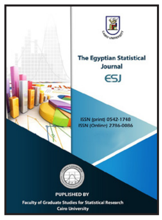
Cairo University

The Egyptian Statistical Journal

No.2 - Volume (66) , 2022

journal homepage: www.esju.journals.ekb.eg

Print ISSN 0542-1748 - Online ISSN 2786-0086



A Direct Bayesian Methodology to Identify the Order of Moving Average Processes using Different Prior Distributions

Mohammed S. AL Bassam[1] Emad E. A. Soliman Sherif S. Ali[2]

Associate Professor Professor Professor

Department of Statistics -King Abdul Aziz university-Jeddah, Saudi Arabia

ABSTRACT

The current study handles the direct Bayesian identification of moving average processes based on different priors. The priors considered in the article are g, natural-conjugate as well as Jeffreys' priors. Posterior probability mass functions for moving average model's order are developed using the above mentioned three priors. Then, the order with maximum probability is selected as the identified model order. In order to ease computations, the likelihood function of the moving average process is approximated. Effectiveness of each posterior mass function in identifying an order for the model is assessed via Mont-Carlo simulations. Numerical results support the use of the proposed Bayesian technique to solve the considered identification problem. Furthermore, the performance of the technique using g prior is better than its performance using the natural-conjugate prior which, in turn, is slightly better than the performance using Jeffreys' prior.

Keywords:

Moving average processes, direct Bayesian identification, indirect Bayesian identification, informative priors, non-informative priors, g prior, Jeffreys' prior.

1. **Introduction**

A direct Bayesian identification approach for autoregressive, AR(p), models was developed by Diaz and Farah (1981). The method assumes the order to be a random variable and derives its posterior probability mass function assuming that the maximum order is a known constant. Then, posterior probabilities of the order's values are computed and the value having maximum probability is chosen to identify the model. Derivation of such posterior mass function is based on the likelihood function of the observations of the time series as well as the chosen prior. Derivations of Diaz and Farah have been employed using a Natural-Conjugate prior (See Broemeling (1985).

An approximate indirect Bayesian method has been developed by Broemeling and Shaarawy (1987) to identify ARMA processes. The approach depends on an assumption that the model's orders are unknown constants but their maximums are known. The technique develops a posterior probability density function for the models' coefficients. To avoid the difficulties in the computations of the exact posterior analysis of ARMA processes, the technique adopts a multivariate t approximation for the posterior density of the coefficients of the model with maximum orders.

[1] Corresponding Author Postal Address: P. O. Box 80203, Jeddah 21589, Saudi Arabia.

[2] The permanent address is: Faculty of Economics and Political Science, Cairo university

Then a sequence of univariate t tests is employed in order to test the significance of coefficients. After removing insignificant coefficients, one can determine the models' order. A Jeffreys' prior was used to develop Broemeling and Shaarawy technique. (see Broemeling and Shaarawy (1987) and (1988)).

The current study is mainly interested in developing a direct Bayesian identification technique of moving average processes using three different well-known priors found in statistical literature. Such priors are the g prior, developed by Zellner (1983 and 1986), the natural-conjugate prior, introduced by Raiffa and Schlaifer (1961) and Jeffreys' prior, introduced by Jeffreys' (1961). The posterior mass function of the order q of MA models will be derived using each prior. The efficiency of each posterior mass function in identifying the first and second order moving average processes is to be investigated using simulation studies.

For well-known reasons, the Bayesian identification of moving average models is complicated because of non-linearity of the errors in the model's coefficients. This problem causes the likelihood function to be complicated and leads to non-standard posterior distributions. A reasonable solution of such problem is to use analytical approximations, such as the Broemeling and Shaarawy one, to give a linear estimate of the errors in the model's coefficients. This simplifies the likelihood function.

The current study is the first attempt to develop a direct methodology for Bayesian identification of MA processes using g prior. Moreover, the study is interested in studying the efficiency of employing g prior in such development using Monte Carlo simulation studies. In addition, the current article is also the first attempt to check the effectiveness of using the natural-conjugate prior to develop the direct Bayesian identification technique for MA processes using simulations. Furthermore, the article aims to compare the goodness of the above-mentioned priors with Jeffreys' prior in developing the proposed identification technique.

The current article is arranged in the following manner: Section 2 introduces a review of the literature. Section 3 discusses the moving average processes. Section 4 defines the three considered priors. In section 5, the direct identification technique is conducted using g prior. Employing the direct technique of MA processes using the natural-conjugate prior as well as Jeffreys' prior is displayed in section 6. Moreover, wide scale simulations to investigate and compare the effectiveness of the considered priors in developing a direct Bayesian identification for MA processes are conducted in section 7.

2. Review of the Literature

There are two well-known Bayesian identification techniques, namely, the direct and the indirect ones. Diaz and Farah (1981) have introduced a direct Bayesian technique to identify AR(p) models. This technique considers the order p of the model a random variable. Then, it develops a posterior mass function for p. The value having maximum probability is selected as a point estimate of the order p (See Broemeling (1985)). Daif et al. (2003) have corrected the derivations of Diaz and Farah. Such corrections enable the investigation of the technique and the comparison of its effectiveness with the indirect Bayesian approach developed by Broemeling and Shaarawy(1988). Using a numerical algorithm, Monahan (1983) developed another direct way to calculate the posterior probabilities for the orders p and q of autoregressive moving average (ARMA) models.

