5doi:10.21608/esju.2024.318938.1044 Research Article
The Egyptian Statistical Journal (ESJ), 68 (2): 129-145

Homepage: https://esju.journals.ekb.eg/

The Egyptian Statistical Journal

Print ISSN 0542-1748– Online ISSN 2786-0086

Bayesian Identification of Seasonal Vector ARMA Processes

Samir M. Shaarawy[1], Sherif S. Ali[2], Emad El-Din A. Salam[2*]

Received 9 September 2024; revised 2 November 2024; accepted 3 December 2024

Keywords

Seasonal multivariate ARMA processes, semi-Bayesian identification, Matrix Normal-Wishart prior, Jeffreys' prior, posterior probability mass function.

Abstract

This research paper uses the Bayesian approach to establish an approximate method to specify the four orders of multivariate seasonal autoregressive moving average (SARMA) models. The proposed methodology consists of four coherent consecutive steps. The first step is the approximation of the likelihood function of the model's parameters by a matrix Normal–Wishart form. The second step is to use a proposed semi-intermediate Bayesian procedure to have initial estimates for the four model orders. The third step is to combine the approximate likelihood function and the initial orders prior with one of the matrix Normal–Wishart prior density or Jeffreys' vague prior to developing an approximate joint posterior probability mass function of the orders of the model in a simple form. The last step is to evaluate the posterior probabilities over the range of the four orders and pick out the values of the orders at which the joint probability mass function reaches the highest probability to be the identified orders of the multivariate seasonal time series being analyzed. To test the adequacy of the proposed methodology, two simulation studies with three different prior orders have been conducted. The numerical results showed that the proposed Bayesian methodology is adequate to identify the orders of multivariate SARMA models for medium and large time series lengths.

1. Introduction

Analysis of seasonal multidimensional time series is one of the most important and major topic and may be found in most fields of scientific research such as economics, meteorology, hydrology, and utilities. An economist might be interested in using a three-dimensional model to check the interdependencies among quarterly sales volume, sales prices, and advertising costs for a product. Using a three-dimensional model, a meteorologist may be willing to model and forecast the hourly temperature, air pressure, and humidity percentages. A hydrologist sometimes wants to employ a four-dimensional model to examine the feedback relationships between stream flow series at four points along a river. In utilities, one may use the two-dimensional model to model and forecast the monthly consumption of electricity and gas. Such variables and others are modeled and predicted jointly using a multidimensional (vector) model to test the dynamic interrelationship between variables and to raise the accuracy of the obtained estimates and forecasts; for more details see Tiao and Box (1981), Tsay (2014), Brockwell and Davis (2016), Box et.al. (2016) and Wei (2019).

There is no doubt that the parametric class of seasonal multidimensional (vector) autoregressive moving average models, abbreviated by SVARMA for short, is the widely accepted class of models to analyze multivariate time series arise in many areas of application, especially in economics,

✉ Corresponding author*: easalam@hotmail.com
[1]Department of Information Systems and Operations Management, College of Business Administration, Kuwait University, Kuwait.
[2] Department of Statistics, Faculty of Science, King Abdul-Aziz University, Jiddah City, Saudi Arabia Kingdom.

since it is the most parsimonious one and is a closed set under linear transformation (see Lütkepohl, 2007). In addition, the goodness of the forecasts achieved by this class outruns the ones obtained by the pure multivariate autoregressive models (See for details Hagnell (1991), Athanasopoulos and Vahid (2008), Raghavan et al. (2009) and Kascha (2012)). Moreover, SVARMA is more compatible with economic theory than the pure vector autoregressive models (See Cooley and Dwyer (1998)). VARMA models were first used by Quenouille (1957). Since then, they have become an important and serious topic for theoretical and applied researchers in multidimensional time domain analysis such as: Tiao and Box (1981), Tiao and Tsay (1983), Tsay (1989), Reinsel (1997) and Athanasopoulos and Vahid (2008).

In *SVARMA* analysis, the first and most important stage is identifying the process orders p, q, P, and Q that might generate the available data set. p is the nonseasonal (regular) autoregressive order and q is the nonseasonal (regular) moving average order; while P is the seasonal autoregressive order, and Q is the seasonal moving average order. This stage is very crucial since the accuracy of all the succeeding stages (estimation, diagnostic checking, and forecasting) depends on its accuracy. The four orders p, q, P, and Q are often not known, and we have to use a convenient algorithm to identify them using the observed time series.

Regarding the seasonal univariate autoregressive moving average processes, Box and Jenkins have explored their predominant non-Bayesian methodology to identify the orders by matching the patterns of both the sample autocorrelation and partial autocorrelation functions with the known theoretical patterns of autocorrelation and partial autocorrelation functions autocorrelation and partial autocorrelation functions. Their methodology has been illustrated and used by many others such as Priestley (1981), Tsay (1984), Harvey (1993), Wei (2005), Liu (2009), Wei (2019), Box et al. (2016), Brockwell and Davis (2016) and Chatfield (2019). Another prevailing non-Bayesian class of methods to identify the orders of univariate time series is known as the exploratory or the objective approach. The approach starts by fitting all possible autoregressive moving average models, assuming the maximum orders are known, and computing a certain criterion for each model; then we may pick out the model with the optimal value of the proposed criterion. However, the researchers do not agree about the form of the criterion that should be optimized. Akaike's Information Criterion (AIC), introduced by Akaike (1974), is considered one of the most preferred and well-known exploratory techniques. Other exploratory techniques were introduced by Schwarz (1978), Risanen (1978), Hannan and Quinn (1979) and Beveridge and Oickle (1994). However, we should observe that the use of any automatic technique for identifying seasonal models is costly and time-consuming.

With respect to the Bayesian approach to specifying orders for univariate processes, Monahan (1983) has introduced a numerical method to solve the identification problem of non-seasonal processes with low orders. Assuming the maximum orders are known, Broemeling and Shaarawy (1988) have introduced an approximate analytical algorithm for identifying the orders of non-seasonal processes depending on t distribution. Shaarawy and Ali (2003) have initiated a solution, based on the Bayesian approach, to identify the orders of pure seasonal autoregressive models. Moreover, Shaarawy et al. (2007) have conducted an approximate algorithm to achieve a joint mass probability posterior distribution of the orders of pure non–seasonal moving average processes.

By comparing the sample cross-correlation functions with their theoretical patterns, Tiao and Box (1981), Tiao and Tsay (1983) and others have studied non-Bayesian identification problems of multidimensional processes. However, the Bayesian techniques to the identification problems of multidimensional processes have recently been studied. Shaarawy and Ali (2008) have laid the foundations of the Bayesian solutions for identifying orders for pure non-seasonal multidimensional autoregressive models by deriving an exact convenient formula for the joint probability mass function of the model order. More recently, Shaarawy and Ali (2015) continued to explore a new approximate algorithm for identifying the orders of pure seasonal multidimensional autoregressive models. Later on, Shaarawy (2021) has developed a complete approximate analysis to solve the problems of identification, estimation, diagnostic checking, and forecasting for non-seasonal multidimensional autoregressive moving average processes. Regarding the seasonal multidimensional autoregressive moving average processes, we can say that a pure Bayesian procedure for the identification stage has not been discovered yet.

