Journal of Crop and Weed, 14(1): 24-29 (2018)

Correlation and regression analysis for describing the inherent variation in yield of rice crop under rice-wheat cropping system on long run basis

V. GUPTA AND [1]H.L. SHARMA

Department of Agronomy, College of Agriculture,
[1]Department of Mathematics and Statistics, College of Agricultural Engineering,
JNKVV, Jabalpur, Madhya Pradesh

Received : 07-08-2017 ; Revised : 19-03-2018 ; Accepted : 20-03-2018

ABSTRACT

In the present paper, prediction of rice 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 organic carbon, nitrogen, phosphorus, potassium, pH and EC over nine years data from 1995-96 to 2003-2004. The results revealed that the organic carbon, nitrogen, phosphorus, potassium, had played important roles for describing the inherent variation in yield of rice crop under rice-wheat cropping system.

***Keywords*:** Correlation, regression, rice seed yield

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. Huge amount 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.

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 rice without degradation of soil health under irrigated production system. Some of the researchers carried out in this direction; are Bajpai *et al.* (2006), Chettri *et al.* (2017), Jat *et al.* (2015), Manjhi (2016), Singh *et al.* (2006), Singh (2006) and Swarup and Yaduvanshi (2000).

In the present paper, prediction of yield in rice 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 organic carbon (X_1), nitrogen (X_2), phosphorus, (X_3), potassium (X_4), pH (X_5) and EC (X_6) over nine years from 1995-96 to 2003-2004.

MATERIALS AND METHODS

The experimental field was situated at irrigated condition and soil was neutral in reaction (soil pH 7.7) and normal in EC (0.38 dS m^{-1}) with medium organic carbon content (6.9 g kg^{-1}) and analyzing medium in available N (260 kg ha^{-1}), P (16 kg ha^{-1}) and high in available K (448 kg ha^{-1}) 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^{-1} 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^{-1} under transplanting with 20 x 15 cm planting geometry. Wheat cv. *Lok-1* was grown by using seeds 100 kg ha^{-1} in rows 20 cm apart. Other cultural practices viz. weed management and plant protection measures were followed as per recommendation in the state.

Email: v_agro@rediffmail.com

Correlation is the study of finding linear relationship between two or more variables and the technique by which it can be measured known as correlation coefficient. 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.

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 regressor. The variable representing effect is known as dependent variable and is denoted by Y. 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.

Table 1: Details of the treatments

Tr. 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)

RESULTS AND DISCUSSION

YEAR 1995-96

In the year of 1995-96, the contribution of individual independent components i.e. coefficient of multiple determination (R^2) were found which are given below:

OC(%)=0.546 N(kg ha^{-1})=0.602, P(kg ha^{-1})=0.001,K(kg ha^{-1})=0.193, pH =0.004 and EC (dS m^{-1}) = 0.000

It indicated that OC(%), N and K contributed the inherent variability of rice yield almost 55, 60 and 19 per cent respectively. The variable OC (%) and N were highly correlated. After deleting the high value of correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 1995-96 are given below. The regression coefficients of the factor X_2 and X_5 are found to be significant through the value of t statistic. Since t statistic is the ratio of concern regression coefficient to the standard error. The value of F statistic was also found to be significant indicating a good forecast for the rice yield giving R^2 value as 70 per cent with help of X_2 and X_5 and X_3 and X_5. Where X_2, X_3, X_5 are nitrogen (kg ha^{-1}) (X_2), phosphorus (kg ha^{-1}) (X_3), and pH(X_5).

$$\bar{Y} = -46.738 + \underset{(63.164)^*}{243.512}\, X_2 - \underset{(0.048)}{0.102}\, X_5$$

$$\bar{Y} = -161.934 + \underset{(0.264)^*}{1.042}\, X_3 - \underset{(0.048)}{0.085}\, X_5$$

Table 2: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	2	727.062	363.531	10.292	0.696
Residual	9	317.906	35.323		

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

Table 3:ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	2	735.530	367.765	10.696	0.704
Residual	9	309.439	34.382		

