DE GRUYTER OPEN

DOI: 10.1515/bile-2016-0003

Biometrical Letters
Vol. 53 (2016), No. 1, 59-67

The dependence of DUS (distinctness, uniformity and stability) decisions concerning white mustard and oilseed rape varieties on the number of measurements

Wiesław Pilarczyk[1], Bogna Kowalczyk[2], Bogna Zawieja[1]

[1]Department of Mathematical and Statistical Methods, Poznan University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland,
e-mail: wpilar@up.poznan.pl, bogna13@up.poznan.pl
[2]The Research Centre for Cultivar Testing at Słupia Wielka, Słupia Wielka 63-022, Poland,
e-mail: b.kowalczyk@coboru.pl

SUMMARY

An analysis is made of the possibility of reducing numbers of measurements in DUS (distinctness, uniformity and stability) trials on varieties of white mustard and oilseed rape, based on the results of DUS trials performed in the years 2005–2009 by the Research Centre for Cultivar Testing. It is shown that, when decisions are taken after three years of testing, it is possible to decrease the number of measurements by 50%, with the decisions practically unchanged for both considered species. For so-called early decisions (after two years of testing) the distinctness decisions are again practically the same for full and reduced data, although the decisions concerning uniformity differ significantly. This means that the number of measurements can be reduced only in cases where DUS testing is continued for three years.

Key words: distinctness and uniformity testing, number of measurements, oilseed rape, white mustard

1. Introduction

A new variety of any cultivated species (often called a candidate variety), before being listed in, for example, the National List of Varieties or the Common Catalogue, is subject to checking for fulfillment of so-called DUS requirements. The acronym DUS denotes distinctness (D), uniformity (U) and stability (S). The candidate variety must be significantly distinct from any other variety in at least one phenotypical characteristic. It must also be satisfactorily uniform (not less

than for all of the varieties with which it is compared) for all characteristics used in the process establishing distinctness. The number of visually observed (qualitative) and measured (quantitative) characteristics is large for many species, and the number of varieties tested simultaneously within one trial (candidate varieties and an established variety used as a reference) is often large. In member states of the UPOV (International Union for the Protection of New Varieties of Plants – more than 80 member states) DUS requirements are checked in special DUS trials conducted according to UPOV Guidelines issued for particular species. For white mustard and oilseed rape (see TG/179/3, 2001 and TG/36/6, 2002) it is stated in the Guidelines that DUS trials ought to be established in RCB designs in two replicates and that 15 observations (measurements) should be taken from every plot, giving in total 30 measurements for every variety for all characteristics. For white mustard 13 continuous (measured) characteristics are involved, whereas for oilseed rape there are 15 such characteristics. Final decisions are usually taken after three years of testing at one location (oilseed rape) or after two years (white mustard). Because of the large numbers of tested objects (varieties), characteristics and measurements, DUS trials are expensive. Therefore various attempts have been made to enable the optimization of DUS testing. Roberts (2011) proposed the testing of different subsets of established varieties. In a series of papers by Pilarczyk et al. (2011, 2012, 2014, 2015), the reduction of the number of measurements within single trials was proposed and the consequences of such reduction investigated. Using real data from DUS trials on several grass species (perennial ryegrass, amenity type of perennial ryegrass, meadow-grass and red fescue) they concluded, that, when decisions are taken after three years of testing, a reduction in numbers of measurements by 50% has very limited impact on decisions concerning the distinctness and uniformity of tested varieties. It should be mentioned that the stability of varieties is usually not tested, since according to UPOV regulations, if a variety is distinct and uniform it is deemed also to be stable. The aim of this paper is to investigate the impact of similar reduction in numbers of measurements on DUS decisions concerning varieties of two dicotyledonous species: white mustard and oilseed rape.

2. Materials and methods

The analyzed set of data consists of results of DUS trials on white mustard and oilseed rape conducted by the Research Centre for Cultivar Testing at Słupia Wielka Experimental Station during the period from 2005 to 2009. Each trial was established in a randomized complete block design (RCB) with two replicates (complete blocks). The trials on oilseed rape were performed in accordance with UPOV Guidelines TG/36/6 (2002), while the trials on white mustard were performed according to the similar Guidelines TG/179/3 (2001). In the trials on oilseed rape there were 24 characteristics observed, including 14 measured (continuous) characteristics, whereas in the trials on white mustard there were 23 characteristics, of which 17 were measured. In this paper only measured characteristics are considered, as the interpretation of such data involves the use of statistical methods. For visually assessed characteristics, other (non-statistical) methods are applied. The list of measured characteristics for the two analyzed species is as follows:

White mustard. *Cotyledon characteristics*: length, width; *Leaf characteristics:* number of lobes, length (blade and petiole), width (widest point), length of petiole; *Flower characteristics:* length of petals, width of petals; *Plant characteristics*: total length; time of flowering; *Siliqua characteristics*: length, width, length of beak, length of peduncle, number of seeds, weight of 1000 seeds.

