



# Statistical methods for identifying traits associated with high yield potential in durum wheat under drought conditions

Reza Mohammadi\*, Alireza Etminan<sup>1</sup> and Lia Shooshtari<sup>1</sup>

Dryland Agricultural Research Institute, Sararood branch, Agricultural Research, Education and Extension Organization (AREEO), Kermanshah, Iran; <sup>1</sup>Department of Plant Breeding, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran

(Received: December 2016; Revised: August 2017; Accepted: September 2017)

## Abstract

The knowledge of key traits contributing to stress tolerance could contribute to increasing the efficiency of the selection process leading to development of drought tolerant genotypes. Over three years (2010-13), 25 durum wheat genotypes were evaluated for different agro-physiological traits under drought conditions. Four statistical procedures including path analysis, stepwise regression, principal component analysis (PCA) and factor analysis (FA) were used to identify the most contributors to grain yield. A mean score index (MSI) based on scaling scores of selection criteria was used for genotypes characterization. The average yield productivity varied between 774 to 2360 kg/ha across years. The statistical procedures confirmed the chlorophyll fluorescence (Fv/Fm), spike length, SPAD reading, plant height, peduncle length and heading date as the most contributors to yield productivity in durum wheat. The methodology of scoring scale provided a simple and easy visualization and identification of resilient, productive and/or contrasting genotypes according to selection criteria. The PCA and FA by justifying the high portion of variability in yield were found to be more efficient for developing proper models for indirect selection.

**Key words:** Drought stress, durum wheat, grain yield, statistical procedures, trait selection

## Introduction

Developing high yielding wheat cultivars under drought conditions in arid and semi-arid regions is an important objective of breeding programs. Durum wheat (*Triticum turgidum* L. var. durum Desf.) is grown on 10% of the world's wheat area. It occupies about 11 million ha in the Mediterranean basin. Terminal drought stress constrains wheat production in rainfed regions of the world including Iran. However, grain yield improvement is the major objective of wheat improvement programs

in those regions. Rainfed wheat covers two-thirds of Iran's total wheat area, but accounts for only about one-third of total wheat production (Mohammadi et al. 2013). Developing plants with suitable advantages under drought conditions is a basic challenge for wheat improvement programs (Trenberth 2011; Staniak and Kocon 2015). Grain yield is a complex quantitative trait that results to the actions and interactions of various component traits (Singh and Diwivedi 2002). However, developing drought-tolerant varieties is an important objective of breeding programs and is expected to play a crucial role in climate change mitigation strategies (Gustafson 2011). Thus, the knowledge of traits associated with drought tolerance would be useful for breeding materials in drought prone-environments (Girdthai et al. 2009; Mir et al. 2012; Mohammadi et al. 2013).

However, appropriate traits used to develop proper models for indirect selection should have significant genetic variability. Different statistical techniques have been used in modeling crops yield. The information on the nature and magnitude of correlation coefficients help breeders to determine the selection criteria for simultaneous improvement of various characters along with yield. However, simple correlation analysis that relates grain yield to a single variable may not provide a complete understanding of the importance of each component in determining grain yield (Dewey and Lu, 1959; Singh et al. 1979). Partitioning the correlation coefficient into direct and indirect effects can be done through path analysis technique (Dewey and Lu, 1959). Many reports have used this technique on wheat (Ehdaei and Waines

\*Corresponding author's e-mail: r.mohammadi@areeo.ac.ir

1989), soybeans (Leilah et al. 1988), rapessed (Huang et al., 2015) and in chickpea (Canci and Toker 2009). Stepwise regression analysis proved to be efficient to determine the predictive equation for yield (Leilah and Al-Khateeb 2005; Mohammadi et al. 2013). Factor analysis suggested by Walton (1972) has been widely used to identify growth and plant traits related to wheat (Moghaddam et al. 1998; Canci and Toker 2009). The principal components analysis (PCA) is a multivariate statistical technique for exploration and simplifying complex data sets. Leilah and Al-Khateeb (2005) used seven different statistical procedures to study the relationship between wheat grain yield and its components.

Identification of genotypic variability for agronomical and physiological traits under drought conditions is of great interest for breeders because selected genotypes with favorable traits can be used as parents in future crosses. Therefore, this study was conducted to show the relationship between grain yield and different plant traits under drought conditions of Iran.

## Materials and methods

### *Plant materials and experimental layout*

Plant genetic material were diverse and consisted of 25 wheat genotypes including 21 durum breeding lines (G1-G21), one new durum cultivar (G22, "Saji" cultivar), two old durum varieties (G23, "Zardak" cultivar; and G24, "Gerdish" cultivar) and one popular bread wheat variety (G25, "Sardari" cultivar). The genotypes evaluated under rainfed conditions during three cropping seasons (2010-13). The experiments were conducted at Dryland Agricultural Research Institute (Sararood station, Kermanshah, Iran; 34°19'N; 47°17'E; 1351 m above sea level). The site representative for moderate winter cold rainfed areas in durum wheat breeding program. The average crop season rainfall of the experimental site is 425 mm with minimum and maximum temperatures of -20 and 45°C, respectively, and 60-100 days of freezing temperatures annually. The soil at the site was clay loam. At each cropping season, experimental layout was a randomized complete block design with three replications. Management practices recommended for each trial were followed in the all yield trials.

### *Measurement of plant traits*

Chlorophyll fluorescence was measured using a fluorometer (OS30, Opti-Science, Hudson, NH, USA)

between 11:00 and 14:00 hours. The clips were placed on the flag leaf and closed to prevent any light from entering into the clipped spot and the clips were left there for at least 30 min. Following dark adaptation, readings were taken by inserting the fluorometer tip, opening the clip shutter and then giving a flash of light from the fluorometer that activated the Photo system-II reaction centers of the photosynthetic apparatus (Ristic et al. 2007). Stomatal conductance (SC) was measured by using a leaf porometer (Decagon Devices, Inc., Pullman, WA, USA). Three random plants were selected in each plot for determining gas exchange parameters. All measurements were made on the flag leaf.

