Advances in Agricultural Technology & Plant Sciences
ISSN: 2640-6586

Academic Strive
Researcher's perception is our reality

Research Article *Volume 7 Issue 1*

A Study of the Fuzzy Linear Regression Model and the Least Square Approach for Depicting the Association between Potato (*Solanum tuberosum* l.) Yield and Growth Attributing Characters

Rana L[1]*, Banerjee H[2], Mishra M[3], Rajput J[4] and Kumar N[1]

[1]Department of Agronomy, Dr. Rajendra Prasad Central Agricultural University, India

[2]Regional Research Station (CSZ), Bidhan Chandra Krishi Vishwavidyalaya, India

[3]Department of Statistics, Babasaheb Bhimrao Ambedkar University (A Central University), India

[4]Division of Agricultural Engineering, Indian Agricultural Research Institute, India

***Corresponding author:** Lalita Rana, Department of Agronomy, Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur, Bihar, 848125, India, Email: lalita@rpcau.ac.in

Received Date: December 23, 2023; **Published Date:** January 10, 2024

Abstract

A field experiment was conducted at Adisaptagram Block Seed Farm in Hooghly, West Bengal, in two consecutive years with the goal of developing a fuzzy linear regression model that used crisp input/output to investigate the relationship between tuber yield (response variable) and its attributing characters (explanatory variables) to assess the mode. Tanaka proposed that the Fuzzy linear regression model be employed in this research, which is based on a linear programming problem to estimate the regression coefficient as a fuzzy integer. When evaluating the Fuzzy linear regression model to the normal multiple linear regression model, the width prediction interval and average width, which are considered as model accuracy for both models, are used. The projected interval computed using the Fuzzy linear regression model has a much narrower average width when compared to the method of least squares. The most important takeaway from this research is that fuzzy linear regression is the best method for determining the relationship between tuber yield and the factors that affect it.

Keywords: Fuzzy Linear Regression; Multiple Linear Regression; Tuber Yield

Abbreviation: Leaf Area Index (LAI); Dry matter accumulation (DMA).

Introduction

Potato is the world's fourth most important food crop, next to maize, rice, and wheat, with annual production of approximately 300 million tons [1], cultivated on around 20 million hectares of arable land [2]. Potato farming produces a much higher value crop per hectare than cereal farming because potatoes are a rich source of starch and have a high biological value protein [3]. After Uttar Pradesh, West Bengal is the second-largest producer of potatoes. The most of West Bengal districts with considerable potato growing areas in winter (*Rabi*) season, contribute to a great extent in total potato production of the state.

Because land competition from industrialization and other sectors makes it impossible to increase potato crop fields, more potato output must be achieved with the same or even

less area. In order to improve potato yield per unit area, this need necessitates the use of proper potato production techniques. Fertilizer usage is becoming more widespread in current agricultural practices [4]. Fertilizer use has skyrocketed, resulting in significant increases in agricultural output all across the world [5]. Regression analysis is used to describe the statistical connection between explanatory and response variables in a range of situations. A statistical regression model may only be performed if the given data are distributed according to the statistical model and the link between explanatory and response variables is apparent.

In addition to developing a system that can cope with ambiguous and vague words or information, fuzzy uncertainty with ambiguity and vagueness introduces fuzzy theory [6]. Tanaka, et al. [7] first propose the fuzzy linear regression model, which uses a linear programming model to calculate the regression coefficient as fuzzy numbers. Scientists continued to improve the method [8-10]. Redden, et al. [11] imply that their methodology sensitive to outliers. The dispersion of the anticipated response expands as more data is supplied to the model. The fuzzy least squares approach, developed by Diamond [12] eliminates the sum of the total error of the squares from the result.