Oilseed rape: *Cotyledon characteristics*: length, width, saddle depth; *Leaf characteristics:* number of lobes, length (blade and petiole), width (widest point), length of petiole; *Flower characteristics:* length of petals, width of petals; *Plant characteristics*: height, total length including side branches; *Siliqua characteristics*: length, width, length of beak, length of peduncle.

For all characteristics there were 60 measurements taken within each trial for every variety, which means that there were 120 measurements after two years and 180 measurements after three years.

The same approach was applied here as in papers by Pilarczyk at al. (2011, 2012, 2014, 2015) for several species of grasses. Namely the data were analyzed

twice: the first time all available measurements (full data) were used, while the second analysis used only half of them (reduced data), randomly chosen. The decisions concerning distinctness and uniformity of tested varieties obtained based on full and on reduced data (after two and three years of testing) were compared.

In testing distinctness the method officially promoted by UPOV (known as the COYD method) was applied in the version implemented in the DUSTNT package (Watson, 2000). In this method two alternative models are utilized. When there is no significant (at a significance level of 0.05) relationship between varietal yearly means and over-years means, an additive model is used in the form

$$y_{ij} = \mu_j + a_i + e_{ij}, \qquad i = 1,2,...,V, \quad j = 1,2,...,L$$

but if such a relationship exists, a modified model of the following form is applied:

$$y_{ij} = \mu_j + b_j a_i + e_{ij}$$

where y_{ij} denotes the mean value of an analyzed characteristic obtained in year j for variety i, μ_j is the mean value for year j, a_i is the effect of variety i, b_j is the coefficient of regression of yearly varietal means on over-years means, and e_{ij} denotes the polled error. However, independently of the model actually used, for testing the significance of differences among varieties the LSD (least significant difference) of Fisher is used. Details of the COYD method can be found in a paper by Talbot (2000).

For testing uniformity the COYU (combined over-years uniformity) method (again in the version implemented in the DUSTNT package) was applied. In the COYU method, the average values of standard deviations calculated for all tested varieties over replicates and years form the basis of decisions. For distributional reasons, the standard deviations are logarithmically transformed, and then the maximum admissible value *UC* (sometimes called the uniformity criterion) of the standard deviation for each variety is calculated according to the formula:

$$UC = SD_r + t_p\sqrt{V[1/Y + 1/(YW)]}$$

where SD_r denotes the average value of standard deviations of the reference set of varieties, V is the variance among transformed standard deviations of all tested varieties, t_p is the one-side value of Student's t-distribution for degrees of freedom as for V and for significance level p, Y is the number of years, and W is the number of varieties in the reference set. If the average standard deviation of a variety is smaller than UC, the variety is accepted as satisfactorily uniform. According to the UPOV recommendation for dicotyledonous species the following levels of significance are to be applied:

for rejecting after three years of testing: 0.2%
for rejecting after two years of testing: 0.2%
for accepting after two years of testing: 2.0%

Throughout this paper we applied the COYD and COYU methods in exactly the form promoted by UPOV, both for full data and for reduced data, after which the decisions on distinctness and uniformity were compared.

3. Results

Decisions concerning the distinctness and uniformity of white mustard varieties are usually taken after two years of testing. The results of testing of distinctness of white mustard varieties in two-year series are presented in Table 1.

Table 1. White mustard. Numbers of tested pairs of varieties and numbers of indistinct pairs

Series	Number of tested pairs	Number of indistinct pairs	
		Full data	Reduced data
2005–2006	378	60	59
2006–2007	861	125	131
2007–2008	903	166	162
2008–2009	1378	0	0
In total	3520	351	352

As can be seen, the numbers of pairs of varieties declared indistinct are practically the same for full data and for data reduced by 50%. For example, in the 2006–2007 period, based on full data 15% pairs are found to be indistinct, while according to the reduced data 16% pairs are indistinct.

DUS testing of oilseed varieties usually lasts for three years, with the possibility of an early decision after two years. Trials on oilseed rape are split into separate experiments on traditional (population) varieties and on hybrid varieties. The results of testing of the distinctness of traditional varieties after two and three years are presented in Table 2. In the case of oilseed rape the numbers of pairs found to be indistinct are similar for both types of data.

Table 2. Oilseed rape – traditional varieties. Numbers of tested pairs of varieties and numbers of indistinct pairs

Series	Number of tested pairs	Number of indistinct pairs	
		Full data	Reduced data
2005–2006	2775	769	813
2006–2007	3240	441	432
2007–2008	3741	260	264
2008–2009	7503	7	11
In total	17259	1477	1520
2005–2007	2145	409	410
2006–2008	3081	518	159
2007–2009	3160	3	6
In total	8386	930	575

Similar results concerning the distinctness of oilseed hybrid varieties after two and three years of testing are shown in Table 3.

After two years of testing, the number of pairs of varieties found to be indistinct when reduced data sets are used is slightly higher (by less than 1% of tested pairs) than for the full data, both for traditional and hybrid varieties. After three years of testing, the number of indistinct pairs of hybrid varieties is exactly the same for full and reduced data. Rather unexpectedly, however, for traditional varieties the number of indistinct pairs is slightly higher for full than for reduced data.

