Australian Journal of Crop Science AJCS
AJCS 4(6):402-407 (2010) ISSN:1835-2707

Determining the most important features contributing to wheat grain yield using supervised feature selection model

Ehsan Bijanzadeh, Yahya Emam, Esmaeil Ebrahimie

Crop production and Plant Breeding Department, Shiraz University, Shiraz, Iran

***Corresponding author: ebrahimie@shirazu.ac.ir**

Abstract

A supervised feature selection algorithm was applied to determine the most important features contributing to wheat grain yield. Four hundreds seventy two fields (as records) from different parts of Iran which were different in 21 characteristics (features) were selected for feature selection analysis. Selection of the wide range of features, including location, genotype, irrigation regime, fertilizers, soil textures, physiological attitudes, and morphological characters, provided the opportunity of precise simultaneous study of a large number of factors in wheat grain yield topic by hand of data mining. The grain yield of each record assumed as target variable. The feature selection algorithm selected 14 features as the most effective features on grain yield. These features included culture type, location, soil texture, 1000 kernel weight, nitrogen supply, irrigation regime, biological yield, organic content of the soil, the amount of rainfall, genotype, plant height, and spike number per unit area. Interestingly, growing season length and plant density were the second most important features for wheat grain yield. Based on the feature selection model, culture type, as dryland farming or irrigated, severely affected wheat grain yield. The soil pH had a marginal effect on wheat grain yield. The results of this investigation demonstrated that feature classification using feature selection algorithms might be a suitable option for determining the important features contributing to wheat grain yield, providing a comprehensive view about these traits. This is the first report in identifying the most important factors on wheat grain yield from many fields using feature selection model.

Keywords: culture type; data mining; plant physiology, wheat genotype, wheat grain yield

Introduction

Data mining is a process of discovering previously unknown and potentially interesting patterns in large datasets (Piatetsky-Shapiro and Frawley, 1991). Nowadays, intelligent data mining and knowledge discovery by artificial neural network, decision trees, and feature selection algorithms have become the important revolutionary issues in prediction and modeling (Roddick et al., 2001; Elson et al., 2004; Schuize et al., 2005). The 'mined' information is typically represented as a model of the semantic structure of the dataset, where the model may be used on new data for prediction or classification (Liu and Motoda, 2008). In data mining, feature selection tools are useful for identifying irrelevant attributes to be excluded from the dataset (Liu and Motoda, 2001). The main idea of the feature selection is to choose a subset of all variables by eliminating a large number of features with little discriminative and predictive information (Blum and Langley, 1997; Beltrán et al., 2005). Usually in a dataset, not all the features are important; some are redundant and some are irrelevant. Data with several irrelevant features can misguide the clustering results and make it hard to explain (Liu and Motoda, 2001 and 2008). There are two ways to reduce the dimensionality: feature transformation and feature selection. Feature transformation reduces the dimension by applying some type of linear or non-linear function on the original features whereas feature selection selects a subset of the original features. One may wish to perform feature selection rather than transformation to keep the original meaning of the features. Furthermore, after feature selection, one does not need to measure the features that are not selected. Feature transformation, on the other hand, still needs all the features to extract the reduced dimensions (Liu and Motoda, 2008). Recently, there is a great interest in employing feature selection algorithms to find the critical features involving in different phenomena including enzyme thermostability (Ebrahimi et al., 2009) and pH tolerance (Ebrahimie et al., 2008). Feature selection allows the variable set to be reduced in size, creating a more manageable set of attributes for modeling (Blum & Langley, 1997). Adding feature selection to the analytical process has several benefits: it simplifies and narrows the scope of the features that is essential in building a predictive model, minimizes the computational time and memory requirements for building a predictive model, because focus can be directed to the subset of predictors that is most essential. It also leads to more accurate and/or more parsimonious models (Dash and Liu, 1997; Liu and Motoda, 1998). Furthermore, it reduces the time for generating scores since the predictive model is based upon only a subset of predictors. Feature selection algorithms have two main components: feature search and feature subset evaluation consists in of screening, ranking and selecting (Liu and Motoda, 2008). There are two types of feature selection algorithms: supervised and unsupervised. Supervised feature selection algorithms rely on measures that take into account the class information. A well-known measure is information gain, which is widely used in both feature selection and decision tree induction (Dash and Liu, 1997). In essence,

