The Pharma Innovation Journal 2021; SP-10(4): 506-509

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2021; SP-10(4): 506-509

www.thepharmajournal.com
Received: 12-02-2021
Accepted: 15-03-2021

Dr. B Sivasankari
Assistant Professor, (Mathematics), Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai, Tamil Nadu, India

Dr. R Vasanthi
Assistant Professor, (Mathematics), Tamil Nadu Agricultural University, Coimbatore Tamil Nadu, India

Dr. G Vanitha
Assistant Professor (Computer Science), Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

Corresponding Author:
Dr. B Sivasankari
Assistant Professor, (Mathematics), Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai, Tamil Nadu, India

Identification of appropriate statistical models for forecasting area, production and productivity of tapioca crop in Salem district of Tamil Nadu

Dr. B Sivasankari, Dr. R Vasanthi and Dr. G Vanitha

Abstract
The study has been undertaken to fit different trend equations like linear, non-linear and time series models for Tapioca and also made the future forecasts by 2022 AD. The study was carried out in Salem district of Tamil Nadu state as a whole using time series data from 1997-1998 to 2018-2019. For forecasting purpose linear and non-linear growth models viz. linear, logarithmic, inverse, quadratic, cubic, power, s-curve, logistic and exponential and time series models like ARIMA models were fitted to the area, production and productivity of Tapioca crop in Salem district of Tamil Nadu. The best fitted model for future projection was chosen based upon least RMSE, MAE and MAPE values. The study revealed that Quadratic model was identified as the best model in Tapioca area and production in Salem district. Cubic model was identified as the best model for Tapioca productivity in Salem district of Tamil Nadu. It was also observed that in Salem district of Tamil Nadu, tapioca showed decreasing trend by 2022 AD in area, production and productivity.

Keywords: Linear model, non-linear model, ARIMA, MAE and RMSE

Introduction
Cassava (*Manihot esculenta* crantz) also known commonly as Tapioca, continues to be a crop of food security for the millions of people especially in the developing countries of the globe. It is an important alternate source of energy to meet the demands of increasing population. This crop has the potential to produce more food per unit area, capacity to withstand adverse biotic and abiotic stresses and adaptability to the conditions of drought and marginal lands. Of the tropical root and tuber crops, Cassava occupies first position in terms of area and production globally. It is found as staple food for those living in several tropical countries of South America and Africa, for eg. Brazil, Nigeria, Ghana etc. Globally cassava is grown in an area of 18.51 million ha producing 202.65 million tonnes with a productivity of 10.95 t/ha. It is grown in 102 countries in the world. African continent occupies first position covering 66.21 per cent of cassava area producing 53.37 per cent of the world cassava as it is a staple in many of the African countries. Even though area is more in Africa, its production is low due to low productivity (8.82 t/ha) which is lower than the world average productivity.India acquires significance in the global cassava scenario due to its highest productivity in the world (27.92 t/ha.) and cultivated in an area of 240,000 ha producing 6.7 million tonnes.

In India, cultivation of tapioca is mostly concentrated in the southern states of Kerala (2 lakh ha area with 3 million tonnes of production) and Tamil Nadu (58,464 ha area with 1 million tonnes of production) and together they account for a little over 90 per cent of the area and production in the country. In India, productivity wise, Tamil Nadu tops the list with a yield of 20.70 tonnes per ha followed by 17.60 tonnes per ha in Kerala and 7.5 tonnes 2 per ha in Andhra Pradesh (Lakshmi *et al*, 1992). In Tamil Nadu, cassava is cultivated in about ten per cent of the area and contributes to more than 70 per cent of the total production in the country. Among the different districts in Tamil Nadu, Salem (31,569 ha in area with 11 lakh tonnes of production), Dharmapuri (31, 564 ha in area with 10 lakh tones of production) and Namakkal (27, 689 ha in area with 13 lakh tonnes of production) are the major cassava growing districts and put together, they account for 92 per cent of the total area planted.

Materials and Methods
The present study was based on secondary data. The time series pertaining to area, production and productivity of Tapioca for Salem district have been collected for the period of 22 years

from 1997 to 2018 from the official website Indiastat.com, Seasonal Crop Report. Data was analyzed by using MINITAB and SPSS version 16.0 software. In this study, the linear trend Model was also used by (Finger, 2007, Broken *et al*., 2000 and Rimi *et al*., 2011) [4].

