

Seasonal autoregressive integrated moving average vector models and their application to quarterly rainfall series

Anthony E. Usoro¹

Abstract

In this study, n-dimensional Seasonal autoregressive integrated moving average vector (SARIMAV) models of additive form are compared with univariate models. Ordinary least squares method is adopted to estimate parameters of the models. The bivariate models obtained are reliable as much as univariate seasonal models. It is established that seasonal time series model is not only applicable in univariate case. Hence, seasonal autoregressive integrated moving average vector models are established, verified valid and useful in modeling multiple seasonal series.

Keywords: Univariate; Bivariate; Seasonal; Autoregressive and Moving average

¹ Department of Mathematics and Statistics, Akwa Ibom State University, Mkpato Enin, Akwa Ibom State. Email: toskila2@yahoo.com

1 Introduction

Multivariate time series analysis is an analysis of multi-variables, with a response vector as a dependent and predictor vectors as independent variables Wei (1990). In multivariate time series analysis, both the response and predictor vectors are modeled with corresponding lagged parameters of the response and predictor vectors. What distinguish multivariate time series from multivariate regression is that, only the response variable is modeled in regression, while all the vectors are modeled in multivariate time series. Dufour (2006) defined an m -dimensional vector process $(X_t: t \in \mathbb{Z})$ as an Autoregressive (p) model or Vector Autoregressive (VAR) of order ' p ' if it satisfies an equation of the form

$$X_t = \mu + \sum \Phi_k X_{t-k} + a_t, \text{ for very } t, \text{ where } \Phi_1, \Phi_2, \dots, \Phi_p \text{ are } m \times m \text{ fixed matrices and } (a_t: t \in \mathbb{Z}) \text{ is a white noise process.}$$

Harrison et al (2003) defined multivariate autoregressive analysis as an analysis of multiple time series whose vector of current values of all variables is modeled as a linear sum of pervious activities. Sims (1996) carried out multivariate time series analysis of consumption and gross national products of United State of America. He considered both maximum likelihood and ordinary least squares approaches in modeling the joint behavior of consumption and gross national products. The models had two lags of each variable and a constant. For the ordinary least squares, the estimated models were,

$$C = (0.9542L + 0.0456L^2)C + (0.1427L - 0.1432L^2)Y + 0.1258$$

$$Y = (0.2991L - 0.2300L^2)C + (1.2139L - 0.2908L^2)Y + 0.9799.$$

For the maximum likelihood method, the estimated models were

$$C = (0.9443L + 0.0364L^2)C + (0.1571L - 0.1409L^2)Y + 0.1385$$

$$Y = (0.2877L - 0.2228L^2)C + (1.2045L - 0.2773L^2)Y + 0.9635$$

The coefficients of maximum likelihood method were closed to that of the ordinary least squares method. Usoro and Omekara (2007) carried out multivariate time series analysis, using three vectors; a response vector and two predictor

vectors. Their analysis involved revenue series from a Local Government Council. In the analysis, every vector whether response or predictor was modeled. Good estimates were obtained from the models. The multivariate models were extended to bilinear time series (Usoro and Omekara, 2008).

In this paper, multivariate seasonal models are applied to quarterly rainfall data, with parameter estimation of bivariate models. In this study, major interest is focused “Seasonal autoregressive integrated moving average vector (SARIMAV) models”. The models are applied to quarterly series.

2 Method of analysis

2.1 Data Collection and Variables Description

The data for the analysis are quarterly rainfall data from CBN statistical bulletin (2010). The data are quarterly data collected for the period 1981-2010.

2.2 Univariate seasonal models

Many time series display seasonality. By seasonality, we mean periodic fluctuations. This can be defined as a pattern of a time series, which repeats at regular intervals every year Oguz and Beyza (2002). Some time series of retail sales will typically show increasing sales periodically every calendar year. Apart from the sharp escalation in most retail series which occurs around December in response to the Christmas period, there is seasonality effect in the water consumption, prices of agricultural products and their supplies, rainfall and many others. A univariate seasonal model for both autoregressive and moving average process is given by $\Phi(B)\Phi(B^s)(1-B)^d(1-B)^D X_t = \Theta(B)\Theta(B^s)E_t$, Box and Jenkins(1976). A pure seasonal model is given by $\Phi(B^s)(1-B^s)^D X_t = \Theta(B^s)E_t$.

