



TWC/33/20 Rev.
ORIGINAL: English
DATE: July 13, 2015

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS
Geneva

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Thirty-Third Session
Natal, Brazil, June 30 to July 3, 2015

**CALCULATED THRESHOLDS FOR EXCLUDING VARIETIES OF COMMON KNOWLEDGE FROM THE
SECOND GROWING CYCLE WHEN COYD IS USED**

Document prepared by experts from the United Kingdom

Disclaimer: this document does not represent UPOV policies or guidance

INTRODUCTION

1. When DUS test are carried out over two or three independent growing cycles, results may be reviewed after the first cycle of testing in order to exclude varieties of common knowledge that are clearly distinct from the candidates (see document TGP/9 "Examining Distinctness"). When COYD is used to assess distinctness for a characteristic, it may be difficult to do this effectively based on experience and no formal mechanism has yet been described to inform such early decisions on distinctness.
2. In document TWC/25/14, a possible approach was proposed. However this approach makes some strong assumptions regarding the properties of the data. Document TWC/28/30 explored the effect of deviations from these assumptions and suggested possible solutions that may be employed if necessary. The conclusions were that whilst deviations from normality were not of great practical concern, the necessary assumption that variety-by-cycle variation is constant from cycle to cycle (homogeneity) is not true in general and the method would be too lenient as a result.
3. This document aims to address the concerns raised in TWC/28/30 by extending the approach to incorporate heterogeneity.

OVERVIEW

4. The aim of this approach is to identify after the first year of test which varieties of common knowledge are so different from the candidate that they do not need to be compared in the second year. To enable this, we estimate the probability that a candidate would be distinct on the 2-year COYD criterion from a particular variety of common knowledge, given the results from the first growing cycle. If the probability is suitably large, the candidate is declared distinct from that variety and does not need to be compared in the second cycle. The method is applied characteristic by characteristic. In order to judge the variability associated with measurements in a particular characteristic we need to have historical data. The approach might be used in combination with processes such as GAIA to arrive at a "Distinctness Plus" threshold.

THE METHOD IN BRIEF

5. The method is based on calculating the probability, p_D , that a candidate would be distinct on the 2-cycle COYD criterion based only on the first cycle's data. If the probability is suitably large, the candidate is declared distinct from that variety and does not need to be compared in the second growing cycle. This process can be inverted to identify thresholds for set probabilities.

6. As well as requiring the first cycle's trial data, the method requires historical data from past DUS trials. At least 10 cycles of trials are needed – more is better. This is used to estimate the variety-by-cycle variance for each characteristic and, importantly, its variability (or level of heterogeneity). The variety-by-cycle variance is a fundamental component of the COYD criterion (see document TGP/8 "Trial Design and Techniques Used in the Examination of Distinctness, Uniformity and Stability").

7. At the moment the method requires use of specialist statistical software to estimate the heterogeneity of the variety-by-cycle variance and the parameters of a gamma distribution. We use GenStat (see Annex). ASREML (perhaps in combination with R) is also capable and possibly so is SAS.

8. Apart from that, the method uses formulae, which whilst being a little complex, should be straightforward to implement in a program. It should not be necessary to update the thresholds every year.

9. Further detail on the method is given in the Annex, though full details are given in a paper that has been submitted to the Journal of Agricultural Science, Cambridge.

AN EXAMPLE

10. The proposed method is exemplified using a data set from the United Kingdom field pea distinctness trials from 1995 to 2013. The semi-leafless group of varieties was considered. The trials were carried out at Science and Advice for Scottish Agriculture (SASA), near Edinburgh. Each trial had two replicates and between 139 and 290 varieties. Thirteen quantitative characteristics were considered. Only those varieties with six or more cycles of data have been retained for this study; this left 222 varieties. A 2% probability level is used for COYD.

11. Table 1 shows the characteristics considered, along with some basic statistics to give an indication of the scales. Note some of these are scored. An index for heterogeneity is included. This is based on changes in deviances between models with and without heterogeneity over cycles divided by the corresponding change in degrees of freedom: the higher the index the greater the importance of the heterogeneity. The greatest heterogeneity was found for characteristics 5 and 28. Note that the level of varietal heterogeneity in the variety-by-cycle variance (not shown) was much lower.

