Economic Affairs
DOI : 10.5958/0976-4666.2015.00101.1

Trend in Area, Production and Productivity of Wheat (*Triticum aestivum*) in Gujarat

L. Netajit Singh[1], V.B. Darji[2] and Y. Santosh Singh[3]

[1]*Department of Agricultural Statistics, N. M. College of Agriculture, Navsari Agricultural University, Navsari, 396450, Gujarat, India*
[2]*V.B. Darji, College of Agricultural Information Technology, Anand Agricultural University, Anand, 388110, Gujarat, India*
[3]*Department of Agricultural Economics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia, 741252, West Bengal, India*

Corresponding author: netajitaau@gmail.com

Paper No.: 291 Received: 27 January 2015 Accepted: 17 November 2015

Abstract

Using polynomial (viz., linear, quadratic and cubic) and exponential models fluctuation in area, production and productivity of wheat crop in Gujarat state was studied. The data for the years 1960-61 to 2006-07 were used for model fitting and efficiency of the fitted model(s) was tested using data of 2007-08 to 2010-11. The models were fitted on original data as well as three, four and five year moving averages data. The most suitable model was selected on the basis of adjusted R^2, significant regression coefficient, root mean square error, mean absolute error, normality (Shapiro-Wilk test) and randomness of residual's (Run test) distribution. The results showed that none of the models was suitable for area prediction whereas third degree polynomial model on original data was found suitable to fit the trend of production and productivity of wheat in Gujarat state with adjusted R^2 values of 53.75 and 85.30 respectively.

Keywords: Area, production, productivity, wheat, polynomial

Time series data of crop statistics are being used to determine the trend. Such analysis helps in planning (at government level) the requirement of agricultural commodities for state/country as a whole. Present study deals with wheat crop grown in Gujarat state where it is grown in 11.05 mha with the production 3.13 mt and productivity of 2986 kg ha^{-1} (Anon., 2013b). Gujarat accounts for 1.75% of the total area and 1.32% of the total production of wheat in the country. Mehsana, Banaskantha, Rajkot, Kheda and Anand districts are the main producers which together contribute about 55% of the state's production of wheat. Others include Ahmedabad, Sabarkantha, Bharuch and Bhavnagar districts where 6 to 10 per cent of the cropped area is devoted to wheat cultivation (Anon., 2012).

In literature, a large number of univariate time series models are available *viz.*, linear, polynomial, exponential, logistic *etc.* for determining the trend i.e. forecasting model(s). In the present investigation linear, quadratic, cubic and exponential model were used.

Database and Methodology

Regression analysis become one of the most widely used statistical tool for analyzing functional relationship among the variables which is expressed in the form of an equation connecting the response or dependent variable Y (area, production and productivity) and time variable t as independent variable. The following models were fitted to the original data as well as moving averages data,

$$Y = a + bt \quad (1)$$

Where, a and b are the regression constant and regression coefficient, respectively to be estimated.

The fitted regression equation was as under,

$$Y = a + bt + ct^2 \quad (2)$$

The unknown parameter *vi.,* a, b and c were estimated by using 'Principle of least square' method (Montgomery *et al*. 2003).

The model for the third degree polynomial fitted to the data of each district was as under,

$$Y = a + bt + ct^2 + dt^3 \quad (3)$$

The constant a coefficient b, c and d were estimated using least square method.

The fitted regression equation was as under,

$$Y = ab^t \quad (4)$$

The goodness of fit (Montgomery *et al*. 2003) of the models was justified by coefficient of determination (R^2) and adjusted (R^2) was calculated as under

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

and

$$\text{Adj.}R^2 = 1 - \frac{(n-1)(1-R^2)}{(n-k)} \quad (5)$$

R^2 indicates the amount of variation in dependent variable accounted due to the model.

To test the overall significance of the model the F test will be used.

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} \quad (6)$$

Which follows F distribution with [k, (n-k-1)] degrees of freedom.

Where, n = number of observations

k = number of independent variable(s)

The individual regression coefficients were tested using the t test,

$$t = \frac{b_i}{\text{S.E.}(b_i)} \text{ with (n-k-1) degrees of freedom} \quad (7)$$

b_i = estimated i^{th} coefficient and

S.E. (b_i) is the standard error of b_i.