2.3 Matrix Presentation of Vector Models

The general non-multiplicative SARIMAV models are presented as:

$$\begin{pmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \\ \vdots \\ X_{nt} \end{pmatrix} = \begin{pmatrix} \Phi_{4.11} & \Phi_{4.12} & \Phi_{4.13} & \cdots & \Phi_{4.1c} \\ \Phi_{4.21} & \Phi_{4.22} & \Phi_{4.23} & \cdots & \Phi_{4.2c} \\ \Phi_{4.31} & \Phi_{4.32} & \Phi_{4.33} & \cdots & \Phi_{4.3c} \\ \vdots & \vdots & \vdots & & \vdots \\ \Phi_{4.r1} & \Phi_{4.r2} & \Phi_{4.r3} & \cdots & \Phi_{4.rc} \end{pmatrix} \begin{pmatrix} X_{1t-4} \\ X_{2t-4} \\ X_{3t-4} \\ \vdots \\ X_{nt-4} \end{pmatrix} \\
 + \begin{pmatrix} \Phi_{8.11} & \Phi_{8.12} & \Phi_{8.13} & \cdots & \Phi_{8.1c} \\ \Phi_{8.21} & \Phi_{8.22} & \Phi_{8.23} & \cdots & \Phi_{8.2c} \\ \Phi_{8.31} & \Phi_{8.32} & \Phi_{8.33} & \cdots & \Phi_{8.3c} \\ \vdots & \vdots & \vdots & & \vdots \\ \Phi_{8.r1} & \Phi_{8.r2} & \Phi_{8.r3} & \cdots & \Phi_{8.rc} \end{pmatrix} \begin{pmatrix} X_{1t-8} \\ X_{2t-8} \\ X_{3t-8} \\ \vdots \\ X_{nt-8} \end{pmatrix} \\
 + \begin{pmatrix} \Phi_{12.11} & \Phi_{12.12} & \Phi_{12.13} & \cdots & \Phi_{12.1c} \\ \Phi_{12.21} & \Phi_{12.22} & \Phi_{12.23} & \cdots & \Phi_{12.2c} \\ \Phi_{12.31} & \Phi_{12.32} & \Phi_{12.33} & \cdots & \Phi_{12.3c} \\ \vdots & \vdots & \vdots & & \vdots \\ \Phi_{12.r1} & \Phi_{12.r2} & \Phi_{12.r3} & \cdots & \Phi_{12.rc} \end{pmatrix} \begin{pmatrix} X_{1t-12} \\ X_{2t-12} \\ X_{3t-12} \\ \vdots \\ X_{nt-12} \end{pmatrix} + \cdots \\
 + \begin{pmatrix} \Phi_{p.11} & \Phi_{p.12} & \Phi_{p.13} & \cdots & \Phi_{p.1c} \\ \Phi_{p.21} & \Phi_{p.22} & \Phi_{p.23} & \cdots & \Phi_{p.2c} \\ \Phi_{p.31} & \Phi_{p.32} & \Phi_{p.33} & \cdots & \Phi_{p.3c} \\ \vdots & \vdots & \vdots & & \vdots \\ \Phi_{p.r1} & \Phi_{p.r2} & \Phi_{p.r3} & \cdots & \Phi_{p.rc} \end{pmatrix} \begin{pmatrix} X_{1t-p} \\ X_{2t-p} \\ X_{3t-p} \\ \vdots \\ X_{nt-p} \end{pmatrix}$$

