



TWC/23/13

ORIGINAL: English

DATE: June 6, 2005

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS
GENEVA

**TECHNICAL WORKING PARTY
ON
AUTOMATION AND COMPUTER PROGRAMS**

**Twenty-Third Session
Ottawa, June 13 to 16, 2005**

EXAMINATION OF STATISTICAL PROCEDURES FOR
CHECKING UNIFORMITY IN VARIETY TRIALS

Document prepared by experts from Germany

EXAMINATION OF STATISTICAL PROCEDURES FOR CHECKING UNIFORMITY IN VARIETY TRIALS

Andreas Büchse ¹⁾, Hans-Peter Piepho ¹⁾ and Uwe Meyer ²⁾

¹⁾Institute for Crop Production and Grassland Research, University of Hohenheim, Fruwirthstr.
23, D-70599 Stuttgart, Germany

²⁾Bundessortenamt, Osterfelddamm 80, D-30627 Hannover, Germany

Abstract

To assess uniformity for quantitative characteristics in trials for distinctness, uniformity and stability (DUS), the International Union for the Protection of New Varieties of Plants (UPOV) has identified a method called Combined-Over-Years-Uniformity (COYU). To account for a dependency between standard deviation and mean, a moving average procedure is used. Standard deviations are transformed and adjusted by the moving averages and new varieties are compared to the mean transformed standard deviations of reference varieties by a one-tailed t-test.

We examine the current moving-average procedure and the test against the mean of transformed standard deviations and compare it to an alternative method, using analysis of covariance and the estimation of quantiles. This is exemplified with a real data set from rapeseed uniformity trials in Germany.

Uniformity – DUS – COYU – moving average – LOESS - national list – standard deviation – analysis of covariance - quantiles - quantitative characteristic

1. Introduction

For the granting of a plant breeder's right, uniformity is one of three important criteria to be met by new varieties. To assess uniformity for quantitative characteristics, the International Union for the Protection of New Varieties of Plants (UPOV) has proposed a method called Combined-Over-Years-Uniformity (COYU) (UPOV 2004). This method assesses uniformity based on within-plot standard deviations (SD) per variety and year using a sample of plants per plot. The SDs are subjected to analysis of variance (ANOVA) with factors year and variety nested in year. Candidate varieties are compared to the mean of a reference set by a one-sided t-test. The uniformity requirement is met when the one-sided t-test does not reject the null hypothesis of uniformity. To better meet the usual assumptions, SDs are log-transformed. In addition it has been found that SDs are often related to the mean, even after log-transformation. To remove this dependency, data are adjusted based on a moving average procedure employing reference varieties.

The purpose of the present paper is to examine the moving-average procedure and to compare it to alternative procedures accounting for an SD-mean relationship. Specifically, we will explore analysis of covariance (ANCOVA) techniques, which allow the use of all reference varieties for adjustment. Finally, we investigate whether testing against the mean of reference varieties is appropriate. We will exemplify this with real data from rapeseed uniformity trials in Germany.

2. Materials and Methods

2.1 Combined Over Years Uniformity according to UPOV (2004)

First the within plot standard deviations for each variety in each year are calculated by averaging the between-plant standard deviations on a plot-basis (SD_k) over replications:

$$SD_k = \sqrt{\frac{\sum_{h=1}^n (z_{hk} - \bar{z}_{\bullet k})^2}{(n-1)}} \quad , \text{ and} \quad (1)$$

$$SD = \frac{\sum_{k=1}^r SD_k}{r} \quad , \quad (2)$$

where z_{hk} is the observation on the h^{th} plant in the k^{th} plot, $\bar{z}_{\bullet k}$ is the mean of the observations from the k^{th} plot, n is the number of plants measured in each plot and r is the number of replicates (plots per variety). The mean standard deviations SD are then log-transformed by $\log(SD+1)$.

To account for a relation between standard deviation and mean, the COYU procedure adjusts log-transformed SD -values for trends as determined from the reference varieties. The adjusted values are then submitted to an ANOVA with factors year and variety nested in year. The adjustment is performed separately for each year and works as follows:

1. Rank the log SD s of reference varieties according to the values of the mean (M), with $M = \bar{z}_{\bullet\bullet}$.
2. For each log SD of a reference variety, compute a nine-point moving average from the ordered log SD s of all reference varieties and denote this as trend value (T). For values ranked 1st and 2nd the trend value is taken to be the mean of the first three values. If the value is ranked 3rd, the mean of the first five values is taken and for a value ranked 4th the mean of the first seven values is used.

3. To calculate the trend value for a candidate variety find the two reference varieties with the next smallest and next largest mean. The trend value for the candidate ($T_{Candidate}$) is computed by linear interpolation between the two reference varieties as

$$T_{Candidate} = \frac{(M_{Candidate} - M_{left})T_{right} + (M_{right} - M_{Candidate})T_{left}}{M_{Right} - M_{Left}}, \quad (3)$$

where M_{left} and M_{right} are the mean values of the reference variety with next smallest and next largest mean, respectively, T_{left} and T_{right} are the corresponding trend values, and $M_{Candidate}$ is the mean value of the candidate.

