Article AR2431

DOI: HTTPS://DOI.ORG/10.23910/1.2021.2431

International Journal of Bio-resource and Stress Management

Crossref

August 2021

Print ISSN 0976-3988
Online ISSN 0976-4038

IJBSM 2021, 12(4):385-392

Research Article *Natural Resource Management*

Statistical Analysis of the Relationships Between Yield and Yield Components in Some Durum Wheat (*Triticum durum* desf.) Genotypes Growing under Semi-Arid Conditions

B. Frih[1*], A. Oulmi[1], A. Guendouz[2], H. Bendada[3] and S. Selloum[3]

[1]Dept. of Biology and Plant ecology, VRBN Laboratory, Farhat Abbas Setif University, Setif (19000), Algeria
[2]National Institute of Agronomic Research of Algeria (INRAA), Setif (19000), Algeria
[3]Technical Institute of Field Crops, Setif (19000), Algeria

Open Access

Corresponding Author

B. Frih
e-mail: benaliafrih@univ-setif.dz

Citation: Frih et al., 2021. Statistical Analysis of the Relationships Between Yield and Yield Components in Some Durum Wheat (*Triticum durum* desf.) Genotypes Growing under Semi-Arid Conditions. International Journal of Bio-resource and Stress Management 2021, 12(4), 385-392. HTTPS://DOI.ORG/10.23910/1.2021.2431.

Data Availability Statement: Legal restrictions are imposed on the public sharing of raw data. However, authors have full right to transfer or share the data in raw form upon request subject to either meeting the conditions of the original consents and the original research study. Further, access of data needs to meet whether the user complies with the ethical and legal obligations as data controllers to allow for secondary use of the data outside of the original study.

Conflict of interests: The authors have declared that no conflict of interest exists.

Abstract

The aim of this study was to know the relation between the grain yield and its components in order to identify the most important attributes in grain yield prediction which would serve as a criterion for the selection of genotypes growing under semi-arid conditions. Fifteen genotypes were studied composed of 11 advanced lines, 03 local landraces and 01 introduced genotype used as controls. Analysis of variance, simple linear correlation multiple linear regression, stepwise multiple linear regression, path analysis, principal component analysis and hierarchical clustering analysis were used to evaluate six traits including grain yield, plant height, days to heading, number of grains spike^{-1}, number of spikes m^{-2} and thousand kernels weight. Analysis of variance showed that genotype effect was significant for the majority of traits studied (p=0.001), advanced line G10 was shown to be most performing (4.723 t ha^{-1}). Simple linear regression revealed that the number of grains spike^{-1}, number of spikes m^{-2} and thousand kernels weight contributed significantly in grain yield changes (R^{2}=43, 17%). Path analysis showed that the number of grains spike^{-1}, number of spikes m^{-2} and thousand kernels weight had a direct and significant effect on grain yield. Principal component analysis showed that thousand kernels weight and negative days to heading were most important factors traits in grain yield. According to these results number of grains spike^{-1} and thousand kernels weight were crucial for the majority of static analysis.

Keywords: Algeria, durum wheat, grain yield, path analysis, semi-arid

1. Introduction

Durum wheat occupies an important place among the cereals in the world. Total food use of wheat is forecast to approach 518 mt, up 1.1% and rising in close tandem with world population growth. However, large supplies and competitive prices are likely to drive up feed use of wheat by 2.8%, a faster rate than was projected earlier, while industrial use is also anticipated to register strong growth (Anonymous, 2019). Climate changes recorded changes in the composition and geographic redistribution of ecosystems in Algeria. This situation has resulted in a shift towards the north of the arid zones, until then confined between the Sahara and the high cereal plains (Frih et al., 2021). In Algeria, the actual production

Article History

RECEIVED on 15th June 2021 RECEIVED in revised form on 10th August 2021 ACCEPTED in final form on 30th August 2021

