E

TWC/27/18
ORIGINAL: English
DATE: June 4, 2009

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Twenty-Seventh Session
Alexandria, Virginia, United States of America
June 16 to 19, 2009

AN ADJUSTMENT TO THE COYD METHOD WHEN VARIETIES ARE GROUPED WITHIN THE DUS TRIAL

Document prepared by experts from France and the United Kingdom

Background

1. In some crops, it is possible to use grouping characteristics to define groups of varieties such that all the varieties within a group will be distinct from all the varieties of any other group ("distinct groups"). This grouping is preserved in trial layouts so that, within a replicate, varieties in the same group are adjacent. (See document TG/1/3, section 4.8 "Functional Categorization of Characteristics"). Field pea is an example of such a crop in the United Kingdom.

2. In the United Kingdom the current method of analysis used in pea is to apply analysis of variance for a randomized complete block design for each trial and then apply COYD (also analysis of variance) to the variety-by-trial means. This takes no account of the grouping except that semi-leafless and conventional types are analyzed separately because they are distinct types with different characteristics. In some crops, different groups are analyzed separately, but in pea many groups are too small.

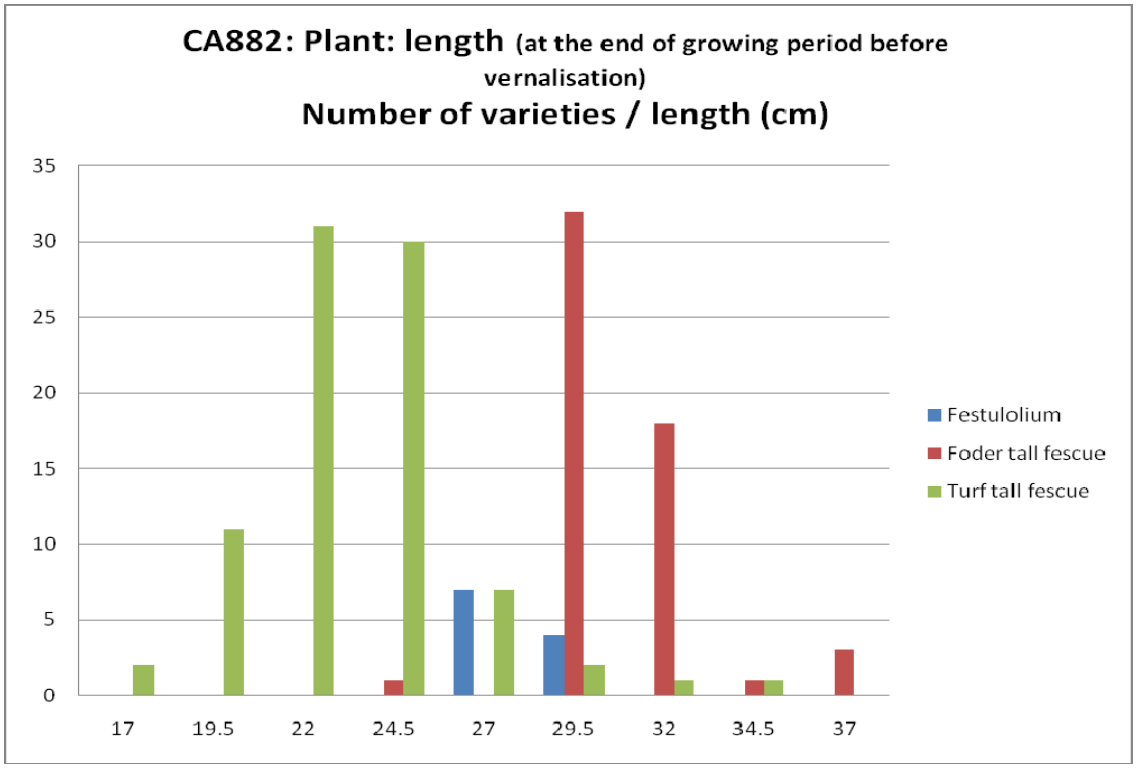
3. In documents TWC/23/8 "A Proposal for an Adjustment to the COYD Method when Varieties are Grouped Within the DUS Trial" and TWC/26/14 "An Adjustment to the COYD

Method when Varieties are Grouped Within the DUS Trial", experts from the United Kingdom proposed an adjustment to the COYD method that is both effective and relatively simple to implement. The benefits of the method were demonstrated by reference to data from 10 years of pea trials in the United Kingdom.

