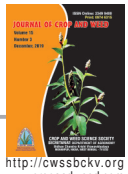
Journal of Crop and Weed, 18(1): 187-195 (2022)
ISSN- O : 2349 9400 ; P : 0974 6315



http://cwssbckv.org
www.cropandweed.com

Detection and impact of outliers on bell pepper (*Capsicum annuum* L. var. *grossum* Sendt.) in field condition

M. ROY, A. SAHA, J. S. BASAK, M. SAHA, MD M. H. MIDDYA, S. ROY AND *A. MAJUMDER

Department of Agricultural Statistics, [1]Department of Vegetable Science, Bidhan Chandra Krishi Viswavidyalaya, P.O. Krishi Viswavidyalaya, West Bengal, India, Pin: 741252

Received : 14.12.2021 ; Revised : 28.02.2022 ; Accepted : 09.03.2022

DOI: https://doi.org/10.22271/09746315.2022.v18.i1.1551

ABSTRACT

The presence of abnormal observations (even a single abnormal observation) in any agricultural field experiment may cause a serious damage to the whole experiment. Actually, the presence of such observations (or a single observation) will make a deviation from the normality assumption of ANOVA model. Such observations are considered as outliers in Statistics. Cook (1977); Andrews and Pregibon (1978) developed methodologies to detect the presence of outliers in regression model. Later Bhar and Gupta (2001) modified those statistics for detection of outliers in experimental design data set. Those modified statistics have been applied in field experiments on bell pepper (Capsicum annuum L. var. grossum Sendt.) for detection of presence of outliers in the data set. In all, more than 10 experiments in RBD layout for different parameters of the crop were judged for presence of outliers and only 6 of them were found with presence of significant influential outliers. After detection of outliers, the outliers were removed and the data set of each experiment was undergone through the missing plot analysis. It has been observed that in all of the cases after removal of detected outliers, the analysis gave more efficient and effective result. The results of outliers, the mean values of treatments (bell pepper) and error mean square (before and after removal of significant outlier) for each experiment have been tabulated.

***Keywords*:** ANOVA, bell pepper, error mean square (EMS), missing plot technique, outlier.

Bell pepper (*Capsicum annuum* L. var. *grossum* Sendt.) belonging to the family Solanaceae is a valuable vegetable crop which is grown for its pleasant flavour, delicate taste and colour across the world. The fruits are generally large, blocky, three to four lobed, thick fleshed and non-pungent. Bell pepper is also known to have high nutritional values due to the presence of different pigments particularly, carotenoids (beta-carotene, capsanthin, leutin, zeaxanthin, *etc*.) and xanthophylls, different vitamins, dietary fibre and several essential minerals (Mitra *et. al.*, 1990; Hazra *et. al*., 2011). Thus, bell pepper needs very careful and efficient experiments for qualitative improvement. The interpretation of analysis of variance (ANOVA) will be valid only when the basic assumptions are fulfilled. But in practice, the experimenter often encounters the problem of departure or deviation from the most important assumption, as the observations not exactly follow the normal distribution. Even a single data point in the set of observations may distort the entire set of observation. These data points which are responsible for the distortion are commonly known as outliers. Outliers may cause a serious problem to the agricultural experiments.

Several statisticians were engaged for detection of outlier or outliers in regression analysis. Among them, most remarkable work was done by Cook (1977). Andrews and Pregibon (1978) defined another test statistic which is also very useful in detecting outlier(s) in experimental data. Some more references in determining the degree of influence of outlier(s) on parameter estimation in block design by using this statistic are noted (Bhar and Gupta 2003; Bhar *et. al.,* 2013; Ojha and Bhar 2015). Pena and Yohai (1995) proposed a method for identifying influential subsets, from the eigen values of an 'influence matrix' (X). This 'influence matrix' is a off-centered covariance of a set of vectors, represent the effect on the fit of the deletion of each data point. This 'influence matrix' is normalized to have the univariate Cook (1979) statistics on the diagonal. Hocking (1984) proposed that the eigen structure of the matrix $(X : y)'$ $(X : y)$ should be computed, where X = design matrix and y = vector of responses. Detection of outliers in linear time series data was done by Chang (1982) and Saha (2016). But the presence of outliers in experimental design had not received much attention to overcome the problems. However, after a long gap, Bhar and Gupta (2001) modified the Cook statistic and AP statistic (Ojha and Bhar, 2015) for detection of outliers in experimental designs. Presently it gathers a huge momentum in subject statistic for detection of outliers in several types of experimental designs (Ojha and Bhar, 2015). Keeping in view the importance of outliers, in experiments on bell pepper for efficient analysis, some field experiments

