E

TWC/34/8
ORIGINAL: English
DATE: May 3, 2016

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

Geneva

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Thirty-Fourth Session
Shanghai, China, June 7 to 10, 2016

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 tests 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/33/20 Rev., an approach was proposed. This method was improved over previous versions by relaxation of the assumption that variety-by-cycle variation is constant from cycle to cycle. It allows for the often-seen material variation in the COYD criterion from cycle to cycle.

3. This document illustrates the application of the method to a field pea DUS data set, showing how beneficial it may be in practice.

OVERVIEW

4. The aim of this approach is to identify after the first test cycle which varieties of common knowledge are so different from the candidate that they do not need to be compared in the second cycle.

5. To achieve this, we estimate the probability that a candidate would be distinct on the 2-cycle 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.

6. 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 (see TGP/8 "Trial Design and Techniques Used in the Examination of Distinctness, Uniformity and Stability", Part II: Selected Techniques Used in DUS Examination, 1 "The GAIA Methodology").

THE METHOD IN BRIEF

7. The method is based on calculating the probability, pD, 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.

8. 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").

9. 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. Here GenStat was used. ASREML (perhaps in combination with R) is also capable and possibly so is SAS.

10. Apart from that, the method uses formula, 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.

11. Further detail on the method is given in document TWC/33/20 Rev. and in a paper (Roberts A.M.I., Nevison I.M., Christie T. (in press) Prediction of variety distinctness decisions under yearly heterogeneity. Journal of Agricultural Science doi: 10.1017/S0021859615001306).

AN EXAMPLE

12. 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.

13. 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.

14. Table 2 shows the first cycle thresholds calculated for each characteristic based on setting distinctness probabilities pD at 90%, 95% 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.

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 tendril (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

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

Table 2. First cycle thresholds allowing for heterogeneity over cycles. For comparison, the long-term 2% COYD criterion, the current first cycle tolerances currently used by the United Kingdom based on experience and proposed new tolerances are included.

Characteristic	Long-term COYD criterion	Threshold with p_D=0.99	Threshold with p_D=0.95	Threshold with p_D=0.9	Current tolerance of the United Kingdom	Proposed new tolerance
5	0.93	4.13	1.81	1.39	3	4.1
15	10.80	23.38	17.90	15.70	25	23.4
16	6.95	14.18	11.15	9.87	20	14.2
21*	0.95	2.01	1.56	1.38	3	3
22	12.61	28.38	21.31	18.56	30	28.4
28	2.39	5.99	4.18	3.56	12	6.0
29*	0.93	1.96	1.54	1.37	2	2
34	19.61	45.63	33.46	28.92	40	45.6
37	5.84	12.56	9.79	8.64	20	12.6
38	0.97	2.00	1.59	1.42	2	2.0
42*	0.83	1.66	1.31	1.16	2	2
46	0.47	1.03	0.77	0.67	2	1.0
57	4.03	9.70	7.01	6.02	8	9.7

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

15. 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 pD.

16. The results above are an update of those presented in TWC/33/20 Rev.. To establish the effect of using tolerances, the pea data set has been used to study the effect of first year decisions based on the existing and calculated tolerances.

17. First-year decisions were compared with COYD decisions in consecutive pairs of years (1995-96, 1996-97, 1997-98 etc.) for each characteristic. To evaluate the different thresholds, error rates were calculated:

- False positive rate: this is the proportion of times for each characteristic that the first-year threshold indicated a variety would be distinct from another variety when the subsequent second-year decision was non-distinct. This indicates the downside of taking early decisions: sometimes a pair of varieties might be declared distinct in the first year when they might later be found non-distinct. The rate of false positives is lower for higher thresholds.

- False negative rate is the proportion of times that the first-year decision was non-distinct when the second-year decision was distinct. This gives an indication of how useful the threshold might be in practice, with lower rates indicating that more pairs of varieties would be found distinct after the first year.

18. The results of this study are shown in Tables 3 and 4. These results should be interpreted with care since typically reference varieties that were clearly distinct from the candidate in at least one characteristic after the first year would have been removed from further comparisons. However the effect of this selection would be to give a pessimistic view of the performance of the calculated thresholds (false negative rates).

