

## **UNDERSTANDING AND UTILIZATION OF GENOTYPE-BY- ENVIRONMENT INTERACTION IN MAIZE BREEDING**

Vojka BABIĆ<sup>1</sup>, Milosav BABIĆ<sup>1</sup>, Mille IVANOVIĆ<sup>2</sup>, Marija  
KRALJEVIĆ-BALALIĆ<sup>3</sup> and Miodrag DIMITRIJEVIĆ<sup>3</sup>

<sup>1</sup>Maize Research Institute, Zemun Polje, Belgrade, Serbia

<sup>2</sup>Institute of Field and Vegetable Crops, Novi Sad, Serbia

<sup>3</sup>Faculty of Agriculture, University of Novi Sad, Novi Sad, Serbia

Babić V., M. Babić, M. Ivanović, M. Kraljević Balalić and M. Dimitrijević (2010): *Understanding and utilization of genotype-by-environment interaction in maize breeding.*- Genetika, Vol 42, No. 1, 79 -90.

Due to the interaction and noise in the experiments, yield trials for studying varieties are carried out in numerous locations and in the course of several years. Data of such trials have three principle tasks: to evaluate precisely and to predict the yield on the basis of limited experimental data; to determine stability and explain variability in the response of genotypes across locations; and to be a good guide for the selection of the best genotype for sowing under new agroecological conditions. The yield

---

*Corresponding author:* Babic Vojka, Maize Research Institute, Zemun Polje, Slobodana Bajica 1, 11185 Belgrade- Zemun, Serbia; Tel: (011)3756704, Fax: 3756707, e-mail: vbabic@mrizp.rs

prediction without the inclusion of the interaction with the environments is incomplete and imprecise. Therefore, a great deal of breeding and agronomic studies are devoted to observing of the interaction via multilocation trials with replicates with the aim to use the interaction to obtain the maximum yield in any environment.

Fifteen maize hybrids were analysed in 24 environments. As the interaction participates in the total sum of squares with 6%, and genotypes with 2%, the interaction deserves observations more detailed than the classical analysis of variance (ANOVA) provides it. With a view to observe the interaction effect in detail in order to prove better understanding of genotypes, environments and their interactions AMMI (Additive Main Effect and Multiplicative Interaction) and the cluster analysis were applied. The partition of the interaction into the principal components by the PCA analysis (Principal Components Analysis) revealed a part of systematic variations in the interaction. These variations are attributed to the length of the growing season in genotypes and to the precipitation sum during the growing season in environments. Results of grouping by the cluster analysis are in high accordance with grouping observed in the biplot of the AMMI model.

*Key words:* AMMI, cluster analysis, interaction, maize hybrids

## INTRODUCTION

Different interests of breeders, as well as, seed producers and distributors, on the one hand, and framers on the other hand, arise an important question: How broadly can a variety be adapted and be able at the same time to have a high yield in a given location? Farmers want a small genotype x year interaction. Breeders, seed producers and distributors want a broadly adapted genotype that will be a great success across a great area (small genotype x location interaction). Dividing broad areas into regions that are, first of all, different units based on climatic and soil conditions, is one of methods to find out a compromising solution for these various interests. Successful breeding for targeted growing areas largely depends on identification of the main sources of phenotypic variation in that region. To obtain variety possessing diminished genotype by environment interaction for those predominant sources of variation means good ratio between the stable and high yield (PETROVIC *et al.*, 2009).

The most often applied model for the analysis of yield trials that is routinely used in commercial breeding programmes, is the analysis of variance (ANOVA), first of all due to the easiness of both, the application and the interpretation of gained results, as well as, due to a concept that is close to the agronomic point of view of field trials. The concepts of stability based on the linear regression were the initial attempts to explain and comprehend the genotype x environment interaction. However, in the course of time it was shown that it was very difficult to explain a complex phenomenon, such as an interaction that included a greater number of

vectors, with one universal parameter. In recent times the statistical methods of multivariate analysis have been applied in order to observe, to a greater extent, the nature of the interaction in the multilocation trials, dividing existing areas into smaller target regions. A target region does not have to be a continuous region in a geographical sense, but can be more in a sense of a similar interaction response of studied genotypes. In some cases, even different levels of cropping practices can lead to dividing into target regions within the same agro-climatic region. Some authors state that narrow adaptation of a genotype in a relatively small region can be used with the aim to increase the yield (ATLIN *et al.*, 2000).