4. To adjust the transformed SDs for candidate and reference varieties, the estimated trend value ($T_{Candidate}$ or T) is subtracted and the grand year mean of reference varieties is added back. The objective of the adjustment is to obtain SD values, which would have been observed if all varieties had the same mean.

5. The variance (V) of adjusted SDs of reference varieties is computed as the residual mean square of a one-way ANOVA with years as classification factor, using the model

$$y_{ij} = \mu + u_j + e_{ij} \quad (4)$$

where

- y_{ij} = SD -value (log-transformed) of i -th variety in j -th year
- μ = constant (grand mean of reference varieties)
- u_j = effect of j -th year
- e_{ij} = residual corresponding to a single variety in a single year

6. The maximum allowable mean adjusted SD of a candidate variety is computed as

$$UC_{\alpha} = \bar{y}_{..} + t_{\alpha} \sqrt{V \left(\frac{1}{k} + \frac{1}{Rk} \right)}, \quad (5)$$

where $\bar{y}_{..}$ is the mean of adjusted log SDs for the reference varieties, t_{α} is the one-tailed t -value for significance probability α with degrees of freedom as for V , k is the number of years and R is the number of reference varieties. The COYU procedure recommends $\alpha = 0.002$ (UPOV, 2004).

2.2 An alternative method

Instead of the moving average procedure we suggest to perform the adjustment for the SD-mean dependence by analysis of covariance (ANCOVA). Inspection of the rapeseed data (Fig. 1) confirms that trends sometimes differ slightly among years. Thus, the model should allow for year-specific trends. The suggested model

$$y_{ij} = \mu_i + u_j + f_j(M_{ij}) + e_{ij} , \quad (6)$$

where

- y_{ij} = SD-value (possibly transformed) of i -th variety in j -th year
- μ_i = effect of i -th variety (reference or candidate)
- u_j = effect of j -th year
- $f_j(\cdot)$ = smooth function for the j -th year (e.g., a quadratic polynomial)
- e_{ij} = residual corresponding to y_{ij}
- M_{ij} = mean corresponding to y_{ij}

An adjusted mean can be defined as the mean of

$$\eta_{ij} = \mu_i + u_j + f_j(\bar{M}_{\bullet j}) \quad (7)$$

over years, where $\bar{M}_{\bullet j}$ is the year mean of means M_{ij} .

Different functions may be fitted for $f_j(M)$, e.g., a quadratic polynomial:

$$f_j(M) = \beta_j M + \gamma_j M^2 .$$

The model assumes a different regression on the mean for each year. An F-test may reveal whether there is significant heterogeneity among years. When there are no significant differences, one may fit a common regression model, e.g.,

$$f_j(M) = \beta M + \gamma M^2 .$$

If the quadratic term is not significant, a linear trend can be fitted.

For candidate varieties we computed least square means ($\bar{\eta}_{i\bullet}$) following (7) and confidence intervals. For approximation of denominator degrees of freedom the Kenward-Roger method was used (KENWARD & ROGER 1997). Candidate variety means can be compared to a specific location parameter of the distribution of reference varieties (e.g. the mean, the median

or a quantile). We calculated a confidence interval for the 95-quantile of reference variety means by using the algorithm of HAHN & MEEKER (1991) which is implemented in the UNIVARIATE procedure of the SAS system. A two-sided $100(1-\alpha)\%$ confidence interval for Q_p , the $100p$ th quantile of a normal distribution, based on a random normal sample x_i ($i = 1, \dots, n$) is

$$[Q_{lcl}, Q_{ucl}] = [\bar{x} - g'_{(1-\alpha/2; 1-p, n)} s, \bar{x} - g'_{(\alpha/2; 1-p, n)} s] \text{ for } 0.00 < p < 0.50 \text{ and}$$

$$[Q_{lcl}, Q_{ucl}] = [\bar{x} + g'_{(\alpha/2; 1-p, n)} s, \bar{x} + g'_{(1-\alpha/2; 1-p, n)} s] \text{ for } 0.50 \leq p < 1.00$$

where \bar{x} and s are the sample mean and standard deviation of x_i , n is the sample size, and the factors $g'_{(y; p, n)}$ are given in Table 1 of ODEH & OWEN (1980) or Table A.12 in HAHN & MEEKER (1991). The method is based on the noncentral t -distribution as given by ODEH & OWEN (1980). Here, we use the least square means ($\bar{\eta}_{i\cdot}$) in place of x_i . The confidence interval for the quantile is approximate in that least square means may show some correlation.

