Multi-response Regression Modeling for an Agricultural Experiment

Nawzad M. Ahmed[1,3] and Handren A. Taher[2]

[1]Department of Statistics & Informatics, College of Administration & Economy, University of Sulaimani,

Kurdistan Region – F.R. Iraq

[2]Department of Statistics & computer, College of Commerce, University of Sulaimani, Kurdistan Region – F.R. Iraq

[3]Department of Finance & Banking, College of Administration and Economics, University of Human Development
Sulaimani, Kurdistan Region – F.R. Iraq

Abstract—During building model process, it is difficult to construct a multiple regression model (MRM) while the response variable(Y) is proposed as a vector of (r.v), $(Y_1, Y_2, Y_3, \ldots, Y_n)$ in an experiment. So that a single response (MRM) is not able to perform multi-response data (MRD) separately (one for each response), this is because of the linear dependency(LD) among responses, then (MRRM) which was proposed by (Len Beirman, & Freidman 1997) has better performance to detect effects and patterns for the factors, (Explanatory)$(X_1, X_2, X_3, \ldots, X_k)$ that are introduced to the (MRRM) system on the (r.v) altogether. This Study was applied (MRRM) on an agricultural experiment through (450m2) in west Sulaimani- Kurdistan Region-Iraq.

Keywords: Multiple regression model (MRM), Response Vector (r.v) Multi-response Data (MRD). Linear dependency (LD), Multi-response Regression Model (MRRM), Kronecker Product, Analysis of Variance (ANOVA), Coefficient of Determination (R2), Eigen Values (EV), Eigen Vector (EV).

I. INTRODUCTION[2], [3], [7]

The plot of land that was experienced upon owns specifics of agricultural land in general and the land area of experiment was (450m2). It is located in the west of Sulaimani, and area called Farmanday-Gshty. In order to apply the experiment, the area was divided into (36) equal blocks. The area of each block was (12m2), and a large area had been chosen to avoid the interaction between the pieces of the experiment units. The factors are (3) different levels of nitrogen (N), (4) different levels of phosphorus (P), and (3) different levels of potassium (K). We took all combinations of three outputs between them which are equal to (36) combinations, each combination consists of three levels of (N, P, and K) and was given to a piece of land. We had to measure the average number of leaves as (Y_1), average height of plant as (Y_2), average circumference as (Y_3), and the average weight of sweet corn flower as (Y_4), all per block of the plants. Each experimental block contained (30) plants, the response variables indicates the average of each experimental block, the circumference of each plant was taken from three different points which are bottom, middle and top , then taking average of them for each plant in the experimental block. The experience duration time was (60) days. The data that had been collected as described in previously are used to perform appropriate (MRRM), and analysis. (matlab v.7.0 and R v.3.2.4 were used).

I-1 Hypotheses

(MRRM) is one of more sensitive type of regression because of the independency between the factors(X_i, i=1,2,…,k), also is a type of regression that deals with more than one response(r.v) as a dependent variable which are linearly dependents by the effect of factors affects these responses as a vector at once. Moreover in order to make sure that the (r.v's) are linearly dependent, this hypothesis must be tested statistically.

I-2 Objective of the study

The aim of this study is to diagnose and define the behaviors and patterns in data experiment by fitting them in a statistical model that is characterized by multi-response regression model. Also using the fitted model to determine the aggregate properties of the material experience (sweet corn) in terms of (Y_1, Y_2, Y_3, Y_4), and at the same time comparing the results of the experiment in both cases (single response, for each Y_1, Y_2, Y_3, Y_4) separately through the use of (MRM) and all responses together(MRRM).

I-3 Statistical Technique for (MRRM)

When more than a response variable exists, it is not easy to estimate postulated models for all responses. (Leo. Breiman, and Jerom. H. Friedman, 1997) proposed a new technique for such studies that have more than one response variables depending on assumptions of linear modeling and its

hypotheses to predict multi-responses in a system of linear equation.

Let (N) be the number of experimental trails and (R) be the number of response variables and (K) be the number of explanatory variables, and then assuming that the responses variables can be shown as polynomial regression model in the value of (Xj) within a certain regression [7], [8], [10], [13].

II. THEORY:

II-1: The response model Estimation formula:[11], [12]

$$Y_i = Z_i B_i + E_i \qquad (1)$$

Where:

Y_i is an (Nx1) vector of observation on (i[th]) response.

Z_i is an (NxP$_i$) matrix of rank (P$_i$) with known function of standardized variables.

