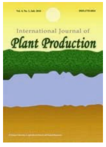

International Journal of Plant Production 9 (3), July 2015
ISSN: 1735-6814 (Print), 1735-8043 (Online)
www.ijpp.info

GUASNR

# Relationships between wheat yield, yield components and physico-chemical properties of soil under rain-fed conditions

A. Ranjbar, A.R. Sepaskhah*, S. Emadi

*Irrigation Department, Shiraz University, Shiraz, I.R. of Iran.*
*Corresponding author. E-mail: sepas@shirazu.ac.ir

Received 25 June 2014; Accepted after revision 10 February 2015; Published online 10 May 2015

## Abstract

This research was conducted to study and classify the physico-chemical properties of soil, yield components of wheat and to determine the significance of these parameters on the grain yield formation. In this research, seven statistical methods consisting of simple correlation analysis (SCA), multiple linear regression (MLR), stepwise regression (SR), factor analysis (FA), principal component analysis (PCA), cluster analysis (CA) and path analysis (PA) have been investigated. The physico-chemical properties of soil, different morphological traits and wheat yield have been obtained from a field with 250×300 meter dimension located in Bajgah (with silty clay loam soil) that consisted of 30 samples. Among statistical analysis performed, MLR has provided more acceptable results. In this method, among the examined characteristics, five traits i.e., the number of stems without spikes per plant, biological yield, harvest index, soil soluble potassium and soil available phosphorus, examined 98.3% of the variations of the yield ($P<0.05$). Lack of soil nitrogen effect on yield is due to drought stress conditions in which the plant growth is less sensitive to nitrogen. Furthermore, the negative effect of phosphorus on the yield of plant may be due to the inverse relationship between the soil phosphorus and micronutrient elements on the plant growth. Generally, among the yield components, biological yield is the most important and effective trait on grain yield, that has presented a significant contribution in different statistical methods. For some of the used statistical methods, the measured traits, like length of spike, the number of spikes per square meter, the number of grains per spike and harvest index showed positive effects on the grain yield and other traits like

1000-grains weight, the weight of grain per spike and the number of tillers without spikes per plant showed negative effects on the grain yield with the highest correlation. Among different soil nutrition, soluble potassium, phosphorus, sulfate and available potassium with positive effects and the clay content with negative effects showed the most correlation with the grain yield.

***Keywords:*** Statistical analysis; Physico-chemical properties of soil; Morphological traits; Wheat grain yield.

## Introduction

In semi-arid regions like Iran, the crops that could be produced without irrigation are fairly limited and none of these crops are able to get priority on wheat which is the most important crops in these regions. Therefore, grain crops, especially wheat in arid and semi-arid regions showed a great economic importance. The total production of world cereal is 1.8 billion Mg and the highest amount of that (about 500 to 600 million Mg) is wheat. Furthermore, in terms of the cultivated area and the annual production, wheat is primarily important. The mean wheat annual consumption per capita is 130 kg in the world and 220 kg in Iran.

Texture and structure of soil, available nutrients in soil, the water holding capacity and hydraulic properties of soil, soil micro-organisms, acidity and available air in soil are the factors that directly affect the quality and quantity of produced wheat. Furthermore, the structure of soil causes the difference in penetration of roots, transmission of water and absorption of nutrients and consequently affects on the percentage of available protein. The soil water content has a direct and indirect effect on the growth of wheat. The direct effect is related to available water around the root that is absorbed by roots. The indirect effect is related to the characteristics of soil for plant water and nutrient absorption ability which is dependent on the soil water content and the spreading of the roots systems. The amount of soil oxygen is one of the important factors in the growth of the plant. Soil pores, particles dispersion and physical properties of soil determines the aeration of soil. The soil water content is one of controller factors of soil oxygen for plants. Wheat is grown up in clay soils, even in soil with saturation condition, in comparison to other plant; however, when saturation prolongs, it vanishes the growth easily. Neutral and alkaline pH is favorable for wheat. Therefore, the wheat growth is sensitive in acidic soils. Wheat grows in clay or sandy soils with enough water content and has satisfactory yield.

Nitrogen, phosphorus and potassium are macro nutrient elements for wheat growth that must be available in soil. Nitrogen is the first element that wheat needs it more than other nutrients. In the rain-fed (dry farming) regions because of weather conditions and low soil organic matter, there is nitrogen deficiency in most of soils. The reason for this phenomenon is improper method of cultivation and harvest of crop. Nitrogen fertilizer is needed for better growth of wheat, grain yield and protein content. By inadequate supply of soil nitrogen both the yield and quality of wheat grain is decreased. Phosphorus plays a very important role in vital and constructional functions of plants. Phosphorus is needed in creation of main roots, tillering, enhancing the seed weight and the quality of cooking, formation of carbohydrate in plants. Furthermore, phosphorus results in increase in nitrogen absorption and resistance of wheat to disease and is a controller of the negative effect of excess nitrogen. Potassium is another element that plant needs it and often there is a sufficient amount of K in heavy soils. Potassium application increases the resistance of plants to frost, drought, lodging and also it increases the storage of starch. In order to determine the effects of soil characteristics on wheat yield, extensive studies have been conducted in all over the world. The results presented the importance of the studying of the relationship between crops yield and properties of soil. Therefore, studying and classification of the physico-chemical properties of soil on wheat yield by using multivariate statistical methods would be a major step towards increasing grain yield. Davatgar et al. (2012) analyzed different soil properties by using principal component analysis (PCA) and fuzzy cluster algorithm. These results indicated that cluster analysis by reducing within- zone variability provided an opportunity for farmers to adopt site-specific nutrient management. Khormali et al. (2007) studied the relationship between soil physico-chemical properties and tea yield in Lahijan area, Guilan province, Iran. They found that the soil properties such as organic carbon, total N, clay, carbonate and exchangeable Mg varied significantly at different slope positions; however, the tea yield was not significantly different at different positions. Grain yield is a complex trait that in addition to soil parameters, it depends on the other different yield components. We could mention morphologic yield components that are measured easily and carefully. Yield components have a lot of inheritability, therefore based on these traits the plant selection would be a confident and quick way for riddling all of the plants and improvement of the yield (Yap and Harvey, 1972). Among morphological characteristics we can name root length, plant height, number of spikes per unit area, number of seeds per spike, number of fertile tillers per plant, 1000-grain weight, stem length, spike weight, stem weight, seed weight per spike.

So far, different models have been proposed to justify the relationships between these traits and grain yield. Correlation studies between yield and yield components in bread wheat under drought conditions were done by Munir et al. (2007) and they concluded that grain yield in the plant has positive and meaningful correlation with leaf area, number of tillers per plant, number of grains per plant and weight of grain per plant and 1000-grain weight. According to various reports, the correlation between grain yield and 1000-grain weight, number of fertile tillers or spikes per plant, number of spikelets per spike, number of grains per spike and spike length is significant (Bahatt, 1973; Bulman and Hunt, 1988; Ledent and Moss, 1979; Mohiuddin and Cory, 1980; Shanahan et al., 1985; Srivatsava et al., 1971; Virk and Vema, 1972).

Recognition and perception of relations between soil and plant trait and yield, especially in rainfed conditions, have an extraordinary effect in crop management. Therefore, in this research seven statistical methods consisting of simple correlation analysis (SCA), multiple linear regression (MLR), stepwise regression (SR), factor analysis (FA), principal component analysis (PCA), cluster analysis (CA) and path analysis (PA) have been investigated in order to classify the physico-chemical properties of soil and yield components on wheat yield and to determine the significance of the traits that are relevant to grain yield.

## Materials and Methods

In this research, in order to study the relationship between soil properties, yield components and grain yield of rain-fed wheat, expertise knowledge and statistical analysis were used. For performing this research, physico-chemical properties of soil and different morphologic characteristics of rain-fed wheat that have been harvested from a field with 250 m×300 m area located in Bajgah (with silty clay loam soil) were used. Data consisted of 30 samples obtained from the 1.25 $m^2$ area and the spacing between each sample was 50 m.

In this field, sardari rainfed local cultivar was planted at a seeding rate of 120 kg $ha^{-1}$ on lines with 0.25 m spacing at second week of November.

*Physical and chemical analysis of soil*

Soil samples were collected from 0-30 cm depth. After air drying of soil samples and passing through a 2 mm sieve, some physico-chemical properties were measured as follows:

• Acidity in saturated paste by pH meter (Thomas, 1996).
• The electrical conductivity of the saturated extract by electrical conductivity meter (Rhoades, 1996).
• The capacity of soil water saturation determined as difference between the soil water of saturation and the dried sample weight at 105 $^{o}$C for 24 hours.
• Sodium and potassium by Flame Photometers apparatus (soluble K with extraction water from the saturated soil and soil available K extracted by ammonium acetate with pH=7) (Helmake and Sparks, 1996).
• Measurement of organic carbon by the method of wet ashing (Nelson and Sommers, 1996).
• Bicarbonate, titrated with sulfuric acid (Loppert and Suarze, 1996).
• Chloride by titrating with silver nitrate (Frankenberger et al., 1996).
• Calcium and magnesium and soluble sulfate (titrating with Na EDTA) (Richards, 1954)
• Total nitrogen by Kjeldahl (Bremner, 1996)
• Available P by extracting with sodium bicarbonate (Olsen et al., 1954).
• Soil texture by hydrometer (Gee and Bauder, 1986).