Email: anurupbckv@gmail.com

were chosen for checking whether the outliers influenced the experiments or not. For the purpose, 10 experiments in RBD layout conducted in 2018-19 on different yield parameters of bell pepper were examined. Out of them, only 6 experiments were detected where the presence of a single outlier influenced the estimation of treatment effects. In all experiments, the above mentioned modified statistics were applied to detect the outlier. After detection, the significant outlier was removed and the data set of each experiment was undergone through missing plot analysis.

MATERIALS AND METHODS

Experimental details

The methodologies of detection and analysis of field experiments were done on bell pepper. Experimental details are given in Table 1.

Software used

Statistical data analysis was performed in R studio (R version 3.4.3) (2013). RBD and Tukey test using general linear model were computed in Agricolae R package (2009). Then outliers were detected in Microsoft Excel. After that the missing plot technique and Tukey test (comparison of the different varieties) were enumerated in Agricolae R package (2009).

Model and statistics used

Let us consider the general linear model for an experimental design,

$$y = X\theta + e; \quad (1)$$

$E(e) = 0$;

$D(e) = \sigma^2 I_n$; $\sigma^2 > 0$, with mean 0 and covariance matrix $\sigma^2 I_n$, where

y = $n \times 1$ vector of observations,

X = $n \times p$ matrix of known constants with full column rank p,

θ = $p \times 1$ vector of unknown parameters,

e = $n \times 1$ vector of independent random variables.

Detection of single outlier through cook statistic

Let the first plot of a block design be considered as an outlier. Consider the design (d) is a block design. The intra-block model for such design is

$$y = \mu 1_n + \Delta'\tau + D'\beta + e \quad (2)$$

Here Δ' = $n \times v$ (0 –1) design matrix for treatment effects.

D' = $n \times b$ (0 –1) design matrix for block effects.

μ = general mean,

τ = v component vector of treatment effects.

β = b-component vector of block effects.

$X = (X_1\ X_2)$,

$X_1 = \Delta_1$, $X_2 = [1_n\ D']$, $\theta_1 = \tau$, $\theta_2 = [\mu\ \beta']'$.

let us define, matrix Φ as, $\Phi = I_n - D'\ k^{-1} D$ (3)

Φ is symmetric and idempotent matrix and

$$C\tau = \Delta\Phi\Delta' \quad (4)$$

C_τ is C-matrix in the block design setup. In S matrix, is the diagonal element.

$$S = \Phi\Delta' C_\tau^{+}\Delta\Phi \quad (5)$$

If r^* = ordinary residual and t_1 = Studentized residuals respectively for the outlying observation, then $r_1^* = y_1 - \hat{y}_1$ and $t_1 = \frac{r_1^*}{\hat{\sigma}\sqrt{v_{11}}}$, where v_{11} = first diagonal element of the matrix V,

$$V = \Phi - \Phi\Delta' C_\tau^{+}\Delta\Phi = \Phi - S \quad (6)$$

Thus D_1 (the outlier statistic for the first plot of the experiment) can be written as-

$$D_1 = \frac{s_{11}}{v_{11}}\ \frac{t_1^2}{v-1} \quad (7)$$

Detection of single outlier through AP statistic

We assume that the first observation in the first block is an outlier, then AP_1 – statistic for a single outlier as,

$$AP_1 = v_{11}\left(1 - \frac{v_{11} r_t^{*2}}{y'Vy}\right) \quad (8)$$

where r_t^* and V defined earlier.

