

H.-P. Piepho

Implication of correlations among some common stability statistics — a Monte Carlo simulations

Received: 4 August 1994 / Accepted: 30 September 1994

Abstract Stability analysis of multilocation trials is often based on a mixed two-way model. Two stability measures in frequent use are the environmental variance (S_i^2) and the ecovalence (W_i). Under the two-way model the rank orders of the expected values of these two statistics are identical for a given set of genotypes. By contrast, empirical rank correlations among these measures are consistently low. This suggests that the two-way mixed model may not be appropriate for describing real data. To check this hypothesis, a Monte Carlo simulation was conducted. It revealed that the low empirical rank correlation among S_i^2 and W_i is most likely due to sampling errors. It is concluded that the observed low rank correlation does not invalidate the two-way model. The paper also discusses tests for homogeneity of S_i^2 as well as implications of the two-way model for the classification of stability statistics.

Key words Phenotypic stability · Genotype × environment interaction · Rank correlation Stability measures · Monte Carlo simulation

Introduction

Several papers have investigated the rank correlations among stability statistics (Becker 1981; Léon 1985; Weber and Wricke 1987; Becker and Léon 1988; Pham and Kang 1988; Hühn 1990; Piepho and Lotito 1992; Helms 1993; Jalaluddin and Harrison 1993). In these investigations the rank correlation between ecovalence (Wricke 1962) and environmental variance was consistently low and non-significant in most cases.

Many stability statistics are derived by the usual two-way model for genotypes × environments. Considering the expected values under this model, a high correlation among the two above stability statistics is expected, which is in contrast to emprical findings. This may be due to one or both of two reasons: (1) either the model assumption is inappropriate or (2) the observed low rank correlation is a result of sampling errors.

The purpose of the present paper is to investigate, via Monte Carlo simulation, whether reason (2) alone can be responsible for the low rank correlation usually found in real data sets. Should the simulation not conform to empirical results, this would have far-reaching consequences for model choice in stability analysis.

Theory

The usual two-way linear model is given by

$$y_{ij} = \mu + g_i + e_j + v_{ij}, \qquad (1)$$

$$(i = 1, \ldots, K; \quad j = 1, \ldots, N)$$

where y_{ij}, μ, g_i, e_j, and v_{ij} are, respectively, the yield of the i^{th} genotype in the j^{th} environment, the grand mean, the effect of the i^{th} genotype, the effect of the j^{th} environment, and a residual corresponding to y_{ij}, which comprises both genotype × environment interaction and errors. It is usually assumed that environments are random, while genotypes are fixed. The random effects are assumed to be stochastically independent with variances Var $(e_j) = \sigma_e^2$ and Var $(v_{ij}) = \sigma_i^2$ (Shukla 1972).

The environmental variance is estimated by

$$S_i^2 = \frac{\Sigma_j (y_{ij} - y_{i\cdot})^2}{N-1}$$

where the dot notation indicates that the mean has be taken across the corresponding index. Wricke's ecovalence is given by

$$W_i = \Sigma_j (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2.$$

This is equivalent for ranking purposes to Shukla's estimator of the stability variance σ_i^2:

$$Q_i = \frac{K(K-1)\,W_i - \Sigma_i W_i}{(K-1)(K-2)(N-1)}.$$

The rank correlation between S_i^2 and W_i is always identical to that

Communicated by P. M. A. Tigerstedt

H.-P. Piepho
University of Kassel, Steinstrasse 19, 37213 Witzenhausen, Germany

between S_i^2 and Q_i. For the ideas to be developed here it will be convenient to consider Q_i along with W_i.

Using model (1) the expectations of S_i^2 and Q_i are found to be

$$E[S_i^2] = \sigma_e^2 + \sigma_i^2$$

and

$$E[Q_i] = \sigma_i^2.$$

Provided that no pair of stability variances is equal, the rank correlation among $E[S_i^2]$ and $E[Q_i]$ equals one.

