Rank correlation among different statistical models in ranking of winter wheat genotypes'

Mozaffar Roostaei[a], Reza Mohammadi[b,*], Ahmed Amri[c]

[a]Dryland Agricultural Research Institute (DARI), Maragheh, Iran
[b]Dryland Agricultural Research Institute (DARI), P. O. Box 67145-1164, Kermanshah, Iran
[c]International Center for Agricultural Research in the Dry Areas (ICARDA), Rabat, Morocco

ABSTRACT

Several statistical methods have been developed for analyzing genotype × environment (GE) interactions in crop breeding programs to identify genotypes with high yield and stability performances. Four statistical methods, including joint regression analysis (JRA), additive mean effects and multiplicative interaction (AMMI) analysis, genotype plus GE interaction (GGE) biplot analysis, and yield–stability (YSi) statistic were used to evaluate GE interaction in 20 winter wheat genotypes grown in 24 environments in Iran. The main objective was to evaluate the rank correlations among the four statistical methods in genotype rankings for yield, stability and yield–stability. Three kinds of genotypic ranks (yield ranks, stability ranks, and yield–stability ranks) were determined with each method. The results indicated the presence of GE interaction, suggesting the need for stability analysis. With respect to yield, the genotype rankings by the GGE biplot and AMMI analysis were significantly correlated ($P < 0.01$). For stability ranking, the rank correlations ranged from 0.53 (GGE–YSi; $P < 0.05$) to 0.97 (JRA–YSi; $P < 0.01$). AMMI distance (AMMID) was highly correlated ($P < 0.01$) with variance of regression deviation (S^2di) in JRA ($r = 0.83$) and Shukla stability variance (σ^2) in YSi ($r = 0.86$), indicating that these stability indices can be used interchangeably. No correlation was found between yield ranks and stability ranks (AMMID, S^2di, σ^2, and GGE stability index), indicating that they measure static stability and accordingly could be used if selection is based primarily on stability. For yield–stability, rank correlation coefficients among the statistical methods varied from 0.64 (JRA–YSi; $P < 0.01$) to 0.89 (AMMI–YSi; $P < 0.01$), indicating that AMMI and YSi were closely associated in the genotype ranking for integrating yield with stability performance. Based on the results, it can be concluded that YSi was closely correlated with (i) JRA in ranking genotypes for stability and (ii) AMMI for integrating yield and stability.

© 2014 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. All rights reserved.

* *Corresponding author.* Tel.: +98 831 2333410 12; fax: +98 831 2333409.
 E-mail address: r.mohammadi@areo.ir (R. Mohammadi).
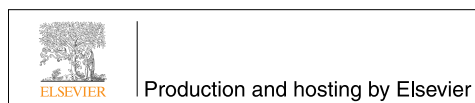
1. Introduction

In crop breeding programs, genotypes are evaluated in multi-environment trials (METs) for testing their performance across environments and selecting the best genotypes in specific environments. Genotype × environment (GE) interaction is an important issue faced by plant breeders in crop breeding programs. A significant GE interaction for a quantitative trait such as grain yield can seriously limit progress in selection. Variance due to GE interaction is an important component of the variance of phenotypic means in selection experiments [1]. GE interactions complicate the identification of superior genotypes [2] but their interpretation can be facilitated by the use of several statistical modeling methods. These methods may use linear models, such as joint regression analysis [3–5], multivariate analytical methods such as AMMI (additive mean effects and multiplicative interaction) analysis [6,7], or GGE (genotype plus GE interaction) biplot analysis [8,9].

The linear regression of genotype values on environmental mean yield [3,4], frequently termed joint regression analysis, is undoubtedly the most popular method for analyzing GE interaction, owing to its simplicity and the ready applicability of its information on adaptive responses to locations other than the chosen test sites. Earlier, Finlay and Wilkinson [4] proposed the use of linear regression slopes as a measure of stability. Eberhart and Russell [5] further proposed that both regression coefficients and deviations from linear regression (S^2di) should be taken into consideration in identifying stable genotypes, and suggested that a genotype with $b = 1.0$ and $S^2di = 0$ would be regarded as stable.

