A comparison of statistical methods for assessing winter wheat grain yield stability

A.F. Cheshkova[1] ✉, P.I. Stepochkin[2], A.F. Aleynikov[1], I.G. Grebennikova[1], V.I. Ponomarenko[2]

[1] Siberian Federal Scientific Center of Agro-BioTechnologies of the Russian Academy of Sciences, Krasnoobsk, Novosibirsk region, Russia
[2] Siberian Research Institute of Plant Production and Breeding – Branch of the Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Krasnoobsk, Novosibirsk region, Russia
✉ e-mail: cheshanna@yandex.ru

Abstract. The multitude of existing methods for assessing the phenotypic stability of plants makes breeders be faced with the problem of choosing an appropriate variant. The purpose of this study was to compare different methods of analyzing the genotype×environment interaction and, on their basis, assess the stability of the yield of 7 varieties of winter wheat. The article compares 17 stability statistics by applying them to data obtained from agrotechnical experiments carried in 2009–2011 for evaluating the grain yield of 7 varieties of winter common wheat of Siberian selection (Novosibirskaya 32, Novosibirskaya 40, Novosibirskaya 51, Novosibirskaya 3, Novosibirskaya 2, Obskaya winter, Omskaya 6). Analysis of variance revealed a significant ($p < 0.001$) genotype×environment interaction in the experiments, which indicates a different reaction of genotypes to changes in environmental conditions. Genotypes were ranked according to the level of stability. Based on the analysis of the rank correlation matrix, the stability statistics were categorized in five groups. Recommendations were made on which group of methods to use depending on the objectives of the study. In the case when the goal of breeding research is the selection of the most biologically stable varieties with the minimum variance across a range of environments, one should use the methods of the static concept. If it is necessary to choose a genotype with a predictable reaction to changes of environmental conditions, corresponding to the calculated level or forecast, the regression approach is the most appropriate. The stability statistics generally identified Novosibirskaya 32 as the most stable variety from a biological point of view. The regression approach showed that Novosibirskaya 3 was the genotype with the smallest deviation from mean yield in all environments, while methods accessing the contribution of each genotype to the genotype×environment interaction defined Novosibirskaya 51 as the most stable variety.
Key words: phenotypic stability; genotype×environment interaction; plant breeding; statistical methods; winter wheat.

Сравнение статистических методов оценки стабильности урожайности озимой пшеницы

А.Ф. Чешкова[1] ✉, П.И. Стёпочкин[2], А.Ф. Алейников[1], И.Г. Гребенникова[1], В.И. Пономаренко[2]

[1] Сибирский федеральный научный центр агробиотехнологий Российской академии наук, р. п. Краснообск, Новосибирская область, Россия
[2] Сибирский научно-исследовательский институт растениеводства и селекции – филиал Федерального исследовательского центра Институт цитологии и генетики Сибирского отделения Российской академии наук, р. п. Краснообск, Новосибирская область, Россия
✉ e-mail: cheshanna@yandex.ru

Аннотация. Многообразие существующих методов оценки фенотипической стабильности растений ставит перед селекционерами проблему выбора подходящего варианта. Целью настоящего исследования было сравнение различных методов анализа взаимодействия генотип×среда и оценка на их основе стабильности урожайности семи сортов озимой пшеницы. Проанализированы 17 статистических показателей стабильности на примере данных полевого опыта 2009–2011 гг. по оценке урожайности зерна семи сортов озимой мягкой пшеницы сибирской селекции (Новосибирская 32, Новосибирская 40, Новосибирская 51, Новосибирская 3, Новосибирская 2, Обская озимая, Омская 6) в шести вариантах сред. Дисперсионный анализ выявил значимое ($p < 0.001$) взаимодействие генотип×среда в опыте, что говорит о различной реакции генотипов на изменение условий среды. Выполнено ранжирование сортов по уровню стабильности и рассчитаны корреляционные связи между параметрами стабильности. На основе анализа корреляционной матрицы рангов проведена классификация методов, разделяющая их на пять групп. Предложены рекомендации по выбору способа определения стабильности генотипов в зависимости от целей исследования. В случае, когда целью селекционных исследований является выбор наиболее стабильных в биологическом смысле сортов, обладающих наименьшей вариацией признака, независимо от меняющихся условий среды, следует использовать методы статической концепции. Если среди набора сортов необходимо выбрать генотип с предсказуемой реакцией на изменения условий среды, соответ-

ствующей расчетному уровню или прогнозу, наиболее оптимален регрессионный подход. В результате оценки методами статической концепции сорт Новосибирская 32 был определен как наиболее стабильный с биологической точки зрения. Сорт Новосибирская 3, имевший наименьшее отклонение от средней урожайности для всех сред, оказался наиболее стабильным при оценке регрессионными методами. Методы, оценивающие вклад генотипа во взаимодействие генотип × среда, определили сорт Новосибирская 51 как самый стабильный.

Ключевые слова: фенотипическая стабильность; взаимодействие генотип × среда; селекция растений; статистические методы; озимая пшеница.

Introduction

Genotype × environment (GE) interaction means that varieties react differently to changes in growing conditions. Relative ranking of the productivity of a group of varieties in different environmental conditions gives different results. Even among specific, regionally adapted breeding forms, it is difficult to choose the best genotype, due to the variability of yields in different years of observations and under different environmental conditions.

Currently, many different ways of evaluating the GE interaction are developed, which are based on statistical methods for calculating certain parameters characterizing the degree of genotype response to changing environmental conditions. Different terms are used: adaptability, homeostaticity, plasticity, stability, etc.

Becker and Leon (1988) consider that, depending on the purpose of the research, there are two main differing concepts of stability: static (biological) and dynamic (agronomic).

The static concept implies a stable genotype possesses an unchanged performance, regardless of variation of the environmental conditions. "This concept of stability is useful for traits the levels of which have to be maintained at all costs, e. g. for quality traits, for resistance against diseases, or for stress characters like winter hardiness". The concept may include methods described by Roemer (1917, in: Becker, Leon, 1988), Francis and Kannenberg (1978), Lin and Binns (1988).

