Journal of Agricultural Engineering and Food Technology
p-ISSN: 2350-0085; e-ISSN: 2350-0263; Volume 4, Issue 4; October-December, 2017, pp. 178-183

http://www.krishisanskriti.org/Publication.html

Forecasting and Modelling of Marine and Inland Fish Production in India

Prema Borkar

*Gokhale Institute of Politics and Economics (Deemed to be University u/s of UGC Act, 1956)
BMCC Road, Deccan Gymkhana, Pune – 411004, Maharashtra. India
E-mail: premaborkar@rediffmail.com, prema.borkar@gipe.ac.in*

Abstract—*The paper describes an empirical study of forecasting and modeling of time series data of marine and inland production of fish in India. Yearly marine and inland production data for the period of 1973-1974 to 2015-2016 of India were analyzed by time-series methods. Autocorrelation and partial autocorrelation functions were calculated for the data. The Box Jenkins ARIMA methodology has been used for forecasting. The diagnostic checking has shown that ARIMA(0, 1,0) is appropriate for both marine and inland production of fish. The forecasts from 2016-2017 to 2024-2025 are calculated based on the selected model. The forecasting power of autoregressive integrated moving average model was used to forecast marine and inland production for nine leading years. This projection is important as it helps to inform good policies with respect to relative production, price structure as well as consumption of fish in the country.*

Keywords*: ACF - autocorrelation function, ARIMA - autoregressive integrated moving average, PACF - partial autocorrelation function, Fish, trends.*

1. INTRODUCTION

Indian fisheries is an important sector of food production, providing nutritional security to the food basket, contributing to the agricultural exports and engaging about fourteen million people in different activities. With diverse resources ranging from deep seas to lakes in the mountains and more than 10 per cent of the global biodiversity in terms of fish and shellfish species, the country has shown continuous and sustained increments in fish production since independence. Contributing about 6.3 per cent of the global fish production, the sector contributes to 1.1 per cent of the GDP and 5.15 per cent of the agricultural GDP. The total fish production of 10.07 million metric tons presently has nearly 65 per cent contributing from the inland sector and nearly the same from culture fisheries. Paradigm shifts in terms of increasing contributions from inland sector and further from aquaculture are significant over the years. With high growth rates, the different facets of marine fisheries, coastal aquaculture, inland fisheries, freshwater aquaculture, cold water fisheries to food, health, economy, exports, employment and tourism of the country.

Forecasts have traditionally been made using structural econometric models. Concentration have been given on the univariate time series models known as auto regressing integrated moving average (ARIMA) models, which are primarily due to world of Box and Jenkins (1970). These models have been extensively used in practice for forecasting economic time series, inventory and sales modeling (Brown, 1959; Holt *et al.,* 1960) and are generalization of the exponentially weighted moving average process. Several methods for identifying special cases of ARIMA models have been suggested by Box-Jenkins and others. Makridakis *et al.* (1982), and Meese and Geweke (1982) have discussed the methods of identifying univariate models. Among others Jenkins and Watts (1968), Yule (1926, 1927), Bartlett (1964), Quenouille(1949), Ljune and Bos (1978) and Pindyck and Tubinfeld (1981) have also emphasized the use of ARIMA models.

In this study, these models were applied to forecast the marine and inland production of fish in India. This would enable to predict expected marine and inland production for the years from 2015 onwards. Such an exercise would enable the policy makers to foresee ahead of time the future requirements for fish storage, import and/or export thereby enabling them to take appropriate measures in this regard.

2. MATERIAL AND METHODS

Time Series data was used for the Study. The data were obtained from the website of Indiastat from 1973-74 to 2015-16. Box and Jenkin (1976) linear time series model was applied on the data. This model is commonly known as Autoregressive Integrated Moving Average Model (ARIMA Model).

One of time series models which is popular and mostly used in ARIMA model. ARIMA (p, d, q) model is a mixture of Autoregressive (AR) model which shows that there is a relation of a value in the present (Z) and value in the past (Z_{t-k}), added by random value and Moving average (MA) model which shows that there is a relation between a value in the present (Z_t) and residuals in the past

(Z_{t-k} k = 1,2...)

