Scientific Innovation Magazine
ISSN : 2584-1157

Statistical Method Used in Stability Analysis for Genotype x Environment Interaction

Abhishek Yadav, Gaurav Shukla and Umesh Chandra

Department of Statistics and Computer Science, CoA, BUAT, Banda (U.P.)

E-Mail: abhishekyadav20121999@gmail.com

Abstract

Stability analysis provides a general summary of the response patterns of genotypes to environmental change. Different approaches have been used to distinguish genotypes for their behaviour in different environmental conditions. There are various numerical and graphically stability methods to analyze G x E interaction and determine most stable genotype under varying environmental conditions. The numerical approaches which used in stability analysis are known as parametric and nonparametric methods. In the context of genotype-environment interaction, stability analysis is an essential part of agricultural research and current crop breeding. Stability analysis leads to the development of resilient and adaptable crop varieties, ultimately improving food security and agricultural sustainability in the face of changing environmental conditions by offering a quantitative framework for evaluating genotype performance across a variety of environments.

Keywords: Stability analysis, Parametric method, Nonparametric method, Environment.

Introduction:

Genotype x environment interaction (GEI) is major problem in the comparison of genotype performance across environments. The researches very often ignore GEI in their recommendation for crop growing. It is necessary to use corresponding statistical techniques for the efficient assessment of interaction. Interaction among genotypes and environment studied and interpreted by a wide genotype of statistical models and methodologies. The performance of any crop genotype actually depends on the effect of its genotype and environment in which it grows. Therefore, the phenotypic variation can be expressed as the sum of the two-component representing genotype and environmental source of variation. Genotypes under assessment are grown in various locations and over a number of years to know the importance of G x E interaction and the stability of performance. Leon defined two concepts of stability based on the goal and on the characteristics under consideration, which are termed "static" and "dynamic" concepts of stability. In the static concept, a specific stable genotype has a performance that is unaffected by the environmental conditions. Furthermore, this concept is analogous to the biological concept of stability such that the yield performance of a stable genotype has an environmental variance near to zero. The dynamic concept states that a stable genotype has no deviation from predictable response to environments. In other words, the performance of a stable genotype is accordance with the estimated level or the prediction for each environment. Thus, the genotypic response to environmental conditions is not equal for all genotypes. A wide array of statistical techniques has been proposed to analyze the adaptability of genotypes. A number of parametric statistical procedures have been developed over the years to analyze genotype x environment interaction and especially yield stability over environments. Some parametric and non parametric methods used in stability analysis are given below:

1. **Parametric Methods**

(a) Eberhart and Russel's regression model

Eberhart and Russel (1966) developed Finlay and Wilkinson's (1963) regression concept of stability and suggested the use of two stability parameters.

$$b_i = \sum_j \frac{(Y_{ij} - \bar{Y}_{i.})(\bar{Y}_{.j} - \bar{Y}_{..})}{(\bar{Y}_{.j} - \bar{Y}_{..})^2}$$

$$= l + [\sum_j \frac{(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})(\bar{Y}_{.j} - \bar{Y}_{..})}{(\bar{Y}_{.j} - \bar{Y}_{..})^2}$$

As a second measure,

$$s_{d_i}^{1^2} = [\sum_j \hat{\delta}_{ij}{}^2/(s-2)] - \bar{S}_e{}^2$$

Where $\bar{S}_e{}^2$, the average error is defined as $\bar{S}_e{}^2 = \left(\frac{\sum_j s_j{}^2}{sr}\right)$, $s_j{}^2$'s being the error mean squares for different experiments, each conducted with the same number of replications r.

(b) Francis and Kannenberg's coefficient of variability ((CV_i))

The mean CVi analysis introduced by Francis (1977) was designed to aid in studies on the physiological basis of yield stability. He introduced a simple graphical approach to assess performance and stability concurrently. It measures the performance and CVi for each genotype over all environments and the mean yield plotted against the CVi. It was found to characterize genotypes in groups rather than individually (Francis and Kannenberg, 1978).

$$s_i^2 = \frac{\sum_{i=1}^{q}\left[X_{ij} - \bar{X}_{i.}\right]^2}{q-1} \qquad CV_i = \frac{s_i^2}{\bar{X}_i} \times 100$$

(c) Lin and Binns cultivar performance measure (Pi)

