Forecasting wholesale price of coarse rice in Bangladesh: A seasonal autoregressive integrated moving average approach

M. F. Hassan[*], M. A. Islam, M. F. Imam, and S. M. Sayem
[1]Department of Agricultural Statistics, Bangladesh Agricultural University, Mymensingh, Bangladesh, *E-mail: fuad_bau016@yahoo.com

Abstract

This article attempts to develop the model and to forecast the wholesale price of coarse rice in Bangladesh. Seasonal Autoregressive Integrated Moving Average (SARIMA) models have been developed on the monthly data collected from July 1975 to December 2011and validated using the data from December 2010 to December 2011. The results showed that the predicted values were consistent with the upturns and downturns of the observed series. The model with non seasonal autoregressive 1, difference 1 and moving average 1 and seasonal difference 1 and moving average 1 that is SARIMA $(1,1,1)(0,1,1)_{12}$ model has been found as the most suitable model with least Root Mean Square Error (RMSE) of 61.657, Normalised Bayesian Information Criteria (BIC) of 8.300 and Mean Absolute Percent Error (MAPE) of 3.906. The model was further validated by Ljung-Box test (Q18=17.394 and p>.20) with no significant autocorrelation between residuals at different lag times. Finally, a forecast for the period January 2012 to December 2013 was made.

Keywords: Coarse Rice, Forecasting, Seasonal Autoregressive Integrated Moving Average

Introduction

Rice is one of the staple foods in Bangladesh. It alone constitutes the lion share (96%) of total food grain produced in Bangladesh (BBS 2010). It is also the most important crop to millions of small farmers who grow it on millions of hectares throughout the region, and to the many landless workers who derive income from working on these farms. Most of the people of our country consume coarse rice. Aus, Aman and Boro in all the three seasons various varieties of coarse rice are cultivated and produced, and government through various operation release mainly coarse rice. Any fluctuation in the price of coarse rice always makes government in anxiety as the fluctuation in coarse rice price has great impact on the millions of its producer and consumers (Brennan 1995).

Forecasting has been very important in decision making at all levels and sectors of the economy. Price forecasts are critical to market participant making production and marketing decisions and to policy makers who administer commodity programs and assess the market impacts of domestic or international events. In agriculture, where the decision environment is characterized by risks and uncertainty largely due to uncertain yields and relatively low price elasticity of demand of the most commodities, decision makers require some information about the future and the likelihood of the possible future outcomes. An early forecasting about the probable price of coarse rice will help the policy makers regarding the probable fluctuation in its price.

Policy makers can get prior indication about possible future prices through price forecasting using time series analysis. Autoregressive integrated moving average (ARIMA) is one of the popular forecasting model. The intrinsic nature of a time series is that successive observations are dependent or correlated and assists in forecasting future values (Abraham *et al.*1983) Results from time series analysis can give valuable information when formulating future policies.

In econometric research, Autoregressive Integrated Moving Average (ARIMA) models as well as Seasonal Autoregressive Integrated Moving Average (SARIMA) models are useful tools for analysing time series data containing ordinary or seasonal trends to develop a predictive forecasting model (Makridakis *et al.* 1998). There have been efforts in forecasting rice prices in different parts of the world using both ARIMA and SARIMA modelling. This study aims at developing univarite time series models to forecast the monthly wholesale price of coarse rice in Bangladesh.

Materials and Methods

Data collection

The monthly data of the wholesale prices of coarse rice (Tk. per quintal) from the year July 1975 to December 2011 have been collected from Department of Agricultural Marketing, Food Planning and Monitoring Unit (FPMU) of the Ministry of Food System Management Division of Bangladesh.

