JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DAIRY PRODUCTION: PREDICTING ANALYSIS USING AUTO REGRESSIVE INTEGRATED MOVING AVERAGE [ARIMA] MODEL, TAMIL NADU

[1]D.Senthilkumar and [2]P.Sabarish
[1]Assosiate Professor Department of Statistics, PSG College of Arts & Science, Coimbatore, India.
[2]Research Scholar Department of Statistics, PSG College of Arts & Science, Coimbatore, India.

ABSTRACT

In every industry, need various predictable solutions for their future planning and developments also they spend huge amount of cost in this division. There are several statistical prediction methods are federalize to analyse their data's in industries. In this article describes Future predicting of Dairy Production using Auto Regressive Integrated Moving Average [ARIMA] Model, Tamil Nadu.

Key words: ACF, ARIMA, Predicting, Moving Average and PACF.

INTRODUCTION

In Indian Market, Tamil Nadu plays a vital role in Dairy production. In Tamil Nadu more than 2,50,000 employees are working Dairy production and related works. Usage of Dairy and its products are placing a vital role in our food in daily. Statistical tools for Predicting purpose started using smooth exponential methods in 1950s. These methods were modified depending upon the trend followed in the data sets, based upon the evaluation purpose. From simple additive to multiplicative effects and then automated functions were used to evaluate the complexity in data for Predicting purpose. In this article, using ARIMA model for prediction purpose, there are several prediction analysis are available but ARIMA is the widely used model.

Umamageswari M et al (2017) have study Economics of dairy production in Tamil Nadu – A comparative study. Box, G. E. P et al (2008) have study Time Series Analysis: Predicting and Control. Box, G. E. P. et al (1975), have developed Intervention analysis with applications to economic and environmental problems. Marek, L (2000), has study Transfer function models. Pankratz, A (1991) has study Predicting with Dynamic Regression Models. Rublíková and E. – Marek, L (2001), have study Linear transfer function model for outflow rates. L. Yu, L. Zhou, et al (2014), " Application of a new hybrid model with seasonal auto regressive integrated moving average (ARIMA) and nonlinear auto-regressive neural network (NARNN) in forecasting incidence cases of HFMD. T. VanCalster et al, (2017) have study Prof ARIMA A profit-driven order identification algorithm form ARIMA models in sales forecasting. P. Ramos et al,(2015) have developed Performance of state space and ARIMA models for consumer retail sales forecasting.

METHODOLOGY

BOX-JENKINS (BJ) MODEL

These models were used as a method of extracting predictable movements in a time series manner. This model has 3 components – autoregressive component, moving average, and white noise. The time-series data to be estimated using the BJ model

must not contain seasonality and must be stationary. The Augmented Dickey-Fuller and the Phillips Perron Unit Root Tests will be used to check if the time-series data is stationary or –non stationary.

METHOD OF MOVING AVERAGE

Method of Moving average models was proposed by Slutsky (1927) and Wold (1938). The series of moving average can be written as

$$Y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \theta_3\varepsilon_{t-3} - ... - \theta_q\varepsilon_{t-q} \quad (1)$$

This is called as method of moving average in the order q (MA q) here $t\ Y$ is the series which contains original data and t is the error term of the series.

METHOD OF AUTO REGRESSIVE PROCESS

The Auto Regressive process was first proposed by Yule (1926) this regressive process is specially satisfies the following equation

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + ... + \phi_p Y_{t-p} \quad (2)$$

The present value Y_t is called as linear combination for p(AR P) also t is the assumption of independent of

$$Y_{t-1}, Y_{t-2}, Y_{t-3} ... Y_{t-q} \quad (3)$$

THE AUTO REGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

The ARIMA methodology was proposed by the author Box and Jenkins in the year 1976. This ARIMA model is also called as Box- Jenkins model. This model is based on the error term of the time series. To get conclude time series the data should be stationary which means, the mean, variance and covariance are constants over the time period. For this we can write the ARIMA model equation as follows. With a help of back shift operator (lag) we write the above equation as follows.

