International Journal of Statistics and Applied Mathematics 2023; 8(4): 36-39

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2023; 8(4): 36-39

https://www.mathsjournal.com
Received: 19-04-2023
Accepted: 24-05-2023

Kumar C
Agricultural Statistics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

Vijaya Bhama M
Department of Basic and Social Sciences, Forestry College and Research Institute, Mettupalayam, Tamil Nadu, India

Selvaraj KN
Department of agricultural economics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

Ravikumar R
Department of Basic and Social Sciences, Forestry College and Research Institute, Mettupalayam, Tamil Nadu, India

Corresponding Author:
Kumar C
Agricultural Statistics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

Spatial and temporal variations of tapioca

Kumar C, Vijaya Bhama M, Selvaraj KN and Ravikumar R

Abstract
This paper explains the spatial and temporal variations of tapioca, panel data of seven districts *viz*, Dharmapuri, Salem, Erode, Namakkal, Tiruchirappalli, Villupuram and Tiruvannamalai were used for the time period of 19 years from 2000-01 to 2018-19. It was found that, tapioca area as a dependent variable and nine independent variables like tapioca productivity, other fallow lands, cultivable waste, barren and uncultivable land, rainfall, cropping intensity, tapioca price, irrigation intensity and maize area used for the analysis. The results showed that the cultivation of tapioca area was low due to the price rate of tapioca and farmers choice for maize cultivation. The price rate of maize was high when compared to tapioca, therefore it played a vital role in farmers choice of maize cultivation over tapioca. OLS regression model and the spatial autoregressive model were used for the variables. Among this SAR model gives better results than the OLS regression model with minimum AIC and BIC values.

Keywords: Spatio-temporal variation, spatio-autoregressive model, ols regression, panel data regression, tapioca

1. Introduction
Tapioca (*Manihot esculenta Crantz*) was a dicotyledonous plant being one among the important crops cultivated in Tamil Nadu. Tapioca is mostly grown in Tamil Nadu, Kerala, Andhra Pradesh, Nagaland, Meghalaya, Assam, and other parts of India (Prabakar *et al.*, 2019) [7]. Tamil Nadu state stands first (64%) in respect of tapioca production (Ragavi *et al.*, 2019) [9]. Namakkal, Salem, Dharmapuri, Villupuram, and Erode districts produce approximately 72 percent of Tamil Nadu's tapioca. Namakkal leads the state in tapioca output and productivity, accounting for roughly 20.15 percent of overall production in Tamil Nadu (Shankar *et al.*, 2021) [10].

Spatial and temporal analysis is an emerging research area because of the development and application of novel computational techniques allowing for the analysis of large spatiotemporal databases (Lee and Yu, 2010) [4]. In the field of agriculture, this Spatio-temporal analysis plays a crucial role. To increase the productivity of agriculture, this analysis will pave the way for factors impacting agricultural production and agricultural conservation problems.

A standard method for analysing data with spatial correlation is the spatial autoregressive (SAR) model (Qu and lee, 2015) [8]. The Spatial Autoregressive (SAR) model is a spatial approach that can be used by considering the spatial effect to explain the relationship between the dependent variable and independent variables (Vijayabhama *et al.*, 2016) [12]. This is used to ensure that the data has a spatial effect. Regular OLS model can be used when there is no spatial effect present in the data (Tognelli and Kelt 2004). If the spatial effect is present in the data, then the SAR model gives a better result than the OLS regression (Permai *et al.*, 2019) [6].

2. Materials and methods
2.1 Area of study
The study area was the major tapioca-growing districts of Tamil Nadu. Seven major tapioca-producing districts are selected. In that, Salem, Dharmapuri, Erode, Namakkal, Tiruchirappalli, Villupuram, and Tiruvannamalai are the majorly Tapioca grown districts of Tamil Nadu.

The panel data was collected for the seven districts (Dharmapuri, Salem, Erode, Namakkal, Tiruchirappalli, Villupuram and Tiruvannamalai) with a time period of 19 years from 2000-2001 to 2018-19.