Logarithmic transformation of the series of coarse rice price

Coarse rice price of Bangladesh is fluctuated over the time (as one moves from left to right on the Fig. 1). This variation in the magnitude of the fluctuation with time is termed as non-stationary in variance of the data series.
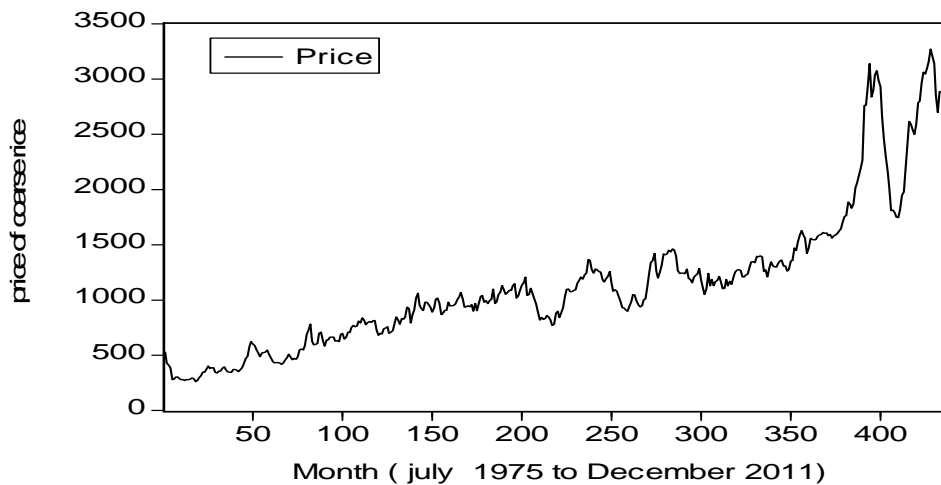
Fig.1. Original series of wholesale price of coarse rice

The main approach for achieving stationary in variance is through a logarithmic or power transformation of the data (Makridakis *et al.* 1998). Fig. 2 shows the logarithmic transformation of the original series of price of coarse rice.
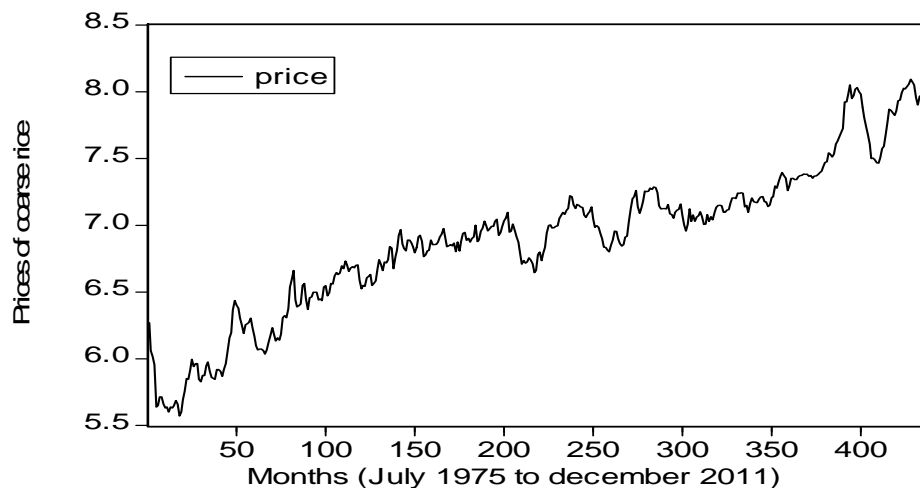
Fig. 2. Log transformed series of wholesale price of coarse rice

It is evident from the Fig. 2 that the magnitude of the fluctuations in the log transformed series does not vary so much with time.

Notations of SARIMA model

The popular Box-Jenkins methodology was applied to fit best seasonal autoregressive integrated moving average model SARIMA $(p,d,q)(P,D,Q)_S$ where p is the order of autoregression, d is the order of integration, q is the order of moving average, P is the order of seasonal autoregression, D is the order of seasonal integration, Q is the order of seasonal moving average and s is the length of seasonal period for forecasting coarse rice price of Bangladesh.

For time series analysis general auto-regressive integrated moving average ARIMA (p,d,q) model is as follows:

$$\left(1 - \varphi_1 B - \cdots - \varphi_p B^p\right)W_t = (1 - \theta_1 B - \cdots - \theta_q B^q)A_t \tag{1}$$

Where, B is the backshift operator, W_t is the first difference of the original time series Y_t and A_t is the white noise process with mean zero. The details of the models are available in O'Donovan and Pankratz (1984). Similarly, the seasonal ARIMA often called SARIMA (p,d,q) (P,D,Q)s in terms of the backward shift operator B can be expressed as-

$$\left(1 - \varphi_s B^s - \cdots - \varphi_{sP} B^{sP}\right)W_t = (1 - \theta_s B^s - \cdots - \theta_{sQ} B^{sQ})A_t \tag{2}$$