Table 1. List of some literatures that were used for feature selection model of wheat grain yield in Iran

Authors	Province-Location	Type of treatment
Abhari et al., (2008)	Golestan- Gorgan	Drought stress
Afiuni et al., (2006)	Isfahan-Rod Dasht	Different genotypes, Salt stress
Akbari et al., (2006)	Khorasan Razavi-Mashhad	Different genotypes
Bijanzadeh et al., PhD thesis - (2008-9)	Fars-Badjgah	Different genotypes, Drought stress
Dastfal et al., (2008)	Fars-Darab	Different genotypes, Drought stress
Emam et al., (2000)	Fars- Badjgah	Different nitrogen level
Emam et al., (2009)	Fars- Badjgah	Different nitrogen level, Different genotypes
Emam, et al., (2007)	Fars- Badjgah	Different genotypes, Drought stress
Farahani and Arzani (2007)	Isfahan- Najaf Abad	Different genotypes
Faraji et al., (2006)	Khozestan- Ramin	Different nitrogen level, Drought stress
Ghodsi et al., (2005)	Khorasan Razavi-Mashhad	Drought stress
Kiani et al., (2004)	Golestan-Agh Ghala	Salt stress, Drought stress
Modhej et al., (2008)	Khozestan-Ahvaz	Different nitrogen level
Momtazi and Emam (2006)	Fars- Badjgah	Growing season length, Plant density
Moussavi Nik et al., (2006)	Sistan and Balochestan-Zabol	Different phosphorous and zinc level
Roustaii et al., (2003)	Western Azarbaijan-Maragheh	Different genotypes
Sadegh Zadeh Ahari et al., (2006)	Kohkiloyeh -Gachsaran	Different genotypes
Sadegh Zadeh Ahari et al., (2005)	Western Azarbaijan-Maragheh	Different genotypes

supervised feature selection algorithms try to find a feature helping to separate data of different classes, named class-based separation. If a feature has no effect on class-based separation, it can be removed. In contrast, a good feature should, therefore, enhance class-based separation (Liu and Motoda, 2008). In the late 90's, research on unsupervised feature selection intensified in order to deal with data without class labels (Dy and Brodley, 2004). It is closely related to unsupervised learning. One example of unsupervised learning is clustering, where similar instances are grouped together and dissimilar ones are separated apart. Similarity can be defined by the distance between two instances. Conceptually, the two instances are similar if the distance between the two is small; otherwise they are dissimilar (Liu and Motoda, 2008). When all instances are connected pair-wisely, breaking the connections between those instances that are far apart will form clusters. Hence, clustering can be thought as achieving locality-based separation. One widely used clustering algorithm is *k*-means. It is an iterative algorithm that categorizes instances into *k* clusters (Dy and Brodley, 2004; Liu and Motoda, 2008). In this study, we tried to examine the efficiency of supervised feature selection on the most critical agronomic trait, wheat grain yield. Wheat grain yield is closely associated with the field characteristics particularly soil properties, available water and nutrient content in the soil, genotype, growing season length etc. (Emam, 2007). The aim of this study was to identify the importance and relationship of wheat field conditions (features) providing a comprehensive view about wheat grain yield by feature selection algorithm. Understanding the importance of features among a large dataset of features can play a key role in improving the wheat grain yield under field conditions.

