E

TWF/46/16
ORIGINAL: English
DATE: July 22, 2015

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

Geneva

TECHNICAL WORKING PARTY FOR FRUIT CROPS

Forty-Sixth Session
Mpumalanga, South Africa, August 24 to 28, 2015

REVISION OF DOCUMENT TGP/8: PART II: SELECTED TECHNIQUES USED IN DUS EXAMINATION, SECTION 9: THE COMBINED-OVER-YEARS UNIFORMITY CRITERION (COYU)

Document prepared by the Office of the Union

Disclaimer: this document does not represent UPOV policies or guidance

EXECUTIVE SUMMARY

1. The purpose of this document is to report on developments concerning the method of calculation of COYU.

2. The TWF is invited to note:

(a) that participants of the exercise to test the software on the new method for the calculation of COYU should;

(i) seek to define probability levels to match decisions using the previous COYU method;

(ii) run the test for rejection probabilities of 1, 2 and 5% levels; and

(iii) assess whether the results are consistent in all crops;

(b) that the expert from the United Kingdom distributed the software module for calculation of COYU and the guidance document to the participants of the exercise;

(c) that the experts from Czech Republic, France, Finland, Germany, Kenya, Poland and United Kingdom will participate in the exercise to test the new software on COYU; and

(d) that a report on the practical exercise and the development of DUST module will be presented at the thirty-third session of the TWC.

3. The structure of this document is as follows:

EXECUTIVE SUMMARY 1
BACKGROUND 2
DEVELOPMENTS IN 2014 2
TECHNICAL WORKING PARTIES 2
DEVELOPMENTS IN 2015 3
TECHNICAL COMMITTEE 3

4. The following abbreviations are used in this document:

TC:	Technical Committee
TC-EDC:	Enlarged Editorial Committee
TWA:	Technical Working Party for Agricultural Crops
TWC:	Technical Working Party on Automation and Computer Programs
TWF:	Technical Working Party for Fruit Crops
TWO:	Technical Working Party for Ornamental Plants and Forest Trees
TWPs:	Technical Working Parties
TWV:	Technical Working Party for Vegetables

BACKGROUND

5. The background to this matter is provided in document TWA/43/16 “Revision of document TGP/8: Part II: Selected Techniques Used in DUS Examination, Section 9: The Combined-Over-Years Uniformity Criterion (COYU)”.

DEVELOPMENTS IN 2014

Technical Working Parties

6. At their sessions in 2014, the TWO, TWF, TWC, TWV and TWA considered documents TWO/47/16, TWF/45/16, TWC/32/16, TWC/32/16 Add., TWV/48/16 and TWA/43/16 “Revision of document TGP/8: Part II: Selected Techniques Used in DUS Examination, Section 9: The Combined-Over-Years Uniformity Criterion (COYU)”, respectively.

7. The TWO, TWF, TWV and TWA noted the developments concerning the method of calculation of COYU, including the development of a demonstration module in DUST and the practical exercise that would be conducted using real data to compare decisions made using the current and the proposed improved method (see documents TWO/47/28 “Report”, paragraph 42, TWF/45/32 “Report”, paragraph 33, TWV/48/43 “Report”, paragraph 38 and TWA/43/27 “Report”, paragraph 35).

8. The TWC received a presentation by an expert from the United Kingdom on the method for improving the calculation of COYU, including a demonstration version of a module for the DUST software in the Annex for document TWC/32/16 Add. “Addendum to Development of the Combined-Over-Year Uniformity Criterion” (see document TWC/32/28 “Report”, paragraph 13).

9. The TWC agreed to request the experts from China, Czech Republic, Finland, France, Germany, Kenya, Netherlands and Poland to test the new software on COYU (see document TWC/32/28, paragraph 14).

10. The TWC also agreed to invite other users of the COYU method to test the new software and agreed that an invitation should be developed by the Leading Expert and sent by the Office of the Union to the users of the DUST software package (see document TWC/32/28, paragraph 15).

11. The TWC agreed that the software module for calculation of COYU developed using the “R” software should be sent to the interested experts that use other systems than DUST (e.g. SAS and GenStat) for testing of the new method (see document TWC/32/28, paragraph 16).

12. The TWC agreed that participants should seek to define probability levels to match decisions using the previous COYU method for continuity in decisions and that the test should be run for rejection probabilities of 1, 2 and 5% levels. The TWC agreed that participants should assess whether the results were consistent in all crops (see document TWC/32/28, paragraph 17).

13. The TWC agreed with the timetable for the development of the new software package for the COYU method, as follows (see document TWC/32/28, paragraph 18):

- By the end of July 2014, the UPOV Office with assistance from the expert of the United Kingdom would invite participants for the practical exercise.

- By the end of September 2014, the expert of the United Kingdom would develop further the DUST module demonstrated at the thirty-second session of the TWC for evaluation by the participants and would prepare code for "R" software for participants that prefer this option to the DUST module.
- By early October 2014, the expert of the United Kingdom would send details of the practical exercise, including access to software, to the participants.
- By March 15, 2015, participants of the practical exercise should send a report on their experiences to the expert of the United Kingdom.
- The expert of the United Kingdom would compile a report on the practical exercise and the development of DUST module for the thirty-third session of the TWC.

