Efficiency of the Approximate Bayesian Prediction of ARMA Models: A Simulation Study

BY

Emad El-Din A. Soliman

Faculty of Economics and Political Science, Department of Statistics
Cairo University, Egypt

ABSTRACT

This paper is interested in the efficiency of the approximate Bayesian prediction of ARMA models. Three different approximate one step-ahead predictive densities using three different approximations are considered. These approximations are proposed by Newbold (1973), Zellner and Reynolds (1978) and Broemeling and Shaarawy (1988). Simulation is employed to study and compare the performance of the three approximate one step-ahead predictive densities using some ARMA models. Simulation results show that, there is no remarkable difference between the three approximate one step-ahead predictive densities in their efficiency. Moreover, their first two moments converge to those of the exact one step-ahead predictive density.

Key words: Prediction - One step-ahead predictive density - ARMA models - Newbold approximation – Zellner and Reynolds approximation – Broemeling and Shaarawy approximation.

1. Introduction

Prediction is the final objective of time series analysis. The predictive density is the Bayesian tool to implement prediction, see Broemeling (1985). The main difficulty encountered with Bayesian prediction of MA and hence ARMA models is that the predictive density is non-standard. Thus, numerical integration methods must

ences and prediction, Monahan(1983).

e analysis of both ARMA and MA models is difficult

functions in the model's coefficients. Therefore, the e

adratic function in the parameters and, hence, the like

ly intractable.

proaches are suggested in the literature to solve this pr

clude the use of numerical integration methods, Mona

irst two

ency criteria.

es are close to

r efficiency in

ive density

Dynamic Models, e
21).
6. DeGroot, M.H. (19
York.
7. Geman, S. and Ge
and the Bayesian r
Analysis Machine
8. Ismail, M.A. (1994

Monte Carlo based methods such as, Gibbs sampling technique, Geman and Geman(1984), and analytic approximations, Shaarawy and El-Shawadfy (1987).

Numerical analysis is difficult and needs a professional use of computers, especially for high order models. On the other side, analytic approximations have the advantage that they lead to standard one step-ahead predictive densities, Shaarawy and Broemeling (1984).

Various approximations are suggested in Bayesian time series literature. However, we confine to three of them. These three approximations were proposed by Newbold (1973), Zellner and Reynolds (1978) and Broemeling and Shaarawy (1988). They will be denoted by (N), (Z-R) and (B-S) respectively.

It is worth noting that, Broemeling and Shaarawy (1988) have used their approximation to derive the posterior and the predictive densities of ARMA models theoretically. Moreover, Shaarawy and El-Shawadfy (1987) have derived the approximate one step-ahead predictive density of SARMA model using the (B-S) approximation. Yet these two studies haven't investigated the efficiency of the asserted approximate densities.

Ismail (1994) has studied theoretically the relationships between the three approximations for a general non-linear autoregressive moving average NLARMA model. Moreover, he investigated and compared the performance of the three approximate posterior densities of the coefficient of the MA(1) model via Monte Carlo simulation. Ali (1998) have extended Ismail's simulation work to the approximate posterior densities of the coefficients of the MA(2) model.

On the other hand, Soliman (1999) has derived the approximate one step-ahead predictive densities using both the (N) and (Z-R) approximations for both the ARMA and seasonal ARMA (SARMA) models. Furthermore, he investigated and compared the performance of these two approximate one step-ahead predictive

ıe (B-S) approximate one step-ahead predictive density,

A(2) models via Monte Carlo simulation.

ıail (1994) and Ali (1998) showed numerically that, although

rors' sum of squares keeps more terms from the exact errors'

N) one, which in turn keeps more terms from the exact

an the (B-S) one, however, the moments of the three

nsities of MA models are similar. Soliman (1999) obtained a

e approximate one step-ahead predictive densities.

(1-

ions

Nevertheless, the above mentioned studies restricted their attention in the study of the efficiency of their proposed approximate densities to pure MA models. This study extends the numerical work of Soliman (1999) to assess and compare the efficiency and performance of the three approximate one step-ahead predictive densities for mixed ARMA models

The structure of the study is designed as follows: Section 2 gives brief definition for the ARMA models. In section 3, the used Bayesian approximations are discussed. Whereas, in section 4 we give the approximate one step-ahead predictive densities of the mixed ARMA model and show their relationships. Section 5 presents the structure and results of the simulation studies. Finally, general conclusions are given in section 6.

