COBISS Code 1.01

DOI: 10.14720/aas.2015.105.1.11

Agrovoc descriptors: genotypes, crop yield, statistical data, statistical methods, plant breeding, models

Agris category code: u10, f30

Identification of the most stable genotypes in multi-environment trials by using nonparametric methods

Naser SABAGHNIA[1]

Received November 14, 2014; accepted February 28, 2015.
Delo je prispelo 14. novembra 2014, sprejeto 28. februarja 2015.

ABSTRACT

Genotype performances in multi-environment trials are usually analyzed by different univariate and multivariate parametric models for assessing yield stability and genotype × environment (GE) interaction investigation. One of the alternative strategies can be nonparametric statistics approach which is particularly useful in situations where parametric statistics fail. For an estimation of yield stability of genotypes in various environments two new nonparametric stability statistics ($NS_i^{(1)}$ and $NS_i^{(2)}$) have been used which are based upon the ranks of the genotypes in each environment. These statistics use median as a nonparametric central tendency, and two nonparametric index of statistical dispersion as inter-quartile range and inter-decile range. The $NS_i^{(1)}$ and $NS_i^{(2)}$ nonparametric stability statistics which presented here is similar to the nature and concept of environmental coefficient of variation. Results indicated that the most stable genotype based on the lowest values of these two nonparametric statistics, had the highest mean yield among studied genotypes. Plotting of mean yield versus $NS_i^{(1)}$ and $NS_i^{(2)}$ verified the above results and indicated that the highest mean yielding genotype is identified as the most stable genotype. These nonparametric statistics would be useful for simultaneous selection for mean yield and stability. They can be very helpful in selection for yield stability and determination of favorable genotypes in plant breeding programs.

Key words: adaptation, GE interaction, stability, yield

IZVLEČEK

DOLOČANJE NAJBOLJ STABILNIH GENOTIPOV V RAZLIČNIH OKOLJIH Z NEPARAMETRIČNIMI METODAMI

Za analizo in ocenjevanje stabilnosti pridelka in interakcije genotipov z različnim okoljem se navadno uporabljajo univariatni in multivariatni parametrični modeli. Eden izmed alternativnih pristopov bi lahko bila uporaba neparametričnih modelov, še posebej v primerih kjer je parametrični pristop težko izvedljiv. Za ocenitev stabilnosti pridelka genotipov v različnih okoljih sta bili uporabljeni dve novi neparametrični metodi ($NS_i^{(1)}$ in $NS_i^{(2)}$), ki temeljita na rangih genotipov v danem okolju. Ti metodi uporabljata mediano kot neparametrično osrednjo tendenco in dva neparametrična indeksa porazdelitve kot inter-kvartilno in inter-decilno območje. Predstavljeni neparametrični metodi, $NS_i^{(1)}$ in $NS_i^{(2)}$, sta podobni konceptu koeficienta okoljske spremenljivosti. Rezultati so pokazali, da so imeli najbolj stabilni genotipi, opredeljeni z najmanjšimi vrednostimi $NS_i^{(1)}$ in $NS_i^{(2)}$ največji pridelek med vsemi analiziranimi genotipi. Primerjanje povprečnega pridelka z $NS_i^{(1)}$ in $NS_i^{(2)}$ je potrdila zgoraj navedene rezultate, kar kaže, da so genotipi, ki dajejo največji pridelek tudi najbolj stabilni. Te neparametrične metode bi lahko bile uporabne za hkratno selekcijo povprečnega pridelka in njegove stabilnosti. Lahko bi bile v pomoč pri selekciji primernih genotipov za stabilnost pridelka v žlahtniteljskih programih.

