



A Bayesian Time Series Analysis of Annual Inflation Growth Percent in Egypt

Sherif S. Ali ¹, Emad E. A. Soliman ^{1*}

¹Department of Statistics, King Abdulaziz University, Jeddah 21589, Saudi Arabia

* Corresponding Email: ehgazzy@kau.edu.sa

Received: 10 August 2026 / Revised: 20 September 2026 / Accepted: 28 September 2026 / Published online: 11 October 2026

This is an Open Access article published under the Creative Commons Attribution 4.0 International (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0/>). © SCOPUA Journal of Applied Statistical Research published by SCOPUA (Scientific Collaborative Online Publishing Universal Academy). SCOPUA stands neutral with regard to jurisdictional claims in the published maps and institutional affiliations.

ABSTRACT

The current manuscript aims to employ the Bayesian approach to analyze a real data time series in order to understand its functional form and give accurate forecasts for its future observations. The time series considered in the study is the data set of annual inflation growth percent in Egypt during the period from 1960 to 2025. The series is modeled and its observations are predicted using a Bayesian approach. Some accuracy measures are used such as Root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The results showed that, ARIMA(1,1,1) model is fitted by the proposed Bayesian technique and passes all the diagnostic checking tests and can be used efficiently in the forecasting process. The model was used to give point and interval forecasts for the succeeding three years.

Keywords: Autoregressive moving average processes; Bayesian time series analysis; Inflation rates; Mean absolute error; Mean absolute percentage error

1. Introduction

Inflation is a crucial macroeconomic indicator for evaluating the performance and stability of an economy. It is characterized by a sustained rise in the general price level of goods and services over time, resulting in a decrease in the purchasing power of money. Romer (2018) has defined inflation as an increase in the average price of goods and services in terms of money, and this increase is measured using a price index, typically a consumer price index (CPI). As prices rise, each unit of currency buys fewer goods and services, so inflation corresponds to a reduction in the purchasing power of money. In other words, inflation is a general and ongoing rise in the level of prices in an entire economy. Inflation does not refer to a change in relative prices. A relative price change occurs when, for instance, the price of tuition rises while the price of laptops falls. Inflation, by contrast, means there is upward pressure on prices across most markets in the economy simultaneously, and the rise must be sustained. (Greenlaw et al. (2026))

In Egypt, inflation is one of the most important macroeconomic variables due to its direct effects on household welfare, production costs, investment choices, exchange rate stability, and monetary policy. The Egyptian economy has faced multiple instances of inflationary pressure stemming from economic reforms, fluctuations in exchange rates, subsidy reforms, and external shocks. Notably, the liberalization of the exchange rate in 2016, along with subsequent adjustments in monetary policy and global economic disruptions, led to significant variations in inflation dynamics. These structural transformations have heightened the necessity for dependable statistical models that can accurately depict the stochastic behavior of inflation and generate precise forecasts to assist policymakers in formulating effective monetary and fiscal strategies. (See Elnagger and Richter (2022))

Time series analysis offers a robust statistical framework for modeling inflation data. The observations are collected sequentially over time. Moreover, they display autocorrelation, persistence, and potential non-stationarity.

One of the most prevalent stochastic models is the autoregressive moving average (ARMA) model, which was introduced by Box and Jenkins in 1970. ARMA models integrate autoregressive (AR) and moving average (MA) components to depict the autocorrelation structure within stationary time series. Due to their simplicity, interpretability, and forecasting capabilities, ARMA models have been widely utilized in the modeling of inflation rates, exchange rates, stock prices, and

various macroeconomic indicators. Numerous empirical investigations have demonstrated that suitably chosen ARMA models frequently yield satisfactory forecasts for inflation rates, particularly after transforming the series to attain stationarity.

The theoretical and methodological dimensions of non-Bayesian autoregressive moving average (ARMA) processes are extensive, as documented in works by Box and Jenkins (1970), Granger and Newbold (1977), Chatfield (1980), Harvey (1981), Priestley (1981), Abraham and Ledolter (1983), Bowerman and O'Connell (1987), Wei (2005), Liu (2009), Box et al. (2016) and Chatfield (2019). Indeed, the methodology proposed by Box and Jenkins for the identification, estimation, diagnostic checking, and forecasting of ARMA processes stands out as the most advantageous non-Bayesian approach for modeling and predicting time series data.

The analysis of ARMA(p,q) generally represents a specific instance of general linear models, where the response or dependent variable is linearly regressed on a finite set of p previous values of the process, as well as a finite set of q previous values of the random disturbances. In many practical applications, data can be effectively modeled using ARMA(p,q) processes.

Nevertheless, the Box-Jenkins methodology presents significant disadvantages. Its identification technique lacks objectivity and necessitates meticulous scrutiny of the raw data. Furthermore, the time series analyst must possess a substantial dataset to effectively identify a suitable model. Additionally, conducting multivariate time series analysis through the non-Bayesian approach is quite challenging and demands a highly skilled time series analyst (Tiao and Box (1981)).

Conversely, the Bayesian analysis of ARMA processes presents certain benefits compared to the Box-Jenkins approach. Bayesian methods can effectively execute the four stages of time series analysis, addressing issues related to identification, estimation, diagnostic checking, and forecasting. Furthermore, the Bayesian framework offers time series analysts a structured and cohesive means to integrate prior information into their analyses, potentially resulting in precise outcomes even with small sample sizes.

Bayesian time series analysis has been evolving, with the majority of its advancements taking place over the last few decades. Although there are no textbooks solely dedicated to Bayesian time series, works by Zellner (1971) and Broemeling (1985) provide an introduction to the topic, particularly concerning certain ARMA processes and econometric models. Nonetheless, two primary Bayesian methodologies for time series analysis can be identified. The first is referred to as the numerical approach, while the second is termed the analytical approach.

The body of literature concerning the numerical Bayesian approach is advancing rapidly. A significant contribution to time series analysis from a numerical Bayesian perspective was made by Monahan (1983), who employed a numerical integration method to carry out the identification, estimation, and forecasting stages of low order ARMA(p,q) processes. Similarly, Lahif (1980) examined the first order autoregressive process in a manner akin to Monahan's work with ARMA models. The assumption of stationarity for the process was established by applying a beta distribution to the model coefficient within the interval (-1,1).

Furthermore, the Markov Chain Monte Carlo (MCMC) simulation, particularly the Gibbs sampling algorithm, has been employed to facilitate Bayesian time series analysis. MCMC methods have been utilized for non-seasonal autoregressive processes by researchers such as Albert and Chib (1993), Chib (1993), and De-Alba (1993). Under the assumptions of stationarity and invertibility, Chib and Greenberg (1994) introduced the Gibbs sampling technique to examine regression models characterized by ARMA-type errors, utilizing a prior distribution for the initial observations and errors. With a prior distribution for the initial values, Barnett et al. (1995) formulated a Bayesian estimation method for non-seasonal autoregressive processes through the Gibbs sampling algorithm. Additionally, Marriot et al. (1996) applied the Gibbs sampling technique to investigate non-seasonal ARMA processes. However, their methodology has been found to be inefficient when outliers are present.