ZOBEL *et al.* (1988) summarised their comparisons of different statistical methods and state: "Yield trials frequently have both significant main effects and a significant genotype x environment (GE) interaction. Traditional statistical analyses are not always effective with this data structure: the usual analysis of variance (ANOVA), having a merely additive model, identifies the GE interaction as a source but does not analyse it; PCA analysis, on the other hand, is a multiplicative model and hence contains no sources for additive genotype or environment main effects; and linear regression (LR) analysis is able to effectively analyse interaction terms only where the pattern fits a specific regression model. The consequence of fitting inappropriate statistical models to yield trial data is that the interaction may be declared insignificant, although a more appropriate analysis would find agronomically important and statistically significant patterns in the interaction. Since ANOVA, PCA, and LR are sub-cases of the more complete AMMI model, AMMI offers a more appropriate first statistical analysis of yield trials that may have a genotype x environment interaction. AMMI analysis can then be used to diagnose whether or not a specific sub-case provides a more appropriate analysis. AMMI has no specific experimental design requirements, except for a two-way data structure."

Although the AMMI analysis of yield trials does not use the data on environmental factors, these factors themselves, such as precipitation, average daily, maximum and minimum temperatures, as well as, their height and amplitudes, nitrogen fertilisers, irrigation and the clay content, very often correlate with the data of the AMMI statistics (GAUCH, 1992; ROMAGOSA *et al.*, 1993).

The objective of the present study was to identify a part of the systematic variation within the interaction effect in the multilocation trial with commercial maize hybrids with the use of the AMMI and cluster analysis. The starting point in the studies was the assumption that the differences in yields of commercial maize hybrids were not just a result of different genetic constitutions, but also of their different specific adaptability to conditions of certain environment.

## MATERIALS AND METHODS

As medium late and late maturity maize hybrids (FAO 400-700) are mainly sown in our regions, 15 the following commercial maize hybrids were selected for these studies: ZPSC-42a, ZPSC-480, Stan.-500, ZPSC-533, ZPSC-570, ZPSC-580,

ZPSC-599, Stan.-600, ZPSC-633, ZPSC-677, ZPSC-701, ZPSC-704, ZPSC-732, ZPSC-735 and ZPSC-753.

Two-year four-replicate trials with two plant densities ( $D_1= 54,900$  plants  $ha^{-1}$ ,  $D_2= 64,900$  plants  $ha^{-1}$ ) were set up according to the randomised complete block design. The experimental plots were sown in the experimental fields of the Maize Research Institute, Zemun Polje, under irrigation and dry land farming conditions; the trials were carried out in another four locations under dry land farming conditions.

The AMMI analysis is done in the programme MATMODEL version 1 (Iowa State University; GAUCH, 1992). A two-way data structure is the principal prerequisite for the application of the AMMI analysis. In order to fulfil this prerequisite each density x year x location combination was taken as an environment.

Based on data on obtained yields, the cluster analysis was performed in order to determine to what extent the results of the AMMI analysis are in accordance with the results of the cluster analysis.

## RESULTS AND DISCUSSION

The first year of investigation was almost ideal for the maize crop from the aspects of precipitation, while the second year was exceptionally arid and affected average yields over both, hybrids and locations. The highest precipitation sum from March to August was recorded in the location of Pančevo during the first year of investigation (495 mm), while the lowest precipitation sum was detected in the same location during the second year of investigation (86 mm).

