Recent Research in Science and Technology 2011, 3(7): 104-108
ISSN: 2076-5061
www.scholarjournals.org

www.recent-science.com

RRST-Statistics

Forecasting Fish Product Export in Tamilnadu – A Stochastic Model Approach

T. Jai Sankar*

Department of Statistics, Bharathidasan University, Tiruchirappalli – 620 024, Tamilnadu, India

Article Info

Article History

Received : 17-05-2011
Revised : 12-07-2011
Accepted : 27-07-2011

*Corresponding Author

Tel : +91-9444469629

Email:
tjaisankar@gmail.com

Abstract

This study aims at forecasting fish product export in Tamilnadu, based on data on inland and marine fish product export during the years from 1969 to 2008. The study considered Autoregressive (AR), Moving Average (MA) and Autoregressive Integrated Moving Average (ARIMA) processes to select the appropriate stochastic model for forecasting fish product export in Tamilnadu. Based on ARIMA (p, d, q) and its components ACF, PACF, Normalized BIC, Box-Ljung Q statistics and residuals estimated, ARIMA (0, 1, 2) was selected. Based on the chosen model, it could be predicted that the fish product export would increase to 1, 14,695 tons in 2015 from 74,549 tons in 2008 in Tamilnadu.

Key Words: Fish product export, BIC, Forecasting, ARIMA.

Introduction

Fisheries play a vital role in providing protein-rich food at an affordable price to the people. The inland fisheries sector in Tamilnadu is spread over 3.71 lakh hectare of water-spread area comprising reservoirs, irrigation and seasonal tanks, ponds, estuaries and backwaters. Besides, the State has 56000 hectare of brackish water area suitable for aquaculture, of which, an area of 4455 hectare has already been developed for aquaculture. In Tamilnadu, maritime sector dominates the fishery sector, as the State has a coastal line of 1076 km (13.3% of the Nation's coast line of 8118 km). The sector provides employment to 10.02 lakhs of marine and inland fisher-folk and contributes around Rs.20,000 crore to foreign exchange which accounts for 27.5% of exports from India. In this background, this study was conducted to forecast the future fish product export in the State, so as to help the policy planners to formulate needed strategies for achieving and sustaining the targets in the sector.

Material and Methods

As the aim of the study was to forecast fish product export, various forecasting techniques were considered for use. ARIMA model, introduced by Box and Jenkins (1970)[1], was frequently used for discovering the pattern and predicting the future values of the time series data. Akaike (1970)[2] discussed the stationary time series by an AR(p), where p is finite and bounded by the same integer. Moving Average (MA) models were used by Slutzky (1973)[3]. Hannan and Quinn (1979)[4] suggested obtaining the order of a time series model by minimizing the errors for pure AR models, and Hannan (1980)[5] for ARMA models. A second order determination method could be considered as a variance of Schwarz's Bayesian Criterion (SBC) which gives a consistent estimate of the order of an ARMA model. Hosking (1981)[6] introduced a family of models, called fractionally differenced autoregressive integrated moving average models, by generalizing the 'd' fraction in ARIMA (p, d, q) model.

Stochastic time-series ARIMA models were widely used in time series data having the characteristics (Alan Pankratz, 1983[7]) of parsimonious, stationary, invertible, significant estimated coefficients and statistically independent and normally distributed residuals. When a time series is non-stationary, it can often be made stationary by taking first differences of the series i.e., creating a new time series of successive differences (Y_t-Y_{t-1}). If first differences do not convert the series to stationary form, then first differences can be created. This is called second-order differencing. A distinction is made between a second-order differences (Y_t-Y_{t-2}).

While Mendelssohn (1981)[8] used Box-Jenkins[1] models to forecast fishery dynamics, Prajneshu and Venugopalan (1996)[9] discussed various statistical modeling techniques viz., polynomial, ARIMA time series methodology and nonlinear mechanistic growth modeling approach for describing marine, inland as well as total fish production in India during the period 1950-51 to 1994-95. Tsitsika *et al.* (2007)[10] also used univariate and multivariate ARIMA models to model and forecast the monthly pelagic production of fish species in the Mediterranean Sea during 1990-2005. Jai Sankar *et al.* (2010)[11] also used stochastic modeling for cattle production and forecast the yearly production of cattle in the Tamilnadu state during 1970-2010.