Often, a regression model is deemed more appropriate for stability analysis (Shukla 1972):

$$y_{ij} = \mu + g_i + \beta_i e_j + \delta_{ij}, \tag{2}$$

where β_i is the regression coefficient of the i^{th} genotype and δ_{ij} is a deviation from the regression corresponding to y_{ij}. The regression coefficients are constrained by $\Sigma_i \beta_i = K$. The regression coefficient β_i is estimated by

$$b_i = \frac{\Sigma_j (y_{ij} - y_{i\cdot})(y_{\cdot j} - y_{\cdot\cdot})}{\Sigma_j (y_{\cdot j} - y_{\cdot\cdot})^2}.$$

Under model (2) the expections of S_i^2 and Q_i are

$$\mathrm{E}[S_i^2] = \beta_i^2 \sigma_e^2 + \sigma_{\delta i}^2$$

and

$$\mathrm{E}[Q_i] = K(K-2)^{-1} (\beta_i - 1)^2 \sigma_e^2 + \sigma_{\delta i}^2 + \text{constant}$$

where $\sigma_{\delta i}^2 = \mathrm{Var}(\delta_{ij})$. In cases where the regression coefficients do not depart much from unity and σ_e^2 is of the same order of magnitude as the $\sigma_{\delta i}^2$-values, the rank correlation between the expected values is still high. If, on the other hand, σ_e^2 is large and/or the regression coefficients depart considerably from unity, the rank correlation among $\mathrm{E}[S_i^2]$ and $\mathrm{E}[Q_i]$ diminishes (provided that β_i is independent of $\sigma_{\delta i}^2$ and vice versa). Conversely, by observing that

$$\mathrm{E}[b_i] \approx c(\beta_i \sigma_e^2 + \sigma_{\delta i}^2 / K)$$

where $c = \sigma_e^2 + \sigma_\delta^2/K$ and $\sigma_\delta^2 = \Sigma_i \sigma_{\delta i}^2 / K$ (Piepho 1993a), a high rank correlation between $\mathrm{E}[b_i]$ and $\mathrm{E}[S_i^2]$ is expected when $\sigma_{\delta i}^2$-values are not very variable. In this context it is useful to consider the following identity (Wricke and Weber 1980)

$$(N-1) S_i^2 = W_i + (2b_i - 1) \Sigma_j (y - y_{\cdots j})^2. \tag{3}$$

Equation (3) shows that when the rank correlation between S_i^2 and W_i (or Q_i) is low, the rank correlation between S_i^2 and b_i should be high. Moreover, when $\Sigma_j (y_{\cdot j} - y_{\cdot\cdot})^2$ (and hence σ_e^2) is large compared to W_i, small chance variations in b_i can lead to high rank correlation between S_i^2 and b_i whereas that between S_i^2 and W_i is low. This may happen even if β_i-values i.e., the true regression coefficients, are all equal, which is a first indication that low rank correlations between S_i^2 and W_i may occur even if model (1) is correct.

A simulation experiment

In order to further investigate the relations discussed in the foregoing section, a Monte Carlo simulation was conducted based on model (1). Random effects were assumed to be normally distributed. In various crops (oats, oilseed rape, sugar beets, fodder beets and faba beans) σ_e^2 was found to be about 5–20 times the value of $\sigma^2 = \Sigma_i \sigma_i^2 / K$ (Hühn et al. 1993). Therefore σ^2 was fixed to $\sigma^2 = 10$, while σ_e^2 took the values (50, 200). To create heterogeneity the stability variance was chosen as $\sigma_i^2 = D * i$ where $D = 2\sigma^2/(K+1)$. Alternatively, the case $\sigma_i^2 = \sigma^2$ was simulated. In addition, simulations were also conducted using model (2). Specifications were the same as under model (1) with σ_i^2 being replaced by $\sigma_{\delta i}^2$. β_i was set to $\beta_i = a + b*i + c*i^2$. The constants a, b, and c were so that $\beta_\cdot = 1$, $\beta_1 = \beta_K = 0.5$. The resulting β_i-values range from 0.5 to 1.33. With this choice there is no rank correlation between the true parameter values of β_i and $\sigma_{\delta i}^2$. Normal deviates were generated by the Box-Muller method which is implemented in the NORMAL function of SAS/IML (SAS 1989). The cases (K, N) = (10, 10), (20, 10), (10, 20), (20, 20) were considered. Thus, a total of 32 different cases was simulated. Rank correlations were computed among the following stability statistics: S_i^2, W_i, b_i and S_{di}^2 (an estimate of $\sigma_{\delta i}^2$) (six rank correlations). The procedure was replicated 1000 times. From these values the mean and quantiles (2.5%, 97.5%) were computed for each of the six rank correlations. The simulation results are shown in Tables 1 to 4.

