The Bioscan 13(2): 277-280, 2018(Supplement on Agronomy)
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
www.thebioscan.com

APPLE YIELD ESTIMATION ON THE BASIS OF MORPHOLOGICAL CHARACTERS: A STATISTICAL APPROACH

ANJU SHARMA*, P. K.MAHAJAN, O. K.BELWAL AND SATISH K. SHARMA
Dr Y S Parmar University of Horticulture and Forestry, Nauni, Solan - 173 230 - HP, INDIA
e-mail: anjusharma_uhf@yahoo.com

KEYWORDS
Apple
Correlation
Forecasting
Linear models

Received on :
22.03.2017

Accepted on :
13.01.2018

***Corresponding author**

ABSTRACT
An attempt has been made to develop models for yield estimation of apple on the basis of morphological characters through correlation and regression techniques at three different locations of Himachal Pradesh viz., Shimla, Kotkhai and Theog. Highest correlation coefficient was obtained between number of spurs and number of flowers (0.98), followed by between age and girth (0.81) at Shimla location. Number of fruits, flowers and spurs were perfectly correlated with each other in Kotkhai and Theog regions. The use of age (exponential), girth (exponential) and spread (cubic) was found to be the best independent variables for estimating the apple yield in Shimla, Kotkhai and Theog locations respectively. Age and length of spurs had significant effect on yield estimation at Shimla location while spread of the tree and length of spurs contributed significantly to the yield of tree at Kotkhai. In Theog location, age, height, number of secondary branches and length of spurs contributed significantly to the yield estimation. The model, $X_{12} = b_0 X_1^{b1} X_2^{b2} X_3^{b3} X_4^{b4} X_5^{b5} X_6^{b6} X_7^{b7} X_8^{b8} X_9^{b9} X_{10}^{b10} X_{11}^{b11}$ fitted best with 0.560 and 0.925 as adjusted R^2 value respectively at Shimla and Kotkhai locations, while $X_{12} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11}$ was found to be best fit at Theog location with 0.956 as adjusted R^2 value.

INTRODUCTION

Among horticultural crops of Himachal Pradesh, apple (*Malus domestica*), grown in temperate and dry temperate zones, is so far the most dominating leading cash fruit crop and accounts for about 49 per cent of total area under fruit crops (2,29,202 hectares) and more than 76 per cent of the total fruits production (6,11,877 MT). The state is well recognized as the 'Apple State of India'. The area under apple and its production has increased from 88,673 hectares and 49,129 MT in 1999-00 to 1,11,896 hectares and 4,68,134 MT in 2016-17 (Anonymous, 2017), respectively.

In Himachal Pradesh, a number of apple growers are absentee landlords and the common practice is to sell the crop about 30 to 40 days before actual harvesting for which expected yield estimates are very important. Proper handling and marketing of the fruits at the state level also demand such estimates of the expected fruit crop yield. Forecasting the yield is of great significance to the orchardists to prepare an estimate for harvest management, auction, transportation, storage, marketing and processing etc. However, in the absence of a proper systematic and scientific procedure for yield forecasting, it becomes difficult to make proper assessment of the yield estimates. A need was, therefore, felt to have probabilistic approach in crop yield forecast and develop an objective model to forecast the apple yield by using existing data to predict future outcomes with a predetermined degree of error built into the model (Thomas, 1998).

Model building is currently applied in many fields *i.e.* Agriculture, Biometrics, Econometrics, Education, Meteorology, Industry, Horticulture and Forestry etc. (Prasad, 2010). Costas *et al.* (2006) obtained yield prediction models by using pruning variables for *Pinus taeda* based on data from an experience in a *Pinus taeda* plantation, including ages between 4 and 7 years old, multiple linear regressions were adjusted to obtain stand yield functions including regression variables associated with prunes intensity and opportunity, besides age, stand density and site index. Raizada *et al.* (2007) developed biomass prediction models for 17 year old *Acacia nilotica* trees raised on salt affected vertisols of the semi and tropics in Karnataka, India. Rizvi *et al.* (2006) developed prediction models for timber weight of *Populus deltoids* planted on farmland in Haryana. They estimated the fresh green timber of poplar tree and evaluated growth process based non-linear models for fresh timber weight. Lamien *et al.* (2007) conducted a study to develop fruit yield prediction models based on dendrometric and fruiting variables, to examine variations in these variables between upland and lowland populations (agroforestry parklands) in Burkina Faso, and associations between these variables. All fruiting variables, number of shoots and crown attributes had the highest influence on the models. Models have been proposed to estimate blueberry yield using regression models that relate the number of fruits to the number of flower buds and to climatic variables (Salvo *et al.*, 2011).