Table 3. Oilseed rape – hybrid varieties

Series	Number of tested pairs	Number of indistinct pairs	
		Full data	Reduced data
2005–2006	780	146	163
2006–2007	741	75	79
2007–2008	666	74	91
2008–2009	990	4	1
In total	3177	299	334
2005–2007	406	47	47
2006–2008	351	28	28
2007–2009	435	2	2
In total	1192	77	77

In order to compare uniformity decisions for full and reduced data, the coefficient of concordance CC as described by Pilarczyk et al. (2015) was calculated. This is the ratio of the number of concordant decisions concerning uniformity to the total number of such decisions. This coefficient is equivalent to the overall proportion of agreement p_0 used often in testing inter-rater agreement; see Kundel and Polansky (2003) or Landis and Koch (1977). The results are presented in Table 4, for the two considered species. Rows marked 'W' contain results concerning all tested varieties (established and candidate), while rows marked 'K' contain the results for candidate varieties only. It is seen that after two years of testing, for both white mustard varieties and for the two types of oilseed rape varieties, the proportion of concordant uniformity decisions is about 80%. For three-year series this proportion reaches 97%.

4. Comments and conclusions

In earlier papers by Pilarczyk and others (2011, 2012, 2014, 2015) it has been shown that for certain species of grasses (all monocotyledonous species) it is possible to reduce the number of measurements in DUS trials by 50% without serious impact on decisions concerning the distinctness and uniformity of the tested varieties. Here, following analysis of several two- and three-year series of

Table 4. Coefficients of concordance *CC* in DUS trials on white mustard and oilseed rape varieties

Series	Set of varieties	White mustard	Oilseed rape hybrid varieties	Oilseed rape traditional varieties
2005–2006	W	79.52%	83.48%	66.77%
	K	76.67%	87.50%	56.15%
2006–2007	W	78.89%	81.50%	79.98%
	K	86.67%	84.42%	79.52%
2007–2008	W	78.29%	81.36%	80.62%
	K	70.00%	80.00%	81.00%
2008–2009	W	78.62%	82.59%	80.31%
	K	75.56%	84.74%	81.82%
Average	W	78.83%	82.23%	76.92%
	K	77.23%	84.17%	74.62%
2005–2007	W	-	99.49%	88.22%
	K	-	98.46%	97.25%
2006–2008	W	-	99.74%	98.35%
	K	-	99.40%	97.77%
2007–2009	W	-	98.10%	99.11%
	K	-	95.45%	99.58%
Average	W	-	99.11%	95.23%
	K	-	97.77%	98.20%

W – results concerning all tested varieties (established and candidate)
K– results for candidate varieties only

DUS trials on white mustard and oilseed rape varieties (both dicotyledonous species), it can be concluded that:

- Decisions concerning distinctness and uniformity after three years of testing are practically the same for full data and for data reduced by 50%;
- Distinctness decisions after two years of testing are again very similar for full and reduced data, but uniformity decisions differ significantly.

Therefore the overall conclusion is that a reduction in numbers of measurements is also possible in DUS trials on oilseed rape and white mustard, but subject to the condition that final decisions are taken after collecting data over three years.

REFERENCES

Kunde H.L., Polansky M. (2003): Measurement of observer agreement. Radiology 228: 303-308.

Landis J.R., Koch G.G. (1977): The measurement of observer agreement for categorical data. Biometrics 33: 159-174.

Pilarczyk W., Kowalczyk B. (2011): The influence of reduction of the number of measurements on decisions concerning distinctness of perennial ryegrass varieties. Biometrical Letters 48(2): 149-157.

Pilarczyk W., Kowalczyk B. (2012): The influence of reduction in measured plant number on distinctness of amenity varieties of perennial ryegrass (*Lollium Perenne* L.) tested in DUS trials. Colloquium Biometricum 42: 35-42.

Pilarczyk W., Kowalczyk B. (2014): The influence of reduction of the number of measurements on decisions concerning distinctness of red fescue varieties. Colloquium Biometricum 44: 79-88.

Pilarczyk W., Kowalczyk B., Bakinowska E. (2015): Comparison of uniformity decisions in DUS testing for full and reduced numbers of measurements. Biometrical Letters 52(1): 45-53.

Roberts A. (2011): Cyclic planting of established varieties to reduce trial size. TWC/29/26. UPOV Geneva.

Talbot M. (2000): The Combined-Over-Years Distinctness and Uniformity criteria. TWC/18/10. UPOV Geneva.

TG/36/6 (2002): Guidelines for the conduct of tests for distinctness, uniformity and stability for Rape Seed. UPOV Geneva.

TG/179/3 (2001): Guidelines for the conduct of tests for distinctness, uniformity and stability for White Mustard. UPOV Geneva.

Watson S. (2000): DUST for Widows (DUSTNT). TWC/18/13. UPOV Geneva.