4. COYD is more commonly used in cross-pollinated crops. In light of this, at the twenty-sixth session of the TWC, held in Jeju, Republic of Korea, from September 2 to 5, 2008, an expert from France "considered that it would be interesting to test the adjusted COYD in a cross-pollinated crop, and offered to prepare a document on that subject. The TWC agreed that France should prepare such a document." (see document TWC/26/29, Report, paragraph 74). Subsequently, the expert from France collated six years of data for Tall Fescue *(Festuca arundinacea* Schreb.). This document further demonstrates the adjusted COYD method by application to this example.

5. In cross-pollinated species, such as Tall Fescue, the situation is generally not as clear as for self-pollinated species. The groups are based on the declaration of the applicant and are more difficult to assess in the field. For example, it is very difficult to prove that a given variety is a turf or a fodder variety. Figure 1 shows that, generally, turf Tall Fescue varieties are shorter than fodder varieties. However, some fodder varieties are shorter than some turf varieties.

Figure 1: Distribution of grouping 1 for CA882 Plant: length (at the end of growing period before verbalization) Years 2006-8



Outline of the method

6. When grouping is possible such that all the varieties within a group will be distinct from all varieties of any other group ("distinct groups"), comparisons are only necessary between varieties in the same group. In principle, it would be possible to analyze groups separately; in practice some groups have too few varieties. Instead, we propose that the over-years analysis of variance (COYD) be adjusted to take into account the group-by-year interaction.

7. Whereas the standard COYD has terms for 'year' and 'variety', the adjusted form has terms for 'year', 'group', 'variety-within-group' and 'group-by-year'. The standard error (and LSD) is then calculated for differences between pairs of varieties within the same group. It is assumed that the same standard error is applicable within all groups.

Details of the method

8. With COYD, the analysis of variance is based on variety-by-year means for two or three years depending on the crop. Usually, only varieties present in all years are considered. The analysis of variance includes effects for year and variety. The standard error, SED_{COYD}, for the difference between two varieties is given by:

$$SED_{COYD} = \sqrt{\frac{2}{n} RSS_{COYD}}$$

where n is the number of years and RSS_{COYD} is the residual sum of squares from the analysis of variance (based on means).

9. We propose that an extra factor, the group-by-year interaction, be included in the analysis of variance. So in GenStat terminology (Payne *et al.*, 2008, *The Guide to GenStat Release 11, Part 2: Statistics*. VSN International, Hemel Hempstead) we have:

Block structure: Year + Year.Group
Treatment structure: Group/Variety

10. The standard error, SED_{adj}, for the difference between two varieties *in the same group* is given by:

$$SED_{adj} = \sqrt{\frac{2}{n} RSS_{adj}}$$

where RSS_{adj} is the residual sum of squares from the analysis of variance that includes the group-by-year interaction term. Unlike SED_{COYD}, SED_{adj} excludes variability due to the interaction between varietal groups and years. We believe this is reasonable as candidate varieties are only assessed for distinctness against varieties in their group.

Application to Tall Fescue

11. This adjustment method has been applied to the Tall Fescue DUS trial data in France from 2003 to 2008.

12. Only quantitative (MS - measurement of a number of individual plants or parts of plants) characteristics were considered – these are listed in Table 1. Two sets of grouping were considered: grouping 1 consists of three groups and grouping 2 of five groups (Table 2).

The Festulolium group in grouping 1 splits into two subgroups for grouping 2 and the fodder Tall Fescue into another two groups.

