The Ecoscan 10(1&2): 105-110, 2016
AN INTERNATIONAL QUARTERLY JOURNAL OF ENVIRONMENTAL SCIENCES
www.theecoscan.com

TIME SERIES MODELS FOR APPLE AREA AND PRODUCTION IN HIMACHAL PRADESH

AJIT SHARMA*, V. K. CHAUDHARY AND SHILPA
Dr. Y.S. Parmar University of Horticulture and Forestry Nauni, Solan - 173 230 (HP)
e-mail: ajit.aziz.sharma@gmail.com

ABSTRACT

In this paper, an attempt has been made to forecast apple area and production using statistical time series modeling techniques - straight line, second degree parabola, exponential, modified exponential, gompertz and logistic. On validation of the forecasts from these models, all the models fitted well to the area under the apple crop but the best fitted model was second degree parabola as it is having highest values for R^2, Adjusted R^2 *i.e.,* 0.963 and 0.961respectively and with least coefficient of variation *i.e.,* 4.91. When the same models are fitted to apple production over the years results showed that none of the model give satisfactory values for R^2, Adjusted R^2 as they are even less than 0.5 and coefficient of variation is too high *i.e.,* above 40.0. Therefore, present study is very helpful to the policy makers, planners, research workers and orchardists for making the future plans of important fruit crops of the state and increased the economy of the state.

KEY WORDS
Time series models
Forecasting models
Area and Production

Received : 13.11.2015
Revised : 17.02.2016
Accepted : 15.03.2016

***Corresponding author**

INTRODUCTION

Apple belongs to family Rosaceae, subfamily Pomoideae, genus Malus and its scientific name is *Malus domestica.* Apple is the main fruit crop of the state and also backbone of Himachal Pradesh as it is covering about 75per cent economy of the state. It occupies an area of 313.0 MH in India with annual production of 2497.7 MT, out of which Himachal Pradesh having area 106.44 MH under the apple and annual production is 412.39 MT during the year 2012-13(NHB, 2014). Himachal Pradesh is known as fruit bowl of India because of its fruit quality and production. Himachal Pradesh is one of the leading states in the apple production in India after Jammu and Kashmir.

The basic idea of this paper revolves around comparison between various statistical models in order to forecast the future values for apple area and production in the coming years. (Sisodia and Kumar, 2004)(Panwar and Kumar, 2003)The best fitted model is used to forecast future values which in turn guide the policy makers regarding their decision making. Among all fruit crops grown in Himachal Pradesh, apple is the major contributor to the farmer's economy as well as state's economy. But the area under the fruit crops is increasing every year however; the production is declining due to climatic factors, attack of insect-pests and diseases etc. (Statistical outline of Himachal Pradesh, 2014). Forecasting models are very helpful to know the real behaviour of past studies and predict the future behaviour of present studies. Forecasting models are mostly used by the policy makers, planners, administrators and research workers for analysing the time series data. Its results are further used for making the future planning's and policies for development of State/Nation. It enables us to study the past behaviour of the phenomenon under consideration i.e, to determine the type and nature of variations in the data. It helps us to compare the changes in the values of different phenomenon at different time. Validation of prediction model(s) for fruit crops is very essential today. This study is also needed for better understanding and to know the suitability of the fruit crops to the farmers of the state in future, so that farmers and policy makers make a plan in advance for its fruit products. Statistical tests are used to carefully examine prior activities and to then use these analyses to make informed predictions about future activities. Regardless of the statistical tests, data are examined in a systematic manner so that decisions can be made with some degree of certainty. The opportunity of using existing data to predict future outcomes is viewed as model building. That is to say, existing data are used to build a model of the future, with a predetermined degree of error built into the model (Thomas, 1998).

A statistical model is an equation or set of equations which represents the relationships among variables. Modelling may lead to less adhoc experimentation, as models sometimes make it easier to design experiments to answer particular questions or to discriminate between alternative mechanisms. Modelling provides powerful tools for investigating the dependence and nature of relationship among the variables of interest. The relationship among variables must be determined for the purpose of predicting the values of one or more variables on the basis of observation on other variables. Model building is currently applied in many fields *i.e,* Agriculture, Biometrics, Econometrics, Education, Meteorology, Industry, Horticulture and Forestry etc. (Prasad, 2010).

The present study was carried out to fulfil two main objectives of identifying the best fitted model for data on area and production of apple and forecasting of future values of area and production of apple.

