International Journal of Agriculture Sciences
ISSN: 0975-3710&E-ISSN: 0975-9107, Volume 8, Issue 53, 2016, pp.-2782-2787.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000217

Research Article

STATISTICAL MODELING TO AREA, PRODUCTION AND YIELD OF POTATO IN WEST BENGAL

DASYAM RAMESH*, BHATTACHARYYA BANJUL AND MISHRA P.

Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia, 741252, West Bengal, India
*Corresponding Author: Email-pradeepjnkvv@gmail.com

Received: October 05, 2016; Revised: October 07, 2016; Accepted: October 08, 2016; Published: November 01, 2016

Abstract- In terms of food and nutritional security, Potato is the crop, which is most important after Wheat, Maize and Rice in the World. The State West Bengal is the second leading producers of India. In this study -Area, Production and Yields of Potato in West Bengal for the period of 1963-2012 have been considered to apply different parametric models-linear, non-linear regression and time series models (Box Jenkins& GARCH); and nonparametric model. Suitable parametric model was selected on the basis of various goodness of fit criterion and assumptions of residuals. In case of nonparametric regression, optimum bandwidth was computed by method of cross validation. Here, epanechnikov-kernel was used as the weight function. Non parametric function was emerged as the one of the best fitted trend function among all selected models, wherein parametric models ARIMA (1,1,0) was identified for both area and production of Potato & ARIMA(1,1,1) was appropriate for Potato's yield. Forecasting was made by selected parametric models up to 2020.

Keywords- GARCH, Bandwidth, Cross validation and Kernel.

Citation: Dasyam Ramesh, et al., (2016) Statistical Modeling to Area, Production and Yield of Potato in West Bengal. International Journal of Agriculture Sciences, ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 8, Issue 53, pp.-2782-2787.

Academic Editor / Reviewer: Basu S., P. Sharmila

Introduction

The two emerging Asian counties *viz.* China and India together contribute nearly 1/3rd of the global potato production. In these densely populated countries, Potato is preferred largely due to its high productivity, suppleness in terms of fitting into many prevailing cropping systems and steady yields under conditions in which other crops may fail. It has become one of the most popular crops in our country for vegetable purposes. It provides a source of low cost energy to the human diet, hence it has always been the "poor man's friend". Potatoes are used for several industrial purposes such as for the production of starch and alcohol & also for dextrin and glucose. Potato production jumped from mere 1.54 MT from 0.23 Million Ha (M. ha) area in the year 1949- 50 to 45 MT from 1.96 M. ha area during 2012-13, thus making India the second largest potato producer in the world after China. West Bengal stands as the second largest producer of potato and accounts for about 24 percent of production in the country, after Uttar Pradesh. The state West Bengal had produced 11291 thousand tones of potato during the year 2012-13. Hence, appropriate trend fit is very vital in an economic system for such important crop as it would be easier to originate and commence suitable policy measures if data with regard to the trend of production is obtained and analyzed in advance. Present investigation is planned to study the trends of area, production and yield of Potato in west Bengal by using parametric models i.e., linear-nonlinear regression, ARIMA& GARCH and nonparametric regression (Kernel) models.

Materials and Methods

Annual data with respect to Potato's area, production and yield in West Bengal for period of 1963-2012 was collected from Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Government of India. The study is dealing with time series, so before analysis only these data sets have been verified initially for existence of outlier and randomness. In this work, Grubbs test was selected for detecting outlier in time series data, as this test is chiefly useful in case of large sample and easy to follow. For checking randomness of the observations, Turning point test was used in the present study. Firstly, number of turning points i.e. peaks and troughs in the series are determined and this value forms the test statistic. For large sample, the dataset may be assumed to follow a normal distribution [1]. Descriptive statistics are used to explain the basic features of the data in any study. The selected descriptive measures along with simple growth rates have been used to explain behavior of each series in this study. Simple growth rate (SGAR) per Annum has been calculated by using the following formula: $SGR(\%) = \frac{X_t - X_0}{X_0 \times n} \times 100$;where X_t is the value of the series for the last period and X_0 is the value of the series for first period and n is the total number of periods [2].