Table 1. Characteristics considered in example data set with statistics

Characteristic (UPOV number)	Mean	Standard deviation	Minimum	Maximum	Over-cycle heterogeneity index
(5) Stem: number of nodes up to and including first fertile node	16.0	1.59	9.6	20.9	13.0
(15) Stipule: length (mm)	82.3	13.48	47.2	121.5	4.4
(16) Stipule: width (mm)	46.3	8.80	23.7	79.0	4.1
(21) Stipule: density of flecking (1-9)	5.3	0.90	2.5	8.0	4.3
(22) Petiole: length from axil to first leaflet or tendrill (mm)	83.2	13.34	34.8	128.6	5.8
(28) Flower: width of standard (mm)	31.8	2.64	23.3	41.1	9.1
(29) Flower: shape of base of standard (1-9)	6.8	1.02	4.0	9.0	3.8
(34) Peduncle: length from stem to first pod (mm)	72.9	24.41	12.0	145.7	4.6
(37) Pod: length (mm)	79.1	6.24	63.3	105.6	4.3
(38) Pod: width (mm)	13.9	1.22	10.5	18.6	3.4
(42) Pod: curvature (1-9)	2.4	0.58	1.0	5.5	2.5
(46) Pod: number of ovules	8.2	0.54	6.0	10.0	7.5
(57) Seed: weight	28.1	5.19	12.2	49.1	5.7

12. Table 2 shows the first cycle thresholds calculated for each characteristic based on setting distinctness probabilities p_D at 90% and 99%. These are compared with an average COYD criterion for the two cycle test (based on long-term data and equal to the long-term LSD). They are also compared to the tolerances based on experience that are currently used in the United Kingdom to exclude varieties of common knowledge after the first cycle.

13. Of course the first cycle thresholds are always larger than the COYD criterion. The degree to which they are larger depends on the degree of heterogeneity present, especially for larger values of p_D .

14. The calculated thresholds compare well with the current tolerances. With p_D at 90%, the calculated thresholds are mostly substantially less than the current tolerances. At 99%, the thresholds are still less for most characteristics, with some being substantially less.

15. Based on the calculated thresholds with p_D at 99%, the crop expert has proposed new first year tolerances to be used in the United Kingdom field pea DUS tests with semi-leafless varieties (Table 2). Note that the tolerances for scored MG characteristics do not change, though the calculations provide useful evidence.

Table 2. First cycle thresholds allowing for heterogeneity over cycles. For comparison, a two-cycle COYD criterion based on long-term variance component and the current first cycle tolerances current used by the United Kingdom (UK) based on experience are included. Proposed new tolerances are given based on the calculated thresholds with $p_D = 0.99$

Characteristic	Long-term COYD criterion	Threshold with $p_D=0.99$	Thresholds with $p_D=0.9$	Current UK tolerance	Proposed new tolerance
5	0.93	3.99	1.39	3	4.0
15	10.80	22.76	15.63	25	22.8
16	6.95	13.81	9.82	20	13.8
21*	0.95	1.96	1.37	3	3
22	12.61	27.63	18.49	30	27.6
28	2.39	5.83	3.55	12	5.8
29*	0.93	1.91	1.36	2	2
34	19.61	44.40	28.82	40	44.4
37	5.84	12.23	8.60	20	12.2
38	0.97	1.95	1.41	2	2.0
42*	0.83	1.61	1.15	2	2
46	0.47	1.00	0.66	2	1.0
57	4.03	9.44	6.01	8	9.4

* These characteristics are scored MG and so an integer tolerance is more appropriate

CONCLUSIONS AND FUTURE WORK

16. A method has been developed for calculating thresholds to identify varieties of common knowledge which are clearly distinct from the candidates, based on the trial for the first growing cycle. The method allows for any heterogeneity over cycles in the variety-by-cycle variance.

17. The method has been demonstrated on a United Kingdom field pea data set. It shows that it may be possible to justifiably reduce the first-cycle tolerances for many characteristics. In any case, the method provides a more systematic and transparent way of deriving the tolerances for quantitative characteristics.

18. In practice, it would be important to verify that characteristics do not have very large levels of skewness or kurtosis (see TWC/28/30). In such cases transformations could be considered or such characteristics could simply be ignored in these first cycle decisions. The pea example showed that such cases should be rare.