Unlike the static concept, the dynamic concept permits a predictable reaction of the genotype to changes in environmental conditions. The stable genotype, in accordance with this concept, does not deviate from the average values of this reaction. For each environment, the productivity of a stable genotype corresponds exactly to the predicted level. Among the methods of the dynamic concept used to assess stability, the following groups can be distinguished:

- methods that assess the stability of the genotype, based on the contribution to the variation of genotype × environment interaction are presented in the articles of Wricke (1962), Shukla (1972);
- a regression approach to assessing the stability of a trait is considered in the articles of Finlay and Wilkinson (1963), Eberhart and Russell (1966), Tai (1971);
- non-parametric methods based on relative data rankings were proposed by Nassar and Huehn (1987), Fox et al. (1990), Kang and Pham (1991);
- multidimensional methods that use the results of a cluster or a component analysis to assess stability were described in Fox and Rosielle (1982), Zobel et al. (1988), Lin and Binns (1991), Purchase (2000), Gauch (2006).

Comparison of different stability statistics and classification of methods was made by Lin and Binns (Lin et al., 1986), Becker and Leon (1988), Flores et al. (1998), Adugna and Labuschagne (2003), Mohammadi and Amri (2008), Moham-

madi et al. (2016). Kang (1997) made a wide analysis of the GE interaction phenomenon and gave recommendations how to use this interaction for crop cultivar development.

But some methods for stability assessing suggested by Khangildin et al. (1979), Dragavtsev (Dragavtsev, Averyanova, 1983), Udachin (1990), Martynov (1990), Kilchevsky and Khotyleva (1989), were not mentioned and compared in that works.

The purpose of this study is to compare 17 different methods of analyzing the GE interaction and, on their basis, assess the stability of the yield of 7 varieties of Siberian winter wheat.

Materials and methods

Experimental data

All analyses were performed on the grain yield data obtained from 7 varieties of winter wheat *T. aestivum* L.: Novosibirskaya 32, Novosibirskaya 40, Novosibirskaya 51, Novosibirskaya 3, Novosibirskaya 2, Obskaya winter, Omskaya 6. The trials were conducted at the experimental fields of the Siberian Research Institute of Plant Production and Breeding (SibNIIRS) – Branch of the Institute of Cytology and Genetics of SB RAS during three cropping seasons (2009–2011). The experiment was carried out in two variants each year: with use of mustard protection coulisses and without them. The meteorological conditions during the years of the experiments were different both in temperature and precipitation. Coulisses contributed to a greater accumulation of snow, as compared with no-coulisse plots, which allowed a long period to retain soil moisture in spring and early summer periods. Thus, contrasting growth conditions were provided, which made it possible to take into account six variants of environments for stability analysis.

The trials were evaluated using randomized complete block design with three replications. Sowing was carried out on the fallow precursor on plots of 25 m². Seeding rate was 600 seeds/m², sowing date was August 20–24. The grain yield was measured for each plot and then converted into tons/ha for further statistical analysis.

Statistical methods

The yield data were subjected to statistical analyses using the R (R Development Core Team, 2014) software. A model of analysis of variance (ANOVA) for *v* genotypes, *n* environments, *r* replications of the following form was considered:

$$Y_{ij} = \mu + d_i + \epsilon_j + g_{ij} + \bar{e}_{ij}, \qquad (1)$$

where Y_{ij} is the mean grain yield of the *i*-th genotype in the *j*-th environment ($i = 1, \dots v$; $j = 1, \dots n$); μ is the grand mean; d_i is the main effect of the *i*-th genotype; ϵ_j is the main effect of the *j*-th environment; g_{ij} is the effect of genotype × envi-

А.Ф. Чешкова, П.И. Стёпочкин, А.Ф. Алейников
И.Г. Гребенникова, В.И. Пономаренко

Сравнение статистических методов оценки
стабильности урожайности озимой пшеницы

2020
24 • 3

ronment interaction; $\bar{e}_{ij}$ is the GE interaction residual of the i-th genotype in the j-th environment.

Normality of model residuals was assessed by Shapiro–Wilk test (Shapiro, Wilk, 1965), and homogeneity of variance was estimated based on the criterion of Levene (1960).

The following methods were applied to experimental data (coded abbreviation was used for convenience).

1. (EV). Roemer (1917, in: Becker, Leon, 1988) used the "environmental variance of genotypes" to determine the stability of a genotype:

$$S_{yi}^2 = \frac{\Sigma_j (Y_{ij} - \bar{Y}_{i.})^2}{n-1}, \tag{2}$$

where Y_{ij} is the mean value of the trait for the i-th genotype in the j-th environment ($i = 1, \ldots v; j = 1, \ldots n$); $\bar{Y}_{i.}$ is the mean value of the trait for the i-th genotype in all environments. The most stable genotype has the lowest value of environmental variance. In fact S_{yi}^2 is unbiased estimation of genotype variation.

2. (CV). Francis and Kannenberg (1978) used a combination of mean yield and standard deviation as a measure of stability. The coefficient of variation of the i-th genotype is estimated by the standard formula:

$$CV_i = \frac{\sigma_i}{\bar{Y}_{i.}} \cdot 100 \%, \tag{3}$$

where σ_i is the standard deviation of the trait for the i-th genotype; $\bar{Y}_{i.}$ is the mean value of the trait for the i-th genotype in all environments.

Genotypes with yield above overall mean yield and CV_i below overall coefficient of variation are considered more stable than the others.

3. (HOM). Khangildin (Khangildin et al., 1979) uses the "coefficient of homeostaticity" as stability index:

$$Hom_i = \frac{\bar{Y}_{i.}^2}{\sigma_i(\bar{Y}_{i(opt)} - \bar{Y}_{i(lim)})}, \tag{4}$$

where σ_i is the standard deviation of the trait for the i-th genotype; $\bar{Y}_{i.}$, $\bar{Y}_{i(opt)}$, $\bar{Y}_{i(lim)}$ are the mean values of the trait for the i-th genotype in all environments, in the optimal and limited environment, respectively. The higher the coefficient of homeostasis is, the more stable is the genotype.