The direct analytical methodology has been extended to seasonal AR processes by Shaarawy and Ali (2003). Another extension to MA processes was conducted by Shaarawy et al. (2007). In that work, the direct technique is compared with both the indirect technique and Monahan's technique. Ali (2009) extended the technique proposed by Shaarawy et al. (2007) in order to identify the mixed ARMA(p, q) processes. Moreover, Shaarawy et al. (2011) have extended the technique to seasonal moving average models. Furthermore, identification of bivariate autoregressive models using a direct Bayesian approach was introduced by Sharaawy et al. (2006). In addition to this, the direct approach has been developed to identify multivariate (vector) AR, MA as well as seasonal AR processes by Shaarawy and Ali (2008), (2012) and (2015) respectively. It might be important to mention that Shaarawy et al. (2007) and (2011) as well as Shaarawy and Ali (2012) have employed the Broemeling and Shaarawy indirect approach as an intermediate stage while developing the direct Bayesian identification approach of non-seasonal, seasonal and multivariate moving average processes. Using Jeffreys' prior, the direct methodology has been successfully extended to the case of nons-easonal and seasonal multivariate ARMA processes by Shaarawy et al. (2018, 2019).

Previous studies employed their theoretical derivations using the natural-conjugate prior as well as Jeffreys' prior. However, in the simulation studies all of them employed Jeffreys' prior only. The g prior has not been employed in developing theoretical derivations or numerical studies by anyone of the above mentioned studies. Furthermore, the goodness of the behavior of the natural-conjugate prior in developing Bayesian identification has not been checked by anyone of these studies.

On the other hand, the prior selection is one of the fundamental problems in Bayesian analysis. Some efforts in the Bayesian time series literature were devoted to this problem. El

Zayat (2007) developed comprehensive surveys for various known informative and non-informative priors. She has developed wide scale simulation studies to check the effectiveness of the employed prior in solving the estimation problem of AR(1) model. Moreover, Shaarawy et al. (2010) considered prior selection problem in the Bayesian prediction process of AR models and investigated the performance of the one step-ahead predictive densities based on different priors. Recently, Al-Bassam et al. (2013) have derived an approximate posterior distribution for the coefficients of MA models using g prior and conducted wide scale simulation studies to investigate and compare the performance of g prior, the natural-conjugate prior as well as Jeffreys' prior in developing an indirect Bayesian approach for the identification of moving average models.

3. Moving average processes

The moving average process of order q, MA(q), is a special case of the ARMA(p, q) process. The MA(q) model can be written using the following formula (Box and Jenkins (1970)):

$$y_t = \Theta(B)\varepsilon_t \qquad (3.1)$$

Where,

$$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

Where, B represents the backshift operator such that $B^j y_t = y_{t-j}$. $\{y_t, t = \cdots, -1,0,1, \ldots\}$ is the time series. $\varepsilon_t{}'$s, are the random errors. They are assumed to be independent identically normally distributed having mean zero and variance τ^{-1}, where $\tau = 1/\sigma^2 > 0$ is known as the precision parameter. θ_i's are the coefficients of the model.

The model can be displayed in the following explicit form:

$$y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q} \qquad (3.2)$$

MA processes are known to be always stationary, whereas they need conditions to be invertible. The conditions of invertibility can be formulated such that the roots of the polynomial equation $\Theta(B)$ should lie outside the unit circle.

Special cases of (3.2) are the MA(1) model given by

$$y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} \qquad (3.3)$$

And the MA(2) model given by

$$y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \qquad (3.4)$$

For MA(1) model the invertibility condition is $|\theta_1| < 1$ and for MA(2) model, the invertibility conditions are $\theta_2 - \theta_1 < 1$, $\theta_2 + \theta_1 < 1$, and $|\theta_2| < 1$. (Box and Jenkins (1970)).

4. Prior Distributions

The g prior is a middle ground of sorts between the informative Natural-Conjugate (NC) prior and non-informative Jeffreys' prior (Karlsson (2001)). It was developed by Zellner (in 1983 and 1986) to avoid a problem that faced the NC prior which is the evaluation of prior covariates of the hyper-parameters. Zellner's motivation was to derive an appropriate prior for the regression parameters in the General Linear Models (GLM).

The joint g prior of the parameters $\underline{\gamma}$ and σ for MA(q) models has been derived by Shaarawy et al. (2010). It has the form

$$p_g(\underline{\gamma},\sigma) \propto \sigma^{-(q+1)} \exp\left\{-\frac{g}{2\sigma^2}(\underline{\gamma}-\overline{\underline{\gamma}})' X'X(\underline{\gamma}-\overline{\underline{\gamma}})\right\}. \qquad (4.1)$$

Where, $\underline{\gamma}=[-\theta_1 - \theta_2 \ldots - \theta_q]'$, σ is the errors variance, $\overline{\underline{\gamma}}$ is an anticipated value of $\underline{\gamma}$ and X represents the matrix of regressors of MA(q) model. It has the t[th] row $\underline{X}_t$ of the form

$$\underline{X}_t = [\varepsilon_{t-1} \ \ldots \ \varepsilon_{t-q}]. \qquad (4.2)$$

Several estimates were suggested for the constant g in the Bayesian literature. Fernàndez et al. (2001) have employed a simulation study and suggested the following estimate for g:

$$g = \begin{cases} \frac{1}{n} & for \quad n > k^2 \\ \frac{1}{k^2} & for \quad n \le k^2 \end{cases} \qquad (4.3)$$

Where n is the sample size and k is the number of the parameters in the model.