Canopy temperature (CT) was measured using a handheld infrared instrument (E200IR, Germany). Three measurements were taken per plot at approximately 0.5 m from the edge of the plot with an approximately 45° from the horizontal position. Canopy temperatures were measured between 12:00 to 14:00 hours on a clear sunny day.

The other physiological traits measured were including relative water loss (RWL, Yang et al. 1991), relative water content (RWC, Barrs 1968) and relative growth rate (RGR, Hoffmann and Poorter 2002).

At maturity, the agronomic traits recorded for each genotype were plant height (PH), peduncle length (PL), flag-leaf length (FL), spike length (SL), days to heading (DH), days to maturity (DM), grain yield (YLD), 1000-kernel weight (TKW) and number of grain per spike (NGPS). Days to heading was designated as the time when 50% of the plants in a plot had at least one open flower. Days to maturity was recorded when 50% of the plants in a plot had yellow leaves. The PH, PL, FL, SL, and NGPS were measured based on five randomly samples for each genotype at physiological maturity. After harvest, the TKW was recorded based on weight of 1000 grains for each genotype. The plot yields were converted to productivity per hectare (kg/ha) and subjected to statistical analyses.

### *Data analyses*

Analysis of variance was carried out using the MSTAT-C statistical program. For each trait, the data of 25 genotypes grown across three years were subject to combined analysis of variance to partition trait variation into year (Y), genotypes (G), and GxY interaction effects. In multi-year trials with  $m$  genotype and  $n$  year, the combined ANOVA of multi-year data is based on the following equation:

$$X_{ijk} = \mu + g_i + y_j + (gy)_{ij} + b_{jk} + e_{ijk}$$

where  $X_{ijk}$  is the phenotypic value of the  $i$ th genotype in the  $k$ th replicate in the  $j$ th year;  $\mu$  is mean of all genotypes over all years;  $g_i$  is the effect of the  $i$ th genotype as fixed factor,  $i=1, 2, \dots, m$ ;  $y_j$  is the effect of the  $j$ th year as random factor,  $j=1, 2, \dots, n$ ;  $(gy)_{ij}$  is effect of the interaction between  $i$ th genotype and the  $j$ th year,  $b_{jk}$  is the effect of the  $k$ th replicate in the  $j$ th year,  $k = 1, 2, \dots, p$ ; and  $e_{ijk}$  is random error deviate on the  $i$ th genotype in the  $k$ th replicate in the  $j$ th year.

Broad-sense heritability for yield and each plant traits were estimated to take account of the more stable traits as: genotypic variance/phenotypic variance:  $h_b^2 = \sigma_g^2 / \sigma_p^2$ , the phenotypic variance was calculated as follow (Nyquist 1991):

$$\sigma_p^2 = \sigma_g^2 + (\sigma_{gy/y}^2) + (\sigma_{e/ry}^2)$$

where  $y$ ,  $g$  and  $r$  are represent for year, genotype, and replication, respectively, and  $\sigma_g^2$  and  $\sigma_e^2$  are the components of variance for genotypes and error, respectively. Path coefficient analysis developed by Wright (1921) and applied by Dewey and Lu (1959), was used to partitioning the correlation coefficients to their direct and indirect effects through other traits. Stepwise multiple linear regression analysis was used according to Montgomery (2006) to develop the prediction model for grain yield and to determine the variables accounting for the majority of total grain yield variability. Principal component analysis (PCA) was used to classify variables into major components and their total variation. In this model, the first principal component accounting for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible (Everitt and Dunn1992).

The factor analysis method (Cattell 1965) consisted of the reduction of a large number of correlated variables to a much smaller number of clusters of variables called factors. Then the matrix of factor loading was submitted to a varimax orthogonal rotation (Kaiser 1958). Thus, factor analysis indicates both groupings and contribution percentage to total variation in the dependence structure. The array of communality, the amount of variance of a variable accounted by the common factors together, was estimated by the highest correlation coefficient in each array as suggested by Seiller and Stafford (1985).

### Scoring scale for selection criteria

The scoring scale for each trait was calculated as described by Thiry et al. (2016). For each trait the minimum and maximum values among genotypes were identified. The difference of these two values gives the range of the scale for each trait. This range is divided into ten parts and each part has a score from 1 to 10. Therefore, each part represents 10%, 20%, ... or 100% of the range value. However generally some traits i.e., grain yield their high values are desirable (Class 1), while some traits i.e., phenological traits their low values are desirable (Class 2). For class 2 of traits, we have inverted the value of traits, so a high value obtained with the original equation will receive a lower score. This allows the two classes of traits to have the same scale, where a high score will always mean a 'good' genotype. For example, a score value of 2 is assigned for the traits in Class 1 for all the values within 10-20% of the range and 80-90% of the range for traits in Class 2. A tool developed in Microsoft Excel has been created to assign a score to each value.

A further combination of the score traits was named as mean score index (MSI):

$$MSI = \frac{X_{1s} + X_{2s} + X_{3s} + \dots + X_{ns}}{N}$$

where  $X_{is}$  is the scoring scale for  $i$ th selection criterion and  $N$  is number of selection criteria.

Prior to pooling of data over years, the homogeneity of error variances was checked. Then statistical analyses were done using the pooled data of three years for physiological and agronomical traits evaluated under drought in three growing seasons, using IBM SPSS Statistics 19.

## Results

### Weather condition

Rainfall distribution pattern remarkably varied among cropping seasons. However, rainfall was contrasting in the cropping seasons, and hence, the genotypes were exposed to drought stress, which is the most limiting factor in moderate cold rainfed areas of Iran. The all three seasons were characterized by lower rainfall levels (342.5, 302.9 and 394.3 mm, respectively in 2010-11, 2011-12 and 2012-13 seasons) than the average long-term rainfall in the station (425 mm rainfall), where the crops experienced severe droughts

with remarkable respective decreases in rainfall of 82.5, 122.1 and 30.7 mm relative to average long-term rainfall. 2012-13 cropping season with rainfall amount of 394.3 mm was relatively close to average long-term of the station. No marked variation in temperatures was observed across cropping seasons, although the 2012-13 season with higher average temperature was warmer than other two seasons in winter.