4. (Pi). Lin and Binns (1988) proposed to use the "superiority measure" P_i as a parameter of stability:

$$P_i = \Sigma_{j=1}^{n}(Y_{ij} - M_j)^2 / 2n, \tag{5}$$

where M_j is the maximal value Y_{ij} for all genotypes in j-th environment. Genotypes with a lower P_i value are considered to be more stable.

5. (UST). Udachin (1990) modified the method of Khangildin, proposing a new indicator – "steadiness of stability index":

$$U_i = \left(1 - \frac{I_{i(opt)} - I_{i(lim)}}{\bar{I}}\right) \cdot 100 \%. \tag{6}$$

Here $I_{i(opt)}$, $I_{i(lim)}$, $\bar{I}$ are stability indexes calculated by formulas:

$$I_{i(opt)} = \frac{\bar{Y}_{i(opt)}^2}{\sigma_{i(opt)}}, \quad I_{i(lim)} = \frac{\bar{Y}_{i(lim)}^2}{\sigma_{i(lim)}}, \quad \bar{I} = \frac{\Sigma_{ij} \bar{Y}_{ij}^2 / \sigma_{ij}}{v \cdot n}, \tag{7}$$

where σ_{ij}, $\sigma_{i(opt)}$, $\sigma_{i(lim)}$ are standard deviations of the trait for i-th genotype in j-th, optimal and limited environments,

respectively; $\bar{Y}_{ij}$, $\bar{Y}_{i(opt)}$, $\bar{Y}_{i(lim)}$ are average values of the trait for i-th genotype in j-th, optimal and limited environments; v is a number of genotypes, n is a number of environments. The higher this indicator is the more stable is the genotype.

6. (MAR). Martynov (1990) proposes to use a "weighted homeostacity index" for stability evaluation:

$$H_i = \Sigma_i a_{j(k)}(Y_{ij} - \bar{Y}_{.j})/\sigma_j, \tag{8}$$

where $a_{j(k)}$ is the calculated weighting factor determined for the j-th environment as follows.

All environments are divided into three groups with similar conditions: favorable ($\bar{Y}_{.j} > \bar{Y}_{..} + LSD$), medium ($\bar{Y}_{..} - LSD < \bar{Y}_{.j} < \bar{Y}_{..} + LSD$) and unfavorable ($\bar{Y}_{.j} < \bar{Y}_{..} - LSD$). The value of LSD (least significant difference) is determined by the formula:

$$LSD = t_\alpha \sqrt{MS_{GE}(n+1)/(nv)}, \tag{9}$$

where t_α is the value of the Student's t-test (for selected significance level α and $df = (v-1)(n-1)$); MS_{GE} is the mean square of G × E interaction.

Let's $N_k (k = 1, 2, 3)$ – number of elements within k-th group of environments, then the weighting coefficients $a_{j(k)}$ for each environment in k-th group are the same:

$$a_{j(k)} = \frac{n}{3} N_k \quad (k = 1, 2, 3). \tag{10}$$

The higher the H_i value, the more homeostatic the genotype.

7. (CAC). Kilchevsky and Khotyleva (1989) call the sum of the effects of the j-th environment and the interaction of genotype × environment as a "specific adaptive ability" of the i-th genotype in the j-th environment.

$$CAC_{ij} = d_j + (vd)_{ij} = \bar{Y}_{.j} - \bar{Y}_{..} + (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = Y_{ij} - \bar{Y}_{i.}. \tag{11}$$

As a measure of the stability of the i-th genotype, they use the CAC_i variance:

$$\sigma_{CAC_i}^2 = \frac{1}{n-1} \Sigma_j (d_j + (vd)_{ij})^2 - \frac{n-1}{1} \sigma^2 = $$
$$= \frac{\Sigma_j (Y_{ij} - \bar{Y}_{i.})^2}{n-1} - \frac{n-1}{n} \sigma^2. \tag{12}$$

8. (Wi). Wricke (1962) proposed to use the sum of squares of environmental effects as a measure of genotype stability. This parameter is called "ecovalence" and is calculated by the formula:

$$W_i = \Sigma_j (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2. \tag{13}$$

The more stable genotype is, the smaller its ecovalence value will be.

9. (SIGMA). Shukla (1972) developed an unbiased estimate using "stability variance" of genotypes:

$$\hat{\sigma}_i^2 = \frac{v(v-1)\Sigma_j(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 - \Sigma_{ij}(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2}{(n-1)(v-1)(v-2)}. \tag{14}$$

10–11. (Bi, S2di). Eberhart and Russell (1966) used a regression approach to assess the stability. The following model is considered:

$$Y_{ij} = \mu_i + \beta_i I_j + \delta_{ij}, \tag{15}$$

where Y_{ij} is the variety mean of the i-th variety in the j-th environment ($i = 1, \ldots v; j = 1, \ldots n$); μ_i is the mean of the i-th

variety over all environments; β_i is the regression coefficient, that measures the response of the i-th variety to varying environments; δ_{ij} is the deviation from regression of the i-th variety at the j-th environment; I_j is the environmental index of the j-th environment, calculated by the formula:

$$I_j = \frac{\Sigma_i Y_{ij}}{v} - \frac{\Sigma_i \Sigma_j Y_{ij}}{vn}. \quad (16)$$

The estimation of the regression coefficient b_i is the first parameter of genotype stability.

$$b_i = \frac{\Sigma_j Y_{ij} I_j}{\Sigma_j I_j^2}. \quad (17)$$

The second parameter of stability is the variance of deviation from the regression line.

$$s_{d_i}^2 = \frac{\Sigma_j \hat{\delta}_{ij}^2}{n-2} - \frac{s_e^2}{r}, \quad (18)$$

where s_e^2 is the estimate of the pooled error; r – number of replications. The sum of the squares of deviations from the regression is calculated by the formula:

$$\Sigma_j \hat{\delta}_{ij}^2 = (\Sigma_j Y_{ij}^2 - (\Sigma_j Y_{ij})^2/n) - (\Sigma_j Y_{ij} I_j)^2/\Sigma_j I_j^2. \quad (19)$$

A variety with $b_i = 1$, $s_{d_i}^2 = 0$ are referred to stable ones.

