Journal of Animal Research: v.8 n.2, p. 259-262. April 2018

DOI: 10.30954/2277-940X.04.2018.14

Modeling of Monthly Arrival of Rohu Fish using ARIMA in Jammu Region of J&K State

Manish Sharma[1*], Nishant Jasrotia[1], Banti Kumar[1], Anil Bhat[2] and Sunali Mahajan[1]

[1]*Division of Statistics & Computer Science, FBSc, SKUAST Jammu, Jammu, J&K, INDIA*
[2]*Division of Agricultural Economics and ABM, FoA, SKUAST Jammu, INDIA*

**Corresponding author: MK Sharma; Email: manshstat@gmail.com*

Received: 07 March., 2018 **Revised:** 20 March, 2018 **Accepted:** 24 March, 2018

ABSTRACT

In this study, the Box-Jenkins methodology has been applied to build Autoregressive Integrated Moving Average (ARIMA) model for monthly arrival of Rohu fish in Jammu region of J&K state. The values ADF and Durbin Watson statistics were -3.65** and 2.11 respectively. Among the ARMA(1,1), ARIMA(1,1,1), ARMA(2,2), ARMA(2,1), ARMA(1,2), ARIMA(2,1,1) ARIMA(1,1,2), ARIMA(1,1,3), ARIMA(2,1,2), ARIMA(2,2,2), the best model obtained was ARMA (2, 2) on the basis of significance of model and parameters. The values of R^2 MAPE and -2 loglikelihood were 0.74, 0.70, 4.23 and 473.80, respectively. The AIC and SBIC of selected model were 483.01 and 459.51.

Keywords: Box-Jenkins approach, AIC, SBIC, ARIMA, forecasting

Fisheries and aquaculture provide direct and indirect employment to over 500 million people across the world. According to Food and Agriculture organization (FAO) of the United Nations, the total number of commercial fishermen and fish farmer is estimated to be 38 million. Fishing in India is a major industry in its coastal states, employing over 14 million people. Fish production in India has increased more than tenfold since its independence in 1947. According to the Food and Agriculture organization (FAO) of the United Nations, fish output in India doubled between 1990 and 2010. India is a major supplier of fish in the world and account for 4.4 per cent of global fish production.

The J&K state is blessed with plenty of aquatic resources in the form of rivers, ponds, reservoirs and wetlands. Since last few decades, fish has been an important food item for the inhabitants of the state. The state has 27,781 km length of river/streams which facilitates fish farming of more than 40 million tonnes of fish. During the year 2012-13, 2 lakh quintals fish production was recorded and revenue receipts from fisheries was ₹ 520.33 lakh (economic survey of J&K 2013-14). As per livestock census 2003 of J&K state, the total fishermen population was around 31,000. It is presently estimated to be 91,984. The fish catch which was 1, 84,667 quintals in 2000-01, has reached to 1, 99,500 quintals in 2012-13. During the year 2012-13, 2 lakh quintals fish production was recorded and revenue receipts from fisheries were ₹ 520.33 lakh as per economic survey of J&K 2013-14. A time series is a sequential set of data points, measured typically over successive times. The procedure of fitting a time series to a proper model is termed as Time Series Analysis. There are two widely used linear time series models in literature, *viz.,* Autoregressive (AR) and Moving Average (MA) models. Combining these two, the Autoregressive moving average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) models can be obtained. The ARIMA (p, d, q) model where p, d and q are non- negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. By keeping all the above into consideration, the study has been conducted with the objective to build up an ARIMA model for predicting the monthly arrivals of Rohu fish for Jammu region.

MATERIALS AND METHODS

The secondary data for fish arrivals in Jammu have been procured from fisheries department of Jammu for the last three years w.e.f. Jan., 2013 to Dec. 2015 on daily basis. The models have been build up and used for forecasting the monthly demand of Rohu fish for Jammu region through ARIMA method. The ARIMA (Box-Jenkins model) has been developed through the iterative three stages as:

Identification and model selection

The first step in developing ARIMA (Box-Jenkins model) is to determine stationarity of the time series data of fish arrival in Jammu region. In case of non-stationarity in the data the differencing approach has been used to achieve the stationarity. Once stationarity and seasonality have been addressed, the next step is to identify the order (*i.e.,* the p and q) of the autoregressive and moving average terms. The partial autocorrelation of an AR (p) process becomes zero at lag p+1 and greater, which can be examined through sample partial autocorrelation function to see if there was evidence of a departure from zero. Further autocorrelation function of a MA (q) process becomes zero at lag q+1 and greater, so we have examined the sample autocorrelation function to see whether it essentially becomes zero (Pankratz, 1991). The stationarity of the data has been tested through the Augmented Dickey Fuller test (ADF) test. The presence of autocorrelation in time series data have been tested through the Durbin Watson test.