A Basic difficulty with the identification problems of the seasonal multidimensional autoregressive moving average processes is that there is no convenient form for likelihood function as a function of the parameters directly. An enormous number of calculations are indispensable to compute the likelihood function for any point in the parameter space; thus, as the time series length becomes large, the computation of the likelihood function becomes increasingly exhausted even for effective computers. On the analytical side, as compared with purely numerical, a convenient Bayesian identification of seasonal multidimensional autoregressive moving average processes is impossible unless we find a way to represent the likelihood function in a simple way to achieve analytically tractable posterior mass function for the model order. This causes a challenging problem, for both Bayesians and those interested in maximum likelihood estimation. The fundamental target of the current paper is to derive an approximate joint posterior probability mass function for the orders of seasonal multidimensional autoregressive moving average processes in a convenient form using a matrix Normal–Wishart or Jeffreys' vague prior. Then we can easily calculate posterior probabilities for all the possible values of the model's orders over their domain and select the values at which the probability reaches its highest value to be a solution to the identification problem. Two Monte Carlo simulation studies, with three different priors, were carried out to demonstrate the idea of using the proposed algorithm and assess its numerical performance. The numerical performance will be checked for the two-dimensional ARMA models based on the parameter's values and sample sizes. The two-dimensional model is quite useful in modeling and forecasting two variables where the dependencies between them are important and where analyzing variables individually loses important aspects of the data.

2. Seasonal Vector ARMA Processes

The vector time series y(t) of dimension k is said to be generated from a seasonal vector autoregressive moving average (*SVARMA*) model with orders p, q, P and Q, which are four positive integers, if it has the form

$$\phi_p(B)\Phi_P(B^s)y(t) = \theta_q(B)\Theta_Q(B^s)\varepsilon(t) \quad , t = 1,2,...,n \ , \tag{2.1}$$

where

$$\phi_p(B) = I_k - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p, \ \Phi_P(B^s) = I_k - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_p B^{Ps},$$

$$y(t) = \left[y(t,1) \quad y(t,2)... \quad y(t,k) \right]', \ \theta_q(B) = I_k - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q,$$

$\Theta_Q(B^s) = I_k - \Theta_1 B^s - \Theta_2 B^{2s} - ...\Theta_Q B^{Qs}$, and $\varepsilon(t) = [\varepsilon(t,1) \quad \varepsilon(t,2)... \quad \varepsilon(t,k)]'$

I_k is the identity matrix of rank k, B is the back-shift operator defined as $[B^r\, y(t)] = y(t-r)$, s is the periodicity of the time series, ϕ_i's, θ_i's, Φ_i's, and Θ_i's are matrices of real unknown coefficients, each of order $k \times k$. Moreover, $\{\varepsilon\ (t)\}$ is a sequence of random errors vectors that are identically independent and normally distributed with zero mean vector and a $k \times k$ unknown precision matrix T. Furthermore, $\phi_p(B)$, $\Phi_P(B^s)$, $\theta_q(B)$ and $\Theta_Q(B^s)$ are $k \times k$ matrix polynomials in B having degrees p, Ps, q and Qs respectively. In the literature on time series analysis, $\phi_p(B)$ and $\Phi_P(B^s)$ are defined as the non-seasonal and seasonal autoregressive operators with orders p and P_S respectively; while $\theta_q(B)$ and $\Theta_Q(B^s)$ are defined as the non-seasonal and seasonal moving average operators with orders q and Q_S respectively. The stationarity condition of the process $y(t)$ is that the zeros of the determinantal polynomials $|\phi_p(B)|$ and $|\Phi_P(B^s)|$ are all outside the unit circle and the invertibility condition is that the zeros of the determinantal polynomials $|\theta_q(B)|$ and $|\Theta_Q(B^s)|$ are all outside the unit circle.

To express the class of models (2.1) in explicit form, we condition on the first $(p+Ps)$ vectors of observations, then:

$$\mathrm{Y} = \mathrm{X}(p,q,P,Q)\Omega(p,q,P,Q) + U \tag{2.2}$$

Where Y is a time series matrix of dimension $(n-p-Ps)\times k$ where the ij-th element is referred to as $y(p+Ps+i,j)$ and $X(p,q,P,Q)$ is a matrix of dimension $(n-p-Ps)\times kh$, where $h = p+P+pP+q+Q+qQ$, defined as,

$$X = \begin{bmatrix} X_1 & X_2 & Z_1 & Z_2 & R_1 & R_2 ... & R_p & H_1 & H_2 ... & H_q \end{bmatrix}$$

where

$$X_1 = \begin{bmatrix} y'(p+Ps) & y'(p+Ps-1) & \cdots & y'(Ps+1) \\ y'(p+Ps+1) & y'(p+Ps) & \cdots & y'(Ps+2) \\ \vdots & \vdots & & \vdots \\ y'(n-1) & y'(n-2) & \cdots & y'(n-p) \end{bmatrix},$$

$$X_2 = \begin{bmatrix} -\varepsilon'(p+Ps) & -\varepsilon'(p+Ps-1) & ... & -\varepsilon'(p+Ps+1-q) \\ -\varepsilon'(p+Ps+1) & -\varepsilon'(p+Ps) & \cdots & -\varepsilon'(p+Ps+2-q) \\ \vdots & \vdots & & \vdots \\ -\varepsilon'(n-1) & -\varepsilon'(n-2) & \cdots & -\varepsilon'(n-q) \end{bmatrix},$$

$$Z_1 = \begin{bmatrix} y'(p+Ps+1-s) & y'(p+Ps-1) & \cdots & y'(p+1) \\ y'(p+Ps+2-s) & y'(p+Ps+2-2s) & \cdots & y'(p+2) \\ \vdots & \vdots & & \vdots \\ y'(n-s) & y'(n-2s) & \cdots & y'(n-Ps) \end{bmatrix},$$

$$Z_2=\begin{bmatrix} -\varepsilon'(p+Ps+1-s) & -\varepsilon'(p+Ps+1-2s) & \dots & -\varepsilon'(p+Ps+1-Qs) \\ -\varepsilon'(p+Ps+2-s) & -\varepsilon'(p+Ps+2-2s) & \cdots & -\varepsilon'(p+Ps+2-Qs) \\ \vdots & \vdots & & \vdots \\ -\varepsilon'(n-s) & -\varepsilon'(n-2s) & \cdots & -\varepsilon'(n-Qs) \end{bmatrix},$$

$$R_i=\begin{bmatrix} -y'(p+Ps-s-i+1) & -y'(p+Ps-2s-i+1) & \cdots & -y'(p-i+1) \\ -y'(p+Ps-s-i+2) & -y'(p+Ps-2s-i+2) & \cdots & -y'(p-i+2) \\ \vdots & \vdots & & \vdots \\ -y'(n-s-i) & -y'(n-2s-i) & \cdots & -y'(n-Ps-i) \end{bmatrix},$$