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

YEAR 1996-97

In the year 1996-97, the contribution of individual independent components *i.e.* coefficient of multiple determination (R^2) were found which are given below:

OC(%)=0.185, N(kg ha^{-1})=0.235, P(kg ha^{-1})=0.006,K(kg ha^{-1})=0.042, pH =0.054 and EC (dS m^{-1})=0.063.It indicated that OC(%), and N contributed the inherent variability of rice yield 42 per cent. After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 1996-97 are given below. The regression coefficients of the factor X_3 and X_4 are found to be non-significant. The value of F statistic was also not significant.

$\bar{Y}$ = -35.726+109.628X3
(72.705)*

$\bar{Y}$ = -102.542+ 0.549 X_4
(0.313)*

Table 4:ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	131.001	131.001	2.274	0.185
Residual	10	576.180	57.618		

Table 5:ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	166.057	166.057	3.069	0.235
Residual	10	541.124	54.112		

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

YEAR 1997-98

In the year 1997-98, the contribution of individual independent components i.e. coefficient of multiple determination (R^2) were found which are given below:

OC(%)=0.123, N(kg ha^{-1})=0.153, P(kg ha^{-1})=0.117, K(kg ha^{-1})=0.080, pH =0.024 and EC(dS m-^{1})=0.039.

It indicated that OC(%), N and P contributed the inherent variability of rice yield almost 20%. .After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 1997-98 are given below. The regression coefficients of the factor X_3, X_4 and X_5 are found to be non-significant. The value of F statistic was also not significant.

$\bar{Y}$ = -32.487+70.734X3+1.628 X5
(36.925) (2.638)

$\bar{Y}$ = -80.860+0.382X4+ 1.502 X_5
(0.421) (2.311)

Table 6: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	2	148.317	74.158	0.898	0.166
Residual	9	742.828	82.536		

Table 7: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	2	170.400	85.200	1.064	0.191
Residual	9	720.744	80.083		

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

YEAR 1998-99

In the year of 1998-99, the contribution of individual independent components i.e. coefficient of multiple determination (R^2) were found which are given below:

OC (%) = 0.625, N (kg ha^{-1}) = 0.710, P (kg ha^{-1}) = 0.730, K(kg ha^{-1}) = 0.762, pH = 0.123 and EC (dS m-^{1}) = 0.001.

It indicated that OC(%), N, P and K contributed the inherent variability of rice yield almost 70%. .After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 1998-99 are given below. The regression coefficients of the factor X_3, X_4, X_5 and X_6 are found to be significant. The value of F statistic was also significant.

$\bar{Y}$ = -60.577+128.237X3
(31.414)

$\bar{Y}$ = -177.622+0.787X4
(0.159)*

$\bar{Y}$ = -36.162+4.515X5
(0.868)

$\bar{Y}$ = -74.696+0.183 X6
(0.032)*

$\bar{Y}$ = -226.116+33.764X7
(28.465)*

Table 8: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	327.112	327.112	16.664	0.625
Residual	10	196.795	19.630		

Table 9: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	371.875	371.875	24.541	0.710
Residual	10	151.532	15.153		

Table 10: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	382.210	382.210	27.069	0.730
Residual	10	141.196	14.120		

Table 11: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	398.894	398.874	32.036	0.762
Residual	10	124.513	12.451		

Table 12: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	64.255	64.235	1.393	0.123
Residual	10	459.152	45.915		

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

YEAR 1999-2000

In the year of 1999-2000, the contribution of individual independent components *i.e.* coefficient of multiple determination (R^2) were found which are given below:

OC(%)=0.454, N(kg ha^{-1})=0.424, P(kg ha^{-1}) = 0.459, K(kg ha^{-1}) = 0.423, pH =0.009 and EC(dS m-^{1}) = 0.007.

It indicated that OC(%), N , P and K contributed the inherent variability of rice yield almost 70 per cent. .After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 1999-2000 are given below. The regression coefficients of the factor X_3, X_4, X_5 and X_6 are found to be significant. The value of F statistic was also significant.

