Journal of Applied Sciences Research, 4(12): 1885-1892, 2008

Bayesian Analysis of the Autoregressive-Moving Average Model with Exogenous Inputs Using Gibbs Sampling

Hassan M.A. Hussein

Department of Statistics, Faculty of Commerce, Zagazig University, Egypt

Abstract: The problem of estimating a set of parameters in the autoregressive moving average model with exogenous inputs (ARMAX) is considered and a numerical Bayesian method proposed. This paper, develops a Bayesian analysis for the ARMAX model by implementing a fast, easy and accurate Gibbs sampling algorithm. The procedure is easy to implement and can be computed also when some priors in the ARMAX are diffuse. The empirical results of the simulated examples and electricity consumption data in the industrial sector in Egypt showed the accuracy of the proposed methodology and has good statistical properties.

Key words: Autoregressive-Moving Average Model With Exogenous Inputs, Gibbs Sampling

INTRODUCTION

Autoregressive Moving Average (*ARMA*) models can be generalized in other ways, such as, Autoregressive Conditional Heteroskedasticity (*ARCH*) models and Autoregressive Integrated Moving Average (*ARIMA*) models. The *ARIMA* models are introduced to the purposes of forecast for time series data by Box-Jenkins[2]. The models are widely used as forecast techniques and developed to various econometric methodologies such as the Autoregressive Moving Average model with Exogenous inputs (*ARMAX*),[7]. The *ARMAX* model just allows explanatory variables in *ARIMA* system.

Because of energy crisis in the 1970s and the price hikes, especially in oil prices, economic growth of developing countries has been negatively affected. The association between energy consumption and economic growth have been extensively investigated since the late 1970s. These studies show that the relationship between energy consumption and economic growth is still a controversial one.

Different results have been found not only for different countries but also for different time within the same country. The empirical results suggest that, the electricity consumption can be considered as a good approximation for the consumed quantities of energy for the industrial sector[1].

Recently, Bayesian time series analysis has been advanced by the emergence of Markov Chain Monte Carlo (MCMC) methods, especially the Gibbs sampling method. Assuming a prior distribution on the initial observations and initial errors. Chib and Greenberg[3] and Marriott *et al.*[12] developed a Bayesian analysis for autoregressive moving average (*ARMA*) models using MCMC technique.

Ismail[9] employed Gibbs sampling to develop Bayesian analysis of the multiplicative, seasonal moving average model. Her approach was based on approximating the likelihood function via estimating the unobserved errors. The approximate likelihood is then used to derive the conditional distributions required for implement the Gibbs sampler.

A similar approach to that of Ismail[9] is used in this study but for developing Bayesian analysis of the Autoregressive Moving Average model with exogenous inputs (*ARMAX*). The procedure has the advantage that can be used under informative and non-informative priors on the parameters of interest. Moreover,this procedure does not require the estimation of two models, one with and the other without the restriction to be tested if the problem of testing a set of restrictions in the underlying model is considered.

In this paper, we present a simple and fast algorithm to estimate the parameters and forecast future values of Bayesian analysis for the *ARMAX* model by implementing a fast, easy and accurate Gibbs sampling algorithm models. The procedure is easy to implement and can be computed also when some priors in the *ARMAX* model are diffuse. The Bayesian approach allows for incorporating uncertainty about the parameters and initial errors. The empirical results of the simulated examples and the electricity consumption data in the industrial sector in Egypt showed the accuracy of the proposed methodology and has a good statistical properties.

Corresponding Author: Hassan M.A. Hussein, Department of Statistics, Faculty of Commerce, Zagazig University, Egypt

The paper is organized as follows. Section 2 briefly describes the autoregressive moving average model with exogenous inputs (*ARMAX*). Section 3 is devoted to summarize posterior analysis, the full conditional posterior distributions and the implementation details of the proposed algorithm. In Section 4, A Monte Carlo study is discussed and moreover, the results of the estimation using a simulated example and the monthly electricity consumption data in Egypt are given. Finally, conclusions are given in Section 5.