The time series when differenced follows both AR and MA models and is known as autoregressive integrated moving averages (ARIMA) model. Hence, ARIMA model was used in this study, which required a sufficiently large data set and involved four steps: identification, estimation, diagnostic checking and forecasting. Model parameters were estimated using the Statistical Package for Social Sciences (SPSS) package and to fit the ARIMA models.

Autoregressive process of order (p) is, $Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t$;

Moving Average process of order (q) is, $Y_t = \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - .... - \theta_q \varepsilon_{t-q} + \varepsilon_t$;

and the general form of ARIMA model of order (p, d, q) is

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where Y_t is fish product export, ε_t's are independently and normally distributed with zero mean and constant variance σ^2 for t = 1,2,..., n; d is the fraction differenced while interpreting AR and MA and ϕs and θs are coefficients to be estimated.

Trend Fitting: The Box-Ljung Q statistics was used to transform the non-stationary data in to stationarity data and to check the adequacy for the residuals. For evaluating the adequacy of AR, MA and ARIMA processes, various reliability statistics like R^2, Stationary R^2, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Bayesian Information Criterion (BIC) were used. The reliability statistics viz. RMSE, MAPE, BIC and Q statistics were computed as below:

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \right]^{1/2}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(Y_i - \hat{Y}_i)}{Y_i} \right|$$

BIC(p,q) = ln v*(p,q) + (p+q) [ln (n) / n]

where p and q are the order of AR and MA processes respectively and n is the number of observations in the time series and v* is the estimate of white noise variance σ^2.

$$Q = \frac{n(n+2)\sum_{i=1}^{k} rk^2}{(n-k)}$$

where n is the number of residuals and rk is the residuals autocorrelation at lag k.

In this study, the data on fish product export in Tamilnadu were collected from the Department of Fisheries, Government of Tamilnadu for the period from 1969 to 2008 and were used to fit the ARIMA model to predict the future product export.

Results and Discussion

Model Identification: ARIMA model was designed after assessing that transforming the variable under forecasting was a stationary series. The stationary series was the set of values that varied over time around a constant mean and constant variance. The most common method to check the stationarity is to explain the data through graph and hence is done in Figure 1.

Figure 1 reveals that the data used were non-stationary. Again, non-stationarity in mean was corrected through first differencing of the data. The newly constructed variable Y_t could now be examined for stationarity. Since, Y_t was stationary in mean, the next step was to identify the values of p and q. For this, the autocorrelation and partial autocorrelation coefficients (ACF and PACF) of various orders of Y_t were computed and presented in Table 1 and Figure 2.

The tentative ARIMA models are discussed with values differenced once (d=1) and the model which had the minimum normalized BIC was chosen. The various ARIMA models and the corresponding normalized BIC values are given in Table 2. The value of normalized BIC of the chosen ARIMA was 16.881.

Table 1. ACF and PACF of fish product export

Lag	Auto Correlation		Box-Ljung Statistic			Partial Auto Correlation	
	Value	Df	Sig.	Value	Df	Value	Df
1	-0.228	0.154	2.185	1	0.139	-0.228	0.160
2	0.253	0.152	4.948	2	0.084	0.212	0.160
3	-0.323	0.150	9.592	3	0.022	-0.254	0.160
4	0.074	0.148	9.844	4	0.043	-0.078	0.160
5	0.115	0.146	10.462	5	0.063	0.276	0.160
6	0.361	0.144	16.780	6	0.010	0.421	0.160
7	-0.141	0.141	17.777	7	0.013	-0.164	0.160
8	0.140	0.139	18.791	8	0.016	0.036	0.160
9	-0.471	0.137	30.622	9	0.000	-0.268	0.160
10	0.259	0.135	34.326	10	0.000	0.018	0.160
11	-0.052	0.132	34.483	11	0.000	0.063	0.160
12	0.178	0.130	36.353	12	0.000	-0.151	0.160

Table 2. BIC values of ARIMA (p, d, q)

ARIMA (p, d, q)	BIC values
(0, 1, 0)	16.945
(0, 1, 1)	16.940
(0, 1, 2)	16.881
(1, 1, 0)	16.955
(1, 1, 1)	17.012
(1, 1, 2)	17.113
(2, 1, 0)	17.069
(2, 1, 1)	17.119
(2, 1, 2)	17.064