Parametric Regression model: Selected parametric regression models like Linear, Quadratic, Cubic, Logarthmic, Exponential, Hyperbolic, Power, Compound and Gompertz have been applied for modeling of area, production and yield of Potato in West Bengal. The models are given in the following equations.

(i) Linear: $Z_t = a + bt + e_t$ (ii) Quadratic: $Z_t = a + bt + ct^2 + e_t$
(iii) Cubic: $Z_t = a + bt + ct^2 + dt^3 + e_t$ (iv) Logarthmic: $Z_t = a + b\ln(t) + e_t$
(v) Exponential: $Z_t = a[\mathrm{Exp}(bt)] + e_t$ (vi) Hyperbolic: $Z_t = a + (b/t) + e_t$
(vii) Power: $Z_t = a\,t^b + e_t$ (viii) Compound: $Z_t = a\,b^t + e_t$
(ix) Gompertz: $Z_t = a[\exp(-\exp(b-ct))] + e_t$

where a is constant; b, c, d represents regression coefficient; t and e_t are time, error term respectively in the models.

ARIMA model a non seasonal ARIMA model is represented by ARIMA (p,d,q) which is a combination of Auto Regressive (AR) and Moving Average (MA) with an order of differencing or integration (d), where p and q are the order of autocorrelation and moving average respectively [3].

ARIMA in general form is as follows: $Z_t = a + (\emptyset_1 Z_{t-1} + \ldots + \emptyset_p Z_{t-p}) - (\theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}) + e_t$

ARIMA methodology consists of four steps *viz.* model identification, model estimation, diagnostic checking and forecasting [4]. Model identification by ARIMA (p, d, q) is based on the concept of time-domain analysis i.e. autocorrelation function (ACF) and partial autocorrelation function (PACF). In this, present study, Augmented Dickey Fuller (ADF) test has been used to find unit root in the time series data under consideration [5] for identification of data stationarity. After identification of the appropriate p and q values for the model, the parameters of the autoregressive and moving average terms have been estimated.Statistical package SAS was used here to estimate parameters.For evaluating the adequacy of selective process, various reliability statistics along with residual plots for ACF and PACF have been used. In the present study, normality and randomness of residuals were tested by Shapiro-Wilk and Run tests respectively. The model with minimum values of Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Akaike Information Criterion (AIC), Schwarz Bayesian Information Criterion (SBC) and with high value of coefficient of determination (R^2) are considered as appropriate to select model of the particular data series [6].

GARCH

It means Generalized Autoregressive Conditional Heteroskedasticity. GARCH is a method that includes past variances in the explanation of future variances. Morespecifically, GARCH is a time series technique that allows users to model and forecast the conditional variance of the errors. If an ARMA model is assumed for the error variance, the model is called GARCH[7].If the sum of ARCH and GARCH coefficientsclose to 1, it indicates that volatility is quite persistent in the selected series.

To measure the extent of series volatility, GARCH (1, 1) model is specified as:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

α_0 - Constant term

ε_{t-1}^2 - ARCH term this is the news about volatility from the previous period, measured as the lag of the squared residual from the mean equation model,

h_{t-1} - GARCH term, it is the last periods forecast variance.

In this present study, initially residuals of mean equation model is tested for ARCH-LM (ARCH-Lagrange Multiplier) test, if found significance then only GARCH to be applied and the same test again applied at end, as to check weather fitted GARCH model has still any ARCH effect. If not, then that selective model has to be further verified for normality and randomness of residuals.

Among the competitive models, best models are selected based on minimum value of Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE), maximum value of Coefficient of Determination (R^2) and of course the significance of the coefficients of the models. Best fitted models are also verified through ACF and PACF plots of the residuals. Finally the selected parametric model is used to forecast up to 2020 by keeping the years 2012 (In-sample) and 2013(Out-sample) are for model validation.

Nonparametric regression

In general, nonparametric regression model is of the form $Z_i = m(x_i) + \varepsilon_i$ where Z is the response variable. The mean response E(Z|X=x) or regression function m(X) is assumed to be smooth and ε is the independently and identically distributed random error with mean zero. In this nonparametric regression, the optimum bandwidth estimation was done by cross validation method and Epanechnikov-kernel was used as the weight function [8]. Here, Matlab software was used to estimate optimum band width and trend. After fitting the model, residual analysis was carried out to test the randomness.