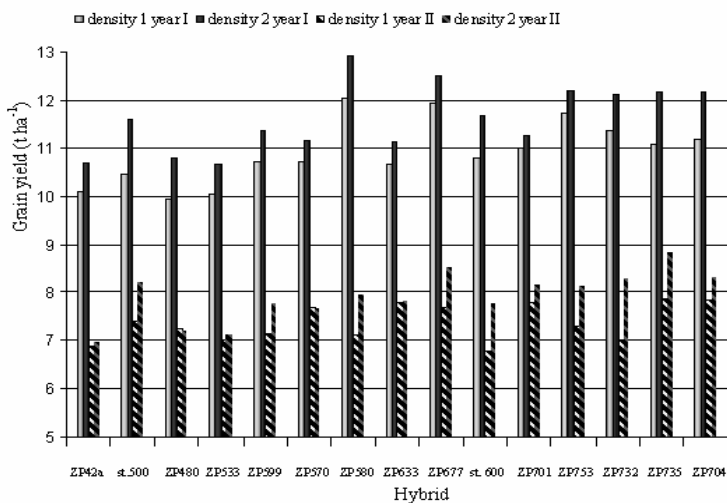


Figure 1. Average grain yields of maize hybrids over years and sowing densities

The highest average yield (15.09 t ha<sup>-1</sup>) was obtained in the location Bečej I in the trial with a greater plant density, while the lowest average yield (3.32 t ha<sup>-1</sup>) was recorded in the location Adaševci II with the greater plant density, too. The average grain yields of hybrids varied from 6.89 t ha<sup>-1</sup> in the hybrid ZP-42a, during the second year of investigation and in a lower plant density, to 12.9 t ha<sup>-1</sup> in the hybrid ZP-580, during the first year of investigation and in a higher plant density. With the exception of hybrids ZP-480 and ZP-570 the remaining hybrids had higher yields in the crops with the greater density and although these differences were statistically significant they did not significantly affect the hybrid ranking across different environments (Figure 1).

Table 1. ANOVA of AMMI2 model for maize grain yield

Source of variation	DF	% SS GxE	SS	%SS tret	% noise	MS
Environments	23		12650.93		0.18	50.04**
Genotypes	14		307.52	2.2	4.52	21.97**
G X E	322		879.43	6.4	6.36	2.73**
PCA 1	36	65.8	368.30	96.3	9.71	10.23**
PCA 2	34	22.5	126.14	97.2	26.76	3.71**
Residual	252		384.98			
Treatments	359		13837.88			
Error	1008		1000.93			0.99
Total	439		15127.70			10.51

The analysis of variance of the AMMI model shows that the main effects and interaction effects are significant. Participation of the genotype variation in the treatments sum of squares (SS) amounted to 2.2%, while the participation of SS of the GxE interaction was 6.4%. Such percentage participation of certain sources of variation in the total sum of squares is not uncommon. GAUCH and ZOBEL (1997) stated in their study that these effects, which are the only relevant effects, very often encompass from 10 to not more than 40% of the total trial variation. On the other hand, as the interaction encompasses the highest number of degrees of freedom, the greatest amount of noise is also incorporated in the interaction (36.4%). The separation of the interaction into the greater number of principal components by the PCA analysis revealed a part of systematic variations in the interaction that could have a known cause. At the same time, the attention has to be paid to the number of axes that should be kept in the analysis, and the number of remaining axes, because useful information can be fast annulled by noise. The principal case is that each subsequent PCA axis is hampered, to a greater extent, by noise. The first, i.e. second PCA component is hampered by noise in the amount of 9.7%, i.e. 26.7%, respectively, whereby they encompass 42 and 14% of the interaction sum of squares (65% and 22.5% of signal without noise respectively). The residual is statistically significant, but hampered by 65% noise. The model with two axes encompasses 97.2% SS treatment and have a residual of 0.517 t ha<sup>-1</sup>, which is 5% of the grand

mean for the grain yield, while the model with one axis encompasses 96.3% SS treatment and have a residual of  $0.596t \text{ ha}^{-1}$ , which is 6% of the grand mean (Table 1).