Table 1: List of characteristics

Characteristic no.	**Characteristic**
CA982	Plant: length (at the end of growing period before vernalization).
CA809	Plant: natural height after vernalization
CA910	Plant: time of inflorescence emergence (in 2nd year)
CA880	Plant: natural height at inflorescence emergence
CA817	Flag leaf: width (same flag leaf as that used for 7)
CA819	Flag leaf: length (flag leaf on representative stem, within 2 weeks after inflorescence emergence)
CA813	Stem: length of longest stem (inflorescence included; when fully expanded)
CA870	Stem: length of upper internode
CA844	Inflorescence: length (when fully expanded)

Table 2: Groups and numbers of varieties

	2003-5	**2004-6**	**2005-7**	**2006-8**
Group 1				
Festulolium	6	8	10	11
Fodder tall fescue	*45	48	55	55
Turf tall fescue	**64	68	76	85
Group 2				
Festulolium (*F.arundinacea* X *L.multiflorum*)	1	3	4	5
Festulolium (*F.pratensis* X *L.multiflorum*)	5	5	6	6
Fodder tall fescue	*43	46	53	53
Fodder tall fescue amphiploid	2	2	2	2
Turf tall fescue	**64	68	76	85

* two varieties less for CA809

** one variety less for CA817, CA819

13. Comparisons were made between standard COYD and the group-adjusted COYD on sets of three consecutive years. Tables 3 and 4 show the decreases in the standard errors (or LSDs) for grouping 1 and grouping 2 respectively.

Table 3: Reduction in SED given by the adjusted-COYD method compared to the standard COYD method: grouping 1

	Set of three years			
Characteristic	2003-5	2004-6	2005-7	2006-8
CA982	25.3%	30.7%	32.9%	10.3%
CA809	7.8%	15.1%	15.5%	24.8%
CA910	3.7%	9.8%	13.4%	5.6%
CA880	21.7%	23.7%	20.8%	12.7%
CA817	2.4%	2.7%	14.8%	11.3%
CA819	2.5%	4.9%	18.8%	18.0%
CA813	8.7%	18.5%	21.3%	26.8%
CA870	8.2%	19.8%	6.8%	5.8%
CA844	10.4%	6.8%	9.9%	8.1%

Table 4: Reduction in SED given by the adjusted-COYD method compared to the standard COYD method: grouping 2

	Set of three years			
Characteristic	2003-5	2004-6	2005-7	2006-8
CA982	25.4%	32.3%	34.2%	11.2%
CA809	17.8%	17.4%	18.0%	25.6%
CA910	5.8%	11.3%	14.6%	6.3%
CA880	22.0%	24.6%	22.3%	14.2%
CA817	3.3%	3.3%	15.6%	11.9%
CA819	5.8%	9.4%	18.8%	17.9%
CA813	13.1%	19.3%	22.5%	27.7%
CA870	9.5%	19.4%	6.8%	5.6%
CA844	10.5%	8.9%	10.5%	9.8%

14. In all cases, there is a decrease in the standard error when the adjustment is used. This was often substantial. In general, the choice of grouping 1 or 2 made little difference (Table 5), especially for the later set of years. However, this was not true for characteristic CA809 in the 2003-5 data set. Here the group “Fodder Tall Fescue amphiploid” made a substantial contribution to the group-by-year interaction.

Table 5: Reduction in SED given by the adjusted-COYD method with grouping 1 compared to grouping 2

	Set of three years			
Characteristic	2003-5	2004-6	2005-7	2006-8
CA982	0.1%	2.4%	1.8%	1.0%
CA809	10.9%	2.7%	2.9%	1.1%
CA910	2.2%	1.6%	1.4%	0.7%
CA880	0.3%	1.2%	1.9%	1.7%
CA817	1.0%	0.6%	0.9%	0.7%
CA819	3.5%	4.7%	-0.1%	-0.2%
CA813	4.9%	1.0%	1.6%	1.3%
CA870	1.4%	-0.5%	0.0%	-0.3%
CA844	0.0%	2.2%	0.7%	1.9%

15. The decrease in the standard error tends to be higher when there are differences between the groups for the characteristic. Figure 2 shows that turf Tall Fescue is shorter than fodder Tall Fescue and, for characteristic CA809, the reduction in standard error is 24.8 %. For the characteristic CA910, the groups overlap and the reduction is only 5.6 % (see Figure 3).