RESULTS AND DISCUSSION

All the experiments were conducted for 12 treatments with 3 replications in RBD layout. The tabulated value of F for 11 and 22 degrees of freedom at 95% *i.e.* (1-α) x100% level of significance is 0.380.

The Cook statistics and AP statistics values were calculated for the experiment on number of fruit/plant of bell pepper. Those values are tabulated in Table 2.

In Table 2, it is observed that the 31st observation, placed in first replication/ block of 'BABY BELL X C/4 (F1)' shows the maximum value of Cook Statistic as well as minimum value of AP statistic (in parenthesis). Thus the value of Cook Statistic is significant to be influential.

Table 3, reveals that the EMS of analysis after removal of outlier is less than the EMS of the analysis of actual observation. Thus, it is clear that outlier removal increases the efficiency of the experiment and the rank of the 'BABY BELL X C/4 (F1)' has been changed from 3rd to 5th.

The Cook statistics and AP statistics values were calculated for the experiment on fruit length (cm) of bell pepper. Those values are tabulated in Table 4.

In Table 4, it is observed that the 4th observation, placed in first replication/ block of '8/4' shows the maximum value of Cook Statistic as well as minimum value of AP statistic (in parenthesis). Thus the value of Cook Statistic is significant to be influential.

Table 5, reveals that the EMS of analysis after removal of outlier is less than the EMS of the analysis of actual observation. Thus, it is clear that outlier removal increases the efficiency of the experiment, but the rank of the '8/4' has not been changed.

Table 1: Experimental details.

Location	Central Research Farm, Gayeshpur, BCKV, West Bengal
Name of the crop	Bell pepper (*Capsicum annuum* L. var. *grossum* Sendt.)
Number of variety	12 variety/ hybrid
Replication number	3
Year	2018-19
Parameters recorded	No. of fruit/ plant, fruit length, fruit width, average no. of seeds/ fruit, 100 seed weight, fruit yield/plant.

Table 2: Computations of Cook statistic and AP Statistic of number of fruits / plant of bell pepper

VARIETIES	REPLICATION 1	REPLICATION 2	REPLICATION 3
BABY BELL	0.009987	0.010261	0.040494
	(0.607672)	(0.607455)	(0.596924)
8/4	0.008547	0.005109	0.00044
	(0.608173)	(0.60934)	(0.610964)
C/4	0.040288	0.001552	0.026025
	(0.596756)	(0.610549)	(0.601877)
C/4(YELLOW)	0.046227	0.056707	0.205332
	(0.594652)	(0.591227)	(0.538585)
AYESHA	0.016368	0.003271	0.034274
	(0.605242)	(0.609981)	(0.598962)
ARYA	0.00722	0.003429	0.000698
	(0.608635)	(0.609926)	(0.610876)
ROYAL WONDER	0.002134	0.002967	0.0000685
	(0.610396)	(0.610087)	(0.611084)
RW X ARYA (F1)	0.000243	0.003576	0.005684
	(0.611038)	(0.609827)	(0.609139)
RW X AYSHA (F1)	0.01261	0.085917	0.032697
	(0.606578)	(0.580723)	(0.599663)
BABY BELL X AYSHA (F1)	0.03578	0.025353	0.121371
	(0.598353)	(0.602242)	(0.568209)
BABY BELL X C/4 (F1)	0.389002*	0.038935	0.181801
	(0.468519)	(0.596513)	(0.54517)
8/4 X RW (F1)	0.013861	0.117492	0.050642
	(0.606133)	(0.569578)	(0.593358)

Note. Values in parentheses are AP statistic values.
*Significant at 95% i.e. (1-α) x100% probability level.

The Cook statistics and AP statistics values were calculated for the experiment on fruit width (cm) of bell pepper and those values are tabulated in Table 6.

In Table 6, it is observed that the 10th observation, placed in first replication/ block of 'C/4(YELLOW)' shows the maximum value of Cook Statistic as well as minimum value of AP statistic (in parenthesis). Thus the value of Cook Statistic is significant to be influential.

Table 7, reveals that the EMS of analysis after removal of outlier is less than the EMS of the analysis of actual observation. Thus, it is clear that outlier removal increases the efficiency of the experiment and the rank of the 'C/4 (YELLOW)' has been changed from 5th to 8th.