with a non-stationary data pattern and d differencing order. The form of ARIMA (p,d,q) is:

$$\Phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t$$

where, p is AR model order, q is MA model order, d is differencing order and:

$$_p(B) = (1-\Phi_1 B - \Phi_2 B^2 - ...-\Phi_p B^p)$$

$$_p(B) = (1-\Phi_1 B - \Phi_2 B^2 - ...-\Phi_q B^q)$$

Generalization of ARIMA model for a seasonal patter data, which is written as:

$$ARIM(p,d,q)(P,D,Q)^2 =$$

$$\Phi_p(B)\Phi_p(B^2)(1-B)^d(1-B^s)(1-B^s)^D Z_t = \theta(B)^{\Theta}Q(B^s)a_t$$

where, s is seasonal period.

$$\Phi_p(B^s) = (1-\Phi_1 B - \Phi_2 B^2 - ...-\Phi_p B^p)$$

$$\theta_q(B) = (1-\theta_1 B - \theta_2 B^2 - ...-\theta_q B^q)$$

Model Identification: To determine whether the series isstationary or not we considered the graph of ACF. If a graph of ACF cuts of fairly quickly or dies down fairly quickly, then the time series value should be considered stationary. Model for non-seasonal series are called Autoregressive Integrated Moving average model, denoted by ARIMA (p,d,q). Here p indicates the order of the autoregressive part, d indicates the amount average of difference and q indicates the order of the moving average part. If the original series is stationary, d = 0 and the ARIMA models reduce to the ARMA models.

The difference linear operator ()), denoted by:

$$\Delta Y_t = Y_t - Y_{t-1} = Y_t - BY_t = (1-B)Y_t$$

The stationary series:

$$W_t = \Delta^d Y_t = (1-B)^d Y_t = \mu + \theta_q(B)\varepsilon_t$$

or$\Phi_p(B)W_t = \mu + \theta_q(B)\varepsilon_t$

Model estimation and checking: Estimate theparameters for a tentative model has been selected. The derived model must be checked for adequacy by considering the properties of the residuals whether the residuals from an ARIMA model is normal and randomly distribution. An overall check of the model adequacy is provided by Ljung-Box Q statistics. The test statistics Q is given in equation below:

$$Q_m = n(n+2)\sum_{k-1}^{n} \frac{r_k^2(e)}{n-k} \quad \chi^2_{m-r}$$

where, r_k (e) = the residual autocorrelation at lag K.

n = the number of residuals

m = the number of time lags includes in the test.

If the p-value associated with the Q Statistics is small (p-value <"), the model is considered inadequate. The analysts should consider a new or modified model and continue the analysis until a satisfactory model has been determined.

3. RESULTS AND DISCUSSION

The maximum marine production of fish was 3439 thousand tons in 2015-16 and was minimium in 1210 thousand tons in 1973-74. For inland production, the maximum production of fish in India was obtained in 2015-2016 year (6132 thousand tons) and minimum in 1973-1974 year (748 thousand tons).

Last 43 years data of marine and inland production of fish in India was used for modeling purpose. In model specification, we looked at the plots of auto-correlation function (ACF) for marine and inland fish production figures (Fig. 1 and 2). Also, partial auto correlation function (PACF) for marine and inland fish production (Fig. 3 and 4). Auto correlation function indicated the order of the auto regression compounds "q" of the model while the partial correlation function gave an indication for the parameter p. The ACF and PACF of the residuals (Fig. 5 and 6) also indicate 'good fit' of the model.

The time series plot (Fig. 7 and 8) of marine and inland fish production showed an increasing trend. ACF of both series showed non-stationary as ACF did not fall as quickly as the log K increased. To check the further stationary, second difference of the original series for marine and inland fish production was taken. The auto correlation formulation of second series and correlogram shows some more stationary than that of the first different. The corellogram of the auto correlation function of first difference series showed that the auto correlation function falls finally after lag 1 for marine and inland fish production, hence the respective values of the parameter "q" decided to be 0.

PAC function of the first differenced series of the cultivation area and production was used to determine parameter "p". Thus, we chose "p" to be 0 for marine and inland fish production respectively which gave good results consequently, the respective value of p,d,q were determined for ARIMA, that is ARIMA (0,1,0) for both marine and inland fish production.

Model estimation

ARIMA(0,1,0) modelwere estimated for both marine and inland fish production. The autocorrelation and partial autocorrelation coefficients of various orders of X_t are computed (Table 1 and 2). Goodness of fit of the model given in as the diagnostic check of the estimated model. The estimate of the parameters with corresponding standard error for ARIMA (0,1,0) is given in Table 3 and 4.

Residual analysis: The time series plot of the residualmarine and inland fish production data showed scattered trend, therefore, models were fitted properly by residual analysis.For normality test, Shapiro-wilk test was used. The test was significant and assumption of normality was accepted. Since the series fitted shows normality, the model is a good fit.

