J. Plant Production, Mansoura Univ., Vol. 4 (3): 375 - 382, 2013

DETERMINING THE RELATIVE CONTRIBUTION OF YIELD COMPONENT IN BREAD WHEAT USING DYFFERENT STATISTICAL METHODS

Mohammad , Wafaa W.* ; A. A. Swelam and Iman K. Abbs***

***Central laboratory for Design and statistical Analysis Research**, Wheat Res. Deprt., Field Crop Res. Inst A R C Egypt.**

ABSTRACT

This investigation was carried out at ‹Kaffr Al-Hmam Experimental Station Sharkia Governorate during 2010/2011 and 2011/2012 to evaluate the performance of seasons eight wheat genotypes namley Sakha 93, Sakha 94, Sids1, Gemmeza 7, Gemmeza 9, Gemmeza 10, Sids10 and Giza 168. The treatments were arranged in randomized complete blocks design with three replications, in order to investigate the relationship between seed yield / plant and its factors using multivariate techniques namley; correlation, stepwise, multiple liner regression ; path –coefficient and factor analysis.

Data showed that cultivar Sakha94 recorded the highest seed / plant, and number of spikes / plant. Moreover cultivar Giza 168 recrded the lowest grain yield plant. Factor analysis grouped the studied variables in two major factors which altogether accounted for 81.00 of the total variation. The first factor include number of spike / plant, number of grains /spike, spike grain weight, and 1000-grain weight. The second factor included the remeaining variables. Multiple linear regression, stepwise and path analysis agreed upon the number of spike / plant, number of grains /spike, spike grain weight, 1000-grain weight as major contribution to seed yield variations. Factor analysis technique was more efficient than other techniques. It provides more information about cluster of intercorrelated variables. Results indicated no significant between the full model regression and stepwise for coefficient determination (R^2) and standard error of stimated value, however, the efficiency expressed is due, in fact, to the reduction in variables number in the equation from all raniables in full model regression to four variables in stepwise.

INTRODUCTION

Wheat (*Triticum aestivum* L.) is the main important food crop grown for grains in Egypt, used for human. It is important to increase the productivity of this crop. This could be achieved by two wayes, improving the variety of wheat from breeding point of view, and improving Agricore tecequenices practices. Yield is the end product of several characters. Relating these characters to define the importance contributing factor to yield is helful as selection aids in breeding programs. Correlation coefficient is not only an important statistical procedure used to facilitate breeding programs for high yield, but it is also important to examine the direct and indirect contribution (Kim and Gary, 1985). Path coefficient analysis could be used, since it divides correlation coefficients into direct and indirect effects through path ways (Deway and Lu, 1959).

Furthermore Walton (1971and 1972) suggested factor analysis as a new technique to identify growth and plant characters related to yield in

spring wheat. Denis and Adams (1978) used factor analysis to search for and identify patters of morphological characters in a set of wheat cultivars which could relate to yield.

Yildirim *et al* (1996) and Leilah and Al Khateed(2005) used factor analysis in wheat and related characters. Stepwis is used to determine the best predictive equation El-Sergany (1992) reported that step wise multiple liner regression was more efficient than the full model regression for yield. Khan and Dar (2010) used correlation and path coefficient analysis of some quantities traits in wheat.

The objectives of this study was to evaluate the relative contribution for some variables of wheat using different statistical techniques.

MATERIALS AND METHODS

Two field experiments were conducted at Kaffr Al-Hamam Experimental station, Sharkia Governorate during the two successive seasons of 20010/2011 and 2011/2012, using eight wheat genotypes namely; Sakha 93, Sakha 94, Sids1, Gemmeiza 7, Gemmeiza 9, Gemmeiza 10, Sids10 and Giza 168.

The eight genotypes were sown in November 15th 2010 and on November 18th 2011. in randomized complete block design with three replicates. Each plot consisted of 10 rows; 2 m long and 0.20 cm apart. The distance between plants was 10 cm. Recommended practices were applied.