*Recording the yield components*

The yield components of wheat were recorded, included the number of spikes per plant, number of tillers per plant, number of tiller without spike per plant, length of spike, 1000-grain weight, number of spikes per $m^2$, number of grains per spike, grain weight per spike, biological yield (straw weight + grain weight), harvest index (grain weight / total weight) and grain yield. Furthermore, the plant density was not very different in different locations; therefore, it was not considered in the analyses.

*Methods of statistical analysis*

The effects of every characteristics of soil and yield components on the grain yield were investigated by seven statistical methods consisting of simple correlations analysis (SCA), multiple linear regression (MLR), stepwise multiple regression (SMR), factors analysis (FA), principal component analysis (PCA), cluster analysis (CA) and path analysis (PA). Leilah and Khateeb (2005) have explained these statistical methods. All calculations and statistical analysis was done by softwares: Spss, Minitab and Path.

## Results and Discussion

*Simple correlation analysis*

The descriptive statistics of all morphologic characteristics of wheat and physico-chemical properties of soil are summarized in Table 1. The mean of grain yield is 121.28 g/m$^2$ that has a broader range as well. The values of variation coefficients (CV) of soil properties showed that the variation percentage in the area is at least 2.46% for acidity and a maximum of 72.34% for soil chloride. According to the classification given by Wilding (1985) and the values of variation coefficients, nitrogen and sulfate showed low variability and other nutritious elements of soil, showed high variability. The highest variation coefficients of morphologic traits belong to the number of tillers per plant and the lowest is related to 1000-grain weight.

The values of simple correlation coefficients among different traits are given in Table 2. The analysis of correlation coefficients for different traits with grain yield helps to decide on the relative importance of these traits and their values as selection criterions (Leilah and Khateeb, 2005). It has shown that there is a positive and significant correlation at 5% of probability level between grain yield and the traits of spike length (r=0.438), number of spikes per $m^2$ (r=0.883), number of grains per spike (r=0.447), biologic weight (r=0.908) and sulfate (r=0.374). Also, negative and significant correlation between 1000-grain weight and grain yield (r= -0.478) indicated that 1000-grain weight and grain yield have opposed behavior with each other. This could result from ontogenic relationships between seed number and weight. The percentage of clay, has a significant negative correlation with the grain yield as well (r= -0.303). However, simple correlation coefficient between yield and other traits are not significant at the 5% level of probability. Therefore, the grain yield has the highest correlation with biological yield. The results of causality analysis in the study of Vrindts et al. (2003) showed that soluble P (r=0.55) and fine sand (r=0.52) have a significant correlation with the yield of wheat grain. Also, Vrindts et al. (2003) achieved in another study, a positive and significant correlation between grain yield and straw with the soil soluble phosphorus, the percentage of coarse silt and sand and a negative and significant correlation with the percentage of coarse silt, clay and potassium. According to Table 2, salinity has no significant relationship with the yield that it might be due to the limited range of salinity in this study. Thus, salinity has not appeared as a negative factor for plant growth. However, it has a positive and significant correlation with soluble potassium (r=0.394), bicarbonate (r=0.519), plant available P (r=0.39) and sand (r=0.383) and a negative and significant correlation with chloride ions (r= -0.422) and silt (r= -0.417). Corwin et al. (2003) have achieved significantly the correlation coefficients between electrical conductivity and the yield of cotton (0.53), acidity (0.33), chloride (0.61), clay (0.76) and saturation capacity (0.77). Kumbhar et al. (1983) and Mohamed (1999), using correlation coefficients showed that grain weight per spike, biological yield and number of spike per $m^2$ have a high correlation with grain yield. These results are similar to our results. Also, in the study of Aycicek and Yildirim (2006) there was a positive and significant correlation between the number of grains per spike and the yield. Since the different relations between yield components with grain yield may be occurred due to environmental effects on plant growth (Assenge et al., 2002), therefore, the lack of a positive and significant correlation between harvest index and grain yield is justified.

Table 1. Basic statistics [minimum and maximum values, arithmetic mean and standard deviation (SD)] for the variable.

| Variables | Minimum | Maximum | Mean | Std. Deviation | Coeff of variation (%) |
|---|---|---|---|---|---|
| Number of spikes per plant ($X_1$) | 4.32 | 15.36 | 7.67 | 2.70 | 35.16 |
| Number of stems per plant ($X_2$) | 4.03 | 16.16 | 8.24 | 3.03 | 36.79 |
| Number of tiller with no spike per plant ($X_3$) | 0.14 | 3.06 | 0.60 | 0.59 | 98.16 |
| Lengths of spike ($X_4$) | 4.72 | 10.06 | 6.96 | 1.05 | 15.07 |
| 1000-grain weight (gr) ($X_5$) | 34.2 | 46.70 | 42.34 | 3.52 | 8.32 |
| Number of spikes/m$^2$ ($X_6$) | 68 | 326.40 | 183.65 | 54.50 | 29.68 |
| Number of grains/spike ($X_7$) | 11.7 | 30.60 | 20.73 | 4.14 | 19.96 |
| Weight of grain/spike (gr) ($X_8$) | 0.5 | 6.50 | 1.07 | 1.04 | 97.11 |
| Biological yield (gr/m$^2$) ($X_9$) | 117 | 644.00 | 388.3 | 137.57 | 35.43 |
| Harvest index ($X_{10}$) | 0.27 | 0.57 | 0.40 | 0.06 | 13.94 |
| pH ($X_{11}$) | 7.72 | 8.51 | 8.24 | 0.20 | 2.46 |
| EC (dS/m) ($X_{12}$) | 0.21 | 0.32 | 0.26 | 0.03 | 11.44 |
| S.P (%) ($X_{13}$) | 40 | 49.00 | 43.93 | 2.72 | 6.18 |
| Na (ppm) ($X_{14}$) | 1.4 | 8.00 | 4.05 | 1.92 | 47.43 |
| K (ppm) ($X_{15}$) | 2.75 | 11.50 | 5.55 | 2.26 | 40.76 |
| O.M (%) ($X_{16}$) | 0.36 | 0.65 | 0.54 | 0.08 | 14.28 |
| $HCO_3^-$ (meq/l) ($X_{17}$) | 3.2 | 5.20 | 4.36 | 0.51 | 11.66 |
| $Cl^-$ (meq/l) ($X_{18}$) | 7.5 | 46.00 | 13.77 | 9.96 | 72.34 |
| $Ca^{2+}$+ $Mg^{2+}$(meq/l) ($X_{19}$) | 5.9 | 9.50 | 7.75 | 0.95 | 12.30 |
| $SO_4^{2-}$ (meq/l) ($X_{20}$) | 3 | 5.50 | 4.15 | 0.66 | 15.80 |
| T.N (%) ($X_{21}$) | 0.073 | 0.24 | 0.12 | 0.03 | 29.21 |
| P (ppm) ($X_{22}$) | 18 | 64.00 | 34.73 | 10.16 | 29.26 |
| $K_{ava}$ (ppm) ($X_{23}$) | 360 | 840.00 | 558.33 | 120.00 | 21.49 |
| Sand (%) ($X_{24}$) | 9.56 | 23.44 | 15.61 | 3.99 | 25.59 |
| Silt (%) ($X_{25}$) | 34.56 | 49.56 | 41.09 | 5.08 | 12.36 |
| Clay (%) ($X_{26}$) | 31 | 51.00 | 43.33 | 6.19 | 14.29 |
| Grain yield (Y) | 53.2 | 217.60 | 121.28 | 37.18 | 30.66 |

Table 2. A matrix of simple correlation coefficients (r) for the variables.

| Variables | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{11}$ | $x_{12}$ | $x_{13}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of stems per plant ($X_2$) | .392(*) | | | | | | | | | | | |
| Number of non fertile tiller per plant ($X_3$) | .410(*) | | | | | | | | | | | |
| Lengths of spike with no awn (cm) ($x_4$) | .371(*) | | | | | | | | | | | |
| Number of spikes/m$^2$ ($X_6$) | | | | .373(*) | -.522(**) | | | | | | | |
| Number of grains/spike ($X_7$) | | | | .889(**) | | | | | | | | |
| Biological yield (g/m$^2$) ($X_9$) | .387(*) | | | .562(**) | -.518(**) | .870(**) | .473(**) | | | | | |
| Harvest index ($X_{10}$) | -.414(*) | | -.569(**) | -.475(**) | | -.418(*) | | | -.623(**) | | | |
| pH ($x_{11}$) | .492(**) | | | | | | | -.424(*) | | | | |
| SP (%) ($x_{13}$) | | | | | | | | | | -.474(**) | | |
| Na (ppm) ($x_{14}$) | | | | | | | | | | .392(*) | | |
| K (ppm) ($x_{15}$) | | | | | | | | | | | .394(*) | .452(*) |
| OM (%) ($x_{16}$) | | | | | | | | | | | | .451(*) |
| $HCO_3^-$ (meq/l) ($x_{17}$) | | | | | | | | | | | .519(**) | |
| $Cl^-$ (meq/l) ($x_{18}$) | | .611(**) | | | | | | | | | -.422(*) | -.476(**) |

Continue Table 2.