The Cook statistics and AP statistics values were calculated for the experiment on average number of seeds/fruits of bell pepper. Those values are tabulated in Table 8.

In Table 8, it is observed that the 3rd observation, placed in third replication/ block of 'BABY BELL' shows the maximum value of Cook Statistic as well as minimum value of AP statistic (in parenthesis). Thus the value of Cook Statistic is significant to be influential.

Table 3: Mean values with corresponding ranks and EMS of ANOVA table before and after removal of significant outlier form the experiment on number of fruits / plant of bell pepper

Varieties	Mean (Actual)	Mean (Considering outlier)	EMS (Actual)	EMS (Considering outlier)
BABY BELL	14.4167(2nd)	14.4167 (2nd)	1.78617	1.160967
8/4	14.4733 (1st)	14.4733 (1st)		
C/4	7.4800 (9th)	7.4800 (9th)		
C/4 (YELLOW)	11.08 (4th)	11.08 (3rd)		
AYESHA	4.4433 (12th)	4.4433 (12th)		
ARYA	8.41 (7th)	8.41 (7th)		
ROYAL WONDER	7.25 (10th)	7.25 (10th)		
RW X ARYA (F1)	7.13 (11th)	7.13 (11th)		
RW X AYSHA (F1)	7.54667 (8th)	7.54667 (8th)		
BABY BELL X AYSHA (F1)	8.9433 (6th)	8.9433 (6th)		
BABY BELL X C/4 (F1)	**11.4900 (3rd)**	**9.82333 (5th)**		
8/4 X RW (F1)	9.8833 (5th)	9.8833 (4th)		

Table 4: Computations of Cook statistic and AP Statistic of fruit length (cm) of bell pepper

VARIETIES	REPLICATION 1	REPLICATION 2	REPLICATION 3
BABY BELL	0.015958	0.008853	0.001039
	(0.607474)	(0.609093)	(0.610874)
8/4	**0.39676***	**0.002918**	**0.331631**
	(0.520672)	(0.610446)	(0.535518)
C/4	0.031278	0.000877	0.042633
	(0.603981)	(0.610911)	(0.601393)
C/4 (YELLOW)	0.000128	0.014877	0.017768
	(0.611082)	(0.60772)	(0.607061)
AYESHA	0.029457	0.00351	0.012631
	(0.604397)	(0.610311)	(0.608232)
ARYA	0.034436	0.264214	0.107878
	(0.603262)	(0.550885)	(0.586521)
ROYAL WONDER	0.008368	0.009184	0.035086
	(0.609204)	(0.609018)	(0.603114)
RW X ARYA (F1)	0.0525	0.038084	0.001154
	(0.599144)	(0.60243)	(0.610848)
RW X AYSHA (F1)	0.018235	0.002918	0.006565
	(0.606955)	(0.610446)	(0.609615)
BABY BELL X AYSHA (F1)	0.032523	0.0000121	0.033792
	(0.603698)	(0.611108)	(0.603408)
BABY BELL X C/4 (F1)	0.005229	0.01093	0.031278
	(0.609919)	(0.60862)	(0.603981)
8"4 X RW (F1)	0.022194	0.003719	0.007743
	(0.606052)	(0.610263)	(0.609346)

Note. Values in parentheses are AP statistic values.

* Significant at 95% i.e. (1-α) x 100% probability level.

Table 5: Mean values with corresponding ranks and EMS of ANOVA table before and after removal of significant outlier form the experiment on fruit length (cm) of bell pepper

