Stochastic Modeling of Wholesale Price of Rohu in West Bengal, India

R. K. Paul*

Central Inland Fisheries Research Institute

Barrackpore, India

*Corresponding author E-mail:ranjitstat@gmail.com

Abstract: Among the Indian major carps, rohu (*Labeo rohita*) is the most widely consumed fish in West Bengal. The wholesale price of rohu in West Bengal fluctuates over seasons due to the variations in production and market arrival. Thus, modelling and forecasting the monthly price behaviour over the years is of much practical importance. To this end, autoregressive integrated moving average (ARIMA) methodology has been successful in describing and forecasting the fishery dynamics of a wide variety of species in the past. In the present study ARIMA stochastic modelling is used for describing monthly wholesale price of rohu in west Bengal. The monthly wholesale price data from January' 1996 to December 2004 has been taken for model building and the data from January' 2005 to December' 2005 has been used for validation of the model. The best ARIMA model is selected based on the minimum Akaike Information Criteria (AIC) and Baysian Information Criteria (BIC) values. It has been found that ARIMA (2,1,1) model described the data satisfactorily with the forecast values lie close to the actual price. The data analysis is carried out using SPSS software package version 18.0.

Keywords: ARIMA model, Forecasting, Seasonality, Stationarity, SPSS, Wholesale price.

Introduction

Fluctuations of Price for the different commodities are a matter of concern among consumers, farmers and policy makers. For the Government, unforeseen variations in wholesale price can complicate budgetary planning. So it's accurate forecast is extremely important for efficient monitoring and planning. Forecasting of the fish production or fish price for that matter is a formidable challenge. In view of globalization, it is imperative to study the trend of price of different commodities by employing sound statistical modelling techniques that, in turn will be beneficial to the planners in formulating suitable policies to face the challenges ahead. Several attempts have been made in the past to develop price forecast models for various commodities. Ghosh and Prajneshu, (2003), have studied the price fluctuation of onion; Paul et al (2009) have studied the fluctuation of export price of spice; Chandran and Pandey (2007) have studied the seasonal fluctuation of potato price in Delhi. Time-series analysis of wholesale fish price has been an important tool for fisheries management and decision making as it reveals hidden trends and seasonality patterns. Among the Indian major carps, rohu (*Labeo rohita*) is the most widely consumed fish in West Bengal. The wholesale price of rohu fluctuates over seasons due to the variations in production and market arrival. Thus, modelling and forecasting the monthly price behaviour over the years is of much practical importance. To this end, autoregressive integrated moving average (ARIMA) methodology has been successful in describing and forecasting the fishery dynamics of a wide variety of species in the past. There are two types of fishery forecasting models: deterministic models and stochastic models. The deterministic models do not have a random variable and each prediction is made under a specific set of conditions that are always the same (William, 1986). These models include the surplus production model and the classic regression model. The stochastic models, in contrast, have a random variable that represents error terms of random factor(s). These models include the autoregressive integrated moving average (ARIMA) model and the transfer function noise model (Box and Jenkins, 2007; Liu and Hanssens, 1982). The temporal variations in fish price can be studied by using Box-Jenkins approach. Here, response variable at any time 't' is assumed to be expressible as a linear function of its values at past epochs t-1, t-2, … Thus, role of various predictor variables enter into the model "implicitly" through response variable observations at past epochs. One advantage of "Implicit modelling" approach is that data requirements are much less.

In the present study ARIMA stochastic modelling is used on the monthly wholesale price of rohu in west Bengal for forecasting purpose.

Materials and Methods

The time series data of wholesale price of rohu in west Bengal from January' 1996 to December' 2005 has been collected from the handbook of fisheries statistics 2006 published by Department of Agriculture and Cooperation, Fisheries Division, Government of India. First, the time series of the wholesale price of rohu in west bengal were analyzed with an exploratory aspect. A seasonal decomposition procedure was applied in order to decompose the seasonal series into a seasonal component, a combined trend and cycle component, and an error component. The seasonal adjustment factor (SAF) was estimated in order to check the distribution patterns on a monthly basis. In the present study, the seasonality of the time series was represented by the 12-month periodicity.