12–13. (Alpha, Lambda). The regression approach is also used by Tai (Tai, 1971; Tai, Young, 1972). A model of analysis of variance with random effects of the environment is considered:

$$y_{ijk} = \mu + d_i + \epsilon_j + \gamma_{k(j)} + g_{ij} + e_{ijk}, \quad (20)$$

where y_{ijk} ($i = 1, \dots v, j = 1, \dots n, k = 1, \dots r$) is the value of the trait of the i-th genotype in the j-th environment in the k-th replication; μ is the overall mean; d_i is the effect of the i-th genotype; ϵ_j is effect of the j-th environment; $\gamma_{k(j)}$ is the effect of replicates within environments; g_{ij} is the effect of genotype × environment interaction; e_{ijk} is the residual variation due to replications.

Tai and Young (1972) proposed to determine the linear response of the genotype to the effect of the environment α_i and the deviation from the linear response λ_i. They are related to the parameters of Eberhart and Russell in the following way:

$$\alpha_i = \frac{MSL(b_i - 1)}{MSL - MSB}, \quad (21)$$

$$\lambda_i = \left(\frac{v}{v-1}\right)\left(\frac{n-2}{n-1}\right)\frac{s_{d_i}^2}{MSE/r} - \alpha_i\left(\frac{b_i-1}{m-1}\right)\frac{MSB}{MSE}, \quad (22)$$

where MSL, MSB, MSE are mean squares, due to environments, replicates within environments and error, respectively. Tai considers that in terms of agricultural usage the preferred variety has ($\alpha = 0$, $\lambda = 1$) and a mean performance which is above-average of all varieties over a series of environments.

14. (Ai). The regression coefficient in the model of Eberhart and Russell depends on the average value of the trait. Varieties with a high average value will have a higher coefficient (scale effect). To eliminate this drawback Dragavtsev (Dragavtsev, Averyanova, 1983) proposed to use the dimensionless "coefficient of multiplicativity":

$$a_i = \frac{\bar{Y}_{i.} - b_i \bar{Y}_{..}}{\bar{Y}_{i.}}. \quad (23)$$

15–16. (S1, S2). Nassar and Huehn (1987) ranked genotypes in the j-th environment ($j = 1, \dots N$), assigning the minimum rank 1 to the genotype with the lowest corrected mean value $Y_{ij}^* = \left(Y_{ij} - (\bar{Y}_{i.} - \bar{Y}_{..})\right)$ ($i = 1, \dots K$), and the maximum rank K to the genotype with the highest corrected mean value of the trait. A genotype is considered stable if its ranks are close for different environments. The most stable genotype has a constant rank. The following parameters are proposed as a measure of stability:

$$S_i^{(1)} = 2 \cdot \Sigma_{j=1}^{N-1} \Sigma_{j'=j+1}^{N} |r_{ij} - r_{ij'}|/[N(N-1)], \quad (24)$$

$$S_i^{(2)} = \Sigma_{j=1}^{N} (r_{ij} - \bar{r}_{i.})^2/(N-1), \text{ where } \bar{r}_{i.} = \Sigma_{j=1}^{N} r_{ij}/N. \quad (25)$$

For the most stable genotypes $S_i^{(1)} = 0$, $S_i^{(2)} = 0$.

17. (ASVI). Zobel, Wright and Gauch (1988) suggested to use the principal component method and a graphical representation of stability based on the AMMI model (Additive Main effects and Multiplicative Interaction). The model has the form:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \Sigma_{k=1}^{N} \lambda_k \zeta_{ik} \eta_{jk} + \theta_{ij}, \quad (26)$$

where Y_{ij} is the mean value of the trait of i-th genotype in j-th environment ($i = 1, \dots v; j = 1, \dots n$); μ is an overall mean; α_i is the genotype mean deviation; β_j is the environment mean deviation; λ_k is the eigenvalue of the PCA axis; ζ_{ik}, η_{jk} are the genotype and environment PCA scores; N is a number of PCA axes, retained in the model; θ_{ij} is the residual.

Based on this model, Purchase (Purchase et al., 2000) proposed to use the ASV (AMMI Stability Value) parameter for ranking genotypes by stability:

$$ASV_i = \sqrt{\left[\frac{SS_{IPCA1}}{SS_{IPCA2}}(IPCA1score)\right]^2 + (IPCA2score)^2}, \quad (27)$$

where SS_{IPCA1}, SS_{IPCA2} are sums of squares of the first and second principal components; $IPCA1score$, $IPCA2score$ are projections of the i-th genotype on the first and second PCA axis. The smaller the value of the ASV parameter is, the more stable the genotype is considered.

18. (MeanY). In addition, values of mean yield were calculated for each genotype using standard formulas.

For comparison of various methods for assessing stability, 17 statistical indexes described above were calculated. Then all varieties were ranked according each of the stability statistics. The ranks were assigned in such a way that the rank 1 received the most desirable variety:

- for statistics *MeanY*, *HOM*, *UST*, *MAR*, *Ai*, rank 1 received varieties with the highest value of the index;
- for statistics *ASVI*, *S1*, *S2*, *Lambda*, *S2di*, *SIGMA*, *Wi*, *CAC*, *Pi*, *CV*, *EV*, rank 1 received varieties with the lowest value of the index;
- for regression coefficient *Bi*, rank 1 was assigned to the variety with the smallest (in modulo) coefficient's deviation from one;
- for *Alpha* coefficient, rank 1 was assigned to the variety with the smallest (in modulo) coefficient's deviation from zero.

А.Ф. Чешкова, П.И. Стёпочкин, А.Ф. Алейников
И.Г. Гребенникова, В.И. Пономаренко

Сравнение статистических методов оценки
стабильности урожайности озимой пшеницы

2020
24 • 3

Results

Mean grain yield and rank of 7 genotypes across 6 environments are shown in Table 1. The fact that genotypes switch ranks from one environment to another means the presence of a crossover GE interaction.