14. On July 21, 2014, circular E 14/193 "TWC/32: Participation in Practical Exercise (COYU)" was sent to the TC and the TWC members. The Czech Republic, Finland, France, Germany, Kenya, Poland and United Kingdom replied that they wished to participate in the exercise. The expert of the United Kingdom developed software modules for calculation of COYU with a guidance document for the exercise: for participants using the "R" software, the module for "R" software and the guidance document were distributed on October 15, 2014, followed by a revised guidance document on October 21, 2014; and for "DUSTNT" software users, a version with a module for the exercise and guidance document were distributed on December 5.

DEVELOPMENTS IN 2015

Technical Committee

15. The TC, at its fifty-first session, held in Geneva, from March 23 to 25, 2015, noted that participants of the exercise to test the software on the new method for the calculation of COYU should:

(i) seek to define probability levels to match decisions using the previous COYU method;
(ii) run the test for rejection probabilities of 1, 2 and 5% levels; and
(iii) assess whether the results are consistent in all crops (see document TC/51/39 "Report on the Conclusions", paragraph 135)

16. The TC noted that the expert from the United Kingdom had distributed the software module for calculation of COYU and the guidance document to participants of the exercise (see document TC/51/39, paragraph 136).

17. The TC noted that the experts from Czech Republic, France, Finland, Germany, Kenya, Poland and United Kingdom would participate in the exercise to test the new software on COYU (see document TC/51/39, paragraph 137).

18. The TC noted that a report on the practical exercise and the development of DUST module would be presented at the thirty-third session of the TWC, to be held in Natal, Brazil, from June 30 to July 3, 2015 (see document TC/51/39, paragraph 138).

19. The TWF is invited to note that:

(a) participants of the exercise to test the software on the new method for the calculation of COYU should:

(i) seek to define probability levels to match decisions using the previous COYU method;

(ii) run the test for rejection probabilities of 1, 2 and 5% levels; and

(iii) assess whether the results are consistent in all crops;

(b) the expert from the United Kingdom distributed the software module for calculation of COYU and the guidance document to the participants of the exercise;

(c) the experts from Czech Republic, France, Finland, Germany, Kenya, Poland and United Kingdom will participate in the exercise to test the new software on COYU; and

(d) a report on the practical exercise and the development of DUST module will be presented at the thirty-third session of the TWC by an expert from the United Kingdom.

[Annex follows]

ANNEX

METHOD OF CALCULATION OF COYU: DEVELOPMENT OF SOFTWARE & PRACTICAL EXERCISE

Document prepared by Experts from the United Kingdom

Background

1. At its thirty-second session, the TWC received a presentation by an expert from the United Kingdom on the method for improving the calculation of COYU, including a demonstration version of a module for the DUST software.

2. The TWC agreed to invite users of the COYU method to test the new method and software. The R software module for calculation of the proposed improved COYU method would be sent to the interested experts that use other systems than DUST.

3. The TWC agreed that participants should seek to define probability levels to match decisions using the previous COYU method for continuity in decisions and that the test should be run for rejection probabilities of 1, 2 and 5% levels. The TWC agreed that participants should assess whether the results are consistent in all crops.

4. The TWC agreed the following timetable for the development of the new software package for the COYU method as follows:

- By the end of July 2014, the UPOV Office with assistance from the expert of the United Kingdom would invite participants for the practical exercise.
- By the end of September 2014, the expert of the United Kingdom would develop further the DUST module demonstrated at the thirty-second session of the TWC for evaluation by the participants and would prepare code for “R” software for participants that prefer this option to the DUST module.
- By early October 2014, the expert of the United Kingdom would send details of the practical exercise, including access to software, to the participants.
- By March 15, 2015, participants of the practical exercise should send a report on their experiences to the expert of the United Kingdom.
- The expert of the United Kingdom would compile a report on the practical exercise and the development of DUST module for the thirty-third session of the TWC.

5. This document reports on progress since the last session of the TWC, including the practical exercise and the further development of software.

Software Development

6. Following last year’s TWC session, the software was improved to make it suitable for release to the Practical Exercise participants for evaluation. The R version was packaged as an R library/package. The DUST version took longer to complete as it essentially calls the R version, requiring refinement of the interface and the development of a new installation process.

7. The Practical Exercise highlighted areas for further improvement, especially in the installation process for the DUST version. These areas are detailed in Appendix B and we plan to address them following this TWC session. In addition the source code for the R package can now be downloaded.

Practical Exercise

8. In July 2014, an invitation to take part in the COYU Practical Exercise was set to TWC members. Software and a document giving guidance on the software and instructions for the Exercise were sent to those expressing an interest in participation. The R software was made available in October 2015 and the DUST software in December 2015. Participants were asked to evaluate the software and to compare the results obtained by the current version of COYU with those produced by the proposed improved version.

9. The following took part in the exercise:

Country	Participant	Software	Crops
Finland	Sami Markannen	DUST	Timothy, meadow fescue, tall fescue, Canarian reed grass, red clover, white clover, turnip rape
France	Christophe Chevalier	R	Fescue
Kenya	Abraham Lagat	R	Wheat
United Kingdom	Sally Watson	DUST	Perennial ryegrass
United Kingdom	Haidee Philpott	DUST	Oilseed rape
United Kingdom	Tom Christie	DUST	Field pea

10. In addition, an expert from Germany informed that they currently use SAS for COYD and COYU, though they are likely to move to R in due course. They are developing software in SAS for COYU using splines and then would be interested to compare results.