Development of ARIMA Models

The development of ARIMA models was based on the methodology described in the classic work of Box and Jenkins. Univariate ARIMA models use only the information contained in the series itself. Thus, models are constructed as linear functions of past values of the series and/or previous random shocks (or errors). Forecasts were generated under the assumption that the past history could be translated into predictions for the future. ARIMA modeling was developed following the standard three-step procedure: (i) identification of the model; (ii) parameter estimation; and (iii) diagnosis and verification of the model. The identification step determines: (i) whether the process is stationary and the possible transformations to obtain stationarity; and (ii) whether the form of the process is autoregressive (AR), moving average (MA) or both (ARMA), and its order(s). Three parameters used in summarizing an ARIMA model are the AR parameter p, integration parameter d, and MA parameter q. Parameters p and q denote the order of AR and MA, while d denotes the degree of differencing the series to obtain stationarity. The autocorrelation (ACF) and partial autocorrelation functions (PACF) of a series together are the most powerful tool usually applied to reveal the correct values of the parameters. The ACF gives the autocorrelations calculated at lags 1, 2 and so on, while PACF gives the corresponding partial autocorrelations, controlling for autocorrelations at intervening lags. Parameter estimation of tentative models was determined using maximum-likelihood methods. The final results included the parameter estimates, standard errors, estimate of residual variance, standard error of the estimate, natural log likelihood, Akaike's information criterion (AIC), and Schwartz's Bayesian criterion (SBC). Model selection was based on the minimization of AIC and SBC. These criteria are descriptors of the model's parsimony as they simultaneously account for the model's fit onto the observed series alongside number of parameters used in the fit. The ability to forecast using ARIMA models was tested by applying the ARIMA methodology to available data (January 1996–December 2004), excluding the monthly data of the last year (January– December 2005), which were used for testing the forecasting power of the established models.

The Autoregressive moving average (ARMA) model, denoted as ARMA(p,q), is given by

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (1)$$

or equivalently by

$$\varphi(B) y_t = \theta(B) \varepsilon_t \quad (2)$$

Where $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$

B is the backshift operator defined by $By_t = y_{t-1}$.

A generalization of ARMA models which incorporates a wide class of nonstationary time-series is obtained by introducing "differencing" into the model. The simplest example of a nonstationary process which reduces to a stationary one after differencing is "Random Walk". A process { y_t } is said to follow Autoregressive integrated moving average (ARIMA), denoted by ARIMA(p,d,q), if $\nabla^d y_t = (1-B)^d \varepsilon_t$ is ARMA(p,q). The model is written as

$$\varphi(B)(1-B)^d y_t = \theta(B) \varepsilon_t \quad (3)$$

where ε_t are identically and independently distributed as $N(0,\sigma^2)$. The integration parameter d is a nonnegative integer. When d = 0, the ARIMA(p,d,q) model reduces to ARMA(p,q) model.

Estimation of Parameters

Estimation of parameters for ARIMA model is generally done through Nonlinear least squares method. Fortunately, several software packages are available for fitting of ARIMA models. To this end, in this paper, SPSS software package is used. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) values for ARIMA model are computed by:

$$AIC = T' \log(\sigma^2) + 2(p+q+1) \quad (4)$$

and

$$BIC = T' \log(\sigma^2) + (p + q + 1) \log T' \tag{5}$$

where T' denotes the number of observations used for estimation of parameters and σ^2 denotes the Mean square error.

Results and Discussion

The graph of the wholesale price of rohu in west Bengal is plotted in figure 1. A perusual of the graph indicates that wholesale price of rohu in west Bengal fluctuates from a minimum of Rs 2700 per quintal to a maximum of Rs 8000 per quintal of fish. The fluctuation of price also indicates that the dataset is non stationary. It can also be visualized from the plot of ACF and PACF (Figure 2(a)) of the series. The decay rate for the acf of the series is very low. But after differencing of the original series the decay rate becomes high (Figure 2(b)) resulting the identification of the order of the model very easy. To this end, Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests were used for the test of stationarity, both seasonal and nonseasonal. On the basis of minimum AIC and BIC values and considering the ACF and PACF of the wholesale price series, ARIMA (2,1,1) model is selected. The AIC and BIC values for the fitted model as computed by equations 4 and 5 are 1220.48 and 1242.57 respectively.

The seasonal adjustment factor (SAF) for the wholesale price series is given in table 1. ARIMA(2,1,1) model was applied on the seasonally adjusted series. Final prediction for the wholesale price of rohu was obtained by adding the seasonal factor of the respective month to the predicted value obtained from ARIMA (2,1,1) model. The estimate of the parameters with corresponding standard error for ARIMA (2,1,1) model is given in table 2.