In addition to the above criteria, two more reliability statistics Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were computed to measure the adequacy of the fitted model. They can be computed as follows:

$$\text{RMSE} = \left[\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2/n\right]^{1/2} \quad (8)$$

and

$$\text{MAE} = \sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|/n \quad (9)$$

The fitted models, which had lower values of these estimates, were considered to be better.

Test for the randomness of the residuals (Sidney and Castellan, 1988)

m = (no. of '+ve' sign residuals),

n = (no. of '-ve' sign residuals)

N = m + n.

r = no. of runs

Mean = $\mu = \frac{2mn}{N} + 1$ and standard deviation = σ_r =

$$\left[\frac{2mn(2mn-N)}{N^2(N-1)}\right]^{1/2} \quad (10)$$

$$Z = \frac{r - \mu_r}{\sigma_r} = \frac{r + h - \frac{2mn}{N} - 1}{\sqrt{\frac{2mn(2mn - N)}{N^2(N-1)}}} \quad (11)$$

Where h = 0.5 if $r < \frac{2mn}{N} + 1$ and

$$h = 0 \text{ if } r > \frac{2mnb}{N} + 1$$

Use normal table (Appendix A by Sidney and Castellan, 1988) for testing Z value. The non significant Z value indicates randomness of the residual.

Test for normality of the residual (Shapiro – Wilk, 1965)

The required test statistics W

$$W = \frac{S^2}{b} \text{ Where } S^2 = \sum a(k)[e(n+1-k) - e(k)] \quad (12)$$

The parameter k takes the values,

$$K = \begin{cases} 1,2,3,4,.....n/2 & \text{when } n \text{ is even} \\ 1,2,3,4,.....(n-1)/2 & \text{when } n \text{ is odd} \end{cases}$$

$$\text{and } b = \sum_{i=1}^{n} (e_i - \bar{e})^2$$

The values of coefficients "a(k)" for different values of n and k are given (Shapiro - Wilk, 1965). When the calculated value of W is non-significant i.e. very close to unity, the null hypothesis regarding normality of residual was accepted.

The model was selected on the basis of following condition:

1. The model should possess significant F value for Coefficient of Determination.
2. The regression coefficient in the model should be significant.
3. The residuals should be normally and independently distributed.

Forecast values for four years (2007-08 to 2010-11) by using selected models were tested by following relation.

$$\% \text{ of } FE = \left(\frac{Y_i - \hat{Y}_i}{Y_i} \right) \times 100 \quad (13)$$

i=1, 2, 3, 4; 1=2007-08, 2=2008-09, 3=2009-10, 4= 2010-11

Where, FE is the forecasting error

Y_i is the observed value of remaining four years

$\hat{Y}_i$ is the forecast value of remaining four years

Results and Discussion

The results of fitted polynomial models are given in Table 1. The results showed that all coefficient of determination (R^2) values were significant in different models, it was maximum (39.09) in cubic regression model on five years moving average approach. The regression constants of all models were significant. Linear regression coefficient values were significant in all cubic models, original and three year moving approach of linear model and quadratic model of four years and five years moving average approaches. Quadratic regression coefficient values were significant in all cubic models and quadratic model of five years moving average approach. Cubic regression coefficients were significant in all the cubic models and all the regression coefficient of exponential model also significant. In original data approach R^2 value was decreased by 1.91% in quadratic regression as compared to the linear regression, while, in case of cubic regression