Estimation Stage

After identifying the order of the tentative model, the parameters of the model has been estimated using the maximum likelihood estimation or ordinary least square to determine the AR and MA parameters, as well as all other parameters reported in the study. Three other penalty function statistics namely the Akaike information criteria (AIC), the Schwarz Bayesian information criteria as well as the corrected Akaike information criteria (AIC) has been used as per the suitability in penalizing fitted model based on the principle of parsimony.

Diagnostic checking

The diagnostic stage of the Box-Jenkins ARIMA process is to examine whether the fitted model follows a white noise process. This can be done by studying the autocorrelation values (r_k) one at a time, and to develop a standard error formula to test whether a particular r_k value is significantly different from zero. For a white noise process, 95 per cent of all sample autocorrelation values (r_k) must lie within a range specified by the mean plus or minus 1.96 standard errors. In this case, since the mean of the process is zero and the standard error is $1/\sqrt{n}$, one should expect about 95 per cent of all sample autocorrelation values (r_k) to be within the range of $\pm 1.96\sqrt{n}$ or $\left(-1.96/\sqrt{n} < r_k < 1.96/\sqrt{n}\right)$.

If this condition has not hold, then the model fitted do not follow a white noise process, or the residuals are not white noise. The Ljung-Box test is a modified version of the portmanteau test statistic developed by Ljung and Box (1978) has been also used. The modified Ljung- Box Q statistic tests whether the model's residuals have a mean of zero, constant variance and serially uncorrelated r_k values (a white noise check). The test statistic is given by; $Q = n(n+2)\sum_{k=1}^{h} \frac{r_k^2}{(n-k)}$ where, n denote the number of data points in the series, r^2_k is the square of the autocorrelation at lag k, and h is the maximum lag being considered. The hypothesis to be tested is formulated in the form;

H_0: *The set of autocorrelations for residual is white noise (model fit data quite well)*

H_1: *The set of autocorrelations for residual is different from white noise*

The test statistic (Q) is compared with a chi-square distribution written as where, α is taken to be 5 per cent (0.05), h is the maximum lag being considered, and p and q are the order of the AR and MA processes respectively.

RESULTS AND DISCUSSION

The behavior of the data of monthly arrival of Rohu fish can be seen from Fig. 1. The scenario of the data through figure have shown that the arrival of Rohu fish during January, 2013 was 3519 kg and arrival of Rohu fish reached to 4603 kg in Jammu by the ending December, 2015. The minimum arrival was found in the third month of 2013 which was, 3313 kg and maximum was observed in the fifth month of 2015 as 4880 kg. The Fig. 1 also showed sudden jump in the demand of Rohu fish which increased rapidly after 2014. This behavior of the data showed trend in the data, so trend modeling is required. The Correlogram of monthly arrival Rohu fish for the period of 2013 to 2015 has been displayed in Fig 2.

The Fig. 2 represented the serial correlation coefficients with Ljung-Box Q (LBQ) value (with p-value) for consecutive lags in a specified range of lags. The ACF values were maximum at lag 1 (0.81) which decreased with the lag but not to zero or equal to zero. At lag 11, the ACF value was negative and gradually decreases whereas, the PACF value showed wave like structure with positive and negative values. The maximum PACF value was at lag 1 which was 0.81. The Augmented Dickey-fuller (ADF) test value (see table 1) was 3.65 and significant. The hypothesis of ADF test was that H_0: the non-stationary of the data against H_1: the data are stationary. The critical value for the rejection of the null hypothesis of unit root test was significant at no difference.

Table 1: Stationary test result and testing of autocorrelation for monthly arrival of Rohu fish through augmented Dickey-Fuller test and Durbin Watson statistics

Test used	Test value (P-value)
Augmented Dickey-Fuller(ADF)	3.65 (0.0402)
Durbin-Watson statistics	2.10

*Significant at 5%

Table 1 showed that the data of monthly arrival of Rohu fish were stationary. The value of Durbin-Watson statistics was 2.10, indicating the presence of very low negative serial correlation in the data. Further, to fulfill the above cited conditions, different models have been applied on the data. Among the 25 models, 10 models have been listed in Table 2. As the data were stationary but some models have been selected with some differencing to see the behaviour of the model. All the models with difference 1 or 2 were non-significant so they have been dropped for further conclusion or forecasting purposes. On the basis of significance of parameters, minimum Akaike Information Criteria (AIC) and minimum Schwartz Bayesian Information Criteria (SBIC), the best model has been selected. Among all the models, the ARIMA (2, 2, 2) having minimum AIC and model ARMA (2,2) have minimum SBIC, but the parameters of ARIMA(2,2,2) was non-significant whereas of ARMA(2,2) was significant and hence selected for further study. Moreover, from the Fig. 2, it has been observed that the two spikes were out of correlogram of ACF, whereas the two spikes of PACF having maximum values. So, the value of p may be 2 and the value of q may be 1 or 2. The value of d was zero because the data was stationary. On the basis of the values of p, d and q, ARMA (2, 2) model was selected.