Exponential, Quadratic and S-Curve Models of trend analysis were applied for this study. Also Logistic, Monomolecular and fuzzy time series models were fitted for the production data. The best fitted Model was selected on the basis of three accuracy measures. These accuracy measures were Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD) and Root Mean Square Error (RMSE). Mean Absolute Percentage Error (MAPE) measures the accuracy of fitted time series values. It expresses accuracy as a percentage. Mean Absolute Deviation (MAD) measures the accuracy of fitted time series values. It expresses accuracy in the same units as the data, which helps conceptualize the amount of error. The Root Mean Square Error (RMSE) is a measure of accuracy, to compare forecasting errors of different models for a particular dataset and not between datasets, as it is scale-dependent. Smaller values of all these measures indicate a good fitted Model with minimum forecasting errors (Karim *et al*., 2010) [3]. The relative measures give equal weight to all errors in contrast to the MSE, which squares the errors and there by emphasizes large errors. It would be helpful to have a measure that considers both the disproportionate cost of large errors and provides are relative basis for comparison with naive methods. This statistics allows a relative comparison of normal forecasting methods with naive approaches and also squares the errors involved so that large errors are given much more weight than small errors. The best fitted Model was applied for forecasting area, production and Productivity of Tapioca in Salem district respectively from the year 2019 to 2022.

Results and Discussion

In the present study, linear, non-linear models viz., linear, logarithmic, inverse, quadratic, cubic, compound, S-curve, growth, power and exponential models have been fitted whereas time series models include Auto Regressive (AR), Moving Average (MA) and Auto Regressive Integrated Moving Average (ARIMA) have been fitted to the data on area, production and productivity of the Tapioca in Salem district by considering 22 years data from 1997-98 to 2018-19. The results obtained for area of Tapioca in Salem District by fitting all the linear, non-linear and time series models were presented in Table1.1.

Table 1: Linear, Non-linear and Time series models of Tapioca area in Salem District of Tamil Nadu

Model	Parameter				Criteria		
	a	b	c	d	RMSE	MAE	MAPE
Linear	27910.81**	-803.02**			5479.74	4363.33	25.03
Logarithmic	28864.62**	-4338.99*			6143.55	5233.06	31.23
Inverse	17853.28**	10857.52			6602.64	5722.85	35.42
Quadratic	**20207.75****	**1397.86**	**-110.04***		**4623.99**	**3758.99**	**21.89**
Cubic	21694.85**	605.38	-13.48	-3.22	4605.85	3822.83	22.65
Power	31364.49**	-0.26*			6467.21	5384.33	29.47
S	9.71**	0.62			6796.14	5844.95	33.25
Exponential	29720.84**	-0.05**			5867.22	4530.28	24.30
Logistic	0.00003**	1.05**			5867.22	4530.28	24.30
Time Series Model							
ARIMA(1,1,0)					5802.41	4437.05	25.09
ARIMA(1,1,1)					5977.86	4416.37	24.79
ARIMA(0,1,1)					5923.64	4436.41	25.02

**,* indicate significant at 1% and 5% level of probability respectively.

From the above Table 1 it was revealed that, in comparison with all other models, Quadratic model has low MAPE (21.89) and MAE values. Hence, the Quadratic model was chosen for future forecasts of tapioca area in Salem district. Among the linear, non-linear growth models quadratic model identified as the best fitted model based on model selection criteria.

Forecasted Quadratic model is $\hat{Y}_{area}$ = 20207.75 + 1397.86 x – 110.04 x^2

The results obtained for production of tapioca in Salem district during the study period by fitting all the models were presented in Table 2.

Table 2: Linear, Non-linear and Time series models of Tapioca Production in Salem District of Tamil Nadu

Model	Parameter				Criteria		
	a	b	c	d	RMSE	MAE	MAPE
Linear	1031919.53**	-33768.93**			232915.59	185132.33	35.85
Logarithmic	1093882.94**	-193020.44*			255824.68	217247.01	43.79
Inverse	596419.19**	523829.09			273881.89	237827.67	49.87
Quadratic	**760825.26****	**43686.57**	**-3872.78**		**208664.38**	**162811.77**	**32.78**
Cubic	856756.15**	-7435.04	2356.50	-207.64	206988.44	163992.06	34.13
Power	1246064.90**	-0.34*			274886.60	234297.65	40.51
S	13.17**	0.87			287799.39	254496.61	46.11
Exponential	1129182.06**	-0.06**			252556.06	189276.74	32.86
Logistic	0.0000009**	1.06**			252556.06	189276.74	32.86
Time Series Model							
ARIMA (1,1,0)					262044.94	210628.09	42.78
ARIMA (0,1,1)					261678.68	201077.51	39.13
ARIMA(1,1,1)					267163.55	208653.66	39.64

**,* indicate significant at 1% and 5% level of probability respectively.

It appears from the above Table 2, in comparison with model selection criteria among all the models quadratic model was with least MAPE (32.78) value. Hence quadratic model was chosen for forecasting purpose.

Forecasted Quadratic model is $\hat{Y}_{production}$ = 760825.26 + 43686.57 x – 3872.78x^2

Among time series models ARIMA (0, 1, 1) lowest MAPE values. But as compared to quadratic model this model is less efficient.

The results obtained for productivity of tapioca in Salem district during the study period by fitting all the models were presented in Table 3.

It is evident from the above Table 4.6. the value of MAE(3.93) and RMSE(5.67) is least for cubic model compared to other growth models.

