Int.J.Curr.Microbiol.App.Sci (2018) *7*(4): 267-276

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 7 Number 04 (2018)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2018.704.030

Utilization of Parametric and Nonparametric Regression Models for Production, Productivity and Area Trends of Apple (*Malus domestica*) in Jammu and Kashmir, India

Yasmeena Ismail*, S.A. Mir and Nageena Nazir

Division of agricultural statistics, Sher-e-Kashmir University of Agricultural Sciences and Technology-Kashmir, India

**Corresponding author*

Keywords

Apple production, Trend analysis, Parametric regression, Nonparametric regression

Article Info

Accepted:
04 March 2018
Available Online:
10 April 2018

A B S T R A C T

Apple (*Malus domestica*) is one of the most important horticultural produce and illustrious worldwide for its health benefits. One of the major apple producing zones in the north India is Jammu and Kashmir and major portion of the state is economically dependent on this horticultural industry. The present study is an attempt to find past trends of apple in Jammu and Kashmir using parametric, nonparametric and semi-parametric regression methods. The performance of each method is compared using higher values of R^2 and lower values of residual criteria. It is found that the nonparametric/semi-parametric regression comes out to be good fit for trends in apple production in comparison to the parametric regression. Even semi-parametric spline regression is selected as the best fitted model for trend analysis. It is inferred that the area under apple cultivation in Jammu and Kashmir is increasing from 1974-2015 and the productivity has also shown an increasing trend except for some years where the trend is found declining. The study advocates for researchers technological breakthrough in apple production in Jammu and Kashmir.

Introduction

The apple cultivation has abundantly colored tranquility of Kashmir`s landscape. Kashmiri apple has lived up to its reputation for being one of the high-quality fruits, Kashmir has for long been considered the home of apples. Hundred and thirteen varieties of apple are found in Jammu and Kashmir. Though the cultivation of apple in India is concentrated in Jammu and Kashmir, Himachal Pradesh, and Uttar Pradesh yet, Kashmir enjoys the distinction of being still hub of apple industry of the country. This is obviously so because the State has not only superiority over Himachal and Uttar Pradesh in the field of production but also in marketing. In order to maintain this superiority emphasis are laid to study the trends in area, production and productivity of apple in Jammu and Kashmir. Analysis of trend is useful for planning and decision making purposes of appropriate policy measures (Sheikh and Tripathi, 2013).

The growth rates of crops are mostly estimated by the linear regression models. However, it might be the case that these models may not fit the data well. Under such

conditions it becomes essential to apply nonparametric and semi-parametric regression, which is based on fewer assumptions. In last few years, nonparametric regression and semi-parametric regression technique for functional estimation has become increasingly popular as a tool for data analysis. These techniques impose only few assumptions about shape of function and therefore it is more flexible than usual parametric regression approaches. Smoothing techniques are commonly used to estimate the function non-parametrically (Härdle, 1990). Nonparametric regression models avoid restrictive assumptions of the functional form of the regression function. Semi-parametric regression model combine the components of parametric and nonparametric regression models, by keeping the easy interpretability of the former and retains some of the flexibility of the latter. Various scientists viz., (Chandran, 2004) has applied nonparametric regression to study the growth rates of total foodgrain production of India during the period 1987 to 2001. Teczan (2010) has studied the nonparametric regression technique to find out the growth rate trends of various crops. Sahu and Pal (2004) used nonparametric regression (Lowess) and semi-parametric (spline) for modeling of pest incidences. (Dhekale *et al.*, 2017) employed the nonparametric regression model to study the trends of tea in India. The current study is aimed to develop appropriate parametric and nonparametric regression models to fit the trends in area, production and productivity.

Materials and Methods

For present study, to study the trends and growth rates, long term data for last 42 years pertaining to the area, production and productivity of apple is collected from Directorate of Horticulture. The descriptive measures of central tendency and dispersions along with the simple and compound growth rates are used to explain the features of the data (Mishra *et al.*, 2012).

Trend Models

Parametric regression models

To find out the path of the production process different parametric trend models are fitted. Among the fitted models, the best model is selected on the basis of their goodness of fit (R^2) value and significance of the coefficients. The dependent variable Y is area, production and productivity and independent variable X is the time points (years).