MATERIALS AND METHODS

From the Statistical outline of Himachal Pradesh, 2011 "Directorate of Economics and Statistics Department, Navbhar, Shimla" the secondary data on apple area and production in Himachal Pradesh for the period of 30 years i.e, 1980-81 to 2009-10 was utilized for model fitting and data for subsequent periods from 2010-11 to 2020-21 (Table 7) was used for validation. The analysis was carried out by using SPSS package and MS-Excel

I.Selection of trend and mathematical equations applied

After ensuring the presence of trend in the data for time series analysis following methods were employed.

Time series Models	Mathematical Equation
Straight line	$Y_t = a + b\,t$
Second degree parabola	$Y_t = a + b\,t + c\,t^2$
Exponential	$Y_t = a\ b^t$
Modified exponential	$Y_t = a + b\,c^t$
Gompertz	$Y_t = a\,b^{c\hat{}t}$
Logistic	$Y_t = k/1 + e^{(a+bt)}$

Straight line

The straight line trend between the time series values (y_t) and time t be given by the equation (Ali and Singh 1995):

$Y_t = a + b\,t$

where,

Y_t = time series values at time t.

t = time period.

a and b are constants.

Second degree parabola

The second degree parabolic trend between the time series values (y_t) and time t be given by the equation (Devi *et al.,* 1990):

$Y_t = a + b\,t + c\,t^2$

where,

Y_t = time series values at time t.

t = time period.

a, b and c are constants.

Exponential

The exponential equation is (Chengappa 1981):

$Y = A\,B^t$

Where, t is the time period and Y is the observation at time t. The exponential model used to compute the interest so that exponential model also known as compound interest model.

$Y = A\,B^t$

$\log Y = \log A + t \log B$

i.e. $y = a + b\,t$

where,

$y = \log Y$

$a = \log A$ and

$b = \log B$

Modified exponential

The modified exponential equation is:

$Y_t = a + b\,c^t$

$Y_{t+h} = a + b\,c^{(t+h)}$

$Y_{t-h} = a + b\,c^{(t-h)}$

$Y_{t+h} - Y_t = b\,c^t (c^h - 1)$

$Y_t - Y_{t-h} = b\,c^{t-h} (c^h - 1)$

$c^h = (Y_{t+h} - Y_t) / (Y_t - Y_{t-h})$ = Constant

From the data three values that are equidistant in time and let these values be y_0, y_1 and y_2. The distance between x_0 and x_1 and between x_1 and x_2 are 'r' years, where r = (n-1) / 2, where n is the number of years *i.e.* period. The constant values are computed by the following formula (Kumar et al. 2005).

$c^r = (y_2 - y_1) / (y_1 - y_0)$ so that, $c = \sqrt[r]{c^r}$

$b = (y_1 - y_0) / (c^r - 1)$ so that, $a = y_0 - b$

Gompertz

(Gaddour et al. 2008) The gompertz equation is:

$Y_t = a\,b^{c\hat{}t}$

Where Y_t is the time series value at time t and a, b and c are its parameters.

$\log Y_t = \log a + c^t \log b$

$y_t = A + B\,c^t$

where,

$y_t = \log Y_t$,

$A = \log a$ and

$B = \log b$

Logistic

The logistic equation is:

$Y_t = k / 1 + (e^{(a+bt)})$

Where Y_t is the value of time series at time t and a, b and k are constants.

$1/Y_t = (1/k)(1 + e^{(a+bt)})$

$= (1/k) + (e^a . e^{bt}) / k$

$= A + bc^t$

Where,

$A = 1/k$ and $B = (1/k)\,e^a$

The logistic curve is fitted by method of three selected points. Three ordinates Y_1, Y_2 and Y_3 are now taken from trend line corresponding to selected equidistant points of time t_1, t_2 and t_3 such that, $t_2 - t_1 = t_3 - t_2$. Then constant values a and b obtained as (Mohanty *et al.*, 2006):

$$k = \frac{y_2^2(Y_1 \quad Y_3) \quad 2Y_1Y_2Y_3}{y_2^2 \quad Y_1Y_3}$$

$$b = \frac{1}{(t_2 \quad t_1)} \log \frac{(k \quad Y_2)Y_1}{(k \quad Y_1)Y_2}$$

$$a = \log \frac{(k \quad Y_1)}{Y_1} \quad bt_1$$

Estimation

The parameters are estimated by the least square technique appropriate to the time series data.