Lin and Binns (1988) defined the superiority measure (Pi) of the ith test cultivar as the MS of distance between the ith test cultivar and the maximum response and it is given as

$$P_i = \sum_{j-1}^{q} (X_{ij} - M_{.j})^2 \,/2q$$

Where, P_i = Superiority measure

X_{ij} = Yield of the i^{th} genotype grown in the j^{th} location,

M_j= Maximum yield in the j^{th} location

(d) Shukla's stability variance parameter

Shukla (1972) defined the genotype's stability variance as its variance across environments following the removal of the primary influences of environmental means. The stability variance in a two-way analysis is determined by the residual (GE_{ij}+e_{ij}) matrix, as the genotype main effect remains constant and it is given as

$$\hat{\sigma}_i^2 = \frac{1}{(G-1)(G-2)(E-2)}\left[G(G\text{-}1)\sum_i(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 - \sum_i\sum_j(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2\right]$$

Where, Y_{ij} = is the mean yield of the i^{th} genotype in the j^{th} environment,

$\bar{Y}_{i.}$ = mean of the genotype i in all environments,

$\bar{Y}_{.j}$ = mean of all genotypes in j^{th} environments

$\bar{Y}_{..}$ =mean of all genotypes in all environments.

(e) Wricke's Ecovalence (W_i)

Wricke's (1962) defined the concept of ecovalence as the contribution of each genotype to the GEI sum of squares. The ecovalence (Wi) or stability of the ith genotype is its interaction with the environments, squared and summed across environments, and express as

$$W_i = [\bar{Y}_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} - \bar{Y}_{..}]^2$$

Where, $\bar{Y}_{ij}$ = Mean performance of genotype I in the j^{th} environment

$\bar{Y}_{i.}$ And $\bar{Y}_{.j}$ = genotype and environment mean deviations

$\bar{Y}_{..}$ = Overall mean

Plant breeders and researchers utilize Wricke's ecovalence, which is comparable to Shukla's stability variance parameter, to evaluate the stability of genotypes and choose those that show good mean performance, reliability, and adaptability in a variety of environmental situations. The development of cultivars that provide steady and high yield which is a prerequisite for sustainable agriculture practice, requires these stability measures.

(f) Finlay and Wilkinson's joint regression analysis (bi)

Finlay and Wilkinson (1963) defined a genotype with bi = 0 as stable. Once the genotype-environment interaction in usual analysis of variance is found significant, next by taking genotypic means of any genotype at different environments as dependent variable and environmental means as independent variable one can frame regression equations for different genotypes on environmental means. Thus, the sum of squares due to interactions is partitioned in to two components viz. sum of square due to regression and deviation from regression.

Let there be i (i= 1,2,.. v) number of genotypes to be tested in j (j=1,2,..s) number of environments, then

$$\bar{g}_i = \frac{\sum_{j=1}^{s} \bar{y}_{ij}}{s}; \quad \bar{e}_j = \frac{\sum_{i=1}^{v} \bar{y}_{ij}}{v} \quad \text{and} \quad b_i = \frac{\text{cov}(\bar{y}_{ij}, \bar{e}_j)}{\text{var}(\bar{e}_j)}$$

According to the model of Finlay and Wilkinson the regression coefficient is the stability parameter.

(g) Pinthus's Coefficient of Determination (R^2)

This statistic is defined as predictability of response suggested by Pinthus as another stability parameter, in which a variation of mean yield was explained by genotype response across environments. This parametric statistic can be described with the following equation:

$$R^2 = \frac{b_i^2 \sum (\bar{x}_{.j} - \bar{x}_{..})^2}{\sum (x_{ij} - \bar{x}_{i.})^2}$$

where bi is slope regression, x_{ij} is the grain yield of i^{th} genotype in j^{th} environment, $\bar{x}_{i.}$ is the mean grain yield of i^{th} genotype; $\bar{x}_{.j}$ is the mean grain yield of the j^{th} environment; and $\bar{x}_{..}$ is the grand mean. A genotype with the highest value is intended to be more stable.