#### **Variance components, variability and heritability of traits**

The results of combined ANOVA for grain yield, phenological, physiological and agronomic traits are given in Table 1. ANOVA indicates that the year effect was significant for all traits. The genotype effect for most traits was significant, but not for relative water loss, relative growth rate, chlorophyll fluorescence and number of grain per spike. Genotype and year interactions were significant for all traits except for CTD.

Estimates of variability for the studied traits within and between years are given in Table 2. The traits with the least changes in quantity across years were Fv/Fm followed by RWC, SL, DH and DM, while those with the highest changes in quantity were PL followed by YLD, RGR, CTD and SC. The results also revealed that the trait values from one year to another were not constant and showed large fluctuations. Some traits showed positive response to a season and negative to another season. The traits including TKW, RWL, RWC, SC and PL showed the highest performance in 2012-13 season (severe drought), while the traits with the least performance were PH, YLD, RGR, and FL. The traits with the best performance in 2010-11 (moderate drought) were SPAD reading, CTD, FL and lesser RWL, showing that they can be regarded as drought-adaptive traits, while the TKW and NGPS were more sensitive to drought. In 2011-12 (mild drought) the drier traits with enhanced performance were YLD, RGR, Fv/Fm, SL and NGPS.

The highest coefficient of variation (CV) was found for CTD followed by RWL, SC and PH. The least value was shown by DM followed by DH, Fv/Fm and RWC. The broad-sense heritability was the highest for plant height (90.8%), while Fv/Fm (31.6 %) and RGR (32.5%) were estimated as having the lowest broad-sense heritability. Plant height was found to be the most important constant character, while Fv/Fm and RGR were estimated to be the most affected traits over the three years in this study.

#### **Path coefficient analysis**

The correlation coefficients were partitioned into direct and indirect effects by path analysis (Table 3). Results indicated PH (0.65) followed by Fv/Fm (0.54), SL (0.47), RWC (0.46), NGSP (0.35) and TKW (0.30) had the highest positive direct effects, while DM (-0.55) and PL (-0.54) had the highest direct and negative effects on yield productivity.

The indirect influence of plant height through PL was positive, while through RWC and SL was negative. The indirect influence of days to maturity through DH was negative, while through Fv/Fm was positive and high. However, PL, SL, RWC and DM had the highest indirect effects on grain yield through plant height. Based on the path coefficient results, high residual (0.49) was observed for the data.

#### **Stepwise multiple linear regression (SMLR)**

The SMLR used to determine the variables accounting for the majority of total yield variability. Results of SMLR showed that the SPAD reading, Fv/Fm and SL with R square of 65.1%, had justified the maximum of yield changes (Table 4). Therefore the following equation can be obtained:

$$YLD = -6898.6 + 36.2*(SPAD) + 7945.8*(Fv/Fm) + 120.3*(SL)$$

The significant R square in the model indicates the effectiveness of these traits to increase grain yield. With respect to the positive and significant regression coefficients of the above mentioned traits, it could be stated that increasing the amount of these traits will cause an increase in the yield. The other variables were not included in the analysis due to their low relative contributions.

#### **Principal component analysis**

According to PCA, an increase in the number of components was associated with a decrease in eigen values (Table 5). This trend reached its maximum at three factors. Accordingly, it is reasonable to assume that the PCA had grouped the estimated wheat variables into five main components captured for 74.2% of the total variation of grain yield. The first component (PC1) accounted for about 25.65% of the variation in grain yield, while PC2 and PC3 accounted for 18.8% and 13.8%, respectively. The results showed that PC1 positively correlated with PH, SL and CTD. The PC2 correlated moderately with PL and Fv/Fm and the PC3 associated with RWL and RGR. The next two PCs

**Table 1.** Variance components for agronomic and physiological traits for 25 durum genotypes tested in three cropping seasons

Traits	Mean squares				
	Year (Y) ( $\sigma_{el}^2 + r\sigma_Y^2$ )	Block/Y $\sigma_{el}^2$	Genotype (G) ( $\sigma_{e2}^2 + r\sigma_{GY}^2 + r\sigma_G^2$ )	GxY ( $\sigma_{e2}^2 + r\sigma_{GY}^2$ )	Error $\sigma_{e2}^2$ E(MSS) <sup>#</sup>
Grain yield (YLD)	51966563.1**	572029.1	286079.1**	144753.3**	85621.8
Plant height (PH)	15704.6**	125.1	941.3**	88.8**	14.1
Thousand kernel weight (TKW)	4225.4**	6.0	19.8**	10.2**	2.6
Spike length (SL)	8.2**	0.3	2.4*	1.3**	0.6
Peduncle length (PL)	4966.9**	7.8	19*	13.2**	1.5
Flag-leaf length (FL)	240.7**	4.7	15.6*	10.4**	1.5
Number of grain per spike (NGPS)	354.0**	20.5	65.3ns	61.5**	4.5
Days to heading (DH)	4233.0**	14.4	36.7**	19.2**	6.1
Days to maturity (DM)	4854.5**	15.3	26.7**	18.8**	6.3
Relative growth rate (RGR)	103.2**	2.0	0.5ns	0.5**	0.2
Relative water loss (RWL)	0.797**	0.01	0.031ns	0.026**	0.006
Relative water content (RWC)	216.1**	17.4	58.7*	31.4**	10.4
Chlorophyll content (SPAD-reading)	685.3**	20.5	82.5**	18.9**	7.7
Canopy temperature dispersion (CTD)	87.8**	10.2	9.5**	4.2ns	4.2
Chlorophyll fluorescence (Fv/Fm)	0.024**	0.001	0.003ns	0.004**	0.001
Stomatal conductance (SC)	16318.1**	45.3	201.0*	135.9**	10.1