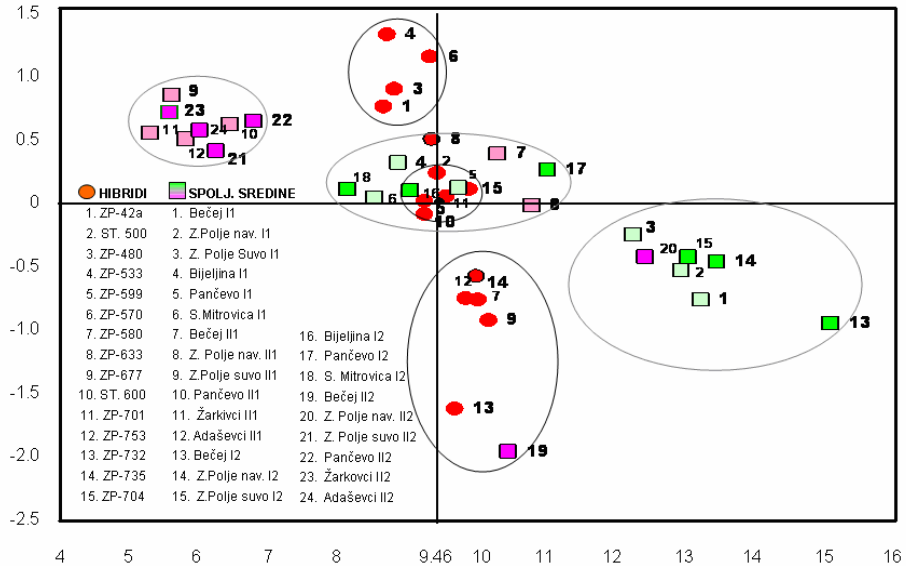


Figure 2. AMMI1 (grain yield-PCA1) biplot for maize grain yield

The AMMI analysis provides a possibility to display results on the biplot diagram, and at the same time, provides the comprehension of the complete complex of genotypes, environments and their interactions (Figure 2). Grouping of both, hybrids and environments, is observable on the biplot. They are grouped into three groups that differ in values of the scores of the first PCA axis and in the average yields. The hybrids 1, 3, 4 and 6 have high positive values of the first PCA axis and the average yields below the grand mean. On the other, hand the hybrids 7, 9, 12, 13 and 14 have the average yields above the grand mean and high negative values of the first interaction component. The remaining hybrids have small values of the first interaction components and yields around or above the grand mean. Since genotypes of the medium early maturity group belong to the first group, and hybrids of the late maturity group belong to the second group, the assumption is that the first component encompasses the part of variability that is attributed to the length of the growing season of the observed hybrids. It is also confirmed by the rank correlation

coefficient of the anthesis date and the height of the PCA1 score that is statistically significant ( $P < 0.005$ ) and amounts to 0.63. This is in accordance with studies performed by GAUCH (1992) in which this author states that the first principal component scores are correlated with the length of the soya bean growing season.

Environments also clearly form three separate groups. Positive values of the PCA1 score and low average yields were recorded in the locations in Zemun Polje under conditions of dry land farming, than in Pančevo, Žarkovci and Adaševci in both densities during the arid year (9, 10, 11, 12, 21, 22, 23 and 24). Yields around the grand mean and relatively small values of the PCA1 score were detected in the locations of Bijeljina, Pančevo and Sremska Mitrovica during the first year and in the both densities, as well as, in the locations of Bečej and Zemun Polje during the second year under conditions of irrigation and a lower crop density (4, 5, 6, 7, 8, 16, 17 and 18). Negative values of the PCA1 score and yields above the grand mean were recorded in the locations of Bečej and Zemun Polje under irrigation conditions and Zemun Polje under conditions of dry land farming during the first year and in both densities, as well as, in the locations of Bečej and Zemun Polje during the second year and in a higher sowing density (1, 2, 3, 13, 14, 15, 19 and 20). The level of the PCA1 score for locations is correlated with the precipitation sum during the growing season, as well as, with the average yield levels across locations ( $P < 0.001$ ) and amounts to 0.88 and 0.66, respectively. This points out that a part of systematic variability related to the available precipitation sum during the growing season and to the average yields is extracted from the interaction in locations by the first interaction component. These results are in accordance with results gained by ANNICCHIARICO AND MARIANI (1996).