Philipp (2006) introduced a Bayesian method for determining the orders of ARMA processes and estimating their coefficients through the MCMC algorithm, presuming a prior distribution for the initial values. His approach is noted for its computational efficiency and relative insensitivity to additive outliers. In a similar vein, Young and Pettit (2006) employed the Gibbs sampling algorithm to compute Bayes factors, which assist in selecting the order of an autoregressive model, while also assessing the impact of a specific set of observations on these factors. Readers are referred to the publications of Chib and Greenberg (1994), Marriott et al. (1996) and Amin (2019) for additional information on MCMC approaches.

In the realm of analysis, as opposed to a solely numerical perspective, Zellner (1971) established the posterior and predictive distributions for first and second order autoregressive processes utilizing Jeffrey's improper prior density. Newbold (1973) introduced an approximate Bayesian analysis for transfer function models, of which an ARMA process is a specific instance. Harison and Stevens (1976) provided a comprehensive review of time series that can be examined through a Bayesian framework. Additionally, a significant advancement in the field of autoregressive processes was made by Chow (1975), who determined the moments of the joint predictive distribution for future observations.

Diaz and Farah (1981) proposed a Bayesian approach for determining the order of autoregressive processes by deriving the posterior probability mass function related to the order. Following this, Shaarawy and Broemeling (1984), along with Broemeling and Shaarawy (1988), enhanced Bayesian techniques for the identification, estimation, and forecasting phases, based on an approximation of the posterior distribution of the coefficients. In their initial research, they developed a normal gamma posterior analysis specifically tailored for pure moving average processes. In their later work, they demonstrated how the posterior and predictive distributions of ARMA processes could be employed to carry out the four stages of time series analysis. The results from the first study were subsequently broadened to include bilinear models by Chen (1992). The numerical efficiency of their approximation was examined by Shaarawy and El-Shawadfy (1994) in relation to the estimation and forecasting issues associated with ARMA processes. Additionally, Shaarawy and Ali (2003) created a direct Bayesian method for determining the orders of seasonal autoregressive processes. Moreover, Shaarawy et al. (2007) developed an approximate Bayesian strategy aimed at identifying moving average models. Extensions to multivariate seasonal and non-seasonal ARMA were conducted by Shaarawy and Ali (2008), Shaarawy and Ali (2015), Ali (2015), Shaarawy (2023), Albassam et al. (2023), Shaarawy et al. (2024), and Soliman and Ali (2026).

The existing body of literature on inflation modeling in Egypt is extensive and has employed a variety of econometric and statistical methodologies. Among these studies, El-Sheikh and Kandil (2016) Evaluate monthly historical inflation data from 1996 to 2012 to account for calendar-based cycles. The paper isolates SARIMA (1,1,0) $\times$ (0,0,1)₁₂ as the optimal fit to account for recurring annual seasonal spikes in Egypt. Moreover, Omar et al. (2022) used the Box-Jenkins methodology to select a univariate ARIMA(1,1,1) model to forecast the inflation in Egypt (2019-2022).

Driven by these factors, the current research employs Bayesian analysis of ARMA models to examine the inflation rate in Egypt. Bayesian identification is utilized to derive posterior distributions of the model orders. Bayesian estimation is also utilized to derive posterior distributions of the model parameters. Diagnostic check is employed to assess model adequacy. And a Bayesian prediction technique produces probabilistic forecasts for future inflation rates. The results of this research are anticipated to enhance the existing literature on Bayesian time series analysis and to furnish policymakers with statistically robust forecasts that aid in the formulation of monetary policy and economic planning in Egypt.

This manuscript's next portions are arranged as follows: ARMA models and their basic characteristics are presented in Section 2. The Bayesian tools and the four phases of Bayesian time series analysis, Identification, Estimation, Diagnostic Check, and Prediction, are briefly explained in Section 3. The application to Egypt's annual inflation growth percentage and the numerical results are shown in section 4. Finally, the study's conclusion is presented in Section 5.

2. Autoregressive Moving Average Models

An Autoregressive Moving Average model, referred to as the ARMA(p,q) model, can be expressed as follows (see Box et al. (2016))

$$\phi(B)y_t = \theta(B)\varepsilon_t \quad (1)$$

where B represents the back shift operator, which is defined as follows: $B^r y_t = y_{t-r}$ and

$$\begin{aligned} \phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, & \phi_i &\in \mathbb{R}, \\ \theta(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q, & \theta_i &\in \mathbb{R}. \end{aligned}$$

The ARMA(p,q) model can be expressed in an equivalent form as

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

where,

y_t is the t^{th} time series observation, $t = 1, 2, \dots, n$.

p and q are the numbers of AR and MA coefficients, respectively.

ϕ_i 's and θ_j 's are the AR and MA coefficients, respectively, where $i = 1, 2, \dots, p$ and

$$j = 1, 2, \dots, q.$$

ε_t is the t^{th} unobserved error term, $t = 1, 2, \dots, n$. It is assumed to be independent identically distributed (i.i.d) following normal distribution $N(0, \tau^{-1})$ where τ is the precision parameter.

The ARMA(p,q) model, as described in (1), suggests that the observation recorded at time t is influenced by two sets of regressors; the first set comprises its values from the preceding p time periods, while the second set consists of the errors from the prior q time periods.

Stationarity and invertibility represent two critical attributes that can influence both the estimation and prediction processes. To achieve stationarity or invertibility, certain constraints are placed on the coefficients of the model.

The conditions for stationarity in the model presented in (1) require that the roots of $\phi(B)$ exceed one, while the conditions for invertibility stipulate that the roots of $\theta(B)$ must also exceed one (Box et al., 2016).



Specific instances of ARMA(p,q) models are derived by replacing p and q with suitable values. For example, if we set $p = 1$ and $q = 1$, the resulting model is ARMA(1,1)

$$y_t = \phi_1 y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

This model is stationary if $|\phi_1| < 1$ and invertible if $|\theta_1| < 1$.

In (2), by setting $q = 0$, we obtain an ARMA(p,0) or AR(p) model, which is referred to as a pure autoregressive model of order p . The structure of an AR(p) model is given by

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t. \quad (3)$$

An example for AR(p) models is AR(1) which has the form

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, n.$$

The AR(1) model is stationary if $|\phi_1| < 1$, whereas, AR models are invertible by definition (Box et al. (2016)). A further illustration can be seen in the AR(2) model, which is expressed in the following form

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \quad t = 1, 2, \dots, n.$$

The AR(2) model will be stationary if its coefficients satisfy the following three conditions $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$, and $|\phi_2| < 1$.

