Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/130106

Estimating Seasonal Moving Average Model Using Bayesian Approach

E. A. Ahmed, H. M. Abd-Elgaber and R. A. farghali*

Department of Mathematics, Insurance and Applied Statistics, Faculty of Commerce and Business Administration, Helwan University, Cairo, Egypt

Received: 25 Nov. 2022, Revised: 4 Jan. 2023, Accepted: 15 Jan. 2023.
Published online: 1 Jan. 2024

Abstract: This paper utilizes the Gibbs sampling technique to develop a Bayesian inference for Seasonal Moving Average (SMA) model, which includes parameters that distinguish between Multiplicative and Non-multiplicative models (referred to as Augmented Seasonal Moving Average hereafter). The construction of Bayesian inference involves several steps. Firstly, the method of Non-linear least squares (NLS) is used to estimate unknown lagged errors, allowing for the approximation of the complex likelihood function. Secondly, both a semi-conjugate prior distribution and a non-informative prior distribution are applied to the unknown parameters and initial errors. Thirdly, the prior distributions are combined with the approximated likelihood function to obtain the joint posterior distribution. Lastly, the full conditional distributions are derived as part of the Gibbs sampling process. The proposed method is notable for its simplicity in assessing the significance of the parameters that distinguish between Multiplicative and Non-multiplicative models, a task that is challenging to accomplish using classical analysis. The convergence of the method was verified, ensuring that it reached stable and reliable results.

Keywords: Multiplicative Seasonal Moving Average Model; Non-multiplicative Seasonal Moving Average Model; Gibbs Sampling.

1 Introduction

Box and Jenkins methodology is a highly regarded technique for analysing time-series data. In their work, Box and Jenkins proposed the concept of Seasonal Autoregressive Integrated Moving Average (SARIMA) model family as a means of forecasting seasonal patterns in time-series data that involves a single variable. The analysis of SMA models, which are a specific subset of SARIMA models, using Bayesian methods presents a challenge due to the nonlinear nature of the errors in the parameters. This results in a complicated likelihood function, which makes defining prior distributions and conducting posterior analysis difficult. The literature provides three distinct approaches for Bayesian time series analysis. The first method for Bayesian time series analysis involves numerical integration, as described in [1]. The second method utilizes analytical approximations, with several examples of these approximations being developed in the literature (see [2,3,4,5,6]). The third method is based on sampling techniques, and the advancement of Markov Chain Monte Carlo (MCMC) techniques, specifically the utilization of the Gibbs sampling algorithm, has greatly improved Bayesian time series analysis. In this paper, we utilized the third approach to Bayesian time series analysis, as it effectively deals with the issue of starting values and considers SMA model as Multiplicative model. This method is also less complicated and less time-consuming compared to the first approach, especially when working with models with multiple parameters.

Several sources employed this method. For instance, [7,8] utilized MCMC techniques to conduct Bayesian analysis on Autoregressive Moving Average (ARMA) models. They assumed prior distributions for the initial observations and errors without considering the seasonality feature. In order to estimate Multiplicative Seasonal Autoregressive (SAR) and ARMA models, Barnett et al. [9,10] utilized MCMC methods, but their procedure was more complex as it relied on sampling functions of partial autocorrelations and restricted the coefficient space to meet the stationarity and invertibility conditions.

In recent years, several studies have utilized MCMC techniques, particularly Gibbs sampling, to develop Bayesian estimates for Multiplicative (SARIMA, Double SARIMA, Triple SARIMA) models as well as their special cases (SAR, SMA, Double SAR, Double SMA, Triple SAR) (see [11,12,13,14,15,16,17,18,19]).

Furthermore, Amin [20] utilized the Gibbs algorithm to predict Multiplicative SARIMA models.

Unlike the approach employed by [8,10], which imposes restrictions on the parameter space to satisfy stationarity and invertibility conditions, this paper adopts the approach utilized by [11,12,21,22], and others. This approach incorporates hyper-parameters that ensure the prior distribution for the model coefficients resides within the region of stationarity and invertibility, thus ensuring the process is both stationary and invertible.

*Corresponding author e-mail: hemat.mustafa.abdel-gaber@commerce.helwan.edu.eg

The purpose of this paper is to extend Amin's [20] algorithm on Augmented Seasonal Moving Average models. Our approach does not rely on the initial errors and allows us to easily construct confidence intervals for the parameters that distinguish between Multiplicative and Non-multiplicative models and assess their significance.

The present paper is organized as follows: In Section 2, we provide an overview of SMA models. In Section 3, we introduce Bayesian analysis of Augmented SMA model. Section 4 presents the proposed Gibbs sampling algorithm for Augmented SMA model. In Sections 5 and 6, the algorithm is assessed through simulation study and real-world application. Finally, the paper concludes in Section 7.

2 Overview of SMA models

It is stated that a time series can be considered to have been generated from SMA model with orders q and Q if it adheres to:

$$y_t = \sum_{j=1}^{q} \varphi_j \epsilon_{t-j} + \sum_{k=1}^{Q} \Phi_k \epsilon_{t-ks} + \sum_{j=1}^{q}\sum_{k=1}^{Q} \beta_{jk}\, \epsilon_{t-j-ks} + \epsilon_t \quad (1)$$

where $\{\epsilon_t\}$ is a series of independent variables that have mean equals zero and variance of a^2. The non-seasonal and seasonal orders of SMA model are symbolized by q and Q, respectively. $\varphi_j (j = 1, \dots, q)$ and $\Phi_k\ (k = 1, \dots, Q)$ indicate the non-seasonal and seasonal coefficients, respectively. $\beta_{jk}(j = 1, .., q; k = 1, \dots, Q)$ represent the interaction coefficients, and the symbol s indicates the number of seasons that occur in a year.