$$i=1,2,\dots,p$$

And

$$H_i=\begin{bmatrix} \varepsilon'(p+Ps-s-i+1) & \varepsilon'(p+Ps-2s-i+1) & \dots & \varepsilon'(p+Ps-Qs-i+1) \\ \varepsilon'(p+Ps-s-i+2) & \varepsilon'(p+Ps-2s-i+2) & \cdots & \varepsilon'(p+Ps-Qs-i+2) \\ \vdots & \vdots & & \vdots \\ \varepsilon'(n-s-i) & \varepsilon'(n-2s-i) & \cdots & \varepsilon'(n-Qs-i) \end{bmatrix},$$

$$i=1,2,\dots,q$$

These matrices show that the columns of the matrix $X(p,q,P,Q)$ consist of the elements of the regressors $y(t-1),y(t-2),\cdots,-y(t-Ps-p),-\varepsilon(t-1),-\varepsilon(t-2),\cdots,\varepsilon(t-QS-q)$, respectively. Furthermore, $\Gamma(p,\ q,\ P,Q)$ is a $kh\times k$ matrix of coefficients defined by

$$\Omega(p,p,q,Q)=\begin{bmatrix} \underset{kp\times k}{\phi(\mathrm{p})} & \underset{kq\times k}{\theta(\mathrm{q})} & \underset{kP\times k}{\Phi(\mathrm{P})} & \underset{kQ\times k}{\Theta(\mathrm{Q})} & \overbrace{\gamma_1(p,P)}^{kpP\times k} & \overbrace{\gamma_2(q,Q)}^{kqQ\times k} \end{bmatrix}',$$

where $\phi(p)=\begin{bmatrix} \underset{kxk}{\phi_1'} \\ \dots\dots \\ \underset{kxk}{\phi_2'} \\ \dots\dots \\ \vdots \\ \dots\dots \\ \underset{k\times k}{\phi_p'} \end{bmatrix}$, $\theta(q)=\begin{bmatrix} \underset{kxk}{\theta_1'} \\ \dots\dots \\ \underset{kxk}{\theta_2'} \\ \dots\dots \\ \vdots \\ \dots\dots \\ \underset{k\times k}{\theta_q'} \end{bmatrix}$, $\theta_i=\begin{bmatrix} \theta_{i.11} & \theta_{i.12} & \dots\theta_{i.1k} \\ \theta_{i.21} & \theta_{i.22} & \dots\theta_{i.2k} \\ \vdots & & \\ \theta_{i.k1} & \theta_{i.k2} & \dots\theta_{i.kk} \end{bmatrix}$, $i=1,2,\dots,q$

$$\phi_i = \begin{bmatrix} \varphi_{i.11} & \varphi_{i.12} & \dots \varphi_{i.1k} \\ \varphi_{i.21} & \varphi_{i.22} & \dots \varphi_{i.2k} \\ \vdots & & \\ \varphi_{i.k1} & \varphi_{i.k2} & \dots \varphi_{i.kk} \end{bmatrix} \quad , i = 1,2,...,p$$

$$\Phi(P) = \begin{bmatrix} {}^{kxk} \\ \Phi_1^{'} \\ \dots\dots \\ {}^{kxk} \\ \Phi_2^{'} \\ \dots\dots \\ \vdots \\ \dots\dots \\ {}^{kxk} \\ \Phi_P^{'} \\ \end{bmatrix}, \ \Phi_i = \begin{bmatrix} \Phi_{i.11} & \Phi_{i.12} \dots \Phi_{i.1k} \\ \Phi_{i.21} & \Phi_{i.22} \dots \Phi_{i.2k} \\ \vdots & \vdots & \vdots \\ \Phi_{i.k1} & \Phi_{i.k2} \dots \Phi_{i.kk} \end{bmatrix}, i = 1,2,...,P$$

$$\Theta(Q) = \begin{bmatrix} {}^{kxk} \\ \Theta_1^{'} \\ \dots\dots \\ {}^{kxk} \\ \Theta_2^{'} \\ \dots\dots \\ \vdots \\ \dots\dots \\ {}^{kxk} \\ \Theta_Q^{'} \\ \end{bmatrix}, \ \Theta_i = \begin{bmatrix} \Theta_{i.11} & \Theta_{i.12} \dots \Theta_{i.1k} \\ \Theta_{i.21} & \Theta_{i.22} \dots \Theta_{i.2k} \\ \vdots & \vdots & \vdots \\ \Theta_{i.k1} & \Theta_{i.k2} \dots \Theta_{i.kk} \end{bmatrix}, i = 1,2,...,Q$$

$$\gamma_1(p,P) = \gamma_1(\phi_i, \Phi_j), i = 1,2\dots, p\,;\ j = 1,2,\dots, P,$$

and

$$\gamma_2(q,Q) = \gamma_2(\theta_i, \Theta_j), i = 1,2\dots, q\,;\ j = 1,2,\dots, Q.$$

Finally, the matrix $U(p, P)$ is the errors' matrix of dimension $(n - p - Ps) \times k$ where the ij-th element is referred to as $\varepsilon(p + Ps + i, j)$.

It should be known that the dimension of the regressor matrix $X(p,q, P, Q)$ depends on the model orders p, q, P and Q. That is, for each specific set of orders, say p_0, q_0, P_0, and Q_0, there is a specific matrix $X(p_0, q_0, P_0, Q_0)$. One may also notice that the model parameters $\gamma_1(p,P) = \gamma_1(\varphi_i, \Phi_j)$

and $\gamma_2(q,Q)=\gamma_2(\theta_i,\Theta_j)$ will be manipulated as free parameters in the proposed Bayesian identification methodology in order to have a tractable likelihood function. If n is large enough, the approximate likelihood function will express a good approximation to the exact one.

3. Initial Estimates of the Model Orders

The seasonal vector autoregressive moving average (*SVARMA*) having the orders p, q, P, and Q, is denoted by *SARMA*(p,q)(P, Q). It is useful for modeling and forecasting multidimensional time series data and usually p, q, P, and Q do not exceed two. In practice, the values of the four orders p, q, P, and Q are not known, and we must use the observed n vectors of observations S_n to identify them. The formal Bayesian procedure to identify these four values is to derive their joint posterior probability mass function. Then we can investigate the probabilities over the range of the four orders and select the combination of values of the orders with maximum probability to be the values of the orders determined by the proposed technique. The technique employed here is slightly different from the pure Bayesian technique. Instead of manipulating the joint posterior distribution of p, q, P, and Q, it is suggested to consider the joint posterior density of the coefficients matrix Ω_0 where

$$\Omega_0^{'}=\left[\varphi_1^{'}\ \ \varphi_2^{'}\ \cdots\ \varphi_l\ \ \theta_1^{'}\ \ \theta_2^{'}\cdots\theta_m\ \ \Phi_1^{'}\ \ \Phi_2^{'}\cdots\Phi_v\ \ \Theta_1^{'}\ \ \Theta_2^{'}\ \ \cdots\Theta_r^{'}\ \ \gamma_{10}^{'}(l,v)\ \ \gamma_{20}(m,r)\right] \tag{3.1}$$

where $\phi_i,\ i=1,\ 2,\ldots,l$; $\theta_i,\ i=1,2,\ldots,\ m$; $\Phi_j, j=1,2,\ldots,v$; $\Theta_j, j=1,2\ldots,r$ are as defined in the previous section and $\gamma_{10}(l,v)=\gamma_{10}(\varphi_i,\Phi_j)$; $\gamma_{20}(m,r)=\gamma_{20}(\theta_i,\theta_j)$. The maximum orders l, m, v and r are assumed to be known.