Forecasting: The last stage in the modeling process is forecasting. ARIMA models are developed basically to forecast the corresponding variable. There are two kinds of forecasts: sample period forecasts and post-sample period forecasts. The former are used to develop confidence in the model and the latter to generate genuine forecasts for use in planning and other purposes. The ARIMA model can be used to yield both these kinds of forecasts. The residuals calculated during the estimation process, are considered as the one step ahead forecast errors. The forecasts are obtained for the subsequent agriculture year from 2016-17 to 2024-25.

4. CONCLUSION

In our study, the developed model for marine and inland fish production was found to be ARIMA (0,1,0). The forecasts of marine and inland fish production, lower control limits (LCL) and upper control limits (UCL) are presented in Table 5. The validity of the forecasted values can be checked when the data for the lead periods become available. ARIMA model being stochastic in nature, it could be successfully used for modeling as well as forecasting the marine and inland fish production of India. The model demonstrated a good performance in terms of explaining variability and predicting power. The supply projection of an agricultural commodity especially fish plays a vital role in the adjustment of supply to demand. These projections help the government to make policies with regards to relative price structure, production and consumption patterns and also, to establish relationship with other countries of the world.

REFERENCES

[1] Bartlett, M.S., 1964. "On The Theoretical Specification of Sampling Properties of Autocorrelated Time Series".*J. Roy. Stat. Soc.,* B 8: 27–41.

[2] Box, G.E.P., and G. M. Jenkins. (1970). Time series analysis: forecasting and Control, Holden Day, San Francisco, CA.

[3] Box, G.E.P. and G.M. Jenkins, 1976.Time Series Analysis: Forecasting and Control. Rev. Ed. San Francisco. Holden-Day.

[4] Brockwell, P.J., and Davis, R. A. (1996). Introduction to time series and Forecasting, Springer.

[5] Brown, R.G., 1959. Statistical Forecasting For Inventory Control. New York, McGraw-Hill,

[6] Iqbal, N., K. Bakhsh, A. Maqbool and A. S. Ahmad. (2005). Use of the ARIMA Model for forecasting Wheat area and production in Pakistan". Journal of Agriculture & Social Sciences., 2: 120-122.

[7] Jenkins, G. M. and D.G. Watts, 1968.Spectral Analysis and its Application, Day, San Francisco, California, USA.

[8] Ljunge, G.M. and G.E.P. Box, 1978. "On a Measure of Lack of Fit in Time Series Models".Biometrika, 65: 67–72.

[9] Kendall, M. G., and A. Stuart. (1966). The advanced theory of statistics. Vol. 3.

[10] Design and Analysis and Time-Series. Charles Griffin & Co. Ltd., London, United Kingdom.

[11] Holt, C.C., F. Modigliani, J.F. Muth and H.A. Simon, 1960.Planning, Production, Inventores, and Work Force. Prentice Hall,Englewood Cliffs, NJ, USA.

[12] Muhammad, F., M. S. Javed and M. Bashir. (1992). Forecasting Sugarcane Production in Pakistan using ARIMA Models, Pak. J. Agri. Sci., 9(1): 31-36.

[13] Makridakis, S, A. Anderson, R. Filds, M. Hibon, R. lewandowski, J. Newton, E.

[14] Parzen and R. Winkler, 1982. The Accuracy of Extrapolation (time series) methods: Results of a Forecasting Competition, Journal of Forecasting Competition.J.Forecasting, 1: 111–53.

[15] Meese, R. and J. Geweke, 1982.A Comparison of Autoregressive Univariate Forecasting Procedures for Macroeconomic Time Series.

[16] Unpublished Manuscript, University of California,Berkeley, CA, USA.

[17] Prindycke, R.S. and D.L. Rubinfeld, 1981.Econometric Models and Economic Forecasts, 2nd Ed.New York, McGraw-Hill.

[18] Quenouille, M.H., 1949. "Approximate Tests of Correlation in Time- Series.J. Roy. Stat. Soc., B11: 68–84.

[19] Saeed, N., A. Saeed, M. Zakria, T. M. Bajwa. (2000). Forecasting of Wheat Production in Pakistan using ARIMA models, International Journal of Agriculture and Biology., 4: 352-353.

[20] Yule, G.U., 1926. "Why Do We Sometimes Get Nonsence-corrleations Between Times Series. A study in Sampling and the Nature of Series".J. Roy. Stat. Soc., 89: 1–69.