Let $\beta_{jk} = \varphi_j \Phi_k + \lambda_{jk}, (j = 1,2, \dots, q; k = 1,2, \dots, Q)$

Model (1) can be restated in the following way:

$$y_t = \sum_{j=1}^{q} \varphi_j \epsilon_{t-j} + \sum_{k=1}^{Q} \Phi_k \epsilon_{t-ks} + \sum_{j=1}^{q}\sum_{k=1}^{Q} \varphi_j \Phi_k\, \epsilon_{t-j-ks} + \sum_{j=1}^{q}\sum_{k=1}^{Q} \lambda_{jk}\, \epsilon_{t-j-ks} + \epsilon_t \quad (2)$$

Model (2)[23] is called Augmented SMA model of orders q,Q, and it may be expressed in matrix form as:

$$y_t = M_t \beta + R_t \lambda + \epsilon_t \quad (3)$$

Where, $M_t = (\epsilon_{t-1}, \dots, \epsilon_{t-q}; o; \epsilon_{t-s}, \epsilon_{t-s-1}, \dots, \epsilon_{t-s-q}; o; \dots; \epsilon_{t-Qs}, \epsilon_{t-Qs-1}, \dots, \epsilon_{t-Qs-q})$,

$\beta = (\varphi_1, \varphi_2, \dots, \varphi_q; o; \Phi_1, \varphi_1 \Phi_1, \dots, \varphi_q \Phi_1; o; \Phi_Q, \varphi_1 \Phi_Q, \dots, \varphi_q \Phi_Q)^T$,

$R_t = (\epsilon_{t-s-1}, \dots, \epsilon_{t-s-q}; o_1; \epsilon_{t-Qs-1}, \dots, \epsilon_{t-Qs-q})$, $\lambda = (\lambda_{11}, \dots, \lambda_{q1}; o_1; \lambda_{1Q}, \dots, \lambda_{qQ})^T$, and both o and o_1 are row vectors composed entirely of zeros and their order are (s-q-1) and (s-q) respectively.

If all $\lambda_{jk}(j = 1, \dots, q, k = 1, \dots, Q)$ are insignificant, model (2) is reduced to Multiplicative model as follows:

$$y_t = \sum_{j=1}^{q} \phi_j \epsilon_{t-j} + \sum_{k=1}^{Q} \Phi_k \epsilon_{t-ks} + \sum_{j=1}^{q}\sum_{k=1}^{Q} \phi_j \Phi_k\, \epsilon_{t-j-ks} + \epsilon_t \quad (4)$$

It should be noted that the model will be Non-multiplicative if at least one of λ_{jk} is significant ($\lambda_{jk} \neq 0$). Therefore, testing the multiplicativity of the SMA model is the same as testing the significance of λ_{jk} This can be accomplished by constructing confidence intervals for the parameters that distinguish between Multiplicative and Non-multiplicative models and thus testing their significance.

3 Bayesian analysis of Augmented SMA model

3.1 *Likelihood Function*

Assume that y is a sequence of observations $(y_1, y_2, \dots, y_n)$ generated from Augmented $SMA(q,Q)_s$ model (2), and denoting the random error as $\epsilon_t \sim Normal(0, a^2)$, one can obtain the likelihood function $L(\varphi, \Phi, \lambda, a^2, \epsilon_0 | y) = \ell$ by transforming ϵ_t into y_t , as follows:

$$\ell \propto (a^2)^{-\frac{n}{2}} exp\{-\frac{1}{2a^2}\sum_{t=1}^{n}\epsilon_t^2\} \tag{5}$$

$$\ell \propto \left(a^2\right)^{-\frac{n}{2}} exp\{-\frac{1}{2a^2}\sum_{t=1}^{n}(y_t - \sum_{j=1}^{q}\varphi_j\epsilon_{t-j} - \sum_{k=1}^{Q}\Phi_k\epsilon_{t-ks} - \sum_{j=1}^{q}\sum_{k=1}^{Q}\varphi_j\Phi_k\,\epsilon_{t-j-ks} - \sum_{j=1}^{q}\sum_{k=1}^{Q}\lambda_{jk}\,\epsilon_{t-j-ks})^2\Big\} \tag{6}$$

$$= (a^2)^{-\frac{n}{2}} exp\{-\frac{1}{2a^2}(y - M\beta - R\lambda)^T(y - M\beta - R\lambda)\} \tag{7}$$

Where, M is a $n \times (q + Qs)$ matrix that has t^{th} row

$$M_t = (\epsilon_{t-1}, \dots, \epsilon_{t-q}; o; \epsilon_{t-s}, \epsilon_{t-s-1}, \dots, \epsilon_{t-s-q}; o; \dots; \epsilon_{t-Qs}, \epsilon_{t-Qs-1}, \dots, \epsilon_{t-Qs-q}),$$

R is a $n \times (q + Qs - s)$ matrix that has t^{th} row

$R_t = (\epsilon_{t-s-1}, \dots, \epsilon_{t-s-q}; o_1; \epsilon_{t-Qs-1}, \dots, \epsilon_{t-Qs-q})$, and β and λ are defined in (3)

We can see from (6) that ϵ_t' is not linear in the model parameters, so the likelihood function (6) is complex concerning the parameters φ, Φ, λ , and ϵ_0, which complicates and makes the likelihood function analytically challenging. To solve this, the unknown errors can be estimated by replacing them with their estimated values through a recursive process as follows:

$$\hat{\epsilon}_t = y_t - \sum_{j=1}^{q}\hat{\varphi}_j\hat{\epsilon}_{t-j} - \sum_{k=1}^{Q}\widehat{\Phi}_k\hat{\epsilon}_{t-ks} - \sum_{j=1}^{q}\sum_{k=1}^{Q}\hat{\varphi}_j\widehat{\Phi}_k\,\hat{\epsilon}_{t-j-ks} - \sum_{j=1}^{q}\sum_{k=1}^{Q}\hat{\lambda}_{jk}\,\hat{\epsilon}_{t-j-ks} \tag{8}$$

