Urbanization on Agriculture: India

Chanchal Pramanik*
Tata Consultancy Services, Vadodara, India – cpramanik@gmail.com

Suhrit K Dey
Eastern Illinois University, USA – skdey@eiu.edu

Abstract

Present study on India is concerned with effect of urbanization, population growth and total cropped area on total food grain production based on 60 years of data (1950-61 to 2008-09). Multiple Linear Regression has been explored to explain the relationship between total food grain production and urbanized area, population growth and total cropped area. Autoregressive Integrated Moving Average (ARIMA) models have been used to develop forecasting models and to forecast the variables involved in the study up to 2020. Later the multiple linear regression equation has been used to forecast the possible total food grain production based on the forecasted figures of the urbanized area, population growth and total cropped area. Separate forecast obtained for the total food grain production from the ARIMA model based on the trend is the expected total food grain production. Differences between possible and expected total food grain productions are obtained for the years 20010-11 to 2020-21. It has been observed that differences have a steep downward linear trend. This is an alert for Indian agriculture, in distant future. Similar results have been found for India with the help of Markov chain modeling.

Keywords: Multiple Linear Regression, ARIMA, Markov Chain Modelling.

1. Introduction

Agriculture is the backbone of Indian economy, which is providing livelihood for about 65 to 70 per cent of total population and employs about 52 per cent of country's workforce and presently contributing nearly 14 per cent to GDP (2012-13). India is the second most populated country in the world after China. Presently the study on India is concerned with effect of urbanization, population growth and total cropped area on total food grain production based on 60 years of data (1950-51 to 2009-10), as per data availability. The data of the above mentioned variables is collected from www.indiastat.com. Here Multiple Linear Regression, Auto Regressive Integrated Moving Average (ARIMA) and Markov Chain modeling techniques are used. Similar studies were carried out on Andhra Pradesh state of India which shows that the urbanization and population growth have adverse effect on the agriculture of the state. It is the aim of this research investigation to see if the condition of overall India is similar to that of Andhra Pradesh.

2. Methodology

The data have been collected online from www.indiastat.com from 1950-51 to 2009-10 on Total Food Grain production (TFG), Total Cropped Area (TCA), Urbanized Area (UA) and Population. Then on the basis of 60 years of data a Multiple Linear Regression (MLR) has been fitted taking TFG as dependent variable and TCA, UA and Population as independent variables. The statistical significance of the model is determined based on Coefficient of Determination (R^2). The significance of the estimated independent variables is tested based on the respective p-values, taking 5 percent level of significance as standard.

Next after fitting the MLR equation ARIMA models are fitted to forecast all the variables involved in the study. Based on the respective ARIMA models TFG, TCA, UA and Population are forecasted for next 11 years (2010-11 to 2020-21).

Now based on the TCA, UA and Population forecasted values TFG values are obtained based on the previous fitted MLR equation for 2010-11 to 2020-21. Here two sets of TFG values have been obtained, one from the time series analysis of ARIMA model and other from fitted MLR equation where independent variables are estimated based on ARIMA model. Differences of two sets of TFGs are calculated and fitted with a suitable trend equation.

The forecasted TFGs based on MLR up to 2020-21 consider the trend of TCA, UA and Population which signifies the possible TFGs. The forecasted TFGs based on ARIMA technique up to 2020-21 signify the expected TFGs considering the present trend of TFG. The differences between two sets of TFGs signify the shortage of Total Food Grain production due to the present trend of UA, TCA and Population.

3. Statistical Methodology

Multiple Linear Regression analysis (MLR) has been explored to establish relation between the TFG and TCA, UA and Population. Ordinary Least Square (OLS) technique has been used to estimate the coefficients of independent variables. Coefficient of Determination (R^2) has been calculated to test the significance of the MLR equation. To test the significance of the estimates of independent variables respective p-values has been compared with 5 percent level of significance. The general form of the MLR equation is as follows:

$$Y = b_0 + b_1X_1 + \cdots + b_kX_k \qquad (1)$$

where,

Y = Dependent variable

X_i = Independent variable (i =1, 2,…, k)

b_0 = Constant

b_i = Regression coefficient of X_i

Auto Regressive (AR) Model

$$Y = a + b_1Y_{t-1} + \cdots + b_pY_{t-p} + U_t \qquad (2)$$

where,

Y_t = The values of the variable for forecasting at time 't'

a = Constant

b_i = i^{th} regression Coefficient, (i = 1,2,…,p)

U_t = Random error

This above model is called AR (p) model or AR (p) process.

Moving Average (MA) Model:

Sometimes residuals with different lags may exhibit relationships with the dependent variable, as follows:

$$Y_t = \mu + \varphi U_t + \varphi_1U_{t-1} + \varphi_2U_{t-2} + \cdots + \varphi_qU_{t-q} \qquad (3)$$

where,

Y_t = The values of the variable for forecasting at time 't'

μ = Constant

ϕ_l = Partial regression coefficient, (l = 0, 1, 2,…,q)

This model is known as MA (q) model.