2.2 Spatial autoregressive model (SAR)
The spatial autoregressive model is also known as Spatial Lag Model and is suitable when the focus of interest is the evaluation of the presence and strength of spatial interaction Zheng *et al*., 2013 and Anselin, 1988) [13, 1]. It states the levels of the dependent variable Y in neighbouring regions depends on the levels of Y. Thus, it is a formulation of the spatial spillover model (Jaisankar *et al*., 2020) [3].
The spatial lag model or the spatial autoregressive model (Cliff and Ord, 1973; Le Sage and Pace, 2009 and Anselin, 1988) [2, 1, 5] is defined as:

$$Y = \rho\, WY + X\beta + e$$

Where
Y= $R \times 1$ dependent variable's vector variation.

R is defined as the no of districts, W= $R \times R$ spatial weight matrix with the zero diagonal elements, ρ = Spatial lag parameter or spatial autoregressive coefficient, WY = average of spatially neighbouring Y values of the spatially lagged dependent variable, X = $R \times k$ exogenous variables matrix observations, with associated with the $k \times 1$ regression coefficient of Y values, e is the error term.
Note that it makes sense because the diagonal elements of W are zero, which means that there is no circular definition so that Y_j on the left is affected by the same Y_j on the right. Because the presence of Y on both the left and the right sides implies that the presence of correlation between the errors and the regressors problem and the results would be biased and inconsistent. But the reduced form can easily be obtained as:

2.3 Econometric Equations Tapioca area
Panel data regression without spatial effect

$$TA_{ij} = a + b_1\, TPDY_{ij} + b_2\, OF_{ij} + b_3\, CW_{ij} + b_4\, BUL_{ij} + b_5\, RF_{ij} + b_6\, CI_{ij} + b_7\, TP_{ij} + b_8\, II_{ij} + b_9\, MA_{ij}$$

Panel data regression with spatial effects

SAR model

$$TA_{ij} = a + \rho\, W_Y\, TA + b_1\, TPDY_{ij} + b_2\, OF_{ij} + b_3\, CW_{ij} + b_4\, BUL_{ij} + b_5\, RF_{ij} + b_6\, CI_{ij} + b_7\, TP_{ij} + b_8\, II_{ij} + b_9\, MA_{ij}$$

Where, ρ = Spatial lag or Spatial autoregressive parameter
W = Contiguity-based row standardized spatial weight matrix

$W_Y TA$ = Spatially lagged dependent variable
$TPDY_{ij}$ = Tapioca Productivity (Kg/ha)
OF_{ij} = Other fallow lands (ha)
CW_{ij} = Cultivable waste (ha)
BUL_{ij} = Barren and uncultivable Land (ha)
RF_{ij}= Rainfall (mm)
CI_{ij} = Cropping Intensity (%)
TP_{ij} = Tapioca Price (Rs/qtl)
II_{ij} = Irrigation Intensity (%)
MA_{ij} = Maize Area (ha)
i = 1, 2, 3…, 7 districts; j = 1, 2, 3…, 19 years

3. Results and Discussion
3.1 Spatial and Temporal Variation
Spatial and temporal variation of tapioca was calculated using the SAR model. For the analysis, the spatial weight matrix has been created. Shape file of the selected district was used for the analysis. Neighbourhood matrix was created using the queen's contiguity method. In this method, all the eight directions were considered as a neighbour. For Rooks contiguity matrix, only four directions were considered as a neighbour and the remaining were not taken to account. Then the queen's contiguity matrix was row standardised and used for further analysis.

