Article AR2024a

DOI: HTTPS://DOI.ORG/10.23910/IJBSM/2020.11.1.2024a

International Journal of Bio-resource and Stress Management

February 2020

Crossref

Print ISSN 0976-3988

Online ISSN 0976-4038

IJBSM 2020, 11(1):014-019

Research Article *Natural Resource Management*

Forecasting the Production and Area of Mango (*Mangifera indica* L.) in Himachal Pradesh by using Different Statistical Models

Arun Kumar and R. K. Gupta*

Dept. of Basic Sciences, College of Forestry, Dr Y S Parmar University of Horticulture and Forestry, Nauni, Solan, Himachal Pradesh (173 230), India

Open Access

Corresponding Author

R. K. Gupta

e-mail: arunkumarpatyal123@gmail.com

Citation: Kumar and Gupta, 2020. Forecasting the Production and Area of Mango (*Mangifera indica* L.) in Himachal Pradesh by using Different Statistical Models. International Journal of Bio-resource and Stress Management 2020, 11(1):014-019. HTTPS://DOI.ORG/10.23910/IJBSM/2020.11.1.2024a

Data Availability Statement: Legal restrictions are imposed on the public sharing of raw data. However, authors have full right to transfer or share the data in raw form upon request subject to either meeting the conditions of the original consents and the original research study. Further, access of data needs to meet whether the user complies with the ethical and legal obligations as data controllers to allow for secondary use of the data outside of the original study.

Conflict of interests: The authors have declared that no conflict of interest exists.

Abstract

A statistical model is a formulization of relationships between variables in the form of the mathematical equation or set of mathematical equations. Statistical model provides tools for investigating the dependence and nature of relationship among the variables of interest. Different statistical models were used for forecasting the production of vegetables and fruits. Forecasts of agricultural production are intended to be useful for farmers, governments, and agribusiness industries. Because of the special position of food production in a nation's security, governments have become both principal suppliers and main users of agricultural forecasts. They need internal forecasts to execute policies that provide technical and market support for the agriculture sector. Government publications routinely provide private decision makers with commodity price and output forecasts at regional and national levels and at various horizons The present investigation were conducted in the Department of Basic Science, Dr Y.S. Parmar University of Horticulture and Forestry, Nauni, Solan (173 230) (H.P.) during 2013-2015. Secondary data on area (ha) and production (MT) for mango for last eighteen years was used and different prediction models viz. autoregressive, straight line, second degree parabola, exponential, modified exponential and gompertz were fitted & tested by using Adj. R^2, Root mean square error and Thiel's inequality coefficient. Second degree parabola and autoregressive models were found to be best models to forecast the area and production of mango crop as per Adj. R^2, Root mean square error and Thiel's inequality coefficient (U).

Keywords: Mango, autoregressive model, Thiel's inequality coefficient

1. Introduction

Mango (*Mangiferal ndica* L.) is known as "King of Fruits" for many years and it is also considered a major fruit crop of various countries such as Philippines, Mexico, India, Brazil, Pakistan, China and Thailand. India ranks first among the mango producing countries of the world accounting for about 50% of the world's mango production.The mango is adapted to both tropical and subtropical conditions. InHimachal Pradesh is known as the fruit bowl of India because of its excellent fruit quality and production. Mango is cultivated in the low hill zone of Himachal Pradesh. In Himachal Pradesh, mango cultivation occupies an area of 40,298 hectare with production of 25,408 MT. A statistical model is a formulization of relationships between variables in the form of the mathematical equation or set of mathematical equations. Statistical model provides tools for

Article History

RECEIVED in 09th August 2019 RECEIVED in revised form 20th January 2020 ACCEPTED in final form 11th February 2020

investigating the dependence and nature of relationship among the variables of interest. Different Statistical models were used for forecasting the production of fruits. Forecasts of agricultural production are intended to be useful for farmers, governments, and agribusiness industries.

Fernondeg and Gomej (2004) employed autoregressive models to predict weekly milk yield in the goat farm. Twenty-eight goats were used to build the model and eight goats were used to validate itShukla and Jharkharia (2011) studied the application of autoregressive integrated moving average (ARIMA) and autoregressive models t o forecast the demand of fresh produce (fruits and vegetables) on a daily basis. Models were built using 25 months sales data of onion from Ahmadabad market in India. Sankar (2014) studied that the egg has a very well balanced amino acid profile with the required minerals and vitamins. the design of the stochastic modeling for egg production forecasting in Tamilnadu, based on data on egg production during the years from 1996 to 2008. The study considered Autoregressive (AR), Moving Average (MA) and Autoregressive Integrated Moving Average (ARIMA) processes to select the appropriate stochastic model for egg production forecasting in Tamilnadu. Based on the chosen model, it could be predicted that the egg production would increase to 19,179 millions in 2015 from 8,960 millions in 2008 in Tamilnadu.