<sup>#</sup>E(MSS): Expectation of mean sum of squares; \*, \*\*Significant at 5% and 1% level of probability, respectively; ns: non-significant

**Table 2.** Descriptive statistics for grain yield and agronomic, phenologic and physiological criteria of 25 durum wheat genotypes across three cropping seasons

Traits	2010-11 (Moderate stress)		2011-12 (Mild stress)		2012-13 (Severe stress)		Overall		
	Average	Range	Average	Range	Average	Range	Reduction (%)*	CV%	H <sup>2</sup> b
YLD	1840.0	1345-2809	2360.0	866.7-3007	773.6	300.3-1567.0	67.2	12.3	53.6
PH	82.0	62-109	65.0	50.0-95.0	50	30.0-85.0	39	16.5	90.8
TKW	18.8	15.5-23	25.5	21.1-33.6	34.1	30.0-39.8	44.9	7.19	60.8
SL	6.9	5.4-9.2	7.4	6.4-9.2	7.3	5.5-11.5	6.8	10.3	54.3
PL	10.2	6.0-18.0	7.0	3.2-20.2	13.46	7.89-22.41	48	14.4	56.1
FL	19.1	15.4-24.3	17.6	14.0-32.8	15.4	11.3-19.0	19.4	9.85	56.3
NGPS	30.0	19-44	35.0	22.0-44.0	26.37	15.9-35.6	32.7	10.7	49.1
DH	194.0	180-200	181.0	176-186	181	176-183	6.7	1.02	59.0
DM	229.0	213-234	213.0	210-222	218	216-223	7	0.68	49.4
RGR	3.2	1.8-4.5	4.2	2.0-5.4	1.5	0.6-2.6	64.3	11.6	32.5
RWL	0.194	0.091-0.457	0.323	0.098-0.978	0.42	0.282-0.588	53.8	24.9	47.2
RWC	72.3	63.3-79.5	73.5	55.5-92.7	76.2	67.2-83.9	5.1	4.4	58.1
SPAD	54.3	38.9-63.7	38.0	30.8-45.4	41.5	37.2-47.1	30	6.22	77.7
CTD	3.2	1.8-4.5	1.2	0.1-4.7	2.4	0.4-5.3	62.5	31.3	48.7
Fv/Fm	0.745	0.716-0.809	0.779	0.704-0.845	0.768	0.70-0.83	4.4	2.04	31.6
SC	23.6	10.2-33.8	20.6	15.7-42.9	49.7	26.9-74.3	58.6	18.8	57.8

For trait codes see Table 1; \*Calculated from yield data of two extreme years

**Table 3.** Direct and indirect genetic effects via various paths of 15 agro-physiological traits on the grain yield of 25 durum wheat genotypes across years

Traits	Direct effect	DH	DM	PH	TKW	GR	RWL	RWC	SPAD	CTD	Fv/Fm	SC	SL	PL	FL	NGPS
DH	0.02		0.02	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	-0.01	-0.01	0.00	-0.01	0.00	0.00
DM	-0.55	-0.41		-0.21	0.07	-0.08	0.05	0.07	0.13	-0.05	0.24	0.33	-0.08	0.16	-0.12	0.02
PH	0.65	0.06	0.25		0.15	0.17	-0.13	-0.48	-0.26	0.16	-0.20	-0.23	0.30	0.34	0.19	-0.05
TKW	0.30	-0.07	-0.04	0.07		-0.02	-0.05	-0.07	-0.03	0.04	0.05	0.01	0.08	0.15	0.06	-0.12
RGR	-0.28	-0.02	-0.04	-0.07	0.02		-0.13	0.04	0.10	-0.09	0.05	0.03	-0.08	0.02	0.00	0.13
RWL	0.12	-0.02	-0.01	-0.02	-0.02	0.06		0.04	0.01	-0.01	0.00	0.01	-0.01	-0.02	-0.02	-0.02
RWC	0.46	-0.02	-0.06	-0.34	-0.11	-0.07	0.17		0.18	-0.08	0.08	0.08	-0.12	-0.17	-0.11	-0.06
SPAD	0.26	-0.10	-0.06	-0.10	-0.02	-0.09	0.03	0.10		-0.12	0.02	0.03	-0.15	-0.02	-0.12	0.10
CTD	0.20	0.00	0.02	0.05	0.02	0.06	-0.02	-0.04	-0.10		0.01	0.02	0.14	0.06	0.01	-0.07
Fv/Fm	0.53	-0.25	-0.23	-0.16	0.09	-0.10	0.00	0.10	0.04	0.03		0.03	0.06	0.16	0.07	0.06
SC	-0.15	0.04	0.09	0.05	0.00	0.02	-0.01	-0.03	-0.02	-0.01	-0.01		0.01	-0.01	-0.01	0.00
SL	0.47	0.01	0.07	0.22	0.13	0.13	-0.04	-0.12	-0.28	0.33	0.06	-0.03		0.22	0.14	-0.17
PL	-0.54	0.25	0.16	-0.29	-0.26	0.04	0.09	0.20	0.05	-0.16	-0.16	-0.04	-0.25		-0.14	0.01
FL	-0.12	-0.02	-0.03	-0.04	-0.02	0.00	0.02	0.03	0.06	-0.01	-0.01	-0.01	-0.04	-0.03		0.00
NGPS	0.35	0.02	-0.01	-0.03	-0.14	-0.17	-0.07	-0.05	0.13	-0.13	0.04	0.00	-0.12	-0.01	-0.01	-0.01

For trait codes see Table 1.

showed low correlations with the studied traits. Traits which significantly correlated with the first three eigen vectors were the variables with the greatest variability. Thus based on the first three PCs, the traits PH, SL, CTD, PL, Fv/Fm, RWL and RGR shown to be the important traits affecting greatly grain yield. However, high correlation between PC1 and a trait indicates that the trait is associated with the direction of the maximum amount of variation in the dataset.