2. ARMA(p,q) Models

The ARMA(p,q) model is defined in Box and Jenkins (1970). It can be written as

$$\Phi(B)y_t = \Theta(B)\varepsilon_t \qquad (2.1)$$

where, p and q are the orders of the model. They represent the numbers of AR and MA coefficients in the model respectively. Moreover, y_t's t=1,2,..... are the time series observations, ε_t's t=1,2,..... are the random errors assumed to be i.i.d. N(0, τ^{-1}), where the precision parameter $\tau = \sigma^{-2}$, and

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots\ldots - \phi_p B^p \text{, and}$$

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

where, B is the backshift operator, such that, $B^s y_t = y_{t-s}$ and the coefficients vector γ of the ARMA model consists of $h = p+q$ elements and is defined as

$$\underline{\gamma} = [\phi_1 \ldots\ldots \phi_p \ \theta_1 \ldots\ldots \theta_q]' \qquad (2.2)$$

Stationarity and invertibility conditions of the model (2.1) are such that the roots of $\Phi(B) = 0$ and $\Theta(B) = 0$ respectively lie outside the unit circle, see Box and Jenkins (1970). The model (2.1) can be written explicitly in the form

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots\ldots + \phi_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots\ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t \qquad (2.3)$$

Special cases of the ARMA(p,q) model (2.3) are obtained by substituting p and q by appropriate values. e.g. ARMA(1,1) model is given by,

$$y_t = \phi y_{t-1} - \theta \varepsilon_{t-1} + \varepsilon_t \qquad (2.4)$$

The stationarity condition of this model is $-1 < \phi < 1$ and the invertibility condition is $-1 < \theta < 1$. Note that, pure AR and MA models are also special cases of (2.3) when either q=0 or p=0 respectively.

The conditional likelihood function of the ARMA(p,q) model, conditioned on the first p observations can be written as, Broemeling (1985),

$$L(\underline{\gamma}, \tau \mid \underline{Y}) \propto \tau^{\frac{n-p}{2}} \exp\{ -\frac{\tau}{2} Q \} \qquad (2.5)$$

where, γ is the coefficients vector (2.2), τ is the precision parameter and the vector of observations $\underline{Y} = [y_{p+1} \ .... \ y_n]'$, n > p . Moreover, Q is the errors sum of squares

$$Q = \sum_{t=p+1}^{n} \varepsilon_t^2 = \sum_{t=p+1}^{n} \left(y_t - \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \right)^2 \qquad (2.6)$$

For convenience we rewrite model (2.1) in matrix notation as

$$\underline{Y} = X\underline{\gamma} + \underline{\varepsilon} \qquad (2.7)$$

where, $\underline{Y}$ and $\underline{\gamma}$ are as defined above, $\underline{\varepsilon} = [\varepsilon_{p+1} \ ... \ \varepsilon_n]'$ is the vector of errors. Whereas, X is a (n-p)×h matrix with the (i-p)$^{\text{th}}$ row vector, X_t, of the form

$$\underline{X}_t = \left[y_{t-1} \ .... \ y_{t-p} \ \ -\varepsilon_{t-1} \ .... \ -\varepsilon_{t-q} \right] \quad t=p+1,\ldots,n \qquad (2.8)$$

Moreover, the initial values of both the observations and the errors are assumed to be zero, i.e. $y_t = \varepsilon_t = 0$ for all $t < 1$.

3. Bayesian Approximations

Several Bayesian approximations have been proposed in the literature. The basic idea of these approximations is to simplify the likelihood function of the model by approximating the errors' sum of squares as a quadratic function in the coefficients' vector $\underline{\gamma}$.

3.1. The (B-S) Approximation

Broemeling and Shaarawy (1988) have approximated the errors as linear functions in $\underline{\gamma}$, by their non-linear least squares estimates (NLSE) as follows :

$$\varepsilon_t^{(1)} = y_t - \underline{\hat{X}}_t \underline{\gamma} \qquad t=p+1,\ldots,n \qquad (3.1)$$

where, y_t and $\underline{\gamma}$ are as defined before and $\underline{\hat{X}}_t$ is obtained from (2.8) by replacing each ε_i by its NLSE $\hat{\varepsilon}_i$ (See Broemeling and Shaarawy (1988)).

3.2. The (N) Approximation

Newbold (1973) has expanded the errors, as linear functions in $\underline{\gamma}$, around their NLSE's using Taylor's expansion as follows :

$$\varepsilon_t^{(2)} = \hat{\varepsilon}_t + \underline{U}_t(\underline{\gamma} - \underline{\hat{\gamma}}) \qquad t=p+1,\ldots,n \qquad (3.2)$$

Where,

$$\underline{U}_t = \left[\frac{\partial \varepsilon_t}{\partial \gamma_1} \quad \frac{\partial \varepsilon_t}{\partial \gamma_2} \quad \ldots\ldots \quad \frac{\partial \varepsilon_t}{\partial \gamma_h} \right] \Big|_{\underline{\gamma}=\underline{\hat{\gamma}}} \qquad t=p+1,\ldots,n \qquad (3.3)$$

3.3. The (Z-R) Approximation

Rather than approximating the errors by linear functions, Zellner and Reynolds (1978) have expanded the errors' sum of squares as a quadratic function in the coefficients vector $\underline{\gamma}$, using Taylor's expansion as follows :

$$Q^{(3)}(\underline{\gamma}) = \hat{Q} + (\underline{\gamma} - \underline{\hat{\gamma}})' \frac{1}{2} R (\underline{\gamma} - \underline{\hat{\gamma}}) \qquad (3.4)$$

where, $\hat{Q}$ is the least errors sum of squares and R is an h×h matrix with ij$^{\text{th}}$ element,