Material and methods

Data collection

One part of data was collected from field experiment conducted during 2008 growing season, at Badjgah experimental station (29° 50´ N and 52 ° 46´ E; elevation 1810 m above mean sea level), Shiraz University, Shiraz, Iran by the authors. The experiment was conducted in a split plot design with four moisture regimes (irrigation according to 125%, 100%, 75%, and 50% field capacity) as the main factor and five wheat cultivars (Shiraz, Bahar, Pishtaz, Yavaros and Sistan) as subplots in three replications. In addition, data from the field study was also extracted from the literature describing effect of field conditions on wheat grain yield in Iran (Table 1). As a result, 472 records with 21 features including location, precipitation (mm), soil texture, soil pH, culture type (dryland or irrigated), irrigation water EC (dS/m), nitrogen, phosphorus and potassium applied to the soil (kg/ha), organic content of the soil (%), growing season length (days), plant height (cm), biological yield (kg/ha), irrigation regime (according to field capacity), genotype, 1000 kernel weight (g), spike number per unit area, plant density (plant/m^2), harvest index (%), grain number per spike and wheat grain yield were prepared in Excel software sheets.

Statistical analyses

Statistical analyses for feature selection were performed using SPSS Clementine 11.1. First, data were transported from Excel software to SPSS Clementine 11.1. Wheat grain yield was set as output variable and the others as input variables. Some features such as biological yield, rainfall, and plant height were classified as continuous variables and features like location, soil type, and genotype were classified as categorical. Finally, features contributed to wheat grain yield were selected.

Feature selection algorithms

Feature selection algorithm was applied to recognize important features showing strong correlation with wheat grain yield. The algorithm considered one attribute at a time to see how well each predictor alone (feature) predicted the target variable (output). The importance value of each variable was then calculated as (1- p) where p was the p value of the appropriate test of association between the candidate predictor and the target variable. The association test for categorized output variables was different from the test for continuous ones. When target value was continuous, p values based on the F statistic were used. If some predictors were continuous and some categorical in the dataset, the criterion for continuous predictors was still based on the p value from a transformation and that for categorical predictors from the F statistic. Predictors were ranked according to sorting by p value in ascending order. If ties occurred, the rules for breaking ties were followed among all categorical and continuous predictors separately, and then these two groups

Table 2. Classification of important features affecting on wheat grain yield (as output) by feature selection method

Rank	Feature	Type	Importance	Value
1	Culture type (dryland or irrigation farming)	Flag	Important	1.0
2	Location	Set	Important	1.0
3	Soil texture	Set	Important	1.0
4	1000 kernel weight (g)	Range	Important	1.0
5	Nitrogen applied to the soil (kg/ha)	Range	Important	1.0
6	Irrigation regime (according to percentage field capacity)	Range	Important	1.0
7	Biological yield (kg/ha)	Range	Important	1.0
8	Organic content of the soil (%)	Range	Important	1.0
9	Precipitation (mm)	Range	Important	1.0
10	Genotype	Set	Important	1.0
11	Plant height (cm)	Range	Important	1.0
12	Spike number per unit area	Range	Important	1.0
13	Growing season length (day)	Range	Important	0.999
14	Plant density (plant/m^2)	Range	Important	0.997
15	Soil pH	Range	Marginal	0.909
16	Harvest index (%)	Range	Unimportant	0.854
17	Irrigation water EC (dS/m)	Range	Unimportant	0.804
18	Phosphorus applied to the soil (kg/ha)	Range	Unimportant	0.538
19	Grain number per spike	Range	Unimportant	0.335
20	Potassium applied to the soil (kg/ha)	Range	Unimportant	0.133

(categorical predictor group and continuous predictor group) were sorted by the data file order of their first predictors. The p value was based on the asymptotic t distribution of a transformation on the Pearson correlation coefficient. The predictors were then labeled as ‘important’, ‘marginal’, and ‘unimportant’ with values above 0.95, between 0.95 and 0.90, and below 0.90, respectively.