| Variables | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{11}$ | $x_{12}$ | $x_{13}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Ca^{2+}$ + $Mg^{2+}$ (meq/l) ($x_{19}$) | -.449(*) | | | | | | | | | | | |
| $SO_4^{2-}$ (meq/l) ($x_{20}$) | | | | .379(*) | | | .396(*) | | .368(*) | | | |
| P (ppm) ($x_{22}$) | | | | | | | | | | | .390(*) | .437(*) |
| $K_{ava}$ (ppm) ($x_{23}$) | | | | | | | | | | | | .545(**) |
| Sand (%) ($x_{24}$) | | | | | | | | | | | .383(*) | -.388(*) |
| Silt (%) ($x_{25}$) | | | | | -.368(*) | | | | | | -.417(*) | |
| Clay (%) ($x_{26}$) | | | | | | | | | | | | .541(**) |
| Grain yield ($g/m^2$) (Y) | | | | .438(*) | -.478(**) | .883(**) | .447(*) | | .908(**) | | | |
| Variables | $x_{14}$ | $x_{15}$ | $x_{16}$ | $x_{17}$ | $x_{18}$ | $x_{19}$ | $x_{20}$ | $x_{21}$ | $x_{22}$ | $x_{23}$ | $x_{24}$ | $x_{25}$ |
| OM (%) ($x_{16}$) | | .367(*) | | | | | | | | | | |
| $SO_4^{2-}$ (meq/l) ($x_{20}$) | | | | | | .552(**) | | | | | | |
| P (ppm) ($x_{22}$) | | .795(**) | .475(**) | | | | | | | | | |
| $K_{ava}$ (ppm) ($x_{23}$) | | .819(**) | | | -.383(*) | | | | .665(**) | | | |
| Sand (%) ($x_{24}$) | .458(*) | -.361(*) | | .588(**) | | | | .421(*) | | -.461(*) | | |
| Silt (%) ($x_{25}$) | | | | | .606(**) | | | | | -.430(*) | | |
| Clay (%) ($x_{26}$) | -.369(*) | .391(*) | | | -.457(*) | | | | .462(*) | .646(**) | -.587(**) | -.759(**) |
| Grain yield ($g/m^2$) (Y) | | | | | | | .374(*) | | | | | |

* and **: means that r is significant at 5% and 1% level of probability, respectively.

### *Multiple linear regression*

The correlations between different factors were discussed in previous section, however, their relative importance and combined effects are not presented. Therefore, in multiple regression analyses, grain yield as a dependent variable against the other traits as independent variables are examined. The achieved equation for yield (Y) was obtained by entering and using the variables of yield components and physico-chemical properties of soil as follows:

$$Y = 26.82 - 8.52X_3 + 0.328X_9 + 282.3X_{10} + 4.464X_{15} - 1.207X_{22} \quad (1)$$

$$R^2 = 0.995,\ \ SE = 7.93, N = 30,\ \ P = 0.011$$

Where Y is the grain yield in g per harvest area (1.25 $m^2$), ($x_3$) is the number of the tillers without spike per plant, ($x_9$) is the biological yield, ($x_{10}$) is the harvest index, ($x_{15}$) is the soluble potassium and ($x_{22}$) is the soil available phosphorous. This might attributed to the rainfed condition in which the crops show lower growth and require less nitrogen. All the studied traits with $R^2$=99.5% justify well the yield variations and only 0.5% depends on the remaining effects on wheat yield. However, as it is shown among all the studied traits only 5 traits are known significant. Therefore, the grain yield as a dependent variable against significant traits as independent variables was investigated and the final equation is as follows:

$$Y = -104.518 - 6.421X_3 + 0.316X_9 + 294.504X_{10} + 4.01X_{15} - 0.979X_{22} \quad (2)$$

$$R^2 = 0.983,\ \ SE = 5.399, N = 30,\ \ P = 0.001$$

It is shown that Eq. (2) explained 98.3% of the variations in grain yield by using the number of the stems without spike per plant ($x_3$), biological yield ($x_9$), harvest index ($x_{10}$), soluble potassium ($x_{15}$) and phosphorous ($x_{22}$).

Grain yield was not dependent on the soil nitrogen in the regression equation. This might be due to the rain-fed conditions for wheat production. In these conditions, plant growth is low and it needs less nitrogen. Also, the negative effect of phosphorus on yield may be due to the inverse relationship between phosphorous and the micro nutrient uptake in the plant. In Iran, highly calcareous soils are frequently used for wheat production;

however, despite of relatively high total soil P, crops often suffer from P deficiency because of interaction with other nutrients especially micronutrients (Sepehr et al., 2009).

Results of the multiple regression in the study of Ayoubi et al. (2009 ) on soil parameters including total nitrogen, available phosphorus, available potassium, the capacity of cation exchange, electrical conductivity, acidity, MWD[1], WSA[2], FC[3], AWHC[4], BD[5] and CCE[6] showed that soil properties justify 78 and 73% of the variations of the grain yield and bio-mass of barley. Leilah and Al-khateeb (2005) in their studies on yield and yield components by using the statistical method of multiple regression showed that according to the t-student test, the traits of spike length, number of spikes per $m^2$, grain weight per spike, harvest index and biological yield have significant effects on wheat yield that they enter in the equation of multiple regression by stepwise linear regression.

*Stepwise linear regression*

Using stepwise regression based on the order of traits importance, three traits were selected and their causalities were analyzed. Biological yield was the first trait included in the model and 82.5% of the yield variation was justified by itself. Next traits entered in the model, were harvest index and number of stems without spikes per plant that totally these traits justified 96.1% of the variations of grain yield. The equations achieved from 3 steps are as follows:

$$Y = 25.974 + 0.245X_9 \qquad R^2 = 0.825 \tag{3a}$$

$$Y = -122.305 + 0.321X_9 + 295.7X_{10} \qquad R^2 = 0.947 \tag{3b}$$

$$\begin{gathered} Y = -92.924 + 0.319X_9 + 237.996X_{10} - 9.256X_3 \\ R^2 = 0.961, \quad SE = 7.738, \quad N = 30, \quad P = 0.001 \end{gathered} \tag{3c}$$

---

1.Mean weight diameter of aggregates
2.Water-stable aggregates
3.Volumetric field capacity
4.Available water-holding capacity
5.Bulk density
6.Calcium carbonate equivalent

Where Y is the grain yield in g per harvest area (1.25 $m^2$), ($x_3$) is the number of the tillers without spikes per plant, ($x_9$) is the biological yield and ($x_{10}$) is the harvest index. According to the results of stepwise regression analysis, physico-chemical parameters of soil have not entered in the equation and the grain yield was dependent only on the morphological traits including number of the tillers without spikes per plant, biological yield and harvest index: This indicated that the higher biological yield and harvest index are, the higher grain yield is obtained. Furthermore, the number of tillers without spikes per plant has a negative relation with the yield. By increasing of this trait, the grain yield decreased. However, since the role of soil elements in plant growth is undeniable and nitrogen, potassium and phosphorus are the macro-nutrient elements of the plant, so using the method of multiple regression in order to study and classify the effects of physico-chemical properties of soil on wheat yield is more suitable. In contrary to the method of stepwise regression analysis, the parameters of soil (soluble K and P) as independent variables that are effective on wheat yield besides yield components (number of the tillers without spikes per plant, biological yield and harvest index) entered in the equation. It is shown that different methods present different results as the main factors affecting the grain yield. However, by comparison between the different models, we will conclude the main factors affecting the grain yield and present the best method. Yan et al. (2007) by using the method of stepwise regression installed a relationship between chemical parameters of soil (6 main limiting factors of growth) and cotton yield. In this method $EC_b$[1], organic carbon, total nitrogen and the exchange capacity of cation entered in the equation and totally made clear 42% of the variations of plant yield.

Jiang and Thelen (2004) investigated the effect of physico-chemical parameters of soil on the yield of corn. According to the results of the stepwise regression, clay with explanation of 32% of the variations of plant yield in the first year and 20% of the variations of soybean yield in the second year, entered in the equation. Furthermore, Stone et al. (1985), Miller et al. (1988) and Wright et al. (1990) achieved a negative correlation between the percentage of clay and plant yield that was obtained due to the slowdown of drainage in the root zone. In the second year, acidity, very fine sand and potassium entered in the equation, respectively and justified 85% of the variations of corn yield. Ezrin et al. (2010) examined through stepwise regression the relation between the yield of rice and $EC_a$[2]. Althogh

1.Bulk electrical conductivity
2.Apparent soil electrical conductivity

$EC_{ad}$[1] is more important than $EC_{as}$[2], results showed that both variables, $EC_{ad}$ and $EC_{as}$, are required to produce the models of linear regression. Mohamed (1999) in his research, by using stepwise regression entered spike length, number of spikes, number of grains per spike, spike weight and straw weight and recognized these traits effectively on grain yield. However, in the studies of Leilah and Al-khateeb (2005), the traits of spike length, number of spikes per $m^2$, grain weight per spike, harvest index and biological yield entered the model of stepwise regression. Harvest index and biological yield were common in both researches of Mohammad (1999) and Leilah and Al-Khateeb (2005).

*Factors analysis*

Factors analysis was investigated in order to classifying the traits, determining the order of traits importance and the relation between each of them in creating the variation of all the data. This investigation in 2 cases was considered: i) including the grain yield and ii) without the grain yield. Damania and Jackson (1986) in factors analysis did not consider the grain yield. While in factors analysis most researchers emphasized on estimating the yield with including it in the other traits (Bramel et al., 1984).