Auto Regressive Moving Average (ARMA) Model

In this model Y_t depends on AR as well as MA variables and can be specified as:

$$Z_t - b_1Z_{t-1} - \cdots - b_pZ_{t-p} = U_t - \varphi_1U_{t-1} - \cdots - \varphi_qU_{t-q} \qquad (4)$$

where,

$Z_t = Y_t \quad \bar{Y}$ (deviation of Y_t from mean $\bar{Y}$)

The above model is ARMA(p,q) model.

Markov Chain Model

The Markov chain model can described as follows: Let us assume we have a set of states, $S = (s_1, s_2, \ldots, s_r)$. The process starts in one of these states and moves successively from one state to another. Each move is called a step. If the chain is currently in state s_i, then it moves to state s_j at the next step with a probability denoted by p_{ij}, and this probability does not depend upon which states the chain was in before the current state.
The probabilities p_{ij} are called transition probabilities. The process can remain in the state it is in, and this occurs with probability p_{ii}. An initial probability distribution, defined on S, specifies the starting state. Usually this is done by specifying a particular state as the starting state. The details of the Markov chain model will be clearer from the following example.
Let us consider two disjoint sets S_1 and S_2.
Let p_{ij} = probability that an element S_i will move into S_j , where i = 1, 2 and j = 1, 2 for each i.
Let us assume $U = S_1 \cup S_2$ = The universe.
Let X_i^0 =value of the X_i at time = 0.
= $P(S_i)|$at t = 0, i = 1, 2.
Then $X_1^0 + X_2^0 = P(S_1)_0 + P(S_2)_0$
$= P(U) = 1$ (5)

Land Acquisition

Let S_1 = Agricultural Land, S_2 = Urban Land.
Let at t = 0, 95% land is farmland. So, only 5% is urbanized.
Hence $X_1^0 = 0.95$ and $X_2^0 = 0.05$.
Let us consider the best possible scenario. Let 98% of S_1 is retained as S_1; so, only 2% of S_1 is urbanized each year. However, once it is urbanized, it is not converted into farmland.
Thus $p_{11} = 0.98$, $p_{12} = 0.02$, $p_{21} = 0$ and $p_{22} = 1$. If this pattern is maintained, then only after 10 years, 77.6% of land will be available to feed the world population.

4. Results and Discussions

A Multiple Linear Regression (MLR) equation has been fitted by taking Total Food Grain (TFG) production as dependent variable and Total Cropped Area (TCA), Urbanized Area (UA) and Population as independent variables based on 60 years of data (1950-51 to 2009-10). The MLR equation is as follows:

*TFG = -159119909.30 + 1.18*TCA -0.63*UA + 0.20*Population* (6)

Table 1: *Model Summary of the MLR*

Parameter	*Estimate*	*Standard Error*	*T-Statistic*	*P-Value*
CONSTANT	159119909.30	37445949.11	-4.25	<0.0001
UA	-0.63	1.77	-0.35	0.7245
TCA	1.18	0.35	3.33	0.0015
Population	0.20	0.03	8.03	<0.0001

Table 2: *ANOVA of the MLR*

Source	*Df*	Sum of Squares	*Mean Square*	*F-Ratio*	*P-Value*
Model	3	$1.70*10^{17}$	$5.68*10^{16}$	696.929	<0.0001
Residual	56	$4.56*10^{15}$	$8.15*10^{13}$		

Now based on the forecasted values of TCA, UA and Population TFGs are estimated on the basis of the basis of the equation (1) for 12 years (2009-10 to 2020-21).

Table 5: *Comparative study of two sets of forecasted TFGs based on MLR & ARIMA*

Year	**TFG** *(in tonnes)* (MLR method)- (I)	**TFG** *(in tonnes)* (ARIMA method) -(II)	**D** (I)-(II) *(in tonnes)*
2010-11	237573030.6	221490030.2	16083000.34
2011-12	240696915.6	224553174.7	16143740.95
2012-13	243767190.7	227552803.2	16214387.51
2013-14	246787024.1	230579631.2	16207392.95
2014-15	249759045.9	233594811.5	16164234.4
2015-16	252685903.6	236614979.7	16070923.91
2016-17	255570036.9	239633011.9	15937024.96
2017-18	258413784.1	242651958.8	15761825.24
2018-19	261219346.9	245670514.1	15548832.78
2019-20	263988814.8	248689237	15299577.71
2020-21	266724163.8	251707888.2	15016275.61

The difference (D) values exhibits a linear trend which is as follows:

$$D=16503853.4 - 107502.6 \times T \qquad (7)$$

where, T = time

Table 6: *D- trend model summary*

Parameter	*Estimate*	*Stnd. Error*	*t*	*P-value*
Constant	16503853.4	132988.9	124.0995	<0.0001
Slope	-107503.6	19608.1	-5.48255	0.0004

Thus it is quite evident that the differences of the possible and expected TFGs are in decreasing linear trend. Hence if the same trend of TCA, UA and Population exists there will be problem for food security in distant future. This study can be verified further with the help of Markov-chain model.