Results And Discussion

Univariate time series data on area, production and yield of Potato in West Bengal from 1963-2012 were examined for randomness and outliers by Turning point test and Grubbs method respectively. From [Table-1], it was clear that the series under consideration having no outliers and only the yield had random pattern.

Table-1 *Test for randomness and outliers for area, production and yield of Potato in West Bengal*

Potato	No. of Observations	No. of Turnings (p)	Mean E(P)	Variance V(P)	Test statistic (T_{cal})	Inference	Outlier
Area	50	24	32	8.567	2.733	Trend	No
Production	50	24	32	8.567	2.733	Trend	No
Yield	50	28	32	8.567	1.366	Random	No

From[Table-2], descriptive statistics were calculated and represented for area, production and yield of Potato.It was observed that area under Potato had varied from 62.4 to 407.9 ('000 ha) with an average of 206.37 ('000 ha), registering a simple growth rate of almost 9.787% per annum. Similarly, the average values of production and yield were 4358.268 ('000 tonne), 190.6 (Q/ha) with simple growth rate of 40.203%, 5.349% per annum respectively.

Table-2 *Descriptive statistics for area, production and yield of Potato in West Bengal*

	Potato Area ('000 ha)	Potato Production ('000 tonne)	Potato Yield (Q/ha)
Mean	206.370	4358.268	190.624
Maximum	407.9	11291.3	299.82
Minimum	62.4	535.1	81.6
Standard Deviation	114.806	3084.608	55.774
Skewness	0.353	0.424	-0.617
Kurtosis	1.695	2.058	2.367
CV (%)	55.631	70.776	29.259
SGAR (%)	9.787	40.203	5.349

Before analyzing by time series models, selected linear-nonlinear regression models were applied to all the datasets under consideration. Estimated parameters and goodness of fit for the models were depicted in [Table-3 & 4] for area and production cultivation. It was revealed from the results that among the fitted models, the maximum R^2value of 95% was observed in case of Cubic model with minimum values of RMSE (20.91) and MAPE (10.17) in comparison to those of the other models. However, the residual analysis confirmed that the assumptions of independence (by Run test) of error terms were failed by all the models employed for both area and production. Hence it was concluded that none of the selected nonlinear regression models was found suitable to fit the cultivable Potato area and its production in West Bengal. Similar kind of findings was reported by [9] who studied the linear and nonlinear models to fit the area of castor in Anand district of Gujarat. In case of yield of Potato – Gompertz model was appeared to be most plausible due to highest R^2value and lowest values of other diagnostic measures. This model also satisfied the normality and randomness properties of residuals as shown in [Table-5].

Before employing of ARIMA technique, stationary of data series was tested first. For this, Augmented Dickey Fuller (ADF) test was applied. From [Table-6], it was revealed that all the three data series were non stationary and became stationary at first difference

As per autocorrelation and partial autocorrelation considerations, suitable ARIMA (p,d,q) models were selected and compared to each other as depicted in [Table-7]. In all of these models, ARIMA (1,1,0) was appropriate in case of area as due to highest value of R^2 and lowest values of other criterion. Normality and randomness properties of residuals were also satisfied, as these were non-significant. From the residual ACF and PACF plots of ARIMA (1,1,0) it was clear that all autocorrelations and partial autocorrelations lie between 95% control limits as represented in [Fig-1]. This also confirmed the 'good fit' of the selected model.

Equation of the ARIMA model was formulated as: Potato Area$_t$ (Z_t)=6.54– 0.32Z_{t-1} + e_t. ARCH-LM test was found to be non-significant for area as shown in [Table-7]. Hence none of GARCH models were developed.

Table-3 *Fitting of nonlinear models for area underPotato production in West Bengal*

Model	Parameter Estimates				Goodness of Fit					
	a	b1	b2	b3	RMSE	MAPE	MAE	R^2	SW test	Run test
Linear	11.31	7.65*			27.02	16.41	22.64	0.93	0.07	0.01
Quadratic	50.82*	3.09*	0.09*		21.29	10.29	16.66	0.95	0.39	0.01
Cubic	84.02	-4.35*	0.45*	-0.01*	20.91	10.17	13.98	0.95	0.12	0.02
Logarthmic	-107.26*	105.61*			65.38	41.14	56.85	0.67	0.01	0.01
Exponential	60.78*	0.04*			28.33	10.24	19.96	0.93	0.51	0.01
Hyperbolic	236.51*	-334.88*			100.89	62.53	88.82	0.21	0.01	0.01
Power	28.23*	0.61*			46.98	24.51	39.77	0.83	0.03	0.01
Compound	60.77*	1.04*			27.99	10.19	20.04	0.94	0.01	0.01
Gompertz	1173.67	1.19*	0.03*		22.12	10.95	17.11	0.95	0.13	0.01