Based on the results, according to Tables 3, standard of KMO is less than 0.5, (KMO=0.392 for the first case and KMO=0.36 for the second case). It stated that the value of correlation for primary variations for performing factors analysis is insufficient. Therefore, in these conditions factors analysis is not a suitable tool for analysis and we disregarded to mention the other results of this analysis; also non-coincidence of available parameters in each group and as a result according to Table 4 the unsuitable classification confirmed the truth of this finding. Mallarino et al. (1999) in studying the relation between the yield of corn and the variables of soil and the plant, used factors analysis for classification of some measured variations and presented 3 hidden factors that are effective on the yield of plant including the conditions of primary growth of the plant, controlling of weed and fertility of soil. Also Ayoubie et al. (2009) by using factors analysis recognized 4 hidden factors that are effective on the yield of grain and biomass of barely including salt and Na of soil, fertility of soil and soil available water.

---

1.Deep ECa
2.Shallow ECa

Table 3. KMO and Bartlett's Test of variable by the factor analysis.

| | | With grain yield (Y) | Without grain yield (Y) |
|---|---|---|---|
| Kaiser-Meyer-Olkin Measure of sampling adequacy | | 0.392 | 0.36 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 773.293 | 682.247 |
| | df | 351 | 325 |
| | Sig. | 0.000 | 0.000 |

Table 4. Summary of factors loading of variable by the factor analysis.

| With grain yield (Y) | | Without grain yield (Y) | |
|---|---|---|---|
| Variables | Loading | Variables | Loading |
| Factor 1 | 4.391 | Factor 1 | 3.488 |
| Number of non fertile tiller per plant ($X_3$) | 0.500 | EC(dS/m) ($x_{12}$) | 0.560 |
| 1000-grain weight (g) ($X_5$) | -0.765 | K(ppm) ($x_{15}$) | 0.895 |
| Number of spikes/m$^2$ ($X_6$) | 0.826 | P(ppm) ($x_{22}$) | 0.794 |
| Biological yield (g/m$^2$) ($X_9$) | 0.848 | $K_{ava}$ (ppm) ($x_{23}$) | 0.795 |
| Harvest index ($X_{10}$) | -0.585 | | 3.436 |
| Na(ppm) (14) | 0.472 | Factor 2 | 0.704 |
| Grain yield (g/m$^2$) (Y) | 0.737 | Number of non fertile tiller per plant ($X_3$) | -0.672 |
| | | 1000-grain weight (g) ($X_5$) | |
| Factor 2 | 3.980 | Number of spikes/m$^2$ ($X_6$) | 0.664 |
| K(ppm) ($x_{15}$) | 0.862 | Biological yield (g/m$^2$) ($X_9$) | 0.689 |
| P(ppm) ($x_{22}$) | 0.787 | Harvest index ($X_{10}$) | -0.662 |
| $K_{ava}$ (ppm) ($x_{23}$) | 0.861 | Na(ppm) ($X_{14}$) | 0.619 |
| Clay(%) ($x_{26}$) | 0.643 | | 2.893 |
| | | Factor 3 | |
| Factor 3 | 2.566 | EC(dS/m) ($x_{12}$) | 0.560 |
| EC(dS/m) ($x_{12}$) | 0.598 | $HCO_3^-$ (meq/l) ($x_{17}$) | 0.876 |
| $HCO_3^-$ (meq/l) ($x_{17}$) | 0.895 | TN(%) ($x_{21}$) | 0.599 |
| TN(%) ($x_{21}$) | 0.513 | Sand(%) ($x_{24}$) | 0.778 |
| Sand(%) ($x_{24}$) | 0.677 | Clay(%) ($x_{26}$) | -0.526 |
| Factor 4 | 2.447 | Factor 4 | 2.809 |
| Lengths of spike (cm) ($x_4$) | 0.900 | Number of stems per plant ($X_2$) | 0.640 |
| Number of grains/spike ($X_7$) | 0.926 | $Cl^-$ (meq/l) ($x_{18}$) | 0.891 |
| | | Silt(%) ($x_{25}$) | |

Continue Table 4.

| With grain yield (Y) | | Without grain yield (Y) | |
|---|---|---|---|
| Variables | Loading | Variables | Loading |
| Factor 5 | 2.444 | Clay(%) ($x_{26}$) | 0.760 |
| Number of stems per plant ($X_2$) | 0.731 | | -0.501 |
| $Cl^-$ (meq/l) ($x_{18}$) | 0.827 | Factor 5 | 2.624 |
| Silt(%) ($x_{25}$) | 0.622 | Lengths of spike (cm) ($x_4$) | 0.929 |
| | | Number of grains/spike ($X_7$) | |
| Factor 6 | 2.270 | | 0.944 |
| Number of spikes per plant ($X_1$) | -0.549 | Factor 6 | 2.259 |
| $Ca^{2+}$+ $Mg^{2+}$(meq/l) ($x_{19}$) | 0.880 | Number of spikes per plant ($X_1$) | -0.521 |
| $SO_4^{2-}$ (meq/l) ($x_{20}$) | 0.655 | $Ca^{2+}$+ $Mg^{2+}$(meq/l) ($x_{19}$) | 0.904 |
| | | $SO_4^{2-}$ (meq/l) ($x_{20}$) | |
| Factor 7 | 2.040 | | 0.728 |
| Weight of grain/spike (g) ($X_8$) | -0.773 | Factor 7 | 1.813 |
| pH ($x_{11}$) | 0.696 | Weight of grain/spike (g) ($X_8$) | -0.794 |
| | | pH ($x_{11}$) | |
| Factor 8 | 1.641 | | 0.636 |
| SP(%) ($x_{13}$) | 0.536 | Factor 8 | 1.702 |
| OM(%) ($x_{16}$) | 0.864 | SP(%) ($x_{13}$) | 0.551 |
| | | OM(%) ($x_{16}$) | 0.868 |

## *Principal components analysis*

The purpose of principal components analysis is analyzing the available variance in multiple variables data into different components so that the first component would be the cause of the most available variance in data and the next components justify slowly less variance of the variations. In this method, each component is independent of the other components, i.e., there is no correlation between the resulting components.

With due regard to the Eigen value above 1.0, according to Table 5 and also the choice of components, according to Figure 1, including the highest slope of graph, 5 principal components were selected. Based on this selection we can explain principal component analysis in 5 groups or elements having the most variance.

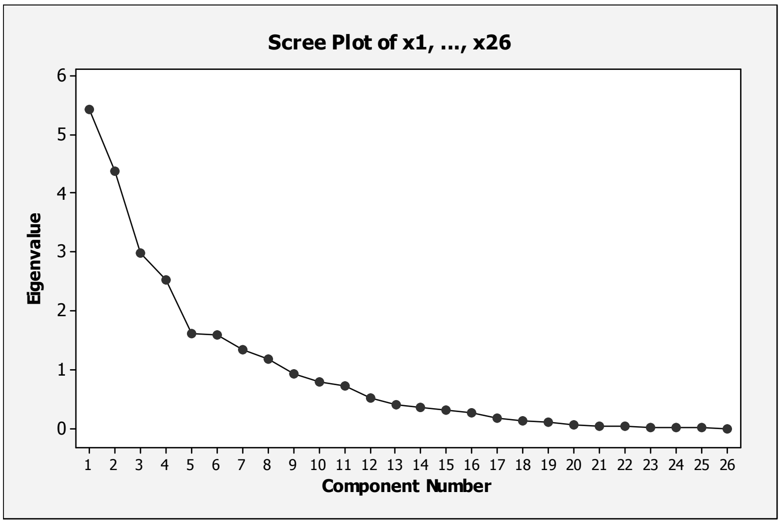


Figure 1. Scree plot showing eigen values in response to number of components for the estimated variables of wheat.

Based on these results, 5 principal components justify about 65% of the variations. In general, an important feature of each major component is its variance. Principal components are often ordered by decreasing of variance. Each component is a linear combination of all the variations with different weights. To select important features for the interpretation of each component, according to the SC (the standard of choice increasing the amount of that from the first component till the last component), weights are extracted that in each component the amount of weight ABS would be more than the amount of SC. For example, in the first component, the variables of number of spikes per $m^2$, biological yield, acidity, the percentage of saturation, sodium, chloride, phosphorus, available potassium, percentage of sand, silt and clay have greater weight of the selection standard for this component (0.21) and therefore, other variables are not important in the interpretation and are not presented in the Table 5. When the Eigen value decreased, the numbers of important traits in the component also declined as the first two components have more traits. It is shown that the first principal component ($PC_1$), the percentage of saturation, phosphorus, available potassium and the percentage of clay have the highest positive weights and number of spikes per $m^2$, biological yield, acidity, sodium, chloride , the percentage of sand and silt have the highest negative weights which this component included 20.9% of the total variance. In the second principal component ($PC_2$) the highest loading is related to spike length, number of spikes per $m^2$, biological yield, soluble potassium, phosphorus and available potassium with the highest positive weight and 1000-grain weight and harvest index with the largest negative weight. This factor includes 16.8% of the total variance. The third principal component ($PC_3$), includes the number of tillers per plant and chloride with the largest positive weight and electrical conductivity, bicarbonate and the percentage of sand with the largest negative weight and has 11.5% of the total variations. The fourth principal element ($PC_4$), included the electrical conductivity with the most positive weight and the ions of calcium, magnesium and sulfate with the highest negative weight. This element explained 7.9% of the total variance. The fifth principal component ($PC_5$) included spike length and number of grains per spike with the most positive weight and have 6.2% of all the variations that totally these five components justify 65% of the total variance. By using this method we can say that the most important parameters affecting on grain yield are: number of grains per spike, spike length, number of tillers per plant, soluble K, the percentage of clay, biological yield, chloride and number of spikes per $m^2$ with positive coefficient and sulfate, the percentage of sand, Ca and Mg, electrical conductivity and carbonate with negative coefficient. The positive coefficient for chloride and the percentage of clay and also the negative coefficient of sulfate are not explicable and the available parameters in each group do not fit together.