11. Impressions of the R software were positive, with no problems noted. However problems with software installation delayed the start of the exercise for all participants using DUST. The main cause of these difficulties was related to installing software on secure government networks. We have identified how this can be improved for the next release (Appendix B). During the exercise some issues related to sorting results and dealing properly with missing data were identified and corrected.

12. A review of the results obtained by the participants is presented in Appendix A. This indicated that a higher probability level for the proposed new COYU method would be needed to match decisions with the current method as closely as possible. More data sets would be required to establish more exactly the probability level required. This review also highlights the need to discuss what action should be taken when the candidate has a level of expression outside that seen the reference varieties. Whilst extrapolation cases are not infrequent, cases where the candidate's level of expression is more than 10% outside the range of are not common.

Conclusions and Future Work

13. The new method works well in practice. The fit of the spline adjustments seem to be fit for purpose.

14. Higher probability levels are likely to be required than for the current method; the example data sets indicate that the probability level required to match the 0.001 level for the current method might be in the range 0.005 to 0.026. More data sets are required to examine this more fully.

15. The Practical Exercise highlighted the need to discuss what action should be taken when the candidate has a level of expression outside that seen the reference varieties. We suggest that cases of minor extrapolation can safely be ignored, but cases of major extrapolation should be considered by the crop expert.

16. The software worked well though some areas for improvement and development have been noted. These will be pursued during 2015-6.

Acknowledgements

17. We are grateful for the essential input of Sally Watson and her AFBI colleagues. We are also very grateful to all those contributing to the Practical Exercise and for their useful suggestions for improvements.

Adrian Roberts and David Nutter
Biomathematics & Statistics Scotland

Appendix A: Results of Practical Exercise

Data sets

1. The table below summarises the data sets considered in the practical exercise:

Country	Crop	Number of data sets	Number of years for each data set	Probability level for COYU	Number of characters	Number of candidates	Number of reference varieties
Finland	Timothy	1	2	0.001	6	3	~60
Finland	Timothy	2	3	0.001	1-7	2-4	63
Finland	Meadow fescue	1	2	0.001	6	2	35
Finland	Meadow fescue	1	3	0.001	6	2	35
Finland	Tall fescue	1	2	0.001	6	1	20
Finland	Canarian reed grass	1	3	0.001	8	1	10
Finland	Red clover	2	2	0.001	6	1	27
Finland	Red clover	2	3	0.001	7	1	15
Finland	White clover	1	2	0.001	9	1	22
Finland	White clover	1	3	0.001	9	1	23
Finland	Turnip rape	1	2	0.001	8	3	13
Finland	Turnip rape	1	3	0.001	8	1	13
France	Fescue	1	2	0.001	11	4	126
Kenya	Wheat	1	2	?	3	2	?
UK	Perennial ryegrass	2	3	0.001	30	30/16	102/74(cyclic)
UK	Oilseed rape	4 (by type)	2	0.001	12	64/15/1/48	444/136/272/217
UK	Field pea (conventional)	5	2	0.001	17-19	6/5/3/1/1	18/39/31/10/21
UK	Field pea (semi-leafless)	5	2	0.001	13-14	4/14/7/2/4	49/67/77/61/72

Fit of splines to data

2. The principal change in the improved COYU method is the use of splines instead of the moving average method to map the relationship between uniformity and level of expression. This method also restricts the flexibility of the spline (see TWC/31/15 corr.). To assess the success of this, participants were asked to review the plots output by the new software. A number of these were also reviewed by the author. It was seen that the curves fitted the data adequately, without any tendency to over-fit. An example plot is shown in Figure A1.