$$R_{ij}(\underline{\hat{\gamma}}) = \frac{\partial^2 Q}{\partial \gamma_i \partial \gamma_j}\Big|_{\underline{\gamma}=\underline{\hat{\gamma}}} = 2\left[\sum_{t=p+1}^{n} \left(\hat{\varepsilon}_t \frac{\partial^2 \varepsilon_t}{\partial \gamma_i \partial \gamma_j} \right) + \sum_{t=p+1}^{n} \left(\frac{\partial \varepsilon_t}{\partial \gamma_i} \right)\left(\frac{\partial \varepsilon_t}{\partial \gamma_j} \right) \right] \Bigg|_{\underline{\gamma}=\underline{\hat{\gamma}}}$$

$$i,j = 1,\ldots,h \qquad (3.5)$$

Note that, the (B-S) approximate errors' sum of squares is a special case of that of (N), which in turn is a special case of that of (Z-R). Equality of the sums of squares holds under some conditions, see Ismail (1994).

4. One Step-Ahead Predictive Densities of ARMA Models

The one step-ahead predictive density of the ARMA(p,q) model can be obtained by multiplying the joint conditional density $f(y_{n+1}, \underline{Y} \mid \underline{\gamma}, \tau)$ of the first future observation y_{n+1} and the time series data given the parameters by the

appropriate prior density $\xi(\gamma,\tau)$. Then, the parameters γ and τ are eliminated by integration, Shaarawy and El-Shawadfy (1987).

Regarding the matrix notation of the ARMA(p,q) model in (2.7), let's define the row vector $\underline{X}_{n+1}$ of the p+q regressors of the future observation by letting $t = n+1$ in (2.8). ie,

$$\underline{X}_{n+1} = \left[y_n \quad \cdots \quad y_{n-p+1} \quad -\varepsilon_n \quad \cdots \quad -\varepsilon_{n-q+1} \right] \tag{4.1}$$

Thus, the prediction error is defined as

$$\varepsilon_{n+1} = y_{n+1} - \underline{X}_{n+1}\,\underline{\gamma} \tag{4.2}$$

In what follows, the Jeffreys' prior will be used. It has the form, see DeGroot (1970),

$$\xi(\gamma,\tau) \propto \frac{1}{\tau} \quad , \ \tau > 0 \tag{4.3}$$

The form of the exact one step-ahead predictive density of the ARMA(p,q) model corresponding to the Jeffreys' prior is given by, see Soliman (1999),

$$\xi\left(y_{n+1}|\underline{Y}\right) \propto \int_{\underline{\gamma}} \left[Q_{n+1}\right]^{\frac{-(n+1-p)}{2}} d\underline{\gamma} \tag{4.4}$$

where, $Q_{n+1} = \sum_{t=p+1}^{n+1} \varepsilon_t^2 = Q + \varepsilon_{n+1}^2$,

such that, Q is the errors sum of squares (2.6) and ε_{n+1} is the prediction error (4.2).

One can conclude that, the exact one step-ahead predictive density is unattainable analytically since Q_{n+1} is non-quadratic in γ. Therefore, it is very difficult to perform this multiple integral analytically. It requires numerical integrations in order to evaluate the density at any value of y_{n+1}. This numerical integration becomes more difficult as the number of coefficients increases.

If the sum of squares Q_{n+1} is approximated as a quadratic function in γ, the approximate one step-ahead predictive density would be a t density. Broemeling and Shaarawy (1988) have derived the (B-S) approximate one step-ahead predictive density of the ARMA(p,q) model. Whereas, Soliman (1999) has derived the (N) and the (Z-R) approximate one step-ahead predictive densities of the ARMA(p,q) model. They used several priors in the derivation. They assumed that the prediction error ε_{n+1} defined in (4.2) is estimated as a linear function in γ by

$$\tilde{\varepsilon}_{n+1} = y_{n-1} - \underline{\hat{X}}_{n+1}\underline{\gamma}$$

where, $\underline{\hat{X}}_{n+1}$ includes the NLSE of the error terms in $\underline{X}_{n+1}$ defined in (4.1).

$$\underline{\hat{X}}_{n+1} = \left[y_n \quad \cdots \quad y_{n-p+1} \quad -\hat{\varepsilon}_n \quad \cdots \quad -\hat{\varepsilon}_{n-q+1} \right] \tag{4.5}$$

Broemeling and Shaarawy (1988) and Soliman (1999) have proved that based on the jeffreys' prior (4.3), the three approximate one step-ahead predictive densities of the ARMA(p,q) model are non-central t with **(n-p-1)** degrees of freedom. However, they have different location and precision. Soliman (1999) has also studied the relationships between the parameters of the three one step-ahead predictive densities, see Soliman (1999).

The (B-S) approximate one step-ahead predictive density has the following location and precision respectively,

$$\begin{aligned} E(y_{n+1} \mid \underline{Y}) &= \underline{\hat{X}}_{n+1} A_1^{-1} \underline{B}_1 \\ P(y_{n+1} \mid \underline{Y}) &= \frac{n-p-1}{d_{11}\hat{Q}} \end{aligned} \tag{4.6}$$

where,

$$\begin{aligned} A_1 &= \hat{X}'\hat{X} \quad , \quad \underline{B}_1 = \hat{X}'\underline{Y} \\ d_{11} &= 1 + \underline{\hat{X}}_{n+1} A_1^{-1} \underline{\hat{X}}'_{n+1} \end{aligned} \tag{4.7}$$

where, $\hat{X}$ is the matrix with **(i-p)**$^{\text{th}}$ row $\underline{\hat{X}}_t$ defined in (2.8). t = p+1, …, n.