Assuming that $\varepsilon(0)=\varepsilon(-1)=\cdots=\varepsilon(1-m-rs)=0,$ the likelihood function of the model parameters Ω_0 and T can be written as

$$L(\Omega_0,\mathrm{T}\,|\,S_n)\propto(2\pi)^{-k(n-l-vs)/2}\,|\,\mathrm{T}\,|^{(n-l-vs)/2}\exp(-\tfrac{1}{2}tr\{\sum_{t=l+vs+1}^{n}\varepsilon(t)\varepsilon'(t)\mathrm{T}\}), \tag{3.2}$$

where

$\Omega_0\in T^{kg\times k},\ \mathrm{T}>0,\ \mathrm{g}=l+v+lv+r+m+rm\ \ l,\ T>0,$

$$\varepsilon'(t)=y'(t)-x'(t-1)\,\Omega_0 \tag{3.3}$$

and

$$\begin{aligned}x'(t-1)=[\,&y'(t-1)\ \ y'(t-2)\cdots y'(t-l)\ \ -\varepsilon'(t-1)\ \ -\varepsilon'(t-2)\cdots\ -\varepsilon'(t-m)\\ &y'(t-s)\ \ y'(t-2s)\cdots\ y'(t-vs)\ \ -\varepsilon'(t-s)\ \ -\varepsilon'(t-2s)\cdots-\varepsilon'(t-rs)\\ &-y'(t-s-1)\ \ -y'(t-2s-1)\cdots\ y'(t-vs-1)\ \ -y'(t-s-2)\ \ -y'(t-2s-2)\\ &\cdots-y'(t-vs-2)\cdots-y'(t-s-l)\ \ y'(t-2s-l)\cdots y'(t-vs-l)\ \ \varepsilon'(t-s-1)\\ &\varepsilon'(t-2s-1)\cdots\ \varepsilon'(t-rs-1)\ \ \varepsilon'(t-s-2)\ \ \varepsilon'(t-2s-2)\cdots\varepsilon'(t-rs-2)\\ &\cdots\ \varepsilon'(t-s-m)\ \ \ \varepsilon'(t-2s-m)\ \cdots\ \ \varepsilon'(t-rs-m)\,]\end{aligned}$$

The expression in equation (3.3) is a recurrence relation in terms of the residuals. This recurrence causes the main problem in the exact analysis of the seasonal multidimensional moving average

process. However, if one knows Ω_0 and the initial values of the residuals, this recurrence formula can be used to evaluate the residuals recursively. The suggested approximation depends on replacing each exact residual $\varepsilon(t-j)$ by its least squares estimate. Moreover, the initial values of the residuals are assumed to equal their unconditional means, namely zero. Thus, one estimates the residuals recursively by

$$\hat{\varepsilon}'(t) = y'(t) - \hat{x}'(t-1)\,\hat{\Omega}_0$$

Where $\hat{\Omega}_0$ are the nonlinear least squares estimates of the coefficients Ω_0 and $\hat{x}(t-1)$ are the same as $x(t - 1)$ but using the estimated residuals instead of the exact ones. Using the estimates of the residuals, we can write the likelihood function (3.2) approximately as follows;

$$L^*(\Omega_0, \mathrm{T} \mid S_n) \propto (2\pi)^{-k(n-l-vs)/2} \mid \mathrm{T} \mid^{(n-l-vs)/2} \exp\Big(-\frac{1}{2} tr \sum_{t=l+vs+1}^{n} [\, y(t) - \Omega_0{}' \hat{x}(t-1)][y(t) - \Omega_0{}' \hat{x}(t-1)\,]\,\mathrm{T}$$

An appropriate selection for the prior distribution of the model's parameters Ω_0 and T is a matrix Normal–Wishart distribution, i.e.

$$\xi(\Omega_0, T) = \xi_1(\Omega_0 \mid T)\; \xi_2(t) \qquad (3.5)$$

where

$$\xi_1(\Omega_0 \mid T) \propto |T|^{kg/2} \exp\Big(-\frac{1}{2}\, tr(\Omega_0 - D)'V(\Omega_0 - D)T\Big)$$

$$\xi_2(\mathrm{T}) \propto |\,\mathrm{T}\,|^{[a-(k+1)]/2} \exp\Big(-\frac{1}{2}\, tr\psi \mathrm{T}\Big)$$

and the hyper-parameters $D \in R^{kg\times k}, V$ is a positive definite matrix of dimensions $kg \times kg$, and ψ is a positive definite matrix of dimensions $k \times k$. If one has no or little prior information about the parameters, one may use Jeffrey's vague priori

$$\xi(\Omega_0, \mathrm{T}) \propto |\,\mathrm{T}\,|^{-(k+1)/2}, \Omega_0 \in R^{kg\times k}, \mathrm{T} > o \qquad (3.6)$$

Theorem 3.1: combining the approximate likelihood function of the model given in (3.4) with the matrix Normal–Wishart prior density given in (3.5), it is found that, the posterior distribution of Ω_0 is a matrix t distribution with parameters $(\hat{\Omega} = A^{-1}B, A^{-1}, C - B'A^{-1}B, n^*)$, where

$$A = V + \sum_{t=l+vs+1}^{n} \hat{x}(t-1)\hat{x}'(t-1),$$

$$B = VD + \sum_{t=l+vs+1}^{n} \hat{x}(t-1)y'(t),$$

$$C = D'VD + \psi + \sum_{t=l+vs+1}^{n} y(t)y'(t) \text{ and } n^* = n - k + a + 1$$

Corollary 3.1: combining the approximate likelihood function (3.4) with Jeffrey's vague prior (3.6), the posterior distribution of Ω_0 is a matrix t with parameters

$(\hat{\Omega} = A^{-1}B, A^{-1}, C - B'A^{-1}B, n^*)$. However, A, B, C, and n^* will be modified by letting $V \to 0\ (kg \times kg)$, $a \to -kg$ and $\Psi \to 0(k \times k)$

Since Ω_0 has a matrix t distribution, any subset of k rows has a matrix t distribution. In addition, the conditional distribution of any subset of rows of the matrix Ω_0 given any other subset of rows is also a matrix t. Furthermore, one can test (marginally or conditionally) any subset of the matrix rows to equal zero using an exact F test statistic for $k = 1, 2$ and using an approximate χ^2 test statistic for $k \geq 3$. See Box and Tiao (1973) for more details on the equations and characteristics of the matrix t distribution.