19. The rate of false positives was very low, especially with the existing tolerance of the United Kingdom and with the calculated threshold with pD being 0.99. Note that it is difficult to attain a 0% false positive rate simply because the two-year COYD criterion is in itself subject to variability.

20. The usefulness of the tolerances is represented by the false negative rates. The calculated tolerances with lower values of pD find more pairs of varieties distinct in the first year. Performance varies widely between characteristics e.g. 29, 46 and 57 have low false negative rates even with pD being 0.99. Remember that these rates are likely to be pessimistic due to selection in the example data set.

21. Choice of a suitable pD value for setting tolerances involves balancing the risks associated with false positives and negatives. In the case of the United Kingdom, a conservative approach has been taken. The calculated thresholds with the distinctness probability set at 0.99 had reasonable levels of false positives and false negatives. So based on the calculated thresholds with pD 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/VG characteristics (21, 29 and 42) are not based on the calculated threshold. However these calculations do give confidence in the tolerances currently used.

Table 3. Proportion of times in the United Kingdom pea data set that the first-year thresholds indicated a variety would be distinct from another variety when the subsequent COYD decision was non-distinct in the field pea data set (false positive)

Characteristic	Calculated threshold with p_D=0.99	Calculated threshold with p_D =0.95	Calculated threshold with p_D =0.9	Current tolerance of the United Kingdom
5	0.00%	0.05%	0.40%	0.00%
15	0.07%	0.62%	1.34%	0.04%
16	0.17%	0.79%	1.59%	0.00%
21	0.01%	0.18%	1.34%	0.00%
22	0.05%	0.41%	0.96%	0.03%
28	0.04%	0.54%	1.17%	0.00%
29	0.15%	0.15%	0.99%	0.15%
34	0.03%	0.40%	1.05%	0.07%
37	0.02%	0.23%	0.57%	0.00%
38	0.04%	0.58%	1.17%	0.05%
42	0.04%	0.56%	0.56%	0.04%
46	0.03%	0.33%	0.82%	0.00%
57	0.00%	0.23%	0.72%	0.08%

Table 4. Proportion of times in the United Kingdom pea data set that the first-year thresholds indicated a candidate variety would be non-distinct from another variety when the second-year decision was distinct in the field pea data set (false negative)

Characteristic	Calculated threshold with p_D=0.99	Calculated threshold with p_D =0.95	Calculated threshold with p_D =0.9	Current tolerance of the United Kingdom
5	82.8%	39.4%	24.5%	67.6%
15	86.8%	66.6%	54.3%	90.1%
16	76.8%	56.6%	46.2%	94.6%
21	81.0%	60.8%	34.6%	88.8%
22	88.5%	67.5%	54.2%	91.3%
28	89.3%	65.4%	51.7%	99.9%
29	57.3%	57.3%	34.4%	57.4%
34	84.7%	59.4%	46.2%	75.6%
37	81.2%	65.1%	54.4%	96.5%
38	77.1%	57.1%	47.4%	76.2%
42	80.9%	58.2%	58.2%	81.1%
46	66.7%	44.1%	34.2%	97.5%
57	58.5%	34.6%	25.2%	43.7%

CONCLUSIONS AND FUTURE WORK

22. The method proposed in TWC/33/20 Rev. has been applied to the United Kingdom field pea data set and its performance evaluated. The results show how different risks may be balanced to select an appropriate value of pD for calculating thresholds. Based on these results, the United Kingdom has now updated the first-year tolerances for pea so the tolerances now have a more transparent basis than previously. In this report we have looked at the effectiveness of the method on a characteristic-by-characteristic basis. In future we intend to examine the effect on individual variety decisions.

23. 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.

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

ACKNOWLEDGEMENTS

25. This work has been funded by the Scottish Government.

Adrian Roberts (Biomathematics and Statistics Scotland (BioSS)), Ian Nevison (BioSS) & Tom Christie (SASA)

[End of document]