Whereas, the (N) approximate one step-ahead predictive density has the following location and precision respectively,

$$\begin{aligned} E(y_{n+1} \mid \underline{Y}) &= \underline{\hat{X}}_{n+1}\hat{\gamma} \\ p(y_{n+1} \mid \underline{Y}) &= \frac{n-p-1}{d_{22}\hat{Q}} \end{aligned} \tag{4.8}$$

where,

$$d_{22} = 1 + \underline{\hat{X}}_{n+1} A_2^{-1} \underline{\hat{X}}'_{n+1} \tag{4.9}$$

and $A_2 = U'U$, where, U is the matrix of order **(n-p)×h** with **(i-p)**$^{\text{th}}$ row $\underline{U}_t$ defined in (3.3), t = p+1, ... ,n.

Moreover, The (Z-R) approximate one step-ahead predictive density has the following location and precision respectively,

$$E(y_{n+1} \mid \underline{Y}) = \underline{\hat{X}}_{n+1}\underline{\hat{\gamma}}$$

$$p(y_{n+1} \mid \underline{Y}) = \frac{n-p-1}{d_{33}\hat{Q}} \qquad (4.10)$$

where,

$$d_{33} = 1 + \underline{\hat{X}}_{n+1} A_3^{-1} \underline{\hat{X}}'_{n+1} \qquad (4.11)$$

and $A_3 = \frac{1}{2}R$, where R is the matrix of order **h**× **h** with **ij**th element defined in (3.5), t = p+1, ... , n.

Note that, the variance of the t distribution is obtained from the precision using the relation, see DeGroot (1970),

$$VAR(Y) = \frac{d}{d-2}\left(\frac{1}{p(Y)}\right)$$

Where d are the degrees of freedom and $P(y)$ is the precision.

5. Simulation Studies

The main objective of these simulation studies is to investigate and compare the performance and the efficiency of the three approximate one step-ahead predictive densities of mixed ARMA models using some criteria.

5.1. Efficiency Criteria

To investigate and compare the performance of the three approximate one step-ahead predictive densities and their closeness to the exact density, two types of measures are considered: Discrepancy measures and Moments.

Discrepancy Measures

Discrepancy measures are usually used to check the closeness of an approximate density to the exact. The smaller the value of the discrepancy measure, the closer the approximate density to the exact.

A famous example of the discrepancy measures is the maximum absolute difference (MAD). The MAD measure is defined as follows:

$$MAD = \max\left|\hat{\xi}(.) - \xi(.)\right| \qquad (5.1)$$

where, $\xi(.)$ is the exact density and $\hat{\xi}(.)$ is the approximate one. For more information about MAD and its properties see Silverman (1986).

ARMA(1,1) model are (0.2, 0.2), (0.3,0.5), (0.2,0.8) and (0.8, 0.7). These combinations represent different positions inside both the stationarity and the invertibility domains.

To perform the integrals involved, the Gaussian numerical integration method is used because of its accuracy (see Atkinson (1985)). Computer programs written in Matlab7 were especially designed to perform the concerned computations.

5.3. Simulation Results

The results of each selected model are displayed in three tables corresponding to the three above mentioned efficiency criteria. Each table is divided into five blocks according to the selected time series lengths. Each block includes rows displaying the frequency distribution of the criteria for one of the considered predictive densities over the 500 generated time series of the corresponding length.

Model 1: ARMA(1,1) with $\varphi = 0.2$, $\theta = 0.2$

This model well lies in both the stationarity and the invertibility domains. Its results are displayed in tables (1), (2) and (3). Table (1) is devoted to the MAD results. The table displays the frequency distributions of the values of the MAD in the 500 replications for all approximate one step-ahead predictive densities at all time series lengths.

It is clear that about 85% of the frequencies of the MAD values are concentrated in the smallest interval [0,1) for all approximate densities at all time series lengths. Moreover, this percentage increases to more than 90% for long time series (n=200 and 250). That is most of the values of the MAD are close to zero regardless the time series length. This result shows the closeness of the three approximate one step-ahead predictive densities to the exact one in the MAD sense. Moreover, there is no observable difference between the three approximate densities in their closeness to the exact one.