Instead of manipulating the distribution of the orders p, q, P, and Q, it is suggested to consider the posterior distribution of Ω_0, given by the previous theorem, and employ a backward elimination procedure to specify initial orders p, q, P and Q for the model as follows:

1. Test H_0: $\Theta_r = 0$ versus $\Theta_r \neq 0$ using the marginal posterior matrix t distribution of Θ_r.
2. If H_0 in the above test is not rejected, test H_0: $\Theta_{r-1} = 0$ versus $\Theta_{r-1} \neq 0$. This test uses the conditional distribution of Θ_{r-1} given $\Theta_r = 0$. It has also a matrix t form.
3. If H_0 in test 2 is not rejected, test H_0: $\Theta_{r-2} = 0$ versus $\Theta_{r-2} \neq 0$. This test uses the conditional distribution of Θ_{r-2} given $\Theta_r = \Theta_{r-1} = 0$. This distribution also has a matrix t form.
4. The above three steps are repeated in the same manner until the hypothesis $H_0: \Theta_{r_0} = 0$ is rejected for a value r_0 where $0 < r_0 \leq r$. Then r_0 is considered an initial Bayesian estimate of the seasonal moving average order r.
5. The four previous steps are repeated for the seasonal autoregressive order v until the hypothesis Φ_{v_0} is rejected for a value v_0 where $0 < v_0 \leq v$. Then v_0 is considered an initial Bayesian estimate for the seasonal autoregressive order v. Then the values v_0 and r_0 are the proposed initial Bayesian estimates of the seasonal orders v and r.
6. The five previous steps are repeated for the non-seasonal (regular) orders l and m until the hypotheses $\theta_{m_0} = 0$ and $\phi_{l_0} = 0$ are rejected for some l_0 and m_0 where $0 < m_0 \leq m$ and $0 < l_0 \leq l$. Then the values l_0 and m_0 are the proposed initial Bayesian estimates of the non-seasonal orders l and m.

Very often the orders l, m, v, and r do not exceed two. Therefore, we restrict attention here to implementing such initial Bayesian technique, described above, to specify the initial orders of $SARMA(p,q)(P, Q)$ models assuming $l = m = v = r = 2$ using Jeffrey's vague prior. Figure 1 gives a binary decision tree for the seasonal part, which displays the 4 paths using which a seasonal order is chosen assuming the maximum seasonal orders $v = r = 2$

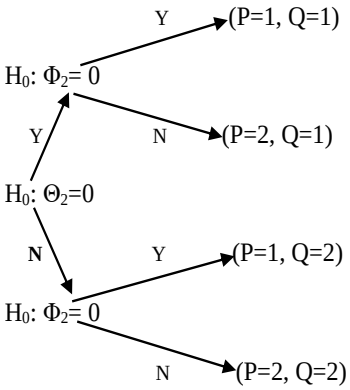

Figure 1. Decision Tree for Initial Specification of Seasonal Order (P, Q) with $v = r = 2$

Note: Y= yes, N= no

For example, one first tests $\Theta_2 = 0$, and if not rejected tests $\Phi_2 = 0$ and if not rejected one concludes that the appropriate order of the seasonal part is (1, 1). The hypothesis $\Theta_2 = 0$ is checked using the matrix t marginal distribution of Θ_2, but the hypothesis $\Phi_2 = 0$ is checked using the conditional matrix t distribution of Φ_2 given $\Theta_2 = 0$. The algorithm starts with the moving average coefficients, and then moves to autoregressive coefficients. The algorithm is arbitrary; to begin with the moving average part of the model instead of its autoregressive part.

In a similar fashion, one can specify the initial values of the non-seasonal order (p, q). The initial specified values of the non-seasonal and seasonal orders, say p_0, q_0, P_0 and Q_0, will be used in our suggested formal Bayesian methodology to derive an approximate posterior probability mass function for the orders p, q, P, and Q of the model in a simple and tractable form.

4. Formal Bayesian Identification

Based on initial values p_0, q_0, P_0 and Q_0, estimated above, the basic goal of this section is to introduce an approximate formal Bayesian approach for identifying the orders p, q, P and Q of the seasonal multidimensional autoregressive moving average models. Unlike the initial technique developed in the previous section, the orders p, q, P and Q are regarded as random variables and the task is to obtain their joint posterior probability mass function in a simple and tractable form. To do that, define S_n as the n vectors of observations generated form a seasonal multidimensional autoregressive moving average process with orders p, q, P and Q defined in equation (2.1) such that p, q, P and Q are positive unknown integers. To get the likelihood function of the parameters $\Omega(p,q,P,Q)$, p, q, P, Q and T, we Condition on the first (p+Ps) observed vectors, then

$$L(\Omega(p,q,P,Q), p,q,P,Q,T \mid S_n) \propto (2\pi)^{-k(n-p-Ps)/2} |T|^{(n-p-Ps)/2} \exp\left(-\frac{1}{2} tr \sum_{t=p+Ps+1}^{n} \varepsilon(t)\varepsilon'(t)T \right) \tag{4.1}$$

where $\Omega(p,q,P,Q) \in R^{k(p+q+P+Q+pP+qQ)\times k}$; $T > 0$; $p = 1, 2, ..., p^*$; $q = 1, 2, ..., q^*$; $P = 1, 2, ..., P^*$; $Q = 1, 2, ..., Q^*$ and (p^*, q^*, P^*, Q^*) are the largest possible orders of the model. From equation (2.2), the vectors $\varepsilon(t)$ of the residuals can be written as

$$\varepsilon'(t) = y'(t) - x'_{p,q,P,Q}(t-1)\Omega(p,q,P,Q), \quad t = 1,2,\ldots,n, \tag{4.2}$$

such that $x'_{p,q,P,Q}(t-1)$ is the row number t in the regressor's matrix $X(p,q,P,Q)$ having the form

$$\begin{aligned} x'_{p,q,P,Q}(t-1) = [\, & y'(t-1) \quad y'(t-2) \cdots y'(t-p) \quad -\varepsilon'(t-1) \quad -\varepsilon'(t-2) \cdots -\varepsilon'(t-q) \\ & y'(t-s) \quad y'(t-2s) \cdots y'(t-Ps) \quad -\varepsilon'(t-s) \quad -\varepsilon'(t-2s) \cdots -\varepsilon'(t-Qs) \\ & -y'(t-s-1) \quad -y'(t-2s-1) \cdots y'(t-Ps-1) \quad -y'(t-s-2) \quad -y'(t-2s-2) \\ & \cdots -y'(t-Ps-2) \cdots -y'(t-s-p) \quad y'(t-2s-p) \cdots y'(t-Ps-p) \\ & \varepsilon'(t-s-1) \quad \varepsilon'(t-2s-1) \cdots \varepsilon'(t-Qs-1) \quad \varepsilon'(t-s-2) \quad \varepsilon'(t-2s-2) \cdots \\ & \varepsilon'(t-Qs-2) \cdots \varepsilon'(t-s-q) \quad \varepsilon'(t-2s-m) \cdots \varepsilon'(t-Qs-q)\,] \end{aligned}$$