Days to heading were recorded on plot basis. At harvest, a random sample of 10 guarded plants was collected from each plot to record the flowwing characters:

Number of tillers/plant; plant height (cm);number of spikes / plant; number of grains / spike; spike grain weight(g.),1000-grain weight (g.), and grain yield / plant (g.)

Statistical procedures:

Single and combined analysis of variance of randomized complete block design over the two seasons were performed according to Sendocor and Cochran (1980). Treatment means were compared by using least significant difference test (L.S.D.)test at 5% and 1% levels of significance. The following procedures indicated :

1- Basic statistics arthimatic mean, standard deviation, standard error were calculated. Simple correlation coefficient was computed between seed yield and its components.

2- Factor analysis:

The factor analysis procedure basically reduces a larg number of correlated variables to a small number of uncorrelated factor (Cattel 1965); when the contribution of a factor to the total percentage of the trace is less than 10%, the process stops.

After exraction, the matrix of factors is transmitted to a varimax orthogonal rotation. The effect of rotation is to accentuate the larger loading in each factor and to discard the minorloading coefficient for improving the opportunity to achieve meaning full biological interpetation of each factor.

Communality (h^2) is the variance amount of a variable accounted for the common factors together. Since the purpose was to determine the way in which yield components are related to each other, yield was not included in this structure.

3- Multiple linear regression analysis was performed as applied by Draper and Smith (1966).

4- The stewise multiple linear regression was used to compute sequance of multiple regression equation in astepwise manner (Draper and Smith 1966). At each step one variable was added to the regression equations, it was the one that caused the maximum reduction in residual sum of sequares. Equivalenty, it was the variables that had the highest partial correlation with the dependent variable adjusted for the variables already added. Similarly, it was the variable which of added, had the highest f value in the regression analysis of variance.

5- Path coefficient analysis was used as applied by Dewey and Lu (1959) .It to is used partition the total correlation coefficients between yield and its components into direct and indirect effects.

RESULTS AND DISCUSSION

1-Performence of variaties:

Mean performance of genotypes for all studied characters indicated diversity as shown in Table (1). With respect for all studied character in the two season and combined, It is noticed that grain yield/plant (gm) was significantly affected by the tested cultivars.However, Sakha94 variety produced the highest grain yield/plant, number of spikes / plant and number of tillers/plant, While; the cultivar Giza 168 produced the lowest for grain yield / plant .In addition, Sids 10 had the highest performance in both seasons and combined analysis for grain yield / plant. compared with the rest of other varieties.

2-Simple correlation coefficients mean values, standard deviation and standard error for the studied variables are presented in Table 2. The results showed that, relationship between seed yield / plant and each of the other seven components was positive and highly significant . This indicated that these characters had great influence on grain yield/plant. These findings in most cases were in accordance with those obtained by Khan and Dar (2010)

3- Factor analysis:

The results of factor analysis are recorded in Table 3. Factor analysis grouped the seven variables into two main factors which accounted for 81.01% of the total variability in dependence structure.

Factor 1 included that five variables accounted for 55.343% of the total variance. These variables were number of spike / plant, number of grains /spike, spike grain weight, 1000-grain weight.

Factor 2 included that three variables which accounted 25.671% of the total variance. These three variables were number of days to heading, plant height and number of tillers/plant. These results were similar to those of according with Leilah and Al Khateed(2005)

Table (1): Mean values of the eight characters as affected by wheat genotypes performance during 2010/2011 and 2011/2012 seasons and its combined analysis.