19. The method can straightforwardly be extended to three cycle situations.

20. The method might be incorporated in a wider system for a “distinctness plus” threshold, such as GAIA.

21. The method could also be modified to give an early indication as to whether a candidate may have distinctness problems as well as guidance on the closest reference varieties. Both of these should be of benefit to COYD users.

22. We would welcome other example data sets so that the new method can be tested further. We also plan to examine software options to make implementation easier.

ACKNOWLEDGEMENTS

23. This work has been funded by the Scottish Government. This is a summary of a paper submitted to the Journal of Agricultural Science, Cambridge.

Adrian Roberts (Biomathematics & Statistics Scotland (BioSS)), Ian Nevison (BioSS) & Tom Christie (Science and Advice for Scottish Agriculture (SASA))

[Annex follows]

MATHEMATICAL DETAIL

This annex gives a little more detail on the method. Further detail, including the derivation of the method, will be available in a paper submitted to the Journal of Agricultural Science, Cambridge. A “pre-paper” is available on request from the first author.

The first step is to fit mixed linear models for each characteristic, with and without heterogeneity. This can be done using a suitable software package such as GenStat or ASREML (the latter may be used in conjunction with R). This may also be possible in SAS – we have not investigated this fully yet.

The basic model assumes that the variety-by-cycle variance is constant over cycles. For this, fit the data to a model with a fixed effect for variety and a random effect for cycle.

The model for heterogeneity extends the basic model by estimating separate residual variances for each cycle. In GenStat this can be done using the variance structure mechanism – here is some example code:

```
VCOMPONENTS [FIXED=variety] RANDOM=cycle+cycle.variety
VSTRUCTURE [TERMS= cycle.variety; FORMATION=direct] MODEL =
diagonal,identity; FACTOR= cycle, variety
REML [PRINT=model,components,deviance,waldTests; MVINCLUDE=y; METHOD=AI]
response
```

In GenStat, it is essential to expand the data to include all possible combinations of variety and cycle (a “full” data set), with missing values for the response when the variety was not tested in a particular cycle. Without this, the algorithm tends not to converge or converge to an incorrect solution.

It is a good idea to check the residuals from these models to verify that the levels of skewness and kurtosis are not overly large (see TWC/28/30).

The deviances from the basic and extended models can be compared to evaluate the degree of heterogeneity present.

The individual residual (cycle-by-variety) variances, s_i^2 , can be extracted from the model output for the extended model.

Then fit a gamma distribution to the reciprocals of the variances, $\frac{1}{s_i^2}$, to estimate shape and rate parameters, $\hat{\alpha}$ and $\hat{\beta}$. This can be done in statistical software packages such as R or GenStat.

The calculated threshold for a set p_D is given by $\frac{-B+\sqrt{B^2-4AC}}{2A}$, where

$$A = K^2 - M\Phi^{-1}(p_D)^2,$$

$$B = -2K\Phi^{-1}\left(1 - \frac{p}{2}\right),$$

$$C = \left[\Phi^{-1}\left(1 - \frac{p}{2}\right)\right]^2 - L * \Phi^{-1}(p_D)^2, \text{ where}$$

$$K = \left(1 + \frac{3}{8} \frac{v_{12} + 2\hat{\alpha} - 2}{v_{12}(\hat{\alpha} - 2)}\right) \sqrt{\frac{\hat{\alpha} - 1}{\hat{\alpha}} \frac{\sum \frac{1}{s_i^2}}{n}},$$

$$L = \frac{1}{2}, \text{ and}$$

$$M = \frac{1}{4} \frac{(\hat{\alpha} - 1) \sum \frac{1}{s_i^2} (v_{12} + 2\hat{\alpha} - 2)}{\hat{\alpha} n v_{12}(\hat{\alpha} - 2)}.$$

Here p is the probability level used for COYD (e.g. 0.02), n is the number of cycles in the historical data set and v_{12} is the degrees of freedom for the 2-cycle COYD test.

Note that if A is negative, the threshold is nonsensical. This may happen when the level of heterogeneity is high and the required p_D value is high. The threshold is also sensitive to the degrees of freedom, v_{12} , especially when the level of heterogeneity is high, at least for larger values of p_D (e.g. 0.999). The threshold decreases as v_{12} increases. Given that v_{12} is not known in advance, it would be best to set a conservative value.

[End of annex and of document]