating the parameters of the bivariate moving average process of order 1.

ACKNOWLEDGMENT

The authors would like to thank the editor and reviewers for their valuable suggestions to improve the quality of the paper, and also would like to thank the DSR at King Abdulaziz University for their technical and financial support.

REFERENCES

- [1] G. C. Tiao and G. E. Box, "Modeling multiple time series with applications," *J. Amer. Statist. Assoc.*, vol. 76, no. 376, pp. 802–816, 1981.
- [2] R. S. Tsay, *Multivariate Time Series Analysis: With R and Financial Applications*. Hoboken, NJ, USA: Wiley, 2014.
- [3] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: Wiley, 2016.
- [4] P. J. Brockwell and R. Davis, *Introduction to Time Series and Forecasting*, 3rd ed. Cham, Switzerland: Springer, 2016.
- [5] C. Chatfield, *The Analysis of Time Series: Theory and Practice, An Introduction With R*, 7th ed. London, U.K.: Chapman & Hall, 2019.
- [6] W. W. S. Wei, *Multivariate Time Series Analysis and Applications*. Hoboken, NJ, USA: Wiley, 2019.
- [7] J. F. Monahan, "Fully Bayesian analysis of ARMA time series models," *J. Econometrics*, vol. 21, no. 3, pp. 307–331, 1983.
- [8] P. Newbold, "Bayesian estimation of Box–Jenkins transfer function-noise models," *J. Roy. Stat. Soc., B, Methodol.*, vol. 35, no. 2, pp. 323–336, Jan. 1973.
- [9] A. Zellner and R. Reynolds, "Bayesian analysis of ARMA models," in *Proc. 16th Seminar Bayesian Inference Econometrics*, Jun. 1978, pp. 1–5.
- [10] L. Broemeling and S. Shaarawy, "Time series: A Bayesian analysis in time domain: Bayesian analysis of time series and dynamic models," in *Studies in Bayesian Analysis of Time Series and Dynamic Models*, J. spall, Eds. New York, NY, USA: Marcel Dekker, 1988, pp. 1–22.
- [11] S. M. Shaarawy and G. A. El-Shawadfy, "Approximate Bayesian densities of future observations for seasonal models," in *Proc. 22nd Annu. Conf. Statist., Comput. Sci. Operation Res.*, vol. 2, Giza, Egypt: Cairo Univ., 1987, pp. 160–178.
- [12] S. M. Shaarawy and M. Ismail, "Bayesian inference for seasonal ARMA models," *Egyptian Stat. J.*, vol. 31, pp. 323–336, 1987.
- [13] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984, doi: 10.1109/TPAMI.1984.4767596.
- [14] G. Barnett, R. Kohn, S. Sheather, and J. Wong, "Markov chain Monte Carlo estimation of autoregressive models with application to metal pollutant concentration in sludge," *Math. Comput. Model.*, vol. 22, nos. 10–12, pp. 7–13, Nov. 1995.

