Int.J.Curr.Microbiol.App.Sci (2020) *9*(10): 75-84

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 9 Number 10 (2020)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2020.910.011

Modeling and Monitoring Walnut (*Juglans regia*) Area & Production Based on Parametric and Non-Parametric Regression Models

Nageena Nazir*, S. A. Mir and B. S. Dhekale

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

**Corresponding author*

Keywords

Walnut (*Juglans regia*), Vitamin E and A

Article Info

Accepted:
04 September 2020
Available Online:
10 October 2020

A B S T R A C T

Walnut (*Juglans regia* L.) occupies an important position in the horticulture industry of Jammu and Kashmir. It has the monopoly of producing excellent quality of walnuts contributing more than 90 per cent of Indian walnut production. The temperate climatic conditions favor its cultivation and offer Jammu and Kashmir an exceptional edge to super pass the other states in terms of walnuts. Being organic in nature (which is its USP), as no fertilizers or sprays are used on walnut plants and its yield, and high in nutrients with immense health benefits, Kashmiri walnut has seen growing demand and acceptability in the domestic and international market. The present study is an attempt to find past trends of walnut in Jammu and Kashmir using parametric, non parametric and semi-parametric regression methods. The performance of each method is compared using high value of R and low value of residual criteria. It is found that non parametric/semi parametric regression comes out to be a good fit for trend in walnut production in comparison to parametric regression. Even semi parametric spline is selected as the best fit model for trend analysis. It is inferred that the area under walnut cultivation in J&K is increasing from 1998-2017 and the productivity has also shown an increasing trend except for some years where the trend is found declining.

Introduction

Walnut (*Juglans regia* L.) is economically important tree species, highly valued for its timber and edible nuts (Pollegioni *et al.*, 2017). Walnut has the age of hundreds of years and it grows up to the height of 25 to 30 meters. Native to Central Asia, Mediterranean climate with cold winters and mild summers is most suitable for this dry fruit. A walnut tree is harvestable after 4 to 6 years and reaches its full productivity by 11 or 12 years of age (Simpson, 2016). Rich in Vitamin E and Vitamin A and a perfect source of Omega 3 is considered best for the human brain.

Jammu and Kashmir with its extensive potential of the production of temperate fruits supported by its vibrant geo-climatic conditions play marvelous role in the development of the walnut industry. The state is also well known for the production of numerous types of fruits like apple, cherry, pear, apricots and in dry fruits category it also produces walnuts (Taufique and *et al.,* 2017) hence takes the lead in the walnut production in India. Walnuts became the viable horticulture industry in India since 1980's particularly in the valley of Kashmir (Pandey and Shukla, 2007). Jammu and Kashmir occupy almost 90 per cent share of walnut industry in India (Mir *et al.,* 2016). Recently Jammu and Kashmir have been declared as an 'Agri-Export Zone' for 'walnuts' (Shah, R.A, 2017). Walnut produced in Jammu and Kashmir is purely organic as it is grown in the naturally induced conditions without the heavy doses of fertilizers and sprays. The aroma, flavor and taste of its excellent capacity make the walnut of Jammu and Kashmir one of the best in Indian and abroad as well.

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. Yasmeen *et al.,* (2018, 2019) employed non parametric and parametric regression models to study the trend of Area, Production and productivity of Apple and cherry in Kashmir. The current study is aimed to develop appropriate parametric and nonparametric regression models to fit the trends in area, production and productivity of Walnut in Kashmir.

Materials and Methods

For present study, to study the trends and growth rates, long term data for last 20 years pertaining to the area, production and productivity of Walnut 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

$$y_i = m(x_i) + e_i , x_i = i / n, i = 1,2,..., n$$

Where, y_i is observation of i^{th} time point, $m(.)$ is trend function which is assumed to be smooth and are e_i random errors with mean zero and finite variance. Since there is no assumption of parametric form of function, 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 i Y th i) (<mi) (<m by Cleveland (1979), is obtained on the basis of the data points around it within a band of certain width.