In the study of Vrindts et al. (2003) the first three factors justified 85% of all the yield variance of winter wheat and the first factor included general texture variations, the second factor included organic carbon and nitrogen distribution in spring and the third factor included distribution of phosphorus, pH of soil and calcium.

The achieved results from principal components analysis in the study of Jiang and Thelen (2004) in two fields of corn demonstrated that in the first field 5 chief parts and in the second field 4 chief parts justified, respectively, 89 and 88% of variations of corn yield. In both fields, $PC_1$ is representative of soil texture variable including sand, fine sand and silt and in $PC_2$ chemical and fertility variables of soil such as acidity, total bases, base saturation and magnesium concentration are effective factors on the yield. Richardson and Bigler (1984) in their study on wet soils reported that $PC_1$

covers chemical potentiality, $PC_2$ covers organic C and $PC_3$ covers soil texture. The difference between two recent studies indicated that the results of principal component analysis on soil parameters depend on conditions of soil and location and special processes of soil forming. In the first field, the percentage of clay in $PC_3$, the percentage of potassium in $PC_4$ and the percentage of very fine sand in $PC_5$ and in the second field the percentage of very fine sand and electrical conductivity in $PC_3$ and the percentage of coarse sand and horizon thickness of soil in $PC_4$ are the most important variables affecting on the yield of corn.

Leilah and Al-Khateeb (2005), in their study on the genotypes of wheat under drought conditions achieved three major factors including 74.4% of the total variations of variance, which according to special amounts in the principal components, harvest index, biological yield, diameter of spike, the number of spikes per $m^2$, length of spike, weight of grain per spike, 1000-grain weight, were the most important parameters affecting on the yield.

## *Cluster analysis*

In order to classify the studied traits, we used the method of aggregation hierarchical called cluster analysis. It is shown in Figure 2, 26 variables were laid in 4 different groups. In the first group the number of spikes per plant, acidity, number of tillers spikes without per plant, Na, number of tillers per plant, chloride and the percentage of silt; in the second cluster, spike length, number of grains per spike, number of spikes per $m^2$, biological yield, the yield, soil calcium, magnesium and sulfate; in the third cluster, 1000-grain weight, harvest index, electrical conductivity, bicarbonate, the percentage of sand and total soil nitrogen; and in the forth cluster, weight of grain per spike, the percentage of saturation, clay, organic carbon, soluble potassium, available potassium and phosphorus are categorized. Classifying by clustering technique is performed based on similarities and distances. As the level of similarity increased, the number of clusters increased too. According to Table 7 and Figure 2, it can be concluded that spike length, number of grains per spike, number of spikes per $m^2$, biological yield, soluble potassium, available potassium and phosphorus have the highest level of similarity and these traits have the greatest effect on the yield.

Table 5. Eigen value of the correlation matrix of the variables using the principal component procedure.

| Variable | $PC_1$ | $PC_2$ | $PC_3$ | $PC_4$ | $PC_5$ |
|---|---|---|---|---|---|
| Number of spikes per plant ($X_1$) | -0.17 | 0.228 | 0.151 | 0.305 | 0.27 |
| Number of stems per plant ($X_2$) | -0.094 | 0.004 | 0.379 | 0.116 | 0.258 |
| Number of non fertile tiller per plant ($X_3$) | -0.09 | 0.232 | 0.136 | 0.149 | -0.243 |
| Lengths of spike (cm) ($x_4$) | -0.16 | 0.27 | 0.005 | -0.244 | 0.423 |
| 1000-grain weight (g) ($X_5$) | 0.117 | -0.268 | -0.072 | -0.022 | 0.19 |
| Number of spikes/m$^2$ ($X_6$) | -0.215 | 0.3 | -0.143 | -0.018 | -0.148 |
| Number of grains/spike ($X_7$) | -0.198 | 0.172 | 0.01 | -0.284 | 0.467 |
| Weight of grain/spike (g) ($X_8$) | 0.115 | 0.05 | 0.067 | -0.097 | 0.068 |
| Biological yield (g/m$^2$) ($X_9$) | -0.233 | 0.338 | -0.081 | -0.09 | -0.074 |
| Harvest index ($X_{10}$) | 0.177 | -0.29 | -0.048 | 0.045 | 0.13 |
| pH ($x_{11}$) | -0.239 | 0.039 | -0.026 | 0.261 | -0.025 |
| EC (dS/m) ($x_{12}$) | 0.024 | 0.13 | -0.382 | 0.334 | 0.174 |
| SP (%) ($x_{13}$) | 0.272 | 0.196 | 0.031 | -0.178 | 0.067 |
| Na (ppm) ($x_{14}$) | -0.229 | -0.005 | -0.147 | 0.069 | -0.355 |
| K (ppm) ($x_{15}$) | 0.19 | 0.368 | -0.041 | 0.092 | -0.021 |
| OM (%) ($x_{16}$) | 0.114 | 0.224 | 0.012 | 0.072 | -0.177 |
| $HCO_3^-$ (meq/l) ($x_{17}$) | -0.076 | -0.075 | -0.332 | 0.245 | 0.144 |
| $Cl^-$ (meq/l) ($x_{18}$) | -0.236 | -0.045 | 0.312 | -0.055 | -0.041 |
| $Ca^{2+}$+ $Mg^{2+}$(meq/l) ($x_{19}$) | 0.025 | 0.042 | -0.228 | -0.386 | -0.226 |
| $SO_4^{2-}$ (meq/l) ($x_{20}$) | -0.136 | 0.056 | -0.241 | -0.418 | -0.034 |
| TN (%) ($x_{21}$) | -0.161 | -0.063 | -0.182 | 0.187 | 0.09 |
| P (ppm) ($x_{22}$) | 0.233 | 0.268 | -0.061 | 0.242 | 0.003 |
| $K_{ava}$ (ppm) ($x_{23}$) | 0.279 | 0.28 | -0.052 | -0.023 | 0.001 |
| Sand (%) ($x_{24}$) | -0.218 | -0.118 | -0.407 | 0.004 | 0.126 |
| Silt (%) ($x_{25}$) | -0.278 | 0.001 | 0.285 | 0.011 | -0.157 |
| Clay (%) ($x_{26}$) | 0.367 | 0.075 | 0.04 | -0.006 | 0.055 |
| | | | | | |
| Eigen value | 5.4255 | 4.3663 | 2.9774 | 2.517 | 1.6192 |
| Proportion | 0.209 | 0.168 | 0.115 | 0.097 | 0.062 |
| Cumulative | 0.209 | 0.377 | 0.491 | 0.588 | 0.65 |
| SC | 0.215 | 0.239 | 0.290 | 0.315 | 0.393 |

Table 6. Similarity Number of clusters and distance levels of the variables using the hierarchical cluster analysis.

| Step | Number of clusters | Similarity level | Distance level |
|---|---|---|---|
| 1 | 26 | 95.408 | 0.09185 |
| 2 | 25 | 94.453 | 0.11094 |
| 3 | 24 | 93.299 | 0.13402 |
| 4 | 23 | 90.947 | 0.18107 |
| 5 | 22 | 85.045 | 0.2991 |
| 6 | 21 | 80.574 | 0.38852 |
| 7 | 20 | 79.408 | 0.41185 |
| 8 | 19 | 77.603 | 0.44794 |
| 9 | 18 | 77.025 | 0.45949 |
| 10 | 17 | 74.578 | 0.50843 |
| 11 | 16 | 70.593 | 0.58814 |
| 12 | 15 | 70.264 | 0.59472 |
| 13 | 14 | 67.915 | 0.64169 |
| 14 | 13 | 64.035 | 0.71931 |
| 15 | 12 | 62.508 | 0.74983 |
| 16 | 11 | 60.723 | 0.78554 |
| 17 | 10 | 58.881 | 0.82239 |
| 18 | 9 | 54.168 | 0.91664 |
| 19 | 8 | 44.435 | 1.1113 |
| 20 | 7 | 40.282 | 1.19436 |
| 21 | 6 | 32.065 | 1.3587 |
| 22 | 5 | 28.661 | 1.42677 |
| 23 | 4 | 25.401 | 1.49198 |
| 24 | 3 | -19.06 | 2.38119 |
| 25 | 2 | -53.8 | 3.07601 |
| 26 | 1 | -111.3 | 4.22599 |

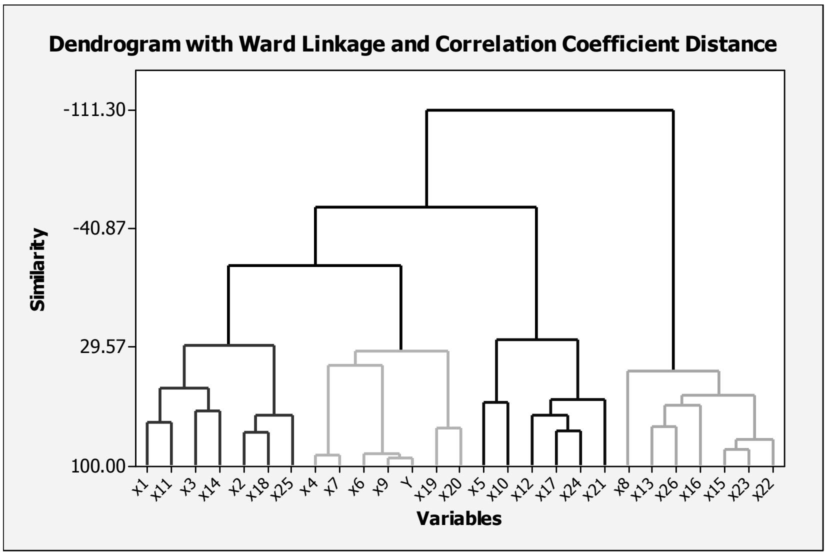


Figure 2. Similarity levels of the estimated 26 characteristics of wheat and soil (variables) using the cluster analysis.