of cereals during the period 2010–2017 is estimated at 4.12 million tonnes on average, an increase of 26% compared to the decade 2000-2009 when production is estimated on average at 3.26 million tonnes. Production consists mainly of durum wheat and barley, which respectively represent 51% and 29% of all cereal production on average 2010–2017 (Anonymous, 2018). Grain yield is a complex trait; it is determined by the interaction of a number of attributes (Singh and Diwivedi, 2002). Better knowledge of the genetic association of traits with yield is of great importance for breeders to develop an appropriate breeding strategy and to improve complex traits like grain yield which have shown little response to direct selection. Several studies have been carried out to provide insight into the factors responsible for the cereal yield of durum wheat in order to develop high yielding genotypes by indirect selection. Correlation and path analysis could be used as an important tool to bring information about appropriate cause and effects relationship between yield and some yield components (Khan et al., 2003). Path analysis showed direct and indirect effects of cause variables on effect variables. In this method, the correlation coefficient between two traits is separated into the components which measure the direct and indirect effects (Farshadfar, 2004; Zakizadeh et al., 2010). With the aid of stepwise regression analysis, ineffective or low-effective traits on yield can be omitted in the regression model, and traits which had significantly accounted for yield alterations can be evaluated .Moghaddam et al. (1998) found a negative correlation between plant height and grain yield of wheat, only the use of the correlation coefficient remains insufficient information on the relationship between grain yield and its components, a multivariate analysis remains essential. Leilah and Al-Khateeb (2005)reported that the number of spike, thousand kernels weight, grain weight spike^{-1} and biological yield were the most effective variables influencing grain yield. Guendouz et al. (2013a) indicated that the number of grain spike^{-1}, thousand kernels weight and biological yield had the most direct and positive effect on grain yield. Iftikhar et al. (2012) indicated that spike length, grains spike^{-1} and thousand kernels weight may be used as direct selection criteria to develop higher yielding varieties due to positive association and direct effects of these traits on the grain yield. The aim of this study is to determine the relationships between the grain yield and its components in order to identify the main attributes based on some statistical analysis which can serve as selection criteria for better durum wheat yielding genotypes under semi-arid conditions.

2. Materials and Methods

2.1. Study sites

The experimental material used in this study consisted of 15 genotypes (Table 1), which 03 were local landraces namely Boussellem, Boutaleb and Oued Bared and one introduced genotype Jupare C 2001, the objective was to test the adaptability of 11 foreign pedigrees and to estimate their yield in the local conditions. All genotypes studied were sown at December 2019 during the 2019/2020 season at Setif Agricultural Experimental Station (ITGC-AES, 36° 12’N and 05° 24’E and 1.081 m asl), Algeria, in a random block design with three replications. Each plot consisted of a row 2.5 m long spaced 20 cm between rows. The different agronomical traits were measured at maturity; grain yield (GY) (t ha^{-1}), plant height (PH) (cm), days to heading (DH) (day) from sowing date, number of spike m^{-2} (NS m^{-2}), number of grains spike^{-1} (NGS) and thousand kernel weight (TKW).

Table 1: Genotypes evaluated in this study

Genotypes	Varieties/pedigrees
G1	RASCON_37/GREEN_2/9/USDA595/3/D67.3/RABI//CRA/4/ALO/5/…
G2	MINIMUS_6/PLATA_16//IMMER/3/SOOTY_9/RASCON_37/9/…
G3	CMH77.774/CORM//SOOTY-9/RASCON-37/3/SOMAT-4
G4	CNDO/PRIMADUR//HAI-OU-17/3/SNITAN/4/SOMAT-3/
G5	RASCON_37/GREEN_2/9/USDA595/3/D67.3/RABI//CRA/4/ALO/5/…
G6	SILVER 14/MOEWE//BISU_l/PATKA_3/3/PORRON_4/YUAN_l/9/...
G7	GUANAY /HU ALITA / 10/PLATA _10/6/MQUE/4/USDA573/...
G8	BCRIS/BICUM//LLARETA INlA/3/DUKEM_12/2*RASCON 21/5/R
G9	Simeto/3/Sora/2*Plata_12//SRN_3/Nigris_4/5/Toska_26/...
G10	Ossl1/StjS5/5/Bicrcderaal/4/BEZAIZSHF//SD19539/Waha/3/St
G11	Stj3//Bcr/Lks4/3/Ter-3/4/Mgnl3/Aghrass2
G12	Jupare C 2001
G13	Boussellem
G14	Boutaleb
G15	Oued Bared

2.2. Statistical data analysis

All statistical analyses were performed by (Anonymous, 1998), (Anonymous, 2020).

2.2.1. Analysis of variance (ANOVA)

Analysis of variance (ANOVA) is a statistical tool used to detect differences between experimental group means. (Sawyer, 2009).

2.2.2. Simple linear correlation

A simple linear correlation was used when there is only one predictor variable, matrix of simple between grain yield and its components was computed according to the formula given by Snedecor and Cochran (1981).

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

r: correlation coefficient, x: first character, y: second character, n: total of number of observations.

2.2.3. Multiple linear regressions

A multiple linear regression was used when there is more than one predictor variable. This model was used to evaluate the relative contribution of related components to the grain yield variations by applying the following equation (Snedecor and Cochran, 1981).

$$Y=a+b_{1*}x_1+b_{2*}x_2+b_{3*}x_3+\ldots+b_{n*}x_n$$

Y: dependent variable, x: independent variables, a: intercept on the y axis, b: slope of the regression line.

2.2.4. Stepwise multiple regression linear

Stepwise regression was used in order to determine the most important variables (independent Variables) significantly contributed to total yield (dependent variable) variability (Draper and Smith, 1981).