Where $\hat{\varphi}_j, \widehat{\Phi}_k$,and $\hat{\lambda}_{jk}$ are estimates obtained using NLS through the minimization of the sum of square errors $ss(\varphi, \Phi, \lambda)$ with respect to φ, Φ and λ over the invertible region. As a result, we may approximate the likelihood function (6) as follows:

$$\ell \propto (a^2)^{-\frac{n}{2}} exp\{-\frac{1}{2a^2}\sum_{t=1}^{n}(y_t - \sum_{j=1}^{q}\varphi_j\hat{\epsilon}_{t-i} - \sum_{k=1}^{Q}\Phi_k\hat{\epsilon}_{t-ks} - \sum_{j=1}^{q}\sum_{k=1}^{Q}\varphi_j\Phi_k\,\hat{\epsilon}_{t-j-ks} - \sum_{j=1}^{q}\sum_{k=1}^{Q}\lambda_{jk}\,\hat{\epsilon}_{t-j-ks})^2\} \tag{9}$$

$$= (a^2)^{-\frac{n}{2}} exp\{-\frac{1}{2a^2}(y - \widehat{M}\beta - \hat{R}\lambda)^T(y - \widehat{M}\beta - \hat{R}\lambda)\} \tag{10}$$

Where, $\widehat{M}$ is a $n \times (q + Qs)$ matrix that has t^{th}row

$\widehat{M}_t = (\hat{\epsilon}_{t-1}, \dots, \hat{\epsilon}_{t-q}; o; \hat{\epsilon}_{t-s}, \hat{\epsilon}_{t-s-1}, \dots, \hat{\epsilon}_{t-s-q}; o; \dots; \hat{\epsilon}_{t-Qs}, \hat{\epsilon}_{t-Qs-1}, \dots, \hat{\epsilon}_{t-Qs-q}), \hat{R}$ is a $n \times (q + Qs - s)$ matrix that has t^{th} row $\hat{R}_t = (\hat{\epsilon}_{t-s-1}, \dots, \hat{\epsilon}_{t-s-q}; o_1; \hat{\epsilon}_{t-Qs-1}, \dots, \hat{\epsilon}_{t-Qs-q})$ and β and λ are defined in (3).

3.2 *Prior and Posterior Distribution.*

In the context of Augmented SMA model (2), it is assumed that the parameters φ, Φ, λ and ϵ_0 are independent a priori, given the error variance parameter a^2.

$$\xi(\varphi, \Phi, \lambda, a^2, \epsilon_0) = \xi(\varphi|a^2) \times \xi(\Phi|a^2) \times \xi(\lambda|a^2) \times \xi(\epsilon_0|a^2) \times \xi(a^2)$$
$$= N_q(\mu_\varphi, a^2\Sigma_\varphi) \times N_Q(\mu_\Phi, a^2\Sigma_\Phi) \times N_{qQ}(\mu_\lambda, a^2\Sigma_\lambda) \times N_{q+Qs}(\mu_{\epsilon_0}, a^2\Sigma_{\epsilon_0}) \times IG\left(\frac{\nu}{2}, \frac{\eta}{2}\right) \quad (11)$$

Where $N_{num}(\mu, a^2\Sigma)$ refers to the multivariate normal distribution with mean vector μ and variance-covariance matrix $a^2\Sigma$ while $Inv-Gamma(\frac{\nu}{2}, \frac{\eta}{2})$ refers to the inverse-gamma distribution with parameters $\frac{\nu}{2}$ and $\frac{\eta}{2}$. The representation of the prior distribution (11) can now be expressed in the following way:

$$\xi(\varphi, \Phi, \lambda, a^2, \epsilon_0) \propto (a^2)^{-\left(\frac{\nu+2q+Q+qQ+Qs}{2}+1\right)} exp\{-\frac{1}{2a^2}[\eta + (\varphi-\mu_\varphi)^T\Sigma_\varphi^{-1}(\varphi-\mu_\varphi) +$$
$$(\Phi-\mu_\Phi)^T\Sigma_\Phi^{-1}(\Phi-\mu_\Phi) + (\lambda-\mu_\lambda)^T\Sigma_\lambda^{-1}(\lambda-\mu_\lambda) + (\epsilon_0-\mu_{\varepsilon_0})^T\Sigma_{\epsilon_0}^{-1}(\epsilon_0-\mu_{\epsilon_0})]\} \quad (12)$$

There were multiple reasons for selecting the prior distribution (12). Firstly, it is versatile enough to be applied in various situations. Secondly, it simplifies mathematical computations, which is beneficial for practical implementation. Finally, it is at least conditionally a conjugate prior, which makes it a useful choice for the given context.[14]

In situations where there is limited or no information available regarding the unknown parameters, the Jeffreys' prior distribution can be employed, which is a special case of the normal-inverse gamma distribution when $\eta = 0, \Sigma_\varphi^{-1} = \Sigma_\Phi^{-1} = \Sigma_\lambda^{-1} = \Sigma_{\epsilon_0}^{-1} = 0$ and $\nu = -p^*$ where $p^* = 2q + Q + qQ + Qs$.

We get the joint posterior distribution $\zeta(\varphi, \Phi, \lambda, a^2, \epsilon_0|y)$ By multiplying (12) with (10) as follows:

$$\xi(\varphi, \Phi, \lambda, a^2, \epsilon_0|y)$$
$$\propto (a^2)^{-\left(\frac{n+\nu+2q+Q+qQ+Qs}{2}+1\right)} exp\{-\frac{1}{2a^2}[\eta + (\varphi-\mu_\varphi)^T\Sigma_\varphi^{-1}(\varphi-\mu_\varphi) + (\Phi-\mu_\Phi)^T\Sigma_\Phi^{-1}(\Phi-\mu_\Phi)$$
$$+(\lambda-\mu_\lambda)^T\Sigma_\lambda^{-1}(\lambda-\mu_\lambda) + (\epsilon_0-\mu_{\epsilon_0})^T\Sigma_{\epsilon_0}^{-1}(\epsilon_0-\mu_{\epsilon_0}) + (y-\widehat{M}\beta-\widehat{R}\lambda)^T(y-\widehat{M}\beta-\widehat{R}\lambda)]\} \quad (13)$$

3.3 *Full conditional distributions*

This sub-section introduces the full conditional distributions (FCDs) for the unknown parameters. The process of obtaining these FCDs involves the following steps:

1) We start with the joint posterior distribution, which represents the distribution of all the unknown parameters given the observed data.
2) We identify the specific unknown parameter for which we want to derive the conditional distribution.
3) Next, we group the terms in the joint posterior distribution that depend on the chosen parameter. This step entails isolating the relevant terms and excluding those that do not involve the parameter of interest.
4) To ensure that the resulting distribution is a proper density function, we determine the appropriate normalizing constant.
5) We then simplify and manipulate the grouped terms to obtain the functional form of the conditional distribution.
6) The characteristics of the grouped terms determine the nature of the conditional distribution. In our case, the study reveals that all the conditional posteriors are either normal or inverse gamma distributions.