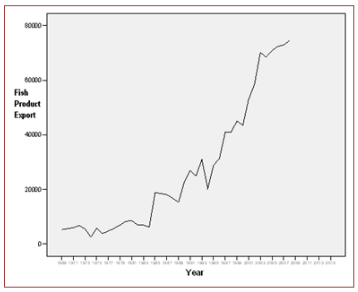

Figure 1. Time plot of fish product export in Tamilnadu

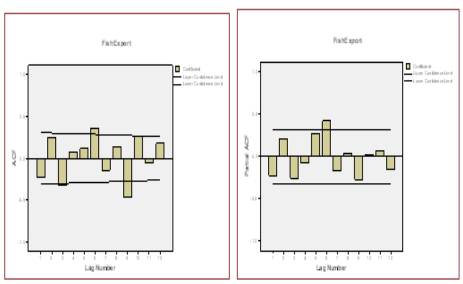

Figure 2. ACF and PACF of differenced data

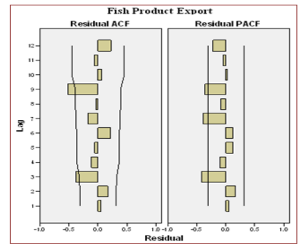

Figure 3. Residuals of ACF and PACF

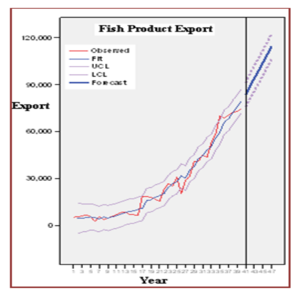

Figure 4. Actual and estimate of fish product export

Model Estimation: Model parameters were estimated using SPSS package and the results of estimation are presented in Tables 3 and 4. R^2 value was 0.97. Hence, the most suitable model for fish production was ARIMA (0, 1, 2), as this model had the lowest normalized BIC value, good R^2 and better model fit statics (RMSE and MAPE).

Table 3. Estimated ARIMA model of fish production

	Estimate	SE	t	Sig.
Constant	-260934.103	24933.056	-10.465	0.000
MA 1	0.739	162.592	0.005	0.996
MA 2	0.261	42.491	0.006	0.995

Table 4. Estimated ARIMA model fit statistics

Fit Statistic	Mean
Stationary R-squared	0.318
R-squared	0.976
RMSE	3837.419
MAPE	17.219
Normalized BIC	16.881

Table 5. Residual of ACF and PACF of fish product export

Lag	ACF		PACF	
	Mean	SE	Mean	SE
Lag 1	0.049	0.160	0.049	0.160
Lag 2	0.168	0.161	0.166	0.160
Lag 3	-0.386	0.165	-0.413	0.160
Lag 4	-0.112	0.187	-0.099	0.160
Lag 5	-0.059	0.188	0.123	0.160
Lag 6	0.214	0.189	0.120	0.160
Lag 7	-0.179	0.195	-0.385	0.160
Lag 8	-0.031	0.199	-0.086	0.160
Lag 9	-0.517	0.199	-0.358	0.160
Lag 10	0.063	0.231	0.029	0.160
Lag 11	-0.062	0.232	-0.051	0.160
Lag 12	0.224	0.232	-0.224	0.160

Table 6. Forecast of fish product export (in tons) in Tamilnadu

Year	Actual	Predicted	LCL	UCL
1969	5220	--	--	--
1970	5637	4601	-5079	14282
1971	5942	4800	-4310	13910
1972	6765	4913	-3742	13569
1973	5511	5384	-3056	13824
1974	2572	4976	-3316	13269
1975	5719	3989	-4202	12181
1976	3800	5378	-2739	13495
1977	4720	4694	-3365	12754
1978	5743	5505	-2509	13519
1979	6945	6155	-1822	14131
1980	8294	7077	-869	15023
1981	8550	8139	219	16059
1982	7000	8952	1054	16850
1983	6872	9307	1428	17186
1984	6252	10219	2357	18082
1985	18792	10900	3052	18748
1986	18456	15880	8045	23715
1987	18053	16379	8555	24202
1988	16745	18026	10213	25839
1989	15330	19107	11303	26910
1990	22768	20287	12492	28082
1991	26851	24141	16353	31928
1992	24949	26634	18853	34414
1993	30963	27869	20095	35642
1994	20311	31772	24004	39540
1995	28831	30257	22495	38019
1996	31330	35527	27770	43284
1997	40878	37592	29840	45344
1998	41052	42699	34951	50446
1999	45026	44600	36857	52344
2000	43464	48486	40746	56225
2001	53005	50391	42655	58127
2002	58482	56031	48298	63763
2003	70147	59822	52093	67551
2004	68462	66061	58335	73787
2005	70809	68141	60418	75864
2006	72418	72438	64717	80158
2007	72883	76025	68307	83743
2008	74549	79533	71818	87248