$$\begin{pmatrix} E_{1t} \\ E_{2t} \\ E_{3t} \\ \vdots \\ E_{mt} \end{pmatrix} = \begin{pmatrix} \Theta_{4.11} & \Theta_{4.12} & \Theta_{4.13} & \cdots & \Theta_{4.1d} \\ \Theta_{4.21} & \Theta_{4.22} & \Theta_{4.23} & \cdots & \Theta_{4.2d} \\ \Theta_{4.31} & \Theta_{4.32} & \Theta_{4.33} & \cdots & \Theta_{4.3d} \\ \vdots & \vdots & \vdots & & \vdots \\ \Theta_{4.s1} & \Theta_{4.s2} & \Theta_{4.s3} & \cdots & \Theta_{4.sd} \end{pmatrix} \begin{pmatrix} E_{1t-4} \\ E_{2t-4} \\ E_{3t-4} \\ \vdots \\ E_{mt-4} \end{pmatrix} \\
+ \begin{pmatrix} \Theta_{8.11} & \Theta_{8.12} & \Theta_{8.13} & \cdots & \Theta_{8.1d} \\ \Theta_{8.21} & \Theta_{8.22} & \Theta_{8.23} & \cdots & \Theta_{8.2d} \\ \Theta_{8.31} & \Theta_{8.32} & \Theta_{8.33} & \cdots & \Theta_{8.3d} \\ \vdots & \vdots & \vdots & & \vdots \\ \Theta_{8.s1} & \Theta_{8.s2} & \Theta_{8.s3} & \cdots & \Theta_{8.sd} \end{pmatrix} \begin{pmatrix} E_{1t-8} \\ E_{2t-8} \\ E_{3t-8} \\ \vdots \\ E_{mt-8} \end{pmatrix} \\
+ \begin{pmatrix} \Theta_{12.11} & \Theta_{12.12} & \Theta_{12.13} & \cdots & \Theta_{12.1d} \\ \Theta_{12.21} & \Theta_{12.22} & \Theta_{12.23} & \cdots & \Theta_{12.2d} \\ \Theta_{12.31} & \Theta_{12.32} & \Theta_{12.33} & \cdots & \Theta_{12.3d} \\ \vdots & \vdots & \vdots & & \vdots \\ \Theta_{12.s1} & \Theta_{12.s2} & \Theta_{12.s3} & \cdots & \Theta_{12.sd} \end{pmatrix} \begin{pmatrix} E_{1t-12} \\ E_{2t-12} \\ E_{3t-12} \\ \vdots \\ E_{mt-12} \end{pmatrix} + \cdots \\
+ \begin{pmatrix} \Theta_{q.11} & \Theta_{q.12} & \Theta_{q.13} & \cdots & \Theta_{q.1d} \\ \Theta_{q.21} & \Theta_{q.22} & \Theta_{q.23} & \cdots & \Theta_{q.2d} \\ \Theta_{q.31} & \Theta_{q.32} & \Theta_{q.33} & \cdots & \Theta_{q.3d} \\ \vdots & \vdots & \vdots & & \vdots \\ \Theta_{q.s1} & \Theta_{q.s2} & \Theta_{q.s3} & \cdots & \Theta_{q.sd} \end{pmatrix} \begin{pmatrix} E_{1t-q} \\ E_{2t-q} \\ E_{3t-q} \\ \vdots \\ E_{mt-q} \end{pmatrix}$$

The above SARIMAV models are reduced to the form;

$$X_{it} = \sum_{a=4}^p \sum_{i=1}^n \sum_{k=1}^r \sum_{f=1}^c \Phi_{a.k.f} X_{it-a} + E_{jt} + \sum_{b=4}^q \sum_{j=1}^m \sum_{u=1}^s \sum_{v=1}^d \Theta_{b.u.v} E_{jt-b}$$

where $\Phi_{a.k.f}$ and $\Theta_{b.u.v}$ are matrices of the coefficients of seasonal vector autoregressive and moving average models respectively.