On the other hand, regarding table (2), which is devoted to the mean results, it is clear that the values of the predictive mean of the exact one step-ahead predictive density concentrates in the interval [-0.5, 0.5) which is centered at zero for all time series lengths. Moreover, the predictive means of the three approximate one step-ahead predictive densities tend to concentrate in the same interval for short and moderate time series lengths ($n \leq 100$) and concentrate in the same interval for longer

Moments

The second type of criteria to check and investigate the performance of the three approximate predictive densities is the group of moments. This group consists of the first two moments, namely, the Predictive Mean

$$E\left(y_{n+1}|\underline{Y}\right)=\int y_{n+1}\xi\left(y_{n+1}|\underline{Y}\right)dy_{n+1} \qquad (5.2)$$

and the Predictive Variance

$$VAR\left(y_{n+1}|\underline{Y}\right)=\int\left(y_{n+1}-E\left(y_{n+1}|\underline{Y}\right)\right)^2\xi\left(y_{n+1}|\underline{Y}\right)dy_{n+1} \qquad (5.3)$$

5.2. Simulation Design

Each simulation run consists of the following steps: First, a time series following a given ARMA model is generated. The data generation process consists of two steps. A sequence of 450 values of a standard normal random variate ε_t is generated. Then, the ARMA model's equation is used recursively to obtain 450 observations of the process. The initial values of both the errors and the observations are taken to be the errors' unconditional mean, i.e. $y_i = \varepsilon_i = 0$ if $i < 1$. The first 200 observations are ignored to remove the initialization effect, that is, we have a time series of 250 observations.

Second, the generated time series is used to evaluate both the exact and the three approximate one step-ahead predictive densities. Moreover, the efficiency criteria are calculated for each approximate density. The time series lengths are chosen to be 30, 70, 100, 200 and 250. These lengths represent short, moderate and long time series lengths. That is, Using the first 30 observations, of the 250 generated observation, the exact and the three approximate predictive densities are calculated. Moreover, the efficiency criteria are obtained for each approximate density. This process is repeated for the first 70 observations (including the first 30). Again, the process is repeated for the first 100, 200 and, finally, for the 250 observations.

Third, 500 replications of the above two steps are done. Finally, the obtained 500 values of each efficiency criteria for each approximate density, at each time series length, are tabulated in a frequency table, see section (5.3).

In this study, four simulation processes are employed using the ARMA(1,1) model, defined in equation (2.4). The combinations of the parameters (φ, θ) of the

time series. This indicates that the predictive means of the three approximate densities are identical and similar to the mean of the exact one.

In addition, regarding table (3) devoted to the variance results, the variance of the exact density tend to concentrate in the interval [0,1) as the time series length increases. While the variances of the three approximate densities appear to be similar at all time series lengths. Although they are different from the variance of the exact density for short and moderate time series ($n \leq 100$), they become similar to it for longer time series.

That is, the three approximate one step-ahead predictive densities are close to the exact one in the sense of the three considered efficiency criteria. Moreover, their behavior is identical in the sense of each of the three criteria.

Table (1)

Frequency Distributions of MAD for ARMA(1,1) with $\varphi = 0.2$ $\theta = 0.2$

n	Dist.	MAD Classes					Total
		0.0-	1.0-	2.0-	3.0-	4.0-	
30	(B-S)	438	31	9	2	20	500
	(N)	438	32	8	2	20	500
	(Z-R)	438	32	8	2	20	500
70	(B-S)	428	46	13	8	5	500
	(N)	429	46	12	8	5	500
	(Z-R)	429	45	13	8	5	500
100	(B-S)	441	28	10	6	15	500
	(N)	442	27	10	6	15	500
	(Z-R)	441	28	10	6	15	500
200	(B-S)	458	19	8	5	10	500
	(N)	458	19	8	5	10	500
	(Z-R)	458	19	8	5	10	500
250	(B-S)	457	24	3	2	14	500
	(N)	457	24	3	2	14	500
	(Z-R)	457	24	3	2	14	500

Table (2)

Frequency Distributions of Predictive MEAN for ARMA(1,1) with $\varphi = 0.2$ $\theta = 0.2$

n	Dist.	MEAN Classes					Total
		<-1	-1.0-	-0.5-	0.5-	1.0-	
30	Ex	0	0	500	0	0	500
	(B-S)	2	10	472	13	3	500
	(N)&(Z-R)	1	10	475	10	4	500
70	Ex	0	0	500	0	0	500
	(B-S)	1	1	494	4	0	500
	(N)&(Z-R)	1	1	495	3	0	500
100	Ex	0	0	500	0	0	500
	(B-S)	0	3	494	3	0	500
	(N)&(Z-R)	0	3	494	3	0	500
200	Ex	0	0	500	0	0	500
	(B-S)	0	0	500	0	0	500
	(N)&(Z-R)	0	0	500	0	0	500
250	Ex	0	0	500	0	0	500
	(B-S)	0	0	500	0	0	500
	(N)&(Z-R)	0	0	500	0	0	500

Table (3)

Frequency Distributions of Predictive VARIANCE for ARMA(1,1) with $\varphi = 0.2$ $\theta = 0.2$

n	Dist.	VARIANCE Classes					Total
		0.0-	1.0-	1.5-	2.0-	2.5-	
30	Ex	485	15	0	0	0	500
	(B-S)	1	0	0	0	499	500
	(N)	1	0	0	0	499	500
	(Z-R)	1	0	1	0	498	500
70	Ex	497	3	0	0	0	500
	(B-S)	1	4	327	166	2	500
	(N)	1	5	351	139	4	500
	(Z-R)	2	6	361	131	0	500
100	Ex	500	0	0	0	0	500
	(B-S)	1	466	33	0	0	500
	(N)	1	470	28	1	0	500
	(Z-R)	2	478	20	0	0	500
200	Ex	500	0	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	500	0	0	0	0	500
250	Ex	500	0	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	500	0	0	0	0	500

Model 2: ARMA(1,1) with $\varphi = 0.3$, $\theta = 0.5$

This model also well lies in both the stationarity and the invertibility domains. However, its coefficients are unequal and slightly farther from zero than the coefficients of model 1. The results of this model are displayed in tables (4), (5) and (6) respectively. Regarding the three tables, the conclusions obtained are similar to those obtained from tables (1), (2) and (3) respectively. That is, the three approximate one step-ahead predictive densities are also close to the exact one in the sense of the three considered efficiency criteria having identical behavior.