Results and discussion

In all 16 simulated cases under model (1) (Tables 1 and 2) the mean rank correlation was high between S_i^2 and b_i as well as between W_i and S_{di}^2, while that between other pairs of stability measures was low. Considering the 95% confindence interval (given by the 2.5% and 97.5% quantiles), these results are well in accordance with empirical results. We investigated 25 data sets from

Table 1 Mean and quantiles (2.5; 97.5%) (in brackets) of rank correlations among stability measures S_i^2, W_i, b_i and S_{di}^2 based on 1000 runs of a Monte Carlo experiment for $\sigma_i^2 = 10$ and $\beta_i = 1$

σ_e^2	K	N	$S_i^2; W_i$	$S_i^2; b_i$	$S_i^2; S_{di}^2$	$W_i; b_i$	$W_i; S_{di}^2$	$b_i; S_{di}^2$
50	10	10	0.24	0.92	0.24	0.00	0.90	0.01
			(−0.36; 0.79)	(0.69; 1.00)	(−0.39; 0.81)	(−0.59; 0.59)	(0.64; 1.00)	(−0.62; 0.65)
	10	20	0.24	0.92	0.24	0.01	0.95	0.01
			(−0.42; 0.78)	(0.72; 1.00)	(−0.43; 0.80)	(−0.61; 0.65)	(0.77; 1.00)	(−0.61; 0.65)
	20	10	0.26	0.93	0.26	0.01	0.92	0.01
			(−0.15; 0.65)	(0.79; 0.99)	(−0.18; 0.67)	(−0.40; 0.41)	(0.76; 0.99)	(−0.42; 0.45)
	20	20	0.26	0.93	0.26	−0.01	0.96	−0.01
			(−0.16; 0.62)	(0.82; 0.98)	(−0.18; 0.64)	(−0.44; 0.41)	(0.85; 0.99)	(−0.46; 0.46)
200	10	10	0.13	0.97	0.13	0.01	0.91	0.00
			(−0.53; 0.67)	(0.85; 1.00)	(−0.57; 0.72)	(−0.62; 0.60)	(0.68; 1.00)	(−0.65; 0.64)
	10	20	0.12	0.97	0.12	0.00	0.95	−0.01
			(−0.57; 0.68)	(0.87; 1.00)	(−0.60; 0.72)	(−0.66; 0.64)	(0.79; 1.00)	(−0.67; 0.62)
	20	10	0.13	0.97	0.13	0.00	0.92	0.00
			(−0.32; 0.52)	(0.92; 1.00)	(−0.35; 0.56)	(−0.43; 0.41)	(0.76; 0.99)	(−0.47; 0.45)
	20	20	0.13	0.98	0.12	−0.01	0.96	−0.01
			(−0.34; 0.53)	(0.93; 1.00)	(−0.35; 0.55)	(−0.44; 0.40)	(0.87; 0.99)	(−0.49; 0.43)

Table 2 Mean and quantiles (2.5%; 97.5%) (in brackets) of rank correlations among stability measures S_i^2, W_i, b_i and S_{di}^2 based on 1000 runs of a Monte Carlo experiment for $\sigma_i^2 = 20i/(K+1)$ and $\beta_i = 1$