	Season	Sakha 93	Sakha 94	Sids1	Gemmeeza7	Gemmeeza9	Gemmeza10	Sids10	Giza168	L.S.D
Days to	S_1	88.33	93.22	90.89	99.55	95.14	96.16	87.20	99.86	2.11
heading	S_2	89.31	95.15	90.00	96.99	94.03	98.41	86.13	98.70	2.79
(day)	Comb.	88.82	94.18	90.44	98.01	94.58	97.28	86.66	99.29	1.88
Plant height	S_1	83.36	94.44	112.16	100.22	102.99	103.06	93.20	102.09	.3.22
(Cm)	S_2	83.09	96.66	112.11	102.11	104.62	112.85	97.13	102.03	2.44
	Comb.	83.22	95.55	112.14	101.16	103.80	107.95	95.16	102.06	1.24
Number of	S_1	10.90	11.22	8.22	5.33	7.36	6.34	5.01	8.22	0.67
Tillers/plant	S_2	12.80	13.00	8.89	7.44	7.56	8.22	4.34	9.00	0.89
	Comb.	11.85	12.11	8.05	6.37	7.96	7.28	4.67	8.61	0.75
Number of	S_1	852	7.00	6.73	5.00	7.15	6.00	4.37	7.45	0.315
Spikes/plant	S_2	7.54	11.44	6.25	7.05	6.81	6.54	4.03	7.13	0.403
	Comb.	8.03	9.22	6.49	6.02	6.98	5.77	4.20	7.29	0.400
Number of	S_1	73.40	66.44	72.40	63.54	66.98	62.75	73.11	52.31	1.282
Grains/spike	S_2	63.00	60.00	72.30	62.91	62.50	63.48	72.99	52.25	1.241
	Comb	68.20	63.22	72.35	63.22	64.14	62.63	73.05	52.28	1.241
Spike grain	S_1	2.88	3.66	3.66	3.11	3..52	3.66	4.09	3.10	0.720
weight	S_2	2.97	3.00	3.39	3.01	3.17	4.01	4.01	2..72	0.243
	Comb.	2.90	3.33	3.52	3.06	3.34	3.83	4.05	2.91	0.445
1000 grain	S_1	54.11	60.10	62.47	55.89	60.07	64.55	68.89	50.57	3.341
Weight (g.)	S_2	5633	60.22	62.05	56.31	60.52	62.02	67.11	51.28	2.452
	Comb.	55.22	60.16	62.26	55.10	60.29	63.28	68.0	50.92	2.651
grain yield	S_1	16.55	18.17	14.26	13.58	14.12	13.33	17.89	12.95	0.212
/plant (g.)	S_2	17.02	19.57	14.09	13.17	13.25	14.18	18.02	12.26	0.401
	Comb.	16.78	18.87	14.17	13.37	13.68	13.75	18.40	12.60	0.351

Table (2): Simple correlation coefficients, means, standard deviations and standard errors for wheat grain yield/ plant and its components over 2010/2011and 2011/2012 seasons.

Components	r value	Mean	Standard deviation	Standard error
1-Days to heading (day)	0.687**	93.68	12.441	1.710
2-Plant height (cm)	0.763**	100.13	15.447	2.073
3- Number of tillers / plant	0.897**	8.36	0.726	0.178
4- Number of spikes/ plant	0.916**	6.75	0.561	0.168
5- Number of grain / spike	0.988**	64.88	10.567	2.583
6- Spike grain Weight (g.)	0.822**	3.36	0.342	0.100
7- 1000- grain weight (g.)	0.891**	59.40	13.955	2.037
8- grain yield/plant (g.)		15.20	1.589	0.273

**** Significant at 0.01 level of significant.**

Table (3): Summary of factor leading for 8 variables of wheat.