Non-parametric and semi-parametric regression models

The model considered here is of the form

$$Y_i = m(x_i) + \varepsilon_i, x_i = i/n, i = 1,2,...n$$

Where, Y_i is observation of i^{th} time point, $m(\cdot)$ is trend function which is assumed to be smooth and ε_i are random errors with mean zero and finite variance. Since there is no assumption of parametric form of function $m(\cdot)$, this approach is flexible and robust to deviations from an assumed model form. To obtain an estimate of the mean response value at a point X, most of the smoothers are averaging the Y – values of observations having predictor values closer to the target value X. The averaging is done in neighborhoods around target value. The main decision to be made in any of the smoothing techniques is to fix the size of neighborhood which is typically expressed in terms of an adjustable smoothing parameter or bandwidth. Intuitively, large neighborhoods will provide an estimate with low variance but potentially high bias, and conversely for small neighborhoods. Lowess regression, introduced

by Cleveland (1979), is obtained on the basis of the data points around it within a band of certain width. The point x_i is the midpoint of the band. The data points within the band are assigned weights in a way so that x_i has the highest weight. The weights for the other data points decline with their distance from x_i according to a weight function. The weighted least squares method is used to find the fitted value corresponding to x_i, which is taken as the smoothed value. The procedure is repeated for all the data points. The spline method of estimation make use of the penalized least squares method (Simonoff, 2012), which balances the fitting of the data closely. The objective is to estimate *m* by means of a function that fits the data well and is as smooth as possible. A measure of smoothness of m is the integral of the square of its second derivative as

$$\sum_{i=1}^{n}(Y_i - m(x_i))^2 + \lambda\int_a^b (m^{"}(x))^2\,dx$$

Where $\lambda > 0$ is a fixed constant and $x_i \in [a,b], i = 1,2,...n$

The first term is the sum of squares of the residuals; it provides a measure of how well the function *m* fits the data. The integral of the above equation is a measure for the roughness/smoothness of the function m. The functions which are highly curved will result in a large value of the integral; straight lines result in the integral being zero. The roughness penalty λ, controls the emphasis which one wishes to place on smoothness. By increasing the value of λ, one places more emphasis on smoothness; as λ becomes large the function approaches a straight line. On the other hand, a small value of λ emphasizes the fit of *m* to the data points: as λ approaches a function that interpolates the data points.

Results and Discussion

The maximum growth rate is observed in production of apple over the years, whereas the minimum growth rate is exhibited by productivity of the apple.

The positive compound growth of production (0.06 percent per annum) reveals that there is no decrease in the production of apple over the years with a maximum of 1.97 million kilogram and minimum of.19 million kilogram. Similarly, the simple growth rate (9.32 per cent per annum) is observed in production indicates an increase in the production of apple in Jammu and Kashmir over the years (Table 1 and 2).

This is due to the fact that a large area of land is being brought under agriculture we have noticed a compound growth rate of area (0.03 per cent per annum) under apple cultivation indicating that a large portion of the land is being utilized for the latter.

Trend analysis of area, production and productivity

Parametric techniques

The linear models used here are the cubic model or the third degree polynomial model $Y_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3$ and the quadratic model or second degree polynomial model $Y_t = b_0 + b_1 t + b_2 t^2$. The value of b_2 for area is negative which indicates that area under apple cultivation decreased in the middle part of the cultivation period and the value of b_1 and b_2 being positive clearly indicates that there was an increase in the cultivation area. Further, the negative value of b_1 for production is an indication of the decrease in the production during the initial period of the study and the positive values of b_2 and b_3 indicates an increase in the production

Table.1 Performance of apple production in Jammu and Kashmir during 1974-2015

Area ('000 hectare)		Production ('000 MT)		Productivity (MT per hectare)	
Maximum	171.00	Maximum	1966.42	Maximum	13.07
Minimum	46.19	Minimum	190.45	Minimum	4.12
Mean	90.57	Mean	8.77	Mean	9.40
SD	36.69	SD	4.30	SD	1.67
CV (%)	4.05	CV (%)	4.91	CV (%)	1.78
Skewness	0.86	Skewness	0.81	Skewness	-0.82
Kurtosis	-0.40	Kurtosis	0.19	Kurtosis	1.64
SGAR (%)	2.50	SGAR (%)	9.32	SGAR (%)	1.94
CGAR (%)	0.03	CGAR (%)	0.06	CGAR (%)	0.02

CV= coefficient of variation, SD= standard deviation, SGAR= simple growth rate per annum, CGAR= compound growth rate per annum

Table.2 Trends in area, production and productivity of apple in Jammu and Kashmir

	R^2	Constant b_0	b_1	b_2	b_3	RMSE	MAPE	MAE	MaxAPE	MaxAE
Area	0.92	345.9	3.092	-0.024	0.041	11.09	2.89	4.09	5.83	32.56
Production	0.91	37.32	-1.980	0.436	0.002	17.45	4.90	19.45	6.21	51.34
Productivity	0.90	645.33	14.98	-0.432		76.90	6.34	34.78	17.98	91.34