Substituting from (4.2) to equation (4.1), the conditional likelihood function is given by

$$L(\Omega(p,q,P,Q), p,q,P,Q,T \mid S_n) \propto (2\pi)^{-k(n-p-Ps)/2} |T|^{(n-p-Ps)/2}$$
$$\exp\left(-\tfrac{1}{2} tr \sum_{t=p+Ps+1}^{n} [y(t) - \Omega'(p,q,P,Q) x_{p,q,P,Q}(t-1)][y(t) - \Omega'(p,q,P,Q) x_{p,q,P,Q}(t-1)]' T \right) \tag{4.3}$$

The likelihood function (4.3) is a complicated function because the errors $\varepsilon(t-j)'s$ are non-linear function in the coefficients of the model φ_i , θ_i, Φ_i and Θ_i. To simplify the form (4.3), we suggest using the initial values p_0, q_0, P_0, and Q_0 to find estimates for the residuals $\varepsilon(t-j)'s$ using the following recurrence formula.

$$\hat{\varepsilon}(t) = y(t) - \hat{\Omega}\, \hat{x}_{p,q,P,Q}(t-1),$$

where $\hat{\Omega}$ the non-linear least square estimate of the matrix coefficients Ω and $\hat{x}_{p,q,P,Q}(t-1)$ is the same as $x_{p,q,P,Q}(t-1)$ replacing the exact residuals by the estimated ones $\hat{\varepsilon}(t-j)$'s . Then, the estimated residuals $\hat{\varepsilon}(t-j)$'s are substituted in the likelihood function (4.3) to get an approximate likelihood function on the form

$$L^*(\Omega(p,q,P,Q), p,q,P,Q,T \mid S_n) \propto (2\pi)^{-k(n-p-Ps)/2} |T|^{(n-p-Ps)/2}$$
$$\exp\left(-\tfrac{1}{2} tr \sum_{t=p+Ps+1}^{n} [y(t) - \Omega'(p,q,P,Q) \hat{x}_{p,q,P,Q}(t-1)][y(t) - \Omega'(p,q,P,Q) \hat{x}_{p,q,P,Q}(t-1)]' T \right) \tag{4.4}$$

An appropriate selection of the conditional prior density for $\Omega(p,q,P,Q)$ given p, q, P, Q and T is

$$\xi_1(\Omega(p,q,P,Q) \mid p,q,P,Q,\mathrm{T}) = (2\pi)^{-h(p,q,P,Q)k^2/2} \mid R(p,q,P,Q) \mid^{k/2} \mid \mathrm{T} \mid^{kh(p,q,P,Q)/2}$$
$$exp(-\tfrac{1}{2} tr[\,\Omega(p,q,P,Q) - D(p,q,P,Q)\,]' R(p,q,P,Q)[\Omega(p,q,P,Q) - D(p,q,P,Q)\,]\mathrm{T}), \tag{4.5}$$

where the hyper-parameters $D(p,q,P,Q) \in R^{h(p,q,P,Q)kxk}$ and $R(p,q,P,Q)$ is a positive definite matrix of orders $kh(p,q,P,Q) \times kh(p,q,P,Q)$. The precision matrix T is assumed to have, a Wishart prior distribution in the form.

$$\xi_2(\mathrm{T}) \propto |T|^{\frac{a-(k+1)}{2}} \exp(-\tfrac{1}{2} tr\, \Psi\, \mathrm{T}), \tag{4.6}$$

where Ψ is a positive definite matrix of orders $k \times k$.

Suppose β_{ijmr} is the prior probability that the four orders of the seasonal multivariate autoregressive moving average process which generated the time series realization S_n are i, j, m and r; i.e.

$$\beta_{ijmr} = P_r[\, p=i, q=j, P=m, Q=r\,] \;, i=1,2,\ldots,p^*;\; j=1,2,\ldots,q^*;\; m=1,2,\ldots,P^*;\; r=1,2,\ldots,Q^* \tag{4.7}$$

The four maximum orders of the process p^*, q^*, P^*, Q^* are assumed to be known. Regarding equations (4.5), (4.6) and (4.7), the joint prior distribution of the models' parameters $\Omega(p,q,P,Q)$, p, q, P, Q and T is

$$f(\Omega(p,q,P,Q), p,q,P,Q, \mathrm{T}) \propto \beta_{ijmr} (2\pi)^{-hk^2/2} |R(p,q,P,Q)|^{k/2} |\mathrm{T}|^{\frac{kh+a-k-1}{2}}$$
$$\exp(-\tfrac{1}{2} tr\{[\Omega(p,q,P,Q) - D(p,q,P,Q)]' R(p,q,P,Q)[\Omega(p,q,P,Q) - D(p,q,P,Q)] + \Psi\}\mathrm{T})$$
(4.8)

If one is not quite confident about the hyper-parameters $D(p, q, P,Q)$, $R(p, q, P,Q)$, and Ψ, one might use Jeffreys' vague prior

$$\xi(\Gamma(\mathrm{p},q,\ \mathrm{P},\ Q), \mathrm{p}, q,\ \mathrm{P}, Q,\ \mathrm{T}) \propto |\mathrm{T}|^{-(k+1)/2} \tag{4.9}$$

Combining the approximate likelihood (4.4), according to Bayes' theorem, with the prior distribution in (4.8), then, the joint posterior distribution of the model's parameters $\Omega(p,q,P,Q)$, p, q, P, Q and T will be

$$g(\Gamma(p,q,P,Q), p,q,P,Q, \mathrm{T} \mid S_n) \propto \beta_{ijmr} (2\pi)^{\frac{k(p+Ps)-hk^2}{2}} |R(p,q,P,Q)|^{k/2} |\mathrm{T}|^{\alpha(p,q,P,Q)/2}$$
$$\exp(-\tfrac{1}{2} tr\{[\Omega(p,q,P,Q) - D(p,q,P,Q)]' R(p,q,P,Q)[\Omega(p,q,P,Q) - D(p,q,P,Q)] + \Psi$$
$$+ \sum_{t=p+Ps+1}^{n} [y(t) - \Omega'(p,q,P,Q) x_{p,q,P,Q}(t-1)][y(t) - \Omega'(p,q,P,Q) \hat{x}_{p,q,P,Q}(t-1)]' T\}\mathrm{T}),$$
(4.10)

where $\alpha(p,q,P,Q) = n - p - Ps + kh(p,q,P,\ Q) + a - k - 1$

Theorem 4.1: based on the approximate likelihood function (4.4) and the joint prior density (4.8), the joint posterior probability mass function of the orders p, q, P and Q is

$$h(p,q,P,Q|S_n) \propto \beta_{ijmr}(\pi)^{k(p+Ps)/2} |R(p,q,P,Q)^{k/2}| A(p,q,P,Q)|^{-k/2}$$
$$|C(p,q,P,Q)|^{-\frac{1}{2}[n-p-Ps+a]} \prod_{j=1}^{k} \Gamma\left(\frac{n-p-Ps+a-k+j}{2}\right), n > k+p+Ps-a-1$$

$$A(p,q,P,Q) = R(p,q,P,Q) + \sum_{t=p+Ps+1}^{n} \hat{x}_{p,q,P,Q}(t-1)\hat{x}'_{p,q,P,Q}(t-1), \text{where}$$

$$B(p,P) = R(p,P)D(p,P) + \sum_{t=p+Ps+1}^{n} x_{p,P}(t-1)y'(t)$$

and

$$C(p,q,P,Q) = D'(p,q,P,Q)R(p,q,P,Q)D'(p,q,P,Q) + \Psi + \sum_{t=p+Ps+1}^{n} y(t)y'(t)$$
$$- B'(p,q,P,Q)A^{-1}(p,q,P,Q)B(p,q,P,Q)$$

Theorem (4.1) can be proved by integrating (4.10) with respect to the parameters Ω and T, respectively. The integral with respect to Ω is done by completing the squares of the exponent (4.10) and then applying the matrix normal integral, see Box and Tiao (1973). The integral concerning the parameter T is done using Wishart density.