Table 1: Fitted polynomial models for wheat area in Gujarat state

Model	Moving Average	Regression constant	Regression coefficients			Adj. R^2	RMSE	MAE	S-W Test	Run test (|Z|)
		a	b	c	d					
Linear	Original	4516.80*	46.72*	—	—	13.38*	1526.28	1233.73	0.969	2.358*
	3 year	4798.66*	34.06*	—	—	12.42*	1103.00	862.58	0.977	2.109*
	4 year	4914.89*	29.20	—	—	11.64*	953.72	949.21	0.968	2.593*
	5 year	5027.42*	24.32	—	—	10.37*	817.94	681.14	0.948	3.7.3*
Quadratic	Original	4425.59*	57.86	-0.23	—	11.47*	1543.02	1254.58	0.968	2.358*
	3 year	4357.46*	90.39	-1.22	—	12.96*	1099.58	837.33	0.976	2.109*
	4 year	4374.43*	99.64*	-1.57	—	14.80*	935.51	714.09	0.951	3.203*
	5 year	4388.01*	109.57*	-1.94*	—	18.38*	780.10	618.60	0.963	2.467*
Cubic	Original	2312.41*	559.95*	-26.11*	0.36*	29.49*	1377.04	1138.90	0.950*	0.882
	3 year	2740.00*	488.04*	-22.60*	0.31*	33.21*	963.50	819.51	0.947*	4.221*
	4 year	2961.42*	456.55*	-21.17*	0.29*	35.92*	811.34	688.81	0.954	4.119*
	5 year	3196.28*	417.83*	-19.25*	0.26*	39.09*	673.89	600.42	0.970	4.012*
Exponential	Original	4532.49*	0.007*	—	—	7.90*	1547.29	1293.10	0.954	2.358*
	3 year	477106*	0.006*	—	—	10.6*	1111.21	872.44	0.965	3.316*
	4 year	4873.49*	0.005*	—	—	11.5*	958.63	754.47	0.953	2.593*
	5 year	4972.82*	0.005*	—	—	11.6*	822.21	678.20	0.946*	3.703*

* Significant at 5% level.

Table 2: Fitted polynomial models for wheat production in Gujarat state

Model	Moving Average	Regression constant	Regression coefficients			Adj. R^2	RMSE	MAE	S-W Test	Run test (|Z|)
		a	b	c	d					
Linear	Original	3412.60*	296.28*	—	—	49.25*	4078.26	3299.74	0.977	1.177
	3 year	4223.60*	256.85*	—	—	62.51*	2685.60	2044.07	0.981	3.316*
	4 year	4561.79*	254.80*	—	—	67.95*	2235.78	1680.07	0.968	3.813*
	5 year	4908.98*	242.86*	—	-	74.36*	1787.88	1409.85	0.962	3.085*
Quadratic	Original	4015.85*	222.41	1.54	—	48.30*	4116.04	3417.12	0.981	1.177
	3 year	8203.38*	319.62*	-1.17	—	61.79*	2711.25	2064.13	0.977	2.712*
	4 year	3839.38*	349.03*	-2.09	—	67.79*	2241.18	1649.03	0.977	3.203*
	5 year	3861.20*	382.56*	-3.17	—	75.27*	2752.47	1376.10	0.957	3.085*
Cubic	Original	-226.97	1230.44*	-50.41*	0.72*	53.75*	3893.42	3118.23	0.965	0.882
	3 year	933.34	1029.36*	-39.32*	0.55*	66.00*	2557.71	2016.18	0.978	3.014*
	4 year	1494.35	941.35*	-34.63*	0.48*	71.20*	2119.14	1621.38	0.972	2.898*
	5 year	2156.80	822.32*	-27.88*	0.37*	77.42*	1674.48	1331.27	0.973	2.158*
Exponential	Original	4202.86*	0.031*	—	—	44.3*	4150.56	3463.16	0.976	1.177
	3 year	4830.61*	0.021*	—	—	65.1*	2796.23	2205.13	0.978	2.712*
	4 year	5073.52*	0.028*	—	—	70.2*	2831.08	1803.29	0.970	3.203*
	5 year	5313.55*	0.027*	—	—	74.1*	1984.23	1528.85	0.978	3.085*

* Significant at 5% level.

model it was improved about 18.02% over the quadratic model. The R^2 values for all the exponential models were very low. By taking five years moving average, coefficient of determination value was improved in second, third degree polynomial and exponential model by 6.91, 9.60 and 3.7% respectively, while in case first degree polynomial model, it was decreased by 3.01% over the original data approaches. Thus, no much improved in coefficient of determination values were observed due to moving average.