On the other hand, by setting $p = 0$ in equation (2), we obtain the ARMA(0,q) or MA(q) model, which is referred to as the pure Moving Average model of order q (MA(q)). This model takes the following form:

$$y_t = - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t. \quad (4)$$

An example for MA(q) models is MA(1) which has the form

$$y_t = -\theta_1 \varepsilon_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, n.$$

The MA(1) model is invertible if $|\theta_1| < 1$, whereas, MA models are stationary by definition (Box and Jenkins (1970)). A further illustration can be seen in the MA(2) model, which is expressed in the following form

$$y_t = -\theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} + \varepsilon_t, \quad t = 1, 2, \dots, n.$$

The MA(2) model will be invertible if its coefficients satisfy the following three conditions $\theta_1 + \theta_2 < 1$, $\theta_2 - \theta_1 < 1$, and $|\theta_2| < 1$.

The conditional likelihood function for the ARMA(p,q) model is conditioned on the initial p observations of the time series. It is noteworthy to state that henceforth, the conditional likelihood function will be simplified to the term likelihood function or LF. The LF can be expressed in the following manner (See Broemeling (1985)).

$$L(\underline{\gamma}, \tau | \underline{S}_n) \propto \tau^{\frac{1}{2}[n-p]} \exp \left\{ -\frac{\tau}{2} \sum_{t=p+1}^n \varepsilon_t^2 \right\} \quad (5)$$

or equivalently

$$L(\underline{\gamma}, \tau | \underline{S}_n) \propto \tau^{\frac{1}{2}[n-p]} \exp \left\{ -\frac{\tau}{2} \sum_{t=p+1}^n (y_t - \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j})^2 \right\}$$

where

$\underline{S}_n$: is the vector of observations such that

$$\underline{S}_n = [y_{p+1} \ y_{p+2} \ \dots \ y_n]', \quad n > p. \quad (6)$$

$\underline{\gamma}$: is the coefficients' vector consisting of $h = p + q$ elements and is defined as

$$\underline{\gamma}' = [\phi_1 \ \phi_2 \ \dots \ \phi_p \ -\theta_1 \ -\theta_2 \ \dots \ -\theta_q]'. \quad (7)$$

τ : is the precision parameter.

ε_t : is the error term defined as

$$\varepsilon_t = y_t - \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j}. \quad (8)$$

It is evident that the exponent on the right side of (5) includes a squared term related to the errors ε_t 's. Nevertheless, the error associated with ARMA models is non-linear with respect to the coefficients of the model due to the presence of the MA component (See Broemeling and Shaarawy (1988)). This characteristic renders the LF intractable, resulting in posterior or predictive distributions that exhibit non-standard forms.

Approximations can simplify the likelihood function (LF) to derive standard posterior and predictive densities, specifically the t density. Numerous approximations are documented in the literature, including those by Newbold (1973), Zellner and Reynolds (1978), and Broemeling and Shaarawy (1988). The fundamental concept underlying these various approximations is the estimation of the sum of squares of errors as quadratic functions of the coefficients. Estimation may be performed directly by treating the sum of squares of errors as a quadratic function of the coefficients, as exemplified by the method of Zellner and Reynolds, or by approximating the errors as a linear function of the coefficients, as is the case with the other two methods.

In the present research, the Broemeling and Shaarawy (B-S) approximation has been chosen due to its simplicity and convenience for purposes of the study. This method allows for the errors to be estimated as linear functions of the coefficients derived from their Nonlinear least squares Estimates (NLSE's).

$$\hat{\varepsilon}_t = y_t - \sum_{i=1}^p \hat{\phi}_i y_{t-i} + \sum_{j=1}^q \hat{\theta}_j \hat{\varepsilon}_{t-j} \quad (9)$$

The initial values of the errors are presumed to be equivalent to their unconditional mean, namely zero, i.e. $\hat{\varepsilon}_0 = \hat{\varepsilon}_{-1} = \dots = \hat{\varepsilon}_{1-q} = 0$. Substituting $\hat{\varepsilon}_t$ from (9) instead of ε_t in the LF (5), the following approximate LF is obtained

$$\hat{L}(\underline{\gamma}, \tau | \underline{S}_n) \propto \tau^{\frac{1}{2}[n-p]} \exp \left\{ -\frac{\tau}{2} \sum_{t=p+1}^n \hat{\varepsilon}_t^2 \right\} \quad (10)$$

3. Bayesian time series analysis

Bayesian time series analysis presents unique benefits compared to the Box-Jenkins approach. Bayesian methods effectively execute the four stages of time series analysis, addressing issues related to identification, estimation, diagnostic checking, and prediction. Furthermore, the Bayesian framework equips time series analysts with a coherent and systematic method to integrate prior knowledge into their analyses, potentially resulting in precise outcomes even with limited time series lengths. The essential tools for conducting Bayesian time series analysis include the prior, posterior, and predictive distributions.

3.1 Prior distributions

The prior distribution signifies the initial beliefs regarding the parameters of the model. There are two categories of priors: informative and non-informative. When information about the model parameters is accessible, it is essential to enhance the accuracy and reliability of the conclusions by utilizing informative priors. Conversely, a non-informative prior is employed in situations where there is minimal or no information concerning the model parameters.

In the context of an ARMA(p,q) process, a suitable choice for an informative prior is the normal-gamma prior, particularly when employing the Broemeling and Shaarawy approximation (See Broemeling and Shaarawy (1988)). The normal-gamma prior for the parameters $\underline{\gamma}$ and τ of the ARMA(p,q) model is represented in the following manner,

$$\xi_1(\underline{\gamma}, \tau | p, q) \propto \frac{|Q_h|^{\frac{1}{2}} \tau^{\alpha + \frac{h}{2} - 1}}{(2\pi)^{\frac{h}{2}}} e^{-\frac{\tau}{2} [2\beta + (\underline{\gamma} - \underline{\mu})' Q_h (\underline{\gamma} - \underline{\mu})]}, \quad \tau > 0, \quad (11)$$

Where, $\alpha, \beta, \underline{\mu} \in R^h$ 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 $h = p+q$. Note that, the normal-gamma prior is conditioned on p and q since the orders are considered random variables in the Bayesian analysis. Moreover, let β_{ij} be the prior probability mass function of the orders

$$p \text{ and } q, \text{ i.e. } \beta_{ij} = \Pr[p = i, q = j]; \quad i = 1, 2, \dots, P; \quad j = 1, 2, \dots, Q \quad (12)$$

Where, P and Q are the assumed upper limits of the orders p and q respectively. β_{ij} is assumed to have one of the following three forms:

Prior 1:

$$\beta_{ij} = \frac{1}{P} \times \frac{1}{Q}, \quad i = 1, 2, \dots, P; \quad j = 1, 2, \dots, Q$$

Prior 2:

$$\beta_{ij} \propto (0.5)^{i+j}, \quad i = 1, 2, \dots, P; \quad j = 1, 2, \dots, Q$$

Prior 3:

$$\beta_{11} = 0.35; \quad \beta_{12} = \beta_{21} = 0.25; \quad \beta_{22} = 0.15 \text{ for } P=Q=2$$



The first prior assigns equal probabilities to all possible values of the orders p and q . The second prior is chosen in such a way to give probabilities that decline exponentially with the orders p and q . While the third prior is chosen in such a way to give probabilities that decrease with an absolute amount 0.1 as the orders increase.