Materials and Methods

A field trial was performed under irrigated situations during the winter (*rabi*) season at Adisaptagram Block Seed Farm, Hooghly, West Bengal, employing potato cultivar var. Kufri Jyoti under irrigated conditions for two consecutive years, 2012-13 and 2013-14 which lies at $23^0 26'$ North latitude and $88^0 22'$ East longitude with an elevation of 12 m above mean sea level. The soil was clay, slightly acidic (*p*H 6.19), medium in organic carbon (0.78%), available nitrogen (84.70 kg/ha), available phosphorus (193.49 kg/ha), available potassium (251.41 kg/ha). Seven treatment (T_1- 50% NPK, T_2- 100% NPK, T_3- 150% NPK, T_4- 100% PK (-N), T_5- 100% NK (-P), T_6- 100% NP (-K), T_7- Control) were laid out in randomized block design replicated thrice. The plots were 3.5 m long and 3 m wide. Fertilizers were applied in each plot with varied NPK doses. The response variable is tuber yield (Y) and its attributing characters are the explanatory variables viz., plant height, no. of haulms/hill, leaf area index (LAI), dry matter accumulation (DMA). Only the data pertaining to maturity level, i.e. 90 days after planting (DAP), are considered for data analysis using SAS, version 9.3 software package and MS Excel software.

Least Square (LS) Method

MLR modelling (multiple linear regressions) is a enormously sturdy method this is regularly implemented in agricultural research. The linear connection among dependent (response) and independent (explanatory) variables is predicted the usage of this method. The version is written as if Xi, i=1,2,...,n are explanatory variables and Y is the response variable:

$$Y = b_0 + b_1X_1 + ... + b_nX_n + e \quad (1)$$

where b's are parameters and e is the error term assumed to be following a normal distribution. The parameters are generally estimated using method of least squares. A good description of various aspects of multiple linear regression methodology is given by Draper, et al. [13]

Fuzzy Linear Regression (FLR) Method

A fuzzy regression model corresponding to equation (1) can be written as:

$$Y = A_0 + A_1X_1 + ... + A_nX_n \quad (2)$$

Here explanatory variables X_i's, as before, are assumed to be precise. However, as mentioned above, response variable Y is not crisp but instead fuzzy in nature. This implies that the parameters are also fuzzy in nature. Our aim is to estimate these parameters, it is assumed that A_i's are symmetric fuzzy numbers (i.e. vagueness is expressible as equidistant from the center) and so can be represented by intervals. For example, A_i can be expressed as fuzzy set given by:

$$A_1 = < a_{1c}, a_{1w} > \quad (3)$$

where a_{1c} is centre and a_{1w} is radius or vagueness associated. The above fuzzy set describes perception of regression coefficient around in terms of symmetric triangular club function. This method is carried out while the underlying phenomenon is fuzzy this means that that the reaction variable is fuzzy and the relationship is likewise taken into consideration to be fuzzy. Equation (3) is sometimes also written as:

$$A_1 = [a_{1L}, a_{1R}] \quad (4)$$

where $a_{1L} = a_{1c} - a_{1w}$ and $a_{1R} = a_{1c} + a_{1w}$ [14]

This can be visualized as the LP problem and solved by using "simplex procedure" [15].

Results and Discussion

The MLR and FLR version are fitted for displaying the connection among tuber yields and its attributing characters the use of SAS, model 9.three software program package deal and following effects have been obtained (Table 1 & Table 2).

Variable	D.f	Parameter Estimate	Standard Error	T Value	Pr > |t|
Intercept	1	-4.75275	2.79965	-1.7	0.1238
Plant height	1	0.30804	0.08246	3.74	0.0047
No. of haulms/hill	1	0.54761	1.02929	0.53	0.6076
LAI	1	0.79688	1.24975	0.64	0.5396
DMA	1	0.00916	0.00543	1.69	0.1257

Source: Leaf Area Index (LAI); Dry matter accumulation (DMA).
Table 1: Multiple linear regression (MLR) estimates using SAS software.

[1]	Ac	Aw
1	-3.914731	0.5670086
2	0.269469	0.0070007
3	-0.092111	0
4	1.093385	0
5	0.013963	0.0017977

Table 2: FLR estimates using SAS software.