Table 7. Path coefficient (direct and indirect effects) of the variables on grain yield variation of wheat.

| Variables | Path coefficient | | Total correlation |
|---|---|---|---|
| | Direct effect | Indirect effect | |
| Lengths of spike (cm) ($x_4$) | -0.067 | 0.504 | 0.437 |
| 1000-grain weight (g) ($X_5$) | -0.008 | -0.47 | -0.478 |
| Number of spikes/m$^2$ ($X_6$) | -0.3 | 1.183 | 0.883 |
| Number of grains/spike ($X_7$) | -0.005 | 0.451 | 0.446 |
| Biological yield (g/m$^2$) ($X_9$) | 1.275 | -0.368 | 0.907 |
| Harvest index ($X_{10}$) | 0.455 | -0.749 | -0.294 |
| $SO_4^{2-}$ (meq/l) ($x_{20}$) | 0.018 | 0.356 | 0.374 |
| Sand (%) ($x_{24}$) | 0.552 | -0.371 | 0.181 |
| Silt (%) ($x_{25}$) | 0.593 | -0.384 | 0.209 |
| Clay (%) ($x_{26}$) | 0.848 | -1.152 | -0.304 |

Residual effect=0.075.

## *Path analysis*

The correlation coefficient does not specify the relationship between traits; therefore analysis of path coefficients is a method for separating the correlation coefficients into direct effects of traits and the indirect through other traits and it can provide useful information in the manner of effectiveness of the traits on each other and their relationships (Dewey and Lu, 1959).

The purpose of conducting analysis of path coefficients is determining the parts from the yield and characteristics of soil that have some traits such as correlation and high direct effect with the yield and also these traits have minimum of negative indirect effects through other traits on the yield; therefore, these traits could be used as the standard traits for selection tool in plant breeding program. The results of path analysis have been brought to investigate more closely the relationships between the traits and the yield in Tables 8 and 9. Path coefficients for the direct effect in Table 8 demonstrated that biological yield, the percentage of clay, the percentage of silt, the percentage of sand and harvest index have respectively, the highest positive and direct effect on the yield. Soil texture itself is a qualitative factor that can not be inserted in the models; therefore, the percentages of soil particles are used. Furthermore, the highest positive and indirect effect is respectively related to the number of spikes per $m^2$ and number of grains per spike and the highest indirect and negative effect is related to the percentage of clay, harvest index and 1000-grains weight.

Based on Table 9, spike length has the highest positive and indirect effect through biological yield and the highest negative and indirect effect through harvest index and clay on the yield and finally this trait has positive and significant correlation (0.437) with the grain yield.

1000-grain weight has the highest positive and indirect effect through harvest index and the percentage of clay and the highest negative and indirect effect through biological yield and the percentage of silt and finally by its high negative and indirect effect and its insignificant direct effect, it has negative and significant correlation (-0.478) with the grain yield.

Number of spikes per $m^2$ has the highest positive and indirect effect through biological yield and the percentage of sand and silt and the highest negative and indirect effect through the harvest index and the percentage of clay on the grain yield. Considering that the negative and indirect effect is lower than the positive and indirect effect, as a result the number of spikes per $m^2$ has the positive correlation (0.883) with the grain yield and its effect on the grain yield is high.

Number of grains per spike has the highest positive and indirect effect through biological yield and the percentage of sand and highest negative and indirect effect through harvest index and the percentage of clay on the yield and it has the effect of positive and significant (0.446) with the yield.

Table 8. Path coefficient analysis showing direct and indirect effects of the variables on grain yield of wheat.

| Variables | Effect | | | | |
|---|---|---|---|---|---|
| | $X_9$ | $X_{10}$ | $X_{24}$ | $X_{25}$ | $X_{26}$ |
| Number of spikes per plant ($X_1$) | 0.493 | -0.189 | -0.04 | 0.173 | -0.15 |
| Number of stems per plant ($X_2$) | -0.007 | -0.022 | -0.136 | 0.188 | -0.06 |
| Number of non fertile tiller per plant ($X_3$) | 0.423 | -0.259 | -0.104 | 0.093 | -0 |
| Lengths of spike (cm) ($x_4$) | 0.716 | -0.217 | 0.072 | 0.056 | -0.14 |
| 1000-grain weight (g) ($X_5$) | -0.661 | 0.113 | 0.032 | -0.219 | 0.229 |
| Number of spikes/$m^2$ ($X_6$) | 1.109 | -0.191 | 0.109 | 0.156 | -0.3 |
| Number of grains/spike ($X_7$) | 0.603 | -0.129 | 0.139 | 0.076 | -0.23 |
| Weight of grain/spike (g) ($X_8$) | -0.036 | 0.044 | -0.062 | -0.059 | 0.128 |
| Biological yield (g/$m^2$) ($X_9$) | 1.275 | -0.284 | 0.075 | 0.175 | -0.29 |
| Harvest index ($X_{10}$) | -0.795 | 0.455 | 0.009 | -0.183 | 0.197 |
| pH ($x_{11}$) | 0.339 | -0.141 | 0.094 | 0.16 | -0.28 |
| EC (dS/m) ($x_{12}$) | 0.193 | -0.013 | 0.211 | -0.248 | 0.077 |
| SP (%) ($x_{13}$) | -0.035 | -0.01 | -0.215 | -0.212 | 0.458 |
| Na (ppm) ($x_{14}$) | 0.353 | -0.097 | 0.253 | 0.051 | -0.31 |
| K (ppm) ($x_{15}$) | 0.364 | -0.114 | -0.2 | -0.116 | 0.331 |
| OM (%) ($x_{16}$) | 0.238 | -0.1 | -0.103 | -0.084 | 0.203 |
| $HCO_3^-$ (meq/l) ($x_{17}$) | 0.051 | 0.068 | 0.324 | -0.024 | -0.31 |
| $Cl^-$ (meq/l) ($x_{18}$) | 0.281 | -0.048 | -0.031 | 0.359 | -0.39 |
| $Ca^{2+}$+ $Mg^{2+}$(meq/l) ($x_{19}$) | 0.265 | -0.039 | 0.074 | -0.092 | 0.02 |
| $SO_4^{2-}$ (meq/l) ($x_{20}$) | 0.469 | -0.113 | 0.194 | -0.017 | -0.18 |
| TN (%) ($x_{21}$) | 0.175 | -0.003 | 0.232 | 0.063 | -0.3 |
| P (ppm) ($x_{22}$) | 0.098 | 0.009 | -0.166 | -0.196 | 0.391 |
| $K_{ava}$ (ppm) ($x_{23}$) | 0.012 | -0.022 | -0.225 | -0.256 | 0.547 |
| Sand (%) ($x_{24}$) | 0.173 | 0.007 | 0.552 | -0.048 | -0.5 |
| Silt (%) ($x_{25}$) | 0.377 | -0.141 | -0.045 | 0.593 | -0.64 |
| Clay (%) ($x_{26}$) | -0.433 | 0.106 | -0.325 | -0.451 | 0.848 |

Table 9. Variables identified as crucial in wheat grain yield with each one of the used statistical methods.

| Variables | Simple correlation analysis | Multiple linear regression analysis | Stepwise multiple linear regression analysis | Factor analysis | Principal component analysis | Cluster analysis | Path coefficient analysis |
|---|---|---|---|---|---|---|---|
| Number of spikes per plant ($X_1$) | | | | † | | | |
| Number of stems per plant ($X_2$) | | | | ‡ | ‡ | | |
| Number of non fertile tiller per plant ($X_3$) | | † | † | ‡ | | | |
| Lengths of spike (cm) ($x_4$) | ‡ | | | ‡ | ‡ | ‡ | ‡ |
| 1000-grain weight (g) ($X_5$) | † | | | † | | | † |
| Number of spikes/m$^2$ ($X_6$) | ‡ | | | ‡ | ‡ | ‡ | ‡ |
| Number of grains/spike ($X_7$) | ‡ | | | ‡ | ‡ | ‡ | ‡ |
| Weight of grain/spike (g) ($X_8$) | | | | † | | | |
| Biological yield (g/m$^2$) ($X_9$) | ‡ | ‡ | ‡ | ‡ | ‡ | ‡ | ‡ |
| Harvest index ($X_{10}$) | | ‡ | ‡ | † | | | |
| pH ($x_{11}$) | | | | ‡ | | | |
| EC (dS/m) ($x_{12}$) | | | | ‡ | † | | |
| SP (%) ($x_{13}$) | | | | ‡ | | | |