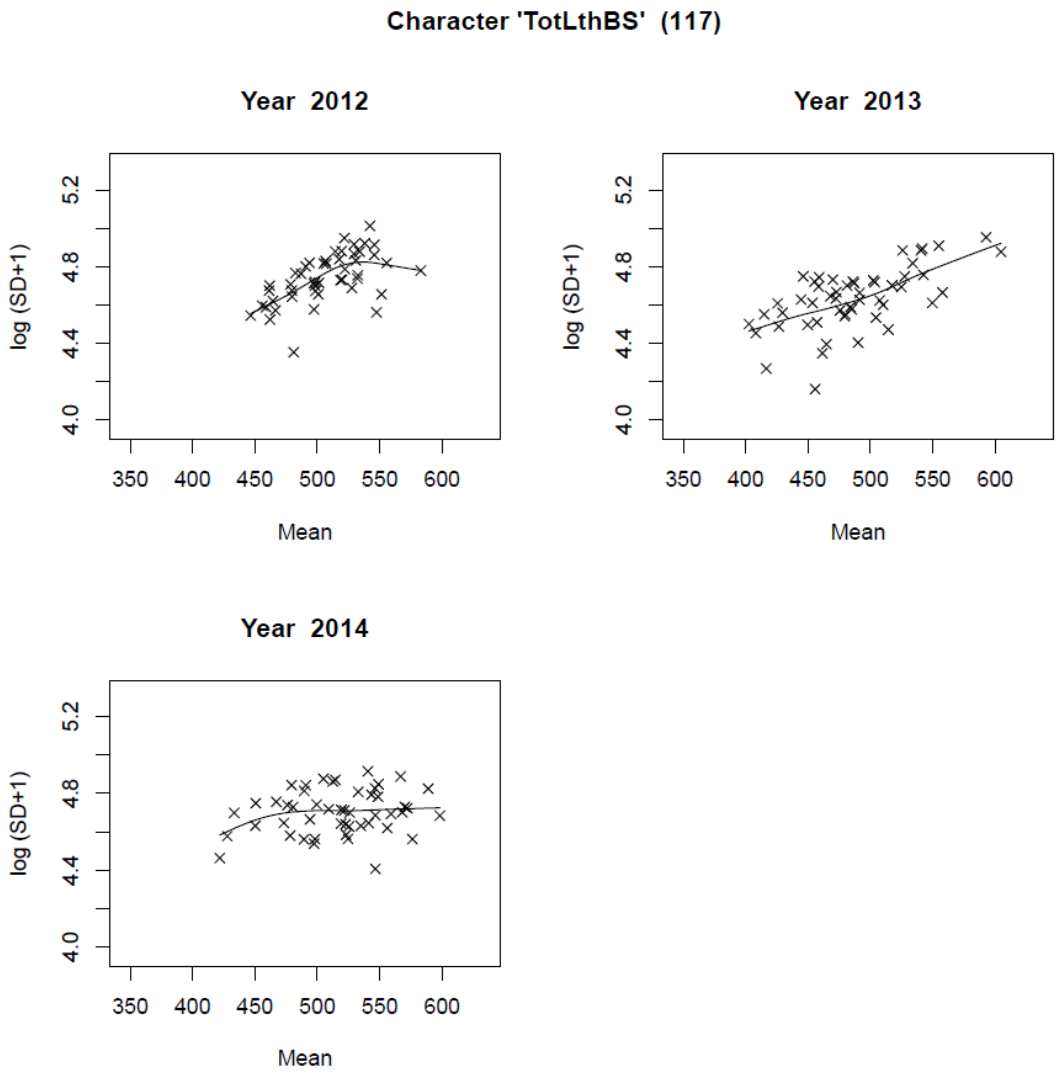


Figure A1: *Spline fits of uniformity to level of expression for characteristic 117 Perennial Ryegrass Tetraploid Intermediates – data from UK*

Matching probability levels between the two COYU methods

3. All participants use the probability level of 0.001 with the current COYU method. The new method is expected to require the use of a higher probability level to achieve the same level of decision-making (TWC/31/15 corr.). In principle, it might be possible to assess what probability level might be required by looking at how many varieties would be found non-uniform with differing levels of probability. In practice, few varieties actually fail the COYU criterion; this means that a comparison of rejection rates would only give a very coarse idea of the probability level required for the new method. Instead, it is better to calculate and compare p-values for each candidate.

4. In the figures below, we compare the p-values for the current and new methods of COYU. The plots on the left hand side show all the candidates – the plots on the right hand side show only those results where the p-value for the current method is less than 0.01. Curves are fitted onto these graphs to give some idea of trend. For Finland, it was necessary to produce graphs covering several species since the number of candidates was low. Candidate varieties that show extrapolation were omitted (see below).

Figure A2: *Comparison of p-values for Finland data sets; solid line is fitted curve, dashed line shows equality between the p-values*

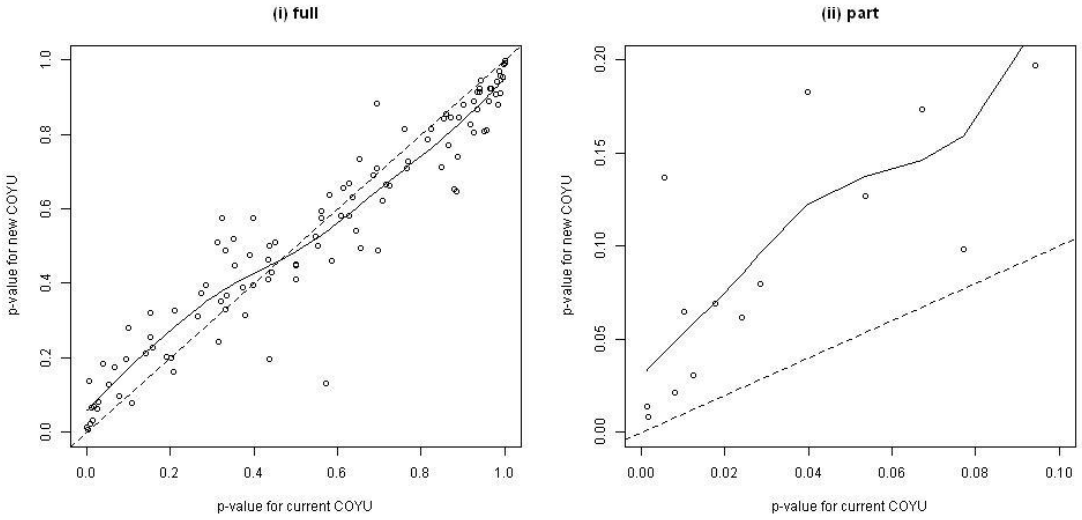


Figure A3: *Comparison of p-values for France data set; solid line is fitted curve, dashed line shows equality between the p-values*