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + ...\phi_p Y_{t-p} + \phi_1 Y_{t-s} \\ &+ \phi_2 Y_{t-2s} + ... + \phi_p Y_{t-ps} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \\ &... - \theta_q a_{t-q} - \Theta_1 a_{t-s} - \Theta_2 a_{t-2s} - ...\Theta_q a_{t-qs} \end{aligned} \quad (4)$$

With a help of back shift operator (lag) we write the above equation as follows:

$$\phi_p(B)\phi_p(B^s) z_t = \theta_q(B)\Theta_q(B^s) a_t \quad (5)$$

Here;

$z_t = (1-B)^d (1-B^s)^D \ln(Y_t)$

$\phi_p(B) - Non\ seasonal\ operator\ of$

$Autoregressive\ process\ AR(p)$

RESULTS AND DISCUSSIONS

GRAPHICAL ANALYSIS

The first method used in the study is to determine whether the time-series data is stationary or not by using the Graphical analysis method. It is shown on the basis of the graph, if the graph comes along in a certain wave, then it is non-stationary. As shown in the below Figure 1 the upward motion of the wave of the time – series data, tells that the data is non-

stationary. A sharp increase in the production of dairy in the year 2003-04. The following Graph (figure1) shows that there is a gradual increase in the motion of Dairy production in Tamil Nadu over a period of 26 years.

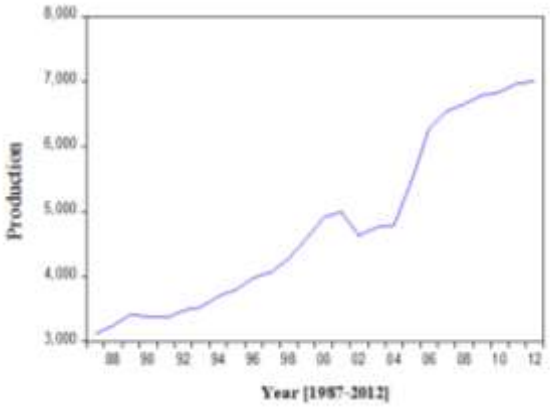


Figure1. Graph of Dairy Production in Tamilnadu (1987-2012)

The following graph (figure 2) shows that there is a gradual decrease of in the dairy production from the starting itself. The first difference of the variable production of Dairy with (p-value= 0.0000) alternatively the stationary of the production is tested by using unit root tests. We hereby, conduct three other different types to check the stationary of the data series.

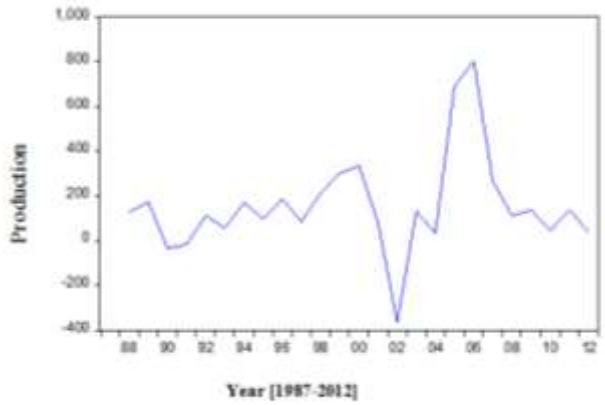


Figure2. Differenced Dairy Production in Tamilnadu (1987-2012)

The stationary of the production is tested by using unit root tests. We hereby, conduct three other different types to check the stationary of the data series. Augmented Dickey – Fuller (ADF) Unit Root Test Using the Augmented Dickey – Fuller (ADF) Unit Root Test, we test if the motion and intercept of the time series data is stationary.

Table.1

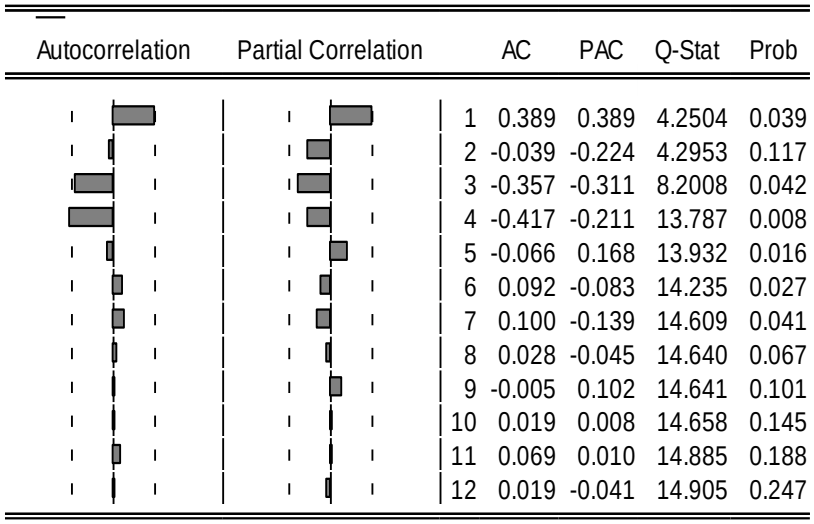
Agumented Dickey – filler unit root test for first order difference