Thus, from (11) and (12), the joint prior distribution of the parameters $\underline{\gamma}$, τ , p and q is

$$\xi_1(\underline{\gamma}, \tau, p, q) \propto \beta_{ij} \frac{|Q_h|^{\frac{1}{2}} \tau^{\alpha + \frac{h}{2} - 1}}{(2\pi)^{\frac{h}{2}}} e^{-\frac{\tau}{2} [2\beta + (\underline{\gamma} - \underline{\mu})' Q_h (\underline{\gamma} - \underline{\mu})]}, \tau > 0 \quad (13)$$

Furthermore, an appropriate selection as a non-informative prior is Jeffrey's prior (See Jeffreys' (1961)). It can be written as

$$\xi_2(\underline{\gamma}, \tau, p, q) \propto \tau^{-1} \quad (14)$$

3.2 Posterior distributions

The posterior distribution serves as the Bayesian instrument for estimating the unknown parameters within the model. The posterior distribution $P(\underline{\gamma}, \tau | \underline{S}_n, p, q)$ is the conditional distribution of the model parameters given the observations where $\underline{\gamma}$, τ , p , q and $\underline{S}_n$ are defined as in (5). Bayes' theorem posits that the multiplication of the chosen prior distribution and the likelihood function (LF) (5) is proportional to the posterior distribution.

$$P(\underline{\gamma}, \tau | \underline{S}_n, p, q) \propto \xi(\underline{\gamma}, \tau | p, q) \cdot L(\underline{\gamma}, \tau | \underline{S}_n, p, q) \quad (15)$$

The marginal posterior distributions of the orders p and q or the other parameters $\underline{\gamma}$ and τ are derived by integrating the joint posterior distribution (15) with respect to the irrelevant arguments.

3.3 Predictive distributions

The Bayesian tool to predict the first future observation y_{n+1} is the one step-ahead predictive distribution $P(y_{n+1} | \underline{S}_n, p, q)$ defined as

$$P(y_{n+1} | \underline{S}_n, p, q) \propto \int \int_{\underline{\gamma}, \tau} P(y_{n+1}, \underline{S}_n | \underline{\gamma}, \tau, p, q) \xi(\underline{\gamma}, \tau | p, q) d\tau d\underline{\gamma} \quad (16)$$

Consequently, the one step ahead predictive distribution for y_{n+1} is derived by multiplying the conditional joint distribution of y_{n+1} and the observations $\underline{S}_n$, given the model parameters $\underline{\gamma}$, τ , p and q , namely, $P(y_{n+1}, \underline{S}_n | \underline{\gamma}, \tau, p, q)$, by the prior distribution, denoted by $\xi(\underline{\gamma}, \tau, p, q)$, then the parameters $\underline{\gamma}$ and τ are eliminated by integration (See Broemeling and Shaarawy (1988)).

By utilizing the aforementioned Bayesian tools, one can implement the four stages of Bayesian time series analysis: identification, estimation, diagnostic checking, and prediction as follow:

3.4 Identification

From a Bayesian viewpoint, the orders p and q in ARMA processes are seen as further unknown parameters. The posterior joint probability mass function must be constructed by employing the approximate conditional likelihood function (10) and an appropriate prior distribution. The first step in this process is to derive the approximate conditional likelihood function for the main parameters $\underline{\gamma}$ and τ and the orders p and q . The second step involves selecting a prior distribution for all parameters, including the orders p and q , which effectively integrates with the approximate conditional likelihood function to facilitate a manageable posterior analysis. The more complex aspect lies in eliminating the parameters $\underline{\gamma}$ and τ to derive the posterior joint probability mass function of the orders p and q in a straightforward manner. Once the joint posterior mass function is established, one can examine the behavior of the posterior probabilities across the grid of orders and identify the pair of orders that yields the highest posterior probability, thereby addressing the identification problem.

3.5 Estimation

After identifying the orders of the model, say as an ARMA $(\hat{p}, \hat{q})$, the parameters $\underline{\gamma}$ and the residual variance $\sigma^2 = \tau^{-1}$ should be estimated. From the Bayesian viewpoint, all statistical inferences are based upon the conditional joint posterior distribution of the parameters $\underline{\gamma}$ and τ given $p = \hat{p}$ and $q = \hat{q}$. This distribution represents our knowledge about the parameters after observing the time series data and estimating the order of the model. Thus, one may investigate the conditional posterior distribution of the parameters and compute certain characteristics such as the expectation, mode, median and variance for this distribution. If τ is a nuisance parameter and the coefficients $\underline{\gamma}$ are our main parameters, the inferences of the main parameters should be based on the conditional joint posterior distribution of the main parameters $\underline{\gamma}$ given $p = \hat{p}$ and $q = \hat{q}$. The conditional expectation, median, or mode of this distribution could be taken as point estimates for the coefficients $\underline{\gamma}$. In addition one may construct HPD regions for the coefficients. If the residual precision τ is of our interest, one may use the expectation (or the mode) of the marginal conditional posterior distribution of τ given $p = \hat{p}$ and

$q = \hat{q}$ as a point estimate of τ . If the error variance $\sigma^2 = \tau^{-1}$ is of our interest, one may use the conditional expectation of τ^{-1} given $p = \hat{p}$ and $q = \hat{q}$ as a point estimate for σ^2 . A confidence interval for the parameter τ (or σ^2) can be constructed using the same distribution of τ given $p = \hat{p}$ and $q = \hat{q}$.

3.6 Diagnostic check

To investigate the adequacy of the model, which was tentatively identified as an $\text{ARMA}(\hat{p}, \hat{q})$, one may over fit the identified model by fitting the series to an $\text{ARMA}(\hat{p}+1, \hat{q})$, an $\text{ARMA}(\hat{p}, \hat{q}+1)$ or an $\text{ARMA}(\hat{p}+1, \hat{q}+1)$ and tests the significance of the additional coefficient $H_0: \phi_{\hat{p}+1} = 0$ versus $H_1: \phi_{\hat{p}+1} \neq 0$ and (or) $H_0: \theta_{\hat{q}+1} = 0$ versus $H_1: \theta_{\hat{q}+1} \neq 0$ using the appropriate marginal conditional posterior distribution. The result will be a model identified as an $\text{ARMA}(p', q')$, where $p' = \hat{p}$ or $\hat{p}+1$ and $q' = \hat{q}$ or $\hat{q}+1$.

3.7 Prediction

The predictive distribution for future observations serves as a Bayesian approach to address prediction challenges. Particular emphasis is placed on forecasting the subsequent observation $Y(n+1)$. Generally, the predictive distribution of $Y(n+1)$ is defined as the conditional distribution of $Y(n+1)$ given the preceding observations $Y(1), Y(2), \dots, Y(n)$. This is achieved by multiplying the conditional distribution of $Y(n+1)$ given the parameters γ, τ and the preceding observations with the posterior distribution of the parameters γ, τ , followed by the elimination of the parameters through integration. Consequently, the predictive distribution of $Y(n+1)$ can be viewed as the average of the conditional predictive distribution of $Y(n+1)$, conditioned on γ, τ , in relation to the posterior distribution of γ, τ . Therefore, it is essential to plot and analyze the predictive distribution of $Y(n+1)$ to derive statistical inferences regarding the subsequent observation. The mean will serve as a point forecast, while the variance, in conjunction with the mean, will provide an interval forecast.