Where, $W_t = (1 - B^s)^D Y_t$, s = 12 monthly data and s =4 for quarterly data. Contrary to (1), the random shocks A_t do not from a white noise process. Combining (1) and (2) we obtain

$$\left(1 - \varphi_1 B - \cdots - \varphi_p B^p\right)\left(1 - \varphi_s B^s - \cdots - \varphi_{sP} B^{sP}\right)W_t = (1 - \theta_1 B - \cdots - \theta_q B^q)\left(1 - \theta_s B^s - \cdots - \theta_{sQ} B^{sQ}\right)A_t \tag{3}$$

Where, $W_t = (1 - B)^d (1 - B^s)^D Y_t$. This is the multiplicative model of order (p,d,q) (P, D, Q)s. Here the term $\left(1 - \varphi_1 B - \cdots - \varphi_p B^p\right)$ is known as the regular autoregressive operator of order p, the term $\left(1 - \varphi_s B^s - \cdots - \varphi_{sP} B^{sP}\right)$ is known as seasonal autoregressive operator of order P, the term $\left(1 - \theta_s B^s - \cdots - \theta_{sQ} B^{sQ}\right)$ is the seasonal moving average operator of order Q.

Data analysis

Accuracy of forecasting depends on the time series data which must be stationary. The stationarity of the series was made by means of seasonal and non-seasonal differencing. Then the order of autoregression and moving average were identified using autocorrelation function (ACF) and partial autocorrelation function (PACF) of the differenced series. A best ARIMA model was fitted for a set of coarse rice price data from July 1975 to December 2011 and the fitted model was used to predict values for a validation period (from December 2010 to December 2011) to evaluate the time series model. Most suitable models were selected on the basis of their ability of reliable prediction. Three measures, namely, Root Mean Square Error (RMSE), Normalised Bayesian Information Criteria (BIC) (Makridakis *et al.* 1998) and Mean Absolute Percent Error (MAPE) were used (Schwarz *et al.* 1978). Lower values of RMSE, Normalised BIC and MAPE were preferable. Furthermore, Ljung-Box test (portmanteau test) was performed to test if the residual ACF at different lag times was significantly different from zero, where not being different from zero was expected. After the best model was identified, forecast for future values from January 2012 to December 2013 was made.

Results and Discussion

The observed series of Wholesale price of coarse rice (July 1975 to December 2011) shows that the series is nonstationary and there are seasonal fluctuations in the dataset (Fig. 1). ACF and PACF of one seasonal differenced series as well as of one seasonal with one non seasonal differenced series (Fig. 3 to Fig. 6) instigated to explore a set of models based on the collected set of data (July 1975 to December 2011), which are listed in Table 1. Among these models, SARIMA $(1,1,1)(0,1,1)_{12}$ has both lowest RMSE (61.657), normalised BIC (8.300) and MAPE (3.906) values and appeared to be the best model. Moreover, the Ljung-Box test suggested that the ACF of residuals for the model at different lag times was

not significantly different from zero (Q18=17.394 and p>.20). All the coefficients of SARIMA (1,1,1)(0,1,1)$_{12}$ model were significant (Table 2). So the final identified model can be written as:

SARIMA $(1,1,1)(0,1,1)_{12}$

$$(1 - \varphi_1 B)(1 - B)(1 - B^{12})Y_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12}) \in_t$$

$$or \ [1 - \varphi_1 B - B + \varphi_1 B^2](1 - B^{12})Y_t = (1 - \theta_1 B - \Theta_1 B^{12} + \theta_1 \Theta_1 B^{13}) \in_t$$

$$or \ [1 - \varphi_1 B - B + \varphi_1 B^2 - B^{12} + \varphi_1 B^{13} + B^{13} - \Theta_1 B^{14}]Y_t = \in_t - \theta_1 \in_{t-1} - \Theta_1 \in_{t-12} + \theta_1 \Theta_1 \in_{t-13}$$

$$Y_t = (1 + \varphi_1)Y_{t-1} - \varphi_1 Y_{t-1} - \varphi_1 Y_{t-2} + Y_{t-12} - (1 + \varphi_1)Y_{t-13} + \varphi_1 Y_{t-14} + \in_t - \theta_1 \in_{t-1} - \Theta_1 \in_{t-12} + \theta_1 \Theta_1 \in_{t-13}$$

The model has been used to predict values from December 2010 to December 2011 (Fig. 7 and Fig. 8 for validation). It appeared that the predicted values could follow the upturn and downturn of the observed series reasonably well. Finally, Fig. 9 represents the forecast values for the period from January 2012 to December 2013, which indicates a certain ups and downs from January 2012 to December 2013.