Variable	Coefficient	Std Error	t- statistics	Probability
AR(1)	0.997984	0.005892	169.3864	0.0000
MA(1)	0.525504	0.107736	4.877683	0.0001
SIGMA SQ	50106.01	9712.898	5.158708	0.0000
R-Squared	0.971267	Mean dependent var		4785.915
Adjusted R- Squared	0.968768	SD dependent var		1346.699
SE of Regression	237.9949	Akaike info criterion		14.14749
Sum Squared residual	1302756	Schwarz criterion		14.29266
Log likelihood	-180.9174	Hannan-qunin criter		14.18930
Durbin-waston stat	1.636324	-		-

CORRELOGRAM

Now, we generate the Correlogram of the time series data to determine whether there is an Autocorrelation Function (ACF) or partial Autocorrelation Function (PACF).

Table.2



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.389	0.389	4.2504	0.039
		2	-0.039	-0.224	4.2953	0.117
		3	-0.357	-0.311	8.2008	0.042
		4	-0.417	-0.211	13.787	0.008
		5	-0.066	0.168	13.932	0.016
		6	0.092	-0.083	14.235	0.027
		7	0.100	-0.139	14.609	0.041
		8	0.028	-0.045	14.640	0.067
		9	-0.005	0.102	14.641	0.101
		10	0.019	0.008	14.658	0.145
		11	0.069	0.010	14.885	0.188
		12	0.019	-0.041	14.905	0.247

CORRELOGRAM FOR GIVEN DATA

The time series data is non stationary because there is a continuous decline in the auto correlation coefficient and these values are significant. we will also get the correlogram of the 1st order difference to see if it becomes stationary

Table.3

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.895	0.895	23.324	0.000
		2	0.773	-0.141	41.449	0.000
		3	0.652	-0.056	54.915	0.000
		4	0.528	-0.091	64.153	0.000
		5	0.408	-0.061	69.920	0.000
		6	0.285	-0.099	72.886	0.000
		7	0.163	-0.094	73.902	0.000
		8	0.073	0.069	74.120	0.000
		9	0.013	0.043	74.127	0.000
		10	-0.047	-0.082	74.226	0.000
		11	-0.104	-0.071	74.756	0.000
		12	-0.182	-0.191	76.487	0.000

Correlogram for 1st Difference

DIAGNOSTIC TEST

Based on the diagnostic test we decide the range of ARIMA model as (1,1,1). The table of the diagnostic test is given below.

Table.4

S.NO	ARIMA	R^2	AIC	BIC
1	**1,1,1**	**0.97126**	**14.14749**	**14.29266**
2	1,1,2	0.96447	14.35991	14.50508
3	2,1,1	0.9598987	14.47765	14.62282
4	2,1,2	0.901634	15.58847	15.73363
5	3,1,1	0.900617	15.66768	15.81284
6	3,1,2	0.876939	15.99833	16.14349
7	3,1,3	0.753886	16.52826	16.67342
8	1,1,3	0.959215	14.4726	14.61782
9	2,1,3	0.866405	15.7794	15.92457
10	4,1,1	0.886198	15.76848	15.91364
11	1,1,4	0.959092	14.47549	14.62066
12	1,1,5	0.961701	14.41988	14.56504

PREDICTION

We have generated a Predict for the next thirteen years (2012 to 2025). This Predict is based on the past time serried data that were used in the identification of the ARIMA Model. The time series data has been used to generate the graph and the Predict is shown in the Figure.