Corollary (4.1): Using the approximate likelihood (4.4) and the Jeffreys' vague prior (4.9), the joint posterior probability mass function of the orders p, q, P and Q is

$$h_1(p,q,P,Q|S_n) \propto (\pi)^{hk^2/2} |A^*(p,q,P,Q)|^{-k/2} |C^*(p,q,P,Q)|^{-\frac{1}{2}[n-hk]}$$
$$\prod_{j=1}^{k} \Gamma\left(\frac{n-hk-k+j}{2}\right), n > k+hk-1,$$

where

$$A^*(p,q,P,Q) = \sum_{t=p+Ps+1}^{n} \hat{x}_{p,q,P,Q}(t-1)\hat{x}'_{p,q,P,Q}(t-1) \quad ,$$

$$B^*(p,q,P,Q) = \sum_{t=p+Ps+1}^{n} \hat{x}_{p,q,P,Q}(t-1)y'(t) \quad ,$$

$$C^*(p,q,P,Q) = \sum_{t=p+Ps+1}^{n} y(t)y'(t) - B^{*'}(p,q,P,Q)A^{*-1}(p,q,.P,Q)B^*(p,q,P,Q) \quad .$$

Note that the form of the joint posterior probability mass function of the model orders is easy to handle with the software packages. Then one may calculate and inspect all posterior probabilities over the range of the orders and select the values p, q, P and Q where the posterior probability mass function reaches its highest value to be the most appropriate orders for the analyzed multivariate time series data.

5. A Numerical Study

We devote this section to assessing the numerical effectiveness of the proposed Bayesian algorithm, via conducting two Monte Carlo simulation studies, in identifying the orders of the bivariate *SVARMA* (1,1) $(1,1)_4$ model. The effectiveness criterion of the proposed algorithm is taken to be the relative frequency of the replicates which succeeds in identifying the true order (1,1) (1,1). Our main goal is to test such effectiveness concerning different time series lengths, and different coefficients' matrices as well as different prior selections. Regarding the variance-covariance matrix of the disturbance term, abbreviated by Σ, and the maximum order of the simulated model, they are fixed at $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and (2,2)(2,2) respectively throughout the whole simulation process. All computations are performed on Pc using SCA package**.**

Consider simulation 1, as an example, it starts by generating 500 data sets of pairs of observations from bivariate normal random variables, each of size 3000, to represent the random errors $\varepsilon(t)$. These 500 data sets are then used to generate 500 realizations of pairs of observation, each of size 2000, from bivariate $SVARMA(1,1)(1,1)_4$ model with coefficient matrices $\phi_1 = \begin{pmatrix} 0.4 & 0.4 \\ 0.4 & 0.4 \end{pmatrix}$, $\Phi_1 = \begin{pmatrix} 0.4 & 0.4 \\ 0.4 & 0.4 \end{pmatrix}$, $\theta_1 = \begin{pmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{pmatrix}$, and $\Theta_1 = \begin{pmatrix} 0.4 & -0.4 \\ -0.3 & 0.2 \end{pmatrix}$. The first 1000 generated observations are ignored to remove the initialization effect. For each prior function, the second step of simulation 1 is to employ all computations, assuming the maximum values of the orders are (2,2)(2,2), required to specify a model for each of the 500 realizations and to find the percentage of the correct specification. The computations are employed for a certain time series length *n* using the first *n* observations of each generated realization. The second step is repeated for each selected time series length and each prior combination. The selected lengths of the time series are taken to be 200, 400, 600, 800, 1500, and 2000. Regarding the prior probability mass function of the orders *p*, *q*, *P* and *Q*, which is combined with Jeffrey's vague prior of $\Omega(p,q,P,Q)$ and *T*, the following three prior distributions are employed:

$$\text{Prior } 1: \xi(p,q,P,Q) = \frac{1}{16} \;;\; p = 1,2 \;;\; q = 1,2 \;;\; P = 1,2 \;;\; Q = 1,2$$

$$\text{Prior } 2: \xi(p,q,P,Q) \propto (0.5)^{p+q+P+Q} \;;\; p = 1,2 \;;\; q = 1,2 \;;\; P = 1,2 \;;\; Q = 1,2$$

$$\text{Prior } 3: \xi(1,1,1,1) = 0.0928,\ \xi(1,2,1,1) = \xi(2,1,1,1) = \xi(1,1,1,2) = \xi(1,1,2,1)$$
$$= 0.07765,\ \xi(1,2,1,2) = \xi(1,2,2,1) = \xi(2,1,1,2) = \xi(2,1,2,1) = \xi(1,1.2,2)$$
$$= \xi(2,2,1,1) = 0.0625,\quad \xi(1,2,2,2) = \xi(2,1,2,2) = \xi(2,2,1,2) = \xi(2,2,2,1)$$
$$= 0.04735,\ \xi(2,2,2,2) = 0.0322$$

The first prior function gives equal probabilities to each combination of orders. The second prior function assigns probabilities that decrease exponentially as the orders increase, while the third prior is selected to give probabilities that decrease by an absolute value 0.01515 when the order

increases. Simulation 2 is conducted similarly with $\phi_1 = \begin{pmatrix} 0.6 & 0.4 \\ -0.5 & 1.11 \end{pmatrix}$, $\Phi_1 = \begin{pmatrix} 0.6 & 0.4 \\ -0.5 & 1.11 \end{pmatrix}$, $\theta_1 = \begin{pmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{pmatrix}$, and $\Theta_1 = \begin{pmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{pmatrix}$.

Table 1 presents the results of simulations 1 and 2. The coefficients in these simulations were selected to satisfy the stationarity and invertibility conditions of the models, see Harvey (1981). Furthermore, percentages of correct specifications using the well-known AIC are also reported in the same table.