Ključne besede: adaptacija, GE interakcije, stabilnost, pridelek

[1] Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Maragheh, Maragheh, Iran. E-mail sabaghnia@maragheh.ac.ir

1 INTRODUCTION

Multi-environment yield trials are conducted in multiple years and test sites, are central to every plant breeding programs to evaluate and improve different crop plants. These trials are the most common and important experiments in agricultural research and different statistical methods for effective analysis of yield trials have received considerable development and discussion (Karimizadeh and Mohammadi, 2010). Although, only simple statistical method would be needed if genotypes performed similarly in all test environments but, in most cases, genotypes and environments interact such that different rankings often exist for the same set of genotypes tested over a range of test environments (Stoilova and Dechev, 2002; Sabaghnia et al., 2008b). This failure of two or more genotypes to respond similarly to a test environment which is known as genotype × environment (GE) interaction complicates their evaluation with respect to relative performance and usefulness. The GE interaction had an important affect on improvement for better genotypes buffering (Allard and Bradshaw, 1964).

The GE interaction prevents the extrapolation of agronomic evaluations from one environment to another, and so requiring more knowledge of the magnitudes of GE interactions and the various sources of variation in GE interaction. Ignoring the GE interaction is problematic when it is larger than the genotype main effect, which is a common issue in multi-environment yield trials (Gauch, 2006; Arslanoglu and Aytac, 2010; Sabaghnia et al., 2012a). Furthermore, GE interaction complicates genotype recommendations because genotypes must be targeted to specific test sites. In most cases, analysis of variance estimates the existence, significance and large magnitude of GE interaction but cannot explain its importance and so, several statistical strategies had been developed to analysis of the GE interaction pattern.

The question, whether the statistical strategy is sufficiently good to explain the GE interaction, is still discussed. The first strategy which is the classical analysis of variance model for the two-way layout of GE matrix is reviewed by Lin et al. (1986). Also, this strategy involves some linear regression models (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966) for yield stability analysis.

The linear regression model has received much attention in the literature and by including further terms using multivariate statistical procuresses as the second strategy. It has been developed into the additive main effects and multiplicative interaction (AMMI) model (Gauch, 1992) and a thorough review of the theory and applications of this model especially versus genotype plus GE interaction (GGE) biplot model has been given by Gauch (2006) and Gauch et al. (2008). All of the stability methods of both mentioned strategies are parametric. In contrast, there are nonparametric stability statistics as the third strategy which are largely unaffected by data distribution. These stability methods are based on ranks and a special genotype is considered stable if its ranking is constant across environments. Several nonparametric stability statistics have been developed to explain the GE interaction (Huehn, 1979; Kang, 1988; Ketata et al., 1989; Fox et al., 1990; Thennarasu, 1995).

The nonparametric stability statistics separate genotypes based on their similarity of response to a range of test environments. The nonparametric strategy is based on ranks of genotypes and provides an important alternative to the parametric strategies including univariate and multivariate statistics. According to Huehn (1990a) and Huehn (1996), the nonparametric strategy has some advantages over the parametric strategies such as: (i) reduction of the bias caused by outliers, (ii) no assumptions are required about the data distribution, (iii) easy to use and to interpret, (iv) additions or deletions of few genotypes or environments do not cause much variation of estimates, and (v) for many applications such as selection in plant breeding programs and cultivar testing trials, the rank order of genotypes is the most essential information.

The good ability of the nonparametric stability statistics for detecting the most stable genotypes as well as GE interaction investigation have been demonstrated in different crops such as lentil (Sabaghnia et al., 2006; Karimizadeh et al., 2012), chickpea (Ebadi-segherloo et al., 2008) and durum wheat (Sabaghnia et al., 2012b). The objective of this study was estimation of stability performance of genotypes in different test environments using two nonparametric stability statistics which are based upon the ranks of the genotypes in each environment.

2 MATERIALS AND METHODS

If x_{ij} is denoted as observed mean value of the *i*th genotype in the *j*th environment (i = 1, 2, .. ., M; j = 1, 2, . . ., N). Then, r*ij* is considered as the rank of genotype *i* in environment *j* which the lowest value is rank 1 and the highest value is rank of K. The concept of yield stability is practicable; a genotype is the most stable over test environments if its ranks are similar over environments, and so maximum stability = equal ranks over all test environments. The two nonparametric stability statistics as $NS_i^{(1)}$ and $NS_i^{(2)}$ which are proposed in this paper are:

$$NS_i^{(1)} = (Q_3 - Q_1) / M_{di}$$

$$NS_i^{(2)} = (D_9 - D_1) / M_{di}$$

In the above nonparametric statistics, $Q_3 - Q_1$ is the inter-quartile range, also called the mid-spread or middle fifty, is a nonparametric index of statistical dispersion, being equal to the difference between the upper and lower quartiles. M_{di} is the median of the genotypes' ranks in the test environments. Also, $D_9 - D_1$ is the inter-decile range is the difference between the first and the ninth deciles. The inter-decile range is another nonparametric index of statistical dispersion of the values in a set of data, similar to the inter-quartile range.