The result of analysis of variance (Table 2) shows that GE interaction is significant ($p < 0.001$), which indicates a significant difference between the responses of genotypes to changes in environmental conditions. Shapiro–Wilk test showed the normality of model residuals ($W = 0.99$, p-value = 0.31). The Levene test proved the homogeneity of variance ($F = 0.57$, p-value = 0.97).

Table 3 provides the estimates of mean grain yield, 17 stability statistics and the result of ranking of the seven winter wheat varieties according these statistics.

Taking mean yield as a first parameter for evaluating the genotypes, Novosibirskaya 3 gave the best mean yield while Novosibirskaya 32 had the lowest mean yield across environments. The highest variation of grain yield had Novosibirskaya 2, the lowest – Novosibirskaya 32. Genotypes Novosibirskaya 2, Novosibirskaya 3, Novosibirskaya 40, Obskaya winter, and Omskaya 6, with regression coefficients b_i higher than one were able to take advantage of favorable environments, whereas Novosibirskaya 32, and Novosibirskaya 51 with $b_i < 1$ were less responsive to environmental changes. According to Shukla's statistic, Novosibirskaya 51 provided the lowest contribution to total GE interaction, while Novosibirskaya 2 had the largest stability variance.

One can see the significant difference between methods in genotypes ranking. Such is the variety Novosibirskaya 32, which has the lowest mean yield and the lowest variance in

the experiment. It is defined as the most stable one by methods focusing on the variation of the yield (*EV*, *CAC*, *CV*, *HOM*, *UST*, *Ai*), and vice versa, as the most unstable, by methods focusing on the mean value of the yield (*Pi*, *MAR*, *Bi*, *Alpha*, *S1*).

For further determination of the interrelationships of the methods, paired correlations of ranks defined by the stability statistics were calculated by Kendall's method. The results are presented in Table 4 and graphically in the Figure.

High level of paired correlations was found between statistics *UST*, *CV*, *CAC*, *HOM*, *Ai* and *EV*. *Pi* and *MAR* were strongly correlated with the average yield of genotype *MeanY*. Another group of methods with significant paired correlations included *Wi*, *SIGMA*, *S2di*, *Lambda*, and *ASVI*.

Some of the considered methods were ranking genotypes identically: *CAC* and *EV*; *Bi* and *Alpha*; *S2di* and *Lambda*; *Wi* and *SIGMA*.

Discussion

According the rank correlation matrix the considered methods can be divided into five groups determined by a high level of paired rank correlations ($r > 0.8$).

Group 1. Includes the regression coefficient of Eberhart and Russell b_i and the similar coefficient of Tai α_i. In fact α_i is equal to $b_i - 1$ when the effect of replicate in environment is not considered. Becker and Leon (1988) pointed a strong correlation between b_i and Roemer's environmental variance S_{yi}^2. But when we proceeded to consideration of ranks then none correlation between these parameters was found. It is explained by the fact that for environmental variance rank 1 received the variety with S_{yi}^2 closest to zero, and for regression coefficient b_i, rank 1 was assigned to the variety with the

Table 1. Mean grain yield (t/ha) and range of 7 winter wheat varieties tested across 6 environments

Variety	2009				2010				2011			
	No-coulisses		In coulisses		No-coulisses		In coulisses		No-coulisses		In coulisses	
	Yield	R	Yield	R	Yield	R	Yield	R	Yield	R	Yield	R
Novosibirskaya 2	5.66	4	5.70	5	3.42	3	3.65	7	5.49	1	5.84	1
Novosibirskaya 3	5.86	2	6.50	1	3.74	1	4.84	1	4.67	4	5.62	2
Novosibirskaya 32	4.93	7	4.71	7	3.21	7	3.96	5	3.78	7	4.41	7
Novosibirskaya 40	5.64	5	6.08	4	3.36	4	3.86	6	4.65	5	5.28	4
Novosibirskaya 51	5.54	6	5.45	6	3.34	5	4.00	4	4.46	6	5.18	6
Obskaya winter	5.90	1	6.34	2	3.72	2	4.35	3	4.72	3	5.20	5
Omskaya 6	5.76	3	6.14	3	3.30	6	4.47	2	5.44	2	5.61	3

Table 2. The results of the variance analysis for grain yield of 7 winter wheat varieties tested across 6 environments

Source of variation	Df	Sum of squares	Mean square	F
Genotype	6	13.51	2.251	171.892***
Environment	5	89.30	17.860	1363.969***
Genotype × environment	30	9.18	0.306	23.378***
Error	84	1.10	0.013	