### Factor analysis

According to factor analysis, the five factors explained 74.2% of the total variance observed in the plant traits (Table 6). Communality data explained by five factors were found to be the highest for PH (92.3%) and the lowest for FL (50.8%). However, PH, DH, DM, RWC and PL had the highest communality and consequently the high relative contribution in the productivity. Factor 1 captured for 19.2% of total variation and comprised plant height and spike length with positive loadings and SPAD reading and RWC with negative loadings. Factor 2 explained 14.6% of the total variance and comprised PL and Fv/Fm with the highest and positive loadings and DH and DM with the highest negative loading effects. The third factor explained 14.2% of the total variance and consisted of RWL and RGR with positive loadings and NGPS with negative loadings. Factor 4 explained 14.0% of the total variation and comprised FL with positive loadings, and PH with negative loading. The last factor explained 12.2% of the total variation and consisted of TKW with positive and SC with negative loadings. Based on the results, the five extracted factors, respectively, are representative for plant height, heading date, water loss, flag-leaf length and 1000-kernel weight.

In Table 7 the most important traits identified by each statistical procedure are presented. The results verified that at least three traits are required to explain the variation of grain yield in durum

**Table 4.** The regression coefficient (b), standard error (SE), T-value and probability of the estimated variables in predicting durum wheat grain yield by the multiple linear regression analysis technique

Traits	Regression slope		t-value		Collinearity statistics	
	b	SE	Value	Probe	Tolerance	VIF
Constant	-6898.6	1388.1	-4.97	0.000		
SPAD	36.2	11.9	3.041	0.006	0.628	1.592
Fv/Fm	7945.8	1734.3	4.582	0.000	0.954	1.048
SL	120.3	44.8	2.686	0.014	0.623	1.606

Model R<sup>2</sup>= 65.1%; SE: standard error; VIF: variance inflation factor  
For trait codes see Table1.

**Table 5.** The correlation coefficients between the traits and the eigenvector values for 25 durum wheat genotypes across three years

Traits	Components				
	PC1	PC2	PC3	PC4	PC5
DH	0.254	-0.821	-0.086	0.352	0.022
DM	0.420	-0.774	-0.096	0.021	0.286
PH	0.795	-0.043	-0.30	-0.435	-0.10
TKW	0.353	0.505	-0.095	-0.093	0.477
RGR	0.409	-0.161	0.672	-0.317	-0.178
RWL	-0.229	-0.027	0.698	-0.331	0.013
RWC	-0.632	-0.037	0.452	0.328	0.338
SPAD	-0.766	0.111	-0.192	-0.35	0.128
CTD	0.605	0.236	0.352	0.23	-0.164
Fv/Fm	-0.167	0.620	-0.038	0.264	0.349
SC	-0.283	0.467	0.158	0.359	-0.585
SL	0.773	0.264	0.208	0.187	0.031
PL	0.447	0.714	-0.271	-0.208	0.02
FL	0.477	0.076	-0.24	0.465	0.04
NGPS	-0.383	-0.116	-0.675	-0.015	-0.296
E.V.	3.85	2.83	2.08	1.31	1.06
Prop.	25.65	18.84	13.85	8.73	7.09
V.C. (%)	25.65	44.5	58.35	67.07	74.17

For trait codes see Table 1; E. V. = Eigen Value; Prop. = Proportion; V. C. = Variance Cumulative

genotypes. However, although the statistical procedures were varied in identify the number of required traits for selection, based on the different statistical procedures the chlorophyll fluorescence (Fv/Fm), spike length (SL), SPAD reading, peduncle length and heading date can be considered as the most contributors to yield productivity in durum wheat under drought conditions.

**Table 6.** Rotated factor loadings and communalities for the estimated variables of 25 durum wheat genotypes across three years

Traits	Factors					TKW
	PH	HD	WL	FLL	TFW	Communalities
PH	<b>0.795</b>	-0.043	-0.300	-0.435	-0.100	0.923
SL	0.773	0.264	0.208	0.187	0.031	0.747
SPAD	-0.766	0.111	-0.192	-0.350	0.128	0.775
RWC	-0.632	-0.037	0.452	0.328	0.338	0.827
CTD	0.605	0.236	0.352	0.230	-0.164	0.626
FL	0.477	0.076	-0.240	<b>0.465</b>	0.040	0.508
DH	0.254	<b>-0.821</b>	-0.086	0.352	0.022	0.870
DM	0.420	-0.774	-0.096	0.021	0.286	0.867
PL	0.447	0.714	-0.271	-0.208	0.020	0.826
Fv/Fm	-0.167	0.620	-0.038	0.264	0.349	0.606
TKW	0.353	0.505	-0.095	-0.093	<b>0.477</b>	0.624
RWL	-0.229	-0.027	<b>0.698</b>	-0.331	0.013	0.651
NGPS	-0.383	-0.116	-0.675	-0.015	-0.296	0.704
RGR	0.409	-0.161	0.672	-0.317	-0.178	0.777
SC	-0.283	0.467	0.158	0.359	-0.585	0.794
LR	2.88	2.20	2.12	2.10	1.83	
Factor	19.17	14.64	14.15	14.01	12.20	
	variance (%)					

For trait codes see Table 1; PH = Plant height; HD = Heading date; WL = Water loss; FLL = Flag leaf length; TKW = 1000-kernel weight; LR = latent roots

#### Validation the methodology of scaling scores

To validate the scaling score for selection criteria, the scores tested against their original value from each selection criteria. Table 8 shows the Pearson's correlation coefficients between the score assigned to selection criteria and the original traits. The

**Table 7.** Traits identified as most influence in durum wheat grain yield under drought condition by each applied statistical model. In regression model the "most important traits" were identified based on significance level, while in three other models based on a cut-off value

Traits	Statistical model			
	Regression analysis	Path coefficient analysis	Principal component analysis	Factor analysis
DH				✓
DM				
PH		✓	✓	✓
TKW				
RGR				
RWL			✓	✓
RWC		✓		
SPAD	✓		✓	✓
CTD			✓	
Fv/Fm	✓	✓	✓	✓
SC				
SL	✓	✓	✓	✓
PL			✓	✓
FL				
NGPS		✓		
Model R-square (%)	65.1	51	74.2	74.2

For trait codes see Table 1.