* Significant at 5% level:

Table-4 *Fitting of nonlinear models for production of Potato in West Bengal*

Model	Parameter Estimates				Goodness of Fit					
	a	b1	b2	b3	RMSE	MAPE	MAE	R^2	SW test	Run test
Linear	-875.96*	205.26*			741.79	27.46	569.28	0.93	0.21	0.01
Quadratic	100.89*	92.55*	2.21*		657.25	17.51	432.72	0.94	0.09	0.01
Cubic	564.56*	-11.44*	7.26*	-0.07*	653.36	16.85	426.25	0.95	0.05	0.03
Logarthmic	-4119*	2855*			1735.1	76.89	1421.9	0.67	0.01	0.01
Exponential	649.44*	0.06*			1323.5	19.95	883.42	0.81	0.01	0.01
Hyperbolic	5178*	-9110*			2701.7	126.61	2336.1	0.22	0.02	0.01
Power	177.66*	0.96*			992.61	27.86	749.50	0.89	0.02	0.01
Compound	649.44*	1.06*			1248.06	19.67	866.06	0.83	0.01	0.01
Gompertz	17656*	1.41*	0.03		659.43	17.77	446.88	0.94	0.01	0.18

* Significant at 5% level:

Table-5 *Fitting of nonlinear models for yield of Potato in West Bengal*

Model	Parameter Estimates				Goodness of Fit					
	a	b1	b2	b3	RMSE	MAPE	MAE	R^2	SW test	Run test
Linear	100.28*	3.54*			20.85	11.12	17.51	0.85	0.41	0.01
Quadratic	69.52*	7.09*	-0.07		16.45	8.71	13.13	0.89	0.83	0.15
Cubic	61.64	8.86	-0.16	0.01	16.33	8.66	13.03	0.89	0.89	0.51
Logarthmic	20.07	57.43*			22.19	12.04	16.36	0.84	0.04	0.03
Exponential	104.65*	0.02*			27.43	13.53	22.91	0.75	0.07	0.01
Hyperbolic	210.32*	-218.88*			43.35	25.04	35.47	0.38	0.02	0.01
Power	61.11*	0.37*			17.84	10.09	14.81	0.89	0.62	0.01
Compound	104.71*	1.02*			27.15	13.46	22.83	0.76	0.06	0.01
Gompertz	263.26	0.29*	0.07*		16.21	8.45	12.87	0.89	0.52	0.07

* Significant at 5% level:

.

Table-6 *Result of ADF test for area, production and yield of Potato in West Bengal*

Potato	Data type	ADF statistic	Critical values at			Decision
			1%	5%	10%	
Area	ADF at level	-0.019	-3.5713	-2.9228	-2.5990	Data Non-Stationary
	ADF at 1st difference	-6.656	-3.5745	-2.9241	-2.5997	Data became Stationary
Production	ADF at level	-0.212	-3.5713	-2.9228	-2.5990	Data Non-Stationary
	ADF at 1st difference	-9.298	-3.5745	-2.9241	-2.5997	Data became Stationary
Yield	ADF at level	-1.543	-3.5713	-2.9228	-2.5990	Data Non-Stationary
	ADF at 1st difference	-7.561	-3.5745	-2.9241	-2.5997	Data became Stationary

Table-7 *ARIMA model fit statistics for area under Potato production in West Bengal*

