Plant Archives Vol. 17 No. 2, 2017 pp. 1495-1500 ***ISSN 0972-5210***

CORRELATION AND REGRESSION ANALYSIS FOR DESCRIBING THE INHERENT VARIATION IN YIELD OF WHEAT CROP UNDER RICE-WHEAT CROPPING SYSTEM ON LONG RUN BASIS

Vikas Gupta[1]* and H. L. Sharma[2]

[1]Department of Agronomy, College of Agriculture, J.N.K.V.V., Jabalpur - 482 002 (Madhya Pradesh), India.
[2]Department of Mathematics and Statistics, College of Agriculture, J.N.K.V.V., Jabalpur - 482 002 (Madhya Pradesh), India.

Abstract

In the present paper, prediction of wheat yield under rice-wheat cropping system from 1995-96 to 2003-04 through correlation and regression analysis has been developed. The attributing factors for predicting yield were OC(%), Nitrogen (Kg/ha), Phosphorus (Kg/ha), Potash (Kg/ha), pH and EC (ds/m) over nine years data of 1995-96, 1996-97, 1997-98, 1998-99, 1999-2000, 2000-01, 2001-02, 2002-03 and 2003-2004. The results revealed that the OC(%), Nitrogen (Kg/ha), Phosphorus (Kg/ha), Potash (Kg/ha) had played an important role for describing the inherent variation in yield of wheat crop under rice-wheat cropping system from 1995-96 to 2003-04.

Key words : Correlation, regression, yield, wheat.

Introduction

The challenges for agriculture is not only to meet the food needs of the world's expanding population, but also to undertake food supply in a manner that is sustainable for present and future generations. Tremendous quantities of nutrients are required to produce the food necessary to feed the world in any given year. Poor as well as indiscriminate management of these nutrients in many parts of the world has led to environment pollution and the degradation of this resource base, particularly in the developing world. To meet agricultural production and sustainable intensification goals over the short and long-term, plant nutrient and soils need to be managed properly.

Integrated Plant Nutrient Supply System (IPNS) is an important component of sustainable agricultural intensification. The goal of IPNS is to integrate the use of all natural and man-made sources of plant nutrients, so as to increase crop productivity in an efficient and environmentally benign manner without diminishing the capacity of the soil to be productive for present and future generations. It seeks to maintain or improve soil fertility for sustaining the desired level of crop productivity through optimization of the benefit from all possible sources of plant nutrients in an integrated manner.

****Author for correspondence:*** E-mail: v_agro@rediffmail.com

Rice-wheat system is the most dominating double cropping system in India and has become mainstay of cereal production. Both rice and wheat requires high quantity of nutrients to harness their potential yield. However, it is unaffordable to poor and subsistence farmers of the country. Application of inadequate and unbalanced quantity of fertilizers to rice and wheat crops results in low crop yields as well as unsustainable productivity. Therefore, a long-term experiment has been initiated on integrated plant nutrient supply system at Jabalpur (MP) since *kharif* season 1987-88 to maintain the sustainable and high grain yield of wheat without degradation of soil health under irrigated production system. Some of the researchers have done work in this direction (Bahudar *et al.*, 2013; Bhandari *et al.*, 1992; Hashim *et al.*, 2015; Patra *et al.*, 2006; Prasad *et al.*, 2010 and Singh *et al.*, 2008).

In the present paper, prediction of yield in wheat crop under rice-wheat cropping system over nine years of agronomical data through correlation and regression analysis approach has been developed. The attributing factors for predicting yield were OC(%) (X1), Nitrogen (Kg/ha)(X2), Phosphorus(Kg/ha) (X3), Potash(Kg/ha)(X4), PH(X5) and EC(ds/m)(X6) over nine years data of 1995-96, 1996-97, 1997-98, 1998-99, 1999-2000, 2000-01, 2001-02, 2002-03 and 2003-2004.