Table (4)

Frequency Distributions of MAD for ARMA(1,1) with $\varphi = 0.3$ $\theta = 0.5$

n	Dist.	MAD Classes					Total
		0.0-	1.0-	2.0-	3.0-	4.0-	
30	(B-S)	450	28	7	5	10	500
	(N)	450	28	7	5	10	500
	(Z-R)	450	28	7	5	10	500
70	(B-S)	450	22	11	4	13	500
	(N)	450	22	11	4	13	500
	(Z-R)	450	22	11	4	13	500
100	(B-S)	440	31	9	6	14	500
	(N)	440	31	9	6	14	500
	(Z-R)	440	31	9	6	14	500
200	(B-S)	453	26	8	6	7	500
	(N)	453	26	8	6	7	500
	(Z-R)	453	26	8	6	7	500
250	(B-S)	430	35	15	4	16	500
	(N)	430	35	15	4	16	500
	(Z-R)	431	34	15	4	16	500

Table (5)

Frequency Distributions of Predictive MEAN for ARMA(1,1) with $\varphi = 0.3$ $\theta = 0.5$

N	Dist.	MEAN Classes					Total
		<-1	-1.0-	-0.5-	0.5-	1.0-	
30	Ex	0	0	500	0	0	500
	(B-S)	3	14	468	14	1	500
	(N)&(Z-R)	3	11	471	14	1	500
70	Ex	0	0	500	0	0	500
	(B-S)	0	5	487	8	0	500
	(N)&(Z-R)	0	4	488	8	0	500
100	Ex	0	0	500	0	0	500
	(B-S)	0	1	498	1	0	500
	(N)&(Z-R)	0	1	497	2	0	500
200	Ex	0	0	500	0	0	500
	(B-S)	0	3	497	0	0	500
	(N)&(Z-R)	0	2	498	0	0	500
250	Ex	0	0	500	0	0	500
	(B-S)	0	1	498	1	0	500
	(N)&(Z-R)	0	1	499	0	0	500

Table (6)

Frequency Distributions of Predictive VARIANCE for ARMA(1,1) with $\varphi = 0.3$ $\theta = 0.5$

n	Dist.	VARIANCE Classes					Total
		0.0-	1.0-	1.5-	2.0-	2.5-	
30	Ex	459	41	0	0	0	500
	(B-S)	1	0	0	0	499	500
	(N)	1	0	0	0	499	500
	(Z-R)	1	0	0	0	499	500
70	Ex	494	6	0	0	0	500
	(B-S)	1	4	328	166	1	500
	(N)	1	6	356	136	1	500
	(Z-R)	1	5	350	142	2	500
100	Ex	498	2	0	0	0	500
	(B-S)	3	456	41	0	0	500
	(N)	3	469	28	0	0	500
	(Z-R)	4	465	30	1	0	500
200	Ex	500	0	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	500	0	0	0	0	500
250	Ex	500	0	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	499	1	0	0	0	500

Model 3: ARMA(1,1) with $\varphi = 0.2$, $\theta = 0.8$

This model well lies inside the stationarity domain. Whereas, it is close to the boundary of the invertibility domain. The results of this model are displayed in tables (7), (8) and (9) respectively. Regarding Table (7), we get a similar conclusion, about the MAD, as that obtained from Tables (1) and (4). That is, the three approximate densities are close to the exact one in the MAD sense. The same conclusion can be said regarding the results of the variance shown in Table (9) which gives similar conclusions as Tables (3) and (6), i.e. the three approximate densities converge to the exact one in the sense of the variance. Note that, their variances become closer to zero than the variance of the exact one.

On the other hand, regarding Table (8), the conclusion is slightly different from that obtained from Tables (2) and (5). The mean of the exact density concentrates in the interval [-0.5, 0.5) for all time series lengths. However, the means of the approximate densities tend to concentrate in the same interval in a slower speed than in the previous two models. A percentage varies approximately from 70 to 80% of the frequencies of the means of the approximate densities concentrate in the interval [-0.5, 0.5). The approximate densities need longer time series to converge to the exact one in the mean sense.