2.2.5. Path analysis

Path analysis is an extension of the regression model, used to test the fitness of the correlation matrix against two or more causal models which are being compared by the researcher. The analysis was done following the method suggested by Dewey and Lu (1959).

$$r_{ij}=P_{ij}+\sum r_{ik}P_{kj}$$

r_{ij}:mutual association between independent character (i) and dependent character (j) as measured by the correlation coefficient, P_{ij}: Direct effects of independent character (i) and dependent character (j) as measured by path coefficient, $\sum r_{ik}P_{kj}$:Summation of components of an indirect effect of a given independent character (i) with all other dependent characters (k).

2.2.6. Principal components analysis (PCA)

Principal components analysis is a mathematical procedure used to classify a large number of variables (items) into major components and determine their contribution to the total variation. The first principal component is accounted for the highest variability in the data, and each succeeding component accounts for the highest remaining variability as possible (Everitt and Dunn, 1992).

2.2.7. Hierarchical clustering analysis (HCA)

Cluster analysis was used for arranging variables into different clusters to find the clusters that their cases within are more similar and correlated to one another comparing to other clusters. This procedure was performed using a measure of similarity levels and Euclidean distance (Everitt, 1993; Eisen et al., 1998).

3. Results and Discussion

3.1. Phenotypic variability

Phenotypic variability was observed in Table 2 composed of arithmetic means, standard deviations, interquartile ranges, minimums, 1st quartiles, median, 3rd quartiles and maximums recorded of the variables measured, grain yield varied from 3.012 to 4.737 t ha^{-1}, days to heading varied from 131 to 140 days, plant height varied from 81.88 to 107.22 cm and number of spikes m^{-2} varied from 296.66 to 408.33 spikes, the number of grains spike^{-1} varied between 31.33 and 49 grains, thousand kernel weight takes the values from 32.06 to 51.92 g.

Table 2: Descriptive statistics of variables measuring durum wheat genotypes

Variables	Basis statics							
	Mean	Sd	IQR	Min	1st Qu	Median	3rd Qu	Max
GY	3.886	0.607	0.872	2.545	3.479	3.827	4.351	4.933
DH	134.71	2.19	3.00	131.0	133.0	135.0	136.0	140.0
HP	90.01	7.35	8.00	76.67	85.33	90.33	93.33	112.33
NGS	37.98	5.66	8.00	28.00	34.00	37.00	42.00	53.00
NS m^2	349.78	53.37	70.00	270.0	310.0	350.0	380.0	510.0
TKW	42.00	5.59	7.28	31.76	37.48	43.32	44.76	52.08

Sd: Standard deviation; IQR: Interquartles range; 1st Qu: First quartile; 3rd Qu: Third quartile

3.2. Analysis of variance (ANOVA)

The results of the analysis of variance (Table 3) showed that the genotype effect was very significant (p<0.001) for the majority of the variables studied excepted for number of spikes m^{-2} (NS m^{-2}) which suggested the presence of big variation among genotypes studied. For (GY) trait, a high significant difference was showed between advanced line G10 and the author genotypes. For non significant traits (NS m^{-2}) this may be due to low variability among genotypes for this trait. Mansouri et al. (2018) reported that genotype effects were significant with spike number, days to heading and plant height. The advanced line G10, which was the best yielding genotype (GY=4.723 t ha^{-1}), was characterized by the shorter duration of days to heading (DH=132 days), a more or less average size (PH=88.55 cm), a number of spikes m^{-2} which

Table 3: The results of the analysis of variance of the genotypes compared to the variables measured

Genotypes	Variables					
	GY	DH	PH	NS m^{-2}	NGS	TKW
G1	4.05^{abc}	135.33^{cd}	89.77^{cd}	360^{abc}	35.33^{de}	47.33^{bc}
G2	3.558^{cd}	133^{efg}	87.55^{cde}	353.33^{abc}	34.66^{de}	42.21^{def}
G3	4.065^{abc}	138^{a}	86.44^{cde}	296.67^{c}	44^{ab}	45.1^{bcd}
G4	4.204^{abc}	136^{bc}	87.55^{cde}	325^{bc}	33.33^{de}	51.92^{a}
G5	4.481^{ab}	135^{cd}	87.55^{cde}	351.67^{abc}	46.33^{a}	36.26^{gh}
G6	3.865^{abcd}	137^{ab}	86.88^{cde}	350^{abc}	49^{a}	32.06^{h}
G7	3.996^{abc}	136^{bc}	84de	406.67^{ab}	41^{bc}	34.78^{gh}
G8	3.012^{d}	134.33^{de}	81.88^{e}	408.33^{a}	33.33^{de}	38.75^{efg}
G9	4.009^{abc}	132.33^{g}	91.88^{bc}	346.67^{abc}	36^{cde}	42.36^{def}
G10	4.737^{a+}	132^{g}	88.55^{cde}	335^{abc}	40.66^{bc}	42.56^{de}
G11	3.786^{bcd}	132.67^{fg}	91.88^{bc}	355^{abc}	37^{cd}	43.47^{cd}
Introduced genotype						
G12 (Jupare C 2001)	3.476^{cd}	132.33^{g}	90^{bcd}	356.67^{abc}	38.33^{cd}	38.18^{fg}
Local landraces						
G13 (Boussellem)	3.592^{bcd}	134^{def}	97^{b}	303.33^{c}	34.33^{de}	43.42^{cd}
G14 (Boutaleb)	3.706^{bcd}	138^{a}	107.22^{a}	336.67^{abc}	31.33^{e}	48.36^{ab}
G15 (Oued Bared)	3.754^{bcd}	134.67^{cd}	92^{bc}	361.66^{abc}	35de	43.22^{cd}
Mean	3.886	134.71	90.01	349.77	37.97	42
Min	3.012	132	81.88	296.66	31.33	32.07
Max	4.737	138	107.22	408.33	49	51.92
LSD (p=0.05)	0.90898	0.70330	7.13600	82.28845	5.23793	4.33975
Genotype effect	Ns	***	***	Ns	***	***