Table 1. RMSE MAPE and normalized BIC of Time Series Models

Models	RMSE	MAPE	Normalised BIC
SARIMA (1, 1, 1) (0, 1 ,1)$_{12}$	**61.657**	**3.906**	**8.300**
SARIMA (1, 1, 1) (1, 1 ,1)$_{12}$	61.708	3.908	8.316
SARIMA (2, 1, 1) (1, 1 ,1)$_{12}$	61.789	3.908	8.333
SARIMA (2, 1, 1) (2, 1 ,1)$_{12}$	61.786	3.914	8.337
SARIMA (3, 1, 1) (0, 1 ,1)$_{12}$	61.806	3.903	8.333
SARIMA (1, 1, 1) (0, 1 ,2)$_{12}$	61.684	3.907	8.315

Table 2. Model parameters of SARIMA $(1,1,1)(0,1,1)_{12}$

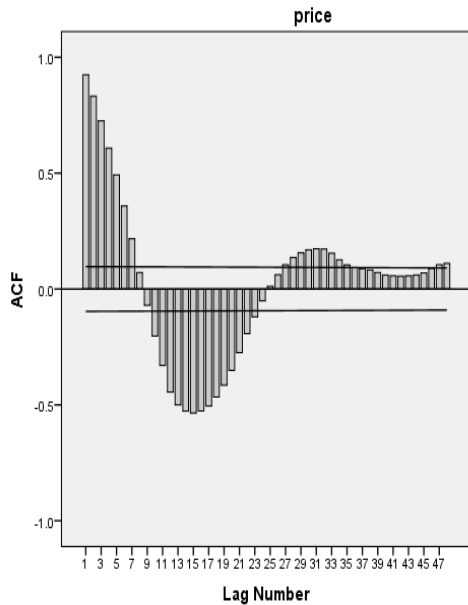| Variable	Coefficient	Standard error	P-value
AR (Lag 1)	0.733	0.140	0.00
Non seasonal difference	1		
MA (Lag 1)	0.586	0.163	0.00
Seasonal difference	1		
MA, Seasonal (Lag-1)	0.893	0.030	0.00

Fig. 3. Plot of ACF of the log transformed series after seasonal differencing
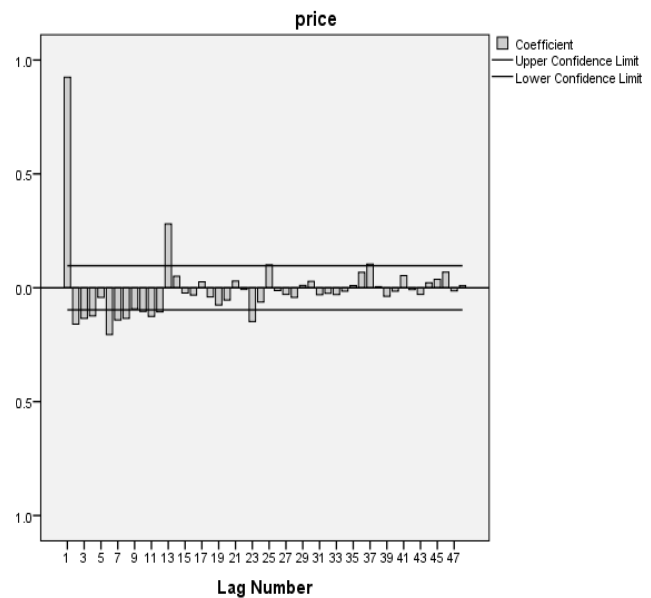