Autoregressive moving-average model with exogenous inputs (*ARMAX*): The notation *ARMAX* (p, q, b) refers to the model with p autoregressive terms, q moving average terms and b exogenous inputs terms. This model contains the *AR*(p) and *MA*(q) models and a linear combination of the last b terms of a known and external time series d_t. The model can be expressed as:

$$y_t = \sum_{i=1}^{p}\gamma_i y_{t-i} + \sum_{i=1}^{q}\xi_i \varepsilon_{t-i} + \sum_{i=1}^{b}\eta_i d_{t-i} + \varepsilon_t, \quad (t=1,..,T) \qquad (1)$$

or

$$y_t = \sum_{i=-\infty}^{\infty} M(t-i)\beta + \varepsilon_t$$

where M?(t-i)= (y(t-i): ε(t-i): d(t-I)) $_{T\times k}$, (k=(p+q+d)), a 1×T row vector of unknown coefficients, β' = (γ_i: ζ_i: η_i) and the error terms ε_t are generally assumed to be independent random variables (i.i.d.) sampled from a normal distribution with zero mean $\varepsilon_t \sim N(0,\sigma^2)$ where σ^2 is the variance. These assumptions may be weakened but doing so will change the properties of the model and $\eta_1,..,\eta_b$ are the parameters of the exogenous input d_t. The model order (p,q,b) are assumed known and after observing the data y_t, t=1,2,..,T, the parameters β and the variance σ^2 are to be estimated. Moreover, *ARMAX* (p, q, b) model can be written as:

$$y_t = \sum_{i=1}^{\infty} \mathrm{M}(t-i)\beta(L)\varepsilon_t, \qquad (3)$$

where L denotes the lag-operator, i.e. $Ly_t = y_{t-1}$ fo $t \in Z, \gamma(L) = \gamma_1 L, \cdots, \gamma_p L^p$ $\xi(L) = \xi_0 + \xi_1 L + \cdots + \xi_q L^q$ and $\eta(L) = \eta_1 L + \cdots + \eta_b L^b$, the time series is as to start at time t = 1 with unknown starting errors $\varepsilon_0 = (\varepsilon_0, \varepsilon_{-1}, \ldots, \varepsilon_{1-p-q-b})$. Since the model is linear in γ, ζ and η which complicates the Bayesian analysis. However, the following sections explain how the Gibbs sampling technique can facilitate the analysis. The *ARMAX* (p,q,b) model (3) is invertible if the roots of the polynomials γ(L), ξ(L) and η(L) lie outside the unit circle. For more details about the properties of *ARMAX* (p,q b) models[7].

Posterior Analysis

Likelihood Function: Suppose $y = (y_1, y_2, \ldots, y_T)$ is a realization of the autoregressive moving average model with exogenous inputs model (1). Assuming that $f(\varepsilon_t) = N(0,\sigma^2)$ and employing a straightforward random variable transformation from ε_t to y_t, the likelihood function L (γ, ζ, η, σ^2 | y) is given by:

$$L(\gamma,\zeta,\eta,\sigma^2 | y) \propto (\sigma^2)^{-\frac{T}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^{T} \left(y_t - \sum_{i=1}^{p}\gamma_i y_{t-1} + \sum_{i=1}^{q}\xi_i \varepsilon_{t-1} + \sum_{i=1}^{b}\eta_i d_{t-1} + \varepsilon_t \right)^2 \right\} \qquad (4)$$

The likelihood function (4) is a complicated function in γ, ζ (2) and σ^2. Suppose the errors are estimated recursively as:

$$e_t = y_t - \sum_{i=1}^{p}\hat{\gamma}_i y_{t-1} + \sum_{i=1}^{q}\hat{\xi}_i \varepsilon_{t-1} + \sum_{i=1}^{b}\hat{\eta}_i d_{t-1},$$

where $\hat{\gamma} \in \Re^p$ $\hat{\xi} \in \Re^q$ $\hat{\eta} \in \Re^b$ and sensible estimates and $e_{1-p-q-b} = \cdots = e_{-1} = e_0$. Several estimation methods, such as the Innovations Substitution (IS) method proposed by Koreisha and Pukkila[10], give consistent estimates for γ, ζ, η and σ^2. The idea of the IS method is to fit a long autoregressive model to the series and obtain the residuals. Then appropriate lagged residuals are substituted into the *ARMAX* model (1). Lastly, the parameters are estimated using ordinary least squares.