Continue Table 9.

| Variables | Simple correlation analysis | Multiple linear regression analysis | Stepwise multiple linear regression analysis | Factor analysis | Principal component analysis | Cluster analysis | Path coefficient analysis |
|---|---|---|---|---|---|---|---|
| Na (ppm) ($x_{14}$) | | | | ‡ | | | |
| K (ppm) ($x_{15}$) | | ‡ | | ‡ | ‡ | ‡ | |
| OM (%) ($x_{16}$) | | | | ‡ | | | |
| $HCO_3^-$ (meq/l) ($x_{17}$) | | | | ‡ | † | | |
| $Cl^-$ (meq/l) ($x_{18}$) | | | | ‡ | ‡ | | |
| $Ca^{2+}$+ $Mg^{2+}$(meq/l) ($x_{19}$) | | | | ‡ | † | | |
| $SO_4^{2-}$ (meq/l) ($x_{20}$) | ‡ | | | ‡ | † | | ‡ |
| TN (%) ($x_{21}$) | | | | ‡ | | | |
| P (ppm) ($x_{22}$) | | † | | ‡ | | ‡ | |
| $K_{ava}$ (ppm) ($x_{23}$) | | | | ‡ | | ‡ | |
| Sand (%) ($x_{24}$) | | | | ‡ | † | | |
| Silt (%) ($x_{25}$) | | | | ‡ | | | |
| Clay (%) ($x_{26}$) | † | | | ‡ | ‡ | | † |

† Correlation is negative
‡ Correlation is positive

Biological yield has the highest positive and indirect effect through the percentage of silt and the highest negative and indirect effect through harvest index and the percentage of clay on the grain yield. Considering that the positive and direct effect is higher than the negative and indirect effect, biological yield has the positive correlation effect (0.907) on the grain yield and also its effect on the yield is high.

Sulfate has the highest positive and indirect effect through biological yield, phosphorus and the percentage of sand and the highest negative and indirect effect through harvest index and the percentage of clay on the yield and finally sulfate has positive and significant correlation effect (0.374) on the grain yield.

The percentage of clay has the highest positive and indirect effect through harvest index and the highest negative and indirect effect through biological yield, phosphorus and the percentage of sand and silt on the yield. Considering the high negative and indirect effect, this trait has the negative and significant correlation (-0.304) with the yield.

The existence of residual effects is representative of other effective traits on the yield of grain. According to the insignificance of this residual (r=0.075) we can realize the sufficiency of studied variables in this investigation.

The results of this method demonstrated that biological yield, number of spikes per $m^2$, number of grains per spike, spike length and sulfate with positive coefficients and 1000-grain weight and clay with negative coefficients (respectively with the most correlation), are effective traits on the grain yield.

Chung et al. (2005) by calculating path coefficients showed that calcium has the highest direct effect (0.52) and magnesium has the highest indirect effect (-0.37) on the yield of rice. In the upland cropping condition, height and slope of land has negative and direct effect (-0.28 and -0.2) on the yield of corn as well. Kozak et al. (2007) studied the relationship between the yield of rye with available potassium, exchangeable potassium and total nitrogen in the form of path analysis. The results showed that total nitrogen has the highest direct effect (0.685) on the yield that leads to a strong correlation (0.76) between the characteristic of soil and the yield. Also exchangeable K has a negative and direct effect.

Talebi et al. (2010) by using the analysis of path coefficient, studied the effects of yield components and some agronomic traits of wheat genotypes under water stress and irrigation conditions. In non-stress conditions, plant biomass, number of grains per spike, the yield of spike (g/plant) and harvest index has the highest positive and direct effect on the yield and the yield of spike has the highest indirect effect through plant biomass and number of grains per spike on the yield. Under water stress conditions, spike length, the yield of spike and harvest index have negative and direct effect on the yield and plant biomass has the highest positive and direct effect on the yield.

Zaeifizadeh et al. (2011) performed path analysis between the yield of wheat and yield components and introduced the highest positive and direct effect on the yield through minimum of number of fertile tillers and spike length and the highest indirect effect through the weight of fertile tillers in the plant and minimum number of fertile tillers.

In the study of Nofouzi et al. (2008) by using this method, the highest positive and direct effect on the yield was related to the number of grains per spike, spike length and 1000-grain. Yagdi (2009) in his study concluded that 1000-grain weight and weight of grain per spike have the highest positive and direct effect on the yield. Aycicek and Yildirim (2006) by using this method realized the positive and direct effect of plant height and weight of grain per spike.

## Conclusions

In this research, using 7 statistical methods resulted in determining the effective traits on the yield of wheat in Bajgah, classifying and distinguishing the relationships between traits; also it represented the importance of traits and the effect of them on the yield of grain. Studying coefficients of correlation between different traits allowed to decide more exactly about the selection of indirect standard traits and removing the ineffective traits. Also, the usage of regression, resulted in the equation between the yield and components that these components have a significant effect on the yield. However, using multivariate methods such as factors analyses to identify the independent factors that separately are effective and important plant characteristics, are very important for improving the plant breeding program. We used the method of gathering hierarchical order named cluster analysis as well to classify the studied traits. By path analysis it can be determined the direct effect of each component of the yield from the indirect effects that is achieved by reciprocal relations between its components.

The results of this study, according to Table 9 are listed as follows:

Among statistical analysis, multivariate regression represents higher acceptable results. In this method, among studied traits number of the tillers without spike per plant, biological yield, harvest index, soluble potassium and phosphorous justify 98.3% of yield variations and have a significant effect on the yield and the other variations at the probability of 5% are not significant and the final equation for the yield is as follows:

$$Y = -104.518 - 6.421X_3 + 0.316X_9 + 294.504X_{10} + 4.01X_{15} - 0.979X_{22} \quad (4)$$

$$R^2 = 0.983,\ SE = 5.399,\ N = 30,\ P = 0.001$$

Where Y is the grain yield in g per harvest area (1.25 $m^2$), ($x_3$) is the number of tillers without spike per plant, ($x_9$) is the biological yield, ($x_{10}$) is the harvest index, ($x_{15}$) is the soluble potassium and ($x_{22}$) is the phosphorous.

No effect of nitrogen on grain yield in the regression equation is due to the rain-fed conditions for wheat. Because in these conditions, plant growth is low and it needs less nitrogen. Also, the negative effect of phosphorus on grain yield may be due to the inverse relationship between phosphorous and the micro-nutrient elements of the plant.

Since the role of soil elements in plant growth is undeniable and nitrogen, potassium and phosphorus are the macro-nutrient elements of the plant, so using the method of multiple regression in order to studying and classifying the physico-chemical properties of soil on wheat grain yield is more suitable. In contrary to the method of stepwise regression, the parameters of soil (soluble K and P) as independent variables that are effective on wheat yield besides the yield components (number of the stems without spikes per plant, biological yield and harvest index) entered in the equation. Also in comparing with the other methods, it represents more logical results. Among the components of the yield, the biological yield is the most important and effective trait on the yield of grain under soil water stress conditions. Also, spike length, number of spikes per $m^2$, number of grains per spike and harvest index with positive coefficients and 1000-grain weight, weight of grain per spike and number of the tillers without spikes per plant with negative coefficients showed the highest correlation with the grain yield, that is observed in the most used statistical methods.

Based on the different statistical methods, among the soil parameters, soluble potassium, phosphorus, sulfate and available potassium with positive coefficients and clay with negative coefficient have the highest correlation with the grain yield.

Only through the principal components analysis, some characteristics like electrical conductivity (salinity), bicarbonate, calcium, magnesium and the percentage of sand showed a negative and significant correlation with the yield and due to the limited range of electrical conductivity in this study, its negative effect on the growth has not appeared in any methods.

According to the standard of KMO and the test of sphercity, the correlation of primary variables, factors analysis is not a suitable tool for the analysis and the obtained results are not justified.

Classifying of the components of the yield and the parameters of soil, performed by principle components analysis, is not suitable and the available parameters in every group are not correlated with each other.

Therefore, as it is conducted that multivariate regression for studying and classifying the effects of components of the yield and the parameters of soil on the yield of wheat is the most suitable method.

## Acknowledgment

This research supported partly by a research project funded by Grant no. 93-GR-AGR 42 of Shiraz University Research Council, Drought National Research Institute and Center of Excellence for On-Farm Water Management.

## References

Asseng, S., Turnera, N.C., Rayb, J.D., Keating, B.A., 2002. A simulation analysis that predicts the influence of physiological traits on the potential yield of wheat. Europ. J. Agron. 17, 123-141.

Aycicek, M., Yildirim, T., 2006. Path coefficient analysis of yield and yield components in bread wheat (*Triticum aestivum* L.) genotypes. Pak. J. Bot. 38, 417-424.

Ayoubi, Sh., Khormali, F., Sahrawat, K.L., 2009. Relationships of barley biomass and grain yields to soil properties within a field in the arid region: Use of factor analysis. Acta Agric. Scand. Section B-Soil Plant Sci. 59, 107-117.

Bahatt, G.M., 1973. Significance of path coefficient Analysis in determining the nature of character associtian. Euphytica. 22, 33-43.

Bremner, J.M., 1996. Nitrogen-Total, P 1082-1122. In: D.L. Sparks et al. (ed.) Methods of soil analysis, Part III, 3rd Ed., Am. Soc. Agron., Madison, WI.