Area in '000 hectares, Production in '000 metric tons, Productivity in metric ton per hectare

Table.3 Trends in area of apple in Jammu and Kashmir using non-parametric and semi-parametric regression

	Loess	Splines
Bandwidth	0.23	0.74
R^2	0.987	0.996
AIC_c	1.77	1.02
RMSE	3.03	2.21
MAPE	1.33	1.22
MAE	1.35	1.14
MaxAPE	9.91	8.43
MaxAE	16.95	14.13

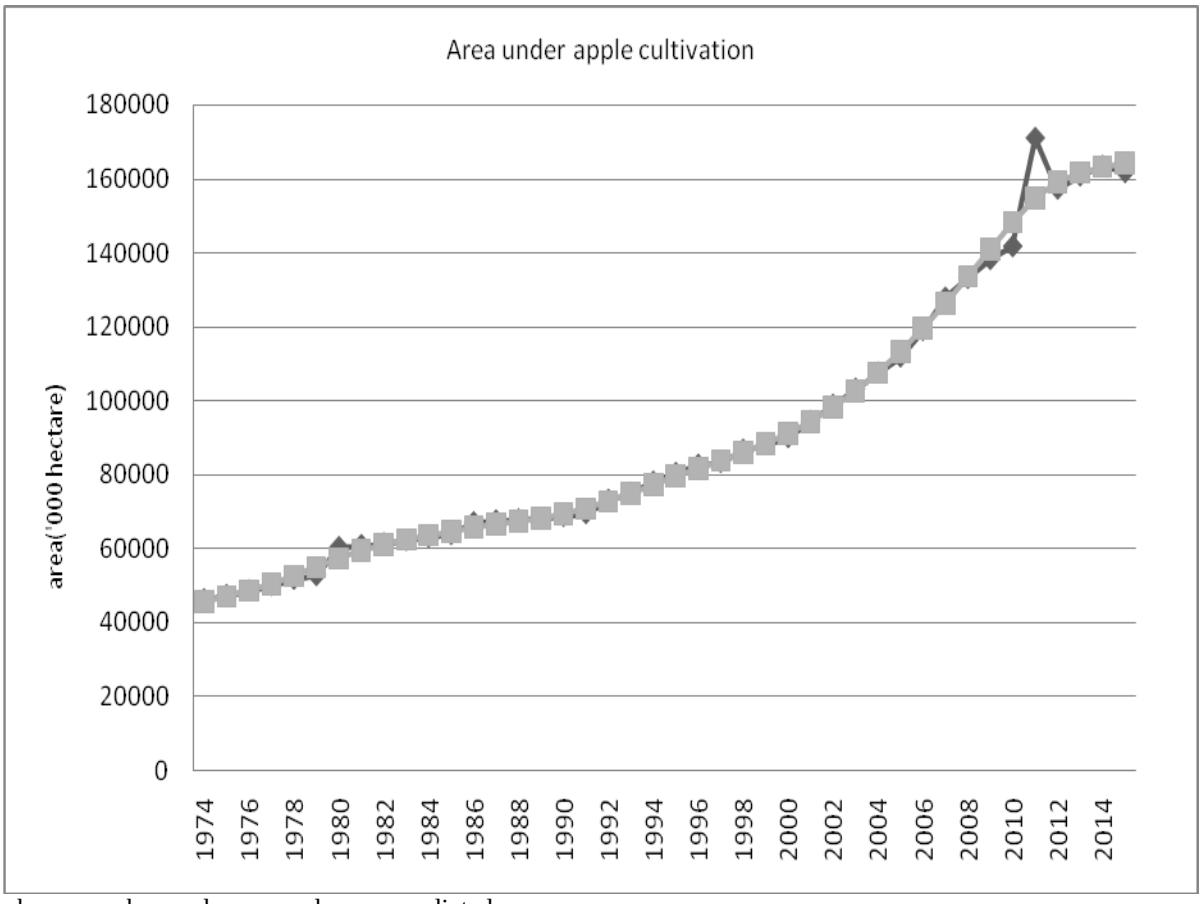
Table.4 Trends in production of apple in Jammu and Kashmir using non-parametric and semi-parametric regression

	Loess	Splines
Bandwidth	0.66	0.165
R^2	0.985	0.998
AIC_c	0.25	0.10
RMSE	1.35	1.04
MAPE	11.73	10.89
MAE	9.15	8.63
MaxAPE	6.18	5.08
MaxAE	4.80	4.03