2.3 Take reference varieties as random

Instead of treating all varietal effects as fixed we can regard candidate varieties as fixed and reference varieties as a random sample of a hypothetical population of reference varieties. This is useful for studying the distribution of reference varieties. The model may be extended as follows:

$$y_{kmj} = \phi_k + x_k a_{km} + u_j + f_j(M_{kmj}) + e_{kmj}$$

where

y_{kmj} = SD-value (possibly transformed) of m -th genotype effect in k -th set of varieties in j -th year; each candidate forms a separate set, while all reference varieties are pooled into a single set

ϕ_k = k -th genotype effect; a separate effect is fitted for each candidate variety, while a common effect is fitted for the reference set.

x_k = dummy variable. $x_k = 1$ for reference set, $x_k = 0$ for candidates

a_{km} = random genetic effect for m -th reference variety ($a_{km} = 0$ for candidate varieties)

u_j = effect of j -th year

M_{kmj} = mean of km -th genotype in j -th year

e_{kmj} = residual associated with y_{kmj}

Based on estimates of $\text{var}(a_{km})$ it would be possible to estimate quantiles of the distribution of reference varieties (see section 4 Discussion for further details).

2.4 Comparison of methods with a rapeseed data set

To study the behaviour of the moving average procedure and to demonstrate the practicability of the ANCOVA method we analysed a three year data set from the official German rapeseed testing program. The trials were conducted in the years 2001-2003 as randomised complete block designs with three replicates. Following quantitative characteristics were observed on 60 plants (3*20 plants):

M11: Cotyledon: length
M12: Cotyledon: width
M14: Leaf: length
M15: Leaf: width
M19: Leaf: number of lobes
M110: Leaf: length of petiole
M21: Plant: total length including side branches
M212: Siliqua: length
M213: Siliqua: length of beak
M214: Siliqua: width
M216: Siliqua: length of peduncle
M31: Time of flowering
M32: Flower: length of petals
M33: Flower: width of petals

2.5 LOESS-Smoother

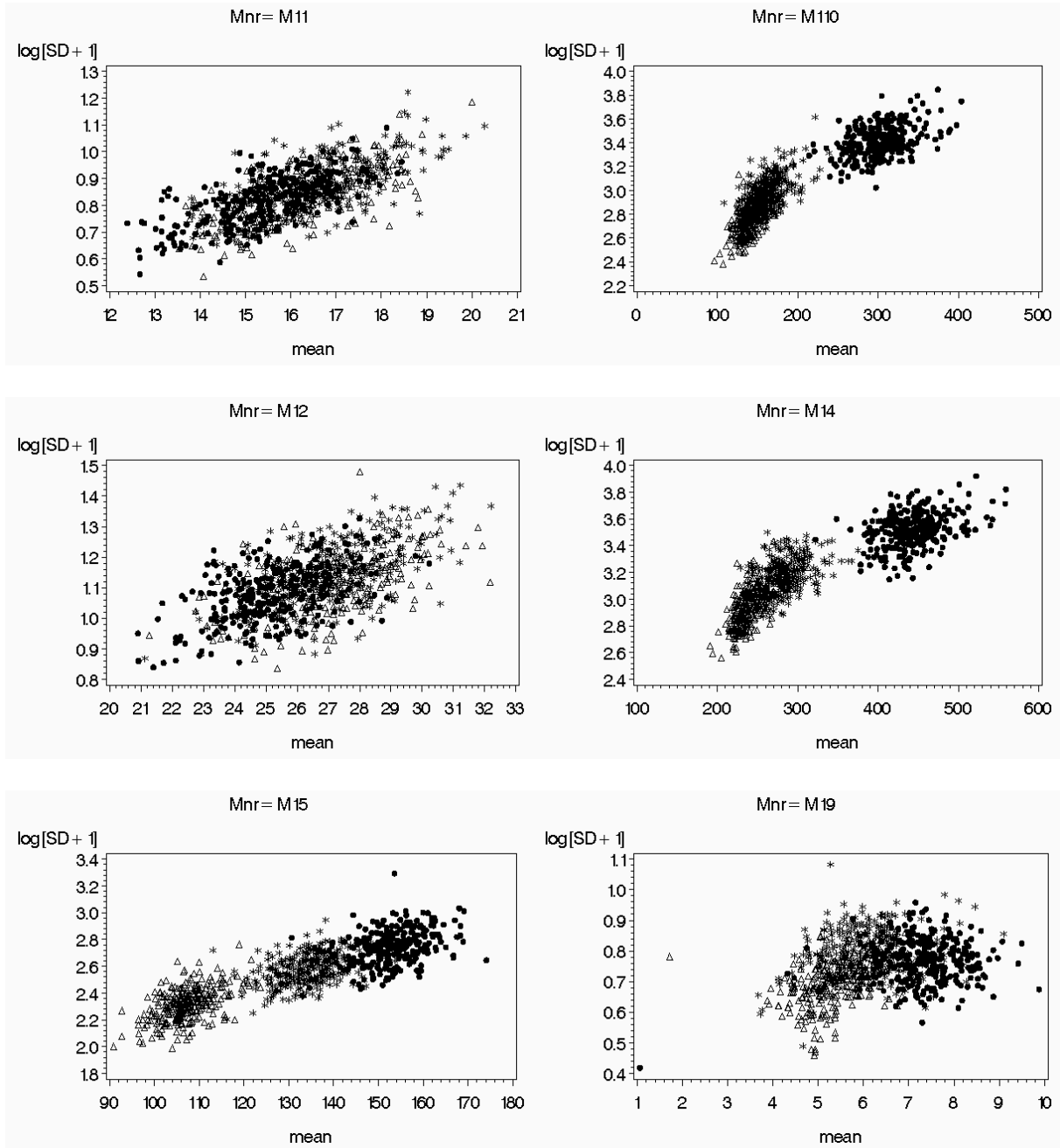
To further investigate the relation between SD and mean a non-parametric LOESS-smoother was fitted to log-transformed data (CLEVELAND 1979). To find the optimal smoothing it is necessary to select a span for the smoother by minimizing a suitable criterion. Here the AICC1-criterion was used (DURBAN et al. 2003, HURVICH et al. 1998). An algorithm to find the optimal parameter is implemented in the LOESS procedure of the SAS System by the macro “%smoothselect”. Setting the smoothing parameter to $s=1$ means that all observations are used to fit a polynomial trend and estimate predicted values for each observation. With $s=0.5$ only 50% of data are used and with $s=0.1$ only 10% respectively. With a total of more than 270 observations using a moving average of only nine observations equals a local fitting with a linear regression and $s = 0.033$.

