



Comparison of uniformity decisions in DUS testing for full and reduced numbers of measurements

Wiesław Pilarczyk¹, Bogna Kowalczyk², Ewa Bakinowska¹

¹Department of Mathematical and Statistical Methods, Poznan University of Life Sciences,
Wojska Polskiego 28, 60-637 Poznań, Poland, e-mail: wpilar@up.poznan.pl,
ewabak@up.poznan.pl

²The Research Centre for Cultivar Testing at Słupia Wielka, 63-022, Poland, e-mail:
b.kowalczyk@coboru.pl

SUMMARY

It is investigated how a reduction in the number of measurements influences uniformity decisions in distinctness, uniformity and stability (DUS) testing. Using real data from DUS trials performed in Poland for several species, it is shown that when final decisions are taken after three years of testing, a reduction in the number of measurements by 50% of the numbers indicated in the Guidelines has very limited impact on decisions (rejection or approval of candidate variety). Decisions taken after one year (or two years) are more dependent on the numbers of measurements.

Key words: DUS testing, grasses, number of measurements, uniformity

1. Introduction

At the end of the breeding process of new varieties of any cultivated species, a new (often called candidate) variety, before being listed on (for example) the National List of Varieties, must be checked to establish whether it is distinct from any other variety (variety of common knowledge) and whether it is sufficiently uniform and stable. According to the UPOV (International Union for the Protection of New Varieties of Plants) Convention, a new variety must be statistically different from all other varieties for at least one phenotypic characteristic and must be sufficiently uniform for all characteristics used for establishing distinctness. Sufficient uniformity means that its uniformity must be at least at the uniformity level of the varieties used as standards (varieties

already listed). If a variety is distinct and uniform it is deemed also to be stable. In some cases stability is also tested using seed of different generations of the same variety within the same trial. In testing distinctness the mean values of characteristics are involved, whereas in testing uniformity, the standard deviations are compared. Uniformity is usually more difficult to achieve (as uniformity concerns all characteristics) than distinctness (where a significant difference for one characteristic is sufficient). Decisions are usually taken after three years of testing in one location (the so-called DUS trials). Depending on species, the number of observed (measured) characteristics varies from a dozen or so up to nearly one hundred. Among them there are usually about 10 quantitative characteristics. Trials are usually established as randomized complete block experiments in two to six replicates. The total number of observed (measured) plants for agricultural species varies from 30 to 60. Because of the large number of compared varieties (established and candidate), the large number of characteristics and measurements (in grasses usually 60 measurements are required for every characteristic) and the fact that the final decisions are usually taken after three years of testing, DUS trials are very expensive. Therefore, under the auspices of UPOV, some optimization efforts have been undertaken (see the documents TWC/26/17, 2008 and TWC/29/26, 2011). In papers by Pilarczyk and Kowalczyk (2011, 2012, 2014) it has been shown that a reduction in the number of measurements by 50% (for quantitative characteristics) has a very limited impact on decisions concerning the distinctness of grass varieties. In this paper the consequences of such a significant reduction in numbers of measurements for uniformity decisions concerning varieties of several species of grasses are reported.

2. Materials and methods

The same DUS trial data as those used by Pilarczyk and Kowalczyk (2011, 2012, 2014) were used here. The data concern red fescue, meadowgrass and perennial ryegrass (including amenity type) varieties. The trials were performed

by the Research Centre for Cultivar Testing in the years 2005–2009 at the Experimental Station in Słupia Wielka. All trials were established as randomized complete block experiments in six replicates. In accordance with UPOV Guidelines (TGP/4/8, 2006; TG/67/5, 2006), 10 plants were taken at random from every plot for making measurements, giving a total of 60 measurements for every variety and characteristic. After three years of testing there were a total of 180 measurements. There were 93, 117, 96, 90 and 96 varieties of red fescue tested in the years 2005–2009, and respectively 91, 90, 88, 88 and 111 varieties of perennial ryegrass, 44, 62, 56, 51 and 55 varieties of meadowgrass, and finally 90, 89, 87, 87 and 110 varieties of the amenity type of perennial ryegrass. Using such extensive data sets, the uniformity of all varieties was tested firstly using all available data and secondly using only half (randomly chosen) of the data. For testing uniformity, the standard procedure COYU (Combined-Over-Years Uniformity) (Talbot, 2000), implemented in the DUSTNT package (Watson, 2000; see also Weatherup, 1992), was applied. The COYU method is officially promoted by UPOV as the best method for testing uniformity of new varieties for cross-pollinated species. Decisions are usually taken after three (sometimes after two) years of testing. The list of observed (measured) quantitative characteristics for the considered species is given in Table 1.

In the COYU method, the uniformity of each variety is verified by comparison of its standard deviation (after logarithmic transformation) with the threshold value UC , calculated as (see Talbot, 2000)

$$UC = SD_r + t_\alpha \cdot \sqrt{V \cdot (1/Y + 1/(Y \cdot W))},$$

where SD_r denotes the average of the standard deviations (also after logarithmic transformation) of W varieties with which the new variety is compared, V denotes the variance among the standard deviations used for calculation of SD_r , Y denotes the number of years, and t_α denotes the one-sided t -Student critical value at significance level α with appropriate degrees of freedom. When the transformed standard deviation of a candidate variety is smaller than UC , the

variety is satisfactorily uniform and can be listed. All the tests were performed at the UPOV-recommended $\alpha=0.002$ significance level.