Materials and Methods

The soil of the experimental field was neutral in reaction (soil pH 7.7) and normal in EC (0.38 dS/m) with medium organic carbon content (6.9 g/kg) and analyzing medium in available N (260 kg/ha), P (16 kg/ha) and high in available K (448 kg/ha) contents. There were 12 treatments (table 1). Different organic manures *viz*. FYM, wheat straw and green leaf manure of sunnhemp were analyzed and their quantities required to substitute a specified amount of N as per the treatments was calculated. Recommended 100% NPK for both crops was 120 kg N + 60 kg P_2O_5 + 40 kg K_2O/ha applied as per the treatment through urea, single super phosphate and muriate of potash, respectively. The experiments were laid out in randomized block design with 4 replications. Rice cv. Kranti was grown by using 40 kg seeds/ha under transplanting with 20 cm × 15 cm planting geometry. Wheat cv. Lok-1 was grown by using seeds 100 kgha^{-1} in rows 20 cm apart. Other cultural practices *viz*. weed management and plant protection measures were followed as per recommendation in the state.

Approach of correlation and regression analysis

Correlation is the study of relationship between two or more variables. Whenever, we conduct any experiment we collect information on more related variables. When there are two related variables their joint distribution is known as bi-variate normal distribution and if there are more than two variables their joint distribution is known as multivariate normal distribution. In case of bi-variate or multivariate normal distribution, we are interested in discovering and measuring the magnitude and direction of relationship between two or more variables. For this we use the technique known as correlation. Suppose we have two continuous variables X and Y and if the change in X affects Y, the variables are said to be correlated. In other words, the systematic relationship between the variables is termed as correlation. When only two variables are involved the correlation is known as simple correlation and when more than two variables are involved the correlation is known as multiple correlation. When the variables move in the same direction, these variables are said to be correlated positively and if they move in the opposite direction they are said to be negatively correlated.

Pearson correlation coefficient

The measure of the degree of linear relationship between two continuous variables is called correlation coefficient. It is denoted by r.(in case of sample). The correlation coefficient r is known as Pearson's correlation coefficient as, it was discovered by Prof. Karl Pearson (1896). It is also called as product moment correlation. The correlation coefficient r is given as the ratio of covariance of the variables X and Y to the product of the standard deviation of X and Y.

Symbolically, which can be simplified as

$$r = \frac{\Sigma xy - \Sigma x \Sigma y/n}{\sqrt{\Sigma x^2 - (\Sigma x)^2} \Big/ \sqrt{\Sigma y^2 - (\Sigma y)^2/n}}$$

This correlation coefficient r is known as Pearson's Correlation coefficient. The numerator is termed as sum of product of X and Y and abbreviated as SP(XY). In the denominator the first term is called sum of squares of X (i.e) SS(X) and second term is called sum of squares of Y (i.e) SS(Y). The denominator in the above formula is always positive. The numerator may be positive or negative making r to be either positive or negative.

Regression

Regression is the functional relationship between two variables and of the two variables one may represent cause and the other may represent effect. The variable representing cause is known as independent variable and is denoted by X. The variable X is also known as predictor variable or repressor. The variable representing effect is known as dependent variable and is denoted by Y. Y is also known as predicted variable. The relationship between the dependent and the independent variable may be expressed as a function and such functional relationship is termed as regression. When there are only two variables the functional relationship is known as simple regression and if the relation between the two variables is a straight line I is known a simple linear regression. When there are more than two variables and one of the variables is dependent upon others, the functional relationship is known as multiple regression. The regression line is of the form y=a+bx where a is a constant or intercept and b is the regression coefficient or the slope. The values of 'a' and 'b' can be calculated by using the method of least squares. An alternate method of calculating the values of a and b are used by the formula:

The regression equation of y on x is given by $y = a_1 + b_1x$

The regression equation of x on y is given by $x = a_2 + b_2y$

Where, $b_1 = \dfrac{\Sigma xy - (\Sigma x)(\Sigma y)/n}{\Sigma x^2 (\Sigma x)^2/n}$ and $b_2 = \dfrac{\Sigma xy - (\Sigma x)(\Sigma y)/n}{\Sigma y^2 (\Sigma y)^2/n}$

Then the equation of regression line of y on x is $y - \bar{y} = b_1(x - \bar{x})$ and the equation of regression line of x

Table 1 : Details of the treatments.