B_i is a (P$_i$ x 1) vector of unknown constant parameters.

$E_i$is a random error vector associated with the (i[th]) response. Also we assume that:

$E(\epsilon_i) = 0$

$Var(\epsilon_i) = \delta_{ii}I_N, \quad i = 1,2,3,\dots,r$

$$Cov(\epsilon_i, \epsilon_j) = \delta_{ij}I_N \; i,j = 1,2,3,\dots,r, \; i\neq j \qquad (2)$$

Then the (rxr) matrix whose (i , and j)[th]element is (δ_{ij}) will be denoted by($\sum$).

Also we can rewrite the equation (2.1) as follow:

$$\tilde{Y} = Z B + \epsilon \qquad (3)$$

Where:

$\tilde{Y} = (\acute{Y}_1 : \acute{Y}_2 : \dots : \acute{Y}_r)'$

$B = (B'_1 : B'_2 : \dots : B'_r)'$

$\epsilon = (\epsilon'_1 : \epsilon'_2 : \dots : \epsilon'_r)'$

$Z = diag(Z_1, Z_2, \dots, Z_r)$

From equation (2) we can see that (ϵ) has the variance-Covariance matrix:

$$Var\text{-}Cov(\epsilon) = \sum \otimes I_N \qquad (4)$$

Where Kronecker product is denoted by $\otimes$, is an operation on two matrices of arbitrary size resulting in a block matrix. It is a generalization of the outer product (which is denoted by the same symbol) from vectors to matrices, and gives the matrix of the tensor product with respect to a standard choice of basis. In order to get best linear unbiased estimator (BLUE) we should apply the following formula:

$$\underline{\hat{B}} = (Z'\Delta^{-1}Z)^{-1}Z'\Delta^{-1}\tilde{Y} \qquad \textbf{(5)}$$

$$Var\text{-}Cov(\underline{\hat{B}}) = (Z'\Delta^{-1}Z)^{-1} \qquad \textbf{(6)}$$

But ($\sum$) must be known in both eq. (5, 6)

If ($\sum$) is unknown as it is usual case then the an estimate of B can be obtained by replacing ($\sum$) in eq.(6)by estimating $\sum^{est}$ provided that this estimate is non-singular , this criteria was proposed by Zellner(1962), and is given by $\sum^{est}=\underline{\delta}_{ij}^{est}$

Where:

$\underline{\delta}_{ij}^{est} = Y_i' [IN - Z_i(Z_i'Z_i)^{-1}Z_i'][IN - Z_j(Z_j'Z_j)^{-1}Z_j'] Yj/N$,

i,j = 1,2,…,r   **(7)**

We note that ($\underline{\delta}_{ij}^{est}$) is computed from the residual vector which resulted from an ordinary least squares fit of the (i[th]) and (j[th]) single response model to their respective data sets, using this estimate of ($\sum$)in eq(6) we get the estimate :

$$\Delta = \sum \otimes I_N \qquad (8)$$

Hence

$$\Delta^{-1} = \sum^{-1} \otimes I_N \qquad (9)$$

Where (Δ) is a kronecker product matrix which diagonal contains ($\underline{\sum}$), and this matrix makes the multi-response model studies the parallel effect of all explanatories or all response variables, then it can be said that the responses act under the explanatories at the same time.

II-2 Linear Dependency among the Responses [1], [4], [6]

Suppose that we have (m) relationships among responses and they are represented as:

$$B (Y_{u1}: Y_{u2}: \dots : Y_{ur})' = C \; , \; : u=1,2,3,\dots,N \qquad (10)$$

B is an (mxr) matrix of rank (m<r) of constant coefficients ($Y_{u1}: Y_{u2}: \dots : Y_{ur}$) is the ($u^{th}$) row of the (Nxr) data matrix (Y).

C is an (mx1) vector of constants.

Then eq.(10)can be shown as :

$$B \acute{Y} = 1_N' \otimes C \qquad (11)$$

We can detect linear dependencies by Eigen value analysis. Let's suppose that rounding errors in the response values exists and they are distributed independently and uniformly over interval (-δ, δ) the quantity (δ) is equal to one half of the last digit reported when all the multi-responses values are rounded to the same number of significant figures.

Let (λ) denote a small Eigen value of (DD´) where (DD´) is a matrix of Eigen values which would be zero if it were not for the rounding error.

Where:

$$D = \acute{Y}[IN - (1_N 1_N')/N] \qquad (12)$$

II-3 Test of Lack of Fit in Multi-response Regression [5], [9],

In regression model lack-of-fit will exist when it fails to fit the functional relationship between the explanatory variables and the response variable.