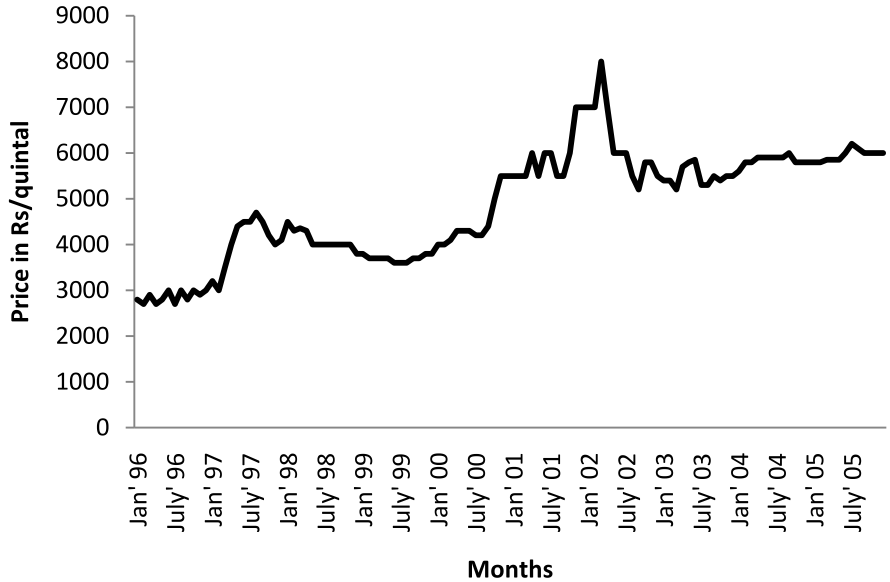

Figure 1: Monthly wholesale price of Rohu in West Bengal

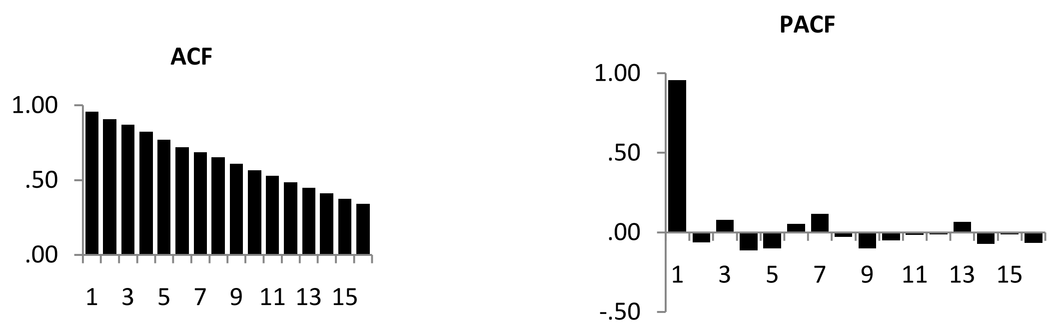

Figure 2(a) ACF and PACF of seasonally adjusted series

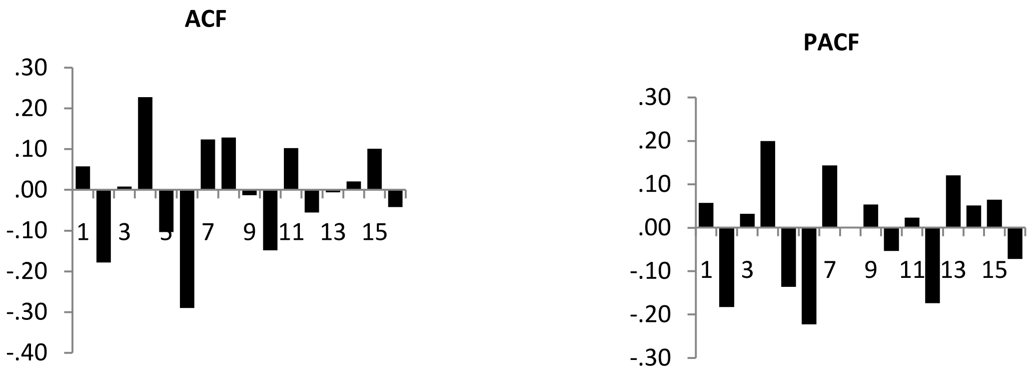

Figure 2(b) ACF and PACF of Ist differenced series of seasonally adjusted series

Table 1: Seasonal factor for the month of January to December

Month	**January**	**February**	**March**	**April**	**May**	**June**
Seasonal factor	66.70	-4.12	148.16	207.33	15.66	54.20
Month	**July**	**August**	**September**	**October**	**November**	**December**
Seasonal factor	-69.39	-174.96	-232.25	-74.96	66.70	-3.08

Table 2: Parameters estimate of the ARIMA(2,1,1) model

Parameters	**Estimate**	**SE**
Constant	28.851	24.348
AR1	0.002	0.541
AR2	-0.176	0.104
MA1	-0.066	0.55

The fitted ARIMA(2,1,1) model along with the actual data points is shown in Figure 3. A perusal of figure 3 indicates that the above ARIMA model has nicely captured the variation in wholesale price of rohu in west Bengal. The forecast of the monthly price of rohu for the period January'2005 to December' 2005 is given in table 3. The forecast values are very near to the actual values indicating the superiority of fitted ARIMA model.