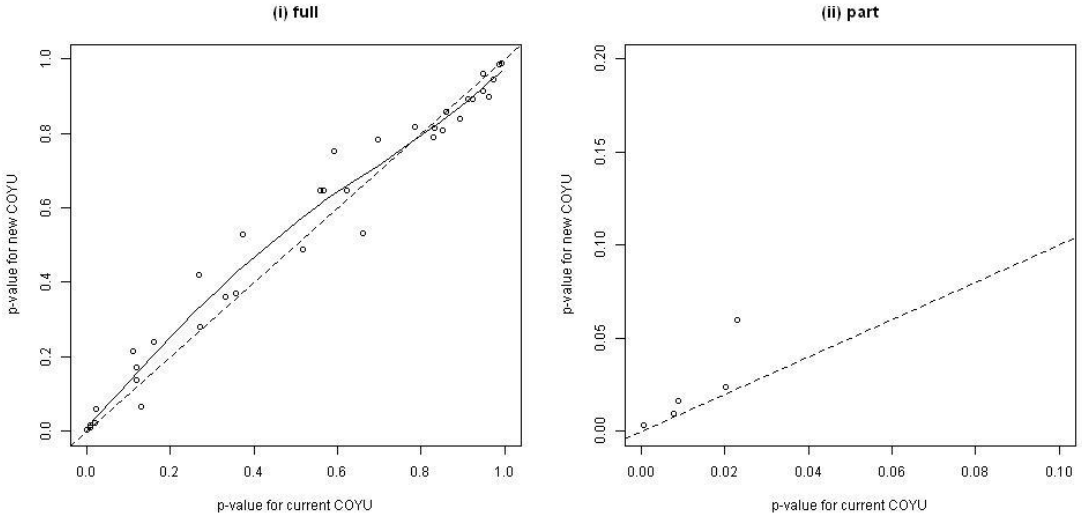


Figure A4: *Comparison of p-values for Kenya data set; solid line is fitted curve, dashed line shows equality between the p-values*

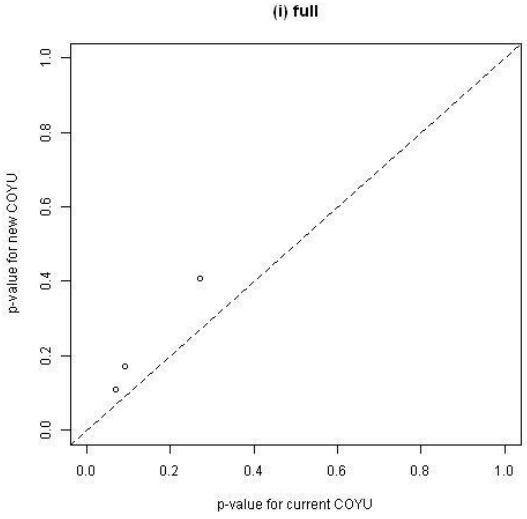


Figure A5: *Comparison of p-values for UK perennial ryegrass amenity data set; solid line is fitted curve, dashed line shows equality between the p-values*

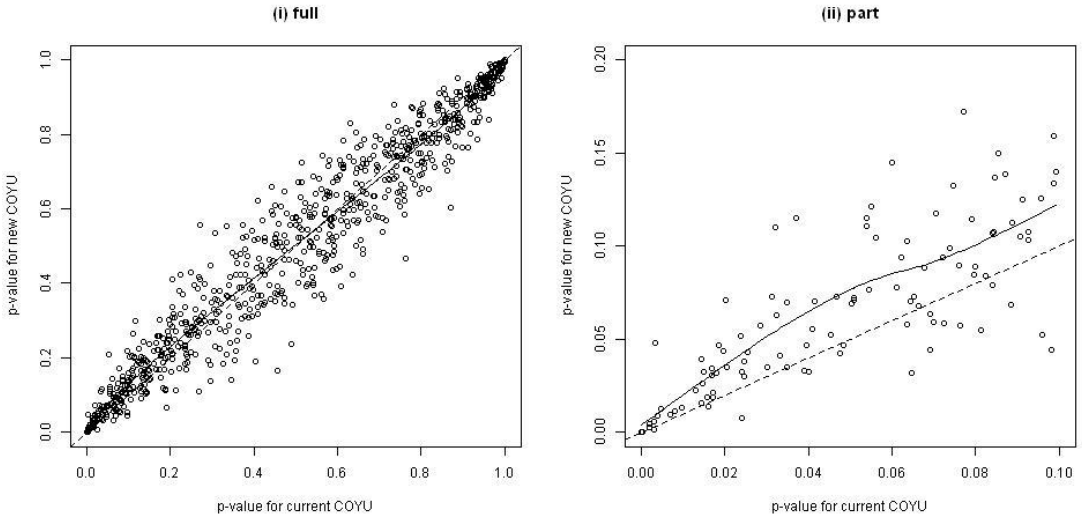


Figure A6: *Comparison of p-values for UK perennial ryegrass tetraploid data set; solid line is fitted curve, dashed line shows equality between the p-values*