Kaabia and Roig (2008) investigated the non-linear adjustments between farm and retail prices in the tomato sector in Spain. The methodology used was based on the multivariate approach to specify and estimate a autoregressive model. Rahman (2010) examined the best fitted ARIMA model that could be used to make efficient forecast boro rice production in Bangladesh from 2008-09 to 2012-13.Chena et al. (2013) suggested that autoregressive model was a popular method for analysing the time dependent data, where selection of order parameter was imperative. Hamjah (2014) suggested that Bangladesh has a large agrarian base country where 77% of total population was living in the rural areas and 90 % of the rural population directly related with agriculture. Banana, Guava, Papaya, Jackfruit, Pineapple, Mango etc. are the major fruits crops in Bangladesh. The main objective of his study was to fit the Box-Jenkins Auto Regressive Integrated Moving Average (ARIMA) and autoregressive Models to forecast the different types of major fruits productions in Bangladesh. Reeves and Xie (2014) forecasted the volatility at horizons such as quarterly were of fundamental importance to asset pricing and risk management. Erfani and Samimi (2009) forecasted stock price index using a fractionally differenced arma model. Ghosh et al (2010) studied functional coefficient of autoregressive nonlinear time-series model for forecasting Indian lac export data

Yusuf and Sheu (2007), they provides the prediction of future production of citrus and mango up to the year 2010 in Nigeria using various forecasting techniques. Khan et al. (2008) predicted the production of Mango in Pakistan. Ali and Singh (1995) studied on the growth pattern using time series data on area, yield and production of wheat for twenty years, Irfan et al. (2011) studied that Autoregressive and moving average models were described for forecasting the yield of rice in four provinces of Pakistan. Padhan (2012) predicted the annual productivity of agricultural crops. Mishra and Kumar (2012) studied the price behavior of major vegetables in hill region of Nepal. Lixin et al (2009) Compare the six statistical approaches in the selection of appropriate fish growth models.

2. Materials and Methods

The secondary data on area (ha) and production (MT) of mango was collected from Department Horticulture Shimla and Department of Economics and Statistics for the period of 18 years i.e. 1996-97 to 2013-14. Keeping in view the objectiveof study, following statistical tools and modelshave been applied. The data analyzed by statistical software like; SPSS, SAS and Curve Expert. Keeping in view the objective of study, following statistical tools and models have been applied.

Statistical Models	Mathematical Equation
Autoregressive	$Y_t=\Phi_1 Y_{t-1}+\Phi_2 Y_{t-2}+\Phi_3 Y_{t-3}+...+\Phi_p Y_{t-p}+e_t$
Straight line	$y_t=a+bt+e_t$
Second degree parabola	$y_t=a+bt+ct^2+e_t$
Exponential	$y_t=ae^{bt}+e_t$
Modified exponential	$y_t=a+bc^t+e_t$
Gompertz	$y_t=ab^{c^t}+e_t$

Several indices viz. Adjusted R^2, RMSE and Thiel's Inequality Coefficient were used to examine the goodness of fit of different time series models.

3. Results and Discussion

3.1. Area of mango crop

In the present study the last eighteen years data from 1996-97 to 2013-14 of area and production of mango crop of Himachal Pradesh has been taken. Area of mango crop was estimated by using various linear and non linear models and results are presented in Table 1. Actual and estimated area of mango for different prediction models from year 1996-97 to 2013-14 are presented in Table 2. The prediction models viz. straight line, second degree parabola, exponential, modified exponential and gompertz were fitted well to predict the area under mango crop. Second degree parabola was found to be best prediction model to forecast the area of mango among all five prediction model because second degree parabola has highest value of Adj. R^2 (0.990) and has lowest value for RMSE (404.887) and Thiel's inequality coefficient (U) (0.0001).