Variables	Factors		Communality
	Factor 1	Factor 2	(h^2)
1-Days to heading (day)	0.122	0.655	0.561
2-Plant height (cm)	0.151	0.767	0.798
3- Number of tillers / plant	0.017	0.907	0.652
4- Number of spikes/ plant	0.873	0.199	0.835
5- Number of grain / spike	0.875	0.163	0.893
6- Spike grain Weight (g.)	0.751	0.153	0.0747
7- 1000- grain weight (g.)	0.876	0.189	0.860
Latent roots	2.899	2.571	5.470
Factor variance ratio %	55.343	25.671	81.014

4- Multiple linear regression analysis:

The perdication equation for grain yield is show in Table (4) and is formulated as follows:

Y = -3.312 + 0.006 X 1 + 0.008 X2 – 0.015 X 3 + 0.444 X 4** + 0.143 X** 5 + 0.193 X 6** + 0.299 X 7**

The relative contribution (R^2%) for yield factor 81.0% of the total variation in grain yield could be linearly related variation in all variables and 19.0 % could be due to residual, number of spike / plant, number of grains /spike, spike grain weight, 1000-grain weight, had the highest partial coefficient of determination (R^2= 7.456%, 6.032, 5.451, 14.245 respectively) the other characters had little contribution in the total yield variance. In addition, given that number of observations were much greater than the number of potential x variables under consideration, the addition of new variable will always increase R^2 but it will not necessary increase the precision of the estiate of the response. Therefore, the stepwise multiple linear regression analysis was carried out to determine the best variables accounted for most avariance in yield. The stepwise multiple linear regression analysis was acceding with results by Leilah and Al Khateed(2005)

Table (4): Summary of factor leading for 7 variables of wheat.

Variables	Loading	Total communality
Factor 1		25.671
1-Days to heading (day)	0.655	
2-Plant height (cm)	0.767	
3- Number of tillers / plant	**0.907**	
Factor 2		55.343
4- Number of spikes/ plant	0.873	
5- Number of grain / spike	0.875	
6- Spike grain Weight (g.)	0.751	
7- 1000- grain weight (g.)	**0.876**	
Commutative variance		81.014

5- Stepwise multiple linear regression analysis:

Either variables acceptance or removal, and relative contributions of variables (R^2%) in predecting grain yield plant are presented in Table 5. According to these results, 81.0% of the total variation to four accepted variables namely; number of spike / plant, number of grains /spike, spike grain weight, 1000-grain weight. While, number of day to heading, plant height and number of tillers/plant, were removed from the equation due to their low relative contributions.

The predication equation was formulated as follow:

Y^2 = -4.201 + 0.444 X1 + 0.143 X 2 + 0.193 X 3 + 0.299 X 4.

These results are in agreement with those reported by El-Rassas *Et .al.* (1990).

Table (5): Relative contributions of 7 components in grain yield variation over both seasons of 2010/2011 and 2011/2012by using multiple linear regression analysis.

Components	Regression coefficient	Standard error	Relative contribution partial (r^2%)
1-Days to heading (day)	-0.006	0.003	0.121
2-Plant height (cm)	0.008	0.002	0.044
3- Number of tillers / plant	0.015	0.007	0.560
4 Number of spikes/ plant	0.444	0.094	7.456**
5- Number of grain / spike	0.143	0.030	6.032**
6- Spike grain Weight (g.)	0.193	0.022	5.451**
7- 1000- grain weight (g.)	0.299	0.025	14.245**

Y intercept = -3.312 standard error of est. = 0.664
Adjusted R squered = 0.799.
R squires= 0.810
Multiple = 0.898.

6- Path coefficient analysis:

Total contribution of yield components with direct, and indirect effect are shown Table (6). Results indicated that, number of spike / plant, number of grains /spike, spike grain weight, 1000-grain weight, were of great importance of yield variation. These results agreed with Hycicek and Yildirim (2006)

Table (6): Accepted and removed variables according to stepwise analysis and the relative contributions (r^2 %) in grain yield variation over both seasons of 2010/2011 and 2011/2012.