The AMMI model uses analysis of variance (ANOVA, an additive model) to characterize genotype and environment main effects and principal component analysis (a multiplicative model) to characterize their interactions (IPCA). The AMMI analysis has been shown to be effective; it captures a large portion of the GE sum of squares, clearly separating the main and interaction effects; and the model often provides an agronomically meaningful interpretation of the data [7]. Another powerful statistical model that addresses some of the disadvantages of AMMI is the GGE biplot. The method is effective for identifying the best-performing cultivar across environments, identifying the best cultivars for mega-environment differentiation, and evaluating the yield and stability of genotypes [8,9]. According to the GGE biplot, a highly stable genotype would have a shorter projection on to the average environment coordinate (AEC) abscissa, irrespective of its direction [9].

Recent review articles [10–12] have compared these statistical models. Gauch [10] and Gauch et al. [12] reviewed the AMMI and GGE literature, favoring AMMI. Yan et al. [11] responded to those articles, favoring GGE. Several studies have also been performed comparing GGE biplots and YSi in bean [13], maize [14], and durum wheat [15]; GGE biplots and JRA in maize [16] and triticale [17]; and JRA and AMMI models in cereal crops [18] for stability analysis. However, little is known about rank correlation among the four statistical methods (AMMI analysis, GGE biplot, JRA, and YSi statistic) applied in a single study. The main objectives of the present study were to (i) compare the statistical methods (AMMI analysis, GGE biplot, JRA, and the YSi statistic) in the ranking of 20 winter wheat genotypes for yield, stability, and yield–stability and (ii) evaluate rank correlations among the statistical methods on the basis of yield ranks, stability ranks, and yield–stability ranks.

2. Materials and methods

2.1. Experimental data

Grain yield data obtained from 20 winter wheat genotypes, consisting of 18 breeding lines (G1–G18) and two check cultivars (G19 and G20, representing the landrace "Sardari" and the released cultivar "Azar-2", respectively), grown in eight test locations representative of winter wheat growing areas in Iran for three consecutive cropping seasons (2003–2005), were subjected to analysis of rank correlation among the four statistical procedures (AMMI, GGE biplot, JRA, and YSi statistic) in the rankings of genotypes. In each environment (location–year combination), the experimental layout was a randomized complete block design with four replicates. The plot size was 7.2 m^2 (6 rows, 6 m long, 20 cm row spacing). The fertilizer rate was 50 kg N ha^{-1} and 50 kg P$_2$O$_5$ ha^{-1} applied at planting stage.

2.2. Statistical analysis

Combined analysis of variance (ANOVA) for grain yield data was performed to determine the effects of environment, genotype, and GE interaction. Four statistical methods were applied to evaluate GE interaction in the wheat MET data. Regression analysis was performed for each of the 20 wheat genotypes based on the method of Eberhart and Russell [5]. The performance of each genotype in each environment was regressed on the means of all genotypes in each environment. Genotypes with regression coefficient (b) of unity and variance of regression deviations (S^2di) equal to zero will be highly stable.

The yield stability (YSi) statistic was generated as described by Kang [19] and applied for selecting high-yielding and stable genotypes. Ranks were assigned for mean yield, with the genotype with the highest yield given a rank of 20. Similarly, ranks were assigned for the stability parameter with the lowest estimated value receiving the rank of 1. Stability ratings were computed as follows: -8, -4, and -2 for stability measures significant at $P < 0.01$, 0.05, and 0.10, respectively; and 0 for the non-significant stability measure. The stability ratings of -8, -4, and -2 were chosen because they changed the genotype ranks from those based on yield alone [19].

AMMI analysis was performed with IRRISTAT 5.1 software [20]. AMMI analysis combines additive components in a single model for the main effects of genotypes and environments, as well as multiplicative components for the interaction effect. Genotypes (or environments) with large IPC scores (either positive or negative) have large interactions, whereas genotypes (or environments) with IPC1 scores near zero have small interactions.

To further describe stability using AMMI analysis, the AMMI statistic coefficient (D) was calculated as follows, [21] and is referred to as AMMI distance:

$$D = \sqrt{\sum_{r=1}^{N} \gamma_{is}^2} \qquad (i = 1, 2, 3, \ldots, n)$$