T. no.	*Kharif* (Rice cv. Kranti)	*Rabi* (Wheat cv.Lok-1)
T_1	No fertilizers, no organic manures (Control)	No fertilizers, no organic manures (Control)
T_2	50% recommended NPK through fertilizers	50% recommended NPK through fertilizers
T_3	50% recommended NPK through fertilizers	100% recommended NPK through fertilizers
T_4	75% recommended NPK through fertilizers	75% recommended NPK through fertilizers
T_5	100% recommended NPK through fertilizers	100% recommended NPK through fertilizers
T_6	50% recommended NPK throughfertilizers + 50% N through FYM	100% recommended NPK through fertilizers
T_7	75% recommended NPK through fertilizers + 25% N through FYM	75% recommended NPK through fertilizers
T_8	50% recommended NPK through fertilizer + 50% N through wheat straw	100% recommended NPK through fertilizers
T_9	75% recommended NPK through fertilizers+ 25% N through wheat straw	75% recommended NPK through fertilizers
T_{10}	50% recommended NPK through fertilizers + 50% N through green leaf manuring (Sunhemp)	100% recommended NPK through fertilizers
T_{11}	75% recommended NPK through fertilizers+25% N through green leaf manuring (Sunhemp)	75% recommended NPK through fertilizers
T_{12}	Farmer's practice (40kg N + 20kg P_2O_5 + 3 tonnes FYM/ha)	Farmer's practice (40kg N + 20 kg P_2O_5/ha)

on y is $x-\bar{x}=b_2(y-\bar{y})$. The regression line indicates the average value of the dependent variable Y associated with a particular value of independent variable X.

Results and Discussion

Year 1995-96

The equation of regression for wheat yield are given below, where the factor X2 and X5 and F ratio are found to be non-significant giving R^2 value as 45% with help of X2 and X5 and X3 and X5.

$$\xi = -39.367 + \underset{(0.162)}{0.339\,X_4} - \underset{(0.030)}{0.040X_6}$$

$$\xi = -0.210 + \underset{(38.839)}{77.109\,X_3} - \underset{(0.030)}{0.046X_6}$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	2	95.122	47.561	3.683	0.450
Residual	9	116.234	12.915		

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	2	91.160	45.580	3.413	0.431
Residual	9	120.196	13.355		

Year 1996-97

The equation of regression for wheat yield are given below, where the factor X3 and X4 and F ratio are found to be non-significant.

$$\xi = -33.669 + \underset{(50.408)}{87.271\,X_3}$$

$$\xi = -77.901 + \underset{(0.222)}{0.403X_4}$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	83.016	83.016	2.997	0.231
Residual	10	276.962	27.696		

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	89.271	89.271	3.298	0.248
Residual	10	270.707	27.071		

Year 1997-98

Similarly, the equation of regression for wheat yield are given below, where the factor X3, X4 and X5 and F ratio are found to be non-significant.

$$\xi = -33.694 + \underset{(0.069)}{0.138X4} + \underset{(0.377)}{0.382\,X5}$$

$$\xi = -17.200+27.429X3+0.407\ X5$$
$$(16.373)\quad (0.400)$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	2	17.650	8.825	4.145	0.479
Residual	9	19.162	2.129		

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	2	15.639	7.819	3.324	0.425
Residual	9	21.174	2.353		

Year 1998-99

Similarly, the equation of regression for wheat yield are given below, where the factor X3, X4, X5 and X6 are significant and F ratio are found to be significant.

$$\xi = -63.256 + 118.612\ X3$$
$$(24.352)$$
$$\xi = -180.743+0.763\ X4$$
$$(0.090)$$
$$\xi = -32.542+3.625\ X5$$
$$(0.899)$$
$$\xi = -75.772+0.168\ X6$$
$$(0.022)$$
$$\xi = -296.486+41.665\ X7$$
$$(32.234)$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	279.854	279.854	23.724	0.703
Residual	10	117.962	11.796		

- The figures in parenthesis indicate about the S.E. of the regression coefficients.