Model	Parameter Estimates				Goodness of Fit							Arch LM test
	a	Autoregressive Coefficient		Moving Average Coefficient	RMSE	MAPE	R^2	AIC	SBC	SW test	Run test	
		AR1	AR2	MA1								
(1,1,1)	6.82*	0.25*		0.65*	20.80	9.60	0.96	303.43	309.11	0.12	0.66	0.77
(1,1,0)	6.54*	-0.32*			18.15	7.92	0.97	301.94	305.73	0.19	0.34	0.94
(0,1,1)	6.66*			0.43*	21.40	8.66	0.96	304.23	308.01	0.13	0.61	0.89
(2,1,0)	6.62*	-0.38*	-0.21		20.88	9.16	0.96	303.81	309.48	0.16	0.66	0.90
(2,1,1)	6.84	0.28	0.03	0.68	20.79	9.58	0.96	305.41	312.98	0.11	0.62	0.79

* Significant at 5% level

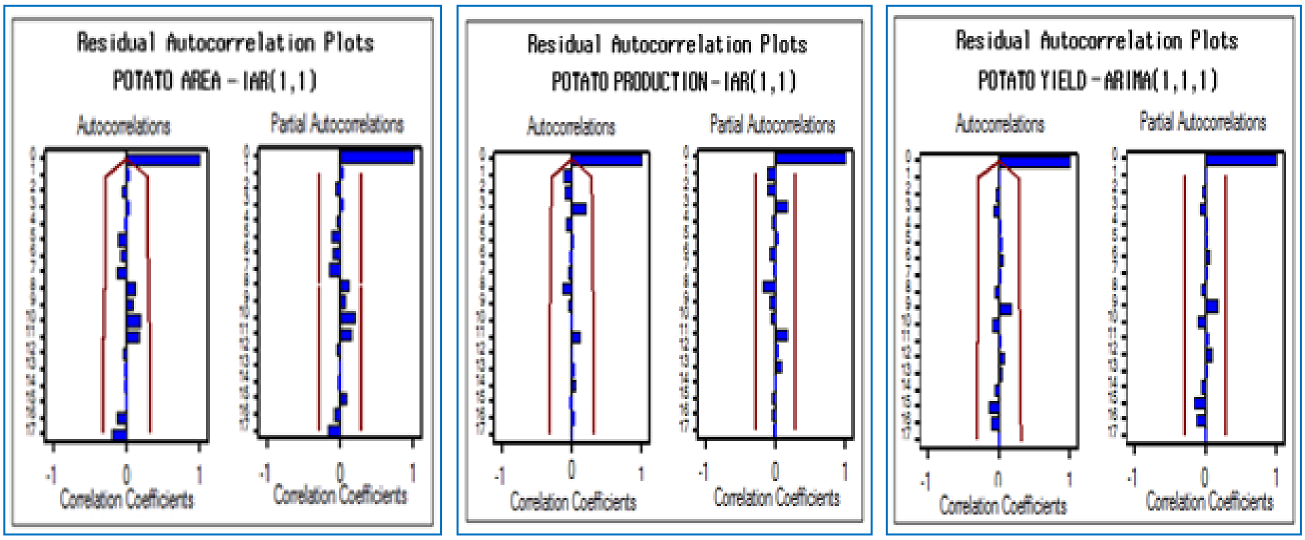


Fig-1 Residual ACF and PACF of ARIMA(1,1,0), ARIMA(1,1,0) and ARIMA(1,1,1) models

After employing selected parametric models, nonparametric regression (Kernel) model was also tried to obtain the trend of Potato area in West Bengal. In this, optimum bandwidth was computed as 0.08 by cross validation method. Using epanechnikov kernel, trend was estimated with the following diagnostic criterion i.e., RMSE (14.36), MAPE (5.61), MAE (10.79), MSE (206.25) and R^2 (0.98). From [Fig-3], it was observed that actual and forecasts were closely related with increasing trend.Residuals were distributed independently as probability value of run test was found to be 0.623 i.e. nonsignificant. Hence nonparametric model was considered as one of the best fit for modeling the area under Potato production in West Bengal.