*** $p < 0.001$.

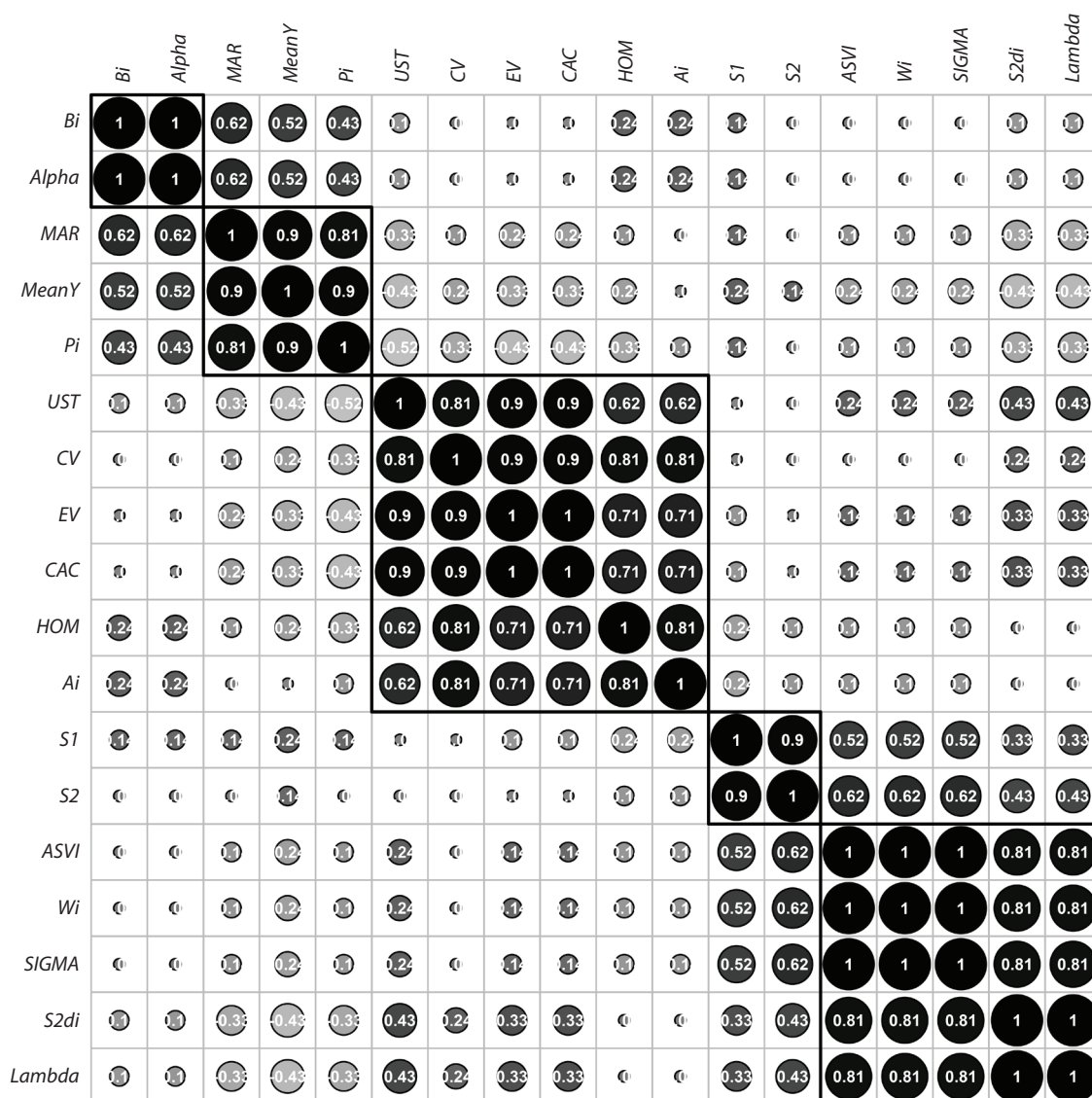**Table 3.** Stability statistics and ranks of 7 varieties of winter wheat tested across 6 environments

Method	Novosibirskaya 2		Novosibirskaya 3		Novosibirskaya 32		Novosibirskaya 40		Novosibirskaya 51		Obskaya winter		Omskaya 6	
	Val	R	Val	R	Val	R	Val	R	Val	R	Val	R	Val	R
MeanY	4.96	4	5.20	1	4.17	7	4.81	5	4.66	6	5.04	3	5.12	2
EV	1.24	7	0.97	4	0.41	1	1.11	6	0.78	2	0.96	3	1.10	5
CAC	1.24	7	0.97	4	0.41	1	1.11	6	0.77	2	0.95	3	1.10	5
Pi	0.18	4	0.06	2	0.85	7	0.20	5	0.30	6	0.11	3	0.04	1
CV	21.15	7	17.92	3	14.55	1	20.68	6	17.88	2	18.32	4	19.36	5
HOM	9.68	5	10.52	3	16.66	1	8.56	7	11.88	2	10.50	4	9.33	6
UST	−4.61	7	0.70	4	3.64	1	−0.98	5	1.01	2	0.85	3	−3.31	6
MAR	0.98	4	5.44	1	−9.10	7	−0.89	5	−2.52	6	3.23	2	2.85	3
Wi	1.06	7	0.39	5	0.76	6	0.15	2	0.08	1	0.25	3	0.37	4
SIGMA	0.05	7	0.01	5	0.03	6	0.00	2	0.00	1	0.01	3	0.01	4
Bi	1.10	4	1.02	1	0.65	7	1.14	6	0.95	3	1.03	2	1.11	5
S2di	0.25	7	0.09	6	0.06	4	0.01	2	0.01	1	0.06	3	0.08	5
Ai	−0.08	6	0.05	2	0.24	1	−0.15	7	0.01	3	0.00	4	−0.05	5
Alpha	0.10	4	0.02	1	−0.35	7	0.14	6	−0.05	3	0.03	2	0.11	5
Lambda	71.24	7	27.06	6	16.46	4	4.40	2	3.68	1	16.44	3	22.12	5
S1	3.07	6	2.47	3	3.13	7	1.27	1	1.80	2	2.80	5	2.53	4
S2	7.07	7	4.30	3	6.97	6	1.37	1	2.17	2	5.47	5	4.67	4
ASVI	2.20	7	1.11	5	1.56	6	0.42	2	0.29	1	0.75	3	0.90	4

N o t e: Refer to text for details of methods. Val is the value of the stability statistic, corresponding to the method; R is the genotype rank, corresponding to the method.