3.1.1 Panel data regression using pooled OLS regression
The linear regression model using pooled OLS regression was done by using the STATA 16 software. For regression analysis, tapioca area was taken as the dependent variable along with nine independent variables. The results obtained by the regression was listed below
The coefficient of determination (R^2) value was 0.51, the adjusted R^2 value of 0.48 and the RMSE value with 6065.8. The result of R^2 showed that 51 percent of the variation was caused by the independent variables. And the remaining 49 percent of the variation was caused by other factors. Among the nine independent variables other fallow lands, cultivable waste and tapioca price were negatively influenced the tapioca area by one percent level of significance. Barren and uncultivable land had positively influenced the area of tapioca with one percent level of significance. Tapioca productivity, irrigation intensity and maize area have negatively influenced the area of tapioca. The equation of the regression model was given below.

$$TA_{ij} = 27385.930 - 0.061\, TPDY_{ij} - 0.066\, OF_{ij} - 1.175\, CW_{ij} + 0.172\, BUL_{ij} + 0.444\, RF_{ij} + 87.008\, CI_{ij} - 3.777\, TP_{ij} - 140.313\, II_{ij} - 0.034\, MA_{ij}$$

3.1.2 Spatial Autoregressive Model
The spread of Tapioca area with and without the spatial effect was presented in Table 4.2. The coefficient of determination (R^2) value was 0.30 within model, 0.54 in between the model and 0.46 was the overall model and the functional form of the model satisfies the goodness of fit measures. The R^2 value of 0.30 indicates that about 30 percent of the variation in tapioca area was influenced by the explanatory variables included within model. 54 percentage of the variation caused by the between model, and 46 percent of the variations for the overall model. The results of the R^2 values indicated that the values of within and overall effects of the models are not the same, In the SAR model, there was proof that individual effects were so important. SAR model was calculated based on the summary statistics of seven districts, irrespective of the time and statistics of the 19 time periods, irrespective of districts, were calculated for within, between and overall effects of the model.
In main effect, the area of the tapioca was spatially and temporally significant, which was positively stimulated for the barren and uncultivable land by five percent level of significance and negatively influenced by other fallow and cultivable waste at one percent level of significance. Irrigation intensity was negatively significant at the five percent level. Tapioca productivity, tapioca price and maize area were also negatively influenced the area of tapioca. In the spatial-panel lag model, the estimate of the parameter ρ was significant at one percent. It suggested that, in 7 districts, there was a

significant spatial correlation between the tapioca area and other explanatory variables. The tapioca area spatial lag has added predictive power to this model. The Theta value was significant at one percent. It suggested that the rate at which the area of tapioca would decrease by 1.49 ha annually, keeping all other variables constant. In the selected sample area, the explanatory variables in the model influenced the area of tapioca.

In direct effect of the model was used to test the hypothesis as to whether a specific variable has a significant impact on the dependent variable in the particular district rather than the coefficient estimation of that variable. Rainfall and barren & uncultivable land were positively significant at the five percent level. Other fallow lands and cultivable wastes were negatively significant at one percent level and irrigation intensity was negatively significant at five percent. Maize area and tapioca price are also showed the negatively influenced on the area of tapioca.

The indirect effect was used to assess whether spatial spillover occurs rather than the spatially lagged dependent coefficient estimate. The spatial spillover effect of the tapioca productivity, other fallow, cultivable waste, tapioca price, irrigation intensity and maize area were negatively influenced the tapioca area. Barren & uncultivable land, rainfall and cropping intensity are positively influenced by the area of tapioca.

The total effect is the sum of the direct and indirect effects. The total effect of the other fallow land and the cultivable waste area was negatively influenced at one percent level of significance. Similarly, tapioca productivity, tapioca price and maize area was negatively influenced, and irrigation intensity was negatively affected the tapioca area with five percent level of significance.

The equation of the SAR model is given below

$$TA_{ij} = 18812.870 + 0.108\, W_Y TA - 0.085\, TPDY_{ij} - 0.036\, OF_{ij} - 0.999\, CW_{ij} + 0.173\, BUL_{ij} + 3.715\, RF_{ij} + 72.609\, CI_{ij} - 2.761\, TP_{ij} - 95.134\, II_{ij} - 0.100\, MA_{ij}$$

3.1.3 Comparison between the OLS regression and the SAR model

The comparison of the linear regression model using OLS regression and Spatial Autoregressive model was given in Table 4.4. Based on the comparison Akaike Information Criterion (AIC) and Bayesian Information Criteria (BIC) values were calculated for both the models. The best model was found out using the smallest values of AIC and BIC values. Based on Table 4.4 values known that AIC and BIC values of the SAR model were smaller than the OLS regression model. Therefore, based on the values, it was concluded that the SAR model was the best model than the OLS regression model.