The cubic model of five years moving average approach showed comparatively lower value of RMSE and MAE with significant regression coefficients and also highest R^2 value. The criteria for testing normality (S-W test) of residual indicated that all the fitted models (except original data approach and three years moving averages of third degree polynomial and five years moving average of exponential models) fulfill the normality of residuals. Run test indicated that only third degree polynomial model on original approach is randomly distributed, but it did not satisfy the normality of residuals. Therefore, none of the polynomial models satisfied normality and randomness of residuals. So, none of the models was found suitable to predict the pattern of wheat area in Gujarat State

The results of fitted polynomial models are given in Table 2. The results indicated that all coefficient of determination (R^2) values were significant in all different models and it was maximum (77.42) in cubic regression of five years moving average approach. The regression constant values were significant in linear, quadratic, exponential models of all data approach. All linear regression coefficients were significant in all models in all approach (except original data approach of quadratic model). Quadratic and cubic regression coefficients were significant in case of the all cubic models. All the regression coefficients of exponential model were also significant. None of the quadratic models were significant on quadratic regression coefficient. In original data approach the value of (R^2) was nearly similar in first and second degree polynomial but in third degree polynomial model it was improved by 5.45 per cent from quadratic model. By taking five years moving average, R^2 values were improved by 25.11, 26.97, 23.67 and 29.8 per cent in case of first, second, third degree polynomial and exponential models, respectively over the original data.

Table 3: Testing of forecast values for remaining four years by using selected model i.e. cubic model in original data approach of wheat production in Gujarat state

Years	Observed Values	Cubic Predicted values	Error Per cent
2007-08	38377	22480.07	41.42
2008-09	25926	23912.46	07.77
2009-10	23192	25456.16	-09.76
2010-11	27915	27115.51	02.86

Table 4: Fitted polynomial models for wheat productivity in Gujarat state

Model	Moving Average	Regression constant	Regression coefficients			Adj. R^2	RMSE	MAE	S-W Test	Run test (|Z|)
		a	b	c	d					
Linear	Original	1119.92*	33.24*	—	—	70.95*	290.32	246.54	0.973	2.062*
	3 year	1184.83*	32.24*	—	—	74.60*	246.14	203.91	0.963	3.014*
	4 year	1217.12*	31.73*	—	—	74.12*	239.87	198.35	0.960	4.424*
	5 year	1253.58*	31.00*	—	—	73.46*	232.97	192.47	0.962	4.629*
Quadratic	Original	731.00*	80.86*	-0.99*	—	80.09*	240.33	196.86	0.981	2.948*
	3 year	793.09*	82.25*	-1.09*	—	86.06*	182.35	153.72	0.950	3.618*
	4 year	825.50*	82.81*	-1.13*	—	86.41*	173.84	147.95	0.951	4.424*
	5 year	861.00*	83.34*	-1.19*	—	86.89*	163.74	141.64	0.962	4.938*
Cubic	Original	326.21*	168.48*	-5.50*	0.06*	85.30*	206.53	171.05	0.991	1.767
	3 year	423.00*	173.80*	-6.01*	0.07*	92.61*	132.73	117.45	0.996	4.825*
	4 year	455.55*	176.25*	-6.27*	0.08*	93.43*	120.85	108.92	0.957	4.729*
	5 year	505.55*	175.05*	-6.34*	0.08*	93.91*	111.59	101.38	0.957	4.321*
Exponential	Original	1113.33*	0.021*	—	—	66.8*	338.96	291.78	0.965	2.653
	3 year	1171.57*	0.020*	—	—	68.2*	294.39	249.26	0.947*	3.618*
	4 year	1200.86*	0.019*	—	—	67.5*	285.13	238.95	0.943*	4.728*
	5 year	1234.70*	0.019*	—	—	66.7*	274.28	231.13	0.939*	4.014*

* Significant at 5% level.

The five year moving average approach data of third degree polynomial model showed comparatively lower value of RMSE and MAE with significant regression coefficient but it did not satisfied the criteria for randomness of residuals. The residuals of all the fitted models were normally distributed but none of the fitted models were randomly distributed except cubic model and exponential model of original data approach. Therefore, cubic model on original data approach was found suitable to explain the trend of wheat production in Gujarat state.

By using selected (cubic of original data approach) model, predicted values and their error percentage are given in Table 4. Predicted values of cubic model were under estimated during the year 2009-10 and remaining were over estimated and their predictions are good in all years (except in the year 2007-08).