[21] Yule, G.U., 1927. "On a method of Investigation Periodicities in Disturbed Series, With Specia; Reference To Wolfer's Sunspot Number". Phill. Trans., A 226: 267–98.

Table 1: Autocorrelations and partial autocorrelations of Marine Fish Production

Lag	Autocorrelation	Std.error	Lag	Partial Autocorrelation	Std.error
1	0.922	0.151	1	0.922	0.156
2	0.848	0.149	2	-0.015	0.156
3	0.776	0.147	3	-0.023	0.156
4	0.709	0.145	4	-0.006	0.156
5	0.644	0.143	5	-0.026	0.156
6	0.580	0.141	6	-0.032	0.156
7	0.512	0.139	7	-0.066	0.156
8	0.442	0.137	8	-0.062	0.156
9	0.367	0.135	9	-0.075	0.156

10	0.292	0.133	10	-0.064	0.156
11	0.215	0.130	11	-0.072	0.156
12	0.148	0.128	12	-0.002	0.156
13	0.084	0.126	13	-0.037	0.156
14	0.013	0.124	14	-0.101	0.156
15	-0.067	0.121	15	-0.133	0.156
16	-0.129	0.119	16	-0.037	0.156

Table 2: Autocorrelations and partial autocorrelations of Inland Fish Production

Lag	Autocorrelation	Std.error	Lag	Partial Autocorrelation	Std.error
1	0.934	0.151	1	0.934	0.156
2	0.868	0.149	2	-0.032	0.156
3	0.802	0.147	3	-0.035	0.156
4	0.734	0.145	4	-0.049	0.156
5	0.663	0.143	5	-0.072	0.156
6	0.589	0.141	6	-0.056	0.156
7	0.518	0.139	7	-0.029	0.156
8	0.450	0.137	8	-0.023	0.156
9	0.384	0.135	9	-0.024	0.156
10	0.315	0.133	10	-0.080	0.156
11	0.243	0.130	11	-0.068	0.156
12	0.178	0.128	12	-0.009	0.156
13	0.113	0.126	13	-0.063	0.156
14	0.052	0.124	14	-0.016	0.156
15	-0.009	0.121	15	-0.063	0.156
16	-0.068	0.119	16	-0.046	0.156

Table 3: Model fit of the fitted ARIMA model for Marine Fish Production

Fit Statistic	Mean	Minimum	Maximum	Percentile						
				5	10	25	50	75	90	95
Stationary R-squared	1.621	1.621	1.621	1.621	1.621	1.621	1.621	1.621	1.621	1.621
R-squared	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970
RMSE	121.36	121.36	121.36	121.36	121.36	121.36	121.36	121.36	121.36	121.36
MAPE	3.713	3.713	3.713	3.713	3.713	3.713	3.713	3.713	3.713	3.713
MaxAPE	17.682	17.682	17.682	17.682	17.682	17.682	17.682	17.682	17.682	17.682
MAE	86.975	86.975	86.975	86.975	86.975	86.975	86.975	86.975	86.975	86.975
MaxAE	402.28	402.28	402.28	402.28	402.28	402.28	402.28	402.28	402.28	402.28
Normalized BIC	9.690	9.690	9.690	9.690	9.690	9.690	9.690	9.690	9.690	9.690

Table 4: Model fit of the fitted ARIMA model for Inland Fish Production

Fit Statistic	Mean	Minimum	Maximum	Percentile						
				5	10	25	50	75	90	95
Stationary R-squared	2.874	2.874	2.874	2.874	2.874	2.874	2.874	2.874	2.874	2.874
R-squared	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997
RMSE	82.526	82.526	82.526	82.526	82.526	82.526	82.526	82.526	82.526	82.526
MAPE	2.782	2.782	2.782	2.782	2.782	2.782	2.782	2.782	2.782	2.782
MaxAPE	12.089	12.089	12.089	12.089	12.089	12.089	12.089	12.089	12.089	12.089
MAE	61.891	61.891	61.891	61.891	61.891	61.891	61.891	61.891	61.891	61.891
MaxAE	200.93	200.93	200.93	200.93	200.93	200.93	200.93	200.93	200.93	200.93
Normalized BIC	8.918	8.918	8.918	8.918	8.918	8.918	8.918	8.918	8.918	8.918

Table 5: Forecasts for Marine and Inland Fish Production (2015-16 to 2024-2025) (In '000' tones)