By following these steps, we can derive the full conditional distributions for each unknown parameter.

3.3.1 FCD for φ

FCD for φ is

$$\varphi^r \sim \zeta(\varphi^r|y, \Phi^{r-1}, \lambda^{r-1}, (a^2)^{r-1}, \epsilon_0^{r-1}) = N_q(\mu_\varphi^*, H_\varphi^*)$$

where, $\mu_\varphi^* = (\Sigma_\varphi^{-1} + A^TA)^{-1}(\Sigma_\varphi^{-1}\mu_\varphi + A^Ty - A^TK\Phi - A^T\widehat{R}\lambda)$, $H_\varphi^* = a^2(\Sigma_\varphi^{-1} + A^TA)^{-1}$

A refers to a $n \times q$ matrix with $A_{tj} = (\hat{\epsilon}_{t-j} + \sum_{k=1}^{Q}\Phi_j\hat{\epsilon}_{t-j-ks})$ and K refers to a $n \times Q$ matrix with $K_{tk} = (\hat{\epsilon}_{t-ks})$.

3.3.2 FCD for Φ

FCD for Φ is

$\Phi^{r} \sim \zeta(\Phi^{r}|y,\varphi^{r},\lambda^{r-1},(a^{2})^{r-1},\epsilon_{0}^{r-1}) = N_{Q}(\mu_{\Phi}^{*}, H_{\Phi}^{*})$

Where, $\mu_{\Phi}^{*} = (\Sigma_{\Phi}^{-1} + W^{T}W)^{-1}(\Sigma_{\Phi}^{-1}\mu_{\Phi} + W^{T}y - W^{T}Z\varphi - W^{T}\hat{R}\lambda)$, $H_{\Phi}^{*} = a^{2}(\Sigma_{\Phi}^{-1} + W^{T}W)^{-1}$

W refers to a $n \times Q$ matrix with $W_{tk} = (\hat{\epsilon}_{t-ks} + \sum_{j=1}^{q}\varphi_{j}\hat{\epsilon}_{t-j-ks})$ and Z refers to a $n \times q$ matrix with $Z_{tj} = (\hat{\epsilon}_{t-j})$.

3.3.3 FCD for λ

FCD for λ is

$\lambda^{r} \sim \zeta(\lambda^{r}|y,\varphi^{r},\Phi^{r},(a^{2})^{r-1},\epsilon_{0}^{r-1}) = N_{qQ}(\mu_{\lambda}^{*}, H_{\lambda}^{*})$

where,

$\mu_{\lambda}^{*} = (\Sigma_{\lambda}^{-1} + \hat{R}^{T}\hat{R})^{-1}(\Sigma_{\lambda}^{-1}\mu_{\lambda} + \hat{R}^{T}y - \hat{R}^{T}\widehat{M}\beta)$
$H_{\lambda}^{*} = a^{2}(\Sigma_{\lambda}^{-1} + \hat{R}^{T}\hat{R})^{-1}$.

3.3.4 FCD for a^{2}

FCD for a^{2} is

$(a^{2})^{r} \sim \xi((a^{2})^{r}|y,\varphi^{r},\Phi^{r},\lambda^{r},\epsilon_{0}^{r-1}) \sim Inv-Gamma(\frac{v^{*}}{2},\frac{\eta + n(s^{2})^{r}}{2})$

Where, $v^{*} = n + v + 2q + Q + qQ + Qs$ and $ns^{2} = (\varphi - \mu_{\varphi})^{T}\Sigma_{\varphi}^{-1}(\varphi - \mu_{\varphi}) + (\Phi - \mu_{\Phi})^{T}\Sigma_{\Phi}^{-1}(\Phi - \mu_{\Phi}) + (\lambda - \mu_{\lambda})^{T}\Sigma_{\lambda}^{-1}(\lambda - \mu_{\lambda}) + (\epsilon_{0} - \mu_{\epsilon_{0}})^{T}\Sigma_{\epsilon_{0}}^{-1}(\epsilon_{0} - \mu_{\epsilon_{0}}) + (y - \widehat{M}\beta - \hat{R}\lambda)^{T}(y - \widehat{M}\beta - \hat{R}\lambda)$

3.3.5 FCD for ϵ_{0}

Utilizing model (2), the equations that describe the elements of ϵ_{0} can be expressed as follows:

$$y_{q+Qs} = L\epsilon_{0} + \Lambda\,\epsilon_{q+Qs}$$

Where, L and Λ are $(q+Qs)\times(q+Qs)$ matrices, $y_{q+Qs} = (y_{1}, y_{2}, \dots, y_{q+Qs})^{T}$ and $\epsilon_{q+Qs} = (\epsilon_{1}, \epsilon_{2}, \dots, \epsilon_{q+Qs})$ which has Normal-distribution with zero mean and variance $a^{2}I_{q+Qs}$ where I_{q+Qs} is the identity matrix of order $(q+Qs)$. The calculation of FCD is performed utilizing the results of linear regression and standard Bayesian methods

$\epsilon_{0}^{r} \sim \zeta(\epsilon_{0}^{r}|y,\varphi^{r},\Phi^{r},\lambda^{r},(a^{2})^{r}) = N_{q+Qs}(\mu_{\epsilon_{0}}^{*}, H_{\epsilon_{0}}^{*})$