Table (7)

Frequency Distributions of MAD for ARMA(1,1) with $\varphi = 0.2$ $\theta = 0.8$

n	Dist.	MAD Classes					Total
		0.0-	1.0-	2.0-	3.0-	4.0-	
30	(B-S)	446	27	9	5	13	500
	(N)	448	25	9	5	13	500
	(Z-R)	448	25	9	5	13	500
70	(B-S)	473	14	4	2	7	500
	(N)	473	14	4	2	7	500
	(Z-R)	473	14	4	2	7	500
100	(B-S)	465	19	6	4	6	500
	(N)	465	19	6	4	6	500
	(Z-R)	465	19	6	4	6	500
200	(B-S)	454	23	7	2	14	500
	(N)	454	23	7	2	14	500
	(Z-R)	454	23	7	2	14	500
250	(B-S)	423	38	11	9	19	500
	(N)	423	38	11	9	19	500
	(Z-R)	423	38	11	9	19	500

Table (8)

Frequency Distributions of Predictive MEAN for ARMA(1,1) with $\varphi = 0.2$ $\theta = 0.8$

N	Dist.	MEAN Classes					Total
		<-1	-1.0-	-0.5-	0.5-	1.0-	
30	Ex	0	0	500	0	0	500
	(B-S)	5	61	376	53	5	500
	(N)&(Z-R)	6	57	387	44	6	500
70	Ex	0	0	500	0	0	500
	(B-S)	10	47	392	50	1	500
	(N)&(Z-R)	8	51	388	52	1	500
100	Ex	0	0	500	0	0	500
	(B-S)	7	53	369	66	5	500
	(N)&(Z-R)	5	51	377	62	4	500
200	Ex	0	0	500	0	0	500
	(B-S)	7	46	380	63	4	500
	(N)&(Z-R)	6	45	389	56	4	500
250	Ex	0	0	500	0	0	500
	(B-S)	9	59	358	69	5	500
	(N)&(Z-R)	9	61	354	71	5	500

Table (9)

Frequency Distributions of Predictive VARIANCE for ARMA(1,1) with $\varphi = 0.2$ $\theta = 0.8$

n	Dist.	VARIANCE Classes					Total
		0.0-	1.0-	1.5-	2.0-	2.5-	
30	Ex	300	187	13	0	0	500
	(B-S)	1	0	0	0	499	500
	(N)	1	0	0	0	499	500
	(Z-R)	2	0	0	2	496	500
70	Ex	318	181	1	0	0	500
	(B-S)	1	5	313	178	3	500
	(N)	1	5	364	129	1	500
	(Z-R)	2	12	353	126	7	500
100	Ex	351	149	0	0	0	500
	(B-S)	2	455	43	0	0	500
	(N)	3	463	34	0	0	500
	(Z-R)	3	466	31	0	0	500
200	Ex	385	115	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	499	0	1	0	0	500
250	Ex	404	96	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	500	0	0	0	0	500

Model 4: ARMA(1,1) with $\varphi = 0.8$, $\theta = 0.7$

This model lies near the boundaries of both the stationarity and the invertibility domains. The results of this model are displayed in tables (10), (11) and (12) respectively. Regarding the three tables, the conclusions obtained are similar to those of Model 1 obtained from tables (1), (2) and (3) respectively.

Table (10)

Frequency Distributions of MAD for ARMA(1,1) with $\varphi = 0.8$ $\theta = 0.7$

n	Dist.	MAD Classes					Total
		0.0	1.0-	2.0-	3.0-	4.0-	
30	(B-S)	441	25	13	1	20	500
	(N)	441	25	13	1	20	500
	(Z-R)	442	25	12	1	20	500
70	(B-S)	434	40	16	5	5	500
	(N)	434	40	16	5	5	500
	(Z-R)	434	40	16	5	5	500
100	(B-S)	439	30	8	5	18	500
	(N)	439	30	8	5	18	500
	(Z-R)	438	30	8	5	19	500
200	(B-S)	438	28	15	5	14	500
	(N)	438	28	15	5	14	500
	(Z-R)	438	28	15	5	14	500
250	(B-S)	414	44	15	6	21	500
	(N)	414	44	15	6	21	500
	(Z-R)	414	44	15	6	21	500

Table (11)

Frequency Distributions of Predictive MEAN for ARMA(1,1) with $\varphi = 0.8$ $\theta = 0.7$

N	Dist.	MEAN Classes					Total
		<-1	-1.0-	-0.5-	0.5-	1.0-	
30	Ex	0	1	497	2	0	500
	(B-S)	8	26	434	27	5	500
	(N)&(Z-R)	7	23	445	20	5	500
70	Ex	0	0	500	0	0	500
	(B-S)	2	7	484	7	0	500
	(N)&(Z-R)	1	6	490	3	0	500
100	Ex	0	0	500	0	0	500
	(B-S)	0	7	488	5	0	500
	(N)&(Z-R)	0	9	486	4	1	500
200	Ex	0	0	500	0	0	500
	(B-S)	0	2	495	3	0	500
	(N)&(Z-R)	0	2	496	2	0	500
250	Ex	0	0	500	0	0	500
	(B-S)	0	1	497	2	0	500
	(N)&(Z-R)	0	1	498	1	0	500

Table (12)