Fig. 4. Plot of PACF of the log transformed series after seasonal differencing
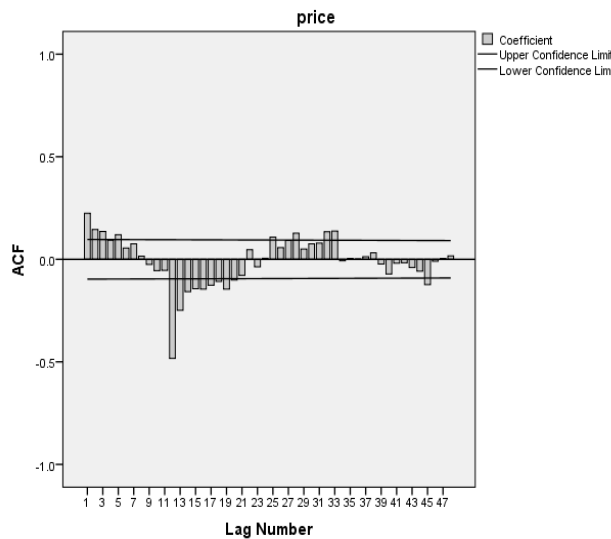
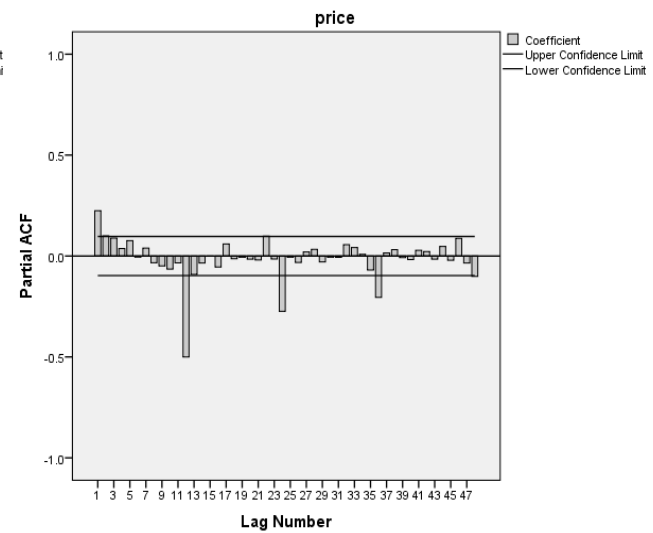Fig. 5. Plot of Autocorrelations of difference of the seasonally differenced series

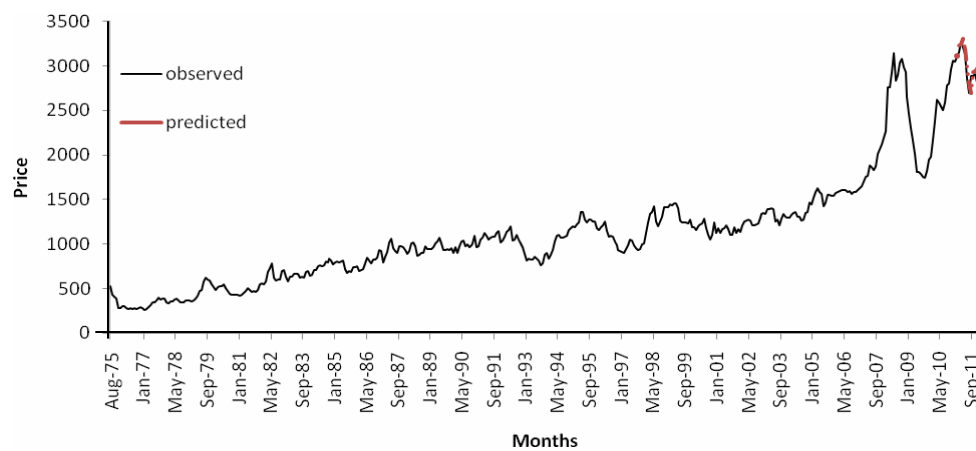Fig. 6. Plot of the PACF of the difference of the seasonally differenced series