- [15] G. Barnett, R. Kohn, and S. Sheather, "Bayesian estimation of an autoregressive model using Markov chain Monte Carlo," *J. Econometrics*, vol. 74, no. 2, pp. 237–254, 1996, doi: [10.1016/0304-4076\(95\)01744-5](https://doi.org/10.1016/0304-4076(95)01744-5).
- [16] G. Barnett, R. Kohn, and S. Sheather, "Robust Bayesian estimation of autoregressive-moving average model," *J. Time Ser. Anal.*, vol. 18, no. 1, pp. 11–28, 1997.
- [17] K. D. S. Young and L. I. Pettit, "The effect of observations on Bayesian choice of an autoregressive model," *J. Time Ser. Anal.*, vol. 27, no. 1, pp. 41–50, Jan. 2006.
- [18] M. A. Tanner and W. H. Wong, "The calculation of posterior distributions by data augmentation," *J. Amer. Stat. Assoc.*, vol. 82, no. 398, pp. 528–550, 1987.
- [19] D. A. Van Dyk and X. L. Meng, "The art of data augmentation," *J. Comput. Graph. Statist.*, vol. 10, no. 1, pp. 1–50, 2001.
- [20] D. B. Rubin, "Using the SIR algorithm to simulate the posterior distribution," in *Bayesian Statistics 3*, J. M. Bernardo, Eds. Oxford, U.K.: Oxford Univ. Press, 1988, pp. 395–402.
- [21] E. J. Hannan, *Multiple Time Series Models*. New York, NY, USA: Wiley, 1970.
- [22] G. T. Wilson, "The estimation of parameters in multivariate time series models," *J. Roy. Stat. Soc., Ser. B*, vol. 35, no. 1, pp. 76–85, 1973.
- [23] D. R. Osborn, "Exact and approximate maximum likelihood estimators for vector moving average processes," *J. Roy. Stat. Soc., Ser. B, Methodol.*, vol. 39, no. 1, pp. 114–118, Sep. 1977.
- [24] G. Box and G. Jenkins, *Time Series Analysis, Forecasting and Control*. San Francisco, CA, USA: Holden-Day, 1970.
- [25] M. S. Phadke and G. Kedem, "Computation of the exact likelihood function of multivariate moving average models," *Biometrika*, vol. 65, no. 3, pp. 511–519, 1978.
- [26] S. C. Hillmer and G. C. Tiao, "Likelihood function of stationary multiple autoregressive moving average models," *J. Amer. Stat. Assoc.*, vol. 74, no. 367, pp. 652–660, Sep. 1979.
- [27] D. F. Nicholls and A. D. Hall, "The exact likelihood function of multivariate autoregressive-moving average models," *Biometrika*, vol. 66, no. 2, pp. 259–264, 1979.
- [28] A. D. Hall and D. F. Nicholls, "The evaluation of exact maximum likelihood estimates for VARMA models," *J. Stat. Comput. Simul.*, vol. 10, nos. 3–4, pp. 251–262, Apr. 1980.
- [29] V. Solo, "The exact likelihood for a multivariate ARMA model," *J. Multivariate Anal.*, vol. 15, no. 2, pp. 164–173, Oct. 1984.
- [30] A. Luceno, "A fast algorithm for the exact likelihood of stationary and partially nonstationary vector autoregressive-moving average processes," *Biometrika*, vol. 81, no. 3, pp. 555–565, 1994, doi: [10.1093/biomet/81.3.555](https://doi.org/10.1093/biomet/81.3.555).
- [31] J. A. Mauricio, "Exact maximum likelihood estimation of stationary vector ARMA models," *J. Amer. Stat. Assoc.*, vol. 90, no. 429, pp. 282–291, Mar. 1995.
- [32] A. C. Monti, "A proposal for estimation of the parameters of multivariate moving-average models," *J. Time Ser. Anal.*, vol. 19, no. 2, pp. 209–219, Mar. 1998, doi: [10.1111/1467-9892.00086](https://doi.org/10.1111/1467-9892.00086).
- [33] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Berlin, Germany: Springer, 2007.
- [34] C. Kascha, "A comparison of estimation methods for vector autoregressive moving-average models," *Econ. Rev.*, vol. 31, no. 3, pp. 297–324, May 2012.
- [35] S. M. Shaarawy, "Bayesian inferences and forecasts with multiple autoregressive moving average models," *Commun. Statist.-Simul. Comput.*, vol. 18, no. 4, pp. 1481–1509, Jan. 1989, doi: [10.1080/03610918908812836](https://doi.org/10.1080/03610918908812836).
- [36] J. Marriott, N. Ravishanker, A. Gelfand, and J. Pai, "Bayesian analysis of ARMA processes: Complete sampling-based inference under full likelihoods," in *Bayesian Analysis in Statistics and Econometrics: Essays in Honor of Arnold Zellner*, D. Berry, K. Chaloner, and J. Geweke, Eds. New York, NY, USA: Wiley, 1996.
- [37] J. C. C. Chan, E. Eisenstat, and G. Koop, "Large Bayesian VARMA," *J. Econometrics*, vol. 192, no. 2, pp. 374–390, Jun. 2016, doi: [10.1016/j.jeconom.2016.02.005](https://doi.org/10.1016/j.jeconom.2016.02.005).
- [38] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. Reading, MA, USA: Addison-Wesley, 1973.



MOHAMMED ALBASSAM received the B.Sc. degree from the Mathematics Department, King Abdulaziz University, in 1990, and the M.Sc. and Ph.D. degrees from the School of Mathematics and Statistics, University of Sheffield, U.K., in 1995 and 2000, respectively. He is currently working as an Associate Professor with the Statistics Department, King Abdulaziz University. He has published many articles in international journals. His research interests include statistical inference, distributions theory, neutrosophic statistics, and time series analysis.



EMAD E. A. SOLIMAN received the B.Sc., M.Sc., and Ph.D. degrees from the Department of Statistics, Cairo University, in 1987, 1993, and 1999, respectively. He is currently working as an Associate Professor with the Statistics Department, King Abdulaziz University. He has published many articles in international journals. His research interest includes time series analysis.



SHERIF S. ALI is a Professor in statistics with Cairo University. He is currently working as an Associate Professor with the Statistics Department, King Abdulaziz University. His research interests include time series analysis, applied statistics, Bayesian Analysis, and data science. He contributed in original researches and authoring books.

...