3.1.4 Summary and Conclusion

The results of the study put forth that the spread of tapioca area was spatially and temporally significant with the significant spatial rho value at a one percent level. Based on the coefficient of determination values (R^2), the independent variables explained 30 percent of the variations influenced by within the model, 54 percent of the variations caused by between models and 46 percent of the variables are explained in, the overall model.

The variations caused by the different variables such as tapioca productivity, other fallow lands, cultivable waste, tapioca price, irrigation intensity and maize area were negatively influenced by the area of tapioca. Meanwhile barren and uncultivable land, rainfall, and cropping intensity positively influenced the area of tapioca in main, direct, indirect and total effects of the model. These results showed that when the possibility for higher irrigation intensity farmers preferred to grow maize crops and when they were dependent on rainfall preference was towards the cultivation of tapioca. SAR model showed that majority of the tapioca growing area were rainfed. Theta value showed that significance at one percent level. Based on the theta value we can conclude that area of the tapioca decreased at the rate of 1.49 ha per annum, keeping the other variables are constant

Comparison of OLS regression and the SAR model has been carried out based on the AIC and BIC values. SAR model gave the minimum of AIC (2620.034) and BIC (2637.377) values when compared to the OLS regression model. The result of the study showed that the SAR model gives better results than the OLS regression when the spatial effect is present in the data.

The overall study showed that the area of the tapioca is in the decreasing trend at the rate of 1.49 ha per annum. This happened because of the negative trend of the tapioca price. On the other hand, the farmers are encouraged to cultivate maize because the market price was far better than that of tapioca. So, the cultivated area of maize showed an increasing trend in contrast to the tapioca area.

Table 1: Pooled OLS regression of tapioca area

Column1	Coefficient	Standard Error	P-Value
Tapioca Productivity	-0.061	0.046	0.188
Other Fallow Land	-0.066**	0.021	0.002
Cultivable Waste	-1.175**	0.184	0.000
Barren and uncultivable Land	0.172**	0.049	0.001
Rainfall	0.444	2.206	0.841
Cropping Intensity	87.008	72.396	0.232
Tapioca Price	-3.777**	1.194	0.002
Irrigation Intensity	-140.313	83.141	0.094
Maize Area	-0.034	0.072	0.636
Constant	27385.93**	7043.442	0.000

* Indicates significance at 5% level of significance
** Indicates significance at 1% level of significance

Table 2: Spatio-temporal variations of Tapioca Area

Variables	Coefficient	Standard Error	Z	P-Value
Main Effect				
Tapioca Productivity	-0.085	0.049	-1.740	0.082
Other Fallow Land	-0.036**	0.012	-2.880	0.004
Cultivable Waste	-0.999**	0.345	-2.900	0.004
Barren and uncultivable Land	0.173*	0.088	1.970	0.048
Rainfall	3.715	1.919	1.940	0.053
Cropping Intensity	72.609	56.232	1.290	0.197
Tapioca Price	-2.761	1.869	-1.480	0.140
Irrigation Intensity	-95.134*	42.608	-2.230	0.026

Maize Area	-0.100	0.116	-0.860	0.392
Constant	18812.870**	6471.395	2.910	0.004
Spatial Rho	0.108**	0.041	2.650	0.008
Variance				
lgt_theta	-1.490**	0.268	-5.570	0.000
sigma2_e	15700000**	3487688	4.500	0.000
Direct Effect				
Tapioca Productivity	-0.083	0.060	-1.390	0.164
Other Fallow Land	-0.036**	0.012	-2.980	0.003
Cultivable Waste	-0.984**	0.332	-2.960	0.003
Barren and uncultivable Land	0.169*	0.082	2.050	0.040
Rainfall	3.695*	1.811	2.040	0.041
Cropping Intensity	76.055	54.313	1.400	0.161
Tapioca Price	-2.673	1.867	-1.430	0.152
Irrigation Intensity	-94.601*	40.779	-2.320	0.020
Maize Area	-0.095	0.111	-0.860	0.389
Indirect Effect				
Tapioca Productivity	-0.009	0.008	-1.200	0.230
Other Fallow Land	-0.004	0.003	-1.680	0.093
Cultivable Waste	-0.120	0.069	-1.730	0.083
Barren and uncultivable Land	0.022	0.016	1.360	0.173
Rainfall	0.443	0.298	1.490	0.137
Cropping Intensity	10.007	9.041	1.110	0.268
Tapioca Price	-0.289	0.235	-1.230	0.219
Irrigation Intensity	-12.352	8.665	-1.430	0.154
Maize Area	-0.011	0.014	-0.790	0.429