On the other hand, the NC prior for the ARMA(p,q) process is the, so called, normal-gamma prior if the Broemeling and Shaarawy approximation is used (See Broemeling and Shaarawy (1988)). A normal-gamma prior of ARMA(p,q) model is displayed in the following form,

$$p_{NC}(\underline{\gamma},\tau) \propto \frac{|Q_h|^{\frac{1}{2}}\tau^{\alpha+\frac{h}{2}-1}}{(2\pi)^{\frac{h}{2}}} e^{-\frac{\tau}{2}[2\beta+(\underline{\gamma}-\underline{\mu})'Q_h(\underline{\gamma}-\underline{\mu})]}, \ \tau > 0, \qquad (4.4)$$

Where, $\alpha, \beta, \mu \in R^h$ and the $h \times h$ positive definite matrix Q_h are called hyper-parameters of the prior distribution, $\underline{\gamma}$ and τ are the model parameters, and $h = p+q$.

Furthermore, Jeffrey's prior is given by (See Jeffreys' (1961)),

$$p(\underline{\gamma},\tau) \propto \tau^{-1} \qquad (4.5)$$

5. Identification of Moving Average Models based on g prior

This section is devoted to developing a direct Bayesian technique to identify MA processes using g prior defined in section 4. This development is one of the basic contributions of the current study. As stated above, the proposed approach is based on assuming the order *q* of MA models to be a random variable with a known upper limit. The technique is interested in obtaining an appropriate posterior probability mass function for *q*.

Assume that $\underline{S}_n = [y_1 \quad y_2 \quad \dots \quad y_n]'$ is a time series with *n* observations. Therefore, the initial residual values $\varepsilon_0, \varepsilon_{-1}, \cdots, \varepsilon_{1-q}$ are unknown. By letting the initial residual values equal their unconditional mean, i.e. $\varepsilon_i = 0,$ where *i*= 0,-1,-2,…,1-*q*, one can get the conditional likelihood function of MA(*q*) model in the form

$$L^*(\underline{\gamma}^{(q)}, q, \tau \mid \underline{S}_n) = \left(\frac{\tau}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{\tau}{2}\sum_{t=1}^{n}\left(y_t + \sum_{i=1}^{q}\theta_{qi}\varepsilon_{t-i}\right)^2\right\} \qquad (5.1)$$

For convenience one should rewrite the form (5.1) in terms of the parameter σ instead of precision τ as follows

$$L^*(\underline{\gamma}^{(q)}, q, \sigma \mid \underline{S}_n) = \sigma^{-n}(2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(y_t + \sum_{i=1}^{q}\theta_{qi}\varepsilon_{t-i}\right)^2\right\} \qquad (5.2)$$

where θ_{qi} is the coefficient of the i^{th} lagged value of ε_t in the q^{th} model, $\underline{\gamma}^{(q)} = [-\theta_{q1} - \theta_{q2} \dots - \theta_{qq}]'$, $\sigma > 0$ and $q=1,2,...,k$ such that *k* is the upper limit of *q*.

Displaying (5.2) in matrix notation we get:

$$L^*(\underline{\gamma}^{(q)}, q, \sigma \mid \underline{S}_n) = \sigma^{-n}(2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\underline{\gamma}^{(q)'} A_1 \underline{\gamma}^{(q)} - 2\underline{\gamma}^{(q)'}\underline{B}_1 + C_1\right]\right\} \qquad (5.3)$$

Where, A_1 is a matrix of order $q \times q$ having ij^{th} element $A_{1ij} = \sum_{t=1}^{n}\varepsilon_{t-i}\varepsilon_{t-j}$, where $i, j = 1,2,\dots,q$. $\underline{B}_1$

Is a vector of order $q \times 1$ having i^{th} element $B_{1i} = \sum_{t=1}^{n} y_t \varepsilon_{t-i}$, where $i = 1,2,\dots,q$. Moreover, $C_1 = \sum_{t=1}^{n} y_t^2$.

The conditional likelihood function (5.3) is analytically intractable since the error terms ε_{t-i}'s as functions in the coefficients of the model are nonlinear. They need an approximation to simplify the form of the likelihood function. However, one can't approximate the unknown ε_{t-i}'s without determining the value of q. Thus, we need another identification method to obtain an initial value of q say q_0, then we can approximate the errors as linear functions of the model's coefficients. The proposed technique to get q_0 is Bayesian. It is the indirect identification approach introduced by Broemeling and Shaarawy (1987, 1988). Such technique was investigated and shows high percentages of correct identification (See Shaarawy et al.(2007) and Al-Bassam et al.(2013)). Consequently, (5.3) is approximated by the following form

$$L^{**}(\underline{\gamma}^{(q)}, q, \sigma \mid \underline{S}_n) = \sigma^{-n}(2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\underline{\gamma}^{(q)'}\hat{A}_1\underline{\gamma}^{(q)} - 2\underline{\gamma}^{(q)'}\underline{\hat{B}}_1 + C_1\right]\right\} \qquad (5.4)$$

And $\hat{A}_1$ is a matrix of order $q \times q$ having ij^{th} element $\hat{A}_{1ij} = \sum_{t=1}^{n} \hat{\varepsilon}_{t-i}\hat{\varepsilon}_{t-j}$, where $i, j = 1,2,\ldots,q$. $\underline{\hat{B}}_1$ is a vector of order $q \times 1$ vector having i^{th} element $\hat{B}_{1i} = \sum_{t=1}^{n} y_t\hat{\varepsilon}_{t-i}$, where $i = 1,2,\ldots,q$.