Substituting the residuals (5) in (4) results in an approximate likelihood function:

$$L^*(\gamma,\zeta,\eta,\sigma^2 | y) \propto (\sigma^2)^{-\frac{T}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^{T} \left(y_t - \sum_{i=1}^{p}\gamma_i y_{t-1} + \sum_{i=1}^{q}\xi_i \varepsilon_{t-1} + \sum_{i=1}^{b}\eta_i d_{t-1} + \varepsilon_t \right)^2 \right\} \qquad (6)$$

Setting the initial errors zeros, i.e. $\varepsilon_0 = \varepsilon_{-1} = \cdots = \varepsilon_{1-p-q} = 0$, model (3) can be expressed in the following matrix form:

$$Y = M\beta + \varepsilon$$

with $Y' = (y_1, \cdots, y_T), \varepsilon' = (\varepsilon_1 \cdots \varepsilon_T)$, matrix with ti-th elements $M' = [y_{t-i} \vdots \varepsilon_{t-j} \vdots \quad d_{t-l}]$ and β is a k×1 vector of unknown parameters. Define $\psi = (\beta_1, \beta^*, \Sigma_\varepsilon, \Sigma^*)$, where β_1 is any element in vector β and β^* is the remain elements of β,under the normality assumptions:

$$N(0, \Sigma_\varepsilon), \quad (8)$$

where Σ_ε is the error term variance–covariance matrix of dimensions T × T.Let $£_0$ is a known matrix of dimensions k × a; relating the parameter β_1 to a parameter vector β^* of dimensions a × 1 ; possibly with a ≤ k; and Σ^* is the k × k variance–covariance matrix of β_1.and model the prior distribution on β_1 as:

$$\beta_1 | \beta^*, \Sigma^* \sim N(£_0 \beta^*, \Sigma^*), \quad (9)$$

This model is of the kind of hierarchical models introduced by Lindley and Smith[11], whose applications abound in fields as different as educational testing[13], medicine[5] and economics[8]. Bayesian estimation of ψ is simple, in principle and works as follows. Given the information contained in the data in the form of a approximate likelihood function:

$$\ell^*(\psi | y) = (\sigma^2)^{-\frac{T}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^{T} (Y - M\beta)' \Sigma_\varepsilon^{-1} (Y - M\beta) \right\}, \quad (10)$$

and a joint prior distribution on the parameters, $P(\psi|y) = P(\beta_1, \beta^*, \Sigma^*)$; the joint posterior distribution of the parameters conditional on the data is obtained through the Bayes' rule:

$$P(\psi|y) = \frac{P(\psi)\pi(y;\psi)}{P(y)} \propto P(\psi)\pi(y;\psi),$$

where ∝ denotes 'proportional to'. Given $P(\psi|y) = P(\beta_1, \beta^*, \Sigma_\varepsilon, \Sigma^*)$, the marginal posterior distributions conditional on the data ($P(\beta_1 | y)$,$P(\beta^* | y)$, $P(\Sigma^* | y)$ and $P(\Sigma_\varepsilon | y)$) can then be obtained by integrating out β_1; β^*; Σ^* and Σ_ε from the posterior distributions. Given that a prior for β_1 conditional on the other parameters have already been defined, a full implementation of the Bayesian approach requires the specification of a prior for β^*; Σ_ε and Σ^* such that:

$$P(\psi) = P(\beta_1, \beta^*, \Sigma_\varepsilon, \Sigma^*) = P(\beta^*, \Sigma_\varepsilon, \Sigma^*)\, P(\beta_1 | \beta^*, \Sigma^*),$$

Assuming independence, as it is customary and considering the inverse of the two matrices, just for convenience, we may take the joint prior distribution:

$$P(\beta^*, \Sigma_\varepsilon^{-1}, \Sigma^{*-1}) = P(\Sigma_\varepsilon^{-1})\, P(\beta^*)\, P(\Sigma^{*-1}),$$

to have, for example, a normal-Wishart–Wishart form;

$$P(\beta^*) = N(£_1\mu, C), \quad (12)$$

$$P(\Sigma_\varepsilon^{-1}) = W((\sigma_\varepsilon S_\varepsilon)^{-1}, \sigma_\varepsilon), \quad (13)$$

$$P(\Sigma^{*-1}) = W((\sigma_\beta S_\beta)^{-1}, \sigma_\beta), \quad (14)$$

where $£_1$ is a known matrix of dimensions a × r; relating the regression vector β^* to a parameter vector μ of dimension r × 1; possibly with r ≤ a; and the hyper parameters μ; C; σ_ε,S_ε, $\sigma_{\beta 1}$ and $S_{\beta 1}$ are assumed to be known. The notation W (Ω, ω) identifies a Wishart distribution with ω degrees of freedom and scale matrix Ω. With this structure, the joint posterior density of all the parameters can be written as:

$$p(\beta_1, \beta^*, \Sigma^{*-1}, \Sigma_\varepsilon^{-1} | Y) \propto |\Sigma_\varepsilon|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(Y - M\beta)' \Sigma_\varepsilon^{-1} (Y - M\beta) \right\}$$

$$\times |\Sigma^*|^{\frac{-1}{2}} \left\{ \exp\left(-\frac{1}{2}(\beta_1 - £_0\beta^*)' \Sigma^* (\beta_1 - £_0\beta^*) \right) \right\}$$

$$\times |C|^{\frac{-1}{2}} \left\{ \exp\left(-\frac{1}{2}(\beta^* - £_1\mu)' C^{-1} (\beta^* - £_1\mu) \right) \right\}$$

$$\times |\Sigma^*|^{\frac{-1}{2}(\sigma_{\beta_1} - k - 1)} \exp\left\{ -\frac{1}{2} tr\left[(\sigma_{\beta_1} S_{\beta_1}) \Sigma^{*-1} \right] \right\}$$

$$\times |\Sigma_\varepsilon|^{\frac{-1}{2}(\sigma_\varepsilon - k - 1)} \exp\left\{ -\frac{1}{2} tr\left[(\sigma_\varepsilon S_\varepsilon) \Sigma_\varepsilon^{-1} \right] \right\}$$

where the first line of the formula corresponds to the approximate likelihood function (eq.(10) and the others represent the prior information[2]. As said above, the marginal posterior densities of the parameters of interest can be obtained by integrating out the hyper parameters from the joint posterior density. The required integration does not provide closed form analytic solution in our case. However, a full Bayesian implementation of the model is still feasible, for instance with the help of the Gibbs sampler, an algorithm that, among other things, allows to derive marginal distributions from the full conditional densities of the parameter vector[6].

Predictive Analysis: In order to obtain the Bayesian forecasts for *h* periods, Let $y_f = (y_{t+1},\ldots,y_{t+h})$ be the forecast vector. The predictive density is

$$p(y_f \mid Y)=\int_{\beta_1}\int_{\beta^*}\int_{\Sigma_\varepsilon}\int_{\Sigma^*} p(y_f,\beta_1,\beta^*,\Sigma_\varepsilon,\Sigma^* \mid Y)\,d\beta_1\,d\beta^*\,d\Sigma_\varepsilon\,d\Sigma^*, \tag{16}$$

Notice that, the predictive distribution does not have a closed analytic form and it is also hard to calculate the integration numerically. However, the Gibbs sampling technique can be employed to compute the predictive distribution. Now, the equations for the *h* future observations being expressed as: y_f = Л $e_{L} + F\varepsilon_f$,

where.

$$\text{JI}=\begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \cdots & \cdots & \beta_{k-1} & \beta_k \\ \beta_2 & \beta_3 & \beta_4 & \cdots & \cdots & \beta_k & 0 \\ \beta_3 & \beta_4 & \beta_5 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_h & \beta_{h+1} & \cdots & \beta_k & 0 & \cdots & 0 \end{bmatrix}_{h\times k}$$

And

$$F=\begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \beta_1 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ \beta_2 & \beta_1 & 1 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_{h-1} & \beta_{h-2} & \cdots & \cdots & \beta_2 & \beta_1 & 1 \end{bmatrix}_{h\times h}$$

with $e_L = (e_t, e_{t-1}, \ldots, e_{t-k+1})'$ and $\varepsilon_f = (\varepsilon_{t+h}, \ldots, \varepsilon_{t+h})'$.

It should be noted that writing the future values in this way facilities writing the predictive distribution for any number of future values conational on the model parameters and initial errors. This enables the generation of any number of future realizations the autoregressive moving average model with exogenous inputs model that is needed for the Gibbs sampler. Since $\varepsilon_f \sim N(0, \Sigma_\varepsilon)$ with ε_f and ε being independent the conditional predictive density $p(y_f, \beta_1, \beta^*, \Sigma_\varepsilon, \Sigma^* \mid Y)$ is

$$p(y_f, \beta_1, \beta^*, \Sigma_\varepsilon, \Sigma^* \mid Y) = N(\text{Л}\, e_L, \Sigma_\varepsilon FF'))$$

The Bayesian forecasts of the future values can be obtained via extending the posterior analysis directly by stretching the parameter vector to include y_f. That is, the above posterior analysis will apply conditionally on the future observations y_f where the elements of y_f are sampled from the conditional predictive density (17).