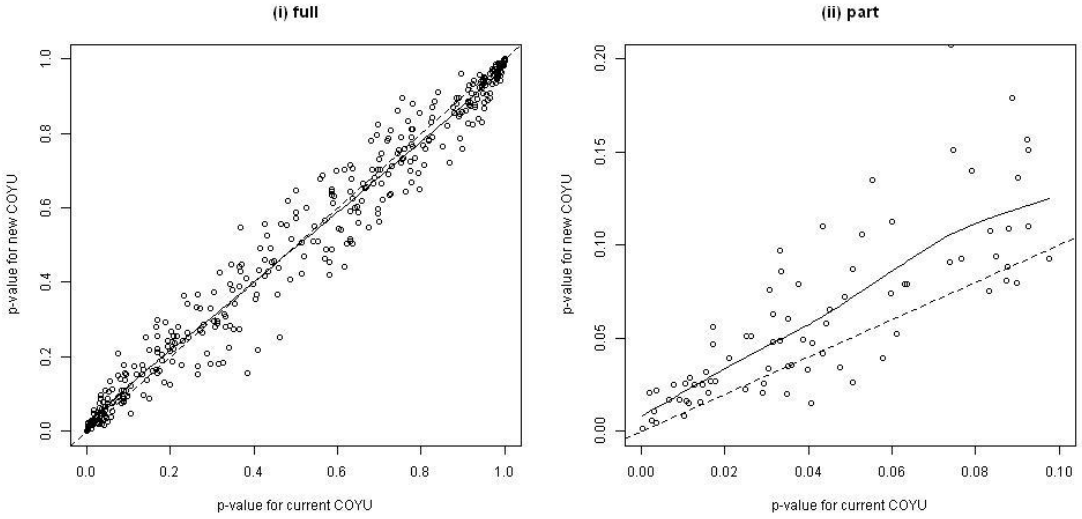


Figure A7: *Comparison of p-values for UK oilseed rape lines data set; solid line is fitted curve, dashed line shows equality between the p-values*

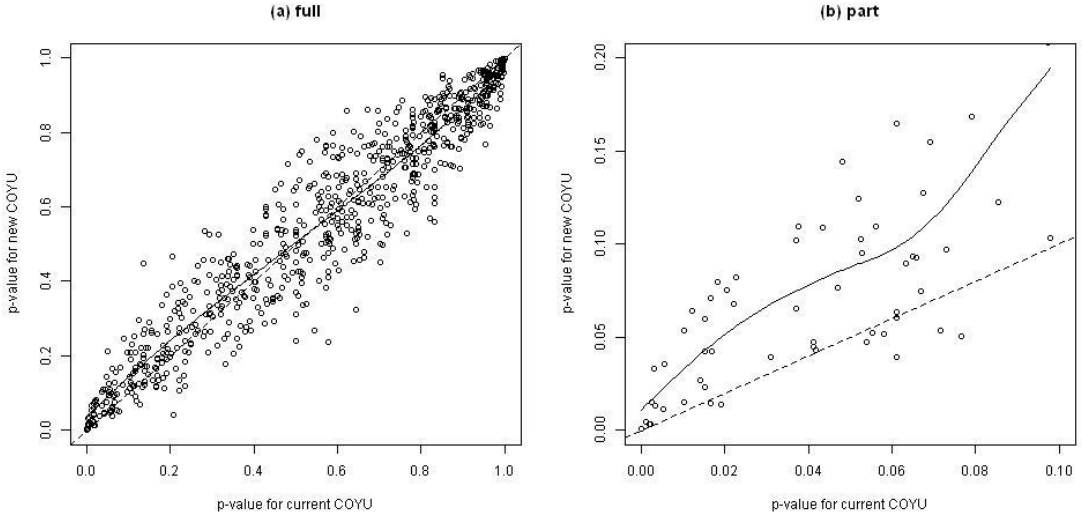


Figure A8: *Comparison of p-values for UK oilseed rape restored hybrids data set; solid line is fitted curve, dashed line shows equality between the p-values*

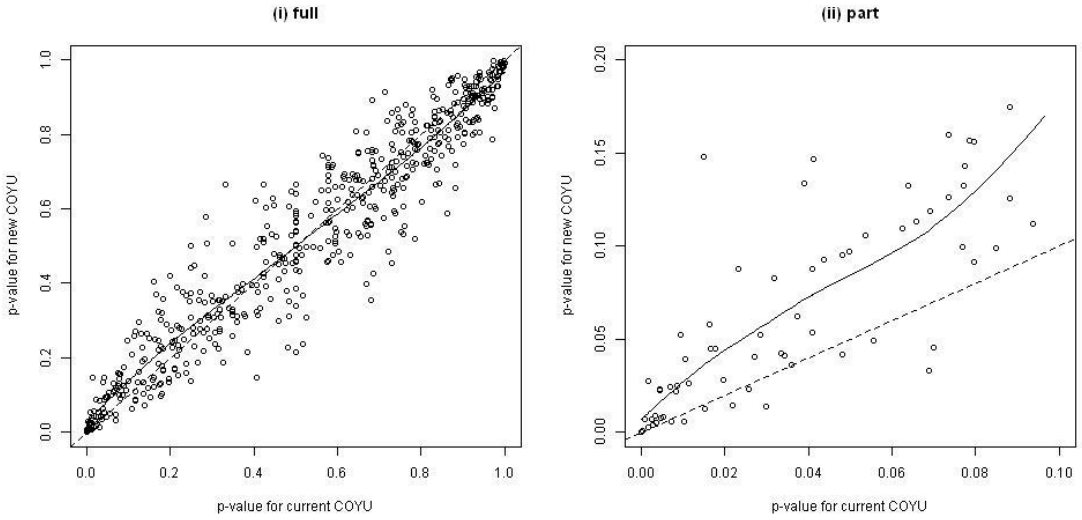


Figure A9: *Comparison of p-values for UK oilseed rape hybrids data set; solid line is fitted curve, dashed line shows equality between the p-values*