Moreover, C_1 is as defined above.

Moreover, the following prior assumptions are considered:

- The conditional joint prior density for $\gamma^{(q)}$ and σ given the order q is assumed to be the joint g-prior defined in (4.1), after approximating the error trems $\boldsymbol{\varepsilon_{t-i}}$ by the corresponding non-linear least squares estimates $\hat{\varepsilon}_{t-i}$. It is given by

$$p_g^*(\underline{\gamma}^{(q)}, \sigma \mid q) \propto \sigma^{-(q+1)} \exp\{-\frac{g}{2\sigma^2}(\underline{\gamma}^{(q)} - \overline{\underline{\gamma}}^{(q)})'\hat{X}'\hat{X}(\underline{\gamma}^{(q)} - \overline{\underline{\gamma}}^{(q)})\} \qquad (5.5)$$

Such that $\hat{X}$ is the estimate of the matrix of regressors $\boldsymbol{X}$ in MA(q) model, where the t[th] row $\underline{\hat{X}}_t$ is defined as

$$\underline{\hat{X}}_t = [\hat{\varepsilon}_{t-1} \ldots \hat{\varepsilon}_{t-q}] \qquad (5.6)$$

- The prior probability mass function of the order q is uniform

$$P_1(q) = k^{-1} \qquad q = 1,2,\ldots,k \qquad (5.7)$$

Using these quantities, we assert the following theorem:

Theorem 5.1

Considering the approximate conditional likelihood function (5.4) in addition to the priors (5.5) and (5.7), the approximate posterior probability mass function of *q*, the order of MA process, is given by

$$\zeta_1(q \mid \underline{S}_n) \propto \left|A_1^*\right|^{-\frac{1}{2}} (2)^{-\frac{n+2}{2}} (2\pi)^{-\frac{n+q}{2}} \Gamma(\frac{n}{2}) [C_1^* - \underline{B}_1^{*'} A_1^{*-1} \underline{B}_1^*]^{-\frac{n}{2}} , \qquad q = 1,2,\ldots,k \tag{5.8}$$

Where

$$\begin{aligned} A_1^* &= \hat{A}_1 + g\hat{X}'\hat{X} \\ \underline{B}_1^* &= \underline{\hat{B}}_1 - 2g\,\hat{X}'\hat{X}\underline{\bar{\gamma}}^{(q)} \\ C_1^* &= C_1 + g\,\underline{\bar{\gamma}}^{(q)'}\hat{X}'\hat{X}\underline{\bar{\gamma}}^{(q)} \end{aligned} \tag{5.9}$$

$\hat{A}_1, \underline{\hat{B}}_1$ and C_1 are as defined above.

Proof

Combining the priors (5.5) and (5.7), one gets the joint prior

$$P_g\left(\underline{\gamma}^{(q)}, q, \sigma\right) \propto \sigma^{-(q+1)} \exp\{-\frac{g}{2\sigma^2}([\underline{\gamma}^{(q)} - \underline{\bar{\gamma}}^{(q)}]'\hat{X}'\hat{X}[\underline{\gamma}^{(q)} - \underline{\bar{\gamma}}^{(q)}])\} \tag{5.10}$$

Multiply the approximate likelihood function (5.4) by the joint prior (5.10), one gets a joint posterior distribution in the form

$$\begin{aligned} \zeta_1\left(\underline{\gamma}^{(q)}, q, \tau \mid \underline{S}_n\right) \propto \frac{\sigma^{-(n+q+1)}}{(2\pi)^{\frac{n}{2}}} \exp\{-\frac{1}{2\sigma^2}([\underline{\gamma}^{(q)'}\hat{A}_1\underline{\gamma}^{(q)} - 2\underline{\gamma}^{(q)'}\underline{\hat{B}}_1 + C_1] + \\ g[\underline{\gamma}^{(q)} - \underline{\bar{\gamma}}^{(q)}]'\hat{X}'\hat{X}[\underline{\gamma}^{(q)} - \underline{\bar{\gamma}}^{(q)}])\} \end{aligned} \tag{5.11}$$

Reorganizing the terms between the brackets ( ) in the exponent of (5.11), one gets

$$(\underline{\gamma}^{(q)'}[\hat{A}_1 + g\,\hat{X}'\hat{X}]\underline{\gamma}^{(q)} - 2\underline{\gamma}^{(q)'}[\underline{\hat{B}}_1 - g\,\hat{X}'\hat{X}\underline{\bar{\gamma}}^{(q)}] + [C_1 + g\,\underline{\bar{\gamma}}^{(q)'}\hat{X}'\hat{X}\underline{\bar{\gamma}}^{(q)}])$$

Let

$$A_1^* = \hat{A}_1 + g\,\hat{X}'\hat{X} \;,\; \underline{B}_1^* = \underline{\hat{B}}_1 - g\,\hat{X}'\hat{X}\underline{\bar{\gamma}}^{(q)} \text{ and } C_1^* = C_1 + g\,\underline{\bar{\gamma}}^{(q)'}\hat{X}'\hat{X}\underline{\bar{\gamma}}^{(q)}$$