Estimation with Gibbs Sampling: Conventional Gibbs sampling can be achieved with the help of Bayes' rule. let $\psi = (\beta_1, \beta^*, \Sigma_\varepsilon, \Sigma^*)$ be our parameter vector. If the prior distribution $P(\psi)$ is chosen to be conjugate to the likelihood then the posterior will be a known distribution. If the parameter vector $\psi = (\beta_1, \beta^*, \Sigma_\varepsilon, \Sigma^*)$ then the Gibbs sampler updates a component by component[3]. The proposed Gibbs sampling algorithm for the *ARMAX* (p, q, b) model can be implemented as follows:

1.Specify starting values $\beta_1^*, \beta^*, \Sigma^*$ and Σ_ε and and set j = 0. A set of initial estimates of the model parameters can be obtained using the IS technique of Koreisha and Pukkila[10].

2.Calculate the residuals recursively using (5) and the IS parameter estimates. 3.Draw the parameter vector ψ as follows:

$$\text{draw } \beta_1^j \text{ from } p\left(\beta_1 \mid y, (\beta^*)^{j-1}, (\Sigma^*)^{j-1}, (\Sigma_\varepsilon)^{j-1}\right) \propto \ell^*\left(\beta_1 \mid y, (\beta^*)^{j-1}, (\Sigma^*)^{j-1}, (\Sigma_\varepsilon)^{j-1}\right) p(\beta_1),$$

$$\text{- draw: } (\beta^*)^j \text{ from } p\left((\beta^*) \mid y, \beta_1^j, (\Sigma^*)^{j-1}, (\Sigma_\varepsilon)^{j-1}\right),$$

$$\text{- draw: } (\Sigma^*)^j \text{ from } p\left((\Sigma^*) \mid y, \beta_1^j, (\beta^*)^{j-1}, (\Sigma_\varepsilon)^{j-1}\right),$$

$$\text{- draw: } (\Sigma_\varepsilon)^j \text{ from } p\left((\Sigma_\varepsilon) \mid y, \beta_1^j, (\beta^*)^j, (\Sigma^*)^j\right),$$

4. Set j= j+1 and repeat step 3 several time.

This algorithm gives the next value of the Markov chain $\left\{\beta_1^{j+1}, (\beta^*)^{j+1}, (\Sigma^*)^{j+1}, and (\Sigma_\varepsilon)^{j+1}\right\}$ simulating each of the full conditionals where the conditioning elements are revised during the cycle. This iterative process is repeated for a large number of iterations and convergence is monitored. After the chain has converged, say at T_0 iterations, the simulated values $\left\{\beta_1^{j-1}, (\beta^*)^{j+1}, (\Sigma^*)^{j+1}, (\Sigma_\varepsilon)^{j+1}, j > T_0\right\}$, as a sample from the joint posterior. Posterior estimates of the parameters are computed direct via sample averages of the simulation outputs. The Gibbs sampler for $\beta_1; \beta^*; \Sigma^*$ Σ_ε and is easily seen to be specified by the following full conditional distributions, obtained from the joint posterior density above:

$$p(\beta_1 \mid Y, \beta^*; \Sigma_\varepsilon^{-1}, \Sigma^{*-1}) = N(\phi(M'\Sigma_\varepsilon^{-1} Y + \Sigma^{*-1} \pounds_0 \beta^*), \phi), \tag{18}$$

$$p(\beta^* \mid \beta_1; \Sigma_\varepsilon^{-1}, \Sigma^{*-1}) = N(\varphi(\pounds_0 \Sigma^{*-1} \beta_1 C^{-1} \pounds_1 \mu), \varphi), \tag{19}$$

$$p(\Sigma^{*-1} \mid Y, \beta_1; \beta^*, \Sigma_\varepsilon^{-1}) = W\left[\begin{pmatrix} (\beta_1 - \pounds_0 \beta^*) \\ (\beta_1 - \pounds_0 \beta^*)' \\ +\sigma_{\beta_1} S_\beta \end{pmatrix}^{-1}, \sigma_{\beta_1} + 1\right] \tag{20}$$

$$p(\Sigma_\varepsilon^{-1} \mid Y, \beta_1; \beta^*, \Sigma^{*-1}) = W\left[\begin{pmatrix} (Y - \mu\beta_1) \\ (Y - \mu\beta_1)' \\ +\sigma_\varepsilon S_\varepsilon \end{pmatrix}^{-1}, \sigma_\varepsilon + 1\right] \tag{21}$$

Where $\phi = (M'\Sigma_\varepsilon^{1} M + \Sigma^{*-1})$ and $\varphi = (\pounds_0' \Sigma^{*-1} \pounds_0 + C^{-1})^{-1}$, to model the data (and some parameters) by normal distributions, the framework is flexible enough to accommodate other distributional forms. Markov Chains Monte Carlo methods can then be easily used to check the sensitivity of inferences to the specific assumptions made.