Lack-of-fit may occur when two different values of response variable have the same quantity of treatment, for example if we have a sample of size (20) or any other size using headache's drug for each person in the sample due to their age. It will be inadequate to give drug (500 mg) to a person whose (10) years old and to a person whose (30) years old, unusually large residuals result from fitting the model might be occurred. That is why lack-of-fit test is required in these cases; also every linear or nonlinear model requires lack of fit test to increase the accuracy of the model under study. In multi-response regression model the test is similar to the test in single response linear regression model in addition to some differences in this model application.

Let we have a multi-response model as follow:

$$\underline{Y} = \underline{w}\underline{\Gamma} + \underline{\delta} \qquad (13)$$

Where:

$\underline{Y} = [Y_1 : Y_2 : \dots : Y_r]$

$\underline{w} = [Z_1: Z_2: \dots : Z_r]$

$\underline{\delta} = [\epsilon_1 : \epsilon_2 : \ldots : \epsilon_r]$

$\underline{\Gamma} = [B_1, B_2, \ldots, B_r]$, and let $\underline{C} = (C_1, C_2, \ldots, C_r)'$ be an arbitrary non zero rx1 vector.

From equation (10) we obtain the model:

$$Yc = w\, \Omega c + \epsilon\, c \qquad (14)$$

Where:

$(\mathbf{Y_c = Yc})$ is a vector of (N) observations on the single response.

The multi-response model in (13) is an adequate if and only if the single models in (14) are adequate for all ($c \neq 0$). Equivalently, if some ($c = 0$) model (14) is inadequate, then the multi-response model (13) is also inadequate. Since (Y_c) has the multivariate normal distribution N (0, $\boldsymbol{\delta^2_c I_N}$).

$Y_c = \sum_{i=0}^{r} CiYi$, $\Omega_c = \Gamma\, C$, $\epsilon_c = \delta_c$ Where [$\epsilon_c \sim$ N $(0, \delta_c\, I_N)$].

The test of lack of fit of eq(14) can be written as :

$$F(c) = V_{pe}\, SSLof(c) / V_{Lof} SSp\epsilon\, (c) \text{ ,where } c \neq 0 \qquad (15)$$

$SSp\epsilon$ (c) and SSLof(c) are pure, and lack of fit error (Lof) sum squares respectively. (V_{pe}, V_{Lof}) are degrees of freedom for pure and lack of fit error respectively.

$$SSp\epsilon\, (c) = \acute{Y}\, K\, Y \qquad (16)$$
$$SSLof(c) = \acute{Y}\, [IN - w_i(w_i'w_i)^{-1}w_i' - k]\, Y \qquad (17)$$

Where K = diag ($k_1, k_2, \ldots, k_n$, 0)

$$K = I_{vi} - (1/vi)\, J_{vi} \, , \quad i=1,2,\ldots,n \qquad (18)$$

Where (J_{vi}) is the matrix of one's of order (v_i x v_i) and (v_i) is the number of observations at the (i^{th}) repeat run side (i=1, 2 ……. n).

The matrix of (W) in eq(17) is not necessary to be a full column rank , then we denote the quadratic form in both eq(17, 18) by (G_1 and G_2) respectively.

$$G_1 = \acute{Y}\, [IN - w_i(w_i'w_i)^{-1}w_i' - k]Y \qquad \mathbf{(19)}$$
$$G_2 = \acute{Y}\, K\, Y \qquad \mathbf{(20)}$$

Then the test statistics of eq(15) will be rewritten as :

$$F_{(c)} = (V_{pe}\, C`G_1\, C) / (V_{Lof}\, C`G_2\, C) \qquad \mathbf{(21)}$$

If the model (14) is correct then F(c) has the F-distribution with (V_{lof} and V_{pe}) degrees of freedom. A large value of F(c), or equivalently, a large value of ($C`G_1\, C$ / $C`G_2\, C$) leads us to believe that eq(14) is inadequate and eq(13) is considered inadequate if at least one of the models in eq(14) is inadequate for some ($c \neq 0$) ,then model(13) has a significant (LOF), if Max ($C`G_1 C/C`G_2\, C$) exceeds a certain critical value.