Results and discussion

Classification of features (showed that in Table 2) indicated that out of 21 features, 14 features were the most important features related to the wheat grain yield including culture type, location, soil texture, 1000 kernel weight, nitrogen applied to soil, irrigation regime (according to field capacity), biological yield, organic content of the soil, precipitation, genotype, plant height, and spike number per unit area, with a value of 1.0, and growing season length (0.999 value), and plant density (0.997 value). The soil pH feature (0.909 value) was recognized to have a marginal effect on wheat grain yield. The rest including of harvest index (0.854 value), irrigation water EC (0.804 value), phosphorus applied to the soil (0.538 value), grain number per spike (0.335 value), and potassium applied to soil (0.133 value), were found to be unimportant (Table 2). In this study, spike number per unit area and biological yield (1.0 value) had important effects on wheat grain yield (Table 2). Farahani and Arzani (2007) reported that spike number per unit area and biological yield were strongly correlated to grain yield of durum wheat. Ghodsi et al., (2005) also reported the significant correlation between grain yield and spike number per unit area. Emam et al., (2009) showed that nitrogen applied to the soil, as a key element in crop nutrition, had an important role in increasing wheat grain yield. Culture type as dryland or irrigated farming severely affected wheat grain yield (Table 2), and mean grain yield in dryland (1966 kg/ha) and irrigated farming (4000 kg/ha) was significantly different ($P<0.01$). The relationship of culture type and wheat grain yield has been reported by some researchers [e.g. Ahmadi and Sio Se Mardeh 2003; Abhari, et al., 2008; Emam et al., 2007].

Feature selection showed that one of the most important features in wheat productivity is location (Table 2). Confirming the feature selection output, results of general linear model showed that wheat grain yield in three locations including Badjgah (3228 kg/ha), Gachsaran (2088 kg/ha) and

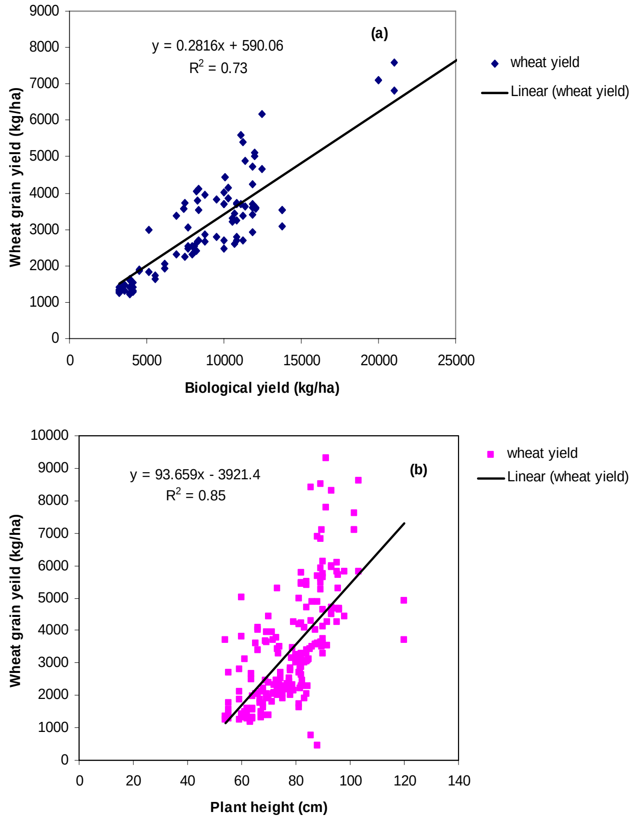

Fig 1. Relationship between wheat yield with biological yield (a) and plant height (b). Best-fit linear regression is plotted in case where the relationship was significant at $P<0.01$

Maragheh (1411 kg/ha) were significantly different under dryland farming (Table 3). In irrigated culture, many of 21 studied locations, including Agh Ghala (3362 kg/ha) with Najaf Abad (6137 kg/ha), Ramin (5333 kg/ha), and Karaj (5265 kg/ha), and Ahvaz (3427 kg/ha (with Najaf Abad (6137 kg/ha), were significantly different ($P<0.01$) in the point of wheat grain yield (Table 3). Biological yield was one of the most relevant features to the grain yield with 1.0 value (Table 2). In addition, biological yield was strongly related to

Table 3. General linear models between location and wheat grain yield under two culture type (dryland farming and irrigated). Statistics are only reported for locations with a significant relationship (P<0.01) with wheat grain yield.