Years	Marine Production	Inland Production
2016-2017	3606+425	9694+309
2017-2018	3662+490	7002+324
2018-2019	3718+548	7009+359
2019-2020	3773+601	7125+367
2020-2021	3829+649	7136+396
2021-2022	3885+694	7156+436
2022-2023	3941+735	7159+453
2023-2024	3996+776	7162+483
2024-2025	4052+815	7168+496

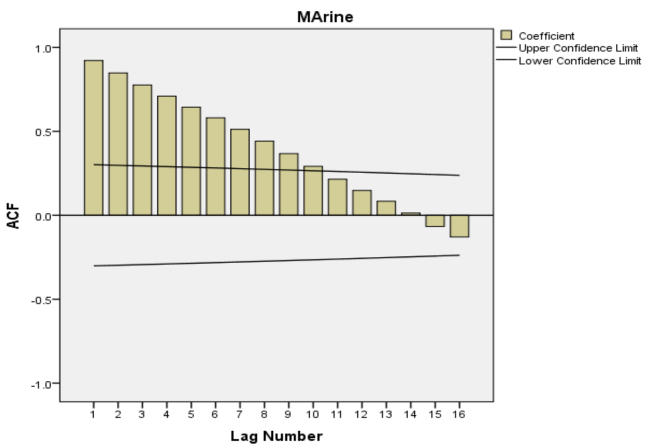

Fig. 1: Autocorrelation function of Marine Fish Production

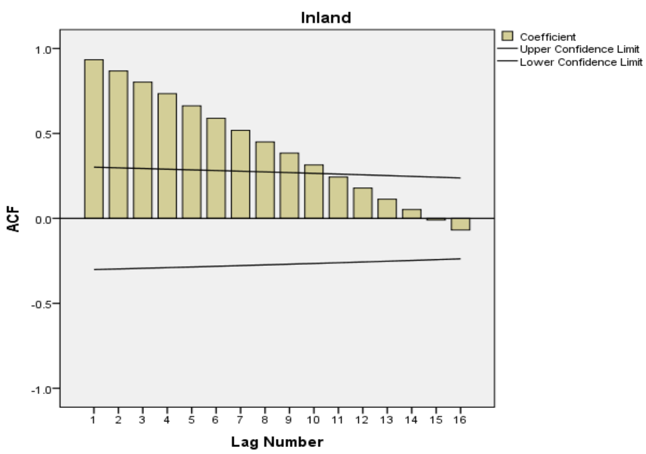

Fig. 2: Autocorrelation function of Inland Fish Production

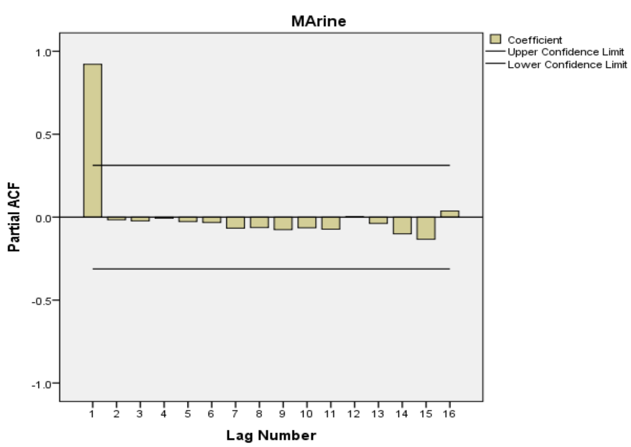

Fig. 3: Partial autocorrelation function of Marine Fish Production