Figure 2: Distribution of grouping 1 for CA809: natural height after vernalization
Years 2006-8

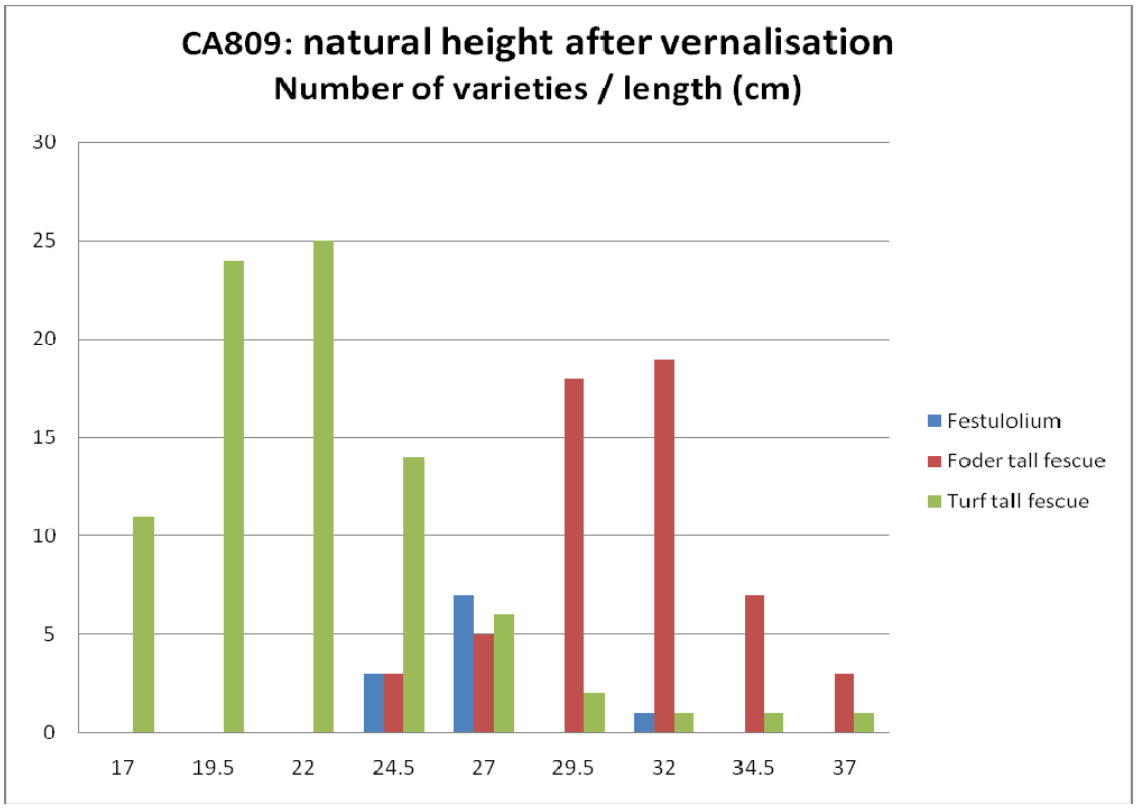
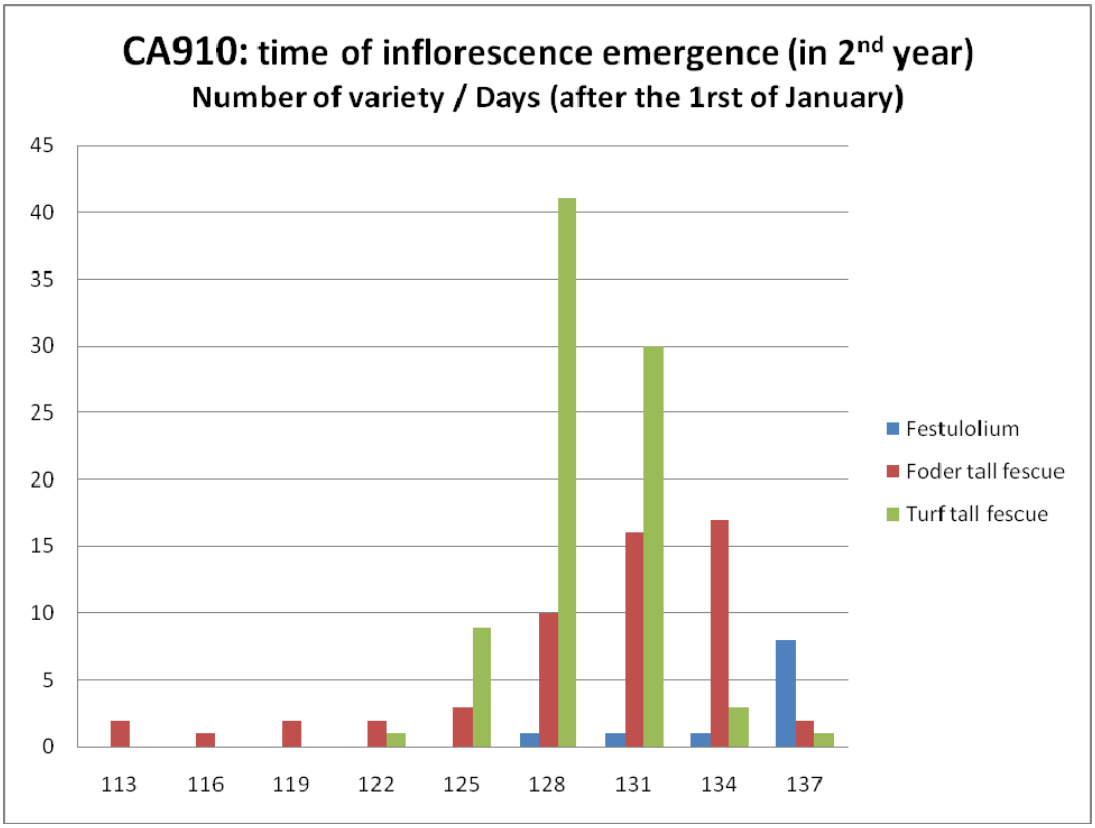