***: p=0.001; Ns: Non-significant; +: significant difference

approaches the average (NS m^{-2}=135 spikes m^{-2}), an average number of grains spike^{-1} (NGS=40.66 grains) and a thousand kernels weight more or less important (TKW=42.56 g). The advanced line G5 which recorded a good score of grain yield (GY=4.481 t ha^{-1}) stayed 03 days more to heading (DH=135 days), its size was about the same (PH=87.55 cm), it had 15 more spikes m^{-2}, its number of kernels spike^{-1} was greater than 6 kernels but its thousand kernels weight was lower by 6 grams this in comparison with the most powerful advanced line G10. The advanced line G4 characterized by the highest thousand kernels weight achieved the 3rd best score in terms of grain yield (GY=4.204 t ha^{-1}). From these observations we can say that a good grain yield may be due to a minimum duration of days to (DH<136 days) heading, a more or less small size (PH<89 cm) and a good thousand kernels weight (TKW>36). Guendouz et al. (2013b) discussed the positive and significant relationship between grain yield and thousand kernels weight (TKW) and the negative relationship of grain yield with days to heading (DH) under semi-arid conditions.

3.3. Simple linear correlation

Our results of the correlation test illustrated in Table 4 showed that the number of grains spike^{-1} was positively and significantly correlated with grain yield (r=0.33 at p=0.05), Previous authors reported similar results between the grain yield, spikes number and number of grains spike^{-1} (Aycicek and Yildirim, 2006), however number of grains spike^{-1} was negatively and significantly correlated with thousand kernel weight (r=-0.51 at p=0.05) and plant height (r=-0.33 at p=0.05), a positive and significant correlation between thousand

Table 4: Matrix of correlations between the different measured traits

Variables	GY	DH	PH	NGS	NS m^{-2}	TKW
GY	1					
DH	-0.05	1				
PH	0.11	0.06	1			
NGS	0.33*	0.18	-0.33*	1		
NS m^{-2}	0.23	-0.15	-0.02	-0.07	1	
TKW	0.18	0.05	0.38*	-0.51*	-0.15	1

*: (p=0.05)

kernel weight and plant height was recorded (r=0.38 at *p*=0.05), similar results were found by Fellahi et al. (2013) and Hannachi et al. (2013). These results indicated that the simple correlation was insufficient to give clear information concerning the relations between grain yield and these components.

3.4. Multiple linear regression analysis

Table 5 presents the regression coefficients of the estimated variables to predicting the grain yield. The dependent variable grain yield could be predicted from a linear combination of the independent variables. The obtained results showed that the prediction model equation for grain yield was formulated as follows:

$GY=21.581-(0.45088\times DH)+(0,126\times PH)+(0,754\times NGS)+(0.039\times NS\ m^{-2})+(0,580\times TKW)$

GY: grain yield; NS m^{-2}: number of spikes m^{-2}; NGS: number of grains by Spike; TKW: thousand kernel weight; PH: plant Height; DH: day to heading.

This model could justify significantly more than 43% (R^2=0.4317) changes in the performance of the grain yield, while the remaining 57% perhaps was due to residual effects. The multiple linear regression analysis for the variables tested revealed that the number of grains spike^{-1}, number of spikes m^{-2} and thousand kernel weight contributed significantly to the grain yield.

Table 5: Regression coefficients of the estimated variables to predicting the grain yield (GY)

	Estimates	Std. error	t value	Pr(>|t|)
Intercept	21.58153	47.04855	0.459	0.648990
DH	-0.45088	0.34823	-1.295	0.203015
PH	0.12625	0.10988	1.149	0.257574
NGS	0.75414	0.15876	4.750	0.0000275***
NS m^{-2}	0.03903	0.01419	2.752	0.008950**
TKW	0.58004	0.16212	3.578	0.000945***