Diagnostic checking

For the adequacy of the model, the best fit model is obtained on the basis of maximum value of R^2 and $\bar{R}^2$ with minimum C.V (%).

R^2

It is one of the indices to measure the goodness of fit of the time series models, it is also known as coefficient of determination.

$$R^2 = 1 \quad \frac{SSE}{SST}$$

Where,

$$SSE = \quad (y_i \quad \hat{y}_i)^2$$

$$SST = \quad (y_i \quad \bar{y})^2$$

$\bar{R}^2$

This statistic used the R^2 statistic, $\bar{R}^2$ is based on the residual d.f and defined as the number of response values 'n' and number of fitted coefficients 'm' estimated from the response value.

Mathematically, V = n - m

V, indicates the number of independent piece of information involving the n data points that are required to calculate the SS. The $\bar{R}^2$ statistic is generally best indicator of the fit quality. When we compare the two models that are nested that is a series of model each of which add previous models.

$$\bar{R}^2 = 1 \quad \frac{SSE(n \quad 1)}{SST(v)}$$

Adjusted R^2 statistic can take all values ≤ 1 closer to 1 indicating better fit, negative values can occur when the model contains terms that do not help to predict the response.

Coefficient of variation

It is the ratio of the S.D to the mean. Usually it is expressed in percentage. The formula for C.V is:

$$C.V = \frac{S.D}{\bar{Y}} \times 100$$

Where, S.D is the standard deviation and $\bar{Y}$ is the mean.

A series or a set of values having lesser coefficient of variation as compared to the other is more consistent.

RESULTS AND DISCUSSION

Area of apple crop

Actual and estimated area of apple crop computed by fitting of six equations namely: straight line, second degree parabola, exponential, modified exponential, gompertz and logistic are presented in (Table 1)

On the basis of statistic R^2, Adjusted R^2 and C.V values (Table 3) (Raizada *et al.,* 2007) (Swetha 2009) all the six prediction models *i.e.* Straight line, Second degree parabola, Exponential, Modified exponential, Gompertz and Logistic was fitted well but the best fitted model was found to be second degree parabola.

Production of apple crop

Actual and estimated production of apple crop computed by fitting of six equations namely: straight line, second degree parabola, exponential, modified exponential, gompertz and logistic are presented in Table 4.

On the basis of R^2, Adjusted R^2 and C.V values (Pandey *et al.,* 1998) (Table 6) it is revealed that none of the model fitted well in the data of apple production, because apple production fluctuate ever year and depends on weather conditions such

Table 1: Data on apple area in Himachal Pradesh for the period 1980-81 to 2009-10.

Year	Actual area (ha)	Straight line	Second degree parabola	Exponential	Modified exponential	Gompertz	Logistic
1980-81	43356	43973.98	39773.83	46318.25	35053.81	37889.65	46199.28
1981-82	45360	45952.95	42621.8	47674.32	39137.88	41027.34	47584.1
1982-83	47354	47931.92	45407.69	49056.09	42999.07	44147.49	49008.3
1983-84	48292	49910.89	48131.52	50506.72	46649.54	47231.62	50472.87
1984-85	49840	51889.86	50793.27	51985.42	50100.79	50263.29	51978.82
1985-86	51103	53868.83	53392.96	53607.4	53363.7	53228.25	53527.16
1986-87	52399	55847.81	55930.57	55073.95	56448.54	56114.35	55118.94
1987-88	54831	57826.78	58406.11	56686.36	59365.03	58911.55	56755.22
1988-89	57447	59805.75	60819.58	58345.97	62122.36	61611.79	58437.08
1989-90	59988	61784.72	63170.97	60054.18	64729.22	64208.87	60165.59
1990-91	62828	63763.69	65460.3	61812.39	67193.81	66698.29	61941.87
1991-92	66767	65742.66	67687.56	64622.08	69523.9	69077.06	63767.03
1992-93	69439	67721.63	69852.74	65484.76	71726.83	71343.59	65642.2
1993-94	72406	69700.6	71955.86	68401.96	73809.53	73497.46	67568.54
1994-95	75469	71679.58	73996.9	69375.3	75778.58	75539.27	69547.2

as rainfall, temperature and humidity etc. as well as attack of insect-pests and diseases viz. wooly aphid, root borer, apple scab, early leaf fall etc. Among all these models second degree parabola seems to be best fit.