**Table 8.** Pearson's correlation coefficients between the score traits and their original values

	PH	SPAD	Fv/Fm	SL	PL	DH
<b>Class 1</b>						
PHs	0.98**	-0.46*	-0.30	0.48*	0.46*	0.18
SPADs	-0.40*	0.99**	0.08	-0.54**	-0.06	-0.40*
Fv/Fm	-0.34	0.10	0.99**	0.05	0.26	-0.44*
SLs	0.44*	-0.61**	0.09	0.98**	0.42*	0.05
PLs	0.54**	-0.08	0.31	0.46*	0.99**	-0.49*
<b>Class 2</b>						
DHs	-0.01	0.30	0.46*	-0.05	0.48*	-0.97**

\*, \*\* significant at 5% and 1% level of probability, respectively.

correlation coefficient between the score heading date and its original trait value is highly negative ( $r=-0.97$   $P<0.01$ ), as the score scale has been inverted. On

the other hand, the correlation coefficients between the original values for plant height, SPAD reading, Fv/Fm, spike length and peduncle length, were highly significant ( $P<0.01$ ). These high correlation coefficient values demonstrate that the score traits can be used as a surrogate of their original trait value.

Table 9 shows the score index of selection criteria and the mean score index (MSI) for 25 durum wheat genotypes. The data used in this table are a mean for each genotype from three cropping seasons.

**Table 9.** Scores of genotypes for each selection criteria and the mean score index (MSI) for the 25 durum wheat genotypes during 2010-13 cropping seasons

Geno- types	Class 1			Class 2			MSI
	Score PH	Score SPAD	Score Fv/Fm	Score SL	Score PL	Score HD	
G1	3	6	5	3	4	3	24
G2	1	7	10	4	3	2	27
G3	4	10	6	2	9	8	39
G4	3	5	5	3	4	6	26
G5	2	5	6	5	4	2	24
G6	3	10	6	3	4	3	29
G7	1	4	6	3	2	3	19
G8	3	5	9	4	8	4	33
G9	2	8	5	3	1	3	22
G10	3	7	5	3	4	3	25
G11	3	6	6	4	6	3	28
G12	2	8	6	2	5	6	29
G13	2	7	8	4	7	7	35
G14	1	6	10	1	4	5	27
G15	3	7	6	3	6	7	32
G16	1	7	2	1	2	3	16
G17	3	6	6	3	2	2	22
G18	1	4	5	3	2	3	18
G19	3	5	5	3	3	2	21
G20	8	4	7	5	7	5	36
G21	8	6	3	5	10	3	35
G22	3	7	9	4	6	10	39
G23	8	3	9	4	9	5	38
G24	10	4	1	2	2	1	20
G25	8	1	4	10	7	3	33

For trait codes see Table 1



The 25 score indices provide an illustration of small differences between plant height, peduncle length and spike length. On the other hand, SPAD reading and Fv/Fm were very similar, but both were slightly different from plant height, peduncle length and spike length. The score index of heading date in class 2 was generally similar to those in plant height, peduncle length and spike length.

According to MSI, the first six top ranking genotypes were G3, G22, G23, G20, G13 and G21 showing that these genotypes possessing some drought-adaptive traits. The superior performance of these breeding lines under drought stress was associated with better the above mentioned selection criteria. A positive correlation between MSI with genotypic mean yield was observed ( $P < 0.01$ ) indicating selection based on the traits involved in MSI produced the higher yield under drought environments over the years.

### Discussion

The genotype by year interaction was significant for most of the evaluated traits indicating that the relative performance of the genotypes changed across the cropping seasons. The multiple statistical procedures which have been used in this study showed that they differ in number of traits affecting grain yield. However, information from this study would be valuable to durum breeder for developing high yielding cultivars.

The four different statistical techniques applied in this study showed that the chlorophyll fluorescence (Fv/Fm), spike length (SL), SPAD reading, peduncle length and heading date were the most important yield traits to be considered under drought conditions. Thus, high yield productivity under drought conditions of Iran can possibly be obtained by selecting breeding materials for these traits. Dogan (2009) found that direct and positive effect of plant height on grain yield in durum wheat. Baranwal et al. (2012) revealed that grains per spike, spike length and 1000-grain weight exhibited the maximum positive direct effect. However, the above results also permit for further study of evolving desirable materials of durum wheat. The UMRL between grain yield and related agronomical and physiological (SPAD and Fv/Fm) traits indicated that under drought conditions, the contribution of the physiological traits was greater than the agronomical traits. Chlorophyll content was positively correlated with grain yield. Drought increases senescence by accelerating chlorophyll degradation leading to a decrease in leaf area and photosynthesis. There is

evidence that stay green phenotypes with delayed leaf senescence (higher SPAD index) can improve their performance under drought conditions (Lopes and Reynolds 2012). In wheat, genotypic variability has been detected in chlorophyll content as well as in the rate of leaf senescence (higher SPAD index) during grain-filling (Harris et al. 2007; Lopes and Reynolds 2012; Mohammadi et al. 2015). In durum wheat stay green genotypes growing under glasshouse conditions remained green for longer and had higher rates of leaf photosynthesis and seed weight (Spano et al. 2003).