***: *p*=0.001; **: *p*=0.01/R=0,6570; R^2: 0.4317; AdjR^2: 0.3589

3.5. Stepwise multiple linear regression

In order to remove effect of non-effective characteristics in regression model on grain yield, stepwise regression was used. In stepwise regression analysis (Table 6), the grain yield as the dependent variable and other traits as independent variables were considered. The results showed that the number of spikes m^{-2}, grains number spike^{-1}, thousand kernel weight with R^2=39.17%, had justified the grain yield changes (Table 7). Due to their low relative contributions, the other variables were not included in the model. Consequently, based on the final step of stepwise regression analyses, the equation for the prediction of grain yield could be obtained as:

$GY=-26.594+(0,\ 0410\times NS\ m^{-2})+(0,680\times NGS)+(0,599\times TKW)$

Table 6: The variables entred by stepwise regression

Step	Variables entred	F to enter	P to enter
1	NGS	5.216*	0.027*
2	TKW	9.108*	0.004*
3	NS m^{-2}	8.399*	0.006*

*: *p*=0.05

Table 7: Multiple linear regression of entred variables by stepwise

	Estimate	Std. error	t value	Pr(>|t|)
Intercept	-26.594	12.93644	-2.056	0.046213*
NSm^{-2}	0.041	0.01423	2.898	0.006000**
NGS	0.680	0.15373	4.426	0.000069***
TKW	0.599	0.15689	3.824	0.000439***

***: *p*=0.001; **: *p*=0.01; *: *p*=0.05; R: 0,6259; R^2: 0.3917; AdjR^2: 0.3472

This result was very consistent with the work of Mohammadi et al. (2011) who founded that 75% of the variation in grain yield was explained by the number of spikes m^{-2}, thousand kernel weight and plant height.

3.6. Path analysis

Path analysis is a form of multiple regression statistical analysis that was used to evaluate causal models by examining the relationships between a dependent variable and two or more independent variables. The results of the path analysis (Table 8) showed that the number of grains spike^{-1}, the number of spikes m^{-2} and thousand kernel weight had a direct and significant effect on the grain yield (0.703, 0.343, 0.534, respectively). The positive direct effects of spikes number, grains number and thousand-kernel weight were previously reported in wheat (Shamsi et al., 2011; Iftikhar et al., 2012).

Table 8: Direct effects of estimated variables for grain yield (GY)

variables	Direct effect for (GY)
DH	-0.162747
PH	0.152957
NGS	0.703020
NS m^{-2}	0.343194
TKW	0.534576

3.7. Principal component analysis (PCA)

Principal component analysis PCA reflects the importance of the largest contributor to total variation at each axis of differentiation (Sharma, 1998). The choice of interpretation axes depends on the importance of their eigenvalues, Kaiser's rule is based on a simple idea, in a standardized PCA, the sum of the eigenvalues being equal to the number of variables, and their mean is equal to 1. We let us, therefore, considered that an axis was interesting if its eigenvalue was greater than 1. The data presented in Table 9 showed the eigenvalues of the six components, the first three components were the most important; they alone accumulate 82.33% of information on variability. Table 10 and Figure 1 showed the correlation of the first three components with our variables. The first component PC1 was strongly and positively correlated with plant height and thousand kernels weight (r=0.779; r=0.894) and negatively correlated with the number of grains spike^{-1} (r=-0.716). PC2 was strongly and positively correlated with the grain yield (R=0.807) andnegatively correlated with the number of spikes m^{-2} (r=-0.648), PC3 was strongly correlated with days to heading (r=0.862). Table 10 and Figure 1 also showed that genotypes G1, G4, G13 (Boussallem) and G14 (Boutaleb) were positively connecting to PC1, these genotypes were characterized by high thousand kernels weight, unlike G5, G6, G7 which were negatively related to PC1 and which were characterized by their low thousand kernels weight. Genotype G3 witch had the lowest (NS m^{-2}) was positively connected to PC2 (NS m^{-2}) unlike genotypes G2, G12 (Jupare C 2001) and G15 (Oued Bared) characterized by similar low (GY) and G8 characterized by the highest (NS m^{-2}) which were negatively connected to PC2. Genotypes G9, G10, G11 were logically and negatively connected with PC3 because of their short duration of days to heading. These results were consistent with our ANOVA results and with the work of Guendouz et al. (2013b) which predicts that the TKW was an important factor in the prediction of GY and that the long duration of DH had a negative effect on the GY.