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	349.402	349.402	72.169	0.878
Residual	10	48.415	4.841		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	246.280	246.280	16.252	0.619
Residual	10	151.536	15.154		

The figures in parenthesis indicate about the S.E. of the regression coefficient

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	337.464	337.464	55.918	0.848
Residual	10	60.352	6.035		

The figures in parenthesis indicate about the S.E. of the regression coefficient

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	98.370	98.370	3.285	0.247
Residual	10	299.446	29.945		

The figures in parenthesis indicate about the S.E. of the regression coefficient.

Year 1999-2000

Similarly, the equation of regression for wheat yield are given below, where the factor X3, X4, X5 and X6 and F ratio are found to be significant giving a higher value of R^2.

$$\xi = -95.753 + 80.508\ X3 + 0.548\ X5 + 0.090X6$$
$$(21.723)\quad (0.636)\quad (0.034)$$
$$\xi = -141.257 +0.377\ X4 + 0.567\ X5 + 0.095X6$$
$$(0.087)\quad (0.565)\quad (0.031)$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	3	203.343	67.781	19.124	0.878
Residual	8	28.354	3.544		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	3	208.565	69.522	24.044	0.900
Residual	8	23.132	2.891		

- The figures in parenthesis indicate about the S.E. of the regression coefficients.

Year 2000-2001

The equation of regression for wheat yield are given below, where the factor X3 , X4 , X5 and X6 and F ratio are found to be non-significant .

$$\xi = -147.312 + 0.376\ X4 + 0.283\ X5 + 0.106\ X6$$
$$(0.200)\quad (1.204)\quad (0.070)$$
$$\xi = -102.048 +80.660\ X3+ 0.259\ X5 + 0.010X6$$
$$(45.797)\quad (1.342)\quad (0.071)$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	3	196.206	65.402	4.319	0.618
Residual	8	121.156	15.145		

- The figures in parenthesis indicate about the S.E. of the regression coefficients.

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	3	191.340	63.780	4.049	0.603
Residual	8	126.022	15.753		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

Year 2001-2002

The equation of regression for wheat yield are given below, where the factor X3, X4 and X5 and F ratio are found to be non-significant.

$$\xi = -51.782 + 0.113\ X5 + 0.049\ X6 + 25.183\ X8 + 0.230 X4$$
$$(0.772) \quad (0.020) \quad (31.789) \quad (0.175)$$

$$\xi = -21.451 + 0.050\ X6 - 29.243\ X8 + 46.487 X3 + 0.044\ X5$$
$$(0.020) \quad (31.536) \quad (37.149) \quad (0.803)$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	4	265.108	66.277	7.557	0.812
Residual	7	61.390	8.770		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	4	264.019	66.005	7.395	0.809
Residual	7	62.478	8.925		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

Year 2002-2003

Similarly, the equation of regression for wheat yield are given below, where the factor X3, X4 and X5 and F ratio are found to be non-significant.

$$\xi = -62.217 + 127.332\ X3$$
$$(36.962)$$

$$\xi = -134.190 + 0.616\ X4$$
$$(0.184)$$

$$\xi = 7.804 + 1.579\ X5$$
$$(0.557)$$

$$\xi = -6.587 + 0.085\ X6$$
$$(0.031)$$

$$\xi = -336.033 + 47.983\ X7$$
$$(37.285)$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	432.362	432.362	11.867	0.543
Residual	10	364.326	36.433		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	422.297	422.297	11.279	0.530
Residual	10	374.391	37.439		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	246.870	246.870	4.490	0.310
Residual	10	549.817	54.982		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	344.815	344.815	7.631	0.433
Residual	10	451.873	45.187		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	113.20	113.20	1.656	0.142
Residual	10	683.488	68.349		

- The figures in parenthesis indicate about the S.E. of the regression coefficients

Year 2003-04

The equation of regression for wheat yield are given below, where the factor X3, X4 and X5 and F ratio are found to be non-significant.