Markov-Chain Model Approach for estimating future Agricultural Area

It has been observed that agricultural area has a growth rate of -0.11 per cent per year whereas urbanized area has a growth rate of 0.99 percent per year. The agricultural area is certainly contributing to urbanized area. Uncultured or barren land is also seen to be has a decreasing trend by converting to urbanized area. At present the share of agricultural area from the human cultured area, including both agricultural and urbanized area, is 84.62 percent. So rest of the human cultured land i.e., 15.38 percent is urbanized area. Again, there is limited possibility of converting uncultured area to agricultural area as that may not be suitable for agriculture to prosper.

Let S_1 = Agricultural Land, S_2 = Urban Land. Let at t = 0, 84.62 percent land is farmland. So, only 15.38 per cent is urbanized. Hence X_1^0= 0.8462 and X_2^0= 0.1538.

Let 99.89 percent of S_1 is retained as S_1; so, only 0.11 percent of S_1 is urbanized each year. However, once it is urbanized, it cannot be converted into farmland.

Thus p_{11} = 0.9989, p_{12} = 0.0011, p_{21} = 0 and p_{22} = 1. If this pattern is maintained, then after 10 years, 83.60% of land will be available to feed the world population. So the issue of conversion of agricultural land to urbanized land may be a serious issue, in distant future, for Indian agriculture with rising population unless dampened by new agricultural technologies.

5. Conclusions

The detailed study on India shows that the urbanization has a long term effect on Indian agriculture. Probably the Green Revolution and innovative agricultural technologies have dampened the effect of urbanization and population growth on agriculture. But in the earlier studies on Andhra Pradesh state of India it was found that the urbanization and population growth has serious effect on the agriculture of the state (which was presented in ICAS-V). Andhra Pradesh besides ranking third in total agricultural production also contributes significantly on industry sector in particular to software industry. This may be the reason why urbanization and population have adverse effect on agriculture in Andhra. Similar trends are being started to be followed in many other states of India following Foreign Deposit Investment scheme granted by Government of India. Hence, the rate of urbanization will increase acquiring more agricultural land, creating threat on Food Security. Hence it is the sole responsibility of the policy maker to keep a balance between Agricultural and Industrial Economy.

References

Bhattacharya A. (2006). A goal Programming Approach for Developing Pre-Harvest Approach in Crop Yield. The Journal of Operational Research Society. 57(8), 1014-1017.

Dey S. K., C Pramanik and C Dey. Mathematical Modelling of Effects of Urbanization and Population Growth on Agricultural Economics. Journal of Intellectual Economics. 2(8), 7-20.

Dobre I and Alexandru A A M (2008). Modeling unemployment rate using Box-Jenkins Procedure. Journal of Applied Quantitative Methods, 3(2), 156-166.

Ferguson C. E. (1958). A Statistical Study on Urbanization. Social Forces, 37(1), 19-26.

Gopinath C. K. (2000). An indirect method for forecasting the annual food production in India. Current Science. 79(1), 23-24.

Iqbal N, Bakhsh K, Maqbool A and Ahmad A S (2005). Use of ARIMA model for forecasting wheat area and production in Pakistan. Journal of Agriculture and Social Sciences. http://www.ijabjass.org. 1813–2235/2005/01–2, 120-126.

Kumari G. K. S. (1997). Forecasting and Estimation of Rainfall. M. Sc. (Ag.) Statistics thesis. Acharya N. G. Ranga Agricultural University.

Nnaji A. O. (2001). Forecasting seasonal rainfall for agricultural decision-making in northern Nigeria. Agricultural and forest meteorology. 107, 193-205.

Pramanik Chanchal (2010). Study of Fertilizer Consumption Pattern in Districts of Northern Telangana Region of Andhra Pradesh. M.Sc. (Ag.) Statistics thesis. Acharya N. G. Ranga Agricultural University.

Pramanik Chanchal, S K Dey and Sarkar A. (2010). Effect of Urbanization on Agriculture: A Special Scenario on Andhra Pradesh, India. International Journal of Applied Science & Computations, 17(2), 121-128.

Weesakul U., Lowanichchai S. (2005). Rainfall Forecast for Agricultural Water Allocation Planning in Thailand. Thammasat International Journal of Science and Technology, 10(3), 18-27.