Area of mango was also estimated by using autoregressive model. Table 3 presents the autocorrelation and standard

Table 1: Coefficients of linear and non-linear models for prediction of area of mango

Statistical Model	Equation	SE of estimate(s)	Adj R^2	RMSE	U
Straight line	27929.85+776.88 X	592.08*, 54.69*	0.92	1168.05	0.001
Second degree parabola	$25115.71-44.43X+1621X^2$	73.41*, 10.48*, 19.55*	0.99	404.88	0.0001
Exponential	28208.46 exp(0.023X)	48.71*,12.04*	0.89	1389.52	0.0014
Modified exponential	$36129.22-12117.10\times0.810^X$	3.70*, 2.90*, 4.50*	0.58	2614.70	0.0535
Gompertz	$55808.45\times0.442^{0.923X}$	3.50*, 5.10*, 3.30*	0.48	2916.90	0.005

Table 2: Trend of five prediction models for the mango area from 1996-97 to 2013-14 in Himachal Pradesh

Year	Actual area (ha)	Straight line	Second degree Parabola	Exponential	Modified exponential	Gompertz
1996-97	26308	28706.73	26692.40	28860.29	26314.39	26296.94
1997-98	28399	29483.61	28180.22	29527.17	28179.21	27848.42
1998-99	29833	30260.49	29579.18	30209.46	29695.05	29360.83
1999-00	30933	31037.37	30889.26	30907.52	30918.88	30828.59
2000-01	32057	31814.26	32110.48	31621.71	31912.48	32257.28
2001-02	33684	32591.14	33242.83	32352.40	32712.20	33630.17
2002-03	33380	33368.02	34286.32	33099.98	33336.53	34908.19
2003-04	35144	34144.90	35240.93	33864.83	33887.56	36219.68
2004-05	36215	34921.78	36106.68	34647.35	34311.66	37430.73
2005-06	37408	35698.66	36883.56	35447.96	34663.05	38585.96
2006-07	38370	36475.54	37571.58	36267.07	34941.75	39685.39
2007-08	37840	37252.43	38170.72	37105.10	35171.97	40723.43
2008-09	38444	38029.31	38681.00	37962.50	35353.73	41716.82
2009-10	38681	38806.19	39102.41	38839.71	35499.13	42648.82
2010-11	39194	39583.07	39434.96	39737.19	35620.30	43530.59
2011-12	39568	40359.95	39678.63	40655.41	35717.24	44362.14
2012-13	39828	41136.83	39833.44	41594.84	35802.60	45143.46
2013-14	40298	41913.71	39899.39	42555.99	35862.64	45880.13

Table 3: Autocorrelation function of mango area

Lags	Autocorrelation	Standard error
1	0.81*	0.21
2	0.63*	0.21
3	0.47*	0.20
4	0.35	0.19
5	0.18	0.19
6	0.07	0.18

error of different lags. Since first three autocorrelations were found statistically significant, hence autoregressive model up to order three were fitted to predict the area of mango. Coefficients, Standard error of coefficients, t-statistic, Adj. R^2 and Root mean square error RMSE of autoregressive model up to third order of mangoarea are presented in Table 4. Perusal of Table 4 shows that coefficients of 2nd order autoregressive and 3rd order autoregressive model are not statistically significant, while the coefficient of 1st order autoregressive model was statistically significant with Adj. R^2 (0.92) and RMSE (666.64). Trend of actual and estimated area of mango by 1st order autoregressive model are presented in Table 5.

1st order autoregressive model:

$y\_t=\Phi_1\ y_{t-1}$

$y_t=(1.018)y_{t-1}$

3.2. Production of mango crop

Production of mango crop was estimated by using various linear and non linear models Coefficients, Standard error of coefficients, t-statistic, Adj. R^2 and (RMSE) of all prediction

Table 4: Coefficients, standard error, t-statistic, Adj. R^2 and RMSE of autoregressive models of mango area

Order of autoregressive model	Coefficients of different Lags	Value of coefficients	Standard error	t-statistic	Adj. R^2	RMSE
1st order	Φ_1	1.018	0.005	207.82*	0.92	666.64
2nd order	Φ_1	1.085	0.26	4.12*	0.92	710.79
	Φ_2	-0.068	0.26	0.25		
3rd order	Φ_2	1.005	0.25	3.95*	0.92	686.64
	Φ_2	0.370	2.21	0.16		
	Φ_3	- 0.366	0.81	0.45		

Table 5: Trend of 1st order autoregressive model for the mango area in Himachal Pradesh