Bramel, P.J., Hinnz, P.N., Green, D.E., Shibles, R.M., 1984. Use of principal factor analysis in the study of three stem termination types of soybean. Euphytica. 33, 387-400.

Bulman, P., Hunt, L.A., 1988. Relationships among tillering, spike number and grain yield in winter wheat in Ontario. Can. J. Plant Sci. 68, 583-596.

Chung, S.O., Sudden, K.A., Chung, T.C., 2005. Path analysis of factors limiting crop yield in rice paddy and upland corn fields. Biosys. Engin. 30, 45-55.

Corwin, D.L., Lesch, S.M., Shouse, P.J., Soppe, R., Ayars, J.E., 2003. Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil Electrical conductivity. Agron. J. 95, 352-364.

Damania, A.B., Jackson, M.T., 1986. An application of factor anlysis to morphological data of wheat and barely landraces from the Bheri River Valley. Nepal. Rachis. 5, 25-30.

Davatgar, N., Neishabouri, M.R., Sepaskhah, A.R., 2012. Delineation of site specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering. Geoderma, 173-174, 11-118.

Dewey, D.R., Lu, K.H., 1959. A correlation and path coefficient analysis of components of crested wheat grass seed production. Agron. J. 51, 515-518.

Ezrin, M.H., Amin, M.S.M., Anuar, A.R., Aimrun W., 2010. Relationship between rice yield and apparent electrical conductivity of paddy soils. Am. J. Appl. Sci. 7, 63-70.

Frankenberger, W.T., Tabatabai, M.A., Adriano, D.C., Doner, H.E., 1996. Bromine, chlorine and fluorine, P 833-868. In: D.L. Sparks et al. (ed.) Methods of Soil Analysis, Part III, 3rd Ed., Am. Soc. Agron., Madison, WI.

Gee, G.W., Bauder, J.W., 1986. P 404-408. Particle size of analysis, hydrometer method. In: A. Klute et al. (ed.) Method of Soil Analysis, Part I, Am. Soc. Agron., Madison, WI.

Helmake, P.A., Sparks, D.L., 1996. Lithium, sodium, potassium, rubidium and cesium, P 551-574. In: D.L. Sparks et al. (ed.) Methods of Soil Analysis, Part III, 3rd Ed., Am. Soc. Agron., Madison, WI.

Jiang, P., Thelen, K.D., 2004. Effect of soil and topographic properties on crop yield in a north-central corn-soybean cropping system. Agron. J. 96, 252-258.

Khormali, F., Ayoubi, Sh., Kanaro-Foomani, F., Fatemi, A., Hemmati, Kh., 2007. Tea yield and soil properties as affected by slope position and aspect in Lahijan area, Iran. Inter. J. Plant Prod. 1, 99-111.

Kozak, M., Kang, M.S., Stepien, M., 2007. Causal pathways when independent variables are co-related: new interpretational possibilities. Plant Soil Environ. 53, 267-275.

Kumbhar, M.B., Larik, A.S., Hafiz, H.M., Rind, M.J., 1983. Interrelationship of polygenic traits affecting grain yield in *Triticum aestivum* L. Wheat Infor. Ser. 57, 42-54.

Ledent, J.F., Moss, D.N., 1979. Relation of morphological characters and shoot yield in wheat. Crop Sci. 12, 31-34.

Leilah, A.A., Al-Khateeb, S.A., 2005. Statistical analysis of wheat yield under drought conditions. J. Arid. Environ. 61, 483-496.

Loppert, R.H., Suarze, D.L., 1996. Carbonate and gypsum, P 437-474. In: D.L. Sparks et al. (ed.) Methods of Soil Analysis, Part III, 3rd Ed., Am. Soc. Agron., Madison, WI.

Mallarino, A.P., Oyarzabal, E.S., Hinz, P.N., 1999. Interpreting within-field relationships between crop yields and soil and plant variables using factor analysis. Precision Agric. 1, 15-25.

Miller, M.P., Singer, M.J., Nielsen, D.R., 1988. Spatial variability of wheat yield and soil properties on Complex Hill. Soil Sci. Soc. Am. J. 52, 1133-1141.

Mohamed, N.A., 1999. Some statistical procedures for evaluation of the relative contribution for yield components in wheat. Zagazig J. Agric. Res. 26, 281-90.

Mohiuddin, S.H., Cory, L.I., 1980. Flag leaf and peduncle area duration in relation to winter wheat grain yield. Agron. J. 72, 299-301.

Munir, M., Chowdhry, M.A., Malik, T.A., 2007. Correlation studies among yield and its components in bread wheat under drought conditions. Int. J.Agric. Biol. 9, 287-90.

Nelson, D.W., Sommers, L.E., 1996. Salinity: Total carbon, organic carbon and organic matter, P 961-1010. In: D.L. Sparks et al. (ed.) Methods of Soil Analysis, Part III, 3rd Ed., Am. Soc. Agron., Madison, WI.

Nofouzi, F., Rashidi, V., Tainejad, A.R., 2008. Path analysis of grain yield with its components in durum wheat under drought stress. International Meeting on Soil Fertility Land Management and Agroclimatology Turkey. pp. 681-686.

Olsen, S.R., Cole, C.V., Watanabe, F.S., Deam, L.A., 1954. Estimation of available phosphorous in soil by extraction with sodium bicarbonate . USDA. Cic, 939. U.S. Gov. Print. Office, Washington, D.C.

Rhoades, J.D., 1996. Salinity: Electrical conductivity and total dissolved solids, P 417-436. In: D.L. Sparks et al. (ed.) Methods of Soil Analysis, Part III, 3rd Ed., Am. Soc. Agron., Madison, WI.

Richards, L.A., 1954. Diagnosis and improvement of saline and alkaline soils. U. S. Salinity laboratory staff. USDA. Hand book NO. 60. Washington, DC, USA.160PP.

Richardson, J.L., Bigler, R.J., 1984. Principal component analysis of prairie pothole soils in North Dakota. Soil Sci. Soc. Am. J. 48, 1350-1355.

Sepehr, E., Malakouti, M.J., Kholdebarin, B., Samadi, A., 2009. Genotypic variation in P efficiency of selected Iranian cereals in greenhouse experiment. Inter. J. Plant Prod. 3, 17-28.

Shanahan, J.F., Donnelly, K.J., Smith, D.H., Smika, D.E., 1985. Shoot developmental properties associated with grain yield in winter wheat. Crop Sci. 25, 770-775.

Srivatsava, K.N., Thakur, A.K., Sinha, T.H., 1971. Path coefficient analysis between yield and yield attributing components in a high yielding wheat variety, S 227. Indian J. Agron. 16, 406-409.

Stone, J.R., Gilliam, J.W., Cassel, D.K., Daniels, R.B., Nelson, L.A., Kleiss, H.J., 1985. Effect of erosion and landscape position on the productivity of piedmont soils. Soil Sci. Soc. Am. J. 49, 987-991.

Talebi, R., Fayyaz, F., Mohammad Naji, A., 2010. Genetic variation and interrelationships of agronomic characteristics in Durum wheat under two constructing water regimes. Braz. Arch. Biol. Technol. 53 (4), 785-791.

Thomas, G.W., 1996. Soil Ph and soil acidity, P 475-490. In: D.L. Sparks et al. (ed.) Methods of Soil Analysis, Part III, 3rd Ed., Am. Soc. Agron., Madison, WI.

Virk, D.S., Vema, M.M., 1972. Relative importance of grain yield components in bread wheat. Wheat Infor. Service. 35, 11-14.

Vrindts, E., Reyniers, M., Darius, P., Baerdemaeker, J.De., Gilot, M., Sadaoui, Y., Frankinet, M., Hanquet, B., Destain, M.F., 2003. Analysis of soil and crop properties for precision agriculture for winter wheat. Biosys. Engin. 85, 141-152.

Vrindts, E., Reyniers, M., Darius, P., Frankinet, M., Hanquet, B., Destain, M.F., Baerdemaeker, J.De., 2003. Analysis of Spatial Soil, Crop and Yield Data in a Winter Wheat Field. An ASAE Meeting Presentation. PP:031080.

Wilding, L.P., 1985. Spatial variability, its documentation, accommodation and implication to soil surveys. In: D.R. Nielsen and J. Bouma (Eds.), Soil Variability. Pudo, Wagenigen, the Netherlands.

Wright, R.J., Boyer, D.G., Winant, W.M., Perry, H.D., 1990. The influence of soil factors on yield differences among landscape positions in an Appalachian cornfield. Soil Sci. 149, 375-382.

Yagdi, K., 2009. Path coefficient analysis of some yield components in durum wheat (*Triticum aestivum* L.). Pak. J. Bot. 41, 745-751.

Yan, L., Zhou, Sh., Feng, L., Hong-Yi, L., 2007. Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land. Computers and Electronics in Agriculture. 56, 174-186.

Yap, T.C., Harvey, B.L., 1972. Inheritance of yield components and morph-physiological traits in barley. Crop Sci. 12, 283-286.

Zaeifizadeh, M., Khayatnezhad, M., ghasemi, M., azimi, J., Vahabzadeh, M., 2011. Path analysis of yield and yield c omponents in synthetic bread wheat (*Triticum aestivum* L.) genotypes. Adv. Environ. Bio. 5, 98-103.