The overall results reflect the importance of the five traits (Fv/Fm, SL, PL, SPAD, and heading date) in durum wheat selection for breeding programs. del Pozo et al. (2016) found a positive relationship between SPAD index and grain yield in durum wheat. It is likely that the breeding lines and new cultivar employed a drought escape strategy to maintain higher yield under the stress condition by shifting its flowering time frame (Table 9). Thus, these genotypes avoided, at least to some extent, the occurrence of the most severe drought (May to early June) during the most sensitive stage. However, selection on flowering time to enhance drought resistance is possible (Kenney et al. 2014).

The results indicated that indirect selection based on the components of yield may lead to different genetic gain in different levels of drought stress conditions. These results were also confirmed with relative efficiency between drought stress conditions which was higher at mild stress condition than severe drought stress condition. However, considerable genetic variation was observed for yield productivity, its components and persistency under variable drought seasons. This indicates that there is high potential for genetic improvement of this genetic materials. Low heritability for Fv/Fm suggests that indirect selection based on components of grain yield such as heading date, which had moderate heritability, would be more effective. However, selection for higher plant height, spike length, peduncle length, SPAD and heading date are associated with higher yield productivity in durum wheat under drought conditions. The path coefficients analysis model explained almost half of total variation (51%) in grain yield productivity. Therefore, some other component traits can be included to enhance utilization of variation in yield. The analysis model explained about 65% of total variability in grain yield. In compared, the two multivariate procedures of PCA and factor analysis explained more than 74% of total variation in grain yield, indicating the potential of PCA and FA for selection purposes. Based on this, the breeder could

use these multivariate selection methods by first determining the combination of traits that constitute an ideal plant. However, PCA and FA may be considered important if their associated scoring coefficients are of relative magnitude or sign consistent with breeding objectives. It has been shown that these statistical procedures can be successfully utilized in cereals (Walton 1972; Lee and Kaltsikes 1973; Godshalk and Timothy 1988; Cagirgan and Yildirim 1990).

The scaling score of selection criteria allowed arithmetic operations to create the MSI, which expresses yield performance under drought stress as a simple score scale value. This expression of yield has demonstrated that yield under stress can usefully be perceived as a function of two major crop characteristics, the resilience capacity and the production capacity (Thiry et al. 2016). The use of this index indicated that the some breeding lines showed a 'good' performance due to possessing high scaling scores of selection criteria. However the scaling score approach for the selection criteria was a useful method for discriminating some genotypes with late phenology, which actually could present better adaptive/tolerant traits to endure the stress although reducing their yield more, compared with early genotypes, and consequently showing a lower performance.

In conclusion, the study has shown the existence of considerable variation among the genetic materials for plant traits under drought conditions. The different statistical procedures which have been used in this study suggested the plant height, spike length, peduncle length, Fv/Fm, SPAD reading and heading date as the best indirect selection criteria for genetic improvement of yield in early generations under drought condition. However, selection of the best genotypes through these traits which have higher heritability than grain yield especially in early generations and associated with these traits have been emphasized for genetic improvement of grain yield. However, these make breeding for drought resistance particularly slow and difficult. Multivariate selection techniques (PCA and FA) by justifying the high portion of variability in grain yield were found to be more efficient for developing proper models for indirect selection in durum wheat.

#### Authors' contribution

Conceptualization of research (RM); Designing of the experiments (RM); Contribution of experimental

materials (RM); Execution of field/lab experiments and data collection (AE, LS); Analysis of data and interpretation (RM); Preparation of manuscript (RM).

#### Declaration

The authors declare no conflict of interest.

#### Acknowledgements

This work was financially supported by the Agricultural Research, Education and Extension Organization (AREEO) of Iran (Project No. 0-15-15-89102).

#### References

- Baranwal D. K., Mishra V. K. Vishwakarma M. K., Yadav and P. S. B. Arun. 2012. Studies on genetic variability, correlation and path analysis for yield and yield contributing traits in wheat (*T. aestivum* L. em Thell.). *Plant Archives*, **12**: 99-104.
- Barrs H. D. 1968. Determination of water deficits in plant tissues. In: Kozolovski TT (ed), *Water Deficits and Plant Growth*. Academic Press, 235-368.
- Canci H. and Toker C. 2009: Evaluation of yield criteria for drought and heat resistance in chickpea (*Cicer arietinum* L.). *J. Agron. Crop Sci.*, **195**: 47-54.
- Cagirgan M. I. and Yildirim M. B. 1990. An application of factor analysis to data from control and macro mutant populations of Quantum barley. *J. Fac. of Agric. of Akdeniz University*, **4**: 125-138.
- Cattell R. B. 1965. *Factor analysis: an introduction to essentials*. 1. The purpose and underlying models. *Biometrics*, **21**: 190-215.
- del Pozo A., Iván A. Y. Matus A. Tapia G. Castillo D. Sanchez-Jardón L. and Araus J. L. 2016. *Physiological Traits Associated with Wheat Yield Potential and Performance under Water-Stress in a Mediterranean Environment*. *Front. Plant Sci.* in press.
- Dewey D. R. and Lu K. H. 1959. A correlation and path coefficient analysis of components of crested wheat grass seed production. *Agro. J.*, **9**: 515-518.
- De Vita P., Nicosia O. L. D. F., Nigro, C., Platani, C. Riefolo, Di Fonzo N. and Cattivelli L. 2007. Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *Eur. J. Agron.*, **26**: 39-53.
- Deyong Z. 2011. Analysis among main agronomic traits of spring wheat (*Triticum aestivum*) in Qinghai Tibet plateau. *Bulgarian J. Agric. Sci.*, **17**: 615-622.
- Dogan R. 2009. The correlation and path coefficient analysis for yield and some yield components of durum wheat (*Triticum turgidum* var. durum Desf.) in west Anatolia conditions. *Pak. J. Bot.*, **41**: 1081-89.
- Ehdaie B. and Waines J. G. 1989. Genetic variation, heritability and path-analysis in landraces of bread wheat from s-w Iran. *Euphytica*, **41**: 183-190.