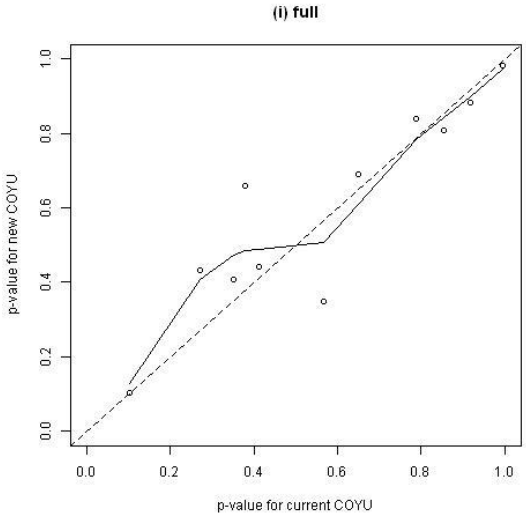


Figure A10: *Comparison of p-values for UK oilseed rape composite data set; solid line is fitted curve, dashed line shows equality between the p-values*

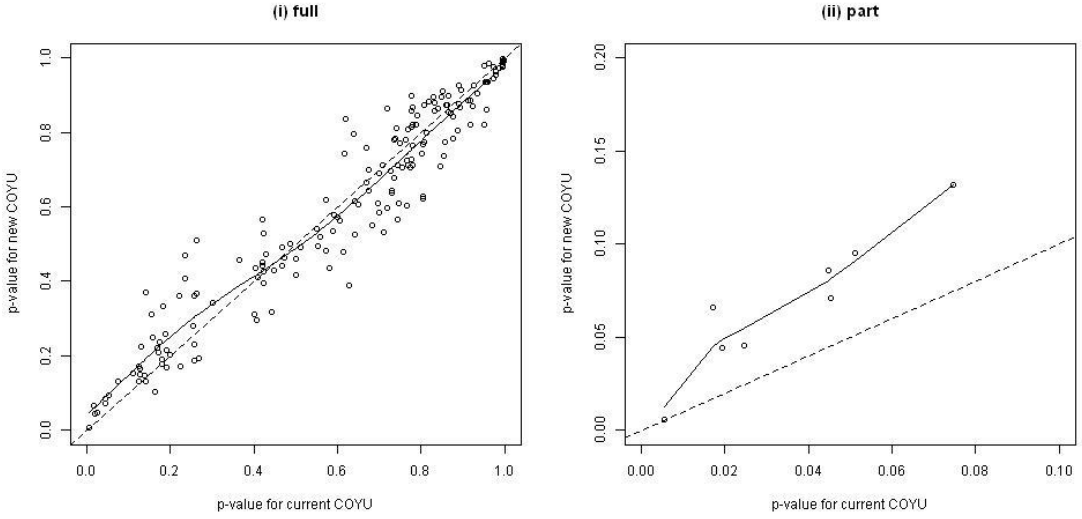


Figure A11: *Comparison of p-values for UK field pea conventional-type data sets; solid line is fitted curve, dashed line shows equality between the p-values*

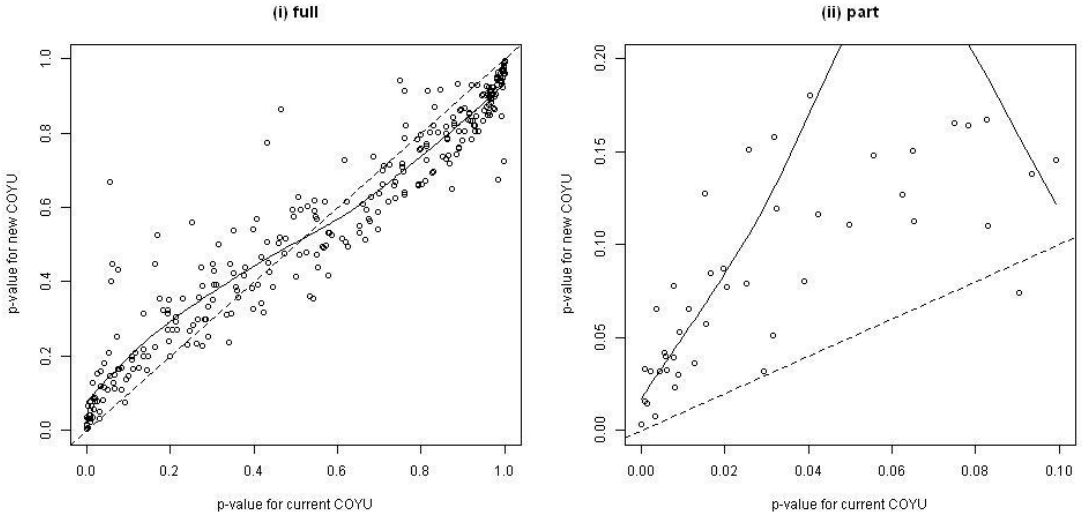
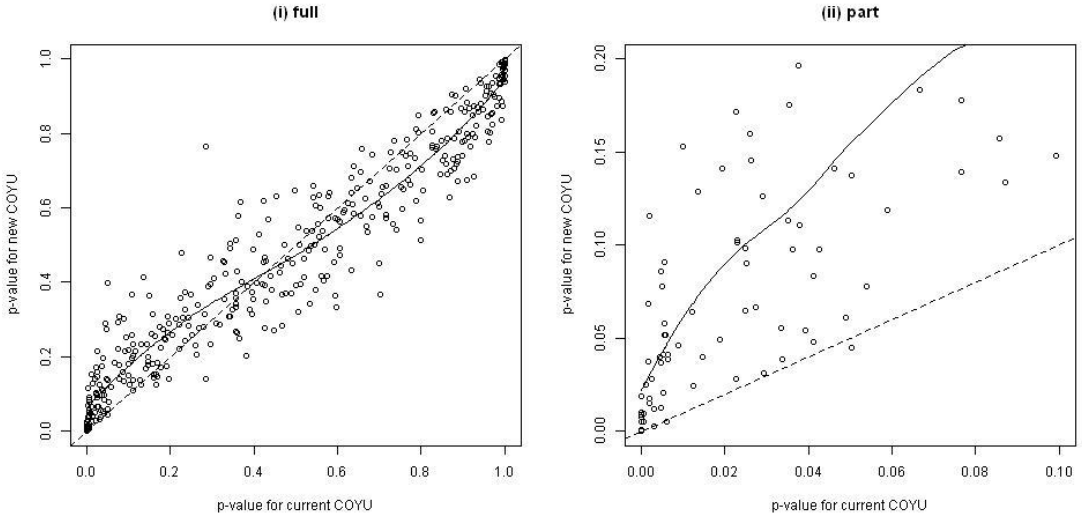