Table 1. List of measured (quantitative) characteristics

Characteristics	Species			
	Red fescue	Meadowgrass	Perennial ryegrass fodder type	Perennial ryegrass amenity type
<i>whole plant characteristics:</i>				
natural height	+	+	+	+
leaf width	+	+		
length of longest stem	+	+	+	+
upper internode length	+	+		
<i>flag leaf characteristics:</i>				
length	+	+	+	+
width	+	+	+	+
size	+		+	+
shape (ratio of length to width)	+		+	+
<i>inflorescence characteristic:</i>				
length	+	+	+	+
number of spikelets			+	+
density			+	+

3. Results

The method described above was applied twice: the first time with the use of all available data (60 measurements for each characteristic in every trial), and the second time with the use of only a randomly chosen half of the data. The random subsample was extracted from the database in such a way that 50% of the measurements from every plot were taken. The results were collected in the form presented in Table 2.

Table 2. Numbers of concordant (n_{11} and n_{22}) and contradictory decisions (n_{12} and n_{21})

		full data	
		uniformity	lack of uniformity
reduced data	uniformity	n_{11}	n_{12}
	lack of uniformity	n_{21}	n_{22}

In Table 2, n_{11} denotes the common number of positive decisions (no reason to reject uniformity), n_{22} the common number of negative decisions (variety rejected as non-uniform), n_{12} the number of cases where the variety was declared non-uniform based on the reduced data set but uniform when the full data set was considered, and n_{21} the number of cases where the reverse situation appeared.

Using data collected in the form presented in Table 2, the so-called coefficient of conformity CC was calculated according to

$$CC = \frac{n_{11} + n_{22}}{n},$$

where $n = (n_{11} + n_{12} + n_{21} + n_{22})$.

Coefficients CC were calculated for all varieties (established and candidate) present in trials and separately for candidate varieties. The results are given in Table 3.

4. Comments and conclusions

In papers by Pilarczyk and Kowalczyk (2011, 2012, 2014) it was shown that reduction of the number of measurements in DUS trials on different species of grasses had very limited impact on distinctness decisions concerning varieties both in single trials and in two- and three-year series of such trials. A slightly different situation exists in the case of the testing of uniformity. As shown in Table 3, the uniformity decisions after three years of testing are practically identical. However, for two-year series there exist serious discrepancies. This means that a reduction in the number of measurements in DUS trials on grasses is possible as long as the final decisions on candidate varieties are taken after three years. This is common practice among UPOV members. However, when a so-called early decision (taken after two years of testing) is expected, the number of measurements should not be changed.

Table 3. Coefficients of conformity *CC* (in percentages) in two- and three-year series

Series		Red fescue	Meadowgrass	Perennial ryegrass – fodder type	Perennial ryegrass – amenity type
2005-2006	W	74.70%	80.49%	74.58%	74.60%
	K	73.33%	77.50%	75.36%	74.44%
2006-2007	W	74.53%	76.57%	77.21%	97.71%
	K	69.70%	78.63%	79.56%	97.22%
2007-2008	W	59.67%	79.95%	75.58%	97.93%
	K	44.00%	78.89%	76.72%	97.22%
2008-2009	W	81.10%	77.04%	64.77%	72.96%
	K	84.44%	76.19%	65.28%	77.78%
Average	W	72.50%	78.51%	73.03%	85.80%
	K	67.87%	77.80%	74.23%	86.67%
2005-2007	W	100%	100%	100%	98.83%
	K	100%	100%	100%	97.22%
2006-2008	W	100%	100%	99.42%	98.36%
	K	100%	100%	100%	100%
2007-2009	W	100%	100%	-	-
	K	100%	100%	-	-
Average	W	100%	100%	99.71%	98.60%
	K	100%	100%	100%	98.61%

Row W – coefficients of conformity for all tested varieties; row K – coefficients of conformity for candidate varieties only.

REFERENCES

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

- TGP/4/8 (2006): Ryegrass. Guidelines for the conduct of tests for distinctness, uniformity and stability, UPOV Geneva.
- TG/67/5 (2006): Red Fescue. Guidelines for the conduct of tests for distinctness, uniformity and stability, UPOV Geneva.
- TWC/26/17 (2008): Some consequences of reducing the number of plants observed in the assessment of quantitative characteristics of reference varieties, UPOV, Geneva.
- TWC/29/26 (2011): Cyclic planting of established varieties to reduce trial size, UPOV, Geneva.
- Weatherup S.T.C. (1992): Distinctness, Uniformity and Stability trial (DUST) analysis system. User manual. Department of Agriculture for Northern Ireland Biometrics Division, Belfast BT9 5PX.
- Watson S. (2000): DUST for Windows (DUSTNT), TWC/18/13 UPOV, Geneva.