* Indicates significance at 5% level of significance
** Indicates significance at 1% level of significance

Table 3: Spatio-temporal variations of tapioca area

Total Effect				
Tapioca Productivity	-0.092	0.066	-1.400	0.162
Other Fallow Land	-0.041**	0.014	-2.840	0.004
Cultivable Waste	-1.103**	0.386	-2.860	0.004
Barren and uncultivable Land	0.191*	0.097	1.980	0.048
Rainfall	4.138*	2.041	2.030	0.043
Cropping Intensity	86.062	62.440	1.380	0.168
Tapioca Price	-2.962	2.060	-1.440	0.150
Irrigation Intensity	-106.953*	48.561	-2.200	0.028
Maize Area	-0.107	0.123	-0.870	0.387

* Indicates significance at 5% level of significance
** Indicates significance at 1% level of significance

Table 4: Comparison of OLS regression and SAR model

	AIC	BIC
OLS regression	2704.012	2732.915
SAR model	2620.034	2637.377

4. References

1. Anselin Luc. Spatial Econometrics: Methods and Models, Dordrecht, the Netherlands. Kluwer Academic Publishers; c1988.
2. Cliff AD, Ord JK. Spatial autocorrelation. Pion, London; c1973.
3. Jaisankar R, Kesavan J, Varadha Raj S. A Spatial Autoregressive Model for Predicting the Dengue Incidences in Tamil Nadu, India. Gedrag & organisatie review. 2020;33(3):178-187.
4. Lee LF, Yu J. Estimation of spatial autoregressive panel data models with fixed effects. Journal of Econometrics. 2010;154(2):165-185.
5. LeSage, Pace. Spatial econometric models Handbook of applied spatial analysis; c2009. p. 355- 376
6. Permai SD, Jauri R, Chowanda A. Spatial autoregressive (SAR) model for average expenditure of Papua Province. Procedia Computer Science. 2019;157:537-542.
7. Prabakar C, Sita Devi K, Ponnarasi T, Stalin P. Estimation of Cost of Cultivation and Evaluation of the Marketing Channels of Tapioca in Tamil Nadu. Plant Archives. 2019;19(2):2810-2814.
8. Qu X, Lee LF. Estimating a spatial autoregressive model with an endogenous spatial weight matrix. Journal of Econometrics. 2015;184(2):209-232.
9. Ragavi G, Kumar S, Rai AK. An Economic Analysis of Production of Tapioca in Namakkal District of Tamil Nadu. International Journal of Innovative Science and Research Technology. 2019;4(5):352-356.
10. Shankar R, Gowri R, Rajesh, Selvam S. A study on resource use efficiency and economic returns for Tapioca in Namakkal district of Tamil Nadu, India. 2021.
11. Tognelli MF, Kelt DA. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. Ecography. 2004;27(4): 427-436.
12. Vijayabhama M, Jaisankar R, Varadha Raj S, Baranidharan K. Spatial–temporal variation of casuarina spread in Cauvery delta and northeastern zone of Tamil Nadu, India: A spatial autoregressive model. Journal of Applied Statistics. 2016;45(1):1-7. DOI: 10.1080/02664763.2016.1247786
13. Zheng X, Li F, Song S, Yu Y. Central government's infrastructure investment across Chinese regions: A dynamic spatial panel data approach, China Econ. Rev. 2012;27:264-276.