The present study was carried out to identify those morphological characters which play an appreciable role in increasing the yield of apple and thereafter an attempt to develop appropriate model for estimation of apple yield based on important morphological characters using correlation and regression techniques has been made.

MATERIALS AND METHODS

Data recorded on various morphological characters viz.,

X_1 = Age (years), X_2 = Girth (m), X_3 = Height (m), X_4 = Spread (m), X_5 = Number of main branches, X_6 = Number of secondary branches, X_7 = Number of spurs per tertiary branch, X_8 = Length of spurs, X_9 = Number of flowers per tertiary branch, X_{10} = Number of fruits per tertiary branch, X_{11} = Fruit weight (g) and X_{12} = Yield of tree (kg) from 300 trees at three different apple growing locations (100 trees at each location) viz., Shimla, Kotkhai and Theog locations (at different elevations) has been used for estimating yield using correlation and regression techniques (Panse and Sukhatme, 1978). The yield (Y) was estimated by fitting following regression models at three different locations by taking each character (X) separately.

Linear : $Y = b_0 + b_1 X$

Logarithmic : $Y = b_0 + b_1 \ln X$

Quadratic : $Y = b_0 + b_1 X + b_2 X^2$

Cubic : $Y = b_0 + b_1 X + b_2 X^2 + b_3 X^3$

Exponential : $Y = b_0 e^{b_1 X}$

The following multiple linear and non-linear regression models were tried to study multiple linear and non - linear regression taking yield (X_{12}) as dependent variable and other characters (X_1 to X_{11}) as independent variables to study the combined effect of different tree growth parameters on apple yield for different locations.

Linear Model

Model 1: $X_{12} = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 + b_6 X_6 + b_7 X_7 + b_8 X_8 + b_9 X_9 + b_{10} X_{10} + b_{11} X_{11}$

Non Linear Models

Model 2: $X_{12} = b_0\ X_1^{b1}\ X_2^{b2}\ X_3^{b3}\ X_4^{b4}\ X_5^{b5}\ X_6^{b6}\ X_7^{b7}\ X_8^{b8}\ X_9^{b9}\ X_{10}^{b10}\ X_{11}^{b11}$

Model 3: $X_{12} = b_0\ b_1^{X1}\ b_2^{X2}\ b_3^{X3}\ b_4^{X4}\ b_5^{X5}\ b_6^{X6}\ b_7^{X7}\ b_8^{X8}\ b_9^{X9}\ b_{10}^{X10}\ b_{11}^{X11}$

RESULTS AND DISCUSSION

Correlation among different agronomic and morphological characters is an important aspect for better planning of selection programs and is also helpful in determining the components of complex traits like yield. Therefore, in the selection process for crop improvement, knowledge of association of various characters is the most important tool (Desai *et al.*, 1994). Karl Pearson's correlation coefficient was worked out between apple yield and morphological characters and is presented in

Table 1: Simple correlation coefficient matrix indicating relationship between yield and various tree growth characteristics at location1: Shimla

Parameters	Age X_1	Girth X_2	Height X_3	Spread X_4	No. of main branches X_5	No. of secondary branches X_6	No. of spurs X_7	Length of spurs X_8	No. of flowers X_9	No. of fruits X_{10}	Fruit weight X_{11}	Yield X_{12}
Age (X_1)	1											
Girth (X_2)	0.81**	1										
Height (X_3)	0.42**	0.48**	1									
Spread (X_4)	0.68**	0.76	0.58**	1								
No. of main branches (X_5)	0.52**	0.55**	0.48**	0.68**	1							
No. of secondary branches (X_6)	0.1	0.09	0.29**	0.09	-0.07	1						
No. of spurs (X_7)	0.12	0.14	-0.01	0.27**	0.19	-0.21	1					
Length of spurs (X_8)	0.09	0.07	0.11	0.14	0.07	0.17	-0.15	1				
No. of flowers (X_9)	0.07	0.1	-0.01	0.22	0.17	-0.19	0.98**	-0.13	1			
No. of fruits (X_{10})	0.09	0.16	-0.12	0.05	0.09	-0.18	0.76**	-0.29	0.77**	1		
Fruit weight (X_{11})	0.33**	0.21	0.27**	0.30**	0.2	0.16	0.12	0.21	0.1	0.04	1	
Yield (X_{12})	0.54**	0.47**	0.36**	0.52**	0.31**	0.1	0.22	0.22	0.18	0.14	0.23	1

** Significant at 1% level of significance

Table 2: Simple correlation coefficient matrix indicating relationship between yield and various tree growth characteristics at location 2: Kotkhai