Fig. 7. Observed values and predicted values of SARIMA$(1,1,1)(0,1,1)_{12}$ for the period from December 2010 to December 2011
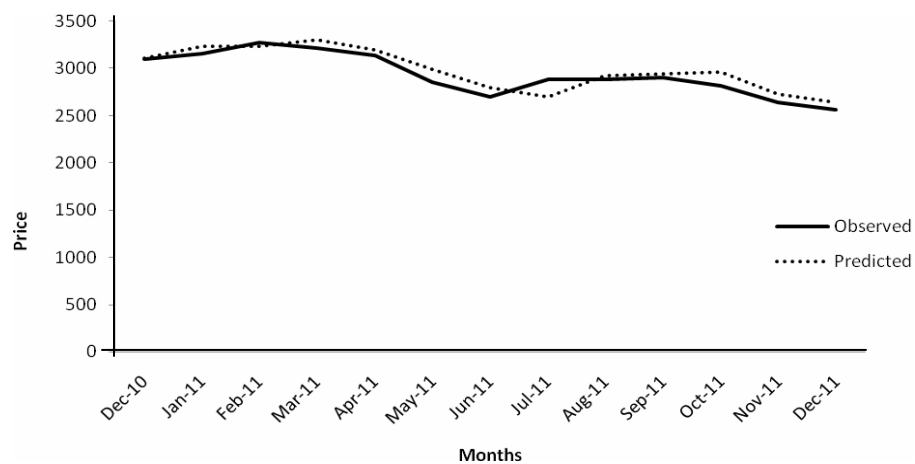

Fig. 8. Forecast of wholesale price of coarse rice from December 2010 to December 2011 by SARIMA $(1,1,1)(0,1,1)_{12}$
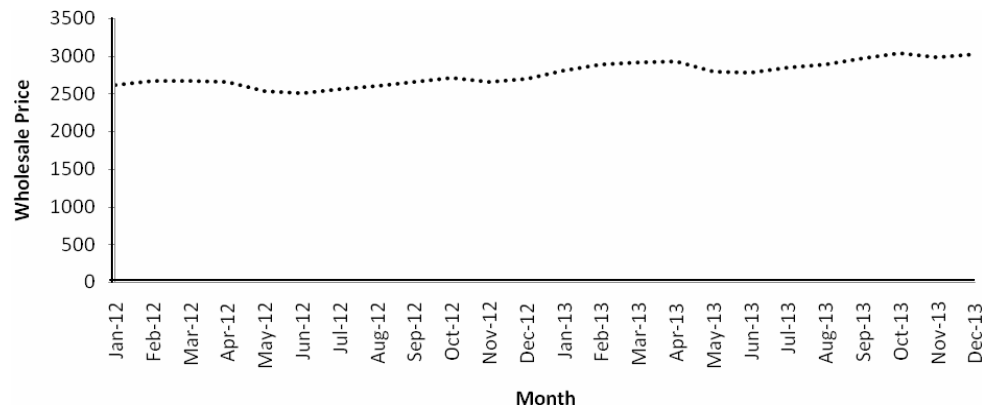
Fig. 9. Forecast of wholesale price of coarse rice from January 2012 to December 2013 by SARIMA $(1,1,1)(0,1,1)_{12}$

Conclusion

Seasonal autoregressive integrated moving average model traces out the seasonal effect of the desired variable. This research identified SARIMA $(1,1,1)$ $(0,1,1)_{12}$ as the best model for forecasting on the basis of latest model selection criteria. On an average the wholesale price of rice in Bangladesh follows an upward trend. This study forecast some future prices which would help the policy planners to take necessary action for the time to come. Basically forecasting accuracy relies on accuracy of the data on the interested variables. Data banks in our country should be well organized to get the best result by the forecasting models. For forecasting purpose mixed regression model i.e. considering various independent variables (political stability and import-export gap etc.) with basic SARIMA can be used.

References

Abraham, B. and Ledolter, J. 1983. Statistical Methods for Forecasting. *Journal of Statistical Methods for Forecasting*. New York: Wiley, 225- 229.

Bangladesh Bureau of Statistics (BBS). 2010. Bangladesh Population Census, Statistics Division, Ministry of Planning, Government of the People's Republic of Bangladesh, Dhaka.

Brennan, D. 1995. Final Research Report of the Bangladesh Foodgrains Management Operations Project: Policies for Stabilizing Rice Prices in Bangladesh, Chemonics International, Mimeo.

Brockwell, P. and Davis, R. 2002. *Introduction to time series and forecasting,* New York: Springer, U.S.A.

Makridakis, S., Wheelwright, S. and Hyndman, R.J. 1998. *Forecasting Methods and Applications*. 3rd edition, John Willey and Sons, U.S.A.

Pankratz, A. and O' Donovon. 1984. *Forecasting with univariate Box-Jenkins models: conceots and cases*, Jhon Wiley, New York.

Schwarz, G. 1978. Estimating the dimension of a model. *Annual Statistics* 6, 461-464.