$\bar{Y}$ = -206.151+166.016X3+0.986X5+0.207 X6
(83.037) (2.433) (0.129)

$\bar{Y}$ = -294.513+0.755X4+1.096X5+0.216X6
(0.365) (2.165) (0.128)

Table 13: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	3	909.240	303.080	5.852	0.687
Residual	8	414.306	51.788		

Table 14: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	3	918.645	306.215	6.050	0.653
Residual	8	404.901	50.613		

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

YEAR 2000-2001

In the year of 2000-2001, the contribution of individual independent components *i.e.* coefficient of multiple determination (R^2) were found which are given below :

OC(%)=0.480, N(kg ha^{-1})=0.471, P(kg ha^{-1}) = 0.368, K(kg ha^{-1}) = 0.256, pH =0.002 and EC(dS m-^{1}) = 0.004.

It indicated that OC(%), N , P and K contributed the inherent variability of rice yield almost 40 per cent. After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 2000-2001 are given below. The regression coefficients of the factor X_3, X_4, X_5 are found to be significant. The value of F statistic was non-significant.

$\bar{Y}$ = -137.020+148.196X3+0.793X5+0.100 X6
(74.923) (2.195) (0.117)

$\bar{Y}$ = -218.996+0.686X4+0.852X5+0.109 X6
(0.326) (2.113) (0.115)

Table 15: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	3	481.869	160.623	3.810	0.588
Residual	8	337.293	42.162		

Table 16: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	3	495.798	165.266	4.089	0.605
Residual	8	323.364	40.420		

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

YEAR 2001-2002

In the year of 2001-2002, the contribution of individual independent components *i.e.* coefficient of multiple determination (R^2) were found which are given below:

OC(%)=0.385, N(kg ha^{-1})=0.367, P(kg ha^{-1}) = 0.548,K(kg ha^{-1}) = 0.423, pH = 0.003 and EC(dS m-^{1})=0.237.

It indicated that OC(%), N , P and K contributed the inherent variability of rice yield almost 70 per cent .After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 2001-2002 are given below. The regression coefficients of the factor X_3, X_4, X_5 and X_6 are found to be significant. The value of F statistic was also significant.

$\bar{Y}$ =-40.280+48.738X3+4.508X5+0.046 X6–56.916X8
(104.249) (2.254) (0.056) (88.499)

$\bar{Y}$ = -62.398+4.638X5+0.047X6+0.201 X4 – 54.264X8
(2.194) (0.057) (0.498) (90.341)

Table 17 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	4	1140.475	285.119	4.056	0.612
Residual	7	492.017	70.288		

Table 18 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	4	1136.697	284.174	4.012	0.699
Residual	7	415.795	70.828		

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

YEAR 2002-2003

In the year of 2002-2003, the contribution of individual independent components i.e. coefficient of multiple determination (R^2) were found which are given below:

OC(%)=0.677, N(kg ha^{-1})=0.661, P(kg ha^{-1})=0.465, K(kg ha^{-1})=0.596, pH=0.087 and EC(dS m-^{1})=0.104.

It indicated that OC(%), N , P and K contributed the inherent variability of rice yield almost 70%. .After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 2002-03 are given below. The regression coefficients of the factor X3, X4, X5 and X6 are found to be significant. The value of F statistic was also significant.

$\bar{Y}$ = -61.626+157.660X3
(34.45)

$\bar{Y}$ = -150.66+0.763X4
(0.173)

$\bar{Y}$ = 22.523+2.144X5
(0.727)

$\bar{Y}$ = 5.008+0.110X6
(0.029)

Table 19 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	662.845	662.845	20.95	0.677
Residual	10	316.392	31.639		

Table 20 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	646.898	646.898	19.465	0.661
Residual	10	332.338	33.234		

Table 21 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	455.454	455.454	8.696	0.465
Residual	10	523.783	52.378		

Table 22 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	538.817	583.817	14.764	0.596
Residual	10	395.419	39.542		

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

YEAR 2003-04

In the year of 2003-04, the contribution of individual independent components i.e. coefficient of multiple determination (R^2) were found which are given below:

OC(%)=0.668, N(kg ha^{-1})=0.756, P(kg ha^{-1}) = 0.570,K(kg ha^{-1})=0.700, pH=0.020 and EC(dS m-^{1}) =0.379.