Finally, the conclusion has been made that the year to year fluctuations in fruit crops production are quite common in Himachal Pradesh. (Ali and Jabbar, 2013) These fluctuations were due to change of climatic factors and weather conditions during the crop season. These fluctuations adversely affect the production, employment and income, thereby hampering the economic growth of the state. (Sharma *et al.*, 2015). Based on the secondary data available from the year 1980-81 to 2009-10 in Himachal Pradesh, different (Deb and Madhugiri, 2012) time series models viz. straight line, second degree parabola, exponential, modified exponential, gompertz and logistic have been developed to forecast the area and production of apple crop for the next upcoming years up to 2020-21.

Table 1: Contd......

Year	Actual area (ha)	Straight line	Second degree parabola	Exponential	Modified exponential	Gompertz	Logistic
1995-96	78292	73658.55	75975.87	71406.41	77640.18	77470.51	71579.36
1996-97	80338	75637.52	77892.77	73496.99	79400.17	79293.39	73666.2
1997-98	83056	77616.49	79747.6	75648.77	81064.13	81010.71	75808.9
1998-99	85631	79595.46	81540.36	77863.55	82637.27	82625.74	78008.68
1999-00	88673	81574.43	83271.04	80143.18	84124.56	84142.1	80266.76
2000-01	90347	83553.4	84939.66	82489.54	85530.69	85563.72	82584.33
2001-02	92820	85532.38	86546.2	84904.6	86860.08	86894.67	84962.65
2002-03	81630	87511.35	88090.68	87390.37	88116.92	88139.16	87402.93
2003-04	84112	89490.32	89573.08	89948.91	89305.17	89301.45	89906.41
2004-05	86202	91469.29	90993.41	92582.36	90428.58	90385.8	92474.34
2005-06	88560	93448.26	92351.67	95292.91	91490.67	91396.44	95107.94
2006-07	91804	95427.23	93647.86	98082.81	92494.81	92337.52	97808.47
2007-08	94726	97406.2	94881.98	100954.5	93444.14	93213.08	100577.2
2008-09	97438	99385.17	96054.02	103910.1	94341.67	94027.07	103415.2
2009-10	99564	101364.2	97164	106952.3	95190.22	94783.27	106323.91
Total	2180072						

Table 2: Coefficients for the equations of apple area.

Models	a	b	c
Straight line	41995.01	1978.97 (86.74)*	-
Second degree parabola	36863.80	2941.07 (311.32)	-31.04
Exponential	45000.76	1.03 (0.002)	-
Modified exponential	109890.50	-79156.50 (0.056)	0.94
Gompertz	104098.4	0.34 (0.004)	0.92
Logistic	0.000021	0.969 (0.001)	-

Values in parenthesis indicate the S.E of coefficient b

Table 3: R^2, Adjusted R^2 and C.V (%) values for the equations of apple area.

Models	R^2	$\bar{R}^2$	C.V (%)
Straight line	0.949	0.947	5.66
Second degree parabola	0.963	0.961	4.91
Exponential	0.933	0.931	7.59
Modified exponential	0.942	0.938	5.34
Gompertz	0.945	0.941	5.38
Logistic	0.935	0.933	7.52

Table 4: Data on apple production in Himachal Pradesh for the period 1980-81 to 2009-10.

Year	Actual production (100 MT)	Straight line	Second degree parabola	Exponential	Modified exponential	Gompertz	Logistic
1980-81	118.01	180.42	216.14	184.44	231.81	216.59	157.93
1981-82	306.8	188.7	217.03	189.21	232.15	216.66	167.06
1982-83	139.09	196.97	218.44	194.1	232.57	216.76	176.5
1983-84	257.91	205.25	220.38	199.11	233.06	216.88	186.24
1984-85	170.63	213.52	222.85	204.26	233.66	217.03	196.27
1985-86	174.62	221.8	225.85	209.54	234.37	217.23	206.57
1986-87	359.32	230.07	229.37	214.95	235.23	217.47	217.13
1987-88	259.28	238.35	233.42	220.51	236.27	217.79	227.91
1988-89	165.16	246.62	238	226.2	237.51	218.19	238.89
1989-90	394.87	254.9	243.11	232.05	239.01	218.71	250.04
1990-91	342.07	263.17	248.74	238.05	240.81	219.36	261.34
1991-92	301.73	271.45	254.91	244.2	242.98	220.2	272.75
1992-93	279.05	279.72	261.6	250.51	245.59	221.27	284.24
1993-94	294.73	288	268.82	256.98	248.72	222.63	295.78
1994-95	122.78	296.27	276.57	263.62	252.5	224.38	307.33
1995-96	276.68	304.55	284.84	270.43	257.03	226.63	318.86
1996-97	288.54	312.82	293.64	277.42	262.49	229.52	330.33

Table 4: Contd.....