2009	--	83443	75730	91156
2010	--	89378	81360	97397
2011	--	94178	86159	102196
2012	--	99109	91091	107127
2013	--	104172	96155	112190
2014	--	109368	101350	117385
2015	--	114695	106679	122712

Diagnostic Checking: The model verification is concerned with checking the residuals of the model to see if they contained any systematic pattern which still could be removed to improve the chosen ARIMA, which has been done through examining the autocorrelations and partial autocorrelations of the residuals of various orders. For this purpose, various autocorrelations up to 12 lags were computed and the same along with their significance tested by Box-Ljung statistic are provided in Table 5. As the results indicate, none of these autocorrelations was significantly different from zero at any reasonable level. This proved that the selected ARIMA model was an appropriate model for forecasting fish product export in Tamilnadu.

The ACF and PACF of the residuals are given in Figure 3, which also indicated the 'good fit' of the model. Hence, the fitted ARIMA model for the fish product export data was:

$$Y_t = -260934.103 - 0.739\varepsilon_{t-1} - 0.261\varepsilon_{t-2} + \varepsilon_t$$

Forecasting: Based on the model fitted, forecasted fish product export (in tons) for the year 2009 through 2015 respectively were 83443, 89378, 94178, 99109, 104172, 109368 and 114695 tons (Table 6). To assess the forecasting ability of the fitted ARIMA model, the measures of the sample period forecasts' accuracy were also computed. This measure also indicated that the forecasting inaccuracy was low. Figure 4 shows the actual and forecasted value of fish product export (with 95% confidence limit) in the State.

Conclusion

The most appropriate ARIMA model for fish product export forecasting was found to be ARIMA (0, 1, 2). From the forecast available from the fitted ARIMA model, it can be found that forecasted product export would increase to 1,14,695 tons in 2015 from 74,549 tons in 2008. That is, using time series data from 1969 to 2008 on fish product export, this study provides evidence on future fish product export in the State, which can be considered for future policy making and formulating strategies for augmenting and sustaining fish product export in the State.

References

[1] Box G E P and Jenkins J M. 1970. Time Series Analysis - Forecasting and Control. Holden-Day Inc., San Francisco

[2] Akaike H. 1970. Statistical Predictor Identification. *Annals of Institute of Statistical Mathematics* 22: 203-270.

[3] Slutzky E. 1973. The summation of random causes as the source of cyclic processes. *Econometrica* 5:105-146.

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

[5] Hannan E J. 1980. The estimation of the order of an ARMA process. *Annuals of Statistics* 8: 1071-1081.

[6] Hosking J R M. 1981. Fractional differencing. *Biometrika* 68(1): 165-176.

[7] Alan Pankratz. 1983. Forecasting with univariate Box–Jenkins models - concepts and cases. John Wiley, New York, Page 81.

[8] Mendelssohn R. 1981. Using Box-Jenkins models to forecast fishery dynamics: identification, estimation and checking. *Fishery Bulletin* 78(4): 887-896.

[9] Prajneshu and Venugopalan R. 1996. Trend analysis in all India marine products export using statistical modeling techniques. *Indian Journal of Fisheries* 43(2): 107-113.

[10] Tsitsika E V, Maravelias C D and Haralabous J. 2007. Modeling and forecasting pelagic fish production using univariate and multivariate ARIMA models. *Fisheries Science* 73: 979-988.

[11] Jai Sankar, Prabakaran R, Senthamarai Kannan K and Suresh S. 2010. Stochastic Modeling for Cattle Production Forecasting. *Journal of Modern Mathematics and Statistics* 4(2): 53-57.