The $NS_i^{(1)}$ and $NS_i^{(2)}$ nonparametric stability statistics which presented here is similar to the nature and concept of environmental coefficient of variation (Francis and Kannenberg, 1978). The important central tendency of ranks is the median and its related measures of dispersion are inter-quartile or inter-decile range. It would be interesting that compare these nonparametric stability statistics with the environmental coefficient of variation (CV). The CV was designed primarily to exploration in investigation on the physiological basis for yield stability (Francis and Kannenberg, 1978), and was found more practical to characterize genotypes on a group basis rather than individually. However, this procedure and its related concept could be used in the plant breeding because it represents a simple and descriptive tool for investigation of genotypes' stability. Considering these benefits of CV concept, using new nonparametric stability statistics ($NS_i^{(1)}$ and $NS_i^{(2)}$) could be useful in GE interaction interpreting and identification of the most stable genotypes especially in nonparametric strategy.

3 RESULTS

The classic dataset of Yates and Cochran (1938) are used in this study and its two-way layout of yield performance for five barley genotypes at six environments yield is shown in Table 1. Also, genotypes mean, the ranks of genotypes in environments and the median of these ranks are given in Table 1. The required statistics for computation of the new nonparametric stability statistics ($NS_i^{(1)}$ and $NS_i^{(2)}$), including the first quartile, the third quartile, the inter-quartile range, the first decile, the third decile and the inter-decile range are shown in Table 2. According to the obtained results, genotype Trebi was the most stable genotype based on the lowest values of these two nonparametric statistics. This genotype had the high mean yield among studied genotypes (Table 1).

Plotting of mean yield versus $NS_i^{(1)}$ (Fig. 1) and $NS_i^{(2)}$ (Fig. 2) verified the above results and indicated that the high mean yielding genotype is identified as the most stable genotype. In other word, these nonparametric statistics would be useful for simultaneous selection for mean yield and stability. Simultaneous selection for mean yield and stability of performance is an important issue in breeding programs (Yan and Kang, 2003). Kang and Pham (1991) have studied several stability methods for simultaneous selection for yield and stability of performance. Also, Kang (1988) proposed a nonparametric stability statistic (rank-sum) using stability variance of Shukla (1972) and genotype mean rank. According to literature, all nonparametric statistics of Huehn (1979) [Sabaghnia et al. 2006]; nonparametric statistics of Thennarasu's (1995) [Ebadi-Segherloo et al. 2008]; and nonparametric statistics of Ketata et al. (1989) [Dehgahni, 2008] could not be useful for simultaneous selection of mean yield and stability.

Table 1: Two-way layout of genotype × environment cited from the paper of Yates and Cochran (1938)

	E1	E2	E3	E4	E5	E6	Mean
Yield							
Manchuria	161.7	247.0	185.4	218.7	165.3	154.6	188.8
Svansota	187.7	257.5	182.4	183.3	138.9	143.8	182.3
Velvet	200.1	262.9	194.9	220.2	165.8	146.3	198.4
Trebi	196.9	339.2	271.2	266.3	151.2	193.6	236.4
Peatland	182.5	253.8	219.2	200.5	184.4	190.1	205.1
Ranks							Median
Manchuria	1	1	2	3	3	3	2.5
Svansota	3	3	1	1	1	1	1
Velvet	5	4	3	4	4	2	4
Trebi	4	5	5	5	2	5	5
Peatland	2	2	4	2	5	4	3

Table 2: Nonparametric despersive statistics for five genotypes across six environments