To predict two steps ahead, one should utilize the conditional predictive distribution of $Y(n+2)$ given that $Y(n+1)$ is equal to its expected value, which corresponds to the prior forecast. To extend the forecast to future observations $Y(n+3), Y(n+4)$, and so forth, a series of forecasts should be employed, each conditional on the preceding forecasts, as indicated by the conditional predictive means (See Broemeling and Shaarawy (1988)).

4. Bayesian analysis of annual growth percent of inflation

The data represents the annual growth percent of inflation (consumer prices) in Egypt from 1960 to 2025. The source of the time series is World Bank with length $n = 66$. We will use the Bayesian analysis to analyze the time series in four steps: Model identification, Model estimation, Diagnostic checking and finally Forecasting the following three years from 2026 to 2028. Due to lack of Bayesian time series analysis packages, the analysis is done using self-developed R programs. Also, Minitab is used to produce some statistical measurements and graphs.

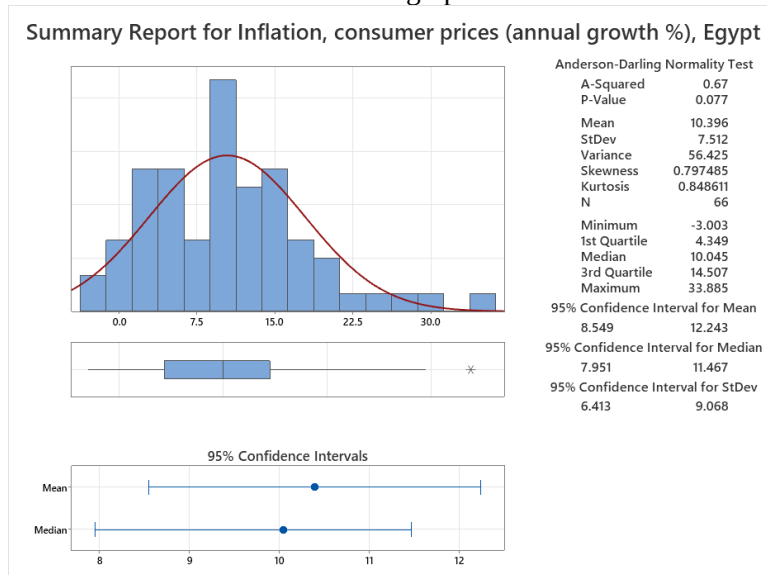


Figure (1): The p-value of Anderson-Darling normality test is greater than the level of significance (0.05) which indicates normality of inflation growth % data.

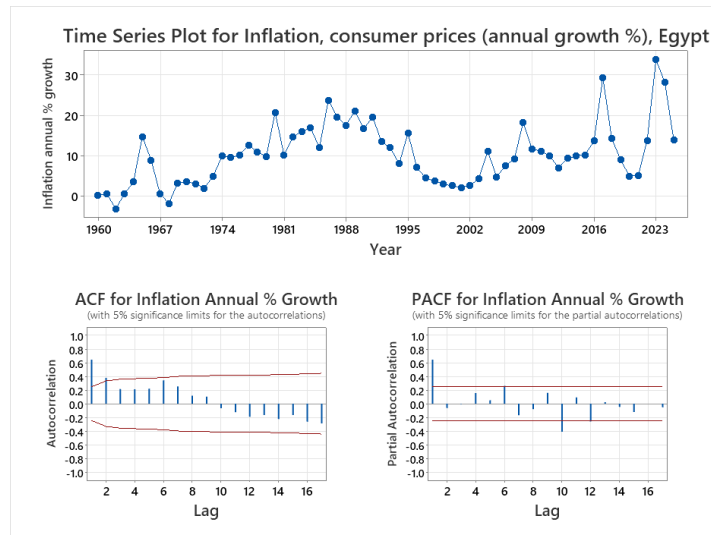


Figure (2): The time series plot shows that the inflation annual percent growth series has an increasing trend over time. Moreover, the inspection of ACF shows a slow decaying pattern and a single large spike in PACF which indicates together to a non-stationary time series. This is also supported by augmented Dickey-Fuller test results.

Augmented Dickey-Fuller Test For Inflation Series

Null hypothesis:		Data are non-stationary
Alternative hypothesis:		Data are stationary
Test Statistic	P-Value	Recommendation
-2.48865	0.118	Test statistic > critical value of - 2.91473.
		Significance level = 0.05
		Fail to reject null hypothesis.
		Consider differencing to make data stationary.

The natural log transformation is used to remove non-stationarity in variance before calculating the first difference of the inflation rate time series.

Augmented Dickey-Fuller Test for 1st Diff. of LN(inflation)

Null hypothesis:		Data are non-stationary
Alternative hypothesis:		Data are stationary
Test Statistic	P-Value	Recommendation
-8.01120	0.000	Test statistic <= critical value of -1.94608.
		Significance level = 0.05
		Reject null hypothesis.
		Data appears to be stationary, not supporting differencing.

It is now clear from Dickey-Fuller test that the first difference of natural logarithm of inflation series is now a stationary time series and no need for further differencing. After preparing the data for analysis, the first step is to primary select the order of the ARIMA model. Using the direct Bayesian identification method as described in section (3), the probability density function of the model order, using the three priors mentioned in section 3, are shown in the following table.

Table 1: Posterior Mass Functions of Models Orders with Three Different Prior Functions

MODEL	Jeffreys' Prior	Geometric Prior	Arithmetic Prior
ARIMA(1,1,0)	0.0248	0.0472	0.0427
ARIMA(0,1,1)	0.3384	0.6443	0.5830
ARIMA(1,1,1)	0.1157	0.1101	0.1423
ARIMA(2,1,0)	0.0003	0.0002	0.0003
ARIMA(0,1,2)	0.1017	0.0968	0.1252
ARIMA(2,1,1)	0.0010	0.0005	0.0007
ARIMA(1,1,2)	0.0060	0.0028	0.0044
ARIMA(2,1,2)	0.4122	0.0981	0.1014

In this analysis, to promote the model parsimonious principle, we will depend on the third prior function of the model order, the arithmetic prior. Therefore, the models with highest posterior probabilities of the model order is ARIMA(0,1,1), ARIMA(1,1,1) and ARIMA(0,1,2) respectively. The second and third steps of the Bayesian time series analysis are model estimations and diagnostic checking.