Culture type	Location	Mean wheat grain yield (kg/ha)	P value
	Badjgah	3228	
	Gachsaran	2088	0.0000
Dry land farming	Marageh	1411	0.0000
	Gachsaran	2088	
	Marageh	1411	0.0000
	Agh Ghala	3362	
	Najaf Abad	6137	0.0000
	Ramin	5333	0.0001
	Karaj	5265	0.0022
	Ahvaz	3427	
	Najaf Abad	6137	0.0000
	Badjgah	4913	
	Najaf Abad	6137	0.0000
	Karaj	5265	0.0021
	Darab	3860	
	Najafabad	6137	0.0000
	Gorgan	2665	
	Mashhad	4369	0.0067
	Najafabad	6137	0.0000
Irrigation regime	Karaj	5265	0.0001
(according to field capacity)	Isfahan	3846	
	Najafabad	6137	0.0000
	Isfahan	3846	0.0000
	Zabol	3002	0.0003
	Mashhad	4369	
	Najafabad	6137	0.0000
	Isfahan	3846	0.0033
	Najafabad	6137	
	Isfahan	3846	0.0000
	Zabol	3002	0.0000
	Rod Dasht	3175	
	Karaj	5265	0.0001
	Zabol	3002	
	Karaj	5265	0.0077

wheat grain yield (R^2= 0.73, P<0.01) (Fig. 1a). With increase in biological yield from 3221 to 21053 kg/ha maximum wheat grain yield (7200kg/ha) was obtained in Najaf Abad (Fig. 1a). The positive relationship between biological yield and wheat grain yield has been demonstrated by Abhari et al., (2008) and Afiuni and Mahlouji (2006). Wheat plant height ranged from 56 to 120 cm and maximum wheat grain yield was achieved at the height of 91 cm in the genotype of Masora. Plant height (R^2= 0.85, P<0.01) was also related to maximum wheat grain yield (Fig. 1b). Genotype was another important feature with a value of 1.0 (Table 2). In this study, comparison of 176 genotypes showed that Buchen (7600 kg/ha), Ostatar (6800 kg/ha), Yavaros (5569 kg/ha), and Shiraz (5044 kg/ha) had the maximum wheat grain yield. On the other hand, wheat yield in the genotypes of Korifla and Sabalan decreased to 1499 and 1668 kg/ha under dryland farming, respectively. In one study with 42 wheat genotypes, there was a significant difference between genotypes for yield and yield components and genotype was considered to be an important factor in determining the final grain yield (Farahani and Arzani, 2007). While biological yield was an important feature in improving wheat grain yield, harvest index feature was found to be less important in modern wheat genotypes in Iran (Table 2). Sharma (1996) suggested that grain yield in wheat may be increased by improving biomass at a given level of harvest index. Austin (1984) also showed that an alternative for achieving increased grain yield is to increase the biomass produced by the crop. In comparison of 36 wheat genotypes, biological yield was found to be the most important factor giving the highest correlation (r=0.75) with wheat grain yield (Farid et al., 2007). Furthermore, improvement in harvest index appears to be difficult (Farid et al., 1996) and recently, increase in wheat grain yield has been attributed to increases in biomass production (Ahmadi and Sio Se Mardeh; 2003, Afuni and Mahlouji, 2006). It can be concluded that the modern wheat cultivars grown in Iran show variation in biomass production, and there might be a scope in improving wheat grain yield by selecting cultivars with a higher biomass.