Varieties	Mean (Actual)	Mean (Considering outlier)	EMS (Actual)	EMS (Considering outlier)
BABY BELL	4.773 (12^{th})	4.773 (12^{th})	2.570573	0.227290
8/4	**13.65 (1^{st})**	**14.307 (1^{st})**		
C/4	9.087 (2^{nd})	9.087 (2^{nd})		
C/4 (YELLOW)	6.507 (7^{th})	6.507 (7^{th})		
AYESHA	8.997 (3^{rd})	8.997 (3^{rd})		
ARYA	7.870 (4^{th})	7.870 (4^{th})		
ROYAL WONDER	7.570 (5^{th})	7.570 (5^{th})		
RW X ARYA (F1)	6.007 (10^{th})	6.007 (10^{th})		
RW X AYSHA (F1)	5.257 (11^{th})	5.257 (11^{th})		
BABY BELL X AYSHA (F1)	6.190 (8^{th})	6.190 (8^{th})		
BABY BELL X C/4 (F1)	6.053 (9^{th})	6.053 (9^{th})		
8/4 X RW (F1)	7.510 (6^{th})	7.510 (6^{th})		

Table 6: Computations of Cook statistic and AP Statistic of fruit width (cm) of bell pepper

VARIETIES	REPLICATION 1	REPLICATION 2	REPLICATION 3
BABY BELL	0.002193	0.003052	0.0000709
	(0.610611)	(0.610415)	(0.611095)
8/4	0.007175	0.009888	0.000217
	(0.609475)	(0.608857)	(0.611062)
C/4	0.000202	0.018569	0.014899
	(0.611065)	(0.606877)	(0.607714)
C/4 (YELLOW)	**0.515763***	0.029962	0.297101
	(0.493507)	(0.604279)	(0.543366)
AYESHA	0.013766	0.002713	0.004256
	(0.607972)	(0.610493)	(0.610141)
ARYA	0.010311	0.000379	0.006736
	(0.60876)	(0.611025)	(0.609575)
ROYAL WONDER	0.045404	0.004325	0.021702
	(0.600758)	(0.610125)	(0.606163)
RW X ARYA (F1)	0.008869	0.021239	0.057559
	(0.609089)	(0.606268)	(0.597987)
RW X AYSHA (F1)	0.244071	0.044957	0.079527
	(0.555458)	(0.60086)	(0.592978)
BABY BELL X AYSHA (F1)	0.061408	0.008477	0.024253
	(0.597109)	(0.609178)	(0.605581)
BABY BELL X C/4 (F1)	0.000146	0.00475	0.003229
	(0.611078)	(0.610028)	(0.610375)
8/4 X RW (F1)	0.034103	0.00000692	0.035082
	(0.603335)	(0.61111)	(0.603112)

Note. Values in parentheses are AP statistic values.

* Significant at 95% i.e. (1-α) x 100% probability level.

Table 7: Mean values with corresponding ranks and EMS of ANOVA table before and after removal of significant outlier form the experiment on fruit width (cm) of bell pepper

Varieties	Mean (Actual)	Mean (Considering outlier)	EMS (Actual)	EMS (Considering outlier)
BABY BELL	4.373 (10^{th})	4.373 (10^{th})	0.186599	0.097602
8/4	4.893 (9^{th})	4.893 (9^{th})		
C/4	6.310 (2^{nd})	6.310 (2^{nd})		
C/4 (YELLOW)	**5.61 (5^{th})**	**4.990 (8^{th})**		
AYESHA	7.673 (1^{st})	7.673 (1^{st})		
ARYA	5.887 (3^{rd})	5.887 (3^{rd})		
ROYAL WONDER	5.670 (4^{th})	5.670 (4^{th})		
RW X ARYA (F1)	3.567 (12^{th})	3.567 (12^{th})		
RW X AYSHA (F1)	3.600 (11^{th})	3.600 (11^{th})		
BABY BELL X AYSHA (F1)	5.140 (7^{th})	5.140 (6^{th})		
BABY BELL X C/4 (F1)	5.097 (8^{th})	5.097 (7^{th})		
8/4 X RW (F1)	5.530 (6^{th})	5.530 (5^{th})		

Table 8: Computations of Cook statistic and AP Statistic of average number of seeds/ fruits of bell pepper