Data and Empirical Evidences: In order to conduct the analysis, the proposed Bayesian methodology of the autoregressive moving average model with exogenous inputs (*ARMAX*) can be evaluated based on a simulated example, moreover, we demonstrate how the proposed Bayesian analysis can be used to analyze the electricity consumption data (y_t) in the industrial sector in Egypt. The empirical results suggest that, the electricity consumption variable is related with the consumed quantities of electricity in a previous period (y_{t-i}) and the number of consumers (dt_{-i}). The data set used in this paper consists of monthly time series observations covers the period (1991:1 – 2004:9)[4] for estimation purpose.

The proposed Bayesian analysis can be used to analyze the electricity consumption data (y_t), the consumed quantities of electricity in a previous period (y_{t-i}) and the number of consumers(d_{t-i}) in the industrial sector in Egypt.. In the empirical study, to deal with stationery series, all of the individual variables(in first difference)are transformed to the nature logarithms. Table 1 presents summary statistics for the three individual series.

As Table 1 shows, the summary statistics suggest that the three individual series that has, a small mean values which is dominated by a larger standard deviation, evidence of non-normality through (positive) skewness and excess kurtosis. Moreover, the results of the ADF test show that, the three individual series in first differences of logarithms are trend stationary at 5% significance level. In conclusion, all the three individual series have a single unit root or are integrated of degree one, I(1).

Monte Carlo Experiments: In this subsection, we present a simulated example to evaluate the efficiency of the proposed methodology. The example deals with generating 300 observations from the following *ARMAX* (1, 1, 1) model by Monte Carlo experiments. By using simulated data that are obtained from the following data generating process:

$$y_t = \gamma y_{t-1} + \xi \varepsilon_{t-1} + \eta d_{t-1} + \varepsilon_t, \quad (t = 1, .., T)$$

where, the error variance Σ_ε was chosen to be 1 and a non informative prior was assumed for γ, ξ, η and Σ_ε via setting $\Sigma_\gamma^{-1} = \Sigma_\xi^{-1} = \Sigma_\eta^{-1} = 0$. A normal p with zero mean and variance Σ_ε was used for the initial error ε_0. The starting values for the parameters (γ, ξ, η and Σ_ε) were obtained using IS method. The starting values for ε_0 and y_f was assumed to be zero.

The Gibbs sampler was run with 10000 iterations and every tenth value in the sequences of the 10000 draws is recorded, to obtain an approximately independent samples. All posterior and predictive estimates were computed as sample averages of the simulated outputs. Table 2 presents the true values and Bayesian estimates of the parameters (γ, ξ, η and Σ_ε) and the next five future values (y_{301}, y_{302}, y_{303}, y_{304}, y_{305}). Moreover, a 95% confidence interval using 0.025 and o.975 percentiles of the simulated draws is constructed for every estimates and future values. Table 2 shows that, Bayesian estimates are closed to the true values of the parameters except the last two future values (y_{304}, y_{305}) which are dominated by a larger standard deviations. In other words, these findings concluded that the proposed Bayesian methodology of the autoregressive moving average model with exogenous inputs (*ARMAX*) is stable and fluctuates in the neighborhood of the true values of the parameters.

Table 1: Descriptive statistics

Measures / variables	The electricity consumption (y_t)	The electricity in a previous period (y_{t-1})	The number of consumers (d_{t-1})
Mean	-0.0102(0.4015)	-0.0101(0.5114)	-0.0167(0.2325)
Median	-0.01227(0.2531)	-0.0117(0.3652)	-0.0192(0.2234)
skewness	0.111(0.4704)	0.054(0.1215)	0.146(0.1114)
kurtosis	- 0.342 (0.1163)	-0. 315(0.2124)	-0.414 (0.1291)
ADF	-17.255*	-28.5757*	-24.8547*

Notes: The null hypothesis in the augmented Dickey Fuller (ADF) test is that there exists a unit root in the time series, that is, the time series is non-stationary. The null hypothesis is rejected if the ADF statistic is greater than the Mackinnon critical values. The ADF test lag length was determined by AIC and BIC criteria.* Test statistic significant in rejecting the null hypothesis at the 5% level. The numbers in parentheses are the standard deviations.