2. Non Parametric Methods

(a) Thennarasu's Statistics

Thennarasu (1995) proposed another set of nonparametric statistics ($NP_i^{(1)}$, $NP_i^{(2)}$, $NP_i^{(3)}$, $NP_i^{(4)}$), based on ranks of adjusted means of the genotypes in each environment, and defined stable genotypes as those whose position remained unaltered in relation to the others in the set of environments assessed. These were calculated as follows:

$$NP_i^{(1)} = \frac{1}{m}\sum_{j=1}^{m} |r_{ij}{}^* - M_{di}{}^*|$$

$$NP_i^{(2)} = \frac{1}{m}\sum_{j=1}^{m} \frac{|r_{ij}{}^* - M_{di}{}^*|}{M_{di}}$$

$$NP_i^{(3)} = \frac{\sqrt{\sum(r_{ij}{}^* - \bar{r}_{i.}{}^*)^2/m}}{\bar{r}_{i.}}$$

$$NP_i^{(4)} = \frac{2}{m(m-1)}[\sum_{j=1}^{m-1}\sum_{j^1=j+1}^{m} \frac{|r_{ij}{}^* - r_{ij^1}{}^*|}{\bar{r}_{i.}}$$

(b) Nassar and Huhn

Six nonparametric techniques for evaluating GEI and stability analysis were put out by Huehn (1979). The phenotypic value of the i^{th} genotype in the j^{th} environment was designated as $x_{ij}$for a two-way dataset consisting of k genotypes and n environments, where i = 1,2,...,k, j = 1,2,...,n, r_{ij} representing the rank of the i^{th} genotype in the j^{th} environment, and r_{ij} representing the mean rank of the i^{th} genotype across all environments. The nonparametric stability statistic$S_i^{(4)}$ is comparable to that of Yau and Hamblin (1994), who employed relative yield to produce a yield stability measure in addition to assigning equal weight to each environment.The yield ranks of genotypes in each environment were used to calculate the following statistics:

$$\boldsymbol{S_i^{(1)} = 2\sum_j^{n-1}\sum_{j^1=j+1}^{n} |r_{ij} - r_{ij^1}|/[n(n-1)]}$$

$$\boldsymbol{S_i^{(2)} = \sum_{j=1}^{n}(r_{ij} - \bar{r}_i)^2 \Big/ \sum_{j=1}^{n}|r_{ij} - \bar{r}_{i.}|}$$

$$\boldsymbol{S_i^{(3)} = \frac{\sum_{j=1}^{n}(r_{ij} - \bar{r}_{i.})^2}{\bar{r}_{i.}}}$$

$$S_i^{(4)} = \sqrt{\frac{\sum_{j=1}^{n}(r_{ij} - \bar{r}_{i.})^2}{n}}$$

$$S_i^{(5)} = \frac{\sum_{j=1}^{n}|r_{ij} - \bar{r}_{i.}|}{n}$$

$$S_i^{(6)} = \frac{\sum_{j=1}^{n}|r_{ij} - \bar{r}_{i.}|}{\bar{r}_{i.}}$$

(c) Kang's Rank

Kang's Rank is based on yield and Shukla's stability variance (σ_i^2)**.** This parameter gives a weight of one to both yield and stability statistic to identify high yielding and stable genotype. The genotype with highest yield and lower σ_i^2 assigned with rank one separately then we added both ranks for each genotype. The genotype with lowest rank sum is known as most desirable.

Conclusion:

The importance of the GEI can be revealed from the relative contributions of the new varieties and followed by improved management to performance increases from direct comparisons of performances of them with old varieties in a single experiment. It is clear that the selection of genotypes for target environment(s) is affected by the GEI effect. For this reason, over the three past decades, numerous statistical models and approaches have been proposed to analyze GEI as well as identify the high-yielding and most stable genotypes. This fact is not unexpected in that each stability parameter or statistic result shows a specialist ranking pattern for genotypes. Hence, in each experiment, it is best that plant breeders compute all statistics and ultimately select the superior genotypes based on their yield performance and stability.

References:

1. Eberhart, S.A.T and Russell, W.A. (1966): Stability parameters for comparing varieties.

Crop Sci., 6, 36–40.

2. Kang, M.S. (1998): Using genotype-by-environment interaction for crop cultivar development. Adv. Agron, 62, 199–252
3. Leon, J. (1985): Beiträge zur Erfassung der Phänotypischen Stabilität; Christian-Albrecht University of Kiel: Kiel, Germany.
4. Nassar, R. and Huehn, M. (1987): Studies on estimation of phenotypic stability: Tests of significance for nonparametric measures of phenotypic stability. Biometrics, 43, 45–53.
5. Pinthus, J.M.(1973): Estimate of genotype value: A proposed method. Euphytica, 22, 121–123.
6. Shukla, G.K. (1972): Some statistical aspects of partitioning genotype-environmental components of variability. Heredity, 28, 237–245.
7. Yates, F. and Cochran, W.G. (1938): The analysis of groups of experiments. J. Agric. Sci., 28, 556–580.