III. Application

Multi-response linear regression was used to study the effects of (Nitrogen, Phosphorus, Potassium) in different levels as explanatory variables on responses [average number of leaves as (Y_1), average height as (Y_2), average circumference as (Y_3) and the average weight sweet corn flower as (Y_4)] of the plants at the same time, multi-response linear regression is different from the other models in technique, it can handle more than a single response at the same time and its output is

a system of linear models according to the numbers of the response variables in the study. The data that had been collected as described in previous sections are used to perform a multiple response regression model analysis. (matlab v.7.0 and R v.3.2.4) were used.

III-1 Multi-response linear regression algorithm:

The procedure of Multi-response linear regression can be shown in following steps

First step: standardize the explanatory and response variables.

Second step: fitting models for each response separately.

Third Step: calculating kronecker product (Δ^{-1}) where its dimensions is (144x144) from eq (9).

Fourth Step: calculating fisher matrix ($Z'\Delta^{-1}Z)^{-1}$ using kronecker product and steps above respectively gets the following matrix.

$$(Z'\Delta^{-1}Z)^{-1} = \begin{pmatrix} 3.45E\text{-}05 & 1.67E\text{-}07 & 1.36E\text{-}03 & 3.17E\text{-}08 \\ 1.67E\text{-}07 & 7.11E\text{-}09 & 9.48E\text{-}06 & 5.33E\text{-}10 \\ 1.36E\text{-}03 & 9.48E\text{-}06 & 2.89E\text{-}01 & 5.65E\text{-}06 \\ 3.17E\text{-}08 & 5.33E\text{-}10 & 5.65E\text{-}06 & 7.63E\text{-}10 \end{pmatrix}$$

Fifth Step: calculating ($\hat{\underline{B}}$) by performing eq(13), where ($\hat{\underline{B}}$) is the estimated parameters of the four models together.

$$\hat{\underline{B}} = \begin{pmatrix} -0.36446 & 0.170529 & -0.49591 & 1.008166 \\ 0.318848 & 0.613716 & 0.620585 & 0.281032 \\ -0.23337 & 0.273046 & -0.80172 & 0.187157 \\ -0.0932 & -0.06955 & 1.603784 & -0.76545 \end{pmatrix}$$

The first column of ($\hat{\underline{B}}$) matrix is the estimated parameters of the average number of leaves in plant per block, the second, third and fourth columns are the estimated parameters of the average height plant per block, average circumference plant per block and average weight of sweet corn flower of plant per block respectively.

Sixth step: Testing Significance of Explanatory variables by using the T-test, the (V-Cov of $\hat{\underline{B}}$) is equal to ($Z'\Delta^{-1}Z)^{-1}$, now it is easy to perform the test by dividing the ($\hat{\underline{B}}$) matrix by the root square of [V-Cov($\hat{\underline{B}}$)] as it is shown below:

As it is obvious from the above table that the calculated values of (t-test) is greater than (-table) value which is equal to (1.36) and this leads to say there is no reason to accept he null hypotheses that says the parameters is equal to zero, all parameters ($\hat{\underline{B}}$) are significant in the system.

Seventh step: Identify Detection of System Models. This step needs to calculate *Order* condition, and *rank* condition which calculated as follow:

Table (1) Summary of t-test for the significance of parameters

Models	$\hat{\underline{B}}$	S.E of $\hat{\underline{B}}$	t-test value
First model	-0.36446	0.005312626	78.03
	0.318848	2.28E-05	26.29
	-0.23337	0.002505092	58.25
	-0.0932	9.95E-07	14.95
Second model	0.170529	2.28E-05	611.08
	0.613716	7.92E-05	7340.31
	0.273046	1.75E-05	3189.18
	-0.06955	3.12E-06	317.23
Third model	-0.49591	0.002505092	4.35
	0.620585	1.75E-05	3.59
	-0.80172	0.537581365	5.46
	1.603784	1.05E-05	3.63
Fourth model	1.008166	9.95E-07	39719.93
	0.281032	3.12E-06	10186.83
	0.187157	1.05E-05	6997.07
	-0.76545	2.53E-05	30251.73

Order condition:
Table (2): Shows the calculation of order condition of the system models

	G	K	M	G-1	K-M
Model-1	4	7	4	3	3
Model-2	4	7	4	3	3
Model-3	4	7	4	3	3
Model-4	4	7	4	3	3

From Table (2) it is clear that the four models are identified, order condition is satisfied (K-M = G-1).

Table (3): The rank condition calculation of system models

	Y1	Y2	Y3	Y4	Intercept	X1	X2	X3
Model-1	1	0	0	0	0.364	-3188	0.233	0.094
Model-2	0	1	0	0	-0.171	-0.614	-0.273	0.069
Model-3	0	0	1	0	0.496	-0.621	0.802	-1.60
Model-4	0	0	0	1	-1.008	-0.281	-0.187	0.765

Since there exist at least, one sub non-singular matrix in the four postulated models, then the system of models is exactly identified.