Genotypes	Q_1	Q_3	Q_3- Q_1	NS^1	D_1	D_9	D_9- D_1	NS^2
Manchuria	1.0	3.0	2.0	0.8	1	3	2	0.8
Svansota	1.0	3.0	2.0	2.0	1	3	2	2.0
Velvet	2.8	4.3	1.5	0.4	2	5	3	0.8
Trebi	3.5	5.0	1.5	0.3	2	5	3	0.6
Peatland	2.0	4.3	2.3	0.8	2	5	3	1.0

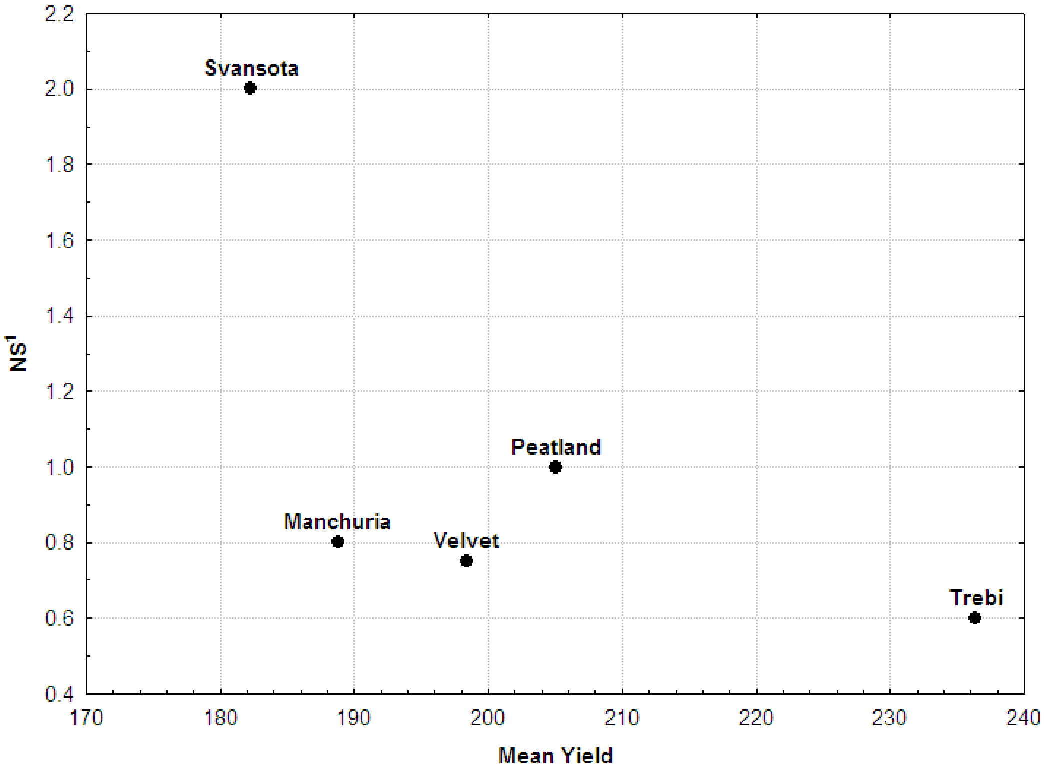

Figure 1: Plot of mean yield versus the first nonparametric stability statistic (NS^1).