Parameters	Age X_1	Girth X_2	Height X_3	Spread X_4	No. of main branches X_5	No. of secondary branches X_6	No. of spurs X_7	Length of spurs X_8	No. of flowers X_9	No. of fruits X_{10}	Fruit weight X_{11}	Yield X_{12}
Age (X_1)	1											
Girth (X_2)	0.77**	1										
Height (X_3)	0.66**	0.67**	1									
Spread (X_4)	0.53**	0.75**	0.79**	1								
No. of main branches (X_6)	0.26**	0.46**	0.22	0.40**	1							
No. of secondary branches (X_7)	0.53**	0.82**	0.59**	0.69**	0.55**	1						
No. of spurs (X_8)	0.39**	0.56**	0.08	0.29**	0.35**	0.50**	1					
Length of spurs (X_9)	0.51**	0.72**	0.44**	0.54	0.50**	0.62**	0.50**	1				
No. of flowers (X_{10})	0.39**	0.56**	0.08	0.29**	0.35**	0.50**	1.00**	0.50**	1			
No. of fruits (X_{11})	0.41**	0.58**	0.11	0.31**	0.36**	0.51**	1.00**	0.51**	1.00**	1		
Fruit weight (X_{12})	0.21	0.15	0.18	0.13	0.04	0.13	0.18	0.21	0.18	0.2	1	
Yield (X_{12})	0.55**	0.81**	0.62**	0.74**	0.51**	0.73**	0.48**	0.81**	0.48**	0.50**	0.2	1

** Significant at 1% level of significance

Table 3: Simple correlation coefficient matrix indicating relationship between yield and various tree growth characteristics at location 3: Theog

Parameters	Age	Girth	Height	Spread	No. of main branches	No. of secondary branches	No. of spurs	Length of spurs	No. of flowers	No. of fruits	Fruit weight	Yield
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
Age (X_1)	1											
Girth (X_2)	0.96**	1										
Height (X_3)	0.83**	0.89**	1									
Spread (X_4)	0.95**	0.95**	0.84**	1								
No. of main branches (X_5)	0.27**	0.11	0.14	0.23	1							
No. of secondary branches (X_6)	0.88**	0.86**	0.72**	0.89**	0.32**	1						
No. of spurs (X_7)	0.60**	0.58**	0.39**	0.68**	0.19	0.63**	1					
Length of spurs (X_8)	0.02	0.05	0.06	0.01	-0.16	-0.11	-0.05	1				
No. of flowers (X_9)	0.60**	0.58**	0.39**	0.68**	0.19	0.63**	1.00**	-0.05	1			
No. of fruits (X_{10})	0.62**	0.60**	0.42**	0.70**	0.21	0.65**	0.99**	-0.06	0.99**	1		
Fruit weight (X_{11})	0.2	0.17	0.12	0.2	0.15	0.26**	0.13	-0.1	0.13	0.16	1	
Yield (X_{12})	0.91**	0.91**	0.83**	0.92**	0.2	0.78**	0.76**	0	0.76**	0.78**	0.14	1

** Significant at 1% level of significance

Table 4: Best fitted function for yield with various tree growth characteristics

Parameters	Best Function and R^2 (SE of estimate)		
	Location 1: Shimla	Location 2: Kotkhai	Location 3: Theog
Age (X_1)	Exponential 0.444 (1.037)	Quadratic 0.555 (86.940)	Exponential 0.858 (0.518)
Girth (X_2)	Exponential 0.352 (1.120)	Exponential 0.774 (0.673)	Cubic 0.849 (27.717)
Height (X_3)	Exponential 0.266 (1.191)	Cubic 0.608 (82.046)	Cubic 0.833 (29.090)
Spread (X_4)	Exponential 0.329 (1.139)	Exponential 0.586 (0.9111)	Cubic 0.869 (25.804)
No. of main branches (X_5)	Exponential 0.126 (1.300)	Cubic 0.280 (111.198)	Quadratic 0.145 (65.511)
No. of secondary branches (X_6)	Quadratic 0.038 (79.895)	Exponential 0.552 (0.948)	Exponential 0.760 (0.674)
No. of spurs (X_7)	Quadratic 0.092 (77.655)	Quadratic 0.472 (94.727)	Quadratic 0.711 (38.094)
Length of spurs (X_8)	Cubic 0.083 (78.037)	Quadratic, Cubic 0.666 (75.306)	Quadratic, Cubic 0.001 (70.801)
No. of flowers (X_9)	Quadratic 0.077 (78.282)	Quadratic 0.472 (94.727)	Quadratic 0.710 (38.184)
No. of fruits (X_{10})	Cubic 0.062 (79.320)	Quadratic 0.466 (95.267)	Quadratic 0.711 (38.100)
Fruit weight (X_{11})	Quadratic 0.098 (77.362)	Logarithmic 0.042 (126.979)	Exponential 0.045 (1.343)