VARIETIES	REPLICATION 1	REPLICATION 2	REPLICATION 3
BABY BELL	0.110389	0.086199	0.391682*
	(0.585931)	(0.591449)	(0.521768)
8/4	0.002025	0.00494	0.000639
	(0.610649)	(0.609984)	(0.610965)
C/4	0.018844	0.036299	0.002835
	(0.606813)	(0.602831)	(0.610464)
C/4 (YELLOW)	0.015487	0.029203	0.002157
	(0.607579)	(0.60445)	(0.610619)
AYESHA	0.001918	0.021521	0.036288
	(0.610674)	(0.606202)	(0.602834)
ARYA	0.015265	0.056886	0.131088
	(0.607629)	(0.598135)	(0.58121)
ROYAL WONDER	0.0000000312	0.016	0.015955
	(0.611111)	(0.607462)	(0.607472)
RW X ARYA (F1)	0.087877	0.035083	0.011911
	(0.591066)	(0.603109)	(0.608394)
RW X AYSHA (F1)	0.11588	0.018398	0.041932
	(0.584679)	(0.606915)	(0.601547)
BABY BELL X AYSHA (F1)	0.004399	0.058593	0.030882
	(0.610108)	(0.597746)	(0.604067)
BABY BELL X C/4 (F1)	0.001906	0.098754	0.073218
	(0.610676)	(0.588585)	(0.59441)
8/4 X RW (F1)	0.014072	0.040861	0.006975
	(0.607901)	(0.601791)	(0.60952)

Note. Values in parentheses are AP statistic values.

* Significant at 95% i.e. (1-α) x100% probability level.

Table 9 : Mean values with corresponding ranks and EMS of ANOVA table before and after removal of significant outlier form the experiment on average number of seeds/ fruits of bell pepper

Varieties	Mean (Actual)	Mean (Considering outlier)	EMS (Actual)	EMS (Considering outlier)
BABY BELL	**60.55333 (9th)**	**50.353 (9th)**	66.41375	62.19868
8/4	42.940 (11th)	42.940 (11th)		
C/4	77.073 (6th)	77.073 (6th)		
C/4 (YELLOW)	69.797 (7th)	69.797 (7th)		
AYESHA	64.947 (8th)	64.947 (8th)		
ARYA	112.303 (1st)	112.303 (1st)		
ROYAL WONDER	25.600 (12th)	25.600 (12th)		
RW X ARYA (F1)	112.023 (2nd)	112.023 (2nd)		
RW X AYSHA (F1)	98.883 (5th)	98.883 (5th)		
BABY BELL X AYSHA (F1)	111.533 (3rd)	111.533 (3rd)		
BABY BELL X C/4 (F1)	104.430 (4th)	104.430 (4th)		
8/4 X RW (F1)	45.270 (10th)	45.270 (10th)		

Table 10 : Computations of Cook statistic and AP Statistic of 100 seed weight (mg) of bell pepper

VARIETIES	REPLICATION 1	REPLICATION 2	REPLICATION 3
BABY BELL	0.0000793	0.0000127	0.000155
	(0.611093)	(0.611108)	(0.611076)
8/4	0.012911	0.094276	0.176965
	(0.608172)	(0.589647)	(0.570821)
C/4	0.074641	0.00000812	0.076206
	(0.594117)	(0.611109)	(0.593761)
C/4 (YELLOW)	0.067799	0.000268	0.059535
	(0.595675)	(0.61105)	(0.597557)
AYESHA	**0.381933***	0.092535	0.098478
	(0.524155)	(0.590044)	(0.58869)
ARYA	0.029111	0.000268	0.023788
	(0.604483)	(0.61105)	(0.605695)
ROYAL WONDER	0.003886	0.005075	0.0000793
	(0.610227)	(0.609956)	(0.611093)
RW X ARYA (F1)	0.059535	0.012992	0.016903
	(0.597557)	(0.608153)	(0.607263)
RW X AYSHA (F1)	0.103917	0.031974	0.020606
	(0.587452)	(0.603832)	(0.60642)
BABY BELL X AYSHA (F1)	0.001678	0.038381	0.056109
	(0.610729)	(0.602373)	(0.598337)
BABY BELL X C/4 (F1)	0.003886	0.032485	0.013901
	(0.610227)	(0.603715)	(0.607946)
8/4 X RW (F1)	0.015277	0.002079	0.028627
	(0.607633)	(0.610638)	(0.604594)

Note. Values in parentheses are AP statistic values.