III-2 Multi-response Model Fitting, Results Summary:
Applying the multi-response linear regression algorithm to carry out a system of models at the same time that is discussed previously after standardizing all variables under consideration, obtaining these estimated system models bellow:

$\hat{Y}1 = -0.36446 + 0.318848\ Z1 - 0.23337\ Z2 - 0.0932\ Z3$

$\hat{Y}2 = 0.170529 + 0.613716\ Z1 + 0.273046\ Z2 - 0.06955\ Z3$ ......(22)
$\hat{Y}3 = -0.4959127 + 0.620585 Z1 - 0.80172 Z2 + 1.603784\ Z3$
$\hat{Y}4 = 1.008166 + 0.281032\ Z1 + 0.187157\ Z2 - 0.76545\ Z3$

The above models represent estimated models for average number of leaves per block, average height plant per block, average circumference plant per block and average yield plant per block respectively. After fitting the four models it is required to test each of them separately as comes from the estimation and testing of multi-response technique, for this purpose(ANOVA) table and coefficients of determination (R^2) is calculated for each model. for the first model, testing the null hypothesis H_0: the model is not significant.

III-3 First model significant testing:

From table (4) it is clear that the value of F-calculated is equal to (205.2217), this implies that there is no reason to accept the null hypothesis, also (R^2) is equal to (0.951) which means that the average number of leaves in plant is changed by the amount (0.951) due to changing in composts (N,P, and K), so this model can depend on predicting the average number of leaves in plant.

Table (4): ANOVA (first model)

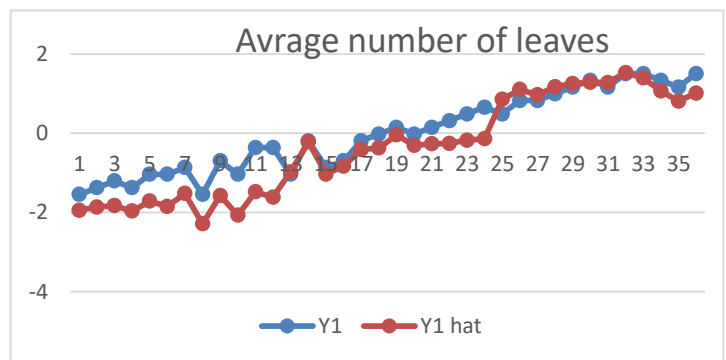| Source	d.f	SS	M.S	F-Test
SSR	3	21098.38	7032.795	205.2217
SSE	32	1096.616	34.26925	
SST	35	22195		

Fig.1: Actual and estimated values of average number of leaves plant per block

III-4 Second model significant testing:

Table (5): ANOVA (second model)

Source	d.f	SS	M.S	F-Test

SSR	3	1241694	413898.1	2927.936
SSE	32	4523.575	141.3617	
SST	35	1246218		

From Table (5) it is clear that the value of F-calculated is equal to (2927.936) and F-table d.f (3,32), and level of sign. (0.95) is equal to (2.92) it is far smaller than F-calculated, this implies that there is no reason to accept the null hypothesis that the model is not significant, also (R^2)is equal to (0.9964) which means that the average height of plant is changed by the amount (%99.64) due to changes in composts. Therefore this model can depend on predicting the average height of plant.
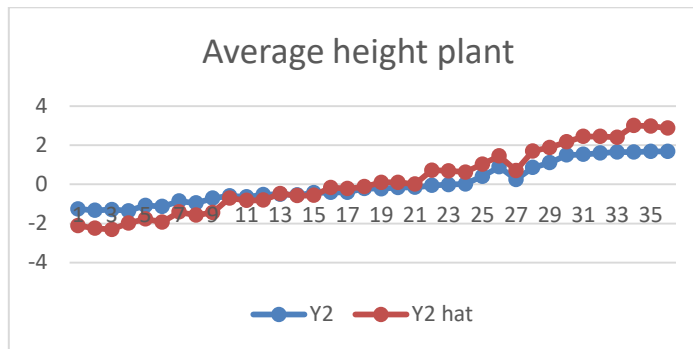

Fig.2: Actual and estimated values of average height of plant per block

III-5 *Third model (Significant testing):*
From table (6) it is clear that the value of F-calculated is equal to (32563009.7), F-table (3,32,0.95) is equal to (2.92) is much smaller than F-calculated this implies that there is no reason to accept the null hypothesis, also (R^2) is equal to(0.9999997) it means that the average circumference of plant is changed by the amount (%99.99997) due to changing in composts. Therefore this model can depended upon to predict the average circumference of plant.