3 Estimation of parameters and analysis

In this section, estimates of the parameters of both the univariate and multivariate are considered. The rainfall data used included quarterly data for

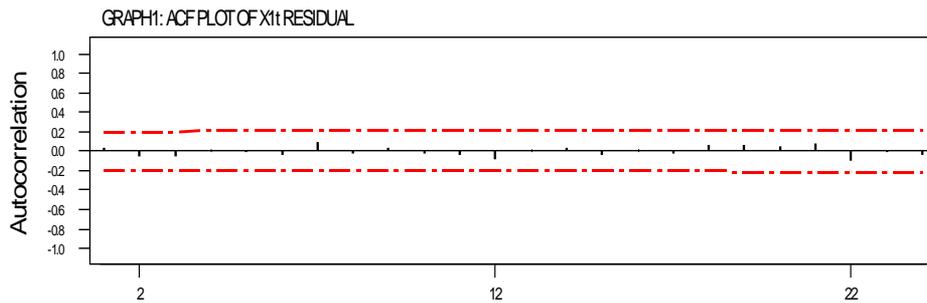
Akwa Ibom State and Borno State; represented by X_{1t} and X_{2t} , respectively. In addition to estimation of the parameters of the univariate seasonal and SARIMAV models, the choice of the rainfall data for the two extreme states (north and south) was motivated by the need to compare the seasonal patterns and the effects of the quarterly rainfall data in the two zones. It also includes performance of the univariate and multivariate models.

3.1 Estimattion of the parameters of univaraiate seasonal models

a. Model for X_{1t} : The distribution of autocorrelation and partial autocorrelation functions of X_{1t} seasonally differenced series suggested pure SARIMA (3,1,1). Therefore, the model with estimated parameters is obtained as:

$$X_{1t} = -0.3842X_{1t-4} - 0.3674X_{1t-8} - 0.2007X_{1t-12} - 0.3158E_{1t-4}.$$

Figure 1 is the autocorrelation function of the residual values of X_{1t} .



Lag	Corr	T	LBQ												
1	0.02	0.23	0.05	8	-0.03	-0.26	1.65	15	-0.05	-0.48	3.30	22	-0.10	-0.98	6.58
2	-0.05	-0.52	0.34	9	0.03	0.30	1.75	16	0.00	0.02	3.30	23	-0.01	-0.10	6.60
3	-0.06	-0.56	0.66	10	-0.04	-0.35	1.90	17	-0.03	-0.29	3.41	24	-0.05	-0.43	6.88
4	0.00	0.02	0.67	11	-0.04	-0.39	2.08	18	0.06	0.60	3.90				
5	-0.01	-0.11	0.68	12	-0.08	-0.80	2.86	19	0.06	0.56	4.33				
6	-0.04	-0.39	0.85	13	-0.01	-0.07	2.87	20	0.04	0.42	4.57				
7	0.08	0.80	1.57	14	0.03	0.32	3.00	21	0.07	0.66	5.19				

Figure 1

b. Model for X_{2t} : The distribution of autocorrelation and partial autocorrelation functions of X_{2t} seasonally differenced series suggested pure SARIMA (5,1,1).

Therefore, the model with estimated parameters is obtained as:

$$X_{2t} = -0.843X_{2t-4} - 0.695X_{2t-8} - 0.500X_{2t-12} - 0.147X_{2t-16} - 0.325X_{2t-20} - 0.107E_{2t-4}.$$

Figure 2 is the autocorrelation function of the residual values of X_{2t} .