σ_e^2	K	N	S_i^2; W_i	S_i^2; b_i	S_i^2; S_{di}^2	W_i; b_i	W_i; S_{di}^2	b_i; S_{di}^2
50	10	10	0.37	0.87	0.38	0.06	0.95	0.06
			(−0.24; 0.84)	(0.55; 0.99)	(−0.31; 0.84)	(−0.55; 0.61)	(0.78; 1.00)	(−0.62; 0.69)
	10	20	0.49	0.82	0.49	0.10	0.98	0.10
			(−0.12; 0.89)	(0.43; 0.99)	(−0.12; 0.89)	(−0.56; 0.71)	(0.91; 1.00)	(−0.55; 0.71)
	20	10	0.40	0.87	0.41	0.03	0.97	0.04
			(−0.03; 0.76)	(0.62; 0.97)	(−0.06; 0.78)	(−0.39; 0.43)	(0.89; 1.00)	(−0.42; 0.45)
	20	20	0.51	0.81	0.52	0.05	0.99	0.05
			(0.11; 0.81)	(0.58; 0.95)	(0.13; 0.80)	(−0.39; 0.46)	(0.96; 1.00)	(−0.42; 0.47)
200	10	10	0.20	0.94	0.20	0.02	0.95	0.02
			(−0.44; 0.73)	(0.78; 1.00)	(−0.50; 0.76)	(−0.60; 0.60)	(0.81; 1.00)	(−0.63; 0.64)
	10	20	0.26	0.93	0.27	0.04	0.98	0.04
			(−0.38; 0.77)	(0.73; 1.00)	(−0.41; 0.78)	(−0.60; 0.65)	(0.92; 1.00)	(−0.61; 0.66)
	20	10	0.21	0.95	0.21	0.01	0.96	0.01
			(−0.22; 0.59)	(0.84; 0.99)	(−0.26; 0.62)	(−0.39; 0.41)	(0.88; 1.00)	(−0.44; 0.45)
	20	20	0.28	0.93	0.28	0.02	0.99	0.02
			(−0.15; 0.65)	(0.80; 0.98)	(−0.16; 0.67)	(−0.39; 0.48)	(0.95; 1.00)	(−0.41; 0.47)

German registration trials, which are described in Table 5. Rank correlations between S_i^2 and W_i and between W_i and S_{di}^2 are shown in Tables 6 and 7, respectively. The rank correlations in these data sets agree well with the simulation results. Hence, the simulation has demonstrated that model (1) conforms to real data. However, this does not neccessarily imply that this model is correct. All that can be said is that the simulations revealed no evidence that it is inappropriate.

For the cases of heterogeneous β_i and $\sigma_e^2 = 200$ (Tables 3 and 4) the mean rank correlations between S_i^2 and W_i became slightly negative and the 97.5% quantile is only barely larger than zero. W_i and S_{di}^2 show a comparatively low rank correlation with the highest 97.5% quantile only at 0.78 (The simulated rank correlations among other pairs of parameters are in agreement with real data). The same tendency, though less marked, was observed for $\sigma_e^2 = 50$. By contrast, empirical data usually show a moderate positive rank correlation of S_i^2 and W_i (Table 6) and a very close rank correlation between W_i and S_{di}^2 (Table 7). Thus, model (2) agrees less well with empirical data than model (1). This is corroborated by the finding that heterogeneity among regression coefficients usually explains only a small fraction of the genotype × environment interaction sum of squares (Wricke and Weber 1980) and is often nonsignificant. It is noted, however, that we have used relatively variable β_i-values (ranging from 0.5 to 1.33). With less variable β_i-values, the difference to the results with model (1) would have been less striking. We re-ran the simulation with β_i from 0.7 to 1.2 (data not shown). The mean rank correlations between S_i^2 and W_i were similar to the those in Tables 3 and 4, but the confidence bands were slightly broader. The rank correlations between W_i and S_{di}^2 were about 0.2 to 0.3 larger than the corresponding values in Tables 3 and 4. Although the mean rank correlations between W_i and S_{di}^2 were rather lower than expected from real data, the confidence band covered the value of 0.9 in most cases.