Year	Actual production (100 MT)	Straight line	Second degree parabola	Exponential	Modified exponential	Gompertz	Logistic
1997-98	234.25	321.1	302.97	284.59	269.05	233.25	341.71
1998-99	393.65	329.38	312.83	291.95	276.95	238.08	352.97
1999-00	49.13	337.65	323.22	299.49	286.44	244.37	364.09
2000-01	376.73	345.93	334.28	307.23	297.86	252.6	375.01
2001-02	180.53	354.2	345.58	315.17	311.6	263.48	385.73
2002-03	348.26	362.48	357.55	323.31	328.12	278	396.22
2003-04	459.49	370.74	370.05	331.67	348	297.64	406.45
2004-05	527.6	379.19	383.07	340.24	371.9	324.62	416.41
2005-06	540.36	387.3	396.63	349.03	400.66	362.51	426.07
2006-07	268.4	395.58	410.71	358.05	435.24	417.15	435.43
2007-08	592.57	403.85	425.32	367.3	476.84	498.71	444.46
2008-09	510.16	412.13	440.46	376.8	526.88	625.9	453.17
2009-10	280.11	420.4	456.13	386.53	587.06	835.56	461.54
Total	9012.51						

Source: NHB, 2014.

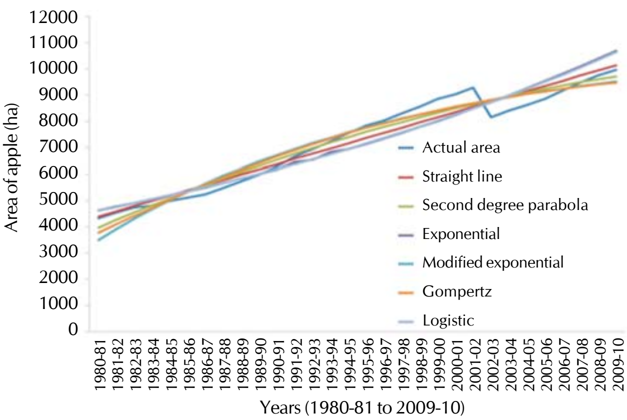

Figure A: Apple area (in ha) with different models over years. Source: NHB, 2014

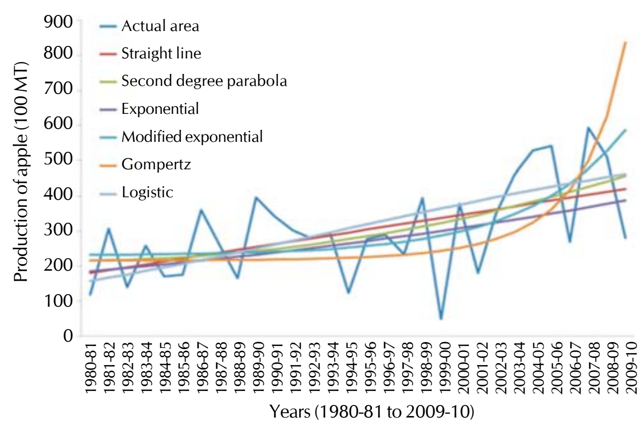

Figure B: Apple production (in MT) with different models over years Source: NHB 2014

Table 5: Coefficients for the equations of apple production

Models	a	B	c
Straight line	172.15	8.27 (2.42)*	-
Second degree parabola	215.79	0.09 (10.08)	0.26
Exponential	179.80	1.03 (0.01)	-
Modified exponential	230.13	1.40 (0.029)	1.20
Gompertz	216.32	1.00 (0.05)	1.27
Logistic	0.005	0.93 (0.02)	-

* Values in parenthesis indicate the S.E of coefficient b.