Table 2: Bayesian results of the simulated examples

Parameters	True values	Mean	Std Dev.	Lower (95%) Limit	Median	Upper (95%)Limit
γ	0.6378	0.6139	9.214E-3	0.5252	0.5997	0.7808
ξ	0.3435	0.3131	0.0087	0.3025	0.3067	0.5454
η	0.4077	0.3656	4.48E-3	0.3298	0.3541	0.4706
Σ	1.0000	0.9717	0.0022	0.7974	0.9686	1.0559
y_{301}	-1.1084	-1.1381	0. 1002	-1.0184	-1.0998	-1.2967
y_{302}	-1.1521	-1.3184	0.2103	-0.9282	-1.2639	-1.4209
y_{303}	0.3596	0.0115	0.1514	1.1283	0.1185	0.6083
y_{304}	1.1024	1.1184	0. 4109	1.0184	1.1298	1.3967
y_{305}	0.6105	0.1024	0. 6109	-0.5487	-0.2302	0.8584

Table 3: Autocorrelations results of the simulated examples

Parameters	Lag 1	Lag 5	Lag 10	Lag 20	Lag 30	Lag 50
γ	-0.2151	-.04221	-0.0541	-0.02282	-0.0095	-0.0027
ξ	-0.0151	0.06991	-0.0341	-0.0252	0.0018	0.0031
η	-0.09151	-0.0213	-0.0341	-0.0252	0.0018	0.0031
Σ	0.0142	0.2391	-0.0212	-0.1233	0.04568	-0.0582
y_{301}	-0.1551	0.0977	-0.0160	-0.02101	-0.0024	-0.0032
y_{302}	-0.2562	-0.06991	-0.0141	-0.0295	-0.0231	0.0081
y_{303}	0.1784	0.02524	-0.0187	-0.0463	-0.0746	0.0128
y_{304}	-0.2248	-0.04254	-0.0813	-0.06441	-0.05811	0.1184
y_{305}	0.1214	0.2104	-0.1146	-0.02158	-0.1077	-0.0143

The convergence of the proposed Bayesian methodology is illustrated based on the autocorrelations for the simulated example are displayed in Table 3. Table 3 shows that the draws for each parameter (γ,ξ,η and Σ_ε) or five future values (y_{301}, y_{302}, y_{303}, y_{304}, y_{305}) had small autocorrelations at lags 1, 5,10, 20 and 30, which means that, the convergence of the proposed Bayesian methodology was achieved. This conclusion was confirmed at lag 50.

The Electricity Consumption Data: The electricity consumption data (y_t) in the industrial sector in Egypt consist of 165 monthly observations. As noted above, all of the individual variables (in first difference) are transformed to the nature logarithms.

In this subsection, we demonstrate how the proposed Bayesian analysis can be used to analyze the electricity consumption data in the industrial sector in Egypt. Using the model *ARMAX* (1,1,1) in eq.(22),

Table 4: Bayesian results for the differenced electricity consumption data using *ARMAX* (1,1,1)

Parameters	True values	Mean	Std Dev.	Lower (95%)Limit	Median	Upper (95%)Limit
γ	-	0.5413	0.0025	0.3235	0.5398	0.6115
ξ	-	03151	0.0008	0.2857	0.3067	0.3127
η	-	0.3324	0.0027	0.3054	0.3299	0.3401
Σ	-	0.0189	0.0054	0.0178	0.0181	0.0201
Z_{161}	-0.1216	0.0544	0.0124	-0.0310	0.0024	-0.0154
Z_{162}	0.1315	0.0421	0.0067	-0.0178	-0.0115	0.0589
Z_{163}	-0.1451	-0.2010	0.0122	-0.1664	-0.1356	0.0313
Z_{164}	0.1165	0.1010	0.0005	-0.06441	0.0991	0.1356
Z_{165}	0.1354	0.1332	0.0242	0.1296	-0.2118	0. 1431

the proposed Bayesian methodology is applied to the first 160 values of the differenced electricity consumption series Z_t and the last five values are set aside for forecasting evaluation. Then all initial values of the parameters (γ,ξ,η and Σ_ε) and the next five future values (Z_{161}, Z_{162}, Z_{163}, Z_{164}, Z_{165}) are chosen as in the simulated examples. Table 4 summarizes the posterior and predictive results for the differenced electricity consumption data in the industrial sector in Egypt series.