Table 3 Mean and quantiles (2.5%; 97.5%) (in brackets) of rank correlations among stability measures S_i^2, W_i, b_i and S_{di}^2 based on 1000 runs of a Monte Carlo experiment for $\sigma_i^2 = 10$ and $\beta_i = a + bi + ci^2$

σ_e^2	K	N	S_i^2; W_i	S_i^2; b_i	S_i^2; S_{di}^2	W_i; b_i	W_i; S_{di}^2	b_i; S_{di}^2
50	10	10	−0.03	0.97	0.09	−0.11	0.55	−0.02
			(−0.46; 0.44)	(0.85; 1.00)	(−0.61; 0.71)	(−0.56; 0.30)	(−0.13; 0.94)	(−0.70; 0.64)
	10	20	−0.11	0.98	0.05	−0.16	0.49	−0.02
			(−0.47; 0.26)	(0.92; 1.00)	(−0.61; 0.67)	(−0.55; 0.20)	(−0.12; 0.89)	(−0.64; 0.62)
	20	10	0.04	0.97	0.14	−0.07	0.66	0.00
			(−0.28; 0.46)	(0.89; 1.00)	(−0.33; 0.60)	(−0.38; 0.27)	(0.27; 0.92)	(−0.46; 0.47)
	20	20	−0.05	0.98	0.09	−0.12	0.61	−0.01
			(−0.32; 0.26)	(0.95; 1.00)	(−0.36; 0.51)	(−0.40; 0.16)	(0.19; 0.88)	(−0.45; 0.42)
200	10	10	−0.11	0.99	0.03	−0.12	0.25	0.00
			(−0.33; 0.12)	(0.96; 1.00)	(−0.59; 0.65)	(−0.35; 0.09)	(−0.42; 0.78)	(−0.63; 0.64)
	10	20	−0.11	1.00	0.01	−0.12	0.19	−0.02
			(−0.28; 0.01)	(0.98; 1.00)	(−0.62; 0.61)	(−0.30; 0.00)	(−0.49; 0.75)	(−0.64; 0.61)
	20	10	−0.09	1.00	0.03	−0.12	0.35	−0.01
			(−0.29; 0.10)	(0.98; 1.00)	(−0.44; 0.46)	(−0.33; 0.07)	(−0.10; 0.77)	(−0.47; 0.43)
	20	20	−0.11	1.00	0.02	−0.12	0.27	−0.01
			(−0.26; 0.04)	(0.99; 1.00)	(−0.45; 0.46)	(−0.27; 0.03)	(−0.18; 0.69)	(−0.47; 0.43)

Table 4 Mean and quantiles (2.5%; 97.5%) (in brackets) of rank correlations among stability measures S_i^2, W_i, b_i and S_{di}^2 based on 1000 runs of a Monte Carlo experiment for $\sigma_i^2 = 20i/(K+1)$ and $\beta_i = a + bi + ci^2$

σ_e^2	K	N	$S_i^2; W_i$	$S_i^2; b_i$	$S_i^2; S_{di}^2$	$W_i; b_i$	$W_i; S_{di}^2$	$b_i; S_{di}^2$
50	10	10	0.12	0.95	0.32	−0.03	0.61	0.15
			(−0.30; 0.65)	(0.79; 1.00)	(−0.11; 0.75)	(−0.40; 0.38)	(0.10; 0.95)	(−0.26; 0.62)
	10	20	0.06	0.96	0.29	−0.07	0.62	0.14
			(−0.30; 0.50)	(0.87; 1.00)	(−0.01; 0.66)	(−0.41; 0.26)	(0.22; 0.92)	(−0.19; 0.52)
	20	10	0.23	0.95	0.37	0.04	0.74	0.17
			(−0.10; 0.64)	(0.81; 0.99)	(0.06; 0.70)	(−0.22; 0.35)	(0.39; 0.95)	(−0.13; 0.51)
	20	20	0.19	0.96	0.33	0.00	0.77	0.13
			(−0.09; 0.51)	(0.88; 0.99)	(0.09; 0.60)	(−0.21; 0.24)	(0.52; 0.93)	(−0.08; 0.39))
200	10	10	−0.09	0.99	0.19	−0.12	0.25	0.14
			(−0.32; 0.15)	(0.95; 1.00)	(−0.20; 0.64)	(−0.36; 0.08)	(−0.25; 0.66)	(−0.24; 0.61)
	10	20	−0.11	0.99	0.17	−0.13	0.24	0.12
			(−0.31; 0.02)	(0.96; 1.00)	(−0.08; 0.50)	(−0.33; −0.01)	(−0.10; 0.52)	(−0.18; 0.47)
	20	10	−0.06	0.99	0.21	−0.10	0.38	0.15
			(−0.23; 0.19)	(0.97; 1.00)	(−0.03; 0.50)	(−0.27; 0.09)	(0.04; 0.73)	(−0.12; 0.45)
	20	20	−0.09	0.99	0.17	−0.12	0.37	0.12
			(−0.22; 0.06)	(0.98; 1.00)	(−0.01; 0.39)	(−0.26; 0.02)	(0.09; 0.61)	(−0.07; 0.34)