Then, the joint posterior distribution (5.11) becomes

$$\zeta_1\left(\underline{\gamma}^{(q)}, q, \sigma \mid \underline{S}_n\right) \propto \frac{\sigma^{-(n+q+1)}}{(2\pi)^{\frac{n}{2}}} \exp\{-\frac{1}{2\sigma^2}[\underline{\gamma}^{(q)'}A_1^*\underline{\gamma}^{(q)} - 2\underline{\gamma}^{(q)'}\underline{B}_1^* + C_1^*]\} \tag{5.12}$$

Consider the exponent in the above equation and complete the squares, then (5.12) can be written as

$$\zeta_1\left(\underline{\gamma}^{(q)}, q, \sigma \mid \underline{S}_n\right) \propto \frac{\sigma^{-(n+q+1)}}{(2\pi)^{\frac{n}{2}}} \exp\{-\frac{1}{2\sigma^2}[(\underline{\gamma}^{(q)} - A_1^{*-1}\underline{B}_1^*)' A_1^* (\underline{\gamma}^{(q)} - A_1^{*-1}\underline{B}_1^*) + (C_1^* - \underline{B}_1^{*'} A_1^{*-1}\underline{B}_1^*)]\} \tag{5.13}$$

Integrating the joint posterior in (5.13) to eliminate $\underline{\gamma}$, one gets the marginal posterior distribution of the parameters q and σ in the form

$$\zeta_1\left(q, \sigma \mid \underline{S}_n\right) \propto (2\pi)^{\frac{-n+q}{2}} \sigma^{-(n+1)} \left|A_1^*\right|^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma^2}[C_1^* - \underline{B}_1^{*'} A_1^{*-1}\underline{B}_1^*]\} \tag{5.14}$$

Then, integrating (5.14) with respect to σ one gets the posterior mass function of q given by (5.8). This completes the proof. ■

This distribution assigns a probability to each value of q or in other words, to each of the k moving average models

$$y_t = \varepsilon_t - \sum_{i=1}^{q} \theta_{qi}\varepsilon_{t-i}\,, \qquad q = 1, 2, \ldots, k\,.$$

When calculating the probabilities of q in (5.8), one should observe that the inverse of the matrix $A_1^* = \hat{A}_1 + g\,\hat{X}'\hat{X}$ and its determinant will be computed for each value of q. One way to select the identified model is to select the value of q with largest posterior probability.

6. Identification of Moving Average Models based on NC and Jeffreys' Priors

Consider the conditional likelihood function (5.1) of MA(q) models, and substitute the errors by their non-linear last squares estimates. One obtains an approximate likelihood function in the form

$$L^1(\underline{\gamma}^{(q)}, q, \tau \mid \underline{S}_n) = \left(\frac{\tau}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{\tau}{2}\sum_{t=1}^{n}\left(y_t + \sum_{i=1}^{q}\theta_{qi}\hat{\varepsilon}_{t-i}\right)^2\right\} \tag{6.1}$$

Shaarawy et al. (2007) have mixed the above approximate likelihood function with the natural conjugate prior defined in (4.4), namely the normal-gamma prior. They have derived the approximate posterior mass function of q. It is given by

$$\zeta_2(q \mid \underline{S}_n) \propto \frac{\left|A_2^*\right|^{-\frac{1}{2}} \Gamma\left(\frac{n+2\alpha}{2}\right)}{\left|Q_q\right|^{-\frac{1}{2}} \pi^{\frac{n}{2}}} [C_2^* - \underline{B}_2^{*\prime} A_2^{*-1} \underline{B}_2^*]^{-\frac{(n+2\alpha)}{2}} \quad , \; q = 1,2,\ldots,k \qquad (6.2)$$

Where

$$A_2^* = \hat{A}_1 + Q_q$$
$$\underline{B}_2^* = \hat{\underline{B}}_1 + Q_q \underline{M}^{(q)} \qquad (6.3)$$
$$C_2^* = C_1 + 2\beta + \underline{M}^{(q)\prime} Q_q \underline{M}^{(q)}$$

and $\hat{A}_1, \hat{\underline{B}}_1$ and C_1 are as defined in section 5.

On the other hand, Shaarawy et al. (2007) have also asserted that, depending on the likelihood function (6.1) and Jeffreys' prior (4.5), the posterior mass function of q is

$$\zeta_3(q \mid \underline{S}_n) \propto \frac{\Gamma\left(\frac{n-q}{2}\right)}{(2\pi)^{\frac{n-q}{2}}} \left|\hat{A}_1\right|^{-\frac{1}{2}} [C_1 - \hat{\underline{B}}_1^{\prime} \hat{A}_1^{-1} \hat{\underline{B}}_1]^{-\frac{(n-q)}{2}} , \qquad q = 1,2,\ldots,k \qquad (6.4)$$

where $\hat{A}_1, \hat{\underline{B}}_1$ and C_1 are as defined in section 5.