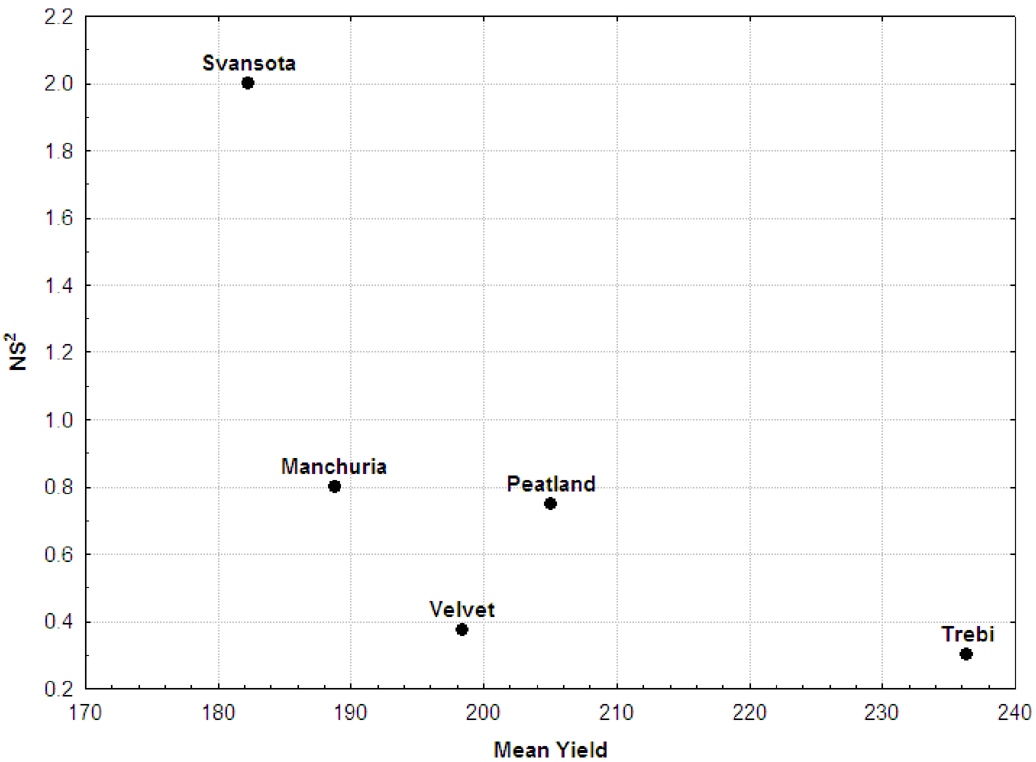

Figure 2: Plot of mean yield versus the second nonparametric stability statistic (NS^2).

4 DISCUSSION

Though several statistical strategies have been proposed for yield stability analysis, they each reflect different aspects of stability nature and maybe no single method can adequately explain genotype performance across test environments. The stability of yield is defined as the ability of a genotype to avoid substantial fluctuations in yield over a range of environmental conditions. The different stability models are broadly classified according to Lin et al. (1986) into there are three types of stability known as Type 1, 2 and 3. Lin and Binns (1991) conclude that stability models of Types 1 and 4 are useful for selection, while those of types 2 and 3 are not useful due to non-heritability. According to Becker and Leon (1988), at least two fundamentally different concepts of stability exist, the static and the dynamic. Both concepts are valuable, but their application depends on the trait considered. It seems that static type of stability is not acceptable for most yield performance breeders, who would prefer a dynamic (agronomic) concept of stability (Becker, 1981; Sabaghnia et al., 2008a). In the agronomic concept of stability, it is not required that the genotypic response to environmental conditions should be equal for all genotypes (Becker and Leon, 1988). For the more important agronomic traits such as grain yield, oil content and etc., the static concept type of stability analysis would not be beneficial for the farmers and is equivalent to type 1 of stability while the dynamic concept of stability is equivalent to type 2 of stability (Lin et al., 1986).

It seems that the new nonparametric stability statistics ($NS_i^{(1)}$ and $NS_i^{(2)}$) have similar nature and concept of environmental CV and so benefits from type 1 of stability while by identification of high mean yield genotype as the most stable genotype benefits from dynamic concept of stability. However, for simultaneous selection of mean yield and stability, it is necessary to use mean yield in the formula of each stability statistic. This concept could be seen in rank-sum (Kang, 1988) as nonparametric stability statistic or desirability index (Hernandez, 1993) as parametric stability statistic. The selection of genotypes for a particular trait depends upon their mean performance and stability statistics. The selected genotypes must have high mean value coupled with stable performance. Most of the nonparametric methods utilized classic stability concept (static or biological concept) for selection of the most favorable genotypes. It seems that there are good potencials in the new introduced nonparametric stability statistics in distinction of favorable genotypes in plant breeding programs. These methods thus provide some flexibility in the hands of plant breeders for simultaneous selection for yield and stability.