Table 2: Posterior Estimates of ARIMA(0,1,1) Model



Parameter	Posterior Estimates	95% HPD Region	Significance
θ	0.2804	(0.0357,0.5251)	(significant)
Error Variance	0.2219		

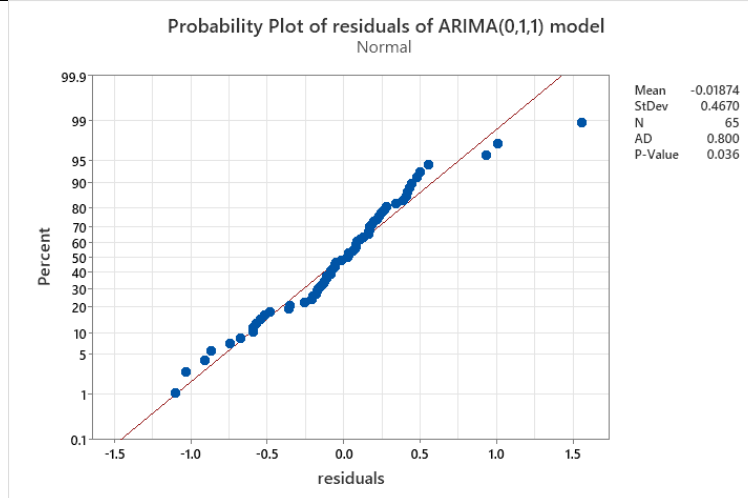


Figure (3): Although the ARIMA(0,1,1) model parameter is significant, the normality plot and Anderson-Darling normality test show that the residuals of ARIMA(0,1,1) lacks the normality condition. So this model fail to fit the data and satisfy the analysis conditions, therefore it is excluded. The next highest posterior probability model is ARIMA(1,1,1) model.

Table 3: Posterior Estimates of ARIMA(1,1,1) Model

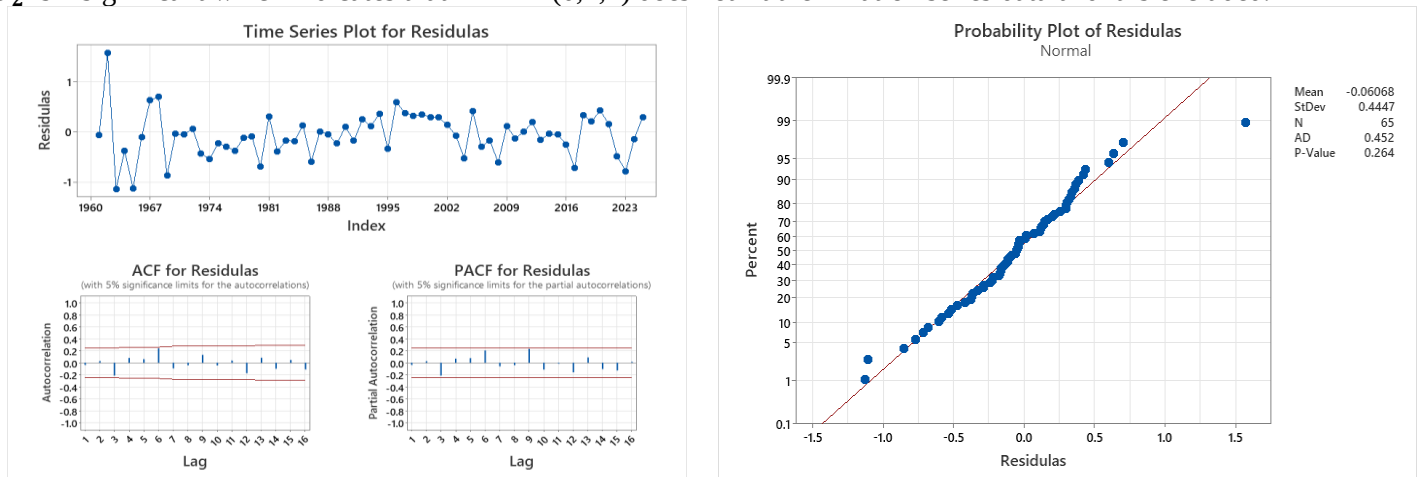
Parameter	Posterior Estimates	95% HPD Region	Significance
ϕ	0.7399	(0.4673,1.0126)	(significant)
θ	1.0106	(0.7537,1.2675)	(significant)
Error Variance	0.2047		

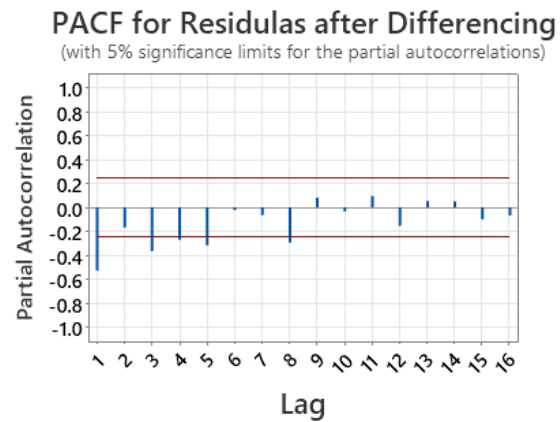
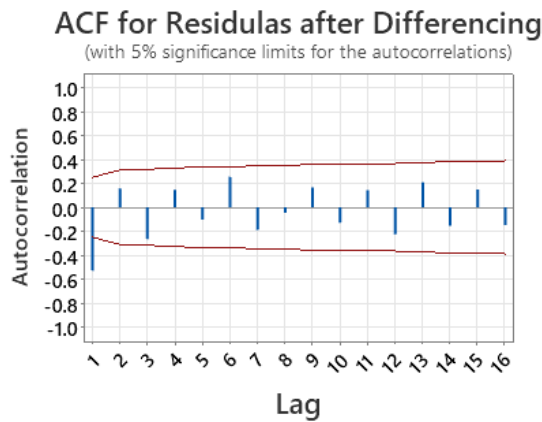
Both ARIMA(1,1,1) model parameters are significant. The preliminary checking of this model is satisfied, so it can be included in further diagnostic checking tests. This includes: Fitting the next model with highest posterior probability, checking randomness and normality of residuals, ACF and PACF of residuals, fitting MA(1) model with parameter almost equal 1 for the first difference of the residuals, Ljung-Box test and calculating accuracy measures to evaluate the performance of the fitted model.

Table 4: Posterior Estimates of ARIMA(0,1,2) Model

Parameter	Posterior Estimates	95% HPD Region	Significance
θ_1	0.2661	(0.0136,0.5186)	(significant)
θ_2	0.2218	(-0.03037,0.4740)	(insignificant)
Error Variance	0.2118		

θ_2 is insignificant which indicates that ARIMA(0,1,2) does not fit the inflation series data and it is excluded.





ARIMA(0,1,1) model for Residuals

Type	Coef	SE Coef	T-Value	P-Value
MA 1	0.9901	0.0380	26.06	0.000

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	13.36	19.35	26.29	36.10
DF	11	23	35	47
P-Value	0.344	0.733	0.882	0.897

Figure (4): Diagnostic Checking of ARIMA(1,1,1) Model

In figure (4), the randomness of the residuals is clear since no significant spikes are found in both ACF and PACF functions. This result is supported by the Ljung-Box test which shows that, there is no significant autocorrelation in residuals. Moreover, the parameter θ in the fitted MA(1) model for the first difference of residuals is almost 1, which show that the residuals follows a white noise process. Also, the normality of the residuals is verified due to the insignificance of Anderson-Darling normality test.