3. Results

3.1 Investigation of SD-mean relation

Fig. 1 shows plots of log SD against the mean for 14 different characteristics of rapeseed. Fig. 2 and 3 show observed data, linear trend, moving average trend and LOESS-smoother for one separate year and characteristic M11.

In 31 of 42 cases smoothing the rapeseed data by using information from all observations and fitting a smooth polynomial trend with few parameters was optimal according to the AICC1-criterion (Tab. 1). At least 55 observations were used to estimate the local LOESS-smoother. For most characteristics a linear trend, allowing for year specific deviations sufficed to describe the SD-mean relation (Tab. 2). For most characteristics the reference varieties also differed in homogeneity (Tab. 3).



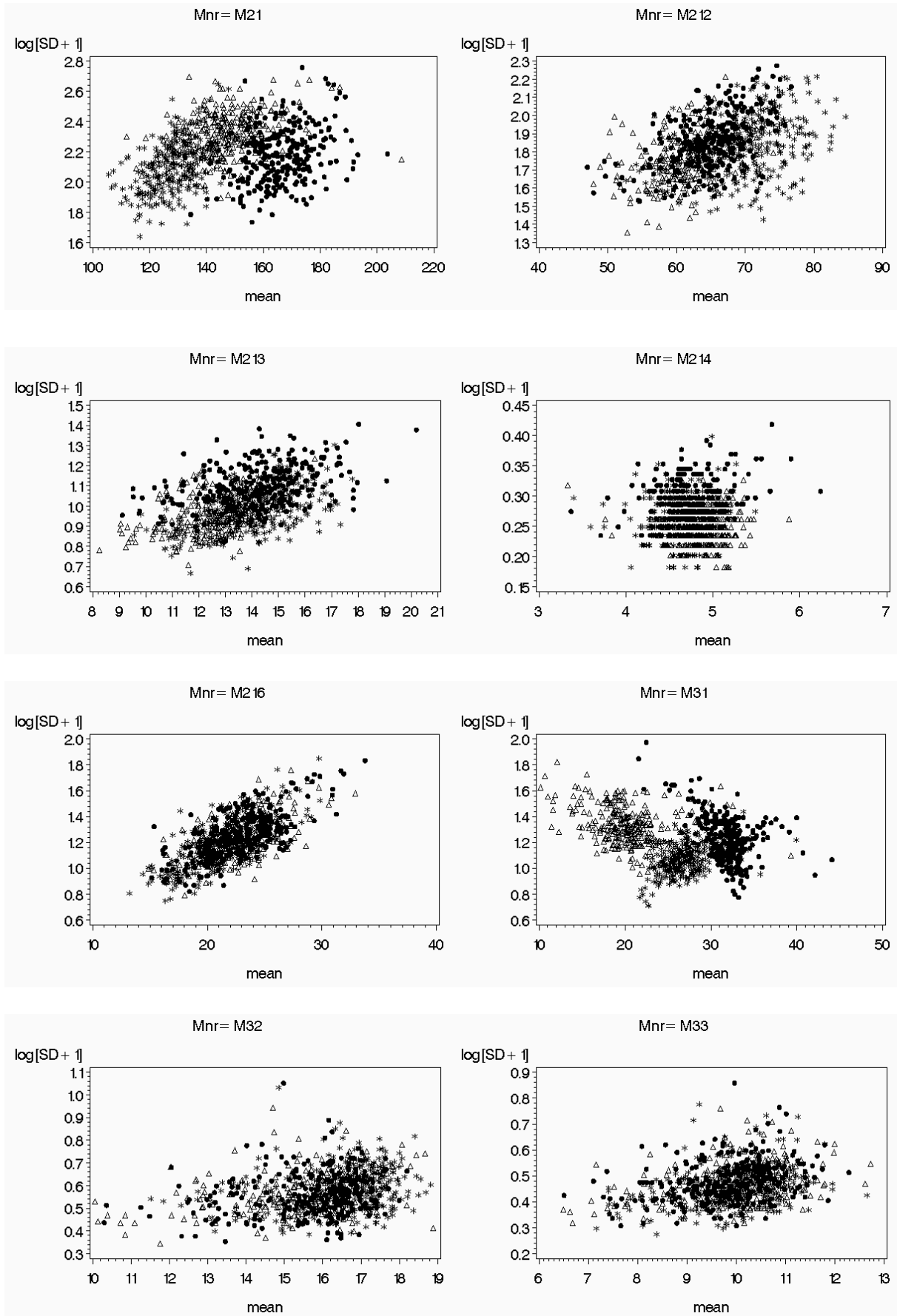


Fig. 1: Log SD plotted against the mean of 14 different characteristics. Dots = first year, triangles = second year, stars = third year.

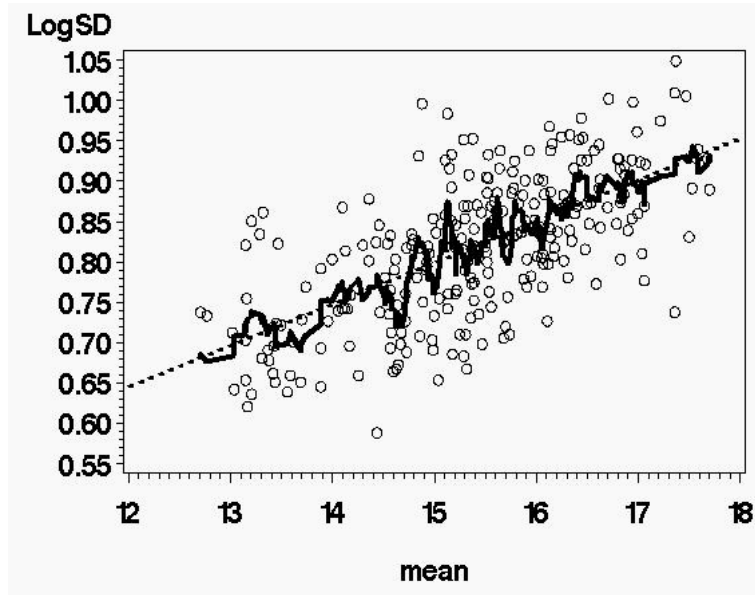


Fig. 2: Characteristic 'M11', year 1. Solid line = nine-point moving averages, circles = observed data. Line fitted to observed data (dashed) is a linear regression.