Table.5 Trends in productivity of apple in Jammu and Kashmir using non-parametric and semi-parametric regression

	Loess	Splines
Bandwidth	0.53	0.49
R^2	0.988	0.997
AIC_c	1.69	1.12
RMSE	1.22	0.99
MAPE	10.12	9.11
MAE	2.84	0.84
MaxAPE	5.37	2.13
MaxAE	6.08	3.16

Fig.1 Observed and expected trends of area under apple cultivation using spline in Jammu and Kashmir



obs_area=observed area, pred_area=predicted area

Fig.2 Fits with specified smooths for area

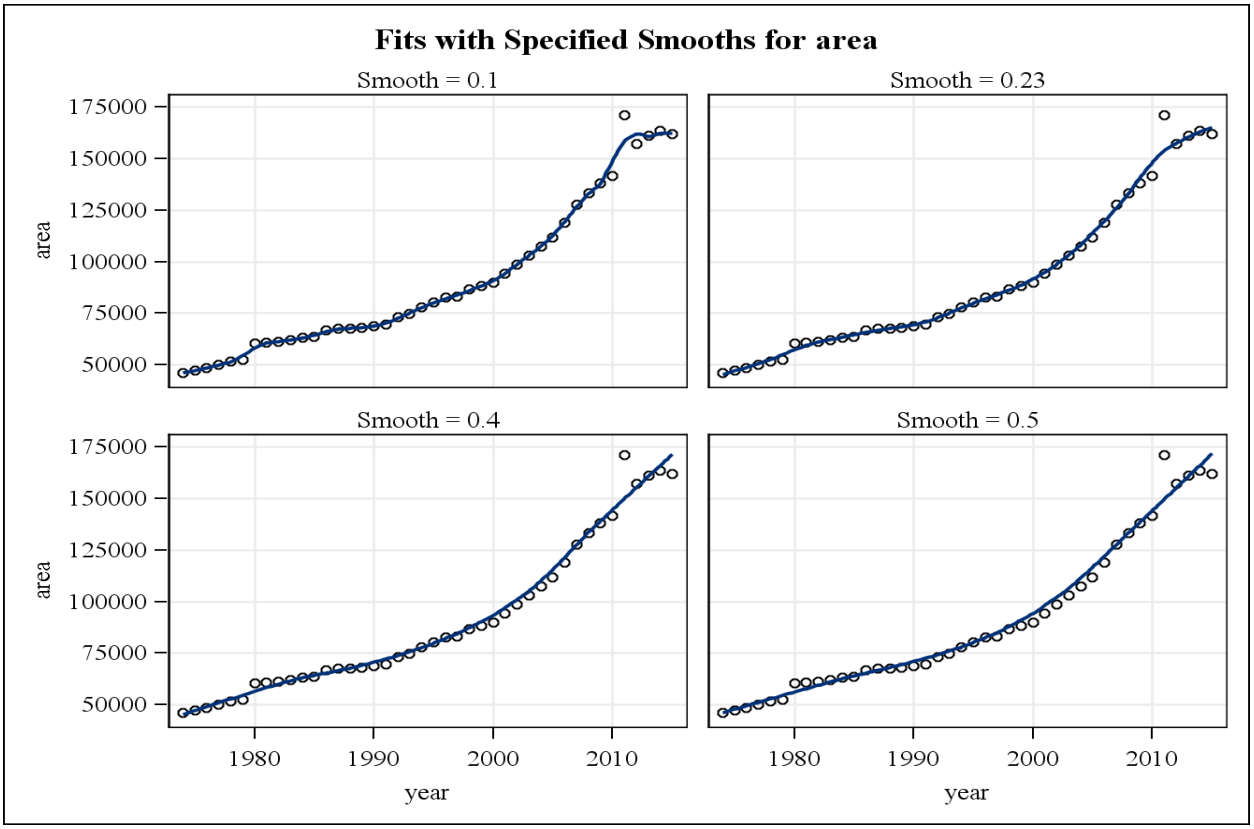
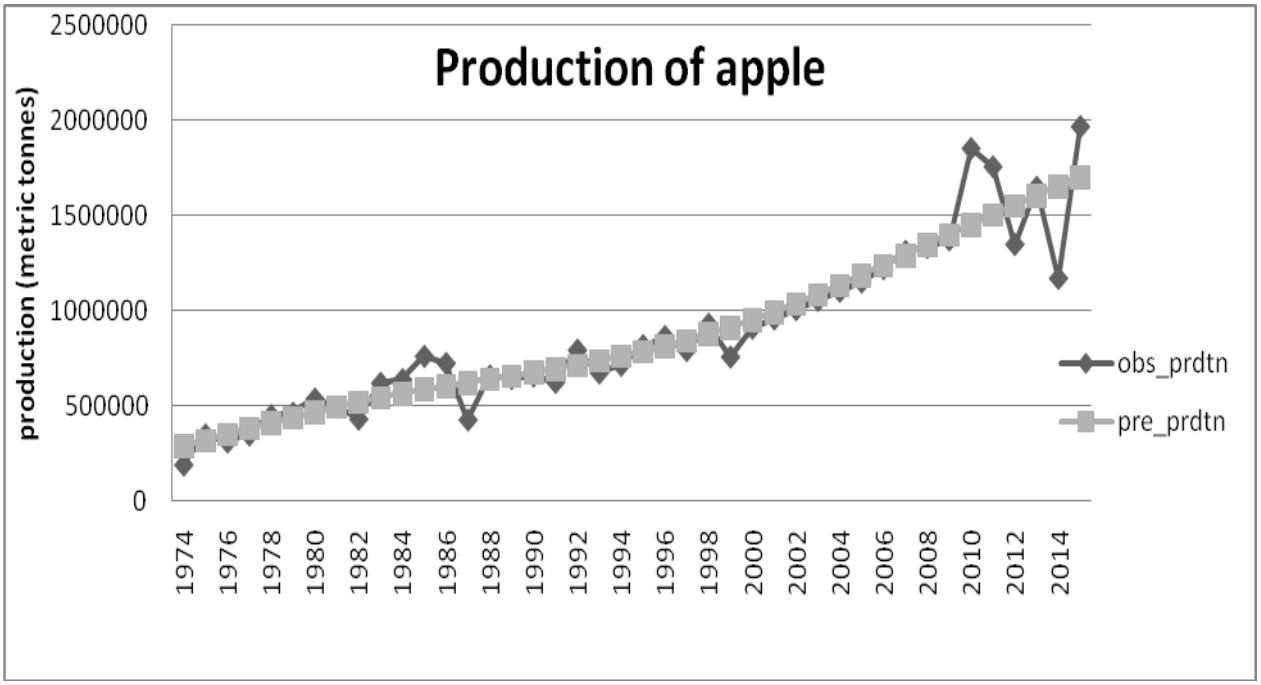