Figure A12: *Comparison of p-values for UK field pea semi-leafless-type data sets; solid line is fitted curve, dashed line shows equality between the p-values*



5. The general pattern was similar across the different data sets, with a high degree of correlation between the two methods but the new method having higher p-values for less uniform varieties (and lower for more uniform varieties). The table below gives a spline-based prediction of the (approximate) equivalent p-value for the new COYU compared to a p-value of 0.001 for the current method (for those data sets with at least 20 observations with current method p-value <0.1). This demonstrates the need to use a larger p-value for the proposed method. The degree of increase would need to be evaluated through analysis of more data sets.

Country	Data set	Approximate equivalent p-value
UK	PRG amenity	0.005
UK	PRG tetraploid	0.009
UK	OSR lines	0.013
UK	OSR restored hybrids	0.009
UK	Field pea conventional	0.020
UK	Field pea semi-leafless	0.026

Extrapolation

6. Both the current and proposed new methods for COYU uses adjustments based on fitting a curve to the relationship between uniformity (represented by the log of the standard deviation plus 1) and the level of expression (represented by the mean) over the reference varieties. This curve is used to adjust uniformity data for both the reference varieties and candidates. As noted in TWC/31/15 corr., there is an issue if a candidate exhibits a level of expression outside the range seen in the reference varieties; this is extrapolation. This issue needs careful consideration and it was an aim of this Practical Exercise to evaluate the frequency of extrapolation cases in practice.

7. The effect of extrapolation is different for the two versions of COYU, current and new. Overall the proposed method is more likely to indicate such a candidate variety as uniform – it gives the benefit of doubt. However perhaps it is better that such cases are evaluated apart and with care by the crop experts. To this end, the new software does indicate cases of extrapolation. A future version may indicate the extent of extrapolation.

8. The table below indicates the frequency of extrapolation cases and, in some cases, the extent of extrapolation in relation to the range of expression in the reference varieties. Note that some example data sets had few varieties (see Table A1) so these provide only a rough indication of more general levels of extrapolation.

Country	Data set	Frequency of extrapolation	Cases > 10% extrapolation	Cases > 20% extrapolation
France	Fescue	0%	n/a	n/a
Kenya	Wheat	50%	n/a	n/a
Finland	Various	19%	n/a	n/a
UK	PRG amenity	9%	2%	<1%
UK	PRG tetraploid	20%	13%	9%
UK	OSR lines	<1%	<1%	0%
UK	OSR restored hybrids	2%	<1%	0%
UK	OSR hybrids	8%	<1%	0%
UK	OSR composite	3%	0%	0%
UK	Field pea conventional	10%	5%	3%
UK	Field pea semi-leafless	7%	3%	1%

9. The large number of extrapolation cases for the UK perennial ryegrass (tetraploid) data set was investigated further. Much of this was due to a single candidate, which was very different to the reference varieties. Most of the remaining large extrapolations in this data set were due to two more candidates in one character.

Appendix B: Planned developments for the new COYU software

Improvements to functionality

- Add candidates to figures showing fit of splines
- Output more information on extent of extrapolation for candidates
- Add capability to analyse candidates with missing years, at least if two years are present
- Deal appropriately with characters with no variation

R-code specific developments

- The code is now freely available on GitHub. This includes the source code. The address is https://github.com/davidnutter/coyus

DUST specific developments

- Improve installation of DUST version. It will be made more robust by, e.g. inclusion of all relevant R packages in install so no need to download from CRAN. It will also consider automating the setting of the R path environmental variable and including DLL/OCX files.
- Review verdict labels on output
- Improve clarity and delivery of errors

[End of document]