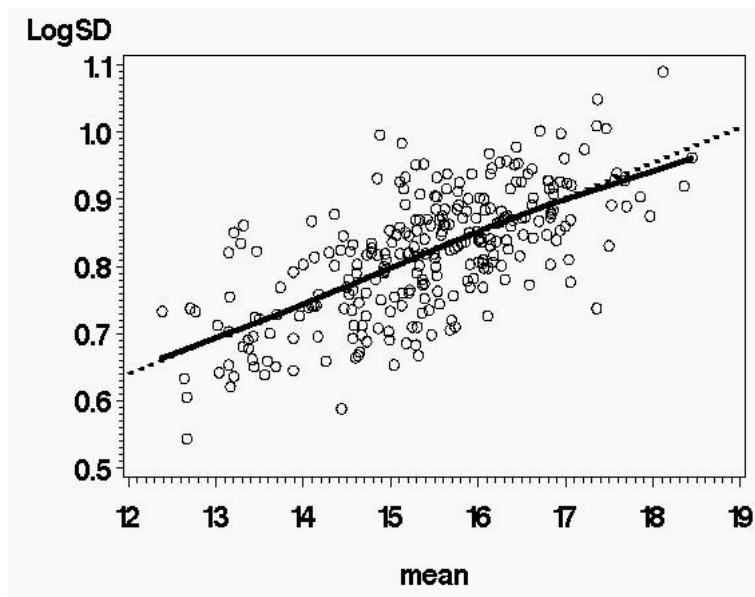


Fig. 3: Characteristic 'M11', year 1. Solid line = optimal LOESS-smoother, circles = observed data. Line fitted to observed data (dashed) is a linear regression.

Table 1: Optimal smoothing parameter values (s) in LOESS-procedure for 14 different characteristics of rapeseed data. $s=1$ is equivalent to using all observations to fit a polynomial trend. Optimality criterion $AICC_1$ (HURVICH et al. 1998).

Characteristic	year 1		year 2		year 3	
	s	points in local neighborhood	s	points in local neighborhood	s	points in local neighborhood
M11	1	274	1	276	1	276
M110	1	275	0.5	138	1	276
M12	1	274	1	276	1	276
M14	1	275	1	276	1	276
M15	1	275	1	276	1	276
M19	0.3	81	0.3	81	1	272
M21	0.4	110	1	271	0.6	165
M212	1	275	1	271	1	275
M213	1	275	1	271	0.5	137
M214	1	275	1	271	1	275
M216	0.7	192	1	271	1	275
M31	0.3	82	0.3	82	0.4	110
M32	1	276	1	276	0.9	246
M33	1	276	1	276	1	274

Table 2: F-values for SD–mean relation. Quadratic polynomial for $f_j()$. Model fitted to $\log(SD+1)$. Kenward-Roger method was used to compute standard errors and degrees of freedom under random effects model (reference varieties random, candidates fixed).