Figures in parentheses denotes the SE of estimate

Tables 1, 2 and 3 for each location. The yield was positively and significantly correlated with all tree growth parameters in all locations under study. Among different tree growth parameters, highest correlation coefficient was obtained between number of spurs and number of flowers (0.98), followed by between age and girth (0.81) at Shimla location (Table1). Number of fruits, flowers and spurs were perfectly correlated with each other in Kotkhai and Theog regions. Least correlation was noticed between fruit weight and number of main branches (0.04) at Kotkhai and between length of spurs and number of main branches at Theog location. This suggests that there was strong inherent relationship between the traits contributing towards the yield. The correlation and regression techniques have been applied to several plant species such as kinnow (Brar and Chander Mohan, 1985), and guava (Lakpathi *et al*., 2013) in order to estimate their yield and production.

On the basis of R^2 value, the use of age, girth and spread remained the best independent variables for estimating the apple yield in Shimla, Kotkhai and Theog locations respectively (Table 4) (Raizada *et al*., 2007). Exponential function was best fitted for tree age to predict the yield of tree, significantly, with Coefficient of Determination (R^2) value maximum (0.444) followed by exponential function having R^2 value of 0.352, using girth as independent variable, at Shimla location. At Kotkhai location, exponential function was best fitted (R^2 = 0.774) for tree girth to predict the yield of tree. Using length of spurs as independent variable, the quadratic and cubic functions exhibited the same significant value for R^2 (0.666). The cubic function was fitted best to predict the yield while taking spread as independent variable with maximum R^2 (0.869) at Theog location.

Table 5: Multiple linear and non-linear regression analysis for yield

Location	Regression model	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	$\bar{R}^2$
Shimla	Linear	-492.722	5.002*	-101.019	14.643	31.003	-6.998	0.029	4.797	310.722*	-1.446	2.231	-0.105	0.333
	Model 1	-170.497	-1.818	-119.584	-9.604	-18.472	-5.918	-1.933	-5.753	-129.376	-1.779	-1.746	-0.21	
	Non linear	0.034	2.058*	0.144	1.870*	0.241	-0.614*	-0.062	0.713	4.242	-0.721	0.364	-0.3	0.56
	Model 2	-2.525	-0.47	-0.722	-0.524	-0.589	-0.31	-0.319	-0.776	-2.391	-0.753	-0.34	-0.354	
	Non linear	0.006	1.113*	0.499	1.650*	1.334	0.857	0.994	1.091	46.307*	0.97	1.039	0.999	0.508
	Model 3	-2.512	-0.027	-1.762	-0.142	-0.272	-0.087	-0.028	-0.085	-1.906	-0.026	-0.026	-0.003	
Kotkhai	Linear	-534.952	-1.268	125.645	2.22	4.216*	5.823	0.668	30.781	306.916*	-	-8.926	0.215	0.798
	Model 1	-76.577	-0.962	-68.246	-1.675	-1.778	-5.16	-2.458	-41.357	-49.163		-13.492	-0.366	
	Non linear	0.366	0.111	0.922*	0.259	0.233	0.037	0.162	-	1.873*	-0.922*	1.317*	0.567	0.925
	Model 2	-2.049	-0.177	-0.263	-0.167	-0.165	-0.119	-0.2		-0.499	-0.35	-0.32	-0.367	
	Non linear	0.484	1.01	3.390*	1.026	1.031*	0.983	1.015	0.205*	4.889*	-	1.769*	1.002	0.885
	Model 3	-0.631	-0.008	-0.562	-0.014	-0.015	-0.043	-0.02	-0.341	-0.405		-0.111	-0.003	
Theog	Linear	49.775	3.125*	37.842	7.713*	1.691	0.844	-3.234*	1.685	-55.653*	0.456	1.356	-0.077	0.956
	Model 1	-42.357	-0.805	-29.416	-1.294	-2.524	-2.16	-0.568	-26.835	-24.944	-5.364	-1.184	-0.096	
	Non linear	0.934	0.901*	0.003	0.509*	0.18	0.278	-0.054	1.008	0.087	-0.3	0.064	-0.165	0.953
	Model 2	-2.485	-0.236	-0.378	-0.222	-0.205	-0.196	-0.225	-1.032	-0.726	-1.024	-0.257	-0.275	
	Non linear	1.305	1.129*	0.352	1	1.014	1.051	1.024	1.093	1.559	0.995	1.018	1	0.94
	Model 3	-0.957	-0.018	-0.665	-0.029	-0.057	-0.049	-0.013	-0.606	-0.564	-0.121	-0.027	-0.002	