Table-8 *Model fit statistics for area, production and yield of Potato in West Bengal by Nonparametric regression*

Potato-Area		Potato-Production		Potato–Yield	
MSE	206.25	**MSE**	335063	**MSE**	194.23
RMSE	14.36	**RMSE**	578.84	**RMSE**	13.93
MAPE	5.61	**MAPE**	12.74	**MAPE**	7.24
MAE	10.79	**MAE**	322.04	**MAE**	11.13
R-square	0.98	**R-square**	0.97	**R-square**	0.93

Table-9 *ARIMA model fit statistics for production and Yield of Potato in West Bengal*

Model	Parameter Estimates					Goodness of Fit							Arch LM test
	a	Autoregressive Coefficient		Moving Average Coefficient		RMSE	MAPE	R^2	AIC	SBC	SW test	Run test	
		AR1	AR2	MA1									
(1,1,1)	209.51*	-0.28*		0.26*		644.62	15.24	0.96	639.71	643.97	0.11	0.27	0.03
(1,1,0)	209.92*	-0.48*				634.01	14.42	0.96	637.42	641.71	0.15	0.25	0.09
(0,1,1)	208.85*			0.49*		642.19	15.23	0.96	639.56	642.94	0.09	0.31	0.10
(2,1,0)	210.71*	-0.58*	-0.24			636.79	15.14	0.96	638.18	642.86	0.08	0.19	0.11
Yield of Potato in West Bengal													
Model	Parameter Estimates					Goodness of Fit							Arch LM test
	a	Autoregressive Coefficient		Moving Average Coefficient		RMSE	MAPE	R^2	AIC	SBC	SW test	Run test	
		AR1	AR2	MA1	MA2								
(1,1,1)	3.86*	0.27*		0.71*		15.56	7.45	0.91	280.31	284.99	0.13	0.19	0.61
(1,1,0)	4.15*	-0.29*				15.95	7.86	0.90	283.67	287.46	0.09	0.08	0.52
(0,1,1)	3.88*			0.52*		15.81	7.63	0.90	281.56	285.34	0.11	0.65	0.76
(2,1,0)	4.03*	-0.35*	-0.22			15.89	7.72	0.89	283.59	289.27	0.07	0.31	0.58
(1,1,2)	3.85*	0.09		0.52	0.12	15.58	7.47	0.90	283.19	290.76	0.05	0.08	0.63

* Significant at 5% level:

It was observed from [Table-6] that both production and yield of Potato became stationary at first difference. So by fixing d=1, different ARIMA models were tried and tabulated in [Table-9]. Then,ARIMA(1,1,0) was selected for Potato production as due to high R^2 (0.96) and low values of MAPE (14.42) AIC (637.42)and SBC (641.71). Similarly, in case of Potato yield also ARIMA(1,1,1) was selected. Residuals of these models were satisfied both the normality and randomness assumptions as shown in [Table-9]. All the estimated parameters, in both the cases, were significant at 5% significant level as well asresidual ACF, PACF also confirmed the 'good fit' of these selected models as shown in [Fig-1]. Equations of the ARIMA model for production and yield of Potato in West Bengal were formulated as:

$PotatoProduction_t\ (Z_t) = 209.92 – 0.48\ Z_{t-1} + e_t$

$PotatoYield_t\ (Z_t) = 3.86 + 0.27\ Z_{t-1} – 0.71 e_{t-1} + e_t$

Residuals of all selective ARIMA models were applied for ARCH-LM test at various lags, but it was found that none of model was significant as shown in [Table-9], except in Potato production. So GARCH model was developed by using this significant model and their estimated parameters and diagnostic checks were depicted in [Table-10]. It was observed that performance of this model by selective diagnostic checks was not superior to earlier discussed models. Hence, among selected parametric models - ARIMA(1,1,0) was considered as appropriate for modeling of potato production in West Bengal.

Table-10 *GARCH model fit statistics for production of Potato in West Bengal*

Model	Mean Equation			Variance Equation	
	Constant	AR(1)	MA(1)	Constant	ARCH effect (α_1)
ARMA(1,1) -GARCH (0,1)	190.478*	-0.688*	0.925*	99741.5*	0.840*
GARCH(0,1)	156.82*			121690.2*	0.735*

* Significant at 5% level:

Goodness of Fit statistics

Model	RMSE	MAPE	MAE	R^2	SW test	Run test	ARCH-LM
ARMA(1,1) -GARCH (0,1)	684.641	17.377	454.99	0.934	0.13	0.652	0.872
GARCH(0,1)	688.764	17.632	469.737	0.921	0.11	0.658	0.593