The point xi is the midpoint of the band. The data points within the band are assigned weights in a way so that xi has the highest weight. The weights for the other data points decline with their distance from xi according to a weight function. The weighted least squares method is used to find the fitted value corresponding to xi, 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 mis 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 $x_i \in [a,b]$ 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. 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, controls the emphasis which one wishes to place on smoothness. By increasing the value of, one places more emphasis on smoothness; as 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 walnut over the years, whereas the minimum growth rate is exhibited by area of the walnut.

The positive compound growth of production (0.078 per annum) reveals that there is no decrease in the production of walnut over the years with a maximum of 0.17 million kilogram and minimum of 0.04 million kilogram. Similarly, the simple growth rate (2.59 per annum) is observed in production indicates an increase in the production of walnut in 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 annum) under walnut cultivation indicating that a large portion of the land is being utilized for the latter.

Table.1 Performance of Walnut production in Kashmir during 1998-2017

Area ('000 hectare)		Production ('000 MT)		Productivity (Quintals per hectare)	
Minimum	32.65	**Minimum**	44.85	**Minimum**	13.26
Maximum	56.98	**Maximum**	169.42	**Maximum**	31.30
Mean	47.58	**Mean**	109.39	**Mean**	22.08
Standard Deviation	8.27	**Standard Deviation**	34.92	**Standard Deviation**	6.206
CV(%)	17.38	**CV(%)**	31.08	**CV(%)**	28.07
Skewness	-0.75	**Skewness**	-0.07	**Skewness**	0.146
SGRA	0.61	**SGRA**	2.59	**SGRA**	1.23
CGRA	.029	**CGRA**	.078	**CGRA**	.047

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 Walnut in Kashmir Division

	R-Square	Constant b0	b1	b2	b3	RMSE	MAPE	MAE	MaxAPE	MaxAE
Area	0.84	29.74	1.73	0.08	-0.005	4.37	0.46	1.25	6.59	3.3
Production	0.82	42.71	-0.2	1.05	0.03	12.42	0.78	4.55	25.28	39.23
Productivity	0.87	13.19	0.14	0.1	0.003	5.56	0.51	0.77	22.25	35.68

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

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

Loess	Splines	
Bandwidth	0.37	0.28
R-Square	0.91	0.94
AICc	1.70	1.23
RMSE	0.75	0.67
MAPE	0.11	0.10
MAE	0.75	0.70
MaxAPE	4.09	3.39
MaxAE	1.6	1.39

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

Loess	Splines	
Bandwidth	0.62	0.43
R2	0.93	0.98
AICc	6.17	3.15
RMSE	9.58	2.05
MAPE	0.246	0.11
MAE	2.55	1.18
MaxAPE	17.28	4.24
MaxAE	26.23	5.46

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

Loess	Splines	
Bandwidth	0.80	0.72
R2	0.90	0.96
AICc	2.76	1.98
RMSE	1.92	0.24
MAPE	0.23	0.08
MAE	0.11	0.39
MaxAPE	20.48	2.89
MaxAE	5.85	0.67