Table 4 shows that,all of the true values of the parameters or the future values are within the 95% confidence interval formed by the 0.025 and 0.975 percentiles. However, the final conclusion confirmed that the proposed algorithm has good statistical properties. It should be noted that uncertainty about the parameters and initial errors were incorporated. However, there exist some non Bayesian papers that are concerned with the prediction error arising from uncertainly about parameter estimation, such as Yamamoto[15].

Conclusion: In this paper, we developed a simple and fast way of estimating the parameters of the autoregressive moving average with exogenous inputs (*ARMAX*) model using the output of the Gibbs sampling. The procedure has the advantage that can be used under informative and non-informative priors on the parameters of interest. Moreover, this procedure does not require the estimation of two models, one with and the other without the restriction to be tested if the problem of testing a set of restrictions in the autoregressive moving average model with exogenous inputs (*ARMAX*) model is considered. The empirical results of the simulated examples and electricity consumption data in the industrial sector in Egypt showed the accuracy of the proposed methodology and has a good has good statistical properties. An extensive check of convergence of the parameters and the future values using autocorrelations at lags 1, 5,10, 20 and 30, which means that, the convergence of the proposed Bayesian methodology was achieved.This conclusion was confirmed at lag 50.

Future work may investigate the problem of testing a set of linear restrictions in the autoregressive moving average model with exogenous inputs (*ARMAX*).

REFERENCES

1. Altinay, G. and E. Karagol, 2004. Structural break, unit root and the causality between energy consumption and GDP in Turkey, Energy Economics, 26: 985-994.
2. Box, G.E.P. and G.M. Jenkins, 1976. Time series analysis". Forecasting and Control. Holden-Day, San Francisco.
3. Chib, S. and E. Greenberg, 1994. Bayes inference in regression models with ARMA (p, q). J. Econom., 64: 183-206.
4. Ciccarelli, M., 2004. Testing restrictions in hierarchical normal data models using Gibbs sampling. Res. Econom., 58: 135-157.
5. DuMouchel, W.H. and J.E. Harris, 1983. Bayes methods for combining the results of cancer studies in humans and other species: with discussion, J. Am. Statistical Assoc., 78: 293-315.
6. Gelfand, A.E. and A.F.M. Smith, 1990. Sampling-based approaches to calculating marginal densities. J. Am. Statistical Assoc., 85: 398-409.
7. Harvey, A., 1994. Time Series Models, 2nd Ed. Cambridge, MA: MIT Press.
8. Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu, 1999. Bayes estimation of short run coefficients in dynamic panel data models". In: Hsiao, C., *et al.*

(Eds.), "Analysis of Panels and Limited Dependent Variable Models": In Honor of G.S. Maddala. Cambridge University Press, Cambridge.
9. Ismail, M.A., 2003. Bayesian Analysis of the Seasonal Moving Average Model: A Gibbs Sampling Approach. Japanese J. Appl. Statistics, 32: 61-75.
10. Koreisha and Pukkila, 1990. Linear methods for estimating ARMA and regression models with serial correlation, Communications in Statistics Simulation, 19: 71-102.
11. Lindley, D.V. and A.F.M. Smith, 1972. Bayes estimates for the linear model: with discussion, J. Royal Statistical Soc., Series B34: 1-41.
12. Marriott, J., N. Ravishanker, A. Gelfand and J. Pai, 1996. Bayesian analysis of ARMA process: 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.
13. Rubin, D.B., 1981. Estimation in parallel randomized experiments, J. Educational Statistics, 6: 377-401.
14. Walker, D.M., F.J.P. Brarberia and G. Marion, 2006. Stochastic modeling of ecological processes using hybrid Gibbs samplers. Ecological Modeling, 198: 40-52.
15. Yamamoto, T., 1981. Prediction of multivariate autoregressive moving average models. Biometrika, 68: 475-492.

[1]Altinay *et al.* (2004).
[2]Ciccarelli, (2004).
[3]Walker *et al.*,(2006).
[4]Source: Egyptian Ministry of Electricity & Energy, Various Issues.