Then, Nonparametric regression model (Kernel) was also employed by computing optimum bandwidth as 0.10 and 0.18, for production and yield respectively. By using Epanechnikov kernel smoothing, trend of Potato production was estimated with the selected diagnostic criteria i.e., RMSE (578.84), MAPE (12.74), MAE (322.04), MSE (335063) and R^2 (0.97); where for Potato yield, these were estimated as RMSE (13.93), MAPE (7.24), MAE (11.13), MSE (194.23) and R^2 (0.93). Diagnostic criteria were slight lower than those of the earlier parametric models in both the cases. Residuals of models were distributed independently as probability value of run test was found to be 0.556 and 0.274 i.e. not significant, for production, yield respectively. Hence nonparametric model was considered as one of the best fit for modeling to the production and yield of Potato in West Bengal.From [Fig-1], it was observed that both production and yield of potato having increasing trend.

Finally among selected parametric models, ARIMA (1, 1, 0) was considered as suitable to area as due to various selected criterion, already discussed. Modeling as well as forecasting was made by this model up to 2020. From [Table-11], it was revealed that both in sample forecast (for the year 2012) as well as out sample forecast (for the year 2013) were very close to actual values of Potato area. It was resulted that predicted area would be decease slightly between the years 2013-15 as compared to actual data of 2013; thereafter having incremental trend up to the year 2020. It was forecasted as 410.67('000 ha) and 443.97('000 ha) for the year 2015, 2020 respectively. Similarly in case of Production, ARIMA (1,1,0) was selected and it was forecasted as 11417 ('000 tonne) and 12519 ('000 tonne) for the year 2015 and 2020 respectively. In case of Yield, it was found that forecasts for the year 2015, 2020 would be 287.84 (Q/ha) and 306.642 (Q/ha) respectively.Forecasted graphs by selected ARIMA models was depicted in [Fig-2], which also indicating the increasing trend.Similarly, nonparametric models also having good criterion, so the trend graphs of area, production and yield were fitted by Kernel regression was depicted in [Fig-3].

Table-11 *Model validation as well as forecasts of area, production and yield of Potato in West Bengal*

	Area of Potato by ARIMA (1,1,0)			Production of Potato by ARIMA (1,1,0)			Yield of Potato by ARIMA (1,1,1)		
Year	Actual Area ('000 ha)	Predicted Area ('000 ha)	Absolute Forecast Error	Actual Production ('000 tonne)	Predicted Production ('000 tonne)	Absolute Forecast Error	Actual Yield (Q/ha)	Predicted Yield (Q/ha)	Absolute Forecast Error
2012	386.61	395.97	0.02	11291	10638	0.05	299.82	260.57	0.13
2013	412.30	397.36	0.04	9030	10229	0.13	219	286.14	0.3
2014		404.02			11364			285.26	
2015		410.68			11417			287.84	
2016		417.34			11703			291.34	
2017		424.00			11876			295.11	
2018		430.65			12104			298.94	
2019		437.31			12305			302.78	
2020		443.97			12519			306.64	

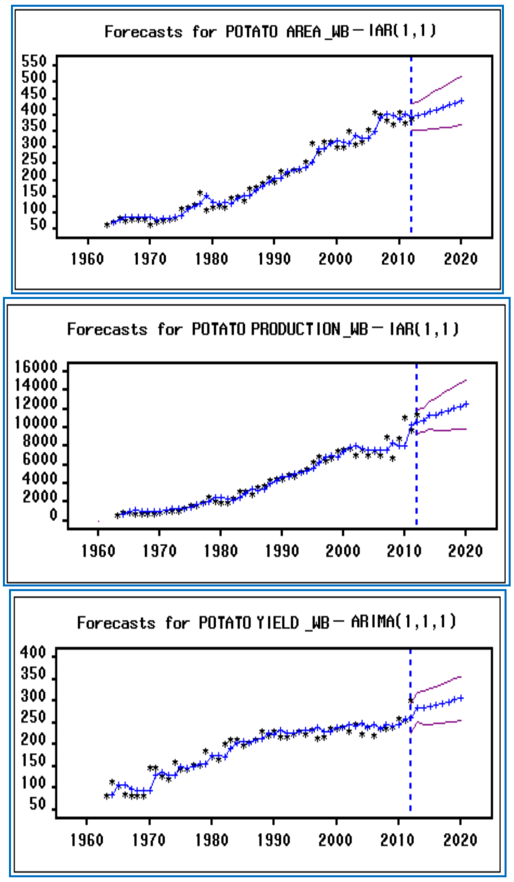
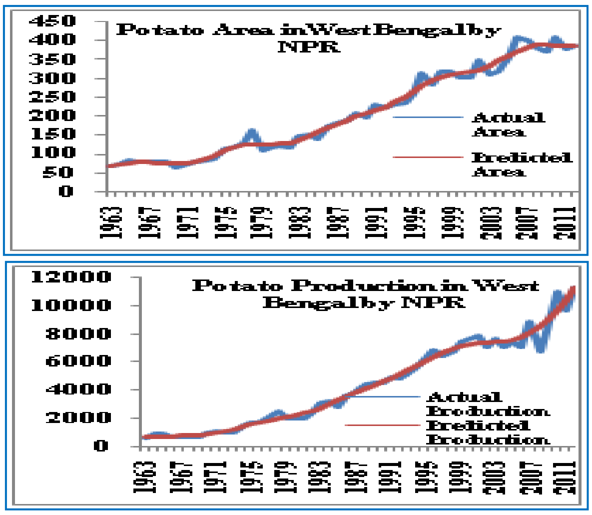