Table 5 Number of genotypes (K) and number of environments (N) in German registration trials (1985–1989) for faba beans, fodder beets, oats, sugar beets and oilseed rape

Year/Crop	1985	1986	1987	1988	1989
Faba beans	14; 9[a]	31; 9	32; 9	35; 10	35; 10
Fodder beets	19; 7	22; 6	21; 8	17; 8	20; 9
Oats	32; 12	33; 12	20; 12	14; 18	33; 12
Sugar beets	78; 11	73; 11	86; 9	71; 9	67; 11
Oliseed rape	32; 8	35; 11	35; 10	42; 9	41; 10

[a] K; N

Table 6 Rank correlations among W_i and S_i^2 in German registration trials (1985–1989) for faba beans, fodder beets, oats, sugar beets and oilseed rape

Year/Crop	1985	1986	1987	1988	1989
Faba beans	0.42	0.22	0.12	−0.24	0.01
Fodder beets	−0.02	−0.03	0.17	0.07	0.37
Oats	0.24	0.39*	0.18	0.24	0.08
Sugar beets	−0.23	0.06	0.34**	0.17	0.35
Oilseed rape	0.32	−0.02	−0.28	0.43**	0.35

*, ** Significantly different from zero at the 5%, 1% levels of probability, respectively

Table 7 Rank correlations[a] among W_i and S_{di}^2 in German registration trials (1985–1989) for faba beans, fodder beets, oats, sugar beets and oilseed rape

Year/Crop	1985	1986	1987	1988	1989
Faba beans	0.96	0.91	0.77	0.88	0.93
Fodder beets	0.83	0.87	0.75	0.99	0.82
Oats	0.94	0.99	0.87	0.99	0.84
Sugar beets	0.96	0.90	0.95	0.95	0.90
Oilseed rape	0.93	0.93	0.98	0.99	0.97

[a] All rank correlations are significantly different from zero at the 0.1% level of probability

It is interesting to observe that S_i^2 and b_i are highly rank correlated even when the expected values $E[S_i^2]$ and $E[b_i]$ do not differ among genotypes (see Table 1). The same is true of the rank correlation between S_{di}^2 and W_i. Conversely, the rank correlation between S_i^2 and W_i is low even when that between $E[S_i^2]$ and $E[W_i]$ equals one, as for the cases shown in Table 2 (Note that the rank order of $E[W_i]$ is the same as that of $E[Q_i]$). The results in Table 2 are based on model (1) with $\sigma_i^2 = D*i$. Thus, all σ_i^2 are distinct and the rank order of $E[W_i]$ and $E[S_i^2]$ is given by the rank order of σ_i^2. Clearly, despite a scaling factor, the difference between the W_i-values of two varieties has the same expectation as the difference between the S_i^2-values of the same two varieties. One may therefore say that in this case W_i (or Q_i) and S_i^2 are measures of essentially the same thing. Nevertheless, the empirical rank correlation of these measures is comparatively low, which is largely a result of the high sampling variances of the stability estimates.

It has been suggested to test for differences among S_i^2 by the usual tests for homogeneity of variances, e.g., the Bartlett test. The F-test has been proposed for comparisons among two genotypes. It should be pointed out that Bartlett's test, as well as the F-test, assumes that samples, i.e., observations y_{ij} for each genotype, are stochastically independent. When model (1) is correct, however, observations in the same environment are positively correlated. The correlation may be considerable when σ_e^2 is large compared to σ_i^2. As has been suggested before, differences among S_i^2-values essentially measure differences among σ_i^2-values. Therefore, it appears to be appropriate to test directly for differences among σ_i^2-values. Various tests are available for this purpose (Piepho 1993b and 1994).