Table (6): ANOVA (third model)

Source	d.f	SS	M.S	F-Test
SSR	3	542.7598	180.9199	32563009.7
SSE	32	0.000178	5.56E-06	
SST	35	542.76		

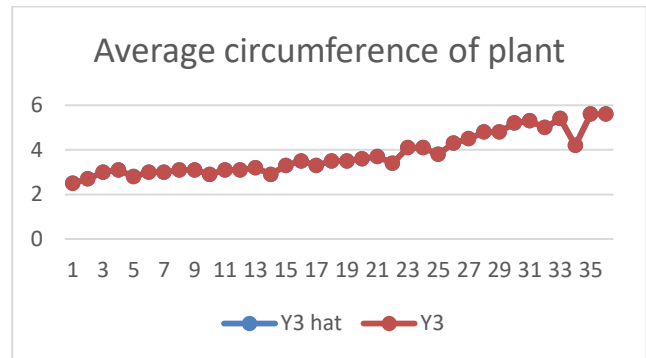
Fig.3:Actual and estimated values of average circumference plant per block

Figure (3) represents the actual and estimated values of average circumference plant per block, it is clear that from the first point to the last point the actual and the predicted values are rather matched approximately.

III-5 *Forth model (Significant testing):*
From table (7) it is clear that the value of F-calculated is equal to (285.9704) and F-table (3, 32, and0.95) is equal to (2.92) it is smaller than F-calculated, this implies that there is no reason to accept the null hypothesis, also (R^2) is equal to(0.964) it means that the average crop of plant is changed by the amount (%96.4) due to changing in composts. Therefore this model can be depended upon to predict the average crop of plant.

Table (7): ANOVA of fourth model

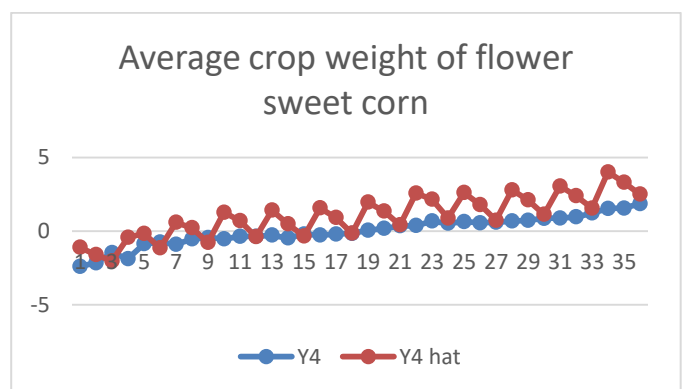| Source	d.f	SS	M.S	F-Test
SSR	3	1367497	455832.4	285.9704
SSE	32	51007.51	1593.985	
SST	35	1418505		

Figure (4): Actual and est. values of average crop weight of flower sweet corn / block

III-6 *Linear dependency Test among Responses:*
The value of (δ_{re}^2 and $\mu\lambda$) is equal to (0.00083, 0.02905) respectively by using (δ=0.05), where (δ_{re}^2)is a rounded error variance, (μ_λ)is the expected value of(μ_λ)is a small Eigen value of matrix (DD`). Then calculating the Eigen value and Eigen vectors of (DD`) matrix which has been calculated.

Table (8): the Eigen values and Eigen vectors of (DD`) matrix

Eigen value	Eigen vectors			
1.4144	0.3752	-0.7746	-0.0710	0.5042
2.4202	-0.8367	-0.0915	0.1834	0.5079
5.0646	0.3895	0.4539	0.6293	0.4962
131.1008	0.0865	0.4308	-0.7519	0.4916

From the above table it is obvious there is no Eigen value of (DD`) matrix which falls within four to five values of ($\delta_\lambda$) from ($\mu_\lambda$), hence, no Eigen value of (DD`) may be regarded as zero. This indicates that no linear relationship exists among the four responses.