Components	Regression coefficient	Standard error	Relative contributions (Partial %)
Accepted variables			
4- Number of spikes/ plant	0.444	0.094	7.456**
5- Number of grain / spike	**0.143**	**0.030**	**6.032****
6- Spike grain Weight (g.)	**0.193**	**0.022**	**5.451****
7- 1000- grain weight (g.)	**0.299**	**0.025**	**14.245****
Removed variables			
1-Days to heading (day)			**0.91**
2-Plant height (cm)			**0.010**
3- Number of tillers /plant			**0.90**

Y- intercept = -4.201
Standard error of est. = 0.661 R squared = 0.813
Adjusted R squared = 0.806
Multiple R = 0.896

1000-grain weight and number of spikes / plant provide to have the highest indirect contributuon to seed yield. Total contribution of studied characters mentioned above tender to the contribution 81.00 variation in seed yield/plant. It could be recommended from the previous results that:

1- The most important variables over all studied statistical procedures were grain yield/plant, number of spike / plant, number of grains /spike, spike grain weight, 1000-grain weight,. Results of factor analysis approach was more efficient than other procedures. It provides more information about cluster

intercorrelated variables . This help plant breeders to determine the mature and important of character in breeding programes. Estimated R^2 and standard error showed no significant various between the full model regression and stepwise, therefore, it could reduce the number of variables in the equation from all variables to four variable in stepwise.

Table (7): Direct and indirect effects for yield factors of seed yield / plant of wheat according to path analysis and percentage of direct effect.

Characters	Direct effect	Indirect effect	Total correlation	Direct effect %
4 Number of spikes/ plant	0.031	0.118	0.149	11.071
5- Number of grain / spike	0.020	0.140	0.160	7.143
6 Spike grain Weight (g.)	0.038	0.141	0.179	13.571
7- 1000- grain weight (g.)	0.191	0.127	0.318	68.215
R^2 %	80. 6			

**** Significant at 0.01 levels of significane**

REFERENCES

Cattell, R.B. (1965) Factor analysis on introduction to essentials The purpose and underlying models Biometrics, 21 : 190 – 215.

Denis, J.C. and M.W. Adams (1978) Afactor analysis of plant variables related to yield in dry beans. I Morphological traits Crop Sci, 18 : 74 – 78.

Dewey, D.R. and KH Lu (1959) A correlation and path coefficient analysis of components of vestal wheat gras seed production Agron J., 51 : 515 – 518.

Draper, N.R. and Smith (1966) Applied regression analysis. John Wilay and Sons, Inc. New York. PP : 171 – 172, 397-402.

El-Sergany, Dawlat, Z. (1992) Comparative study between full modern and stepwise regression analysis for maize and sugar beet yield components. J. Agric. Sci. Mansoura, Univ., 17 (7): 2291 – 2295.

Hycicek M . and T. Yildiring (2006) Path coefficient analysis of yield and yield components in bread wheat *(Triticum aestivum L* .) genotypes .Pakistan J. of Botany , 20 :256-267 .

Khan M. H. and A . N. Dar (2010) correlation and path coefficient analysis of some quantitave traits in wheat. African Crop , Sci. J. .973-1021.

Kim, N.I and M.P. Cary (1985) Response of yield attributes of isogenic tall, semidwarf, and doubled waif winter wheats to nitrogen fertilizer and seeding rates. Deparment of Agronomy, Throckmorton Hall, Kansas State, Univ., Monbottan, Kansa, 66506.U.S.A. contribution No. 86-37 J. Department of Agronomy, Kansas State Univ.

Leilah A . A and S . A . Al khateed (2005) Statistical analysis of wheat yield and drought condition. J. of Arid Envi . 61 :483 – 496 .

Sendecor, G.W. and W.G. Cochran (1980) Statistical methods Oxford and I.B. H puplisihing G. 6th Ed., 299-310.

Walton, P.D. (1971) The use of factor analysis in determining characters for yield selection in wheat. Euphytica J. 20 : 416 - 621 .

Walton, P.D. (1972) Factor analysis of yield in spring wheat (*Triticum aestivum* L.) Crop. Sci., 12 : 731 – 733.