Fig-2 Forecasting for area, production and yield of Potato in West Bengal by ARIMA model.



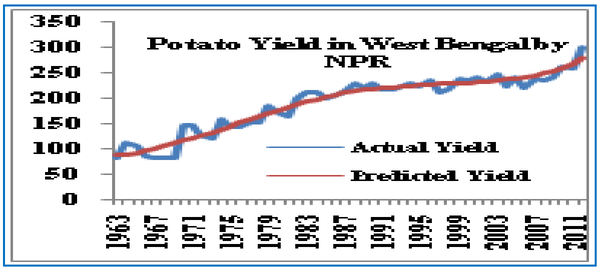


Fig-3 Trends for area, production and yield of Potato in West Bengal by Nonparametric model

Conclusions

In case of Potato, both the area and production of series under consideration having definite trend; where for yield, pattern is random. From calculated CV%, it can be observe that production is highly fluctuating as compared to area and yield. Simple growth rate per annum (SGAR) is higher in case of production over the study period. Here also, none of linear and nonlinear models are found suitable for both area and production; but for yield, gompertz is emerged as one of the fitted model based on residual properties (normality & randomness). In case of area and yield, no ARCH effect is identified; hence none of ARCH/GARCH models are formulated; only ARIMA (1,1,0) and ARIMA (1,1,1) models are found to be suitable based on diagnostics. But for production, in addition to ARIMA (1,1,0), ARMA (1,1)-GARCH(0,1) model also developed; but forecasting is made by ARIMA (1,1,0) only as due to its better diagnostics. Finally, forecasts of potato area, production and yield for the year 2020 are obtained as 443.97 ('000 ha), 12519 ('000 tonne) and 306.64 (Q/ha) respectively by selected parametric models.Based on estimated trends of nonparametric regression, all three series under consideration have increasing trend.

Acknowledgement

The financial assistance received by the first author in the form of Senior Research Fellowship (PGS) from Indian Council of Agricultural Research (ICAR) is highly acknowledged.

Conflict of Interest: None declared

References

[1] Kanji G. K. (2006) 100 Statistical tests, Sage publications, New Delhi.
[2] Sahu P.K. (2010) *International Journal of Statistical Sciences,* 10, 71-92.
[3] Gujarati D., Porter D. and Gunasekar S. (2012) Basic Econometrics, Mc Graw Hill, New Delhi.
[4] Sankar J. (2011) *Recent Research in Science and Technology,* 3, 104-108.
[5] Dickey D.A. and Fuller W.A. (1979) *Journal of American Statistical Association,* 74, 427-431.
[6] Shafaqat M. (2012) *Journal of Contemporary Issues in Business Research,* 1, 41-54.
[7] Bollerslev T. (1986) *Journal of Econometrics,* 31(3), 307-327.
[8] Hardle W. (1990) Applied Nonparametric Regression, Cambridge University Press, New York, USA.
[9] Rajarathinam A. and Parmer R.S. (2011) *Asian Journal of Applied Statistics,* 4(1), 42-52.