Frequency Distributions of Predictive VARIANCE for ARMA(1,1) with $\varphi = 0.8$ $\theta = 0.7$

n	Dist.	VARIANCE Classes					Total
		0.0 -	1.0-	1.5-	2.0-	2.5-	
30	Ex	478	22	0	0	0	500
	(B-S)	1	0	0	0	499	500
	(N)	1	0	0	0	499	500
	(Z-R)	1	1	0	0	498	500
70	Ex	498	2	0	0	0	500
	(B-S)	1	4	328	164	3	500
	(N)	1	4	359	134	2	500
	(Z-R)	3	6	361	127	3	500
100	Ex	499	1	0	0	0	500
	(B-S)	1	463	36	0	0	500
	(N)	1	471	28	0	0	500
	(Z-R)	5	476	19	0	0	500
200	Ex	500	0	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	500	0	0	0	0	500
250	Ex	500	0	0	0	0	500
	(B-S)	500	0	0	0	0	500
	(N)	500	0	0	0	0	500
	(Z-R)	500	0	0	0	0	500

6. Discussion and Conclusions

This study is interested in investigating the efficiency of the approximate Bayesian prediction of mixed ARMA models. Three approximate one step-ahead predictive densities using three different approximations were considered using the Jeffreys' prior. The three approximations were proposed by Newbold (1973), Zellner and Reynolds (1978) and Broemeling and Shaarawy (1988). Moreover, the three corresponding approximate one step-ahead predictive densities were derived by Soliman (1999) and Broemeling and Shaarawy (1988) respectively.

The three approximate one step-ahead predictive densities are non-central t having the same degrees of freedom but different location and precision parameters. There are some relationships between the location and precision parameters of the three approximate densities.

Four large-scale Monte Carlo simulation studies were employed to study and compare the performance of the three approximate one step-ahead predictive densities of ARMA models. Four ARMA(1,1) models are used for

illustration. The maximum absolute difference (MAD) and the first two moments, the predictive mean and variance, were used as efficiency criteria.

Simulation results indicated the following:

1- The three approximate one step-ahead predictive densities are close to the exact one in the MAD sense. They are similar in their efficiency in this sense.
2- The predictive mean of the exact one step-ahead predictive density concentrates around zero for all time series lengths.
3- The predictive means of the three approximate densities were similar to each other and converge to the predictive mean of the exact density as the time series length increases.
4- The predictive variance of both the exact and the three approximate densities converge to zero as the time series length increases.
5- The predictive variances of the three approximate densities were found to be similar to each other and converge to the predictive variance of the exact density as the time series length increases.

The results obtained support the selection of the (B-S) approximate one step-ahead predictive density to get point and interval predictions of the first future observation of mixed ARMA models. This recommendation is due to the fact that the first two moments of this approximate density are very similar to those of the other two approximate densities. Moreover, the (B-S) approximate density is simpler in its computations.

References

1. Ali, S.S. (1998). *Approximations for posterior densities of moving average models: (a comparative study)*. Unpublished M.Sc. thesis, Dept. of Statistics, Faculty of Economics and Political Science, Cairo University.
2. Atkinson, K. (1985). *Elementary numerical analysis*. John Wiley and Sons, USA.
3. Box, G. and Jenkins, G. (1970). *Time series analysis: forecasting and control*. Holden day, New York.
4. Broemeling, L.D. (1985). *Bayesian analysis of linear models*. Marcel Dekker, New York.
5. Broemeling, L.D. and Shaarawy, S.M. (1988). Time series: A Bayesian analysis in time domain. *Studies in Bayesian Analysis of Time Series and*

Dynamic Models, edited by J. Spall, Marcel Dekker Inc., New York, pp.(1-21).

6. DeGroot, M.H. (1970). *Optimal statistical decision*. McGraw -Hill, New York.
7. Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of the images. *IEEE Transaction pattern Analysis Machine Intelligence* . Vol. 6, pp.(721-741).
8. Ismail, M.A. (1994). *Bayesian forecasting for nonlinear time series*. Unpublished Ph.D. thesis, Dept. of Mathematics, University of Wales, Aberystwyth, UK.
9. Monahan, J.F. (1983). Fully Bayesian analysis of ARMA time series models. *J. of Econometrics*, Vol. 21, PP.(307-331).
10. Newbold, P. (1973). Bayesian estimation of Box and Jenkins transfer function model for noise models. *JRSS B*, Vol. 35(2), PP.(323-336).
11. Shaarawy, S.M. and Broemeling, L. (1984). Bayesian inferences and forecasts with moving averages processes. *Communications in Statistics*, Vol. 13(15), PP.(1871-1888).
12. Shaarawy, S.M. and El-Shawadfy, G.A. (1987). Approximate Bayesian predictive densities of future observations for seasonal ARMA models. *The 22nd Annual Conference in Statistics, Computer Science and Operations Research*, Institute of Statistical Studies and Research, Cairo University, Vol. 2 Stat.,pp.(160-178).
13. Silverman, B.W. (1986). *Density estimation: for statistics and data analysis*. Chapman and Hall, London.
14. Soliman, E.E.A. (1999). *On Bayesian Time Series Analysis*. Unpublished Ph.D. thesis, Dept. of Statistics, Faculty of Economics and Political Science, Cairo University.
15. Zellner, A. and Reynolds, R. (1978). Bayesian analysis of ARMA models. *Presented at the Sixteenth Seminar on Bayesian Inference in Econometrics*, June (2-3), pp.(1-5).