Where,

$\mu_{\epsilon_{0}}^{*} = (\Sigma_{\epsilon_{0}}^{-1} + L^{T}(\Lambda\Lambda^{T})^{-1}L)^{-1}(\Sigma_{\epsilon_{0}}^{-1}\mu_{\epsilon_{0}} + L^{T}(\Lambda\Lambda^{T})^{-1}y_{q+Qs)}$, $H_{\epsilon_{0}}^{*} = a^{2}(\Sigma_{\epsilon_{0}}^{-1} + L^{T}(\Lambda\Lambda^{T})^{-1}L)^{-1}$

4 The Proposed Gibbs Sampler

For Augmented SMA model (2), the Gibbs sampling algorithm is as follows:

Step 1: Define initial values for the parameters $\varphi^{0}, \Phi^{0}, \lambda^{0}, (a^{2})^{0}$ and ${\epsilon_{0}}^{0}$ and set the iteration counter to one (r=1). Initial estimates can be obtained by fitting SMA model using Non-linear least squares and setting the initial values of the parameters ϵ_{0} to zero.

Step 2: Repeatedly calculate the residuals utilizing the given formula (8)**.**

Step 3: Get the FCDs for the parameters.

Step 4: Set the iteration counter to one (r=1) and run the simulation.

$\varphi^{r} \sim \zeta(\varphi^{r}|y,\Phi^{r-1},\lambda^{r-1},(a^{2})^{r-1},\epsilon_{0}^{r-1})$
$\Phi^{r} \sim \zeta(\Phi^{r}|y,\varphi^{r},\lambda^{r-1},(a^{2})^{r-1},\epsilon_{0}^{r-1})$
$\lambda^{r} \sim \zeta(\lambda^{r}|y,\varphi^{r},\Phi^{r},(a^{2})^{r-1},\epsilon_{0}^{r-1})$
$(a^{2})^{r} \sim \xi((a^{2})^{r}|y,\varphi^{r},\Phi^{r},\lambda^{r},\epsilon_{0}^{r-1})$
$\epsilon_{0}^{r} \sim \zeta(\epsilon_{0}^{r}|y,\varphi^{r},\Phi^{r},\lambda^{r},(a^{2})^{r})$

Step 5: Increase the iteration counter by 1 $(r = r + 1)$ and return to step 4.

The process of obtaining the next value of the Markov chain $\{\varphi^{r+1}, \Phi^{r+1}, \lambda^{r+1}, (a^2)^{r+1}, \epsilon_0^{r+1}\}$ is performed by iteratively drawing samples from each of the FCDs, while updating the conditions at each iteration of the algorithm. This process is iteratively carried out for a significant number of iterations, and the convergence is verified. When the chain converges, the resulting values from the simulations $\{\varphi^{r+1}, \Phi^{r+1}, \lambda^{r+1}, (a^2)^{r+1}, \epsilon_0^{r+1}, \forall r \succ n_0\}$ are treated as samples from the joint posterior distribution, and the average of these samples is used to obtain the posterior estimates for the parameters. The method for monitoring the convergence of the Gibbs sampling sequence has been widely researched and documented in the literature, as seen in references [24,25,26].

5 Simulation study

5.1 *Simulation design*

A simulation study is presented to assess the precision of our suggested Bayesian methodology for Augmented SMA models.

One thousand sets of time series data are created by using two Augmented SMA models with different values of λ. Simulation design details are presented in Table (1), which comprises actual parameter values, model variance, seasonal periods, and sample size.

Table 1: Design of the Simulation

Model	φ	Φ	λ	a^2	s	n
I	0.4	0.6	0,0.3 and 0.4	1	4	400
II	-0.5	-0.3	0,0.1,0.3 and 0.4	0.5	12	300

5.2 *Simulation steps*

To assess the accuracy of our proposed methodology for Augmented SMA model, we conducted a simulation study following the steps outlined below.

1) Creation of 1000 time series datasets: The study begins by generating 1000 time series datasets from the Augmented SMA models.

2) Bayesian analysis with non-informative prior: The Bayesian analysis is conducted by employing a non-informative prior distribution for the parameters φ, Φ, λ *and* a^2 by sitting , $\Sigma_{\varphi}^{-1} = \Sigma_{\Phi}^{-1} = \Sigma_{\lambda}^{-1} = 0, \eta = S^2$ and $\nu = 3$, while the initial errors ϵ are assumed to follow normal distribution with zero mean and variance-covariance matrix $a^2 I_{q+Qs}$.

3) Selection of starting values: The starting values for φ, Φ, λ *and* $a^2$are determined based on estimates derived from the NLS technique applied to the Augmented SMA model. As for the starting values of ϵ, they are assumed to be zero.

4) Execution of Gibbs sampler: The Gibbs sampler is executed 31,000 times for each dataset. The first 1,000 outputs are discarded as burn-in, and subsequently, every 10th value from the remaining 30,000 outputs is retained to produce an approximately independent sample.

5) Computation of posterior estimates: Summary statistics, such as the mean μ, standard deviation sd, median $med.$, and 95% credible interval (CI) limits, are directly computed from the output of the Gibbs sampler as posterior estimates of the model parameters.

6) Calculation of average summary statistics: The averages of the summary statistics are computed and reported based on the posterior outcomes got from the 1000 generated time series datasets.

Further Discussion on Results and Convergence Diagnostics: Examination of the outcomes obtained from the proposed Gibbs sampling algorithm and their convergence diagnostics is presented in a subsequent sub-section.

5.2 *Simulation outcomes*

Tables (2) to (4) exhibit the results of Model I based on the 1000 time series datasets generated. These tables comprise the Bayesian estimates of the model parameters along with their true values and the 0.025 and 0.975 percentiles of the simulation draws used to create a 95% credible interval for each parameter.