7. Numerical studies

Large-scale simulation studies are conducted to investigate the performance and evaluate the accuracy of the proposed direct Bayesian methodology in solving the problem of identification for MA models. The proposed methodology is based on the above mentioned three priors, the g, the natural-conjugate and Jeffreys' prior. Some MA models, with different sets of coefficients' values, are used for illustration. Different time series lengths are determined to represent short, medium, and long time series to check the effect of the time series length on the behavior of the proposed methodology. Moreover, different values of the upper limit for the MA model's order are assumed. Requested simulations are conducted using computer programs especially developed, in MATLAB 7, for the study.

Observe that, the involved simulation studies are the first attempt to assess the performance of both g prior and the natural conjugate prior in developing a direct Bayesian technique to identify models for MA processes. Employing an informative prior through a simulation study is a very complicated process since the hyper-parameters exist in the informative prior need to be estimated.

Evaluation and comparison of the performance of the adopted identification approach, depending on each one of the above mentioned three considered priors, are done using the percentage of correct identification as an effectiveness criterion. Such percentage has the form

$$p_0 = \frac{n_0}{N} \times 100 \qquad (7.1)$$

Where n_0 is the number of times where the technique selects the correct model and N is the number of series generated from the MA model.

7.1. Design of Simulation Studies

Four simulation studies were conducted, using some MA(1) and MA(2) models having different coefficients. They are referred to as Case I, Case II, Case III and Case IV. The coefficient of MA(1) model is selected to be 0.5 and 0.8. While, the selected coefficients for MA(2) model are (0.9,-0.3) and (0.1,0.8). Some coefficients values are selected at the center of the domain of invertibility of the model, whereas other selections are near its boundaries. Different lengths are selected for the time series as short, medium and long time series; these lengths are 50, 100, 150, 200 and 300. The maximum order of the model, k, is considered once to be 3 and another time to be 4.

For illustration, the procedure of Case I starts by generating a set of data consisting of 500 random variables $\boldsymbol{\varepsilon_t}$ following the normal distribution with mean zero and precision parameter 2. Then the MA(1) model's equation with $\theta = 0.5$ is employed recursively to generate an artificial time series of 500 observations following the model. The initial values of ε_t are assumed to be zero. In order to eliminate the influence of the initial values, we delete the earliest 200 observations from the generated time series, thus we obtain a time series of 300 observations following MA(1) model with $\theta = 0.5$. In the second step, consider only the earliest 50 observations as the generated time series. The adopted direct identification approach is employed to identify a model for the time series. This process is done three times using each of the above mentioned three priors. Moreover, it is repeated for each of the above mentioned two upper limits of the order. The second step, explained above, is employed again for the earliest 100 observations including the earliest 50. The process is repeated again for the earliest 150 and 200 observations, and finally, for the whole generated time series. Third, the above two steps are repeated 500 times. In the last step, one computes the percentage of correct identification (7.1) for the adopted approach using each selection of the prior, time series length and maximum assumed order, k.

It is worth noting that, it is required to estimate the hyper-parameters in order to employ g prior and the natural conjugate prior. To estimate the hyper-parameters a training sample is used (See El-Zayat (2007) and Shaarawy et al (2010)). The training sample is considered as the first 10 observations of the time series (for short time series) or the first 10% of the time series observations, if the series is longer than 100.

To employ the proposed direct identification technique given a certain maximum order k for the moving average model, one evaluates the approximate posterior mass function of q using each selected prior. The posterior probabilities of each value of q are computed in each case. Then, the technique selects the value of q with largest posterior probability as an order for the model. Finally, percentage of correct identification (7.1) is calculated in each case using 500 replications of the simulation steps. Case II, III and IV are structured similarly, however, using some different models and coefficients.

7.2. Numerical Results

After conducting the above mentioned four simulations, the values of the effectiveness criterion computed for the proposed identification methodology are summarized in table (7.1). The lengths of the generated time series as well as the prior functions are presented over the rows of the table, whereas the considered four source models and the assumed two maximum orders are presented in the columns. The entries of the table are the percentages of correct identification in each case.

Table (7.1)
Effectiveness Results of the Proposed Direct Bayesian Identification Technique Using G, N-C and Jeffreys' Priors

		MA(1) Models				MA(2) Models			
		$\theta_1 = 0.5$		$\theta_1 = 0.8$		$\theta_1 = 0.9$, $\theta_2 = -0.3$		$\theta_1 = 0.1$, $\theta_2 = 0.8$	
Time Series length	***PRIOR***	**k = 3**	**k = 4**	**k = 3**	**k = 4**	**k = 3**	**k = 4**	**k = 3**	**k = 4**
50	**G**	**93.0**	**93.4**	**92.0**	**92.2**	**34.8**	**33.4**	**90.0**	**90.2**
	N-C	**85.6**	**87.0**	**83.4**	**82.0**	**47.8**	**46.2**	**85.4**	**85.6**
	JEFFREYS'	**78.0**	**73.8**	**76.4**	**68.8**	**50.4**	**44.8**	**79.6**	**68.4**
100	**G**	**97.8**	**97.4**	**94.4**	**92.8**	**61.8**	**58.0**	**92.2**	**92.0**
	N-C	**92.6**	**91.6**	**88.8**	**83.6**	**73.6**	**70.0**	**88.8**	**88.8**
	JEFFREYS'	**86.8**	**83.8**	**82.2**	**77.4**	**74.2**	**65.8**	**84.8**	**78.6**
150	**G**	**97.4**	**97.6**	**94.2**	**93.0**	**78.0**	**75.4**	**94.0**	**93.6**
	N-C	**90.0**	**89.4**	**85.4**	**82.8**	**84.4**	**78.4**	**88.6**	**85.8**
	JEFFREYS'	**90.4**	**88.6**	**83.8**	**80.0**	**83.0**	**76.0**	**86.8**	**81.0**
200	**G**	**97.6**	**98.2**	**95.2**	**94.4**	**88.2**	**85.2**	**95.0**	**93.8**
	N-C	**90.6**	**89.2**	**86.0**	**83.0**	**88.8**	**83.6**	**86.2**	**81.4**
	JEFFREYS'	**92.0**	**90.8**	**86.4**	**82.4**	**87.2**	**81.0**	**85.4**	**78.8**
300	**G**	**97.8**	**97.8**	**95.2**	**94.4**	**95.2**	**92.8**	**96.6**	**95.8**
	N-C	**87.4**	**85.6**	**84.0**	**80.6**	**89.8**	**85.4**	**86.4**	**81.4**
	JEFFREYS'	**89.8**	**88.0**	**84.8**	**83.0**	**90.2**	**86.0**	**86.8**	**83.0**