The results of fitted polynomials are given in Table 4. The results indicated that the coefficient of determination (R^2) values were significant in all the models on original data as well as moving averages approach and it ranges from 66.8 to 93.91 per cent. The regression constants of all models were significant. Regression coefficient values were significant in all polynomial and exponential models in all data approaches. In original data approach the value of R^2 was increased by 9.14 and 14.35 per cent in quadratic and cubic regression as compared to linear regression model, respectively. By taking five year moving averages, R^2 value was increased by 2.51, 6.81 and 8.61 per cent in cases of first, second and third degree polynomial models, respectively over the original data. But in exponential model R^2 values were not improved due to moving averages. Thus, very less improvement in R^2 value was observed due to moving average approached in productivity trend of wheat in Gujarat state.

Table 5: Testing of forecast values for remaining four years by using selected model i.e. cubic model in original data approach wheat productivity in Gujarat state

Years	Observed Values	Cubic Predicted values	Error Per cent
2007-08	3013	2695.15	10.55
2008-09	2375	2771.97	-16.71
2009-10	2701	2856.21	-05.77
2010-11	2603	2948.26	-13.25

The third degree model on five years moving averages showed comparatively lower value of RMSE and MAE with significant regression coefficients, but it did not satisfied the criteria of residuals for randomness. The assumption of normality of residuals was satisfied by all the polynomial models in all approach except the moving average approaches of exponential model but randomness assumption was satisfied only by third degree polynomial model on original data approach. Therefore Cubic regression model on original data approach is only model to explain the trend of wheat productivity in Gujarat state.

By using cubic model of original data approach, predicted values and their error percentage are given in Table 5. Predicted values of cubic model was over

estimated during the year 2007-08 and remaining were under estimated and their predictions are good in all years. For the productivity of wheat in Gujarat state prediction values were very good by using selected models.

Table 6: Fitted models for area, production and productivity trend of wheat in Gujarat state

Aspect	Fitted model	R^2
Area	—	—
Production	$Y_t = -226.97 + 1230.44^{*}t - 50.41^{*} t^2 + 0.72^{*}t^3$	53.7
Productivity	$Y_t = 326.21^{**} + 168.48^{*}t - 5.51^{*} t^2 + 0.06^{*}t^3$	85.30

Y_t corresponds to area, production or productivity at time t.

The best fitted polynomial models for area, production and productivity of wheat in Gujarat state are given in Table 6. The table indicate somewhat increasing trend in production and productivity of wheat over the study period with R^2 values of 53.75 and 80.30 respectively. But, for the area, the best polynomial model can't be fitted due to the lack of fulfillment of their residual assumptions of normality and randomness.

Conclusion

The trend analysis indicated that the production and productivity of wheat crop increased significantly through the year 1960-61 to 2006-07 in Gujarat State. This might be attributed to the availability of good quality seed of high yielding variety and expansion of irrigation facility.

It is suggested that the model based on original data with reasonably good R^2 can be used for future prediction while model based on moving averages can be used to predict average trend value.

References

Anonymous, 2012. Complete information on Area and Production of wheat in India, Preserve Articles. http://www.preservearticles.com/2012020422629/complete-information-on-area-and-production-of-wheat-in-india.html.

Anonymous, 2013a. The Pocket book on Agricultural Statistics 2013, Directorate of Economics and Statistics, New Delhi.http://eands.dacnet.nic.in/Publication12-122013/AgricultralStats %20 inside_ website%20book.pdf.

Anonymous, 2013b. Wheat Scenario – A Snippet, Directorate of Wheat Research Karnal, Haryana (India). http://www.dwr.res.in/sites/default/files/wheatscenarioe-newsletis-1-13.pdf.

Makridakis, S., Wheelwright, S.C. and Hyndman, R.J. 2003. Forecasting Methods and Applications (3rd edition). *John Wiley and Sons*, Inc. pp. 312-372.

Montgomery, D.C., Peck, E.A. and Vining, G.G. 2003. In: Introduction to Linear Regression Analysis. *John Wiley and sons*, Inc. pp. 221-258.

Pankratz, A. 1983. Forecasting with univariate Box-Jenkins Models - Concepts and Cases. *John Wiley and Sons*, New York.

Rangaswamy, R. 2006. In: A text book of agricultural Statistics. *New Age International (P) Limited Publishers,* pp. 158-210.

Shapiro, S.S. and Wilk, M.B. 1965. An analysis of variance test for normality (complete samples). *Biometrika* **52**: 591-611.

Sidney Siegel and John Castellan, N. Jr. 1988. An analysis of variances test for normality (complete samples). *Biometrica* **52**: 591-611.