There are several statistical models for investigating stability and determination of GE interaction. Each of them reflects different aspects of stability and usually no single approach can fully explain genotype performance across environments. The nonparametric stability statistics seem to be useful alternatives to parametric methods (Huehn 1990b; Yue et al., 1997), although they do not supply information about genotype adaptability. There are several reasons why the use of nonparametric stability statistics could be preferred. These statistics avoid the bias of outliers and no assumptions are required about the distribution of the observations. Furthermore, these methods are easy to use and to interpret; therefore, estimation of stability seems to be an adequate strategy. Many parametric (univariate and multivariate) and nonparametric statistics of stability have been presented and compared in the literature (Lin et al., 1986; Flores et al., 1998; Sabaghnia et al. 2006). For making practical recommendations, it would be necessary to analyse the relationship among these statistics and compare their powers in different stability situations. This topic will be considered in detail in a subsequent paper.

5 REFERENCES

Arslanoglu F., Aytac S. 2010. Determination of stability and genotype × environment interactions of some agronomic properties in the different soybean (*Glycine max*. (L) Merrill) cultivars. Bulgarian Journal of Agricultural Science, 16: 181–195.

Becker, H.C., Leon J. 1988. Stability analysis in plant breeding. Plant Breeding, 101:1–23; DOI: 10.1111/j.1439-0523.1988.tb00261.x

Becker, H.C. 1981. Correlations among some statistical measures of phenotypic stability. Euphytica, 30: 835–840; DOI: 10.1007/BF00038812

Dehghani H., 2008. Estimating yield stability by nonparametric stability analysis in maize (Zea mays L.). Plant Breeding and Seed Science, 58: 61–77.

Ebadi-Segherloo A., Sabaghpour S.H., Dehghani H., Kamrani M. 2008. Non-parametric measures of phenotypic stability in chickpea genotypes (*Cicer arietinum* L.). Euphytica, 2: 221-229; DOI: 10.1007/s10681-007-9552-x

Eberhart S.A. Russell W.A. 1966. Stability parameters for comparing varieties. Crop Science, 6: 36–40; DOI: 10.2135/cropsci1966.0011183X000600010011x

Finlay K.W. Wilkinson G.N. 1963. The analysis of adaptation in a plant breeding program. Australian Journal of Agriculture Research, 14: 742–754; DOI: 10.1071/AR9630742

Flores F., Moreno M.T., Cubero, J.I. 1998. A comparison of univariate and multivariate methods to analyze G × E interaction. Field Crops Research, 56: 271–286; DOI: 10.1016/S0378-4290(97)00095-6

Fox P.N., Skovmand B., Thompson B.K., Braun H.J., Cormier R. 1990. Yield and adaptation of hexaploid spring triticale. Euphytica, 47: 57–64; DOI: 10.1007/BF00040364

Francis T.R. Kannenberg L.W. 1978. Yield stability studies in short-season maize: I. A descriptive method for grouping genotypes. Canadian Journal of Plant Science, 58: 1029–1034; DOI: 10.4141/cjps78-157

Gauch H.G. 1992. Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs. Elsevier, Netherlands, Amsterdam.

Gauch H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. Crop Science, 46: 1488–1500; DOI: 10.2135/cropsci2005.07-0193

Gauch, H.G., Piepho H.P., Annicchiaricoc P. 2008. Statistical analysis of yield trials by AMMI and GGE. Further considerations. Crop Science, 48: 866–889; DOI: 10.2135/cropsci2007.09.0513

Hernández C.M., Crossa J., Castillo A. 1993. The area under the function: an index for selecting desirable genotypes. Theoretical Applied Genetics, 87: 409–415; DOI: 10.1007/BF00215085

Huehn M. 1979. Beitrage zur Erfassung der phänotypischen Stabilität. EDV Medizin und Biologie, 10: 112–117.

Huehn M. 1990a. Non-parametric measures of phenotypic stability: Part 1. Theory. Euphytica, 47: 189–194.

Huehn M. 1990b. Non-parametric measures of phenotypic stability: Part 2. Application. Euphytica, 47: 195–201.