Based on these results, the most desirable genotype for a certain group of environments is the one with the most similar interaction effect, and which at the same time achieves the highest yields. For instance, the genotype 9 or 14 would be the most desirable for the environments 1, 2, 3, 13, 14, 15 and 20. On the other hand, genotypes with a small value of the interaction and high yields (genotype 15) can be recommended when a stable yield is more important than maximum yields, as is case with small-scale farmers.

The AMMI2 biplot diagrams present only the interaction effect and show that hybrids 4, 13, 8 and 14 expressed the highest interaction, pointing out to their narrow adaptability to certain environments, while hybrids 2, 15, 10, 1 and 5 expressed the lowest interaction, pointing out to their stability or broad adaptability. The highest interactions were expressed by environments 19, 20, 7, 17, 16, 5 and 13 (Figure 3). The smaller angle between interaction vectors, is the greater similarity in the interaction response is. A hybrid achieving the highest average yields and with the most similar interaction response is a desirable hybrid for a given environment. The second interaction axis could not be related with any recognisable cause. This is in agreement with results obtained by many authors (GAUCH, 1992; CROSSA, 1990) who state that even in case that axes of a higher order, based on F test, are statistically significant, it is enough to consider only the first or the first two

components, first of all, because of a great hampering of axes of a higher order by noise.

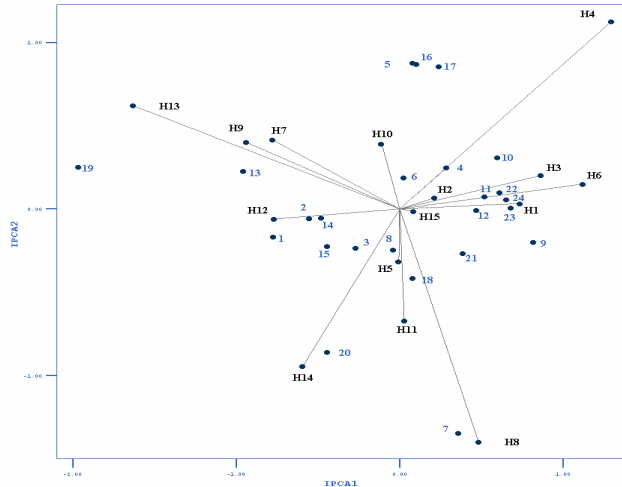


Figure 3. AMMI2 biplot (PCA1-PCA2) for maize grain yield with vectors of genotypes

One of methods to approach the problem of the interaction is to divide the unit of observation, whether it is dealt with genotypes or environments, into smaller homogenous groups within which the interaction is smaller. Therefore, the cluster analysis of both, genotypes and environments, was performed, and based on the yield data, results were presented in dendrograms. Cluster analysis has many features that make it attractive to plant breeders, but also it has some disadvantages. The most frequently it is being used for characterization of genetic distances but it also can be used for environments and genotypes grouping with similar interaction response in multi locations yield trials (BABIC *et al.*, 2009). Grouping of genotypes and environments by the cluster analysis highly corresponds with grouping that is observed in the AMMI1 biplot. There are two clusters, A and B, in the genotype dendrogram, whereby the cluster B divides into subclusters b1 and b2. Hybrids 1, 3, 4 and 6 are included into the cluster A, which is in accordance with the group of hybrids in the AMMI1 biplot that have high positive values of the PCA1 axis score and yields below grand mean. Hybrids 9, 12, 7 and 13 are included into the subcluster b1, which is analogous to the group of hybrids with average yields above the grand mean and with high negative values of the PCA1 score. The subcluster b2 encompasses hybrids 5, 15, 8, 11, 14, 2 and 10, which is, with the exception of the



hybrid 14, analogous to the group of hybrids with small values of the first interaction component score and yields around the grand mean (Figure 4).