It indicated that OC(%), N , P and K contributed the inherent variability of rice yield almost 70%. .After deleting the higher ordered correlation coefficients to avoid the multicollinearity, a regression equation along with ANOVA table for the year 2003-04 are given below. The regression coefficients of the factor X3, X4, X5 and X6 are found to be significant. The value of F statistic was also significant.

$\bar{Y}$ = -47.091+138.243X3
(30.795)

$\bar{Y}$ = -122.068+0.656X4
(0.118)

$\bar{Y}$ = 28.058+1.712X5
(0.470)

$\bar{Y}$ = 13.503+0.090X6
(0.019)

$\bar{Y}$ = -78.784+236.651X8
(222.123)

Table 23 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	383.975	383.975	20.152	0.688
Residual	10	19.054	19.054		

Table 24 : ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	434.205	434.205	30.947	0.756
Residual	10	140.306	14.031		

Table 25: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	327.621	327.621	13.270	0.570
Residual	10	246.890	24.689		

Table 26: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	402.222	402.222	23.314	0.700
Residual	10	172.289	17.229		

Table 27: ANOVA

Source of Variation	df	SS	MS	F	R^2
Regression	1	217.948	217.948	6.112	0.379
Residual	10	356.563	35.656		

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

The inherent variability in rice yield over nine years agronomical data from 1995-96 to 2003-04 was explained by organic carbon(%), N (kg ha^{-1}), P(kg ha^{-1}), K(kg ha^{-1}), mostly. There was no role of pH and EC. The results revealed that the organic carbon, nitrogen, phosphorus, potash had played an important role for describing the inherent variation in yield of rice 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

Bajpai, R. K., Chitale, S., Upadhyay, S. K. and Urkurkar, J. S. 2006. Long-term studies on soil physico-chemical properties and productivity of rice-wheat system as influenced by integrated nutrient management in inceptisol of Chhattisgarh. *J. Indian Soc. Soil Sci.*, **54**: 24-29.

Chettri, P., Maiti, D. and Rizal, B. 2017. Studies on soil properties as affected by integrated nutrient management practice in different cultivars of local scented rice (*Oryza sativa* L.). *J. Crop Weed*, **13**: 25-29.

Jat, A. L., Srivastava, V. K. and Singh, R. K. 2015. Effect of crop establishment methods and integrated nitrogen management on productivity of hybrid rice (Oryza sativa)-wheat (Triticum aestivum) cropping system. *Indian J. Agron*., **60:** 341-46.

Manjhi, R.P., Mahapatra, P., Shabnam, S. and Yadav, M. S. 2016. Long term effect of integrated nutrient management practices on performance of quality protein maize under maize (*Zea mays*) – wheat (*Triticum aestivum*) cropping sequence. *Indian J. Agron*., **61**: 436-42.

Pearson, K. 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.

Singh, V. 2006. Productivity and economics of rice (*Oryza sativa*)-wheat (*Triticum aestivum*) cropping system under integrated nutrient supply system in reclaimed sodic land. *Indian J. Agron.*, **54**: 81-84.

Singh, G., Singh, O. P., Singh, R. G., Mehta, R. K., Kumar, V. and Singh, R. P. 2006. Effect of integrated nutrient management on yield and nutrient uptake of rice (*Oryza sativa*)–wheat (*Triticum aestivum*) cropping system in lowlands of eastern Uttar Pradesh. *Indian J. Agron*. **51**:85-88.

Swarup, A. and Yaduvanshi, N. P. S. 2000. Effect of integrated nutrient management on soil properties and yield of rice in Alkali soils. *J. Indian Soc. Soil Sci*. **58**: 200-04.