Table 9: Eigen values for all principal components

	PC1	PC2	PC3	PC4	PC5	PC6
Var	2.272	1.682	0.986	0.606	0.426	0.028
Var %	37.872	28.027	16.435	10.103	7.096	0.467
CV	37.872	65.899	82.334	92.437	99.533	100.000

Var: Variance; Var %: % of var; CV: Cumulative % of var

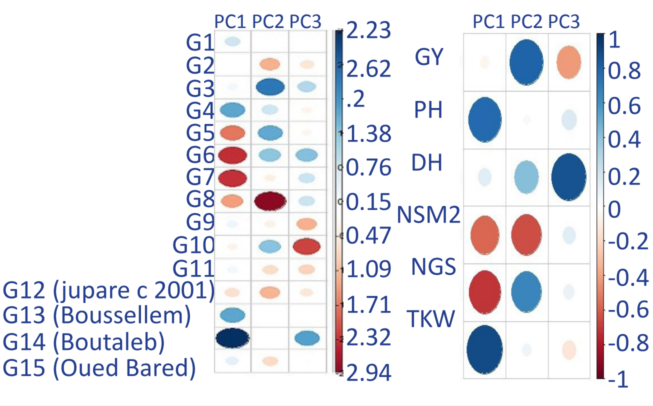

Figure 1: Visualization of the relations of genotypes and traits studied with the 3 first's components

Table 10: Correlations of variables and coordinates of genotypes in first's 3 components

	PC1	PC2	PC3
Variables			
GY	-0.055	0.807©	-0.422
PH	0.779©	0.036	0.156
DH	0.124	0.424	0.862©
NS m^{-2}	-0.579	-0.648©	0.121
NGS	-0.716©	0.652	0.086
TKW	0.894©	0.062	-0.134
Genotypes			
G1	0.726©	-0.034	-0.047
G2	0.030	-1.195©	-0.532
G3	0.284	2.345©	0.982
G4	1.720©	0.741	-0.272
G5	-1.743©	1.706	-0.259
G6	-2.383©	1.346	1.410
G7	-2.359©	-0.374	0.773
G8	-1.366	-2.940©	0.775
G9	0.326	-0.342	-1.195©
G10	-0.317	1.375	-2.170©
G11	0.284	-0.660	-0.786©
G12 (jupare c 2001)	-0.631	-1.147©	-0.494
G13 (Boussellem)	1.738©	-0.102	-0.110
G14 (Boutaleb)	3.233©	-0.029	1.785
G15 (Oued Bared)	0.457	-0.690©	0.140

©: connected to component

The relations of measured traits and genotypes tested with 3 first's components are graphically summarized in Figure 2.

3.8. *Hierarchical clustering analysis (AHC)*

3.8.1. *For traits*

The six variables examined were separated into 03 groups (Figure 3). The first group consisted of a number of spikes m^{-2}; group-II included thousand kernels weight, number of grains spike^{-1} and the grain yield and group-III consisted by plant height, days to heading. Our data reflected the tendency of each variable group in a cluster to be closely related to each other. Therefore, the results of the study showed that the number of grains spike^{-1} and thousand kernels weight were the traits most related to the grain yield.

3.8.2. *For genotypes*

Figure 4 showed that genotypes tested were separated into 03

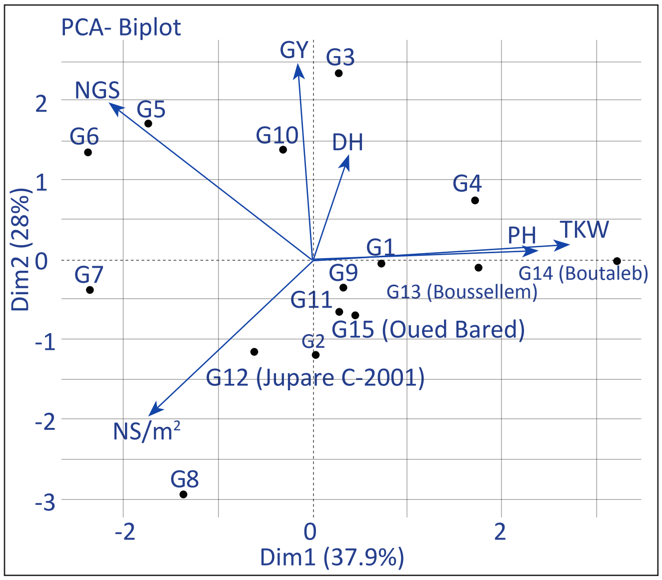

Figure 2: Biplot of genotypes and traits studied with 3 first's components

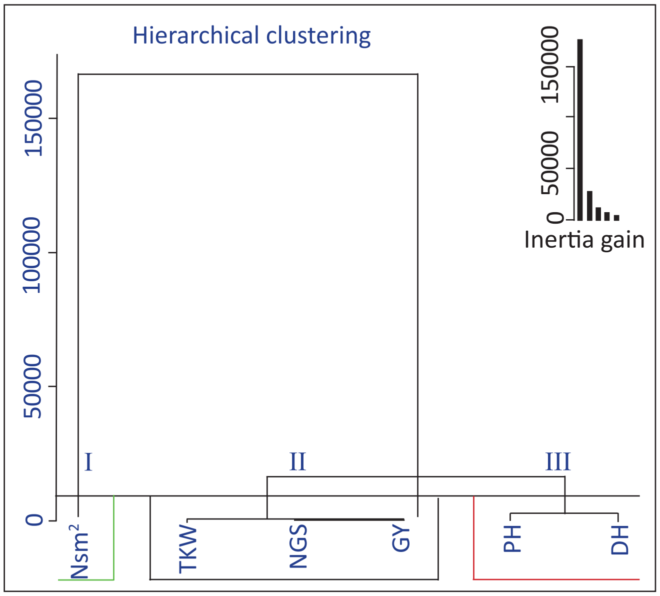

Figure 3: Similarity levels of the estimated variables using Hierarchical clustering analysis