$$\xi = -55.387 + 118.087\ X3$$
$$(38.284)$$

$$\xi = -121.166 + 0.567\ X4$$
$$(0.158)$$

$$\xi = 10.810 + 1.314\ X5$$
$$(0.584)$$
$$\xi = -3.203 + 0.075\ X6$$
$$(0.024)$$
$$\xi = -84.653 + 206.141\ X8$$
$$(102.552)$$

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	280.170	280.170	9.514	0.488
Residual	10	294.484	29.448		

• The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	324.289	324.289	12.952	0.564
Residual	10	250.365	25.037		

• The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	192.979	192.979	5.056	0.336
Residual	10	381.674	38.167		

• The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	285.797	285.797	9.894	0.497
Residual	10	288.856	28.886		

• The figures in parenthesis indicate about the S.E. of the regression coefficients

ANOVA

Source of variation	df	SS	MS	F	R^2
Regression	1	165.373	165.373	4.041	0.288
Residual	10	409.280	40.928		

• The figures in parenthesis indicate about the S.E. of the regression coefficients

Conclusion

The inherent variability in Wheat yield over nine years agronomical data from 1995-96 to 2003-04 was explained by OC(%), N(Kg/ha), P(kg/ha) ,K(kg/ha), mostly. There was no role of pH and EC(ds/mm) In the present paper, prediction of wheat yield under rice-wheat cropping system from 1995-96 to 2003-04 through correlation and regression analysis has been developed. The attributing factors for predicting yield were OC(%), Nitrogen (Kg/ha), Phosphorus (Kg/ha), Potash (Kg/ha), pH and EC (ds/m) over nine years data of 1995-96, 1996-97, 1997-98, 1998-99, 1999-2000, 2000-01, 2001-02, 2002-03 and 2003-2004. The results revealed that the OC(%), Nitrogen (Kg/ha), Phosphorus (Kg/ha), Potash (Kg/ha) had played an important role for describing the inherent variation in yield of wheat crop under rice-wheat cropping system from 1995-96 to 2003-04 which was evident through the values of R^2 and the SE of the regression coefficients and ANOVA table.

References

Bahudar Lal, D. D. Tiwari, J. Mishra and B. R. Gupta (2013). Evaluation of integrated nutrient management in rice (*Oryza sativa*)-wheat (*Triticum aestivum*) cropping system in reclaimed sodic land. *Indian Journal of Agronomy,* **58(2)** : 137-145.

Bhandari, A. L., Anil Sood, K. N. Sharma and D. S. Rana (1992). Integrated Nutrient Management in a Rice-Wheat System. *Journal of the Indian Society of Soil Science*, **40(4)** : 742-747.

Hashim, Mohammad, S. Dhar, A. K. Vyas, V. Pramesh and Bipin Kumar (2015). Integrated nutrient management in maize (*Zea mays*)-wheat (*Triticum aestivum*) cropping system. *Indian Journal of Agronomy*, **60(3)** : 352-359.

Patra, A. K., B. C. Nayak andM. M. Mishra (2000). Integrated nutrient management in rice (*Oryza sativa*)-wheat (*Triticum aestivum*) cropping system. *Indian Journal of Agronomy*, **45(3)** : 453-457.

Pearson, K. (1896). Mathematical contributions to the theory of evolution on a form of spurious correlation, which may arise when indices are used in the measurement of organs *Proc. R. Soc. Lond* on correltion in 1st January, 1896, London. pp. 489-98.

Prasad, J., S. Karmakar, R. Kumar and B. Mishra (2010). Influence of Integrated Nutrient Management on Yield and Soil Properties in Maize-Wheat Cropping System in an Alfisol of Jharkhand. *Journal of the Indian Society of Soil Science*, **58(2)** : 200-204.

Singh, Fateh, Ravindra Kumar and Samir Pal (2008). Integrated Nutrient Management in Rice-Wheat Cropping System for Sustainable Productivity. *Journal of the Indian Society of Soil Science*, **56(2)** : 205-208.