Fig.3 Observed and expected trends of production of apple using splines



obs_prdtn=observed production, pre_prdtn=predicted production

Fig.4 Fits with specified smooths for production

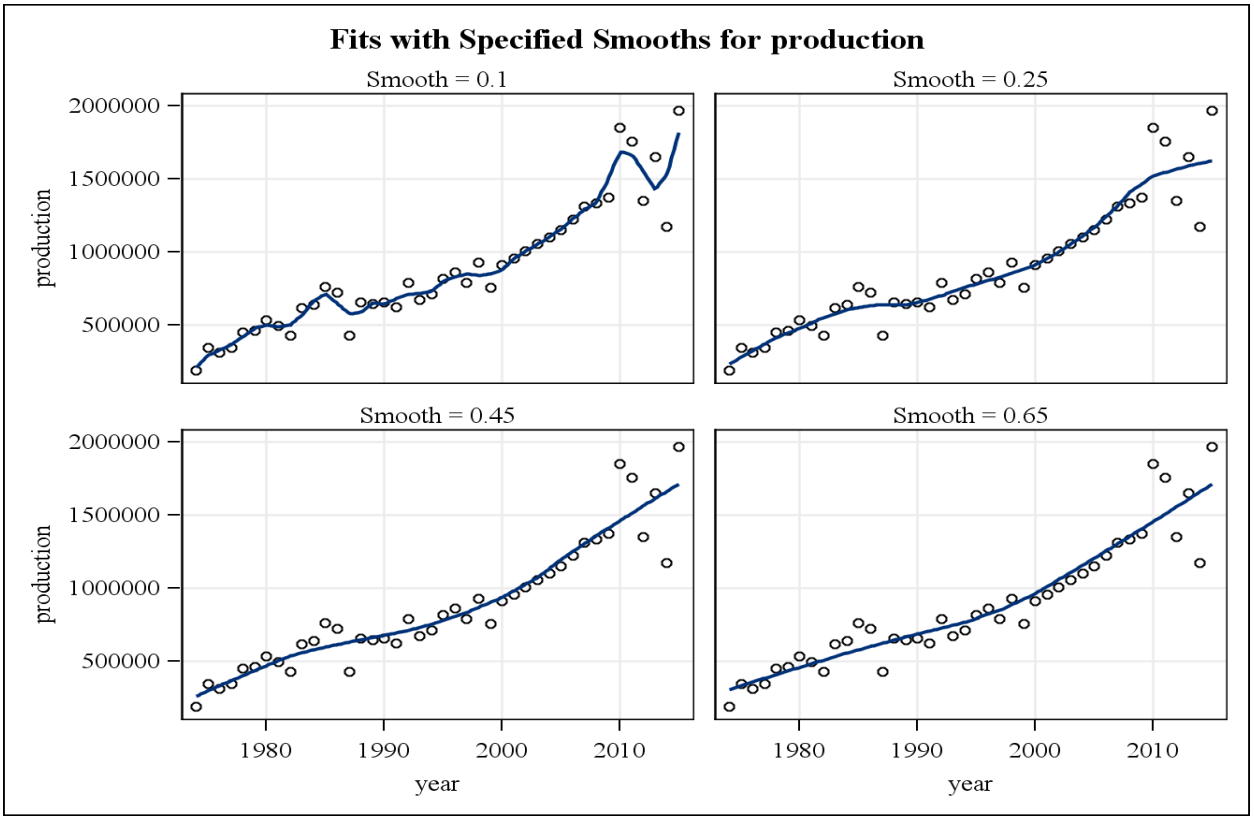
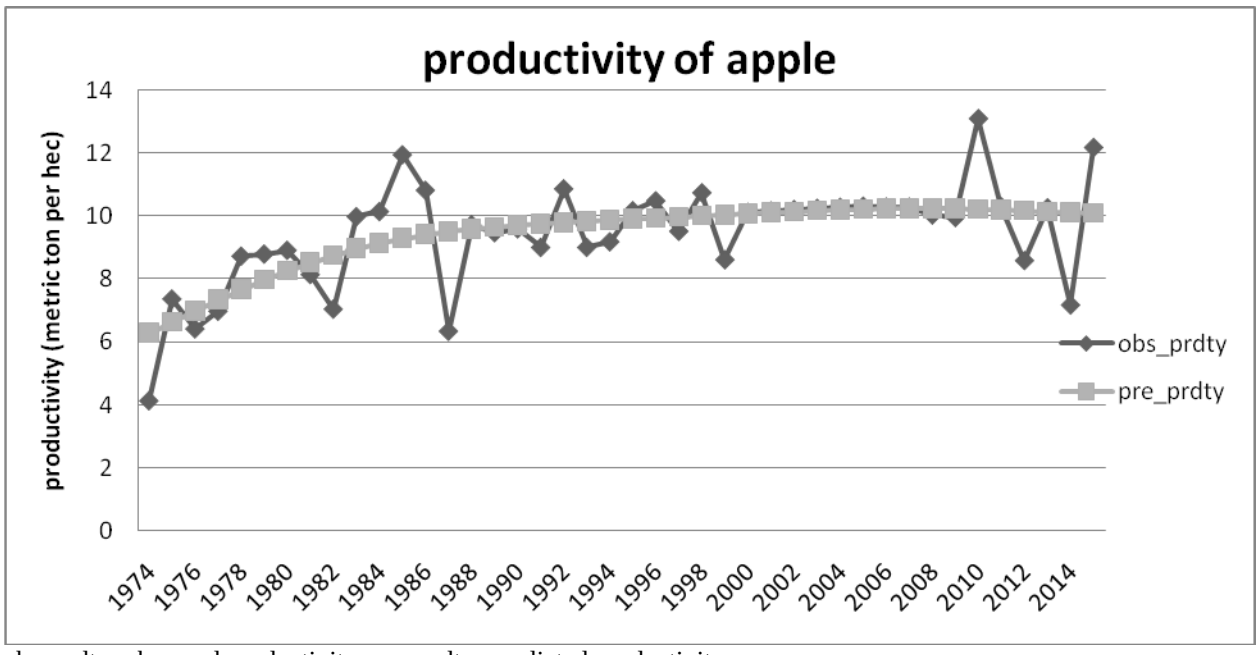
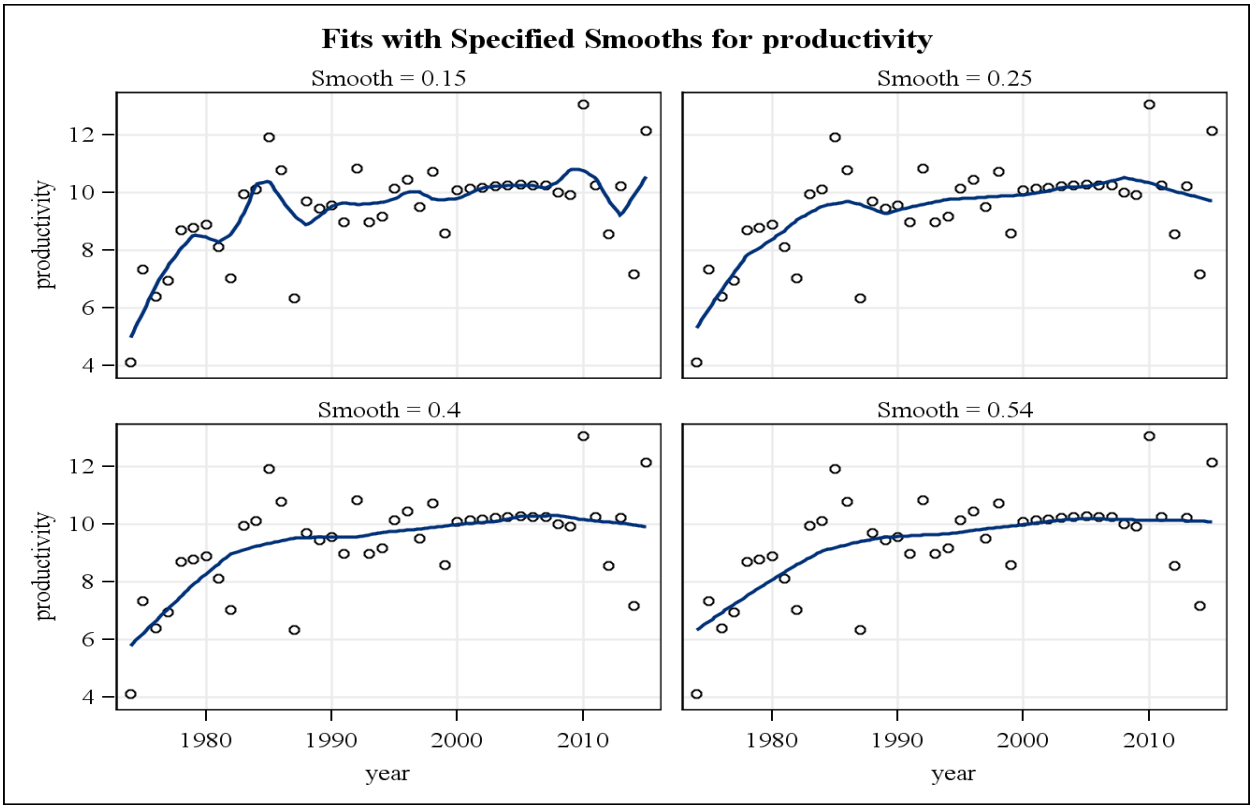