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

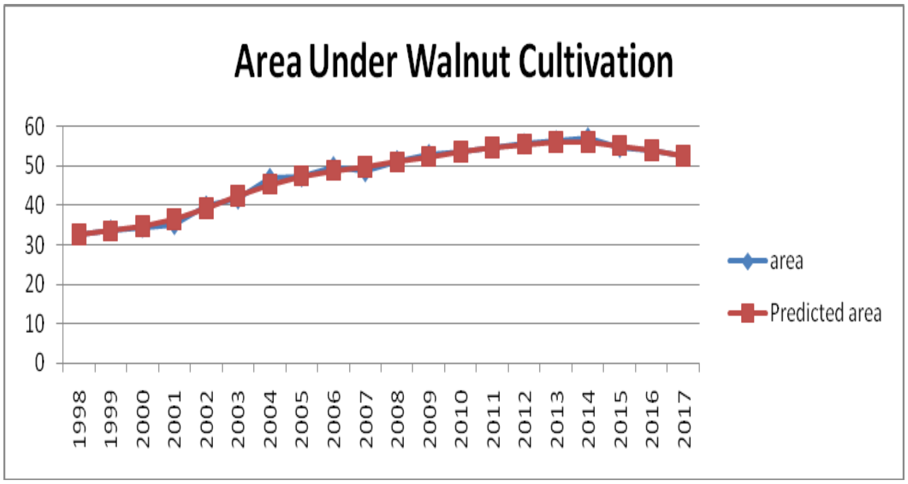


Fig.2 Fits with specified smooths for area

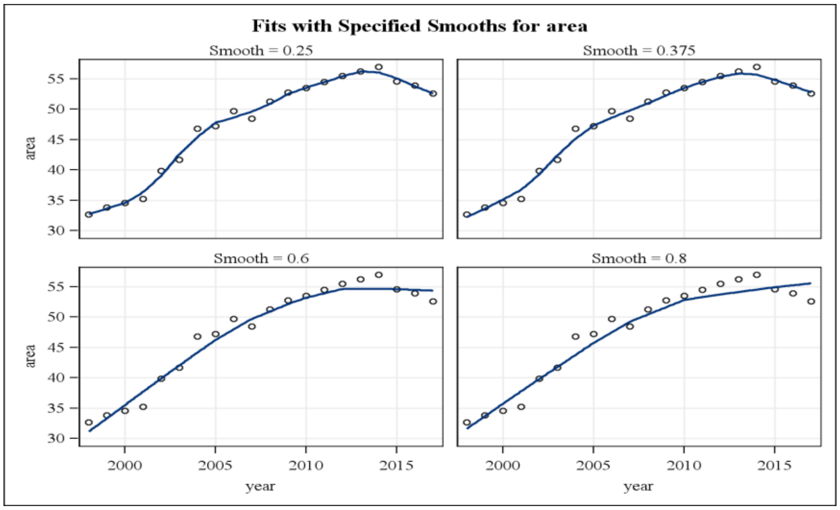


Fig.3 Observed and expected trends of Production under Walnut cultivation using spline in Kashmir

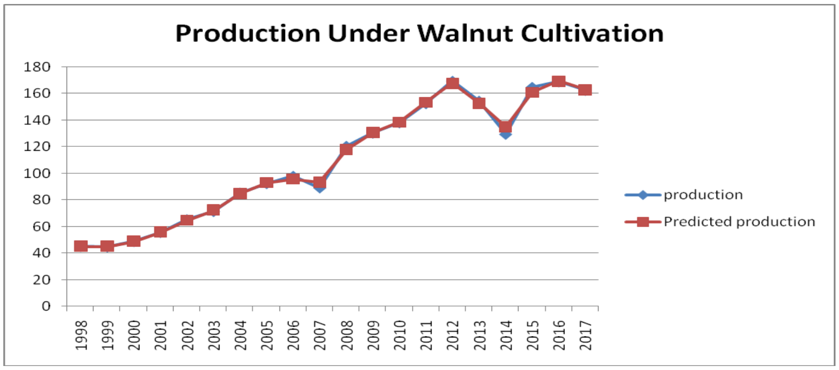


Fig.4 Fits with specified smooths for production

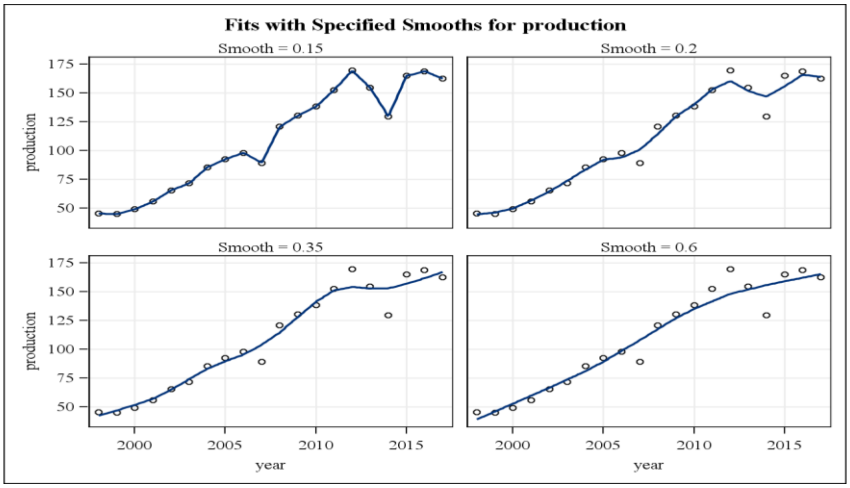


Fig.5 Observed and expected trends of Productivity under Walnut cultivation using spline in Kashmir