Source: Simulated Data

Regarding table (7.1), some general conclusions can be observed. First, regarding MA(1) sources: The values of effectiveness criterion for the proposed technique using the three prior functions are greater than 68% in all cases. They become larger when the number of observations becomes larger. Application of the identification method using G prior appears to give the highest percentages of correct identification for all considered time series lengths and all maximum orders. Moreover, the application using N-C prior appears to give values of effectiveness criterion higher than the application using Jeffreys' prior for short and moderate time series lengths for all maximum orders. Nevertheless, the application of the technique using both N-C and Jeffreys' priors give equivalent results for long time series. Furthermore, the results of $k = 3$ are always better than those of $k = 4$.

On the other hand, regarding MA(2) sources: The position of the coefficients of the model inside the invertibility domain affects the obtained results. The MA(2) model with coefficients (0.9, -0.3) lies far from the boundaries of the invertibility domain. For this model, the values of effectiveness criterion are less than 50% for short time series. They increase as time series length becomes larger. For short and moderate time series the application of the technique using both N-C and Jeffreys' priors achieve higher percentages of correct identification than the application of the technique using G prior. Whereas, for long time series, the performance of the G prior dominates the other two priors. Finally, The MA(2) model with coefficients (0.1, 0.8) lies near the boundaries of the invertibility domain. For this model, the obtained values of effectiveness criterion are similar to those of the MA(1) sources.

The results obtained in table (7.1) support the adequacy of employing each of the three considered priors, especially the G one, to introduce a solution for the Bayesian identification problem of pure MA processes.

It might be suggested for future work to use the G prior and the Natural-Conjugate prior - rather than the classical use of the Jeffreys'one - in the Bayesian identification of mixed autoregressive moving average processes, and in Bayesian estimation and prediction of the class of ARMA, seasonal ARMA and multivariate ARMA models. And to compare the effect of the selected prior on the performance of the technique.

Acknowledgement

This Project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant no. (95 / 130 / 1434). The authors, therefore, acknowledge with thanks DSR technical and financial support.

The authors also would like to express their deep thanks to Professor Samir Shaarawy for his valuable suggestions and proofreading.

References

[1] Akaike, H. (1973). "Information Theory and an Extension of the Maximum Likelihood Principal". *Budapest: Akademia Kiado,* pp.267-281.

[2] Akaike, H. (1974). "A New Look at Statistical Model Identification". *IEEE Transaction on Automatic control*, Vol. **19**, pp.716-723.

[3] AL Bassam, M.S., Soliman E.E.A. and Ali, S.S. (2013). Selection of Prior Distribution in the Indirect Bayesian Identification of Moving Average Models. *Canadian Journal on Computing in Mathematics, Natural Sciences, Engineering and Medicine*, Vol. **4**, No. 2, pp.205-212.

[4] Ali, S.S. (2009). "An Effectiveness Study of Bayesian Identification Techniques for ARMA Models". *The Egyptian Statistical Journal*, Vol. **53**, No. 1, pp.1-13.

[5] Beveridge, S. and Oickle, C. (1994). " A Comparison of Box-Jenkins and Objective Methods for Determining the Order of Non-Seasonal ARMA Model". *Journal of Forecasting*, Vol. **13**, pp.413-434.

[6] Box, G. and Jenkins, G. (1970). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

[7] Broemeling, L. D. (1985). *Bayesian Analysis of Linear Models.* Marcel Dekker Inc., New York.

[8] Broemeling, L. and Shaarawy, S. (1987). "Bayesian Identification of Time Series". *The 22nd Annual Conference in Statistics, Computer Science and Operation Research, Institute of Statistical Studies and Research, Cairo University*, Vol. **2**, pp. 146-159.

[9] Broemeling, L. and Shaarawy, S. (1988). "Time Series: A Bayesian Analysis in Time Domain". *Bayesian Analysis of Time Series and Dynamic Models*, Edited by Spall, pp. 1-22.

[10] Daif, A. L., Soliman, E. A. and Ali, S. S. (2003). "On Direct and Indirect Bayesian Identification of Autoregressive Models". *The 14th Annual Conference on Statistics*

and Computer Modeling in Human and Social Science, Faculty of Economics and Political Science, Cairo University.