Fig.5 Observed and expected trends of productivity of apple using spline



obs_prdty=observed productivity, pre_prdty=predicted productivity

Fig.6 Fits with specified smooths for productivity



Nonparametric and semi-parametric regression

Trend analysis of area, production and productivity using nonparametric (Loess) and semi-parametric (spline) regression are presented in the tables 3, 4 and 5. In Table 3 the value of R^2 is 0.987 for Loess and 0.996 for Spline regression. The AIC_c, RMSE, MAPE, MAE, MaxAPE and MaxAE values comes out be small for Spline regression for the area under apple cultivation. The area under the apple cultivation has increased over the years of study and is shown in figure 1.

On comparing the values of AIC_c, RMSE, MAPE, MAE, MaxAPE and MaxAE for production and productivity the spline regression has the smallest values. According to the state's horticulture department, around 1.5 million tonnes of apples are produced in Kashmir annually. The production of apples in the state is growing every year as a result the percentage share of Jammu & Kashmir in national production has also been increasing steadily; it has increased from, 63.5% in FY2006 to 77.2% in FY2010. The apple production in the year 2004-05 was 10933.33 MT and in year it reached to 1852.41 in the year 2010-11 (Sheikh and Tripathi, 2013). The increasing trend in the production and productivity over the years of study is shown in the figure 3 and 6. It can be observed that upto 2014-15 there is sharp increase in production and productivity. However, a decline in production and productivity can also be observed during the year 2015-16 is observed which is due to the floods that occurred during the said year (Islam and Shrivastava, 2017).