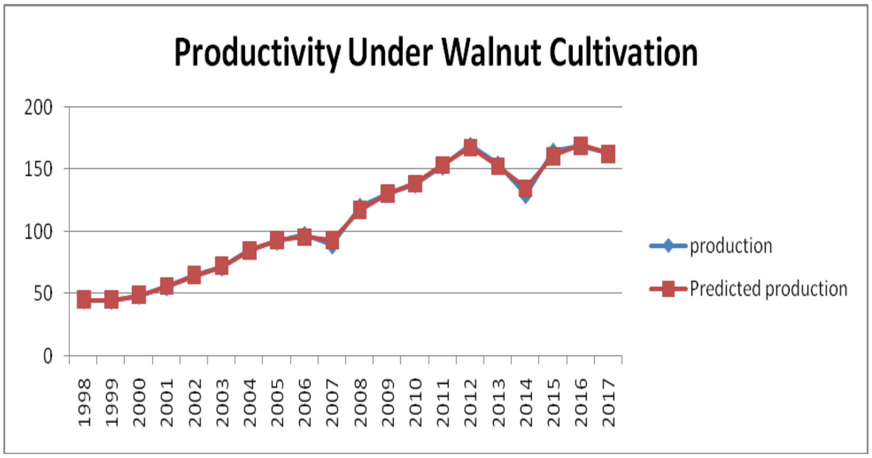
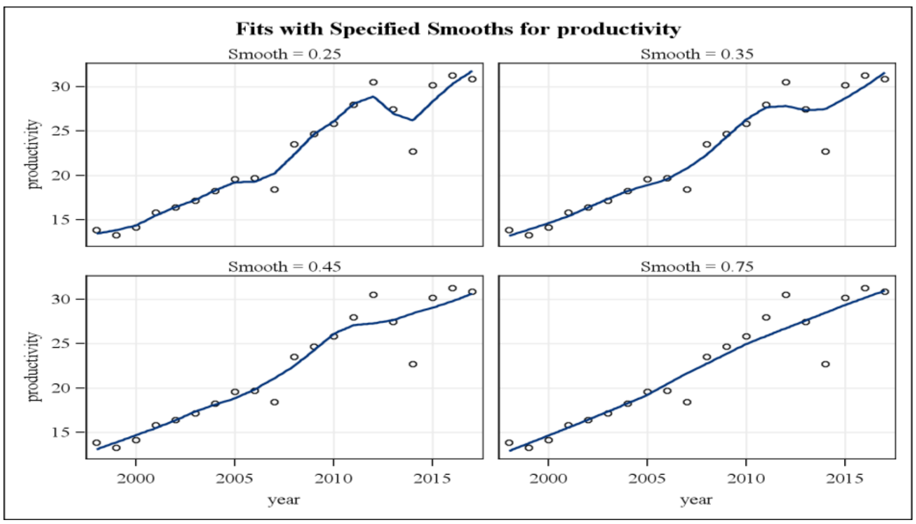


Fig.6 Fits with specified smooths for productivity



Trend analysis of area, production and productivity

Parametric techniques

The linear models used here are the cubic model or the third degree polynomial model and the quadratic model or second degree polynomial model. The value of *b3* for area is negative which indicates that area under walnut cultivation decreased in the last part of the cultivation period and the value of *b1* and *b2* being positive clearly indicates that there was an increase in the cultivation area. Further, the negative value of *b1* for production is an indication of the decrease in the production during the initial period of the study and the positive values of *b2* and *b3* indicates an increase in the production.

Non-parametric 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 R2 is 0.91 for Loess and 0.94 for Spline regression. The AICc, RMSE, MAPE, MAE, MaxAPE and MaxAE values comes out be small for Spline regression for the area under walnut cultivation. The area under the walnut cultivation has increased over the years of study and is shown in figure 1.