III-7 *Lack of Fit Test:* Testing the multi-response regression in order to detect if the models suffer from lack of fit or not we should calculate lack of fit and pure error as (G_1, G_2) matrices respectively according to eq.(19, 20) to test the hypophysis below:
H$_0$: The model doesn't suffer from lack of fit.
H$_1$: The model suffers from lack of fit.
The matrices of (G_1), (G_2) and ($G_2^{-1}G_1$) are showed below:

$$G_1 = \begin{Bmatrix} -23.1149 & -20.6098 & -19.9665 & -22.1829 \\ -20.6098 & -17.8999 & -17.1215 & -19.2101 \\ -19.9665 & -17.1215 & -15.2866 & -18.5552 \\ -22.1829 & -19.2101 & -18.5552 & -21.3507 \end{Bmatrix}$$

Where (G_1) is the lack of fit error which is calculated by performing the equation (20).
Where (G$_2$) is the lack of fit error which is calculated by performing the equations (19, 20).

$$G_2^{-1}G_1 = \begin{pmatrix} -0.36723 & -0.57386 & -0.74056 & -0.52834 \\ 0.026295 & 0.049945 & -0.29956 & 0.201096 \\ -0.11951 & 0.056884 & 0.598074 & -0.04324 \\ -0.50600 & -0.35004 & -0.30418 & -0.52882 \end{pmatrix}$$

Also the calculated Eigen values and Eigen vectors of ($G_2^{-1}G_1$) is shown in table (9) below.

Table (9): Shows the Eigen values and eigenvectors of ($G_2^{-1}G_1$)

Eig-values	Eig-vectors			
-0.9324	0.64783828	-0.41684	0.141741	-0.60046
0.6261	-0.1460923	-0.40645	0.686737	0.738479
-0.1202	0.07702793	0.807697	-0.07349	-0.25327
0.1784	0.7436594	0.093086	-0.70915	0.173044

Table (10): Shows the combinations of the models test of lack of fit

No	Subset	emax($G_2^{-1}G_1$)	F-Value	P-Value
1	Y1,Y2,Y3,Y4	0.6261	0.175308	0.854
2	Y1,Y2,Y3	2.7302	0.764456	0.652
3	Y2,Y3,Y4	1.5630	0.43764	0.423
4	Y1,Y2,Y4	7.9853	2.235884	0.114
5	Y1,Y3,Y4	0.6732	0.188496	0.823
6	Y1,Y2	10.4033	2.912924	0.102
7	Y1,Y3	4.5003	1.260084	0.22
8	Y1,Y4	1.8600	0.5208	0.478
9	Y2,Y3	6.3404	1.775312	0.1067
10	Y2,Y4	9.9763	2.793364	0.1007
11	Y3,Y4	2.4500	0.686	0.7604
12	Y1	3.4000	0.952	0.248
13	Y2	3.8600	1.0808	0.2835
14	Y3	8.6523	2.422644	0.127
15	Y4	2.4500	0.686	0.743

From Table (10) it is obvious that none of the responses are contributing lack of fit then there is no reason to reject H$_0$. The table (10) shows the lack of fit test according to the number of the responses in the procedure of multi response linear regression analysis, in the first step the four responses have been inputted and tested, and then the combinations of all three responses at once have been inputted and tested, then the combinations of all two responses at once have been inputted and tested. Finally the single response has been inputted one by one and tested, the matrix of (**G$_1$**) and (**G$_2$**) have been calculated in each of the fifteen steps and the results of the fifteen combinations model analysis tests were accepting the null hypothesis which means the model doesn't suffer from lack of significance.

$$G_2 = \begin{pmatrix} 26.41577 & 22.91441 & 22.22107 & 22.45285 \\ 22.91441 & 22.47612 & 21.29689 & 20.23854 \\ 22.22107 & 21.29689 & 22.36547 & 19.15687 \\ 22.45285 & 20.23854 & 19.15687 & 24.0715 \end{pmatrix}$$

IV. RESULTS ANALYSIS

After using multi-response regression in this study, the researcher introduced some conclusions due to the application part of the agricultural experiment. So later in the study the researcher added some important points as follow:
1-Multi response regression is a pathological technique to carry out the production of more than one linearly dependent response (Multi-response) because it can produce more than one response parallel at the same time.
2- The researcher concluded that from the first model which was for the average number of leaves in plant per block, the min. number of leaves in plant per block was approximately(17) at the time that the composts (N, P, K) equaled (30, 0, 0) grams respectively. The maximum is approximately (26) as an average per block for each plant corresponding to the composts skill (N, P, K) by (50, 0, 20) grams per plant.
3-From the previous conclusion it is clear that nitrogen has a positive relationship with the average number of leaves in

plant per block by (0.318848). See first postulated system model equations (22).

4-From the second model which was for the average height of plant per block, minimum average is equal to (161.5815 cm) when the composts (N, P, K) equal (0, 0, 50) grams per plant respectively. The maximum average height is (215.624 cm) when (N, P, K) equal (50, 50, 0) grams.