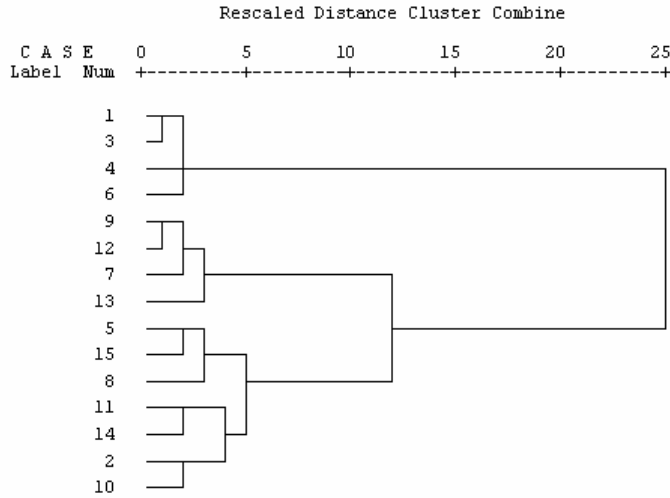


Figure 4. Cluster analysis dendrogram for maize grain yield for genotypes

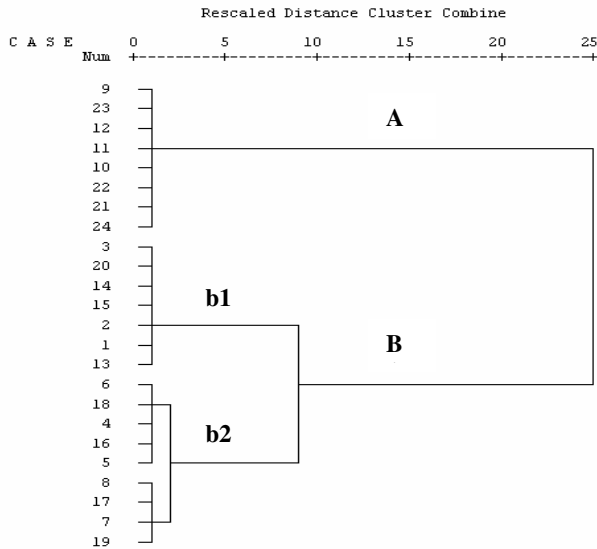


Figure 5. Cluster analysis dendrogram for maize grain yield for environments

Grouping of environments by the cluster analysis is also analogous to groups formed on the AMMI1 model biplot diagram with the exception of the environment 19 (Figure 5).

The principal goal of agronomists and breeders is to maximally increase the yield of a crop in different environments regardless of problems created by the interaction. Although scientists agree that GE interaction is important and should be studied, their approaches and methods in solving this problem differ. Better understanding of genotypes, environments and complexes of their interactions helps in a more precise prediction, and provides better answers to questions asked by breeders. Since the interaction is of a multivariate nature, scientists have been trying for a long time to summarise a great number of vectors, included into the interaction, into one universal parameter (EBERHART-RUSSELL, 1966; FINLY – WILKINSON, 1963). More recent studies point out that such a concept has to be abandoned and the interaction should be presented with two or more parameters depending on each actual case. The AMMI analysis is a very applicable for such an approach.

#### ACKNOWLEDGEMENTS

I should like to thank those taking part in Training Workshop at The Faculty of Agriculture, University of Belgrade, Serbia, for their helpful comments on the version of this paper presented there, and for partial support by the project “Education, Research and Training for Global Environmental Change and Sustainable Management of Natural Resources in Western Balkans”.

Received October 29<sup>th</sup>, 2009

Accepted January 19<sup>th</sup>, 2010

#### REFERENCES

- ANNICCHIARICO, P. and M. PEREZIN (1996): Prediction of adaptability and yield stability of durum wheat genotypes from yield response in normal and drought-stressed conditions. *Field Crops Res.* **46**, 71-80.
- ATLIN, G.N., R.J. BAKER, K.B. MCRAE and X. LU (2000): Selection response in subdivided target regions. *Crop Sci.* **40**, 7-13.
- BABIC, M., I. BAUER, V. BABIC, V. ANDJELKOVIC (2009): Genetic divergence of different temporal cycles of selection. *Maydica* **54**, 221-228.
- CROSSA J. (1990): Statistical analyses of multilocation trials. *Advances in Agron.* **44**, 55-85.
- EBERHART, S.A. and W.A. RUSSELL (1966): Stability parameters for comparing varieties. *Crop Sci.* **6**, 36-40.
- FINLAY, K.W. and G.N. WILKINSON (1963): The analysis of adaptation in a plant breeding programme. *Aust. J. Agric. Res.* **14**, 743-754.