Year	Actual area (ha)	1st Order autoregressive model
1996-97	26308	-
1997-98	28399	26781.54
1998-99	29833	27263.61
1999-00	30933	27754.36
2000-01	32057	28253.94
2001-02	33684	28762.51
2002-03	33380	29280.23
2003-04	35144	29807.28
2004-05	36215	30343.81
2005-06	37408	30889.99
2006-07	38370	31446.01
2007-08	37840	32012.04
2008-09	38444	32588.26
2009-10	38681	33174.85
2010-11	39194	33772.00
2011-12	39568	34379.89
2012-13	39828	34998.73
2013-14	40298	35628.71
2014-15		36270.02
2015-16		36922.88
2016-17		37587.50
2017-18		38264.07
2018-19		38952.82
2019-20		39653.97
2020-21		40367.75

models are presented in Table 6. Actual and estimated production of mango for three different prediction models namely straight line, second degree parabola and exponential from year 1996-97 to 2013-14 are presented in Table 7

Table 6: Coefficients of linear and non-linear models for prediction of production of mango

Statistical model	Equation	SE of estimate (s)	Adj R^2	RMSE	U
Straight line	15038.69 + 1503.14 X	7045.79*, 650.85*	0.20	13898.45	0.345
Second degree parabola	-1530.75– 261.62 X+ 6473.99 X^2	510.59*, 110.90*, 2520.59*	0.33	12310.62	0.286
Exponential	11997.05$_{exp}$ (0.076X)	3221.55*, 0.025*	0.33	15296.84	0.420
Modified exponential	17766.60 -1172.77× 1.17	11104.10*, 2931.93*, 0.90*	0.09	15779.64	-

because these models were fitted well.Second degree parabola has highest value of Adj. R^2 (0.33) and has lowest value for RMSE (12310.62) and Thiel's inequality coefficients (U) (0.149).

Production of mango was also estimated by using autoregressive model. Table 8 presents the autocorrelation and standard error of lags up to order six. Since first two autocorrelations were found statistically significant, hence autoregressive model up to order two were fitted to predict the production of mango. Perusal of Table 9 shows that coefficients of 2nd order autoregressive model are not statistically significant, while the coefficient of 1st order autoregressive model was statistically significant with Adj. R^2 (0.81) and RMSE (16028). Trend of actual and estimated production of mango by 1st order autoregressive model are presented in Table 10.

1st order autoregressive model:

$y_t=\Phi_1 y_{t-1}$

$y_t=(0.910)y_{t-1}$

Table 7: Trend of three prediction models for the mango production from 1996-97 to 2013-14 in Himachal Pradesh

Year	Actual Production (MT)	Straight line	Second degree Parabola	Exponential
1996-97	19144	16541.85	4681.61	12943.38
1997-98	4024	18045.00	10370.70	13964.36
1998-99	16892	19548.15	15536.60	15065.88
1999-00	9414	21051.29	20179.20	16254.28
2000-01	13098	22554.44	24298.60	17536.43
2001-02	26744	24057.59	27894.70	18919.71
2002-03	25311	25560.74	30967.60	20412.11
2003-04	22110	27063.89	33517.30	22022.23
2004-05	59739	28567.04	35543.70	23759.35
2005-06	63091	30070.19	37046.80	25633.50
2006-07	40159	31573.33	38026.70	27655.49
2007-08	29252	33076.48	38483.40	29836.96
2008-09	38751	34579.63	38416.80	32190.52
2009-10	24162	36082.78	37826.90	34729.72
2010-11	31463	37585.93	36713.90	37469.22
2011-12	28972	39089.08	35077.50	40424.81
2012-13	50001	40592.23	32918.00	43613.54
2013-14	25408	42095.37	30235.10	47053.80

Table 8: Autocorrelation function of mango production

Lags	Autocorrelation	Standard error
1	0.51*	0.21
2	0.45*	0.21
3	0.15	0.20
4	0.03	0.19
5	-0.26	0.19
6	-0.16	0.18

Table 9: Coefficients, Standard error, t-statistic, Adj. R2 and RMSE of autoregressive models of mango production

OAGM	CL	VC	SE	t-statistic	Adj. R^2	RMSE
1st order	Φ_1	0.910	0.11	7.78*	0.81	16028
2nd order	Φ_1	0.737	0.28	2.58*	0.81	16355
	Φ_2	0.204	0.30	0.66		

OAGM: Order of autoregressive model; CL: Coefficients for Lags; VC: Value of coefficients; SE: Standard error

Table 11 represents the estimated area and production of mango in Himachal Pradesh for year 2020-2021 by using different prediction models.