The values of area are initially fitted at the smoothing parameters in order to obtain the best fit of the data points we obtain the graph

of the data points in the neighborhood of the smoothing parameters and look for the curve which covers all the points of the data. The one which covers maximum points is the best fit of the data points. In figure 2 the smooth curve fits are obtained for area in the neighborhood of smoothing parameters i.e., at 0.10, 0.23, 0.40 and 0.50. It is observed that the best fit is obtained at smooth=0.23. In figure 4 smooth fits for production are plotted in the neighborhood of the smoothing parameter at 0.10, 0.25, 0.45 and 0.65 and it is observed that the best fit obtained for smooth=0.45.

Figure 6 provides the fits for productivity in the neighborhood of the smoothing parameters i.e., at smooths equal to 0.15, 0.25, 0.40 and 0.54. The best fit is observed to be at the smooth=0.54

Even values of RMSE, MAE, MAPE, MaxAE and MaxAPE for area production and productivity of Jammu and Kashmir for non-parametric regression has observed lower values than the parametric regression (Tables 3, 4, 5). This is clear indication of the superiority of these techniques over the parametric models. These models perform very well in visualizing the past trends where the parametric models fails to.

Among the nonparametric and semi-parametric regression, the spline regression has shown the lowest values of AIC_c, RMSE, MAPE, MAE, MaxAPE and MaxAE for area, production and productivity of apple in Jammu and Kashmir hence spline regression is the best fitted model for apple production in Jammu and Kashmir (Fig. 5). Various scientist viz. Aydin (2007) and Pal (2011) observed similar results where the spline gave the better results than the Loess smoothing.

From the above study, it is observed that there is dramatic increase in the area under apple cultivation and in the production as well as productivity. In order to maintain the trend more and more land is to be brought under the apple cultivation. Parametric regression usually utilized in studying the trend seems not to perform better than the nonparametric and semi-parametric regression. And out of the nonparametric and semi-parametric regression methods the semi-parametric regression (spline) is the best fit for the trend analysis of the apple production of Jammu and Kashmir.

References

Aydin D 2007. A comparison of the nonparametric regression models using smoothing spline and kernel regression. *World Academy of Science, Engineering and Technology* 36: 253-57.

Chandran, K. (2004), *A Study on Some Aspects of Nonlinear Statistical Models in Agriculture,* Indian Agricultural Statistics Research Institute Indian Agricultural Research Institute: New Delhi.

Cleveland WS 1994. *The elements of graphing data (revised edition)*. Hobart Press, Summit, NJ, USA.

Dhekale, B., Sahu, P., Vishwajith, K., Mishra, P., and Narsimhaiah, L. (2017), "Application of Parametric and Nonparametric Regression Models for Area, Production and Productivity Trends of Tea (Camellia Sinensis) in India," *Indian Journal of Ecology*, 44, 192-200.

Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge university press.

Islam, R. T., and Shrivastava, S. (2017), A Study on Area, Production and Productivity of Apples in J&K from 2006-07 to 2015-16, *International Journal of Scientific Research and Management*, 5, 6513-6519.

Mishra, P., Sahu, P., Bajpai, P., and Nirnjan, H. (2012), Past Trends and Future Prospects in Production, and Export Scenario of Tea in India, *International Review of Business and Finance*, 4, 25-33.

Pal S 2011. *Modeling Pest Incidences in Some Agricultural Crops*, Thesis *Submitted to the Bidhan Chandra Krishi Viswavidyalaya under Department of* Agricultural Statistics

Sahu PK and Pal S 2004. *Modelling Pest Incidence - Nonparametric Approach*. Paper presented in 24th International Symposium on Forecasting, Sydeny

Sheikh, S. H., and Tripathi, A. (2013), Socio-Economic Conditions of Apple Growers of Kashmir Valley: A Case Study of District Anantnag, *International Journal of Educational Research and Technology*, 4, 30-39.

Simonoff, J. S. (2012), *Smoothing Methods in Statistics*, Springer Science & Business Media.

Teczan N 2010. *Analysis of the Factors Affecting Productivity Using Non Parametric Regression Method*. 10th Global Conference on Business & Economics, Rome, Italy

How to cite this article:

Yasmeena Ismail, S.A. Mir and Nageena Nazir, S.A. 2018. Utilization of Parametric and Nonparametric Regression Models for Production, Productivity and Area Trends of Apple (*Malus domestica*) in Jammu and Kashmir, India. *Int.J.Curr.Microbiol.App.Sci*. 7(04): 267-276. doi: https://doi.org/10.20546/ijcmas.2018.704.030