Forecasted Cubic model is $\hat{Y}_{productivity}$ = 38.68 – 1.20 x + 0.13 x^2 – 0.01 x^3

Among the time series models ARIMA (1,0,0) was identified as best model based on model selection criteria but as compared to ARIMA (1, 0, 0) cubic model was best fitted.

Table 3: Linear, Non-linear and Time series models of Tapioca Productivity in Salem District of Tamil Nadu

Model	Parameter				Criteria		
	a	b	c	d	RMSE	MAE	MAPE
Linear	37.84**	-0.39			5.73	3.84	11.88
Logarithmic	39.20**	-2.54			5.78	3.72	11.59
Inverse	32.49**	7.79			5.87	3.87	12.10
Quadratic	36.37**	0.03	-0.02		5.71	3.95	12.26
Cubic	**38.68****	**-1.20**	**0.13**	**-0.01**	**5.67**	**3.93**	**12.33**
Power	39.73**	-0.08			5.82	3.70	11.31
S	3.46**	0.25			5.91	3.99	12.24
Exponential	37.99**	-0.01			5.77	3.76	11.46
Logistic	0.026**	1.01**			5.77	3.76	11.46
Time Series Model							
ARIMA (1,0,1)					6.58	4.62	14.24
ARIMA (0,0,1)					6.38	4.61	14.24
ARIMA(1,0,0)					6.38	4.60	14.21

**,* indicate significant at 1% and 5% level of probability respectively.

Forecasting of Tapioca Area, Production and Productivity using best fitted Model

The future forecasts of area, production and productivity of tapioca crop by 2022 AD were calculated in Salem district of Tamil Nadu state based on the best identified fitted model in Linear, Non-Linear models and time series models the tapioca area was forecasted and tabulated in the Table 4.

Table 4: Forecasted values of Tapioca Area, Production and Productivity of Salem district

YEAR	Forecasted Area (hectares)	Forecasted Production (tonnes)	Forecasted Productivity (Kgs/ha)
2019	11518.48	338161.83	26.26
2020	10985.29	318377.68	24.03
2021	10476.79	299751.00	21.43
2022	9991.83	282214.08	18.43

Among all the models quadratic model was found to be best fitted model for area under tapioca in Salem district of Tamil Nadu as it has exhibited least MAE and MAPE values. The forecasted area of tapioca in Salem district by 2022 AD is expected to be 9991.83 hectares. The forecasts showed a decreasing trend by 2022 AD from the average of study period. Regarding the production of tapioca in Salem district, Quadratic model was selected as the best for future forecasts by 2022 AD as it had the least RMSE, MAE and MAPE values and the forecasted production would be 282214.08 tonnes in 2022 AD using Quadratic model. The forecasts showed a decreasing trend by 2022 AD is above from the average of study period. Productivity of tapioca in Salem district was forecasted by using cubic model which has exhibited least Root Mean Square Error and Mean Absolute Error values. By using Cubic model the forecasted productivity of Salem district would be 18.43 Kgs/ ha in 2022 AD.. Based on the best identified fitted (cubic) model the Tapioca productivity of Salem district was forecasted and tabulated in the Table 4.

Conclusion

Based on the least RMSE, MSE and MAPE value, among all the linear, nonlinear and time series models for Tapioca area, production and Productivity cubic model was best fitted model for Salem district. The average area, production and productivity of tapioca in Salem district of Tamil Nadu during the study period were 19880.63 hectares, 694239 tonnes and 33.95 Kgs/ ha respectively. The future forecasts of area and production was worked out by Quadratic model. The forecasts indicated that there would be considerable decrease in area and production of Tapioca. But regarding productivity of tapioca, it was found that Cubic model was best fitted for future forecasts and it revealed that there would be a decreasing trend in productivity. It was observed that the forecasted area of Tapioca by 2022 AD would be 9991.83 hectares in Salem district, the forecasted production and productivity would be 282214.08 tonnes and 13.30Kgs/ ha respectively by 2022 AD.

References

1. Boken VK. Forecasting spring wheat yield using time series analysis: a case study for the Canadian Prairies, Agronomy Journal 2000;92(6):1047-1053.
2. Finger R. Evidence of Slowing Yield Growth- The example of Swiss Cereal Yield. Agri-food and Agri-

Environmental Economics Group, ETH Zürich, Switzerland. Government of Pakistan. 2012. Economic survey of Pakistan, economic advisory wing, finance department, Islamabad 2007.
3. Karim Md R, Awal Md A, Akter M. Forecasting of wheat production in Bangladesh. Bangladesh Journal of Agricultural Research 2010;35(1):17-28.
4. Rimi RH, Rahman SH, Karmaker S, Hussain G. Trend Analysis of Climate Change and Investigation on its Probable Impacts on Rice Production at Satkhira, Bangladesh. Pakistan Journal of Meteorology 2011, 6.