Table 10: Trend of 1st order autoregressive model for the mango production in Himachal Pradesh

Year	Actual production (MT)	1st Order autoregressive model
1996-97	19144	-
1997-98	4024	19086.57
1998-99	16892	19029.31
1999-00	9414	18972.22
2000-01	13098	18915.3
2001-02	26744	18858.56
2002-03	25311	18801.98
2003-04	22110	18745.58
2004-05	59739	18689.34
2005-06	63091	18633.27
2006-07	40159	18577.37
2007-08	29252	18521.64
2008-09	38751	18466.07
2009-10	24162	18410.68
2010-11	31463	18355.44
2011-12	28972	18300.38
2012-13	50001	18245.48
2013-14	25408	18190.74
2014-15		18136.17
2015-16		18081.76
2016-17		18027.51
2017-18		17973.43
2018-19		17919.51
2019-20		17865.75
2020-21		17812.16

Table 11: Predicted area and production of mango in Himachal Pradesh for year 2020-2021

Statistical Models	Area (ha)	Production (MT)
Straight line	47351.88	52617.41
Second degree parabola	37872.69	30001.49
Exponential	49935.45	80058.63
Modified exponential	36066.77	-
Gompertz	49867.71	-
1st order autoregressive	40367.75	17812.16

4. Conclusion

Different prediction models namely; Straight line, Second degree parabola, Exponential, Modified exponential, Gompert and Autoregressive were tried to predict the area and production of mango. All the models fitted well on the basis of Adj. R^2, Root Mean Square Error (RMSE) and Thiel's inequality coefficient (U). Second degree parabola and 1st order autoregressive were best models to forecast the area and production of mango as high Adj. R^2, low RMSE and Thiel's inequality coefficient (U).

5. References

Ali, M.A., Singh, A.K., 1995. Growth and fluctuation in area, production and productivity of wheat in Chhattisgarh region of Madhya Pradesh. Agriculture Situation in India59(9), 608–614.

Chena, C.S., Leeb, Y.H., Hsua, H., 2013. Adaptive order selection for autoregressive models. Journal of Statistical Computation and Simulation 84(9), 1963–1974.

Erfani, A., Samimi, A.J., 2009. Long memory forecasting of stock price index using a fractionally differenced arma model, Journal of Applied Sciences Research 5(10), 1721–1731.

Fernondeg, C., Gomej, J., 2004. Prediction of weekly goat milk yield using autoregressive model. South African Journal of Animal Science 34, 169–177

Ghosh, H., Paul, R.K., Prajneshu, 2010. Functional coefficient autoregressive nonlinear time-series model for forecasting Indian lac export data. Model Assisted Statistics and Applications 5(2), 101–108.

Hamjah, M.A., 2014. Forecasting Major Fruit Crops Productions in Bangladesh using Box-Jenkins ARIMA Model. Journal of Economics and Sustainable Development 5(7), 137–142.

Irfan, M., Irfan, M., Tahir, M., 2011. Modelling "calcots" (*Allium cepa* L.) growth by the Gompertz function. SORT 37(1), 95–106.

Kaabia, M.B., Roig, J.M.G., 2008. Price transmission asymmetries in the Spanish tomato sector. Economia-Agraria-y-Recursos-Naturales 8(1), 57–82.

Khan, M., Mustafa, K., Shah, M., Khan, M., Khan, J.Z., 2008.Forecasting mango production in Pakistan an econometric model approach.Sarhad Journal of Agriculture 24(2), 363–370.

Lixin, Z., Lifang L., Zhenlin, L., 2009. Comparison of six statistical approaches in the selection of appropriate fish growth models, Chinese Journal of Oceanology and Limnology 3, 57–467.

Mishra, R., Kumar, D.A., 2012. Price behaviour of major vegetables in hill region of Nepal: An econometric analysis. SAARC Journal of Agriculture 10(2), 107–12.

Padhan, P.C., 2012. Application of ARIMA model for forecasting agricultural productivity in India. Journal of Agriculture and Social Sciences 8(2), 50–56.

Rahman, N.M.F., 2010. Forecasting of boro rice production in Bangladesh. Journal of Bangladesh Agriculture University 8(1), 103–112.

Reeves, J.J., Xie, X., 2014. Forecasting stock return volatility at the quarterly frequency: an evaluation of time series approaches. Applied Financial Economics 24(5), 347–356.

Sankar, T.J., 2014. Design of a stochastic forecasting model for egg production. International Journal of Innovative Science, Engineering & Technology 1(6), 319–325.

Shukla, M., Jharkharia, S., 2011. ARIMA models to forecast demand in fresh supply chains. International Journal of Operational Research 11(1), 1–18.

Yusuf, S.A., Salau, A.S., 2007. Forecasting mango and citrus production in Nigeria: trend analysis.Nigerian Agricultural Development Studies 1(2), 1–19.