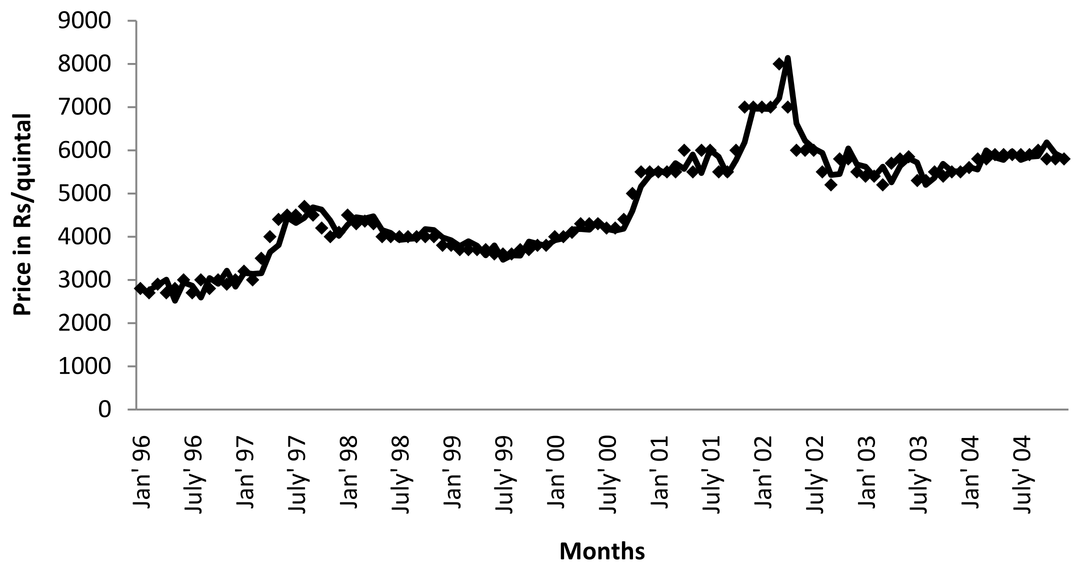

Figure 3: Fitted ARIMA(2,1,1) model with actual data points

Table 3: Forecast of monthly wholesale price of rohu in West Bengal

Months	Actual	Forecast by ARIMA(2,1,1)	UCL	LCL
Jan' 05	5800	5927.549	6476.123	5378.976
Feb' 05	5800	5878.415	6681.318	5075.512
mar' 05	5850	6054.446	6994.887	5114.005
April' 05	5850	6143.713	7200.847	5086.579
May' 05	5850	5981.801	7151.139	4812.462
June' 05	6000	6048.975	7321.147	4776.803
July' 05	6200	5954.066	7320.264	4587.868
Aug' 05	6100	5877.388	7331.467	4423.309
Sept' 05	6000	5848.975	7386.09	4311.86
Oct' 05	6000	6035.111	7651.013	4419.209
Nov' 05	6000	6205.623	7896.618	4514.629
Dec' 05	6000	6164.684	7927.573	4401.794

Conclusion:

ARIMA model being stochastic in nature; it could be successfully used for modeling as well as forecasting of wholesale monthly price of rohu in west Bengal. The model demonstrated a good performance in terms of explained variability and predicting power. The relevant forecast interval for the wholesale price can help both the fish farmers as well as the planners for future planning.

References:

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (2007). *Time-Series Analysis: Forecasting and Control*. 3rd edition. Pearson education, India.

Chandran, K. P. and Pandey, N. K. (2007). Potato price forecasting using seasonal ARIMA approach. *Potato Journal*. **34** (1-2) : 137-138.

Ghosh, H. and Prajneshu. (2003). Non-linear time series modelling of volatile onion price data using AR(p)-ARCH(q)-in mean. *Calcutta Statistical Association Bulletin*. **54**: 231-47

Liu, L. M. and D. M. Hanssens (1982). Identification of multiple-input transfer function models. *Communications in Statistics - Theory and Methods*, **11**(3): 297-314.

Ministry of Agriculture (Eds.), 2006. Handbook of Fisheries Statistics. Department of Agriculture and Cooperation. Fisheries Division, Government of India. New Delhi.

Paul, R. K., Prajneshu, and Ghosh, H. (2009). GARCH Nonlinear Time Series Analysis for Modeling and Forecasting of India's Volatile Spices Export Data. *Journal of the Indian Society of Agricultural Statistics*, **63** (2), 123-131.

William, E.G., (1986). *Systems analysis and simulation in wildlife and fisheries sciences*. John Wiley and Sons, New York, NY, 338 pp.