From the above discussion, ARIMA(1,1,1) model passes all the diagnostic checking tests and can be used efficiently in the forecasting process.

Table 5: Accuracy Measures for Fitted ARIMA(1,1,1) Model

Measure	RMSE	MAE	MAPE
Value	0.4454	0.3343	8.7624

Table 6: Forecasts of inflation growth % for years 2026 to 2028

95% Limits			
Year	Forecast	Lower	Upper
2026	12.0204	11.0732	12.9675
2027	8.1971	7.2658	9.1283
2028	5.5119	4.5971	6.4266

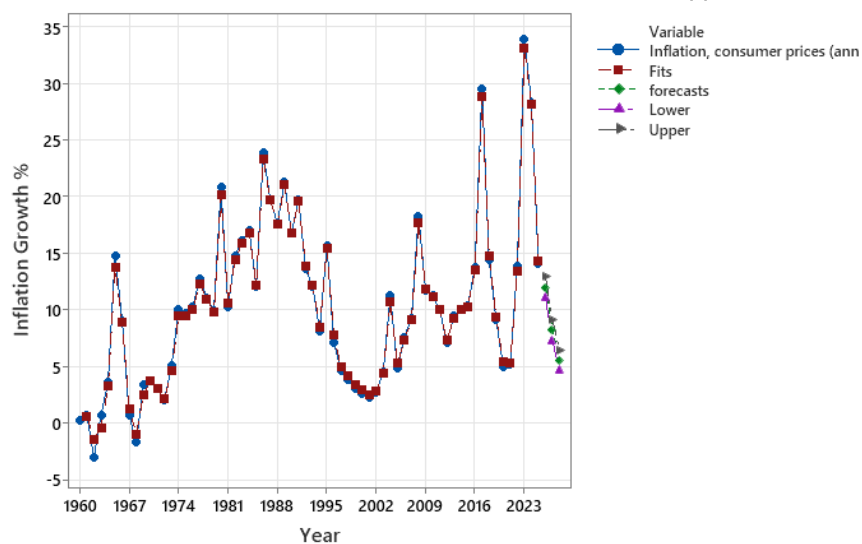


Figure (5): The point forecast of the annual inflation growth percent in 2026 is found to be 12.0204% with a 95% HPD region (11.0732, 12.9675). Then remarkable decreases appear in both 2027 and 2028 where the point forecasts of the annual inflation growth percent are 8.1971 and 5.5119 respectively. It is worth noting that the forecast process assumes conditions remain the same as they are and no external factors interfering.



5. Conclusion

The primary aim of the present study is to utilize the Bayesian methodology to analyze a real-time series of annual inflation growth percentages in Egypt from 1960 to 2025. This analysis seeks to comprehend its functional structure and provide precise forecasts for future observations. The series is modeled, and its observations are predicted through the Bayesian approach. The accuracy metrics employed in this study include Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). An examination of the series revealed its non-stationarity, necessitating certain transformations to achieve stationarity. Transformations involving the natural logarithm and first-order differences were applied. Following the procedures of Bayesian time series analysis, the findings indicated that the ARIMA(1,1,1) model accurately represents the inflation data. The coefficients of this model were found to be significant. Furthermore, the model successfully passed all diagnostic checks and can be effectively utilized to forecast future values of the annual inflation growth percentage. It was employed to generate point and interval forecasts for the subsequent three years, from 2026 to 2028.

References

- Abraham, B. and Ledolter, J. (1983). *Statistical Methods for Forecasting*. John Wiley and Sons Inc. New York.
- Albassam, M.S., Soliman E.E.A. and Ali, S.S. (2023). An Effectiveness Study of the Bayesian Inference with Multivariate Autoregressive Moving Average Processes. *Communications in statistics- Simulation and Computation*, 52(10):4773-4788. <https://doi.org/10.1080/03610918.2021.1967986>.
- Albert, J. and Chib, S. (1993). Bayesian Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. *Journal of Business and Economic Statistics*, 11(1):1-15.
- Ali, S.S. (2015). Bayesian Forecasting Of Vector Moving Average Processes. . *The Egyptian Statistical Journal*, 59(1):68-89. <https://doi.org/10.21608/ESJU.2015.314457>.
- Amin, A. (2019). Gibbs Sampling for Bayesian Prediction of SARMA Processes. *Pakistan Journal of Statistics and Operation Research*, 15(2):483-499. <https://doi.org/10.18187/pjsor.v15i2.2174>.
- Barnett, G., Kohan, R. and Sheather, S. (1995). Markov Chain Carlo Estimation of Autoregressive Models with Application to Metal Pollutant Concentration in Sludge. *Mathematical and Computational Modeling*, 22(1):7-13.
- Bowerman, B. and O'Connell (1987). *Time Series Forecasting: Unified Concepts and Computer Implementation*. Boston: PWS Publishers Dexbury Press.
- Box, G. and Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- Box, G., Jenkins, G., Reinsel, G. and Ljung, M. (2016). *Time series analysis: Forecasting and control*. 5th ed. Hoboken, NJ: John Wiley & Sons.
- Broemling, L. (1985). *Bayesian Analysis of Linear Models*. Marcel Inc., New York.
- Broemeling, L., Shaarawy, S. (1988). *Time series: A Bayesian analysis in time domain: Bayesian Analysis of Time Series and Dynamic Models*. Studies in Bayesian Analysis of Time Series and Dynamic Models, Edited by J. spall, Marcel Dekker Inc., New York:1-22.
- Chatfield, C. (1980). *The Analysis of Time Series: Theory and Practice*. Chapman and Hall Ltd., London.
- Chatfield, C. (2019). *The analysis of time series: Theory and practice*. An introduction with R, 7th ed. Boca Raton, FL: Chapman & Hall.
- Chen, C. (1992). Bayesian Inferences and Forecasting in Bilinear Time Series Models. *Communication in Statistics – Theory and Methods*, 21(6):1725-1743. <https://doi.org/10.1080/03610929208830875>.
- Chib, S. (1993). Bayes Regression with Autoregressive errors. *Journal of Econometrics*, 63(3):275-294.
- Chib, S. and Greenberg, E. (1994). Bayes Inference in Regression Models With ARMA(p,q) Errors. *Journal of Econometrics*, 64(1):183-206.
- Chow, G. (1975). Multi-period Prediction from Stochastic Difference Equations by Bayesian Methods .*Chapter 8 of Studies in Bayesian Econometrics and Statistics* , edited by S.E. Fienberg and A. Zellner. North-Holland, Amsterdam.
- De Alba, E. (1993). Constrained Forecasting in Autoregressive Time Series Models: Bayesian Analysis, *International Journal of Forecasting*, 9(1):95-108.
- Diaz, J. and Farah, J. (1981). Bayesian Identification of Autoregressive Process. *Presented at the 22nd NBER-NSF Seminar on Bayesian Inference in Econometrics*.
- Elnagger, A., & Richter, C. (2022). Exchange rate pass-through, inflation, and monetary policy in Egypt. *Applied Economics Quarterly*, 68(1):1-36. <https://doi.org/10.3790/aeq.68.1.1>.
- El-Sheikh, A. A. and Kandil, S. M. (2016). Time Series Analysis of Inflation in Egypt. *Presented at the 48th Annual Conference on Statistics, Computer Science and Operation Research*, Institute of Statistical Studies and Research, Cairo University.
- Granger, C. and Newbold, P. (1977). *Forecasting Economic Time Series*. Academic Press, New York.
- Greenlaw, S. A., Shapiro, D., & MacDonald, D. (2026). *Principles of economics*. 3rd edition, OpenStax, (an initiative of Rice University).
- Harison, P. and Stevens, C. (1976). Bayesian Forecasting (with discussion). *Journal of the Royal Statistical Society, Series B*, 38(2):205-242.
- Harvey, A. (1981). *Time Series Models*. Wiley, New York.