Table 6: R^2, Adjusted R^2 and C.V (%) values for the equations of apple production

Models	R^2	$\bar{R}^2$	C.V (%)
Straight line	0.293	0.268	38.30
Second degree parabola	0.311	0.270	38.51
Exponential	0.175	0.146	37.91
Modified exponential	0.170	0.108	42.44
Gompertz	0.001	0.0001	48.24
Logistic	0.275	0.249	41.15

From the forecast analysis, the following results were observed. The area of apple may be increased whereas there may be fluctuation in the production of apple. It is revealed that all the forecasts taken for apple production do not show a better trend due to climatic change. In this circumstance based on the present findings, one has to seriously view the various courses of action to be implemented on climatic factors to increase the apple production by taking appropriate positive measures.

Table 7: Forecasts of apple area and production in Himachal Pradesh using second degree parabola model

Year	Forecast of apple area (ha)	Forecast of apple production (100 MT)
2010-11	98207.53	468.44
2011-12	99193.08	484.91
2012-13	100116.6	501.9
2013-14	100977.9	519.41
2014-15	101777.3	537.44
2015-16	102514.5	555.99
2016-17	103189.6	575.06
2017-18	103802.7	594.65
2018-19	104353.7	614.76
2019-20	104842.6	635.39
2020-21	105269.4	656.54

REFERENCES

Ali, M. A. and Singh, A. K. 1995. Growth and fluctuation in area, production and productivity of wheat in Chhattisgarh region of Madhya Pradesh. *Agriculture Situation in India*. **59(9):** 608-614.

Ali, Sajid and Jabbar, A. 2015. Growth and variability in area production and yield ofselected fruit crops in Khyber Pakhtunkhwa. *Pakistan J. Agric. Res.* **28(1):** 64-69.

Anonymous. 2014. Status of apple area and production in India and HP. National Horticultural Board, Gurgaon. http://www.nhb.org.in/

Chengappa, P. G. 1981. Growth rates of area, production and productivity of coffee. *Indian J. Coffee Research*. **11(2):** 19-26.

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

Deb, Paramita and Madhugiri, M. J. 2012. Optimization of apple-pomace based medium for pigment production by *Micrococcus Flavus. The Bioscan*. **7(1):** 57-60.

Devi, P. I., Thomas, E. K . and Thomas, J. K . 1990. Growth and supply response of Banana in Kerala. *Agriculture Situation in India*. **45(4):** 239-242.

Gaddour, A. and Najari, S. 2008. Adjustment of the kid's growth curve in pure goat breeds and crosses under southern Tunisian conditions. *J. Applied Animal Research*. **34(2):** 117-120.

Kumar, T. V., Bhat, A. R. S. and Patil, S. J. 2005. Price forecasting of different classes of teak by the application of exponential smoothing model. *Karnataka J. Agricultural Sciences*. **18(1):** 55-59.

Kumar, A. and Panwar, S.2003. Non-Linear Growth Modelling of Apple Producth'ity of Different States in India. *J. Agriculture Sciences*. **74(5):** 418-423.

Kumar, T. V., Bhat, A. R. S. and Patil, S. J. 2005. Price forecasting of different classes of teak by the application of exponential smoothing model. *Karnataka J. Agricultural Sciences*. **18(1):** 55-59.

Mohanty, M., Painuli, D. K, Misra, A. K., Bandyopadhyaya, K. K and Ghosh, P. K. 2006. Estimating impact of puddling, tillage and residue management on wheat (Triticum aestivum, L.) seedling emergence and growth in a rice-wheat system using nonlinear regression models. *Soil and Tillage Research*. **87(1):** 119-130.

Pandey, Rajiv, Dhall, S. P., Kunwar, B. S., Bharadwaj, S. D. and Panday, R. 1998. Some models for predicting volume of *Populus deltoides*. *Indian Forester*. **124(8):** 629-632.

Prasad, H. 2010. Statistical modelling 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.

Sharma, N. C., Sharma, S. D., Verma, S. and Sharma, C. L. 2013. Impact of changing climate on apple production in Kotkhai area of Shimla district, Himachal Pradesh. *International J. Farm Sciences*. **3(1):** 81-90.

Sisodia, B. V. S. and Sunil Kumar, N. D. 2004. Modelling of Time Series Data for Describing Oilseed Production in Uttar Pradesh. P.hd. thesis. University ofAgriculture & Technology, Faizabad. 87p.

Statistical outline of Himachal Pradesh. 2010. p. 47.

Swetha, K. S. 2009. Statistical study on the impact of rain water harvesting on farming economy. M. Sc. Thesis, University of Agriculture Sciences, Dharwad (Karnatka).

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