[11] DeGroot, M. H. (1970). *Optimal Statistical Decision*. McGraw-Hill, New York.

[12] Diaz, J. and Farah, J. L. (1981). "Bayesian Identification of Autoregressive Process". *Presented at the 22nd NBER-NSC Seminar on Bayesian* Inference in *Econometrics.*

[13] El-Zayat, N. I., (2007). *On the selection of prior distributions in Bayesian analysis*. Unpublished M.Sc. Thesis, Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt.

[14] Fernández, C., Ley, E., and Steel, M. F. J. (2001). Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics*, *Elsevier*, Vol. 100, pp. 381-427.

[15] Hannan, E. J. and Quinn, B. G. (1979). " The Determination of the Order of an Autoregression". *Journal of the Royal Statistical Society*, pp. 190-195.

[16] Jeffreys, H. (1961). *Theory of Probability*. (3rd Ed.), Oxford University Press, London.

[17] Karlsson, S. (2001). *Bayesian Methods in Econometrics: Linear Regression*. (A course to a graduate level class), Stockholm School of Economics, Swedish School of Economics and Business Administration (Helsinki).

[18] Mills, J. and Prasad, K. (1992). A Comparison of Model Selection Criteria. *Econometric Reviews*, vol. 11, p. 201-233.

[19] Monahan, J. F. (1983). "Fully Bayesian Analysis of ARIMA Time Series Models". *Journal of Econometrics*, Vol. **21,** pp. 307-331.

[20] Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory*. Division of Research, Graduate School of Business Administration, Harvard University, Boston.

[21] Shaarawy, S. M. and Ali, S. S. (2003). "Bayesian Identification of Seasonal Autoregressive Models". *Communications in Statistics – Theory and Methods,* issue 5, Vol. **32,** pp. 1067-1084.

[22] Shaarawy, S. M., Al Bassam, M. and Ali, S. S. (2006). "A Direct Bayesian Technique to Select the Order of Bivariate Autoregressive Process". *The Egyptian Statistical Journal*, Vol. **50**, No. 1, pp.1-22.

[23] Shaarawy, S. M., Soliman, E. A. and Ali, S. S. (2007). "Bayesian Identification of Moving Average Models". *Communications in Statistics- Theory and Methods*, Issue 12, Vol. **36**, pp. 2301-2312.

[24] Shaarawy, S. and Ali, S. (2008). "Bayesian Identification of Multivariate Autoregressive Processes". *Communications in Statistics- Theory and Methods*, Issue 5, Vol. **37**, pp. 791-802.

[25] Shaarawy, S. M., Soliman, E. A. and Shahin, H.E.A. (2010). "Bayesian Prediction of Autoregressive Models Using Different Types of Priors". *The Egyptian Statistical Journal*, Vol. **54**, No. 2, pp.108-126.

[26] Shaarawy, S.M., Soliman, E.E.A. and El-Souda, R.M. (2011). "Bayesian identification of seasonal moving average models". *The Egyptian Statistical Journal*, Vol. **55**, No.1, pp.40-52.

[27] Shaarawy, S. and Ali, S. (2012). "Bayesian Model Order Selection of Vector Moving Average Processes". *Communications in Statistics- Theory and Methods*, Issue 4, Vol. **41**, pp. 684-698.

[28] Shaarawy, S. and Ali, S. (2015). "Bayesian Identification of Seasonal Multivariate Autoregressive Processes". *Communications in Statistics- Theory and Methods*, Vol. **44**, pp.823-836.

[29] Shaarawy, S.M., Soliman E.E.A. and Ali, S.S. (2018). “Indirect and Direct Bayesian Techniques to Identify the Orders of Vector ARMA Processes” *The Egyptian Statistical Journal*, Vol. **62**, No.1, pp.15-34. DOI: 10.21608 / ESJU.2018.244222.

[30] Shaarawy, S.M., Soliman, E.E.A. and Shahin, H. (2018). Bayesian Prediction of Moving Average Processes using Different Types of Priors. *The Egyptian Statistical Journal*, Vol.62, No.1, pp(35-52). DOI: 10.21608/esju.2018.244260

[31] Shaarawy, S.M., Ali, S.S. and Soliman, E.E.A. (2021). A Baysian Procedure to Identify the Orders of Vector Moving Average Processes with Seasonality. *The Egyptian Statistical Journal*, Vol. **64**, No.1, pp.1-20. DOI:10.21608/esju.2021.189433.

[32] Albassam, M.S., Soliman E.E.A. and Ali, S.S. (2021). An Effectiveness Study of the Bayesian Inference with Multivariate Autoregressive Moving Average Processes. *Communications in statistics*- Simulation and Computation, DOI:10.1080/03610918.2021.1967986.

[33] Albassam, M.S., Soliman E.E.A. and Ali, S.S. (2022). Bayesian Estimation of Multivariate Pure Moving Average Processes. IEEE Access, Vol.10, pp (14225-14235), DOI:10.1109/ACCESS.2022.3146724.

[34] Zellner, A. (1983). "Application of Bayesian Analysis and Econometrics". *The Statistician*, Vol. **32**, pp.23-34.

[35] Zellner, A. (1986). On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior distributions. *Bayesian inference and Decision Techniques*. *Essays in Honor of Bruno de Finetti, P.d.f. Goel and A. Zellner*, eds., North-Holland, pp.233-243, New York.