The fitted model for MLR is

$$Y = -4.75 + 0.31 * plant\ height + 0.55 * No.\ of\ haulms/hill + 0.80 * LAI + 0.01 * DMA \quad (5)$$

Standard Errors (2.80) (0.08) (1.03) (1.25) (0.005)

Upper and lower limits of prediction interval for MLR models are computed from the prediction equation (5) by taking the coefficient as their corresponding estimated values plus or minus standard error, i.e.

$$Y = (-4.75 + 2.80) + (0.31 + 0.08) * plant\ height + (0.55 + 1.03) * No.\ of\ haulms/hill + (0.80 + 1.25) * LAI + (0.01 + 0.005) * DMA$$

And

$$Y = (-4.75 - 2.80) + (0.31 - 0.08) * plant\ height + (0.55 - 1.03) * No.\ of\ haulms/hill + (0.80 - 1.25) * LAI + (0.01 - 0.005) * DMA$$

The fitted model for FLR is

$$Y = <-3.91, 0.57> + <0.27,\ 0.007> plant\ height + <-0.09, 0> No.\ of\ haulms/hill + <1.09, 0> LAI + <0.01, 0.002> DMA \quad (6)$$

Upper and lower limits of prediction interval for FLR model are computed from the prediction equation (6) respectably as,

$$Y = (-3.91 + 0.57) + (0.27 + 0.007) * plant\ height + (-0.09 + 0) * No.\ of\ haulms/hill + (1.09 + 0) * LAI + (0.01 + 0.002) * DMA$$

And

$$Y = (-3.91 - 0.57) + (0.27 - 0.007) * plant\ height + (-0.09 - 0) * No.\ of\ haulms/hill + (1.09 - 0) * LAI + (0.01 - 0.002) * DMA$$

The width of prediction intervals in respect of MLR and FLR version similar to every set of determined explanatory variables is computed in MS Excel and the outcomes are suggested in Table 3 and similarly supported through (Figure 1). From Table 3, common width for MLR version was found to be 30.43, even as that for FLR version changed into only 4.07, indicating thereby the superiority of fuzzy regression technique.

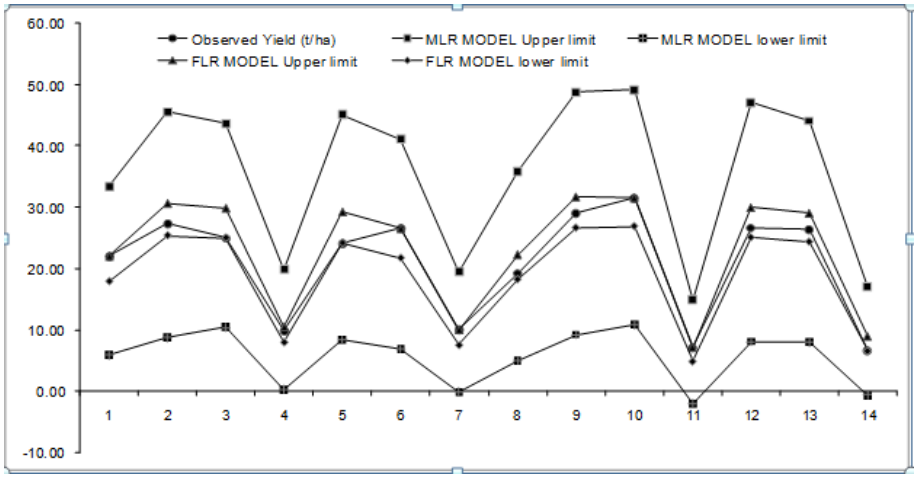

Figure 1: Comparison of Multiple linear regression (MLR) and fuzzy linear regression model (FLR) for prediction of tuber yield using upper and lower width prediction.