Conclusion

The results of this study suggest that feature classification by feature selection algorithm may be a suitable option for determining the important features such as culture type, location, soil texture, water content in soil, plant height, genotype, and the nutrient content of the soil. The weak point of analyzing results of just one experiment (one field) is that commonly, the outcomes of this experiment are robust just about that specified condition. In fact, in many cases, it is not possible to extend the results of separate experiments. As a result, obtaining a comprehensive view about the critical attitudes involving in grain yield is not possible. This study also indicated feature classification by using feature selection might be a suitable option for determining the important features contributing to wheat grain yield providing a comprehensive view about this trait. Also, the main advantages of feature selection method are the reduction of the data processing time, decrement in the requirements of data storage space, decreasing in the cost of data acquirement and the most important, it allows to select a subset of the original features which contribute with the largest amount of information for a particular problem (reduction in the dimensionality of the input data). Determining of important features is useful for wheat grain yield increasing in Iran. This is the first report in identifying the most important factors on wheat grain yield from many fields in Iran using feature selection model.

Acknowledgment

We would like to greatly thank Department of Crop Production & Plant Breeding of Shiraz University for supporting this research and Dr. Noorali Abad Dadkhodaei for editing the manuscript.

References

Abhari A, Galeshi S, Latifi N, Kalateh N (2008) The Effects of some growth parameters on grain yield of wheat genotypes yield under drought stress conditions. J Agric Nat Res 14: 104-112

Afiuni D, Mahlouji M (2006) Correlation analysis of some agronomic traits in wheat (*Triticum aestivum* L.) genotypes under salinity stress. Seed and Plant 22: 186-197

Ahmadi A, Sio Se Mardeh A (2003) Relationships among growth indices, drought resistance and yield in wheat cultivars of different climates of Iran under stress and non stress conditions. Iranian J Agric Sci 34: 667-679

Akbari A, Foghi B, Admin H, Mokhtassi Bidgili A, Rahimian Mashhadi H, Zand E (2006) Investigation of morphological aspects wheat cultivars on their yield augmentation released during past years. J Agric Nat Res 13: 96-103

Austin RB (1984). Crop characteristics and the potential of wheat yield. J Agric Sci Camb 98: 447-453

Beltrán NH, Duarte-Mermoud MA, Salah SA, Bustos MA, Peña-Neira AI, Loyola EA, Jalocha JW (2005) Feature selection algorithms using Chilean wine chromatograms as examples. J Food Eng 67: 483-490

Blum AL, Langley P (1997) Selection of relevant features and examples in machine learning. Arti Intell 97: 245–271

Dash M, Liu H (1997) Feature selection methods for classifications. Intell Data Anal 3: 131–156

Dastfal M, Tavakolli A, (2008) Evaluation of grain yield, its components and stress susceptibility and tolerance indices in wheat genotypes under post anthesis drought stress. Iran Agric Res 27: 113-125

Dy J, Brodley C (2004) Feature selection for unsupervised learning. J Mach Lear Res 5: 845–889

Ebrahimie E, Ebrahimi M, Ebrahimi M, Zinati Z, Delavari A, Mohammadi-dehcheshmeh M (2008) Determining the main factors contributing to xylanase resistance to different pH by applying the asymptotic distribution of a transformation on the Pearson correlation coefficient. Third IAPR International Conference on Pattern Recognition in Bioinformatics 2008, Melbourne, Australia, October 15 – 17, 2008.

Ebrahimi M, Ebrahimie E (2009) Application of neural networks methods to define the most important features contributing to xylanase enzyme thermostability. Proceedings of the Eleventh conference on Congress on Evolutionary Computation. Pages: 2885-2891. ISBN: 978-1-4244-2958-5.

Elson A, Tailor S, Salim RB, Hillaby K, Jurkovic D (2004) Expectant management of tubal ectopic pregnancy: prediction of successful outcome using decision tree analysis. Ultra. Obstet Gyn 6 : 552–556

Emam Y (2007) Cereal Production. Shiraz University Press. Shiraz, Iran. 191 Pp.