Huehn M. 1996. Non-parametric analysis of genotype × environment interactions by ranks. p. 213–228. In M.S. Kang and H.G. Gauch (ed.) Genotype by environment interaction. CRC Press, Boca Raton, F; .DOI: 10.1201/9781420049374.ch9

Kang M.S., Pham H.N. 1991. Simultaneous selection for high yielding and stable crop genotypes. Agronomy Journal, 83: 161–165; DOI: 10.2134/agronj1991.00021962008300010037x

Kang M.S. 1988. A rank–sum method for selecting high-yielding, stable corn genotypes. Cereal Research Communications, 16: 113–115.

Karimizadeh R., Mohammadi M. 2010. AMMI adjustment for rainfed lentil yield trials in Iran. Bulgarian Journal of Agricultural Science, 16: 66–73.

Karimizadeh R., Mohammadi M., Sabaghnia N., Shefazadeh M.K., Pouralhossini J. 2012. Univariate stability analysis methods for determining genotype × environment interaction of durum wheat grain yield. African Journal of Biotechnology, 11: 2563–2573.

Ketata H., Yan S.K., Nachit M. 1989. Relative consistency performance across environments. Int. Symposium on Physiology and Breeding of Winter Cereals for stressed Mediterranean Environments. Montpellier, July 3–6, 1989.

Lin, C.S. Binns M.R. 1991. Genetic properties of four types of stability parameters. Theoretical Applied Genetics, 82: 505–509; DOI: 10.1007/BF00588606

Lin, C.S., Binns M.R., Lefkovitch L.P. 1986. Stability analysis: Where do we stand? Crop Science, 26: 894–900; DOI: 10.2135/cropsci1986.0011183X002600050012x

Sabaghnia N., Dehghani H., Sabaghpour S.H. 2006. Nonparametric methods for interpreting genotype · environment interaction of lentil genotypes. Crop Science, 46: 1100–1106 ; DOI : 10.2135/cropsci2005.06-0122

Sabaghnia N., Dehghani H., Sabaghpour S.H. 2008a. Graphic analysis of genotype by environment interaction for lentil yield in Iran. Agronomy Journal, 100: 760–764; DOI: 10.2134/agronj2006.0282

Sabaghnia N., Karimizadeh R., Mohammadi M. 2012a. Genotype by environment interaction and stability analysis for grain yield of some lentil genotypes. Žemdirbystė=Agriculture 99: 305–312.

Sabaghnia N., Karimizadeh R., Mohammadi M. 2012b. The use of corrected and uncorrected nonparametric stability measurements in durum wheat multi-environmental trials. Spanish Journal of Agricultural Research 10: 722–730; DOI: 10.5424/sjar/2012103-384-11

Sabaghnia N., Sabaghpour S.H., Dehghani H. 2008b. The use of an AMMI model and its parameters to analyse yield stability in multi-environment trials. Journal of Agricultural Science, 146: 571–581; DOI: 10.1017/S0021859608007831

Shukla G.K. 1972. Some aspects of partitioning genotype-environmental components of variability. Heredity 28:237–245; DOI: 10.1038/hdy.1972.87

Stoilova A. Dechev D. 2002. Genotype environment interaction and phenotypic stability of economic traits in cotton lines. Bulgarian Journal of Agricultural Science, 8: 485–491.

Thennarasu K. 1995. On certain non-parametric procedures for studying genotype–environment interactions and yield stability. Ph.D. thesis. P.J. School, IARI, New Delhi, India.

Yan W., Kang M.S. 2003. GGE-biplot analysis:A graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL, USA.

Yates F. Cochran W.G. 1938. The analysis of groups of experiments. Journal of Agricultural Science, 28: 556–580; DOI: 10.1017/S0021859600050978

Yue, G.L., Roozeboom K.L., Schapaugh W.T., Liang G.H. 1997. Evaluation of soybean cultivars using parametric and nonparametric stability estimates. Plant Breeding, 116: 271–275; DOI: 10.1111/j.1439-0523.1997.tb00995.x