Figure 3: Distribution of grouping 1 for CA910: Time of inflorescence emergence (in 2nd year) Years 2006-8



16. It should be noted that, where it is necessary to compare varieties from different groups, the standard error will be larger than for comparisons within the same group. It may also be larger than for the standard COYD approach. Therefore, this method should be applied only once the groups have been assigned.

Software

17. A software module has been written to allow SASA (Science and Advice for Scottish Agriculture), United Kingdom, to apply this method routinely. This is based on the DUST package's TVRP module. This is about to be introduced in DUST's Windows interface.

Conclusions

18. We have proposed a method for adjusting COYD when grouping characteristics are used to identify distinct groups of varieties. It is appropriate when some group sizes are too small to allow separate analyses. The benefits of the approach have now been demonstrated in both a self-pollinated crop; Pea, using data from the United Kingdom; and a cross-pollinated crop; Tall Fescue using data from France. The method should not be used if the resulting residual degrees of freedom are less than 12. The method is relatively easy to implement.

19. In cross-pollinated species, this method permits a flexible approach, when an applicant submits a new variety type for DUS testing or when several types exist in a trial. All varieties

in the trial can still be compared and we may have a better discrimination in some characteristics.

18. Further consideration is required as to whether the adjustment should always be applied for a characteristic or if its use might depend on the size of the group-by-year interaction. Currently, the former approach has been adopted. Consideration is also needed with regard to the role of modified joint regression.

[End of document]