* Significant at 5% level of significance; Values in parentheses are standard error of estimate

Table 5 shows the fitting of multiple linear and non-linear regression models. Fitting of linear equation reveal that variables like age and length of spurs had significant effect on yield estimation at Shimla location while spread of the tree and length of spurs contributed significantly to the yield of tree at Kotkhai. In Theog location, age, height, number of secondary branches and length of spurs contributed significantly to the yield estimation. $\bar{R}^2$values showed that the non-linear model analyses (Kumar and Panwar, 2003; Shastry *etal.*,2017), $X_{12}=b_0 X_1^{b1}X_2^{b2}X_3^{b3} X_4^{b4}X_5^{b5}X_6^{b6}X_7^{b7}X_8^{b8}X_9^{b9} X_{10}^{b10} X_{11}^{b11}$ fitted best at Shimla ($\bar{R}^2 = 0.560$) and Kotkhai ($\bar{R}^2 = 0.925$) locations, while the linear model, $X_{12}=b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11}$ was found to be best fit at Theog location with $\bar{R}^2$ =0.956.

The outcome of the prediction models will assist horticultural agencies in providing orchardists with valuable information as to which factors contribute to high apple yield. Therefore, using these models research workers can obtain more precise fruit yield predictions, thereby providing valuable inferences in increasing the apple production and optimizing the economy of the State (Himachal Pradesh).

REFERENCES

Anonymous. 2017. Directorate of Horticulture, Shimla, Himachal Pradesh, India.

Brar, W. S. and Chander Mohan. 1985. Correlation and regression studies in kinnow mandarin. *South Indian Hort*. **33(5):** 323-325.

Costas, K., Mac-Donagh, P., Weber, E., Figuerodo, S., Gomez, C. and Irschik, P. 2006. Yield prediction models using pruning variables for *Pinustaeda*. *Borque*. **27(2):** 98-107.

Desai, U. T., Musade, A. M., Ranpise, S. A., Choudhari, S. M. and Kale, P. N. 1994. Correlation studies in acid lime. *J. Maharashtra Agric. Univ.***19(1):** 162-163.

Lamien, I. N., Tigabu, M., Guinko, S. and Oden, P. C. 2007. Variation in Dendromatric and fruiting characters of *Vitellari paradoxa* population and multivariate models for estimation of fruit yield. *Agro-For. Sys.* **69(1):** 1-11

Lakpathi, G., Rajkumar, M., Chandrasekhar, R, Kiran Kumar A. and Ramesh, T. 2013. Correlation Studies in Guava (*PsidiumguajavaL.*) cv. Allahabad Safeda. *International J. Recent Scientific Research.***4(12):** 1942-1945.

Kumar, A. and Panwar, S. 2003. Non-linear Growth Modelling of Apple productivity of different States in India. *J. Agriculture Sciences*. **74(5):** 418-423.

Panse, V. G. and Sukhamate, P. V. 1978. Statistical methods for agricultural workers, third revised edition, Indian Council of Agricultural Research, New Delhi. p. 359.

Prasad, H. 2010. Statistical modeling technique on export of fruit crops in India. *Indian J. Agricultural Sciences*. **74(5):** 619-625

Raizada, A., Rao, M. S. R. M., Nambiar, K.T.N. and Padmaiah, H. 2007. Bio-mass production and prediction models for *Acacia niloticain* salt affected vertisols in Karnataka. *Indian For.* **133(2):** 239-246

Rizvi, R. H. and Khare, Diwakar. 2006. Prediction model for timber weight of *Populus deltoids* planted on farmlands in Haryana. *Indian J. Agro. For.* **8(1):** 77-85

Salvo, S., Munoz, C., Avila, J., Bustos,J., Valdivia, M. R., Silva,C. and Vivallo, G. 2012. An estimate of potential blueberry yield using regression models that relate the number of fruits to the number of flower buds and to climatic variables. *Sci. Hortic.***133:** 56-63.

Shastry, A., Sanjay, H. A. and Bhanusree, E.,2017. Prediction of crop yield using regression techniques. *International J. Soft Computing*. **12(2):** 96-102.

Thomas, W. Macfarland. 1998. Regression, prediction and model building. *J. Agriculture Sciences*. **74(5):** 418-423