- 
- GAUCH, H.G.JR. (1992): Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier science publishers B. V., Amsterdam, The Netherlands.
- GAUCH, H.G.JR. (2006): Winning the Accuracy Game. *American Scientist* (online) March-April 2006. [www.americanscientist.org](http://www.americanscientist.org)
- GAUCH, H.G.JR. and R.W. ZOBEL (1997): Identifying mega-environments and targeting genotypes. *Crop Sci.* 37, 311-326.
- PETROVIC, S., M. DIMITRIJEVIC, M. BELIC, B. BANJAC and M.VUKOSAVLJEVIC (2009): Spike stability parameters in wheat grown on solonetz soil. *Genetika*, Vol. 41, No. 2, 199-205.
- ROMAGOSA I., P.N. FOX, L.F. GARCIA DEL MORAL, J.M. RAMOS, F. ROCA DE TOGORES and J.L. MOLINA CANO (1993): Integration of statistical and physiological analyses of adaptation of near-isogenic barley lines. *TAG* 86, 822-826.
- ZOBEL R., M.J. WRIGHT and H.G. GAUCH (1988): Statistical analysis of yield trial. *Agronomy Journal* 80, 388-393

## RAZUMEVANJE I ISKORIŠĆAVANJE Gx E INTERAKCIJE U OPLEMENJIVANJU KUKURUZA

Vojka BABIĆ<sup>1</sup>, Milosav BABIĆ<sup>1</sup>, Mille IVANOVIĆ<sup>2</sup>,  
Marija KRALJEVIĆ-BALALIĆ<sup>3</sup> i Miodrag DIMITRIJEVIĆ<sup>3</sup>

<sup>1</sup>Institut za kukuruz „Zemun Polje“, Beograd

<sup>2</sup>Institut za ratarstvo i povrtarstvo, Novi Sad

<sup>3</sup>Poljoprivredni fakultet, Univerzitet u Novom Sadu, Novi Sad

### I z v o d

Ogledi za ispitivanje varijeteta se izvode u brojnim lokacijama i u toku više godina i u osnovi imaju tri glavna zadatka: da precizno procene i predvide prinos na osnovu ograničenih eksperimentalnih podataka; da determinišu stabilnost i objašnjivu varijabilnost u odgovoru genotipova kroz lokacije; i da budu kvalitetan vodič za odabir najboljeg genotipa za setvu u novim agro-ekološkim uslovima. Procena prinosa bez uključivanja interakcije sa spoljnom sredinom je nekompletna i neprecizna. Zbog toga je značajan deo oplemenjivačkih i agronomskih istraživanja posvećen istraživanju interakcije, kroz višelokacijske oglede sa ponavljanjima, u cilju iskorišćavanja interakcije za dobijanje maksimalnog prinosa u svakoj sredini.

U radu je analizirano 15 hibrida kukuruza u 24 spoljne sredine. Obzirom da interakcija učestvuje u ukupnoj sumi kvadrata sa 6%, a sami genotipovi sa 2% ona zaslužuje detaljnije razmatranje nego što nam to nudi klasična analiza varijanse (ANOVA). Sa ciljem da se detaljnijim uvidom u interakcijski efekat omogući bolje razumevanje genotipova, spoljnih sredina i njihovih interakcija primenjene su AMMI (Additive Main Effect and Multiplicative Interaction) i klaster analiza. Raščlanjujući interakciju na glavne komponente PCA (Principal Components Analysis) analizom, otkriva se deo sistematskog variranja koji se nalazi u interakciji, a koji je kod genotipova vezan za dužinu vegetacije, a kod spoljnih sredina za količinu padavina u toku vegetacije. Rezultati grupisanja klaster analizom su u visokoj saglasnosti sa grupisanjem koje se uočava na biplotu AMMI modela.

Primljeno 29. X. 2009.

Odobreno 19. I. 2010.