Observed yield (t/ha)	MLR Model			FLR Model		
	Upper limit	lower limit	width	Upper limit	lower limit	width
22.01	33.39	5.85	27.53	22.01	17.93	4.09
27.38	45.54	8.75	36.79	30.64	25.37	5.27
24.99	43.55	10.38	33.17	29.85	24.99	4.87
9.81	19.88	0.26	19.62	10.47	8.02	2.45
24.17	45.05	8.37	36.68	29.24	24.17	5.07
26.53	41.06	6.87	34.19	26.53	21.71	4.82
10.04	19.53	-0.15	19.68	10.03	7.56	2.47
19.24	35.83	4.97	30.87	22.29	18.24	4.05
29.1	48.73	9.21	39.53	31.7	26.6	5.1
31.54	49.12	10.84	38.28	31.54	26.86	4.68
7.12	14.9	-2.07	16.97	7.12	4.91	2.21
26.66	46.98	8.05	38.94	29.96	25.05	4.91
26.41	44.05	7.96	36.09	29.03	24.33	4.7
6.61	16.97	-0.69	17.66	8.93	6.61	2.33
Average width			30.43	Average width		4.07

Table 3: Fitting of Multiple linear regression (MLR) and fuzzy linear regression model FLR and the least square method.

Similar kind of findings was reported Kandala VM, et al. [16] who validated the applicability of FLR technique while the 2 explanatory variables (viz. plant height and leaf area index) and response variable (dry matter accumulation) are all crisp however underlying phenomenon is thought to be fuzzy in nature. It was proven that widths of prediction intervals in respect of FLR version have been much less than the ones for MLR model.

References

1. Naz F, Ali A, Iqbal Z, Akhtar N, Asghar S, et al. (2011) Effect of different levels of NPK fertilizers on the proximate composition of potato crop at Abbottabad. Sarhad Journal of Agriculture 27(3): 353-356.

2. Jasim AH, Hussein MJ, Nayef MN (2013) Effect of foliar fertilizer (high in potash) on growth and yield of seven potato cultivars (*Solanum tuberosom* L.). Euphrates Journal of Agricultural Science 5(1): 1-7.

3. Stephen D, Jackson A (1999) Multiple signaling pathways control tuber induction in potato. Plant Physiology 119(1): 1-8.

4. Wang W, Lu J, Ren T, Li X, Su W, et al. (2012) Evaluating regional mean optimal nitrogen rates in combination with indigenous nitrogen supply for rice production. Field Crops Research 137: 37-48.

5. Cassman KG, Dobermann A, Walters DT, Yang HS (2003) Meeting cereal demand while protecting natural resources and improving environmental quality. Annual Review of Environmental Resources 28: 315-358.

6. Zadeh LA (1965) Fuzzy sets. Inform and Control 8(3): 338-353.

7. Tanaka H, Uejima S, Asai K (1982) Linear Regression Analysis with Fuzzy Model. IEEE Transactions on Systems Man and Cybernetics 12(6): 903-907.

8. Tanaka H (1987) Fuzzy data analysis by possibilistic linear models. Fuzzy Sets and Systems 24(3): 363-375.

9. Tanaka H, Watada J (1988) Possibilistic linear systems and their application to the linear regression model. Fuzzy Sets and System 27(3): 275-289.

10. Tanaka H, Hayashi I, Watada J (1989) Possibilistic linear regression analysis for fuzzy data. European Journal of Operational Research 40(3): 389-396.

11. Redden DT, Woodall WH (1994) Properties of certain fuzzy linear regression methods. Fuzzy Sets and Systems 64(3): 361-375.

12. Diamond P (1988) Fuzzy Least Squares. Information Sciences 46(3): 141-157.

13. Draper NR, Smith H (1998) Applied Regression Analysis.

In: 3rd (Edn.), Wiley Publishers, New York, USA, pp: 736.

14. Kacprzyk J, Fedrizzi M (1990) Multiperson Decision Making Models using Fuzzy sets and Possibility Theory. Kluwer Academic Publishers, Boston, USA, pp: 356.

15. Taha HA (1997) Operational Research: An Introduction. Prentice Hall Publishers, United States, pp: 916.

16. Kandala VM, Prajneshu (2003) Application of fuzzy regression methodology in agriculture. Indian Journal of Agricultural Sciences 73: 456-458.