Emam Y, Ranjbari AM, Bahrani M (2007) Evaluation of yield and yield components in wheat genotypes under Post–anthesis drought stress. J Sci Technol Nat Res 11: 317-328

Emam Y, Borjian R (2000) Yield and yield components of two winter wheat (*Triticum aestivum* L.) cultivars in response to rate and time of urea application. J Agric Sci Technol 2: 263-270

Emam Y, Salimi S, Shekofa A (2009) Effect of nitrogen level on grain yield and yield components of wheat (*Triticum aestivum* L.) under irrigation and rainfed conditions. Iranian J Field Crops Res 7: 323-334

Farahani E, Arzani A (2007) Study of genetic diversity of cultivars and F1 hybrids of durum wheat using agronomic and morphological traits. J Sci Technol. Nat Res 10: 341-355

Faraji H, Siadat A, Fathi G, Emam Y, Nadian H, Rasekh AR (2006) Effect of nitrogen on wheat yield under drought stress. J Agric Sci 29: 99-111

Farid M, Niknejad M, Emam Y (1996) Analysis of phytomass in 36 Iranian winter wheat (*Triticun aestivum* L.) genotypes. Iran Agric Res 15: 69-77

Ghodsi M, Chaeichi MR, Mazaheri D, Jalal Kamali MR (2005) Determination of susceptibility of developmental stages in bread wheat to water stress and its effect on yield and yield components. Seed and Plant 20: 489-509

Kiani AR, Mirlatifi M, Homaee M, Cheraghi A (2004) Effect of different irrigation regimes and salinity on wheat yield in Gorgan region. J Agric Nat Res 11: 79-89

Liu H, Motoda H (2008) Computational Methods of Feature Selection. Chapman & Hall/CRC Taylor & Francis Group 6000 Broken Sound Parkway NW

Liu H, Motoda H (2001) Feature Extraction, Construction and Selection: A Data Mining Perspective. Boston, Kluwer Academic Publishers

Liu H, Motoda H (1998) Feature Selection for Knowledge Discovery and Data Mining. Boston, Kluwer Academic Publishers

Modhej M, Naderi A, Emam Y, Ayneband A, Normohamadi G (2008) Effect of post-anthesis heat stress and nitrogen levels on grain yield in wheat (*Triticum durum and T. aestivum*) genotypes. Inter J Plant Prod 2: 257-267

Momtazi F, Emam Y (2006) Effect of sowing date and seeding rate on yield and yield components in winter wheat CV. Shiraz. Iranian J Agric Sci 37: 1-11

Moussavi Nik, S M, Dorji Naseri B, Valizadeh J, Safaei M (2006) Effect of phosphorous and zinc application on yield quantity and quality of wheat cultivars in Sistan. J Agric Nat Res 12: 39-48

Piatetsky-Shapiro G, Frawley WJ (1991) Knowledge Discovery in Databases, Menlo Park, CA, AAAI Press. Pp 201.

Roddick JF, Hornsby K, Spiliopoulou M (2001) An Updated Bibliography of Temporal, Spatial, and Spatio-temporal Data Mining Research. Lecture Notes in Computer Science, Springer, Berlin

Roustaii; M, Sadegh Zadeh Ahari D, Hesami A, Soleimani K, Nader Mahmodi K, Abedi Asl G (2003) Study of adaptability and stability of grain yield of bread wheat genotypes in Maragheh. Seed and Plant 19: 263-275

Sadegh Zadeh Ahari D, Hoseini SK, Hosein Pour T, Jafar Bay JA, Khalil Zadeh GR, Ali Zadeh K (2006) Study on adaptability and stability of grain yield in durum wheat lines in warm dryland areas. Seed and Plant 21: 561-576

Sadegh Zadeh Ahari D, Pasha Pour H, Bahrami S, Hagh Parast R, Aghaei M, Azimzadeh M, Abedi G (2005) Adaptability and stability of grain yield in durum wheat lines in cold dryland areas. Seed and Plant 21: 1-22

Schuize FH, Wolf H, Jansen H, Vander VP (2005) Applications of artificial neural networks in integrated water management: fiction or future? Water Sci Technol 52: 21-31

Sharma R C, Smith EL (1996) Selection for high and low harvest index in three winter wheat population. Crop Sci 26: 1147-1150