groups, the group-I composed from advanced lines G7, G8 for their very highs number of spikes m^{-2}. Advanced line G3 follow G13 (Boussallem) were regrouped in group-II, its characterized by their very lower number of spikes m^{-2}. Group-III was formed by 02 sub-groups; sub-group A composed of advanced lines G4, G10 which follow the local land race G14 (Boutaleb) this sub-group was characterized by a more or less low number of spikes m^{-2}, the subgroup B formed by the advanced lines G1, G2, G5, G6, G9, G11 which followed the introduced genotype G12 (Jupare-C2001) and the local landrace (Oued Bared) for its number of spikes m^{-2} more or less plant height (Table 11).

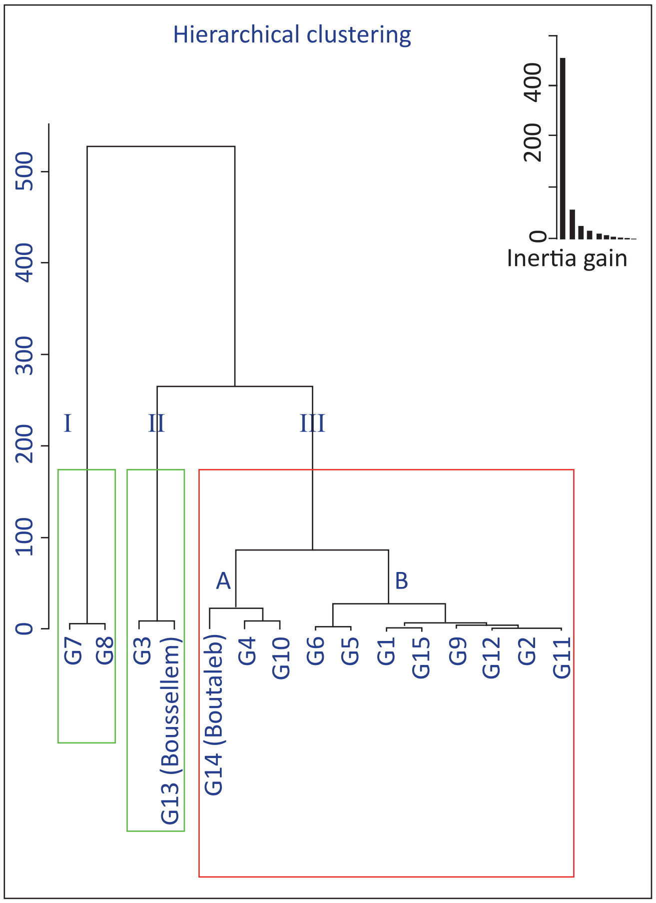

Figure 4: Similarity levels of genotypes tested using Hierarchical clustering analysis

Table 11: The most important characteristics of durum wheat by each statistical method

Variables	Statistical analysis							
	ANOVA	SC	MLR	SR	PA	PCA	CAH	Total of scores
DH	*					*		2
PH	*							2
NS m^{-2}			*	*	*			3
NGS	*	*	*	*	*		*	6
TKW	*		*	*	*	*	*	6

SC: Simple correlation; MLR: Multiple linear regression; SR: Stepwise regression; PA: Path analysis; PCA: Principal components analysis; HCA: Hierarchical clustering analysis

4. Conclusion

ANOVA showed that the genotype effect was significant for the majority of traits studied, G10 was shown to be the most performing. MLR revealed that NGS, NS m^{-2} and TKW contributed significantly to the GY changes. PA showed that NGS, NS m^{-2} and TKW had a direct and significant effect on GY. PCA showed that TKW and negatively DH were the most important factors traits in GY. On the basis of this results, NGS and TKW were the best selection criteria for better grain yield under semi-arid conditions.

5. References

Anonymous, 2019. Food outlook - Biannual Report on Global Food Markets, FAO–November 2019. Rome. Accessed on Tuesday, September 1, 2020.

Anonymous, 2018. Statistiques serie B-Ministere de l'agriculture et du developpement rural. MARD. Accessed on Saturday, September 12, 2020.

Anonymous, 1998. Costat 6.400, 1998. Copyright © 1998–2008, CoHort Software798 Lighthouse Ave BMP 320, Montery, CA 93940, USA, Available at: http://www.cohort.com. Accessed on Sunday, November 15, 2020.

Anonymous, 2020. R Core team. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. Available at: http://www.R-project.org. Accessed on Saturday , December 05, 2020.