- Girdthai T., Jogloy S., Kesmala T., Vorasoot Akkasaeng C., Wongkaew S., Holbrook C. C. and Patanothai A. 2009. Relationship between root characteristics of peanut in hydroponics and pot studies. *Crop Sci.*, **50**: 159-167.
- Godshalk E. B. and Timothy D. H. 1988. Factor and principal component analyses as an alternative to index selection. *Theor. Appl. Genet.*, **76**: 352-360.
- Gustafson D. I. 2011. Climate change: a crop protection challenge for the twenty-first century. *Pest Manag. Sci.*, **67**: 691-916.
- Harris K., Subudhi P. K., Borrel A., Jordan D., Rosenow D., Nguyen H. et al. 2007. Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence. *J. Exp. Bot.*, **58**: 327-338.
- Hoffmann W. A. and Poorter H. 2002. Avoiding Bias in Calculations of Relative Growth Rate. *Annals of Botany*, **90**: 37-42.
- Huang Q., Zhao Y., Liu C., Zou X., Cheng Y., Fu G., Xu J., Zhang X. and Lu G. 2015. Evaluation of and selection criteria for drought resistance in Chinese semiwinter rapeseed varieties at different developmental stages. *Plant Breed.*, **134**: 542-550.
- Kenney A. M., McKay J. K., Richards J. H. and Juenger T. E. 2014. Direct and indirect selection on flowering time, water-use efficiency (WUE, delta C-13), and WUE plasticity to drought in *Arabidopsis thaliana*. *Ecol. Evol.*, **4**: 4505-4521.
- Kaiser H. F. 1958. The varimax criterion for analytic notation in factor analysis. *Psychometrika*, **23**: 187.
- Lee J. and Kaltsikes P. J. 1973. Multivariate statistical analysis of grain yield and agronomic characters in durum wheat. *Theor. Appl. Genet.*, **43**: 226-231.
- Leilah A. A., Badawi M. A. and El-Moursi S. A. 1988. Yield analysis of soybean. *J. Agric. Sci.*, **13**: 2344-2351.
- Leilah A. A. and Al-Khateeb S. A. 2005. Statistical analysis of wheat yield under drought conditions. *J. Arid. Environ.*, **61**: 483-496.
- Lopes M. S. and Reynolds M. P. 2012. Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. *J. Exp. Bot.*, **63**: 3789-3798.
- Mir R. R., Zaman-Allah M., Sreenivasulu N., Trethowan R. M. and Varshney R. K. 2012. Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theor. Appl. Genet.*, **125**: 625-645.
- Moghaddam M., Ehdaie B. and Waines J. G. 1998. Genetic variation for and inter-relationships among agronomic traits in landraces of bread wheat from southwestern Iran. *Journal of Genetics and Breeding*, **52**: 73-81.
- Mohammadi R., Heidari B. and Haghparast R. 2013. Traits associated with drought tolerance in spring durum wheat (*Triticum turgidum* var. durum Desf.) breeding lines from international germplasm. *Crop Breed. J.*, **3**(2): 87-98.
- Mohammadi R., Amri A., Ahmedi H. and Jafarzadeh J. 2015. Characterization of tetraploid wheat landraces for cold tolerance and agronomic traits under rainfed conditions of Iran. *The Journal of Agricultural Science (Cambridge)*, **153**: 631-645.
- Montgomery D. C. 2006. Introduction to linear aggression analysis. John Wiley and Sons.
- Nyquist W. E. 1991. Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.*, **10**: 235-322.
- Ristic Z., Bukovnik U. and Prasad P. V. V. 2007. Correlation between heat stability of thylakoid membranes and loss of chlorophyll in winter wheat under heat stress. *Crop Sci.*, **47**: 2067-2073.
- Rymuza K., Turska E., Wielogórska G. and Bombik A. 2012. Use of principal component analysis for the assessment of spring wheat characteristics. *Acta Scientiarum Polonorum-Agriculture*, **11**: 79-90.
- Seiller G. J. and Stafford R. E. 1985. Factor analysis of components in Guar. *Crop Sci.*, **25**: 905-908.
- Singh, S. P. and Diwivedi V. K. 2002. Character association and path analysis in wheat (*Triticum aestivum* L.). *Agric. Sci. Digest*, **22**: 255-257.
- Staniak M. and Kocon A. 2015. Forage grasses under drought stress in conditions of Poland. *Acta. Physiol. Plant*, **37**: 116.
- Singh D., M. Singh and K. C. Sharma. 1979. Correlation and path-coefficient analysis among flag leaf-area, yield and yield attributes in wheat (*Triticum aestivum* L.). *Cereal Research Communications*, **7**(2): 145-152.
- Snedecor G. W. and Cochran W. G. 1981. *Statistical Methods*, Iowa State University, Ames, Iowa, USA, 7th edition.
- Spano G., Fonzo N. Di, Perrotta C., Platani C., Ronga G., Lawlor D. W. et al. 2003. Physiological characterization of 'stay green' mutants in durum wheat. *J. Exp. Bot.*, **54**: 1415-1420.
- Thiry A. A., Chavez Dulanto P. N., Reynolds M. P. and Davies W. J. 2016. How can we improve crop genotypes to increase stress resilience and productivity in a future climate? A new crop screening method based on productivity and resistance to abiotic stress. *J. Exp. Bot.*, **67**(19): 5593-5603.
- Trenberth K. E. 2011. Changes in precipitation with climate change. *Clim. Res.*, **47**: 123-138.
- Walton P. D. 1972. Factor analysis of yield in spring wheat (*Triticum aestivum* L.). *Crop Sci.*, **12**: 731-733.
- Yang R. C., Jana S. and Clarke J. M. 1991. Phenotypic diversity and associations of some potentially drought responsive characters in durum wheat. *Crop Sci.*, **31**: 1484-1491.
- Wright S. 1921. Correlation and causation. *J. Agric. Res.*, **20**: 557-559.