Table.5

YEAR	PRODUCTION	YEAR	PRODUCTION	YEAR	PRODUCTION
1987	3109	2000	4907	2013	7106
1988	3238	2001	4988	2014	7249
1989	3410	2002	4622	2015	7400
1990	3375	2003	4752	2016	7553
1991	3357	2004	4784	2017	7705
1992	3468	2005	5474	2018	7858
1993	3524	2006	6276	2019	8010
1994	3694	2007	6540	2020	8163
1995	3791	2008	6651	2021	8315
1996	3976	2009	6787	2022	7106
1997	4061	2010	6831	2023	7249
1998	4273	2011	6968	2024	7400
1999	4574	2012	7005	2025	7553

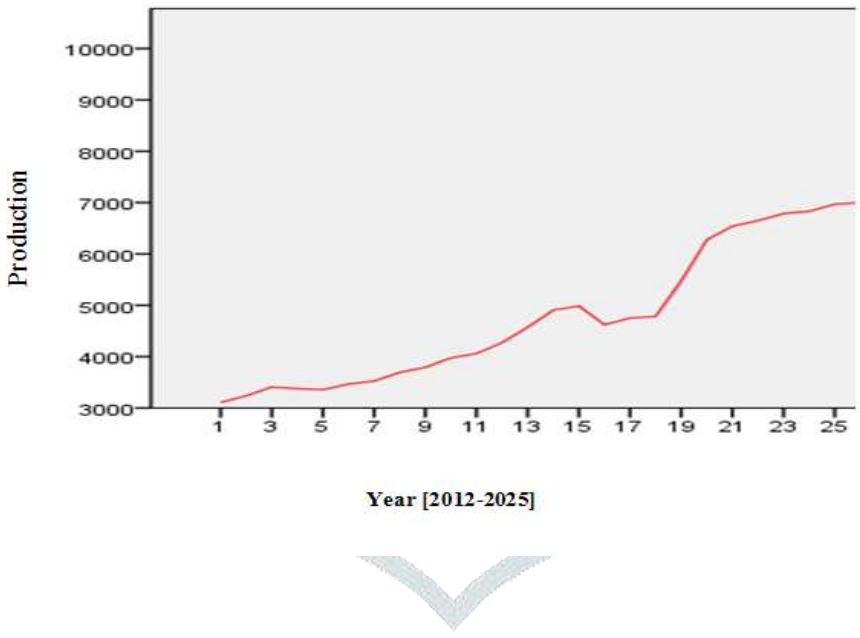


Figure.3 Predicted graph of Milk Production of the year (2012-2025)

CONCLUSION

In this article, we constructed a model called ARIMA (1,1,1). Rooted in this model we conclude that Predicting value of Dairy production 2012 to 2025. The figure-3 is clearly explains the future production of dairy also Table-5 explains that the amount of change in future production. This indicates that there will be a gradual increase production of dairy we should take the necessary steps in order to improve the production.

REFERENCE

1. Box, G. E. P et al (2008), “TimeSeries Analysis: Predicting and Control” John Wiley & Sons Inc., New York, 2008.

2. Box, G. E. P. et al (1975), "Intervention analysis with applications to economic and environmental problems" Journal of the American Statistical Association. vol.70, no. 349, pp. 70-79.

3. Marek, L (2000), "Transfer function models", Acta Oeconomica Pragnesia. vol. 8, no.3, pp. 83-94.

4. Pankratz, A (1991) "Predicting with Dynamic Regression Models". Wiley-Interscience, 1991.

5. Rublíková and E. – Marek, L (2001), "Linear transfer function model for outflow rates" Ekonomické rozhľady. vol,30,no. 4, pp.457 – 466.

6. L. Yu, L. Zhou, et al (2014), " Application of a new hybrid model with seasonal auto regressive integrated moving average (ARIMA) and nonlinear auto-regressive neural network (NARNN) in forecasting incidence cases of HFMD in Shenzhen,China," PLoSONE,vol.9,no.6.

7. T. VanCalster et al, (2017) " Prof ARIMA A profit-driven order identification algorithm form ARIMA models in sales forecasting," Applied Soft Computing, vol. 60, pp. 775– 785.

8. Umamageswari M et al (2017), " Economics of dairy production in Tamil Nadu – A comparative study". Indian J Dairy Sci, No 70(2), pp 221-227.

9. P. Ramos et al,(2015) " Performance of state space and ARIMA models for consumer retail sales forecasting," Robotics and Computer-Integrated Manufacturing, vol. 34, pp. 151– 163.