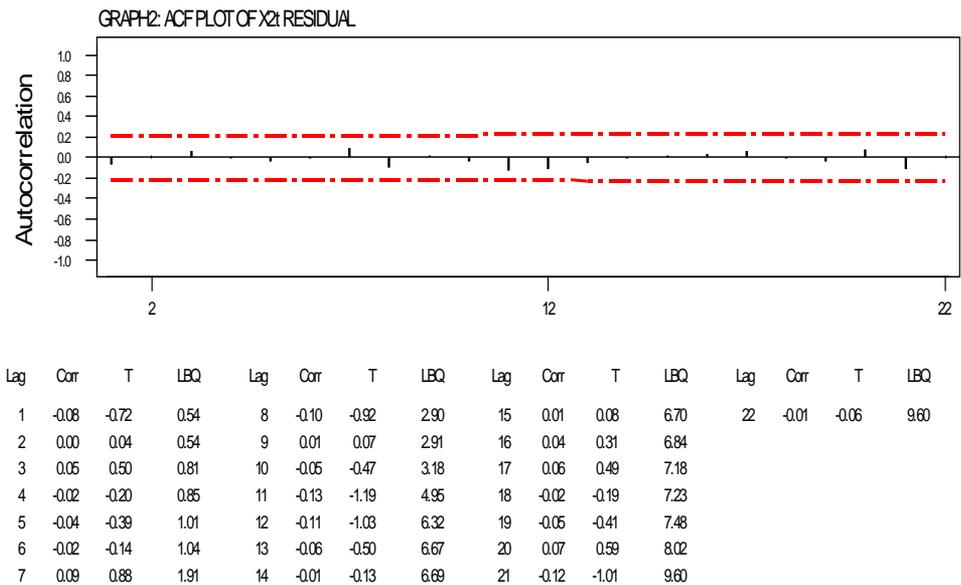


Figure 2

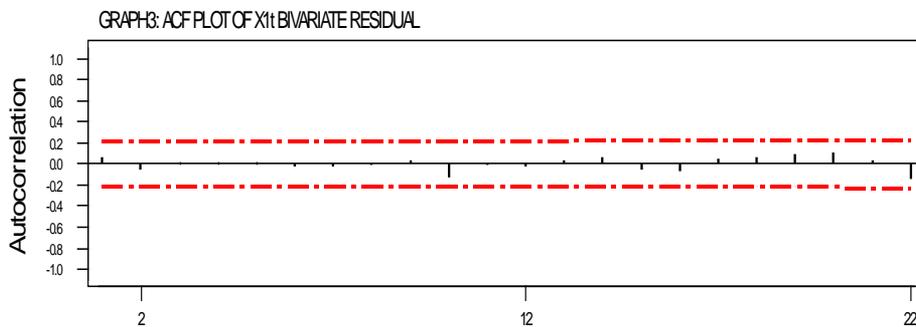
3.3 Estimation of the parameters of multivariate seasonal models

The use of the two vectors X_{1t} and X_{2t} introduced us to bivariate, as a special case of the SARIMAV models. In univariate time series, a process say, X_t is modeled as a function of its past values. Given two vectors X_{1t} and X_{2t} in bivariate time series, each vector is modeled as function of the distributed lags of

the two vectors. This expresses linear relationship of each vector with the past values of the two vectors in the bivariate models. Least squares estimates of the coefficients are provided in the following matrices:

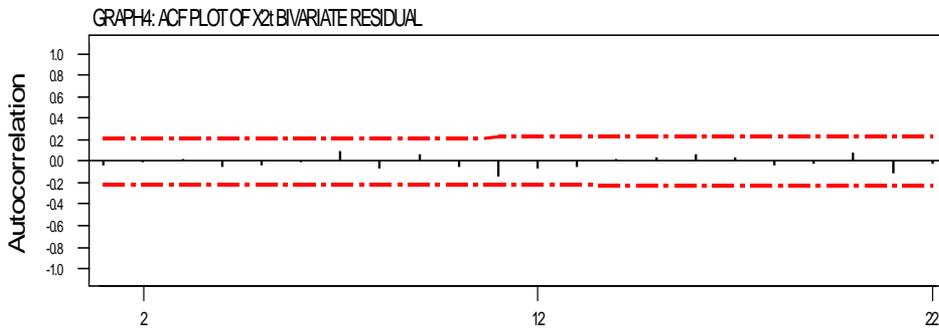
$$\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} -0.095 & -0.124 \\ 0.112 & -0.838 \end{pmatrix} \begin{pmatrix} X_{1t-4} \\ X_{2t-4} \end{pmatrix} + \begin{pmatrix} -0.097 & -0.379 \\ -0.0303 & -0.641 \end{pmatrix} \begin{pmatrix} X_{1t-8} \\ X_{2t-8} \end{pmatrix} \\ + \begin{pmatrix} -0.027 & -0.513 \\ -0.0118 & -0.450 \end{pmatrix} \begin{pmatrix} X_{1t-12} \\ X_{2t-12} \end{pmatrix} + \begin{pmatrix} 0 & 0.255 \\ 0 & -0.166 \end{pmatrix} \begin{pmatrix} X_{1t-16} \\ X_{2t-16} \end{pmatrix} \\ + \begin{pmatrix} 0 & -0.089 \\ 0 & -0.383 \end{pmatrix} \begin{pmatrix} X_{1t-20} \\ X_{2t-20} \end{pmatrix} + \begin{pmatrix} -0.518 & 0.237 \\ -0.122 & 0.117 \end{pmatrix} \begin{pmatrix} E_{1t-4} \\ E_{2t-4} \end{pmatrix}$$

Expansion of the above matrices provides bivariate seasonal autoregressive moving average models for the two vector series X_{1t} and X_{2t} . Figures 3 and 4 are the ACF of X_{1t} and X_{2t} residual series.