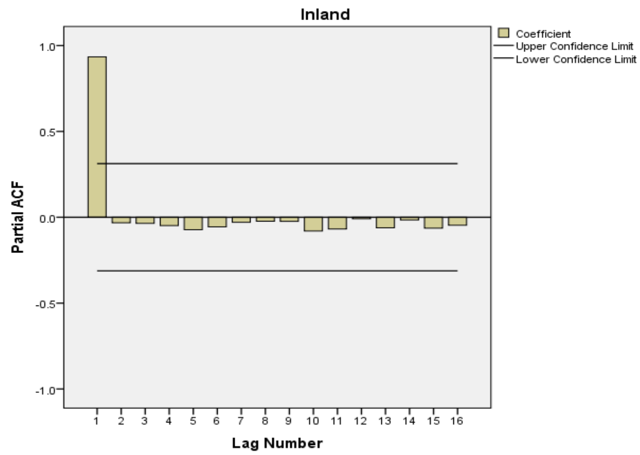

Fig. 4: Partial autocorrelation function of Inland Fish Production

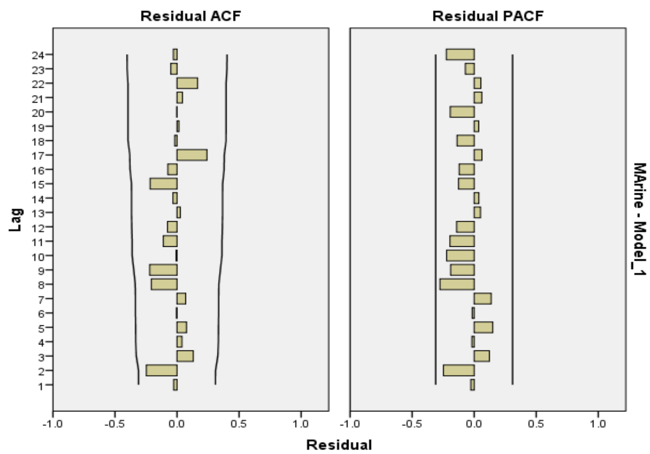

Fig. 5: ACF and PACF of residuals of fitted ARIMA model of Marine Fish Production

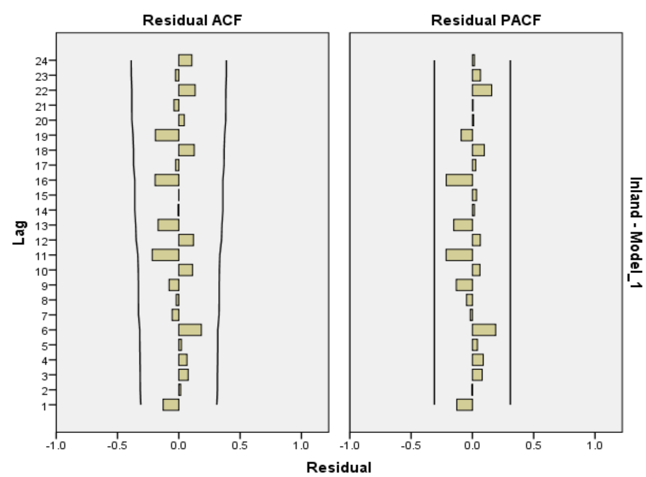

Fig. 6: ACF and PACF of residuals of fitted ARIMA model of Marine Fish Production

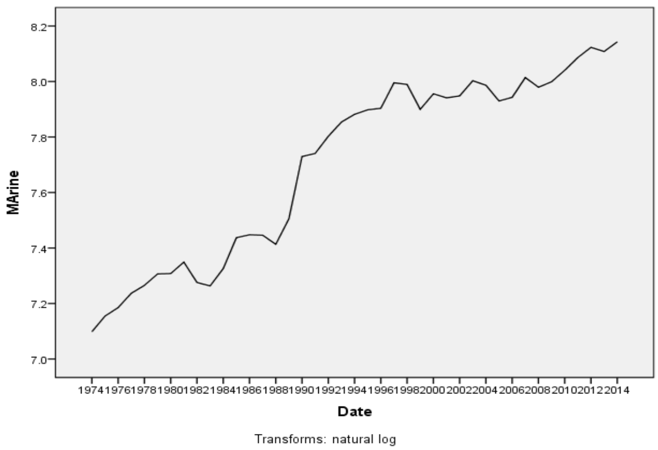

Fig. 7: Time series plot for Marine Fish Production

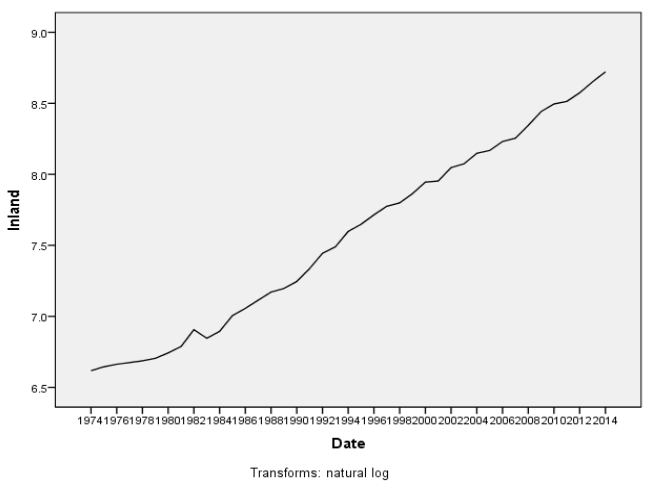

Fig. 8: Time series plot for Inland Fish Production