Table 2: Bayesian analysis outcomes for Model I with $\lambda = 0$

par	True value	μ	Sd.	LL 95% CI	Med.	UL 95% CI
φ	0.4	0.400	0.050	0.302	0.400	0.498
Φ	0.6	0.602	0.050	0.503	0.602	0.700
λ	0	-0.004	0.062	-0.127	-0.004	0.117
a^2	1	0.997	0.071	0.868	00.994	1.144

Table 3: Bayesian analysis outcomes for Model I with $\lambda = 0.3$

par	True value	μ	Sd.	LL 95% CI	Med.	UL 95% CI
φ	0.4	0.402	0.050	0.304	0.402	0.500
Φ	0.6	0.597	0.050	0.499	0.597	0.695
λ	0.3	0.293	0.062	0.169	0.293	0.413
a^2	1	1.001	0.071	0.871	0.998	1.149

Table 4: Bayesian analysis outcomes for Model I with $\lambda = 0.4$

par	True value	μ	Sd.	LL 95% CI	Med.	UL 95% CI
φ	0.4	0.404	0.050	0.306	0.404	0.502
Φ	0.6	0.592	0.050	0.494	0.592	0.690
λ	0.4	0.388	0.062	0.265	0.389	0.509
a^2	1	1.006	0.071	0.875	1.003	1.155

Based on the information presented in Tables (2) to (4), the estimates of the model parameters exhibit remarkable proximity to their actual values. Additionally, the average of the 95% CI limits encompasses the actual value for each parameter. To assess convergence diagnostics, we generate a single time series dataset using Model I with a value of $\lambda = 0.3$. In Figure 1, we present trace plots (traceplot) for each parameter in Model I, accompanied by their corresponding marginal posteriors. Additionally, Table (5) displays autocorrelations (Autocorr) and Raftery-Lewis[25,26] diagnostics, while Table (6) presents Geweke[24] diagnostics.

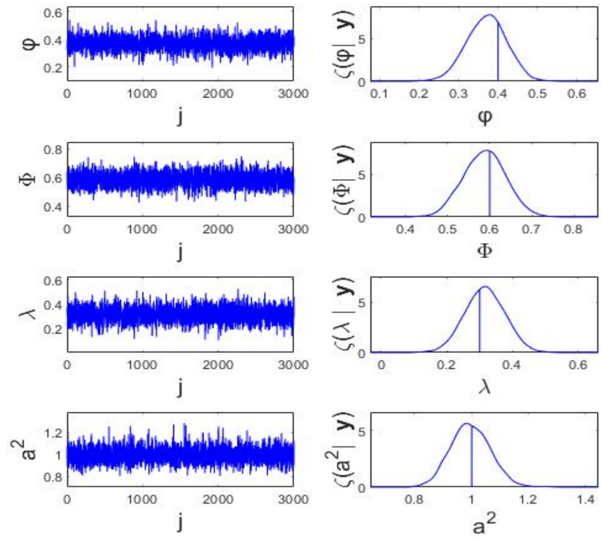

Fig.1: traceplot and marginal posterior densities of Model I when $\lambda = .3$

As depicted in Figure 1, the Gibbs algorithm's posterior draws remain stable and vary around the actual values. The marginal posterior densities display that the actual value of each parameter lies within the 95% CI.

Table 5: Autocorre and Raftery.Lewis.Diag for Model I when $\lambda = 0.3$

Par.	Autocorr			
	1ag1	lag5	lag10	lag50
φ	0.003	0.013	0.031	-0.025
Φ	-0.046	-0.027	0.005	-.007
λ	0.022	0.011	0.022	-00.015
a^2	-0.008	-0.026	-0.022	-0.017
Par.	Raftery.Lewis.Diag			
	Burn-in	Total N	N-min	I-stat
φ	2	892	937	0.952
Φ	2	892	937	0.952
λ	2	892	937	0.952
a^2	2	892	937	0.952

Table (5) indicates that there is no convergence issue as the autocorrelations for each parameter at lags 1, 5, 10, and 50 are low. This is further provided by the Raftery.Lewis.Diag, specifically the I-stat value, which is approximately 1, indicating a value lower than 5.

Table 6: Geweke.diag for Model I when $\lambda = 0.3$

Par.	φ	Φ	λ	a^2
NSE iid	0.001	0.001	0.001	0.001
RNE iid	1	1	1	1
NSE 4%	0.001	0.001	0.001	0.001
RNE 4%	0.6955	1.2767	1.1424	1.1622
NSE 8%	0.0010	0.0008	0.0010	0.0011
RNE 8%	0.7809	1.4478	1.3067	1.3785
NSE 15%	0.0009	0.0006	0.0009	0.0011
RNE 15%	1.054	1.968	1.517	1.281
χ^2	0.584	0.607	0.600	0.578

Table (6) indicates that there is no convergence issue as the χ^2 probabilities indicate that the hypothesis of equal means cannot be rejected, and there are no significant differences in the NSE estimates. Additionally, the RNE estimates being close to one suggests that the output sample has an independent and identically distributed (iid) nature.

The outcomes of a Bayesian analysis performed on Model II are presented in Tables (7) to (9). These results demonstrate that Model II yields comparable findings to Model I, thereby confirming the efficacy and accuracy of the proposed Gibbs sampling algorithm.

Table 7: Bayesian analysis outcomes for Model II with $\lambda = 0$

Par.	True value	μ	Sd.	LL 95% CI	Med.	UL 95% CI
φ	-0.5	-0.493	0.058	-0.606	-0.493	-0.381
Φ	-0.3	- 0.297	0.059	-0.412	-0.297	-0.183
λ	0	-0.002	0.068	-0.136	-0.001	0.130
a^2	0.5	0.500	0.041	0.426	0.497	0.586

Table 8: Bayesian analysis outcomes for Model II with $\lambda = 0.4$

Par.	True value	μ	Sd.	LL 95% CI	Med.	UL 95% CI
φ	-0.5	-0.490	0.058	-0.604	-0.490	-0.378
Φ	-0.3	-0.293	0.058	-0.408	-0.293	-0.179
λ	0.4	0.395	0.067	0.262	0.396	0.527
a^2	0.5	0.504	0.041	0.429	0.502	0.591

Table (9): Bayesian analysis outcomes for Model II with $\lambda = -0.3$

Par.	True value	μ	Sd.	LL 95% CI	Med.	UL 95% CI
φ	-0.5	-0.480	0.058	-0.593	-0.480	-0.368
Φ	-0.3	-0.286	0.058	-0.400	-0.286	-0.172
λ	-0.3	-0.295	0.067	-0.427	-0.294	-0.164
a^2	0.5	0.506	0.041	0.431	0.504	0.593

6 Real-world application

To demonstrate the suggested Bayesian analysis of Augmented SMA model, one of the frequently cited time series examples in the literature was employed as an illustration. This series was chosen because it was really modelled using the Multiplicative Seasonal Moving Averages model. This series is known as the "Airline Series" [23].