* Significant at 95% i.e. (1-α) x100% probability level.

Table 11:Mean values with corresponding ranks and EMS of ANOVA table before and after removal of significant outlier form the experiment on 100 seed weight (mg) of bell pepper

Varieties	Mean (Actual)	Mean (Considering outlier)	EMS (Actual)	EMS (Considering outlier)
BABY BELL	8.4600 (7^{th})	8.460 (7^{th})	0.180935	0.180681
8/4	8.6867 (5^{th})	8.687 (5^{th})		
C/4	8.5400 (6^{th})	8.540 (6^{th})		
C/4 (YELLOW)	7.9300 *(10^{th})	7.930 (10^{th})		
AYESHA	**10.1300 (1^{st})**	**9.603 (1^{st})**		
ARYA	6.7400 (11^{th})	6.740 (11^{th})		
ROYAL WONDER	9.2767 (2^{nd})	9.277 (2^{nd})		
RW X ARYA (F1)	8.0933 (8^{th})	8.093 (8^{th})		
RW X AYSHA (F1)	8.6867 (5^{th})	8.687 (5^{th})		
BABY BELL X AYSHA (F1)	8.0500 (9^{th})	8.050 (9^{th})		
BABY BELL X C/4 (F1)	9.2067 (3^{rd})	9.207 (3^{rd})		
8/4 X RW (F1)	8.7067 (4^{th})	8.707 (4^{th})		

Table 12: Computations of Cook statistic and AP Statistic of fruit yield/ plant (g) of bell pepper

VARIETIES	REPLICATION 1	REPLICATION 2	REPLICATION 3
BABY BELL	0.000311	0.0000179	0.00018
	(0.61104)	(0.611107)	(0.61107)
8/4	0.222852	0.040891	0.072822
	(0.560374)	(0.601801)	(0.594532)
C/4	0.021073	**0.555803***	0.360428
	(0.606314)	**(0.484571**)	(0.529052)
C/4 (YELLOW)	0.001078	0.063602	0.081244
	(0.610866)	(0.596631)	(0.592614)
AYESHA	0.000195	0.003727	0.002216
	(0.611067)	(0.610263)	(0.610607)
ARYA	0.003	0.045509	0.02514
	(0.610428)	(0.60075)	(0.605388)
ROYAL WONDER	0.018149	0.012353	0.000556
	(0.606979)	(0.608299)	(0.610985)
RW X ARYA (F1)	0.01424	0.007175	0.001199
	(0.607869)	(0.609478)	(0.610838)
RW X AYSHA (F1)	0.017936	0.025843	0.00072
	(0.607028)	(0.605228)	(0.610947)
BABY BELL X AYSHA (F1)	0.005074	0.000858	0.001759
	(0.609956)	(0.610916)	(0.610711)
BABY BELL X C/4 (F1)	0.011871	0.01253	0.00000889
	(0.608408)	(0.608258)	(0.611109)
8/4 X RW (F1)	0.003453	0.002467	0.0000827
	(0.610325)	(0.61055)	(0.611092)

Note. Values in parentheses are AP statistic values.

* Significant at 95% i.e. (1-α) x100% probability level.

Table 13:Mean values with corresponding ranks and EMS of ANOVA table before and after removal of significant outlier form the experiment fruit yield/ plant (g) of bell pepper

Varieties	Mean (Actual)	Mean (Considering outlier)	EMS (Actual)	EMS (Considering outlier)
BABY BELL	549.3067 (10^{th})	549.307 (10^{th})	1520.503	1517.527
8/4	1435.503 (1^{st})	1435.503 (1^{st})		
C/4	**1127.993 (2^{nd})**	**1186.133 (2^{nd})**		
C/4 (YELLOW)	659.5433 (8^{th})	659.543 (8^{th})		
AYESHA	539.61 (11^{th})	539.610 (11^{th})		
ARYA	885.21 (4^{th})	885.210 (4^{th})		
ROYAL WONDER	331.98 (12^{th})	331.980 (12^{th})		
RW X ARYA (F1)	830.36 (5^{th})	830.360 (5^{th})		
RW X AYSHA (F1)	748.3167 (6^{th})	748.317 (6^{th})		
BABY BELL X AYSHA (F1)	727.4033 (7^{th})	727.403 (7^{th})		
BABY BELL X C/4 (F1)	953.8067 (3^{rd})	953.807 (3^{rd})		
8/4 X RW (F1)	625.21 (9^{th})	625.210 (9^{th})		

Table 9, reveals that the EMS of analysis after removal of outlier is greater than the EMS of the analysis of actual observation. Thus, it is clear that outlier removal increases the efficiency of the experiment, but the rank of the BABY BELL' has not been changed.