Table 2: Different models for monthly arrival of Rohu fish in Jammu region

Model	Significance of Parameters/ Model	Akaike Information Criteria (AIC)	Bayesian Information Criteria
ARMA(1,1)	Significant	530.14	534.81
ARIMA(1,1,1)	Non-significant	464.80	469.38
ARMA(2,2)	Significant	483.01	459.51
ARMA(2,1)	Significant	483.29	491.58
ARMA(1,2)	Significant	483.80	490.03
ARIMA(2,1,1)	Non-significant	463.22	469.33
ARIMA(1,1,2)	Non-significant	464.73	470.84
ARIMA(1,1,3)	Non-significant	466.59	474.22
ARIMA(2,1,2)	Non-significant	463.64	471.27
ARIMA(2,2,2)	Non-significant	459.14	466.26

The model selected was ARMA (2,2). The AIC and SBIC of selected model were 483.01 and 459.51. From the Fig. 3, it can be observed that the LBQ values increases and non-significant indicated there was zero autocorrelation. The estimation of significant model ARMA (2, 2) with parameters has been displayed in the table 3. Based on the parameters as shown in table 4, the values of AR (1), AR (2), MA (1) and MA (2) were -0.03, 0.93, -0.68 and 0.22, respectively. The values of AR (1) and MA (1) were significant whereas for MA (1) it was negative and significant. The value MA (2) was non-significant.

Thus, the null hypothesis of parameters was or equal to zero is rejected for MA (1) and MA (2). The values of R^2, $\overline{R}^2$, MAPE and -2loglikelihood were as 0.74, 0.74, 4.23 and 473.80, respectively. Since, the model ARMA (2, 2) was significant, AIC and SBIC having minimum value and high R^2 and hence chosen for study the forecasting

Table 3: Parameter estimates of ARMA (2,2) for monthly arrival of Rohu fish in Jammu region

Term	Lag	Estimate (S.E.)	t-test (p-value)	R^2		MAPE	-2logliklehood	F-value
Intercept	0	4069.402	11.37					
		(357.8068)	(<.0001)	0.74	0.74	4.23	473.80	25.34**
AR	1	-0.033	-0.43					
		(0.0768)	(0.0041)					
AR	2	0.935	12.73					
		(0.0735)	(<.0001)					
MA	1	-0.685	-3.37					
		(0.2032)	(0.0021)					
MA	2	0.221	1.20					
		(0.1839)	(0.2384)					

trend for monthly arrival of Rohu fish. The ARMA (2, 2) was fitted as per the observed fact; the model was used to forecast the values. Thus the proposed model is as:

$$\hat{Y}_t = 621.861 - 0.033Y_{t-1} + 0935Y_{t-2} + 0.685e_{t-1} - 0.221e_{t-2}$$

Table 4: The forecasted value values of Rohu fish arrival

Sl. No	Time	Forecasted value (in Kg)
1	Oct., 2017	6213.54
2	Oct., 2018	6572.33
3	Oct., 2019	6768.04
4	Oct., 2020	6874.91

Table 4 shows the predicted value of demand of Rohu fish on the basis of proposed model for the coming years during the month of October which is in increasing trend continuously. Thus, there is a good scope for fishermen to adopt fishery enterprise and Government should bring some good policy for them.

CONCLUSION

It has been observed that the Box-Jenkins methodology applied for forecasting the monthly demands of Rohu fishes by using ARIMA method reveals that the overall demand of Rohu fish was increasing at slow rate. On the basis of above all it has been concluded that the forecasted ARIMA model for arrival of Rohu fish on monthly basis ARMA (2, 2) was the best among the 25 models. The model was significant and their parameters were also significant and also have minimum AIC and SBIC. The forecasted values of Rohu fish has been reached to 3519 Kg during Jan, 2013 to 6874.91 Kg during Oct., 2020. Thus, in future, there is a good scope in fishery enterprise in Jammu and Kashmir.

REFERENCES

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. 1994. *Time Series Analysis, Forecasting and Control,* 3rd ed. Prentice Hall, Englewood Clifs, NJ.

Garret, H.E. and Woodworth, R.S. 1969. Statistics in psychology and education. Vakils, Feffer and Simon Pvt. Ltd., Bombay, pp.329.

Ljung, G.M. and G.E.P. Box 1978. On a measure of lack of fit in time series models. *Biometrika,* **65**: 297-303.

Pankratz, A. 1991. Forecasting with univariate Box-Jenkins models: concepts and cases, John Wiley, New York.

Singh, A.K., Singh, S.P., Dwivedi, M.C., Singh, H. and Parihar, P. 2013. Natural Resources and Environmental management: Is there a need for special emphasis for Priority setting in fisheries research. *Ind. Jr. App. Res.,* **3**(12): 1-3.