Characteristic	Polynomial model terms			
	Linear	Linear*Year	Quadratic	Quadratic*Year
M11	434.34 ***	4.74 **	0.23 ns	3.07 *
M110	196.73 ***	51.46 ***	2.08 ns	0.01 ns
M12	265.20 ***	0.80 ns	0.15 ns	0.56 ns
M14	158.04 ***	48.77 ***	0.01 ns	1.30 ns
M15	178.20 ***	11.61 ***	0.47 ns	1.50 ns
M19	38.28 ***	27.93 ***	43.83 ***	6.17 **
M21	48.34 ***	8.32 ***	12.66 ***	1.03 ns
M212	64.68 ***	6.43 **	0.46 ns	1.02 ns
M213	134.00 ***	5.82 **	0.00 ns	1.82 ns
M214	5.90 *	5.57 **	5.83 *	0.61 ns
M216	297.89 ***	3.53 *	0.32 ns	0.66 ns
M31	10.54 **	125.85 ***	25.71 ***	36.56 ***
M32	32.40 ***	0.44 ns	0.19 ns	1.73 ns
M33	59.15 ***	0.24 ns	0.39 ns	0.05 ns

*, **, *** = significant F-test according to $\alpha = 0.05, 0.01, 0.001$ respectively; ns = not significant

Table 3: *F*-tests for genetic effects of reference varieties (a_{km}), variance components for a_{km} and e_{kmj} . Quadratic polynomial for $f_j()$. Model fitted to $\log(SD+1)$. Kenward-Roger method used to compute standard errors and degrees of freedom.

Characteristic	F-test for a_{km}		$\text{var}(a_{km})$	$\text{var}(e_{kmj})$	§s.e.d.	
	F-value	p-value			fixed	random
M11	2.10	<0.0001	0.00470	0.01362	0.0671	0.0677
M12	1.97	<0.0001	0.00403	0.01252	0.0648	0.0649
M14	1.06	0.2665	0.00022	0.01747	0.0786	0.0773
M15	0.95	0.6850	0.00000	0.01746	0.0799	0.0778
M19	1.31	0.0029	0.00135	0.01903	0.0791	0.0799
M110	1.23	0.0174	0.00102	0.01714	0.0759	0.0760
M21	3.30	<0.0001	0.01337	0.01830	0.0773	0.0787
M212	3.11	<0.0001	0.01253	0.01841	0.0785	0.0790
M213	1.78	<0.0001	0.00483	0.01867	0.0792	0.0792
M214	1.97	<0.0001	0.00517	0.01583	0.0729	0.0730
M216	3.11	<0.0001	0.01401	0.02193	0.0849	0.0863
M31	2.86	<0.0001	0.01373	0.02307	0.0883	0.0884
M32	2.22	<0.0001	0.01305	0.03192	0.1045	0.1040
M33	2.98	<0.0001	0.01550	0.02332	0.0887	0.0888

§ Standard error of a difference between mean of reference varieties and a candidate variety with three years. Depending on purpose, reference varieties assumed fixed (F-test) or random [estimation of $\text{var}(a_{km})$].

3.2 Analysis of characteristic “M11” (cotyledon length)

We fitted model (6) to characteristic “M11” of the rapeseed data using all reference varieties and candidate varieties with two or three years of testing. Variety means of adjusted $\log SD$ of reference varieties followed a normal distribution (Fig. 4). The adjusted SD means of candidate varieties showed a similar distribution (Fig. 5).

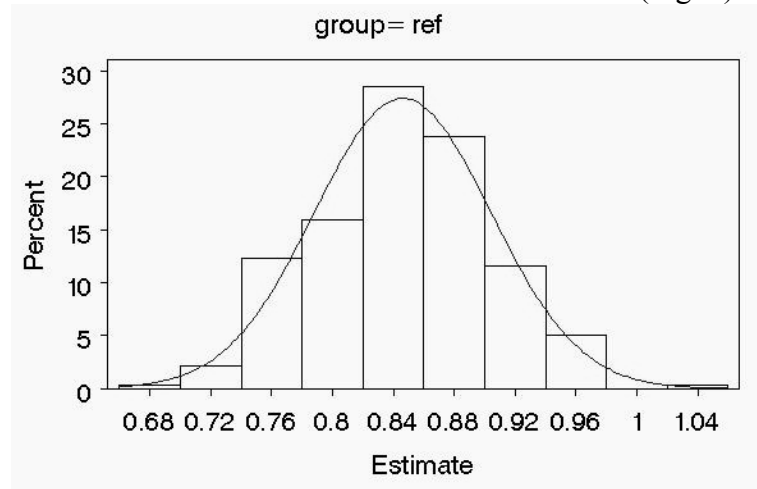


Fig. 4: Histogram for least square means of $\log(SD+1)$ of reference varieties; Characteristic 'M11' - cotyledon length, solid line = expected density under normality