Table 1. Percentages of Correct Identification of the AIC and the Bayesian Techniques for bivariate SVARMA(1,1)(1,1)$_4$ Processes

PARAMETERS	*n*	PRIOR1	PRIOR2	PRIOR3	AIC
$\phi_1 = \begin{bmatrix} 0.4 & 0.4 \\ 0.4 & 0.4 \end{bmatrix}$, $\Phi_1 = \begin{bmatrix} 0.4 & 0.4 \\ 0.4 & 0.4 \end{bmatrix}$	**200**	1.6	73.6	53.4	57.0
	400	12.6	85.0	66.6	63.2
	600	26.6	91.0	77.0	64.2
$\theta_1 = \begin{bmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{bmatrix}$, $\Theta_1 = \begin{bmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{bmatrix}$	**800**	38.0	92.6	81.4	65.8
	1500	55.0	95.4	88.2	63.6
	2000	65.0	97.0	90.4	64.0
$\phi_1 = \begin{bmatrix} 0.6 & 0.4 \\ -0.5 & 1.11 \end{bmatrix}$, $\Phi_1 = \begin{bmatrix} 0.6 & 0.4 \\ -0.5 & 1.11 \end{bmatrix}$	**200**	32.8	81.2	64.8	57.8
	400	47.8	89.4	75.6	62.2
	600	50.4	92.2	79.0	63.4
$\theta_1 = \begin{bmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{bmatrix}$, $\Theta_1 = \begin{bmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{bmatrix}$	**800**	52.0	95.0	82.0	62.8
	1500	48.4	96.8	85.2	66.6
	2000	53.2	96.6	86.8	65.8

Inspection of the numerical results shows that the percentage of correct identification increases as the time series length *n* increases for the two models and all priors. For the uniform prior (prior 1), the percentages of correct identification are reasonable, being greater than 48 %, for time series length 1500 or longer regardless of the values of the coefficients. Moreover, the percentage of correct identification achieved by the second and third priors is very high, being greater than 81%, for a sample size of 800 or more for the two models. Moreover, the percentages of correct specification obtained by the technique using the second prior is higher than that obtained using prior 3, which in turn is better than that obtained using prior 1. For the matter of comparison, the percentages of correct identification obtained using the AIC procedure are greater than those obtained by using prior 1 for a short time series length; however, the results get closer to each other when the sample size increases. For the other two priors and a sample size greater than 200, the percentages of correct specification achieved by the proposed procedure are much higher than that achieved by the AIC no matter what the coefficients are. Considering the previous comments, we

might conclude that the simulation results support the goodness of using the proposed Bayesian approach to identify orders for seasonal vector autoregressive moving average processes for moderate as well as long time series.

Declaration of interests:
The authors declare that they have no conflict of interest.

References

Akaike, H. (1974). A New Look at Statistical Model Identification. *IEEE Transaction on Automatic Control*, 19: 3716–3723.

Athanasopoulos, G. and Vahid, F. (2008). VARMA versus VAR for macroeconomic forecasting. *Journal of Business and Economic Statistics,* 26(2): 237–252.

Box, G.E.P., Jenkins, G.M, Reinsel, G.C. and Ljung, G.M. (2016).Time Series Analysis: Forecasting and Control. Fifth Edition, John Wiley & Sons.

Beveridge, S. and Oickle, C. (1994). A Comparison of Box–Jenkins and objective methods for determining the order of a non-seasonal ARMA model. *Journal of Forecasting,* 13: 419–434. doi.org/10.1002/for.3980130502

Box, G. E. P., Tiao, G. C. (1973). Bayesian Inference in Statistical Analysis. Reading, MA: Addison Wesley.

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.

Brockwell**,** P. J. and Davis**,** R. A. (2016). Introduction to Time Series and Forecasting. Third edition. Springer International Publishing.

Chatfield, C. (2019). The Analysis of Time Series: Theory and Practice. An Introduction with R. Seventh Edition. Chapman & Hall.

Cooley, T. F. and Dwyer, M. (1998). Business cycle analysis without much theory: a look at structural VARs. *Journal of Econometrics*, 83: 57–88. doi.org/10.1016/S0304-4076(97)00065-1

Hagnell, M. (1991). A Multivariate time series analysis of fertility, adult mortality, nuptiality and real wages in Sweden 1751 – 1850: A comparison of two different approaches. *Journal of Official Statistics,* 7(4): 437-455.

Hannan, E. J. and B. G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, Ser. B, 41: 190-195.

Harvey. A., C. (1993). Time Series Models. Second Edition. The MIT Press.

Kascha, C. (2012). A comparison of estimation methods for vector autoregressive moving-average models. *Econometric Reviews*, 31: 297–324. doi: 10.1080/07474938.2011.607343

Liu, L. M. (2009). Time Series Analysis and Forecasting. Second Edition, Villa Park, IL: Scientific Computing Associates.

Lütkepohl. (2007). New Introduction to Multiple Time Series Analysis. Springer-Verlag Berlin Heidelberg Springer Publishing Company, Incorporated.

Monahan, J. F. (1983). Fully Bayesian analysis of ARIMA Time Series Models. *Journal of Econometrics,* 21: 307–331. doi.org/10.1016/0304-4076(83)90048-9

Priestley, M. (1981). Spectral Analysis of Time Series. London: Academic Press.

Quenouille (1957). The analysis of multiple time series. *Journal of Mathematical Sciences and Applications*, 5 (1): 1-16.

Raghavan, M. Athanasopoulos, G. and Silvapulle, P. (2009). VARMA models for Malaysian Monetary Policy Analysis. Monash Econometrics and Business Statistics Working Papers.

Reinsel, G. C. (1997). Elements of Multivariate Time Series. Springer Series in Statistics, Second Edition. Springer Verlag, New York.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14: 465-471.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6: 461-464.

Shaarawy, S.M. and Ali, S.S. (2003). Bayesian identification of seasonal autoregressive models. *Communications in Statistics-Theory and Methods*, 32: 1067–1084.

Shaarawy, S.M., Soliman, E.E.A. and Ali, S.S. (2007). Bayesian identification of moving average models. *Communications in Statistics-Theory and Method*, 36(12), 2301-2312.

Shaarawy, S.M. and Ali, S.S. (2008). Bayesian identification of vector autoregressive processes. *Communications in Statistics-Theory and Methods*, 37: 791–802.

Shaarawy, S. M, and Ali, S. S. (2015). Bayesian identification of seasonal vector autoregressive. *Communications in Statistics-Theory and Methods*, 44: 823-838.

Shaarawy, S.M., (2021). Bayesian modeling and forecasting of vector autoregressive moving average processes. *Communications in Statistics-Theory and Methods.* doi.org/10.1080/03610926.

Tiao, G. C., Box, G. E. P. (1981). Modeling multiple time series with applications. *Journal of the American Statistical Association*, 76: 802–816. doi.org/10.2307/2287575

Tiao, G. C. and R. S. Tsay (1983). Multiple time series modeling and extended sample cross-correlation, *Journal of Business Economics and Statistics*, 1: 43-56. doi.org/10.2307/1391772

Tsay, R. (1984). Order selection in non-stationary autoregressive models. *Annals of the Institute of Statistical Mathematics*, 12: 1425-1433.

Tsay, R. S. (1989). Parsimonious parameterization of vector autoregressive moving average models. *Journal of Business Economics and Statistics*, 7: 327-341. doi.org/10.2307/1391530

Tsay, R. S. (2014). Multivariate Time Series Analysis: With R and Financial Applications. John Wiley and Sons.

Wei, W. W. S. (2005). Time Series Analysis: Univariate and Multivariate Methods. Second edition, Reading, MA: Addison Wesley.

Wei. W. W. S. (2019). Multivariate Time Series Analysis and Applications. John Wiley & Sons.