Table 4. Correlations of ranks between mean yield and stability measures for 7 varieties of winter wheat

Parameter	MeanY	EV	CAC	Pi	CV	HOM	UST	MAR	Wi	SIGMA	Bi	S2di	Ai	Alpha	Lambda	S1	S2
EV	−0.33																
CAC	−0.33	1.00															
Pi	0.90	−0.43	−0.43														
CV	−0.24	0.90	0.90	−0.33													
HOM	−0.24	0.71	0.71	−0.33	0.81												
UST	−0.43	0.90	0.90	−0.52	0.81	0.62											
MAR	0.90	−0.24	−0.24	0.81	−0.14	−0.14	−0.33										
Wi	−0.24	0.14	0.14	−0.14	0.05	−0.14	0.24	−0.14									
SIGMA	−0.24	0.14	0.14	−0.14	0.05	−0.14	0.24	−0.14	1.00								
Bi	0.52	−0.05	−0.05	0.43	0.05	0.24	−0.14	0.62	0.05	0.05							
S2di	−0.43	0.33	0.33	−0.33	0.24	0.05	0.43	−0.33	0.81	0.81	−0.14						
Ai	−0.05	0.71	0.71	−0.14	0.81	0.81	0.62	0.05	−0.14	−0.14	0.24	0.05					
Alpha	0.52	−0.05	−0.05	0.43	0.05	0.24	−0.14	0.62	0.05	0.05	1.00	−0.14	0.24				
Lambda	−0.43	0.33	0.33	−0.33	0.24	0.05	0.43	−0.33	0.81	0.81	−0.14	1.00	0.05	−0.14			
S1	0.24	−0.14	−0.14	0.14	−0.05	−0.24	−0.05	0.14	0.52	0.52	0.14	0.33	−0.24	0.14	0.33		
S2	0.14	−0.05	−0.05	0.05	0.05	−0.14	0.05	0.05	0.62	0.62	0.05	0.43	−0.14	0.05	0.43	0.90	
ASVI	−0.24	0.14	0.14	−0.14	0.05	−0.14	0.24	−0.14	1.00	1.00	0.05	0.81	−0.14	0.05	0.81	0.52	0.62

N o t e: Refer to text for details of methods.

А.Ф. Чешкова, П.И. Стёпочкин, А.Ф. Алейников
И.Г. Гребенникова, В.И. Пономаренко

Сравнение статистических методов оценки
стабильности урожайности озимой пшеницы

2020
24 • 3

Graphic representation of the rank correlation matrix ordered by the cluster method.

smallest (in modulo) coefficient's deviation from one. And so we consider b_i not as a measure of stability but as additional information on the average response of a genotype to advantageousness of environmental conditions.

Group 2. Includes the superiority index Pi of Lin and Binns and Martynov's weighted homeostaticity index H_i. These statistics were strongly correlated with yield. Breeding based on these methods would favor breeding for yield, as Kang and Pham (1991) and Flores (Flores et al., 1998) found.

Group 3. Includes steadiness of stability index of Udachin U_i, coefficient of variation of Francis and Kannenberg CV_i, variance of the specific adaptive ability of Kilchevsky and Khotyleva $\sigma^2_{CAC_i}$, Roemer's environmental variance S^2_{yi}, Khangildin's coefficient of homeostaticity Hom_i, Dragavtsev's coefficient of multiplicativity a_i. All these parameters are connected to genotype variance. The identicall genotypes ranking was received for $\sigma^2_{CAC_i}$ and S^2_{yi}, because the variance of specific adaptive ability of Kilchevsky and Khotyleva is

nothing more than the Roemer's environmental variance, reduced by error of experiment.

Group 4. Includes non-parametric estimates of stability of Nassar and Huehn $S_i^{(1)}$, $S_i^{(2)}$. Becker and Leon (1988), Adugna and Labuschagne (2003), Mohammadi and Amri (2008), reported that these statistics were highly correlated with Shukla's stability variance $\hat{\sigma}_i^2$ and with Eberhart and Russell's $s^2_{d_i}$. But we received only medium correlation of ranks for them.

Group 5. Includes Purchase's parameter based on AMMI model ASV_i, ecovalence of Wricke W_i, Shukla's stability variance $\hat{\sigma}_i^2$, Eberhart and Russell's variance of deviation from the regression $s^2_{d_i}$, and Tai's λ_i variance. The W_i and $\hat{\sigma}_i^2$ parameters produced the same ranking of genotypes wich is in agreement with Kang et al. (1987). The ranks established by W_i and $s^2_{d_i}$ were highly corrrelated because the ecovalence and the deviation from regression are mathematically linked (Flores et al., 1998). Also, $s^2_{d_i}$ and λ_i identically ranged the genotypes because they are mathematically linked. As expected, the statistics

$\hat{\sigma}_i^2$ and $s_{d_i}^2$ were highly rank-correlated too. The multivariate parameter ASV_i produced the same ranging of genotypes as W_i and $\hat{\sigma}_i^2$ in our experiment, which is in agreement with strong correlation of these parameters demonstrated by Purchase et al. (2000).

According to Becker and Leon's (1988) classification, the Group 3 presents methods related to the static (biological) concept of stability, and all other groups of methods refer to dynamic (agronomic) concept of stability.

Lin et al. (1986) suggested the classification of methods which contains three concepts of stability:

Type 1: A genotype is considered to be stable if its among-environment variance is small.

Type 2: A genotype is considered to be stable if its response to environments is parallel to the mean response of all genotypes in the trial.

Type 3: A genotype is considered to be stable if the residual MS from the regression model on the environmental index is small.

According to this classification the statistics of Group 1 corresponds to Type 2 stability, those of Group 3 to Type 1, and those of Group 4 and Group 5 to Type 3 stability.

The choice of the appropriate method for evaluation the stability of genotypes depends on the objectives of the study. In the case when the goal of breeding research is the selection of the most biologically stable varieties with the minimum variance across a range of environments, the stability is defined in the sense of homeostasis, one should use the methods of the static concept from Group 3. It should be remembered that the most stable varieties in this concept may have relatively low yields. Among the varieties of winter wheat under study, the variety Novosibirskaya 32 was the most stable from the biological point of view. This variety has the smallest variance, the smallest coefficient of variation, and the greatest coefficient of homeostaticity. At the same time, this variety showed the lowest yield in the experiment.

If it is necessary to choose a genotype with a predictable reaction to changes of environmental conditions, corresponding to the calculated level or forecast, the regression approach (Group 1) is the most appropriate. The most stable will be the genotype, which has the smallest deviation from the mean yield in all environments. The variety Novosibirskaya 3 with the smallest deviation of b_i from one fits this definition.

To compare the contributions of each genotype to the GE interaction, one should use the methods presented in Group 5. In these methods, the genotype that makes the least contribution to the interaction is considered the most stable. In our experiments the variety Novosibirskaya 51 was defined as the most stable from this point of view.