5-After predicting ($\hat{Y}2$ (average height of plant per block), (N, P, and K) and were being sorted, we found out that (N and P) have a positive relationship with the average height of plants by (0.613716, 0.273046).See the second postulated model.

6-From the third model which is the average circumference plant per block, the minimum average of circumference plant per block is equal to (2.502 cm) when composts (N, P, K) equal (0, 0, 0) grams for per plant respectively. The maximum is (5.6123 cm) when (N, P, K) equal (50, 50, 50) grams.

7-The composts (N, P and K) also have a positive relationship with the average circumference plant per block.

8-The researcher concluded from the forth model which was for the average crop weight of sweet corn flower, the minimum weight is (183.27) grams per flower when the composts (N, P, K) equal (0, 0, 50) grams per plant respectively and the maximum is (258.8443) grams when (N, P, K) equal (50, 50, 0).

9-After sorting ($\hat{Y}4$, N, P, and K) we found out that the composts (N and P) have a positive relationship with the average weight crop by (0.281032, 0.187157).See the forth postulated model.

10-Both of Y2 and Y4after being compared with each other were minimum (N, P, K) equal (0, 0, 0) grams per plant respectively and maximum when (N, P, K) equal (50, 50, 0) grams for both.

11-Since (Δ^{-1}) is a function of (V-COV) matrix of residuals then one can say that (Δ^{-1}) is a solution to remove autocorrelation problem if it exists in the system or in other word (Δ^{-1}) is similar to (GLS) method.

12-Multi-response regression have the best performance when it is compared with a single response multiple regression model because the single model is unable to study the parallel multi-response in the model so it is not able to study the side-effects of the reflections for explanatory. At the same time the multi-response regression model is able to detect all patterns (good or bad), (positive or negative) effects after the experimental data under consideration.

13- As the researcher's suggestion, the results of this study can be tried to expand into another (system of linear programming), and optimizing these products, by solving the linear system using simplex method of solution.

V. REFERENCES

Chow, G. C. (1960). "Tests of equality between sets of coefficients in two linear regressions". Econometrica, 28, 591–605.

Dobson, Annette J. "An Introduction to Generalized Linear Models". 2nd edition, (Chapman & Hall/CRC texts in statistical science series). Includes bibliographical references and index. ISBN 1-58488-165-8 (alk. paper), 2002.

Grabill, F. A. (1961). "An Introduction to Linear Statistical Models", Vol. I, McGraw- Hill Book Company, Inc. New York.

Jump up_ Pearl, J. (2000). "Causality: Models, Reasoning, and Inference". New York: Cambridge University Press. ISBN 0521773628.

Khuri, A. I. (1985). A Test for Lack of Fit of a Linear Multiresponse Model, Technometrics.

Khuri, A. I. and Conlon, M (1981). "Simultaneous optimization of Multiple Responses Represented by Polynomial Regression Functions", American Statistical Association and American Society for Quality.

Kumar, V. and Pant, A. K. (2011). "Statistical Analysis of Multi-response Plant Growth Data", Int. J. Agriculture. Stat. Sci., Vol. 7, No. 1. 289-300.

Liao, H. c. (2004). "A Data Envelopment Analysis Method For optimizing Multi- Response Problem with Censored Data in the Taguchi Method", Technical Report, Department of Health Services Administration, Chung-Shan Medical University.

Manelwijesinha Khuri .A.I (1987) "The Sequential Generation of Multi-response d- optimal Design When the Variance-Covariance Matrix is Not Known". American Statistical Association and American Society for Quality.

Ren, S. and Kim, H. (2003). "Comparative Assessment of Multi-response Regression Methods for Predicting the Mechanisms of Toxic Action of Phenols", J. Chem. Inf. Computer. Sci., 43(6), 2106-2110.

Seber, G. A. F. (1984). Multivariate Observations, John Wiley, New York. Shroff. R. (2000), Chairman's address in Pesticide information annual issue. (New Delhi: Pesticide association of Indian Publication).

Simila.T,.and Tikka, J. (2007)."Input Selection and Shrinkage in Multi-response Linear Regression", Technical Report, Laboratory of Computer and Information Science, Helsinki University of Technology, Helsinki, Finland.

Stewart, W. E., Caracotsios M. and Sorensen J. P. (1992). "Parameter Estimation from Multi-response Data", Dept. Of Chemical Engineering, University of Wisconsin, Madison