On comparing the values of AICc, RMSE, MAPE, MAE, MaxAPE and MaxAE for production and productivity the spline regression has the smallest values. The increasing trend in the production and productivity over the years of study is shown in the figure 3 and 5. It can be observed that upto 2013-14 there is sharp increase in production and productivity. However, a decline in production and productivity can also be observed during the year 2014-15 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 0.25, 0.375, 0.60 and 0.80. It is observed that the best fit is obtained at smooth=0.375. In figure 4 smooth fits for production are plotted in the neighborhood of the smoothing parameter at 0.15, 0.20, 0.35 and 0.62 and it is observed that the best fit obtained for smooth=0.62. Figure 6 provides the fits for productivity in the neighborhood of the smoothing parameters i.e., at smooths equal to 0.25, 0.35, 0.45 and 0.75. The best fit is observed to be at the smooth=0.75

Even values of RMSE, MAE, MAPE, MaxAE and MaxAPE for area production and productivity of 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 AICc, RMSE, MAPE, MAE, MaxAPE and MaxAE for area, production and productivity of walnut in Kashmir hence spline regression is the best fitted model for walnut production in 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.

In the above study, three types of modeling are discussed parametric, semipara-metric, nonparametric modeling. Nonparametric and semi-parametric regression models are flexible compared to parametric models. Semi-parametric is hybrid of both parametric and nonparametric which allow to have the best of both worlds: a model that is understandable and offering a fair representation of the data in the real life. However, semi/nonparametric regression requires larger sample sizes than regression based on parametric models because the data must supply the model structure as well as the model estimates (Mahmoud, 2019). From the above study it is observed that there is dramatic increase in the area under walnut 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 walnut 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 walnut production of 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.

Bryon J. Parman and Allen M. Featherstone 2019. A comparison of parametric and nonparametric estimation methods for cost frontiers and economic measures Journal of Applied Economics Volume 22, 2019 - Issue 1

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.

G.M.Mir, Uzma Iqbal and Owais Nisa (2016). Portable Walnut Grading Machine Best Replacement for perforated Cylindrical Grader. *International Journal of Research and Innovation Vol. 3(8), 82-86.*

G. Pandey and S. K. Shukla (2007). The Walnut Industry in India-Current Status and Future Prospects. *International Journal of Fruit Science, 6:4, 67-75*

Hamdy F. F. Mahmoud 2019. Parametric versus Semi and Nonparametric Regression Models. *arXiv:1906.10221v1 [stat.ME]*.

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 from2006-07 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.

Mohammad Taufique, Vajahat Khursheed (2018). Walnut Industry in Jammu and Kashmir: A Geographical Analysis. *Res. J. Humanities and Social Sciences. 2018; 9(4): 793-798.*

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

Paola Pollegioni, Keith Woeste, Francesca Chiocchini, Stefano Del Lungo, Marco Ciolfi, Irene Olimpieri, Virginia Tortolano, Jo Clark, Gabriel E. Hemery, Sergio Mapelli, Maria Emilia Malvolti (2017). Rethinking the history of common walnut (*Juglans regia* L.) in Europe: Its origins and human interactions https://doi.org/10.1371/journal.pone.0172541.

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

Shah U.N, Mir J.I, Ahmed N, Fazili K.M. 2016. Assesment of Germplasm

Diversity and Genetic Relationships among walnut Genotypes through Microsatellite Markers. J.Saudi Soc. Agri. Sci., 1-12.

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 and Business Media.

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

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.

Yasmeena Ismail, S.A. Mir and Nageena Nazir, M.H. Wani, S.A. Wani and M.S. Pukhta 2019. Trend Analysis of Area, Production and Productivity of Cherry in Jammu and Kashmir. *Int.J.Curr.Microbiol.App.Sci.* *8*(2): 2135-2144.

How to cite this article:

Nageena Nazir, S. A. Mir and Dhekale, B. S. 2020. Modeling and Monitoring Walnut (Juglans Regia) Area and Production Based on Parametric and Non-Parametric Regression Models. *Int.J.Curr.Microbiol.App.Sci.* 9(10): 75-84. doi: https://doi.org/10.20546/ijcmas.2020.910.011