Aycicek, M., Yildirim, T., 2006. Path coefficient analysis of yield and yield components in bread wheat (*Triticum aestivum* L.) genotypes. Pakistan Journal of Botany 38, 417–424.

Dewey, D.R., Lu, K.H., 1959. A correlation and path coefficient analysis of components of crested wheat grass seedproduction. Agronomy Journal 51, 515–518.

Draper, N.R., Smith, H., 1981. Applied regression analysis. 2nd edition. Wiley series in probability and mathematical statistics. John Wiley & Sons. N.Y., 709.

Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D., 1998. Cluster analysis and display of genome-wide expression ilatteriis. Proceedings of the National Academy of Science USA 95, 14863–14868.

Everitt, B.S., 1993. Cluster analysis. Wiley, New York, NY.

Everitt, B.S., Dunn, G., 1992. Applied multivariate data analysis. Oxford University Press, New York, NY.

Farshadfar, E., 2004. Multivariate principles and procedures of statistics. Taghbostan Pub. Kermanshah, Iran, 734.

Frih, B., Oulmi, A., Ali, G., 2021. 2021. Study of the drought tolerance of some durum wheat (*Triticum durum* Desf.) genotypes growing under semi-arid conditions in Algeria. International Journal of Bio-resource and Stress Management 12(2), 137–141. https://doi.org/10.23910/1.2021.2171a.

Guendouz, A., Djoudi, M., Guessoum, S., Maamri, K., Fellahi, Z., Hannachi, A., Hafsi, M., 2013a. Durum wheat (*Triticum durum* Desf.) evaluation under semi arid conditions in eastern Algeria by path analysis. Journal of Agriculture and Sustainability 3(1), 56–64.

Guendouz, A., Guessoum, S., Maamri, K., Benidir, M., Hafsi, M., 2013b. Performance of ten durum wheat (*Triticum durum* desf.) cultivars under semi arid conditions (North Africa-Algeria). Indian Journal of Agricultural Research 47(4), 317–322.

Hannachi, A., Fellahi, Z., Bouzerzour, H., Boutekrabt, A., 2013. Correlation, path analysis and stepwise regression in durum wheat (*Triticum durum* Desf.) under rainfed conditions. Journal of Agriculture and Sustainability 3(11), 122–131.

Iftikhar, R., Khaliq, I., Ijaz, M., Rashid, M.A., 2012. Association analysis of grain yield and its components in spring wheat (*Triticum aestivum* L.). American-Eurasian Journal of Agricultural and Environmental Sciences 12(3), 389–392.

Khan, A.S., Ashfaq, M., Asad, M.A., 2003. A correlation and path coefficient analysis for some yield components in bread wheat. Asian Journal of Plant Sciences 2(8), 582–584.

Leilah, A.A., Al–Khateeb, S.A., 2005. Statistical analysis of wheat yield under drought conditions. Journal of Arid Environments 61(3), 483–496.

Mansouri, A., Oudjehih, B., Benbelkacem, A., Fellahi, Z., Bouzerzour, H., 2018. Variation and relationships among agronomic traits in durum wheat [*Triticum turgidum* (L.) Thell. ssp. turgidum conv. durum (Desf.) MacKey] under South Mediterranean growth conditions: stepwise and path analyses. International Journal of Agronomy, 1–11.

Moghaddam, M., Ehdaie, B., Waines, J.G., 1998. Genetic variation for and interrelationships among agronomic traits in landraces of bread wheat from south western Iran. Indian Journal of Genetics and Plant Breeding 52(1), 73–81.

Mohammadi, M., Karimizadeh, K., Kazem Shafazadeh, M., Sadeghzadeh, B., 2011. Statistical analysis of durum wheat yield under semi-warm dryland condition. Australian Journal of Crop Science 5(10), 1292–1297.

Sawyer-Steven, F., 2009. Analysis of variance: the fundamental concepts. Journal of manual & Manipulative Therapy 17(2), 27E–38E.

Singh, S.P., Diwivedi, V.K., 2002. Character association and path analysis in wheat (*Triticum aestivum* L.). Agricultural Sciences Digest 22(4), 255–257.

Shamsi, K., Petrosyan, M., Noor-Mohammadi, G., Haghparast, A., Kobraee, S., Rasekhi, B., 2011. Differential agronomic responses of bread wheat cultivars to drought stress in the west of Iran. African Journal of Biotechnology 10(14), 2708–2715.

Sharma, J.R., 1998. Stastical and biometrical techniques in plant breeding new age. International (P) Limited Publishers, New Delhi, 432.

Snedecor, G.W., Cochran, W.G., 1981.Statistical Methods, seventhed. Iowa State University Press, Iowa, USA.

Zakizadeh, M., Esmaeilzadeh, M., Kahrizi, D., 2010. Study on genetic variation and relationship between plant characteristics and grain yield in long spike bread wheat (*Triticum aestivum* L.) genotypes using multivariate analysis. Iranian Journal of Crop Sciences 12(2), 18–30.