S_{di}^2 and W_i were often reported to be poorly repeatable (Lin and Binns 1988, 1991; Pham and Kang 1988; Helms 1993; Jalaluddin and Harrison 1993), while b_i and S_i^2 were found to have comparatively good repeatability in some cases (Helms 1993; Lin and Binns 1991). Léon and Becker (1988) found comparable repeatabilities for all four of these measures with slight advantages for b_i and S_i^2. This suggests that S_i^2 is preferable to W_i for selection purposes.

S_i^2 and W_i are interchangable, provided that model (1) adequately describes the data. In this connection it may be worth reconsidering the classification of stability statistics. Lin et al. (1986) distinguish three types of statistics. They consider S_i^2 as a Type-1 statistic, while W_i is classified as a Type-2 statistic. Similarly, Becker and Léon (1988) regard S_i^2 as a static measure of stability, whereas W_i is characterized as a dynamic measure. The statistical considerations presented here suggest that despite their intuitive appeal these classifications may be misleading.

Acknowledgements The Bundessortenamt Hannover (Germany) is thanked for kindly providing the data used for the computations presented in Table 6 and Table 7.

References

Becker HC (1981) Biometrical empirical relations between different concepts of phenotypic stability. In: Gallais A (ed) Quantitative genetics and breeding methods. I.N.R.A., Versailles, pp 307–314

Becker HC, Léon J (1988) Stability analysis in plant breeding. Plant Breed 101:1–23

Helms TC (1993) Selection for yield and stability among oats lines. Crop Sci 33:423–426

Hühn MM (1990) Nonparametric measures of phenotypic stability. Part 2. Applications. Euphytica 47:195–201

Hühn M, Lotito S, Piepho HP (1993) Relationships between genotype-environment interactions and rank orders for a set of genotypes tested in different environments. Theor Appl Genet 86:943–950

Jalaluddin Md, Harrison SA (1993) Repeatability of stability estimators for grain yield in wheat. Crop Sci 33:720–725

Léon J (1985) Beiträge zur Erfassung der phänotypischen Stabilität unter besonderer Berücksichtigung unterschiedlicher Heterogenitäts- und Heterozygotiegrade sowie einer zusammenfassenden Beurteilung von Ertragshöhe und Ertragssicherheit. Dissertation, Kiel University

Léon J, Becker HC (1988) Repeatability of some statistical measures of phenotypic stability. Plant Breed 100:137–142

Lin CS, Binns MR, Levkovitch LP (1986) Stability analysis: where do we stand? Crop Sci 26:894–900

Lin CS, Binns MR (1988) A method of analysing cultivar × location × year experiments: a new stability parameter. Theor Appl Genet 76:425–430

Lin CS, Binns MR (1991) Genetic properties of four types of stability parameter. Theor Appl Genet 82:505–509

Pham HN, Kang MS (1988) Interrelationships and repeatability of several stability statistics estimated from international maize trials. Crop Sci 28:381–385

Piepho HP (1993a) Note on bias in estimates of the regression coefficient in the analysis of genotype-environmental interaction. Heredity 70:98–100

Piepho HP (1993b) Testing for homoscedasticity. Letter. Biometrics 49:1279–1280

Piepho HP (1994) Tests for pairwise commparison of variances in a two-way layout for the analysis of phenotypic stability. Informatik, Biometrie und Epidemiologie in Medizin and Biologie, 25:172–180

Piepho HP, Lotito S (1992) Rank correlation among parametric and nonparametric measures of phenotypic stability. Euphytica 64:221–225

SAS Institute (1989) SAS/IML software: usage and reference. Version 6. Cary, North Carolina

Shukla GK (1972) Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29: 237–245

Weber WE, Wricke G (1987) Wie zuverlässig sind Schätzungen von Parametern der phänotypischen Stabilität? Vorträge Pflanzenzücht 12:120–133

Wricke G (1962) Über eine Methode zur Erfassung der ökologischen Streubreite. Z Pflanzenzücht 47:92–96

Wricke G, Weber WE (1980) Erweiterte Analyse von Wechselwirkungen in Versuchsserien. In: Köpcke W, Überla K (eds) Biometrie – heute und morgen. Springer-Verlag, Berlin, pp 87–95