Yildirim ,M. B. ,N. Budak and Y . Arshad (1996) .Path coefficient analysis in wheat . Pakistan J. of Sci. and Industrial Res . 37:474- 476.

تقدير المساهمة النسبية لمكونات المحصول في قمح الخبز باستخدام بعض الطرق الإحصائية

وفاء وهبه محمد* وعبد الله عبد المحسن سويلم ** وإيمان خليل عباس *

***المعمل المركزي لبحوث التصميم والتحليل الإحصائي**قسم بحوث القمح . معهد بحوث المحاصيل الحقلية- مركز البحوث الزراعية- الجيزة - مصر**

أجريت هذه الدراسة بمحطة البحوث الزراعية بكفر الحملم محافظةالشرقيه خلال موسمى ٢٠١٠/٢٠١١ و ٢٠١١/٢٠١٢ لتقييم المساهمة النسبية لأهم مكونات محصول الحبوب فى القمح كما أستخدم فى هذه الدراسة ثمانيه أصناف من قمح الخبز المصري وهى سخا ٩٣ وسخا ٩٤ وسدس ١ وجميزة ٧ وجميزة ٩ وجميزة ١٠ وسدس١٠ وجيزة ١٦٨ وأستخدم تصميم القطاعات الكاملة العشوائية فى ثلاث مكررات وكانت الصفات المدروسة هى عدد الأيام حتي طرد الاشطاء وارتفاع النبات وعدد الاشطاء للنبات وعدد السنابل وعدد حبوب السنبلة ووزن حبوب السنبلة ووزن الالف حبه ومحصول الحبوب للنبات.

وأجريت التحليلات الآتية: تحليل التباين- الارتباط البسيط وتحليل العامل ومعادلات التنبؤ بالمحصول باستخدام النموذج الكامل والمرحلي للانحدار المتعدد ومعامل المرور.

أظهرت الدراسة النتائج التالي

١- حقق الصنف سخا ٩٤ أعلى محصول حبوب للنبات وعدد الأفرع للنبات بينما كان الصنف جيزة ١٦٨ الأقل في محصول النبات في كلا الموسمين .

٢- دلت النتائج على وجود ارتباط عالي موجب ومعنوي بين محصول الحبوب للنبات وجميع الصفات تحت الدراسة.

٣- أظهرت نتائج تحليل العامل أن المكونات المدروسة تقع فى عاملين ويضم العامل الأول عدد السنابل للنبات وعدد حبوب السنبلة ووزن حبوب السنبلة ووزن الالف حبه ، بينما يضم العامل الثاني باقي الصفات المدروسة.

٤- كانت نتائج تحليل الانحدار المتعدد المرحلي والانحدار المتعدد ومعامل المرور متفقة على أن عدد السنابل للنبات وعدد حبوب السنبلة ووزن حبوب السنبلة ووزن الإلف حبه هي المكونات ذات الإسهامات الأكبر فى المحصول.

٥- أظهرت نتائج تحليل العامل أن التحليل يتسع لدراسة مزيد من المكونات حيث أن المكونات المدروسة لم تفسر سوى ٨١.٠٠% من التباين الكلى. وكانت نتائج تحليل العامل أكثر الطرق كفاءة وأنها تعطى معلومات أكبر وأسهل عن المتغيرات المتداخلة.

٦- أظهرت الدراسة أن طريقة الانحدار المتعدد المرحلي تعطى أفضل معادلة للتنبؤ بالمحصول مع إعطاء أهمية نسبية لكل متغير فى كل معادلة وأنه لا يوجد فرق بين قيمة معامل التحديد والخطأ القياسي فى كلا الطريقتين وهذه ميزة للإنحدار المتعدد المرحلي فى تحقيق أهمية المتغيرات والتى تخدم مربى النبات.

قام بتحكيم البحث

أ.د / عادل عبد الجواد سلامه — كلية الزراعة – جامعة المنصورة

أ.د / سليمان محمد جمعه — مركز البحوث الزراعيه