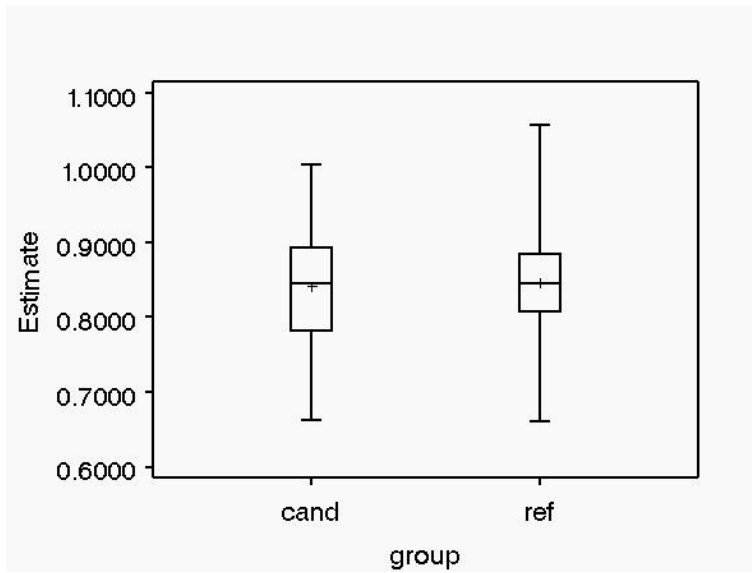


Fig. 5: Boxplot of least square means of $\log(SD+1)$ of reference and candidate varieties; Characteristic 'M11'. cotyledon length

Table 4: Mean of reference varieties, mean of candidates, 95%-quantile of reference varieties, COYU-uniformity-criterion and fraction of rejected candidates according to different criteria

Characteristic	reference varieties						candidates		candidates rejected [%] according to ...		
	Mean	Std Dev	95% quantile	95%-confidence limits for 95%-quantile		UC ***	Mean	Std Dev	UC ***	crit 1 *	crit 2 **
M11	0.846	0.058	0.943	0.932	0.953	0.958	0.841	0.073	6.3	9.1	0.0
M110	2.904	0.082	3.040	3.025	3.055	3.244	2.829	0.091	0.0	2.1	0.0
M12	1.121	0.061	1.221	1.211	1.234	1.245	1.128	0.072	4.9	7.7	0.7
M15	2.524	0.069	2.642	2.626	2.652	2.722	2.461	0.091	0.0	1.4	0.0
M19	0.755	0.054	0.827	0.834	0.853	0.866	0.765	0.074	6.3	18.2	0.7
M21	2.150	0.147	2.378	2.368	2.420	2.442	2.129	0.138	3.5	3.5	0.0
M212	1.818	0.130	2.048	2.009	2.056	2.043	1.779	0.133	2.8	2.8	0.0
M213	1.016	0.070	1.137	1.119	1.144	1.163	0.966	0.079	0.7	2.1	0.0
M214	0.266	0.025	0.312	0.302	0.311	0.317	0.265	0.031	5.6	8.4	0.7
M216	1.227	0.132	1.464	1.422	1.470	1.427	1.190	0.134	4.2	0.7	0.7
M31	1.337	0.124	1.522	1.519	1.564	1.403	1.381	0.128	9.1	15.4	0.7
M32	0.571	0.070	0.684	0.675	0.701	0.716	0.567	0.085	4.9	5.6	2.1
M33	0.470	0.060	0.574	0.559	0.581	0.588	0.474	0.070	5.6	6.3	2.8
mean									4.1	6.4	0.6

* criterion 1 = lsmean of candidate > 95% quantile of reference varieties

** criterion 2 = upper 95% confidence limit of lsmean of candidate > 95% quantile of reference varieties

*** COYU-criterion according to UPOV (2004), see eq. 5

4. Discussion

4.1 Critical comments on COYU

Parametric smoother versus moving averages (COYU-procedure step 3): In most cases in the rapeseed data there is an SD-mean trend, but it is a rather smooth one (Fig. 1). Thus, one might consider fitting a smooth trend. There are many options for this, e.g., linear or polynomial regression or LOESS (CLEVELAND 1979). Searching for the best smoothing parameter shows, that a LOESS-smother should use information from all observations (Tab. 1). A linear or quadratic regression will usually suffice. Because of the high portion of random error, moving average procedures do not seem appropriate. The uneven and jagged shape of the moving average raises the question whether an adjustment based on moving averages produces overfitting and makes the adjusted data more noisy than the original data (Fig. 2). Rather than just using a subset of the reference varieties, as is the case with the current moving average procedure, it may be better to use all reference varieties to fit the trend, as with the procedure proposed here. In the present paper we have used all varieties, including the candidate varieties, for computing the trend. This was done in order to have as much data as possible and because there did not seem to be critical outliers. In routine application, one would want to exclude candidates from trend computations. This may be done in a straightforward fashion by extending model (7) using dummy coding.