The airline series comprises 144 monthly observations of US airline passengers from 1949 to 1960. The series exhibits both trend and seasonal patterns, and its variance increases over time, making it non-stationary Figure (2). However, by taking the natural logarithm of the series and using non-seasonal and seasonal differences, the series becomes stationary Figures (3) to (5). Box et al. [23] identified this logged and differenced series as Multiplicative $SMA(1,1)_{12}$ without testing the multiplicativity of this series, i.e., without testing the significance of λ, the parameter that distinguishes between Multiplicative and Non-multiplicative models, so we estimate Augmented SMA for this series by using the proposed Bayesian analysis and then test the significance of parameter λ by constructing the credible interval of this parameter as shown below. The selection of hyper-parameters and starting values follows the same way as that used in the simulation study.

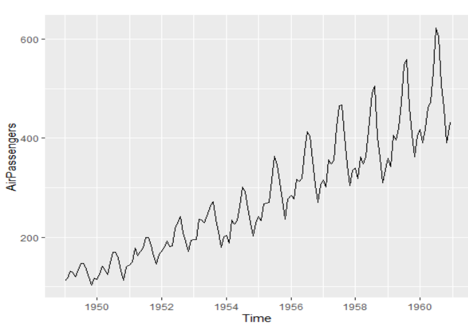

Fig. 2: monthly values of numbers of passengers.

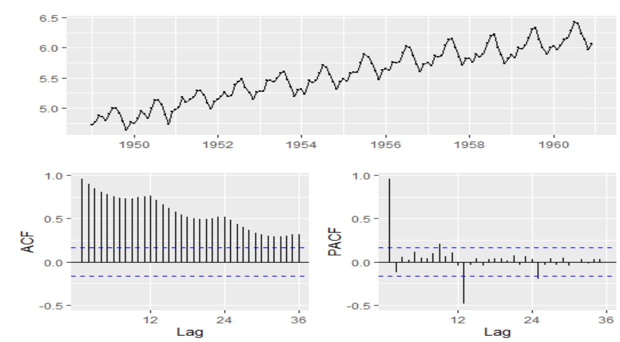

Fig. 3: "ACF" and "PACF" for log (airline data).

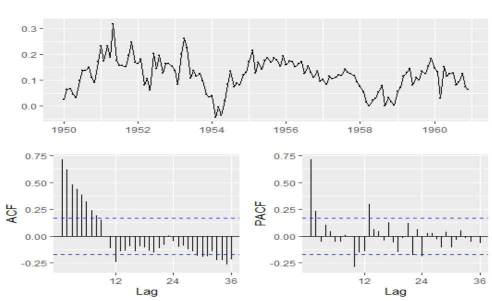

Fig. 4: "ACF" and "PACF" for log(airline data) after taking seasonal differences.

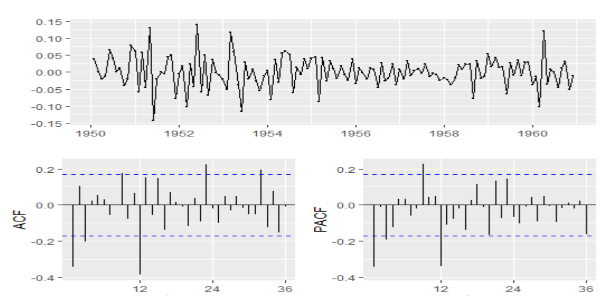

Fig. 5: "ACF" and "PACF" for log (airline data) after taking first differences for seasonal differences.

Table (10) presents the Bayesian analysis outcomes for the logged and differenced airline series. Trace plots and marginal densities of the logged and differenced airline series are displayed in figure (6).

Table 10: Bayesian analysis outcomes for the logged and differenced airline series

Par.	Mean	Sd.	LL 95% CI	Med.	UL 95% CI
φ	-0.3694	0.0888	-0.5499	-0.3698	-0.1937
Φ	-0.6040	0.0952	-0.7894	-0.6028	-0.4244
λ	0.0128	0.1106	-0.2046	0.0151	0.2264
a^2	0.0014	0.0002	0.0011	0.0014	0.0018

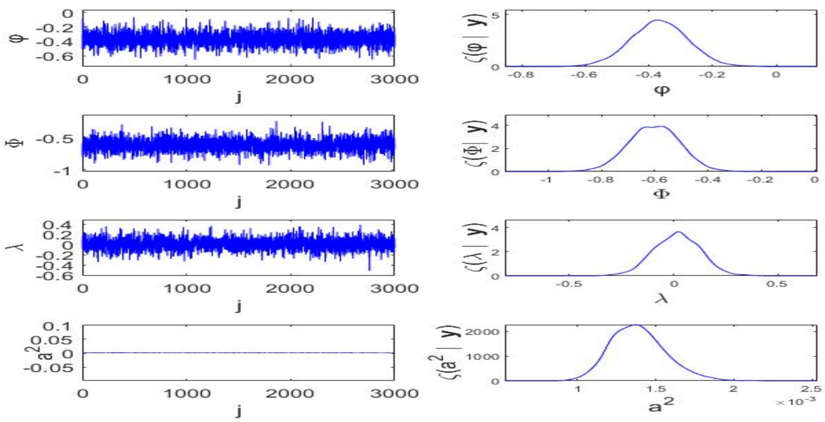

Fig. 6: traceplot and marginal posterior densities for the logged and differenced airline series

From table (10) the 95% credible interval for the parameter λ contains zero, which means that we accept that the used model is Multiplicative SMA model, and this agrees with the reported statistical results in the literature [23].