In the nonparametric methods of Group 4, the concept of stability is the same as in the methods of Group 5. Nonparametric methods should be used in the case when the initial data do not meet the requirements of normal distribution and homogeneity of variance. The most suitable genotype for this group was Novosibirskaya 40.

The methods from Group 2 cannot be recommended for stability evaluation, because they are strongly correlated with the mean yield and breeding based on these methods would favor breeding for yield.

Conclusions

The considered methods reflect different aspects of GE interaction. As Flores et al. (1998) mentioned: "…all the methods examined here to study the stability of genotypes are valid, although they are based on very different concepts of stability". The breeder can choose the appropriate method depending on whether the breeding is to be based primarily on yield, primarily on stability, or simultaneously on yield and yield stability.

In our experiment the methods were applied to a small number of unique cultivars and environments. In order to find stable genotypes of winter wheat for breeding purpose farther field experiments need to be conducted.

References

Adugna W., Labuschagne M.T. Parametric and nonparametric measures of phenotypic stability in linseed (*Linum usitatissimum* L.). Euphytica. 2003;129:211-218. DOI 10.1023/A:1021979303319.

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

Dragavtsev V.A., Averyanova A.F. Redefinition of genetic formulas of quantitative traits in different environmental conditions. Genetika = Genetics. 1983;19(11):1806-1810. (in Russian)

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

Finlay K.W., Wilkinson G.N. The analysis of adaptation in a plant breeding programme. Aust. J. Agric. Res. 1963;14:742-754. DOI 10.1071/AR9630742.

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

Fox P.N., Rosielle A.A. Reducing the influence of environmental main-effects on pattern analysis of plant breeding environments. Euphytica. 1982;31:645-656. DOI 10.1007/BF00039203.

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

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

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

Kang M.S. Using genotype-by-environment interaction for crop cultivar development. Adv. Agron. 1997;62:199-252. DOI 10.1016/S0065-2113(08)60569-6.

Kang M.S., Miller J.D., Darrah L.L. A note on relationship between stability variance and ecovalence. J. Hered. 1987;78:107. DOI 10.1093/oxfordjournals.jhered.a110322.

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

Khangildin V.V., Shayakhmetov I.F., Mardamshin A.G. Homeostasis of crop components and prerequisites for creating a model of a spring wheat variety. In: Genetic Analysis of Quantitative Traits of Plants. Ufa, 1979;5-39. (in Russian)

Kilchevsky A.V., Khotyleva L.V. Genotype and Environment in Plant Breeding. Minsk: Nauka i Tekhnika Publ., 1989. (in Russian)

Levene H. Robust tests for equality of variances. In: Olkin I., Ghurye S.G., Hoeffding W., Madow W.G., Mann H.B. (Eds.). Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford Univ. Press, 1960;278-292.

Lin C.S., Binns M.R. A superiority measure of cultivar performance for cultivar × location data. Can. J. Plant Sci. 1988;68:193-198. DOI 10.4141/cjps88-018.

А.Ф. Чешкова, П.И. Стёпочкин, А.Ф. Алейников
И.Г. Гребенникова, В.И. Пономаренко

Сравнение статистических методов оценки
стабильности урожайности озимой пшеницы

2020
24·3

Lin C.S., Binns M.R. Assessment of a method for cultivar selection based on regional trial data. Theor. Appl. Genet. 1991;82:379-388. DOI 10.1007/BF02190626.

Lin C.S., Binns M.R., Lefkovitch L.P. Stability analysis: where do we stand? Crop Sci. 1986;26:894-900. DOI 10.2135/cropsci1986.0011 183X002600050012x.

Martynov S.P. A method for the estimation of crop varieties stability. Biom. J. 1990;7:887-893.

Mohammadi R., Amri F. Comparison of parametric and non-parametric methods for selecting stable and adapted durum wheat genotypes in variable environments. Euphytica. 2008;159:419-432. DOI 10.1007/s10681-007-9600-6.

Mohammadi R., Farshadfar E., Amri A. Comparison of rank-based stability statistics for grain yield in rainfed durum wheat. N. Z. J. Crop Hortic. Sci. 2016;44:25-40. DOI 10.1080/01140671.2015.1100126.

Nassar R., Huehn M. Studies on estimation of phenotypic stability: tests of significance for nonparametric measures of phenotypic stability. Biometrics. 1987;43:45-53. DOI 10.2307/2531947.

Purchase J.L., Hatting H., Van Deventer C.S. Genotype × environment interaction of winter wheat in South Africa: II. Stability analysis of yield performance. S. Afr. J. Plant Soil. 2000;17:101-107. DOI 10.1080/02571862.2000.10634878.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, 2014. Available at: http://www.R-project.org/

Shapiro S.S., Wilk M.B. An analysis of variance test for normality (complete samples). Biometrika. 1965;52(3-4):591-611. DOI 10.1093/biomet/52.3-4.591.

Shukla G.K. Some statistical aspects of partitioning genotype-environmental components of variability. Heredity. 1972;29:237-245. DOI 10.1038/hdy.1972.87.

Tai G.C.C. Genotypic stability analysis and application to potato regional trials. Crop Sci. 1971;11:184-190. DOI 10.2135/cropsci1971.0011183X001100020006x.

Tai G.C.C., Young D.A. Genotypic stability analysis of eight potato varieties tested in a series of ten trials. Am. Potato J. 1972;49:138-150. DOI 10.1007/BF02861594.

Udachin R.A. Methods of assessing the ecological plasticity of wheat varieties. Selektsiya i Semenovodstvo = Selection and Seed Production. 1990;5:2-6. (in Russian)

Wricke G. Tjber eine Methode zur Erfassung der okologischen Streubreite in Feldversuchen. Z. Pflanzenzuchtg. 1962;47:92-96.

Zobel R.W., Wright M.J., Gauch H.G., Jr. Statistical analysis of a yield trial. Agron. J. 1988;80:388-393. DOI 10.2134/agronj1988.0002196 2008000030002x.

ORCID ID

A.F. Cheshkova orcid.org/0000-0003-2265-7129