The Cook statistics and AP statistics values were calculated for the experiment on 100 seed weight (mg) of bell pepper and those values are tabulated in Table 10.

In Table 10, it is observed that the 13^{th} observation, placed in first replication/block of 'AYESHA' shows the maximum value of Cook Statistic as well as minimum value of AP statistic (in parenthesis). Thus the value of Cook Statistic is significant to be influential.

Table 11, reveals that the EMS of analysis after removal of outlier is less than the EMS of the analysis of actual observation. Thus, it is clear that outlier removal increases the efficiency of the experiment, but the rank of the 'AYESHA' has not been changed.

The Cook statistics and AP statistics values were calculated for the experiment on fruit yield/ plant (g) of bell pepper. Those values are tabulated in Table 12.

In Table 12, it is observed that the 8^{th} observation, placed in second replication/ block of 'C/4' shows the maximum value of Cook Statistic as well as minimum value of AP statistic (in parenthesis). Thus the value of Cook Statistic is significant to be influential.

Table 13**,** reveals that the EMS of analysis after removal of outlier is greater than the EMS of the analysis of actual observation. Thus, it is clear that outlier removal increases the efficiency of the experiment, but the rank of the 'C/4' has not been changed.

REFERENCES

Andrews, D.F. and Pregibon, D. 1978. Finding the outliers that matter. *J. Roy. Statist. Soc.*, **B 40**: 87-93.

Bhar, L. and Gupta, V.K. 2001. A useful statistic for studying outliers in experimental designs. *Sankhya*, **B63**: 338-350.

Bhar, L. and Gupta, V.K. 2003. Study of outliers under variance-inflation model in experimental designs. *J. Ind. Soc. Agril. Stat.*, **56(2)** : 142-154.

Bhar, L., Gupta, V.K. and Parsad, R. 2013. Detection of Outliers in designed experiments in the presence of masking. *Statistics and Applications*, **11(1&2)**: 147-160.

Chang, I. 1982. Outliers in Time Series. Ph.D. Thesis, Department of Statistics, University of Wisconsin, Madison.

Cook, R.D. 1977. Detection of influential observations in linear regression. *Technometrics*, **19 :** 15-18.

Cook, R.D. 1979. Influential observations in linear regression. *J. Amer. Statist. Assoc*. **74 :** 169-174.

De Mendiburu Delgado, F. Una herramienta de análisis estadístico para la investigación agrícola, 2009.

Hazra, P., Chattopadhyay, A. and Dutta, S. 2011. *Modern Technology in Vegetable Production.* New India Publishing Agency, New Delhi, pp.128-140

Hocking, R. R. 1984. Discussion of Gray and Ling Paper. *Technometrics*, **26**: 321-323.

Mitra, S.K., Sadhu, M.K. and Bose, T.K. 1990. *Nutrition of Vegetable Crops*. Naya Prokash, Calcutta, pp. 75-100

Ojha, S. and Bhar, L. 2015. Detection of outliers in designed experiments with correlated errors. *J. Indian. Soc. Agril. Statist*., **69(1)** : 57-63.

Pena, D. and Yohai, V.J. 1995. The detection of influential subsets in linear regression by using an influence matrix. *J. Roy. Statist. Soc*., **57:** 145-156.

R. Core Team. R: A language and environment for statistical computing, 2013. https://cran.r-project.org/bin/windows/base/old/3.4.3/

Saha, G. 2016. Detection of outliers in linear time series in Agricultural data. M.Sc. Thesis, Indian Agricultural Research Institute, New Delhi.