Adjustment by subtracting moving average (COYU-procedure step 4): The adjustment introduces two types of correlations. The subtraction of moving averages causes a correlation between reference varieties with similar character means and additionally the addition of the year grand mean adds a correlation within the same year. The ANOVA in steps 5 and 6 of the COYU-procedure ignores these correlations. It is not clear how robust the procedure is to these correlations, which constitute a violation of the usual ANOVA assumptions.

Because of the moving average procedure the resulting adjusted *SDs* also have a lower variance than the original data. This underestimation of variance diminishes the magnitude of the uniformity criterion (step 5) and therefore influences the probability of rejecting a candidate.

Assumption of homogeneous adjusted means: Model (4), which is used in the COYU procedure to estimate the variance V , has no variety main effect. This implies the strong assumption that all reference varieties have the same mean. The results in Table 3 show that this assumption is not generally tenable.

Validity of uniformity criterion (COYU-procedure step 5 and 6): The calculation of the uniformity criterion in eq. (5) yields a confidence limit for the difference between the mean of transformed standard deviations of the reference varieties and a single candidate variety,

$$UC_{\alpha} = \bar{y}_{..} + t_{\alpha} \sqrt{V \left(\frac{1}{k} + \frac{1}{Rk} \right)}.$$

If a mean adjusted *SD* of a candidate variety is higher than the criterion UC_{α} , the variety is rejected. The criterion implies a test of the null hypothesis that the candidate variety has a mean log *SD* below the mean of the reference set. This is a rather stringent criterion, which is expected to notably increase uniformity of new varieties in the long run. Table 3 shows that uniformity differs significantly between reference varieties in most characteristics. Thus, if the mean of the reference varieties were used as a threshold for uniformity of the reference varieties themselves, then 50% of the reference varieties would not be accepted because of a lack of uniformity! A further critical point is the dependence of UC_{α} on the power of the test: the larger the number of reference varieties, the better the power of the test, and the more candidate varieties will be rejected. We think that a criterion of uniformity should be defined so that it is independent of the power of a test and of the number of reference varieties.

4.2 Proposal for a different criterion

We would suggest to use a higher value than the mean of transformed standard deviations of reference varieties as threshold. The COYU-model assumes that there is a common *SD*-mean

of all reference varieties, but Table 3 shows that reference varieties differ significantly in their mean for most characteristics. Inspection of *SD*-means of single reference varieties shows that they are more or less normally distributed around the mean. Consequently following this idea it would be appropriate to take reference varieties as random, estimate their mean and variance and calculate a critical value (e.g. the 95-quantile). Also it is possible to regard reference varieties as fixed and calculate least square means. Independent from the question of whether reference varieties are random or fixed in a linear model analysis, one can calculate quantiles of the distribution of reference variety means of transformed standard deviations. The quantile-concept allows the definition of criterion for the selection of uniform varieties. We would propose that candidates should be at least as homogenous as the p 100%-quantile of reference varieties, with a suitable choice of the fraction p . According to the COYU-procedure of early rejection or acceptance (UPOV, 2004), a candidate variety may be accepted/rejected after a two or three years of testing if its mean is significantly (e.g. with 5% type I error) lower/higher than the estimated critical quantile.

In the rapeseed example 6,4% of candidates*characteristic combinations exceeded 95%-quantile and 4,1% of combinations exceeded the original COYU-criterion (Table 4). Both thresholds probably lead to a high fraction of rejected candidates, because non-uniformity in one characteristic is enough to give a reason for rejection. In this context we should pay attention to the problem of multiple testing. If many are measured, the probability of a significant non-uniformity in a single characteristic will increase. It may therefore be appropriate to work with a somewhat higher value for the quantile fraction p , e.g., $p = 0.975$ or $p = 0.99$.

References

- CLEVELAND, W.S. (1979): Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74, 829-836
- DURBAN, M., HACKETT, C.A., McNICOL, J.W., NEWTON, A.C., THOMAS, W.T.B. & CURRIE, I.D. (2003): The practical use of semiparametric models in field trials. *Journal of Agricultural, Biological and Environmental Statistics* 1, (8), 48-66
- HAHN, G.J. & MEEKER, W. Q. (1991): *Statistical intervals: A guide for practitioners*, New York: John Wiley & Sons, Inc.
- HURVICH, C.M., SIMONOFF, J.S. & TSAI, C.L. (1998): Smoothing parameter selection in nonparametric regression using an improved Akaike Information Criterion. *Journal of the Royal Statistical Society, Series B*, 60, 271-293.
- KENWARD, M.G. & ROGER, J.H. (1997): Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53, 983-997.
- ODEH, R.E. & OWEN, D.B. (1980): *Tables for normal tolerance limits, sampling plans and screening*. Marcel Dekker.
- UPOV (2004): *Examining Uniformity*. Document TGP/10.3.1, Statistical methods, COYU. International Union for the Protection of New Varieties of Plants.

[End of document]