7 Conclusion

This study demonstrates that the full conditional distributions of Augmented SMA model adhere to standard probability distributions. Specifically, the full conditionals of the parameters φ, Φ, λ and ϵ_0 follow multivariate normal distributions, while the full conditional of the parameter a^2 is an inverse-gamma distribution.

Taking advantage of the standard nature of the full conditional distributions, a Bayesian method was developed to estimate the parameters of Augmented SMA model. This method utilizes the simple MCMC Gibbs sampling algorithm.

To evaluate the accuracy of the proposed algorithm, both a simulation study and a real-world application were employed. The stability of the proposed Gibbs sampling algorithm was demonstrated by several convergence measures.

References

[1] J.F. Monahan, Fully Bayesian analysis of ARMA time series models, *Journal of Econometrics*, **21**, 307-331 (1983).

[2] P. Newbold, Bayesian estimation of Box Jenkins transfer function noise models, *JRSSB* ,**35**, 323-336 (1973).

[3] A. Zellner and R. Reynolds, Bayesian analysis of ARMA models, Presented at the Sixteenth Seminar on Bayesian Inference in Econometrics **23**, 1-5 (1978).

[4] L.D. Broemeling, and S. Shaarawy, Bayesian inferences and forecasts with moving average processes, *Communications in Statistics: Theory and Methods*, **13**, 1871–1888 (1984).

[5] A. Amin, Bayesian Identification of Double Seasonal Autoregressive Time Series Models, *Communications in Statistics: Simulation and Computation* ,**48**, 2501-2511 (2018a).

[6] A. Amin, Kullback-Leibler Divergence to Evaluate Posterior Sensitivity to Different Priors for Autoregressive Time Series Models, *Communications in Statistics: Simulation and Computation*, **48**, 1277-1291 (2018b).

[7] S. Chib, and E. Greenberg, Bayes inference in regression models with ARMA(p,q) errors, *Journal of Econometrics*, **64**, 183-206 (1994).

[8] J. Marriott, N. Ravishanker, A. Gelfand, and J. Pai, Bayesian analysis of ARMA processes: Complete sampling based inference under full likelihoods. In Bayesian Statistics and Econometrics: Essays in honor of Arnold Zellner, D., Berry, K., Chaloner, and J. Geweke," (eds). New York, Wiley,(1996).

[9] G. Barnett, R. Kohn, and S. Sheather, Bayesian estimation of an autoregressive model using Markov Chain Monte Carlo, *Journal of Econometrics*, **74**, 237-254 (1996).

[10] G. Barnett, R. Kohn, and S. Sheather, Robust Bayesian estimation of an autoregressive moving-average models, *Journal of Time Series Analysis*, **18**, 11-28 (1997).

[11] M. A. Ismail, Bayesian Analysis of Seasonal Autoregressive Models, *Journal of Applied Statistical Science*, **12**, 123-136 (2003a).

[12] M. A. Ismail, Bayesian Analysis of Seasonal Moving Average Model: A Gibbs Sampling Approach, *Japanese Journal of Applied Statistics*, **32**, 61-75 (2003b).

[13] A. Amin, *Bayesian Inference for Seasonal ARMA Models: A Gibbs Sampling Approach*, unpublished master's Thesis, Cairo university, Egypt, (2009).

[14] M.A. Ismail, and A.A. Amin, Gibbs Sampling For SARIMA Models, *Pak. J. Statist*, **30**, 153-168 (2014).

[15] A.A. Amin, and M.A. Ismail, Gibbs sampling for double seasonal autoregressive models, *Communications for Statistical Applications and Methods*, **22**, 557-573 (2015).

[16] A. Amin, Bayesian inference for double seasonal moving average models: A Gibbs sampling approach, *Pakistan Journal of Statistics and Operation Research*, **13**, 483–499 (2017a).

[17] A. Amin, *Gibbs sampling for double seasonal ARMA models*, Proceedings of the 29th Annual International Conference on Statistics and Computer Modeling in Human and Social Sciences. At: FEPS, Cairo University, Egypt, (2017b).

[18] A.A. Amin, Bayesian analysis of double seasonal autoregressive models, *Sankhya B*, **82**, 328-352 (2020).

[19] A.A. Amin, Gibbs sampling for Bayesian estimation of triple seasonal autoregressive models, *Communications in Statistics-Theory and Methods*, 1-20 (2022).

[20] A. Amin, Gibbs sampling for Bayesian prediction of SARMA processes, *Pakistan Journal of Statistics and Operation Research*,**15**, 397-418 (2019).

[21] L. D. Broemeling, *Bayesian Analysis of Linear Models*, 1st ed, Marcel Dekker Inc., New York, 1- 472, (1985).

[22] R. E. McCulloch, and R. S. Tsay, Bayesian Analysis of Autoregressive Time Series via the Gibbs Sampler, *Journal of Time Series Analysis*, **15**, 235-250 (1994).

[23] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, and G.M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed, Join Wiley & Sons, Inc., New York, 1-712, (2016).

[24] J. Geweke, Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments, In J. M. Bernardo, J. O. Berger, and A. P. Dawid, ed., Bayesian Statistics 4. USA: Oxford University Press, (1992).

[25] A. E. Raftery, and S. Lewis, One long run with diagnostics: Implementation strategies for Markov Chain Monte Carlo, *Statistical Science*, **7**,493-497 (1992).

[26] A. E. Raftery, S. Lewis, The number of iterations, convergence diagnostics and generic Metropolis algorithms. In W. R. Gilks, D. J. Spiegelhalter, and S. Richardson, ed., Practical Markov Chain Monte Carlo. London: Chapman and Hall, (1995).