Lag	Corr	T	LBQ												
1	0.07	0.61	0.39	8	-0.01	-0.10	0.89	15	-0.07	-0.59	4.02	22	-0.15	-1.29	9.93
2	-0.06	-0.54	0.70	9	0.03	0.30	1.00	16	-0.07	-0.65	4.58				
3	0.00	0.00	0.70	10	-0.14	-1.29	2.95	17	0.04	0.40	4.80				
4	-0.00	-0.02	0.70	11	-0.02	-0.16	2.98	18	0.06	0.52	5.19				
5	0.00	0.04	0.70	12	-0.03	-0.23	3.05	19	0.09	0.77	6.03				
6	-0.04	-0.33	0.83	13	0.03	0.23	3.12	20	0.10	0.91	7.26				
7	-0.02	-0.22	0.88	14	0.06	0.58	3.56	21	0.02	0.21	7.33				

Figure 3



Lag	Corr	T	LBQ												
1	-0.05	-0.49	0.24	8	-0.08	-0.74	2.43	15	0.03	0.25	6.24	22	-0.03	-0.30	9.46
2	-0.01	-0.12	0.26	9	0.06	0.57	2.81	16	0.05	0.49	6.57				
3	0.02	0.20	0.30	10	-0.06	-0.51	3.13	17	0.04	0.32	6.71				
4	-0.06	-0.54	0.62	11	-0.14	-1.29	5.19	18	-0.05	-0.42	6.97				
5	-0.05	-0.48	0.88	12	-0.08	-0.70	5.83	19	-0.03	-0.29	7.10				
6	-0.01	-0.10	0.89	13	-0.05	-0.48	6.14	20	0.07	0.60	7.65				
7	0.10	0.89	1.78	14	0.01	0.09	6.15	21	-0.12	-1.04	9.31				

Figure 4

4 Conclusion

Most of the rainfall models are univariate in terms of location, because correlation of other locations is not considered. This is due to model complexity as a result of the distribution pattern of seasonal series. What distinguishes ARIMAV models from SARIMAV models is the seasonal component in the later models. The generalized models presented in matrix form take care of the multivariate seasonal series, irrespective of the number of vectors. Reliability of the SARIMAV models is established in time series analysis.

References

- [1] G.P. Box and G.M. Jenkins, *Time series analysis, forecasting and control*, Holden-Day, San Francisco, 1976.
- [2] Jean-Marie Dufour, *Multivariate Time Series Modelling*, Cambridge University press, Cambridge, U.K, 2006.
- [3] L. Harrison, W. D. Penny and K. J. Friston, Application of multivariate time series in the functional network in the brain region, *Neuroimage*, **19**(4), (2003), 1477-1491.
- [4] A. Oguz and P.U. Beyza, Seasonal Adjustment Method: An application to the Turkish Monetary Aggregate, Department of Statistics, Central Bank of the Republic of Turkey, (2002).
- [5] C.A. Sims, Multivariate Time Series Modeling of Gross National Products of United State of America, American Statistical Association Meetings, (1996).
- [6] W.S. Wei, *Time Series Analysis; Univariate and Multivariate Methods*, California, Addison-Wesley, 1990.
- [7] A.E. Usoro and C.O. Omekara, Estimation of Pure Autoregressive Models for Revenue Series, *Global Journal of Mathematical Sciences*, **6**(1), (2007), 31-37.
- [8] A.E. Usoro and C.O. Omekara, Bilinear Autoregressive Vector Models and their Application to Estimation of Revenue Series, *Asian Journal of Mathematics and Statistics*, Asian Network for Scientific Information, **1**(1), (2008), 50-56.