- Lahif, M. (1980). *Time Series Forecasting with non-informative Prior Distribution*. Technical Report No. 111, Department of Statistics, University of Chicago.
- Liu, L. (2009). *Time Series Analysis and Forecasting*. 2nd edition. Scientific Computing Association Corp, USA.
- Marriott, J., Ravishanker, N., Gelfand, A. and Pai, J. (1996). Bayesian Analysis of ARMA Processes: *Complete Sampling based Inference under Full Likelihoods*, In *Bayesian Statistics and Econometrics*, edited by Zellner, A., Berry, D., Chaloner, K. and Geweke, J., John Wiley & Sons, Inc., New York.
- Monahan, J. F. (1983). Fully Bayesian analysis of ARIMA time series models. *Journal of Econometrics*, 21(3):307–331. [https://doi.org/10.1016/0304-4076\(83\)90048-9](https://doi.org/10.1016/0304-4076(83)90048-9).
- Newbold, P. (1973). Bayesian Estimation of Box and Jenkins Transfer Function Model for Noise Models. *Journal of the Royal Statistical Society, Series B*, 35(2):323–336.
- Omar, M.A., Shahin, S.E. and Roshdy, M. (2022). Forecasting Inflation in Egypt (2019-2022) by using AutoRegressive Integrated Moving Average (ARIMA) Models. *Journal of Advanced Veterinary Research*, 12(6):670–676.
- Philippe, A. (2006). Bayesian Analysis of Autoregressive Moving Average Processes with Unknown Orders. *Computational Statistics & Data Analysis*, 51(3): 1904–1923. <https://doi.org/10.1016/j.csda.2005.12.005>.
- Priestley, M. (1981). *Spectral Analysis of Time Series*, Academic Press, London.
- Romer, D. (2018). *Advanced Macroeconomics*. 5th edition, McGraw-Hill Education.
- Shaarawy, S. and Ali, S. (2003). Bayesian Identification of Seasonal Autoregressive Models. *Communications in Statistics-Theory and Methods*, 32(5):1067–1084. <https://doi.org/10.1081/STA-120019963>.
- Shaarawy, S., Ali, S. (2008). Bayesian identification of multivariate autoregressive processes. *Communications in Statistics—Theory and Methods*, 37(5):791–802. <https://doi.org/10.1080/03610920701504370>.
- Shaarawy, S. and Ali, S. (2015). Bayesian Identification of Seasonal Multivariate Autoregressive Processes. *Communications in Statistics—Theory and Methods*, 44(4): 823–836. <https://doi.org/10.1080/03610926.2012.752850>.
- Shaarawy, S. and Broemeling, L. (1984). Bayesian Inference and Forecasts with Moving Average Processes. *Communications in Statistics-Theory and Methods*, 13(15):1871–1888. <https://doi.org/10.1080/03610928408828800>.
- Shaarawy, S. and EL-Shawadfy, G. (1994). On the Adequacy of Using T Approximation in Bayesian Inferences of ARMA Models. *The Egyptian Statistical Journal*, 43(2):193–208.
- Shaarawy, S., Soliman, E. and Ali, S. (2007). Bayesian Identification of Moving Average Models, *Communications in Statistics-Theory and Methods*, 36(12):2301–2312. <https://doi.org/10.1080/03610920701215449>.
- Shaarawy, S.M. (2023). Bayesian modeling and forecasting of vector autoregressive moving average processes. *Communications in Statistics-Theory and Method*, 52(11):3795–3815. <https://doi.org/10.1080/03610926.2021.1980047>.
- Shaarawy, S.M., Ali, S.S. and Soliman, E.E.A. (2024). Bayesian Identification of Seasonal Vector ARMA Processes. *The Egyptian Statistical Journal*, 68(2): 129–145. <https://doi.org/10.21608/ESJU.2024.318938.1044>.
- Soliman, E.E.A. (2008). Efficiency of the approximate Bayesian prediction of ARMA models: A simulation study. *The Egyptian Statistical Journal*, 52(2):56–75. <https://doi.org/10.21608/ESJU.2008.315419>.
- Soliman E.E.A. and Ali, S.S. (2026). Assessment of the performance of the Bayesian Forecasting for Vector ARMA Processes. *SCOPUA: Journal of Applied Statistical Research*. 2(1). <https://doi.org/10.64060/JASR.v2i1.2>.
- Tiao, G. and Box, G. (1981). Modeling Multiple Time Series with Application. *Journal of the American Statistical Association*, 76(4):802–816. <https://doi.org/10.1080/01621459.1981.10477728>.
- Young, K. and Pettit, L. (2006). The Effect of Observations on Bayesian Choice of an Autoregressive Model. *Journal of Time Series Analysis*, 27(1):41–50. <https://doi.org/10.1111/j.1467-9892.2005.00449.x>.
- Wei, W.W.S. (2005). *Time Series Analysis: Univariate and Multivariate Methods*. Addison Wesley, Reading, MA.
- Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. John Wiley and Sons Inc., New York.
- Zellner, A. and Reynolds, R. (1978). Bayesian Analysis of ARMA Models. *Presented at the Sixteenth Seminar on Bayesian Inference in Econometrics*, June 2–3.

Declaration

Conflict of Study: The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contribution Statement: Both authors conceived the idea and designed the research; Analyzed and interpreted the data and wrote the paper.

Availability of Data and Material: Data will be made available by the corresponding author on reasonable request.

Author(s) Bio / Authors' Note

Sherif S. Ali is an Associate Professor in Statistics at the Department of Statistics, King Abdulaziz University, Jeddah 21589, Saudi Arabia. Email: ssali@kau.edu.sa

Emad E. A. Soliman is an Associate Professor in Statistics at the Department of Statistics, King Abdulaziz University, Jeddah 21589, Saudi Arabia. Email: ehgazzy@kau.edu.sa



