TWF/45/9
ORIGINAL: English
DATE: May 19, 2014

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

Geneva

TECHNICAL WORKING PARTY FOR FRUIT CROPS

Forty-Fifth Session
Marrakesh, Morocco, May 26 to 30, 2014

ASSESSING UNIFORMITY BY OFF-TYPES ON BASIS OF MORE THAN ONE SAMPLE OR SUB-SAMPLES

Disclaimer: this document does not represent UPOV policies or guidance

1. The purpose of this document is to report on developments concerning a summary of approaches used by members of the Union in assessing uniformity by off-types on the basis of more than one sample or sub-sample.

2. 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

3. The structure of this document is as follows:

BACKGROUND

4. The complete background to this matter is provided in documents TC/48/14 “Assessing Uniformity by Off-Types on the Basis of More than One Sample or Sub-Sample” and TC/50/12 “Assessing Uniformity by Off-Types on the Basis of More than One Sample or Sub-Sample”.

DEVELOPMENTS IN THE TECHNICAL WORKING PARTIES IN 2013

Technical Working Party for Ornamental Plants and Forest Trees
Technical Working Party for Fruit Crops

5. The TWO and the TWF noted the information in documents TWO/46/22 Rev. and TWF/44/22, respectively (see documents TWO/46/29 “Report”, paragraph 63, and TWF/44/31 “Report”, paragraph 61).

Technical Working Party for Vegetables

6. The TWV considered document TWV/47/22 (see documents TWV/47/34 “Report”, paragraphs 62 to 64).

7. The TWV noted that with regard to the situation B, as set out in Annex II of document TWV/47/22, the expert from France indicated that France was now considering each cycle to be independent and was no longer combining the results of two locations, therefore the reference to France and cauliflower was no longer appropriate and that the text should be amended accordingly.

8. The TWV agreed that the preferred approach, for the assessment of uniformity for vegetables, was to use the individual results rather than the combined results and requested the TWC to consider the following example when considering the different approaches:

Trial 1: 50 plants / 2 off-types → the variety is considered uniform
(based on a population standard of 1% and an acceptance probability of 95% from a sample of 36 to 82 plants)

Trial 2: 50 plants / 2 off-types → the variety is considered uniform
(based on a population standard of 1% and an acceptance probability of 95% from a sample of 36 to 82 plants)

Trial 1+2: 100 plants/ 4 off-types → the variety is considered non-uniform
(based on a population standard of 1% and an acceptance probability of 95% from a sample of 83 to 130 plants)

If the two trials are considered independent, the variety is considered to be uniform. If the two trials are combined, the variety is considered not uniform.

Technical Working Party on Automation and Computer Programs

9. The TWC considered document TWC/31/22.

10. The TWC noted that it had agreed that more detailed information and further analysis were needed in order to give guidance on consequences on the use of the different approaches presented in Annexes I to IV of document TWC/31/22, and that France, Germany and the Netherlands would present one or more concrete situations in their countries and the statistical basis of their analysis for its next session.

11. The TWC considered the statistical basis for the acceptable number of off-types in the subsample of 20 plants used in the context of a sample size of 100 plants, as provided in Annex V to document TWC/31/22, which was introduced by an expert from Germany by electronic means.

12. The TWC requested the experts from Germany to explain the Type I and Type II errors used in the statistical basis for the acceptable number of off-types in the subsample of 20 plants used in the context of a sample size of 100 plants. The TWC agreed that the statistical basis for this approach would continue to be discussed at its thirty-second session. The TWC agreed that it would not be appropriate to recommend this approach for other crops or sample sizes before it had agreed on the statistical basis.

13. The TWC noted that, with regard to the approach combining the results from two growing cycles, as set out in Annexes I and II to document TWC/31/22, Situation A and B, the TC had agreed that care would be needed when considering results that were very different in each of the growing cycles, such as when a type of off-type was observed at a high level in one growing cycle and was absent in another growing cycle.

14. The TWC noted information on testing of uniformity of Apple varieties arising from mutation in New Zealand would be presented in document TWC/31/26 (see document TWC/31/32 "Report", paragraphs 82 to 87).

Technical Working Party for Agricultural Crops

15. The TWA considered document TWA/42/22.

16. The TWA requested the Office of the Union to further develop the Annexes to document TWA/42/22 to be presented at its forty-third session with regard to providing the information requested by the TWC for the analysis of consequences of different approaches. The TWA agreed that the experts from France, Germany, the Netherlands and the United Kingdom, should assist the Office in the preparation of the document.

17. The TWA agreed that the different situations should be presented in one example without mention to individual members of the Union. The TWA also agreed that it should be clarified if the two approaches in situation A were inconsistent, or if one of them was more appropriate, and that situation C should refer to the use of samples or subsamples instead of "tests/trials" (see document TWA/42/31 "Report", paragraphs 68 to 70).

DEVELOPMENTS IN 2014

Technical Committee

18. The TC considered document TC/50/12.

19. The TC noted that the TWC had proposed to provide more detailed information and further analysis on the consequences of the use of the approaches presented in situations A, B, C and D at its session in 2014, as set out in document TC/50/12, paragraph 33. The TC agreed that the document should be modified in order to explain that, in Situations A and B, a variety might be rejected after a single growing cycle under certain circumstances.

20. The TC noted that the TWC had proposed that the statistical basis for the acceptable number of off-types in situation D be considered further at its session in 2014, as set out in document TC/50/12, paragraphs 34 and 35.

21. The TC agreed to develop guidance in document TGP/10 for situations A, B, C and D, as set out in document TC/50/12, Annexes I to IV, after consideration of the comments by the TWC.

22. The TC agreed that document TGP/10, paragraph 6 should be considered when considering the possible development of guidance on the approach presented in situation C, as set out in document TC/50/12, Annex III (see document TC/50/36 "Report on the Conclusions", paragraphs 123 to 127).

SUMMARY OF APPROACHES

23. Annexes I to IV to this document summarize different situations when different samples are combined for the overall assessment of uniformity of a variety in accordance with the conclusions of the TC at its fiftieth session, as follows:

Annex I: Situation A: Two growing cycles in a single location

Approach 1: Third growing cycle in the case of inconsistent results
Approach 2: Combining the results of two growing cycles

Annex II: Situation B: Two growing locations in the same year

Approach 1: Third growing cycle in the case of inconsistent results
Approach 2: Combining the results of two locations

Annex III: Situation C: More than one sample/sub-sample for a characteristic in the same growing cycle

Approach: Additional growing cycle in the case of inconsistent results

Annex IV: Situation D: Assessing sub-samples within a single test/trial

Approach*:* Sub-sample as a first step of assessment

24. The summary in Annexes I to IV only relates to situations where more than one sample, or sub-sample, concern the examination of the same characteristic. In the case of different samples, or sub-samples (e.g. special test), to examine a different characteristic there is no requirement to combine the results because a variety is required to be uniform for all relevant characteristics.

25. Annex V to this document presents a memorandum provided by the experts from Germany on the use of a stepwise approach in the off-type procedure within the same growing cycle in Situation D.

26. The TWF is invited to consider Annexes I to IV to this document as a basis to develop guidance in document TGP/10.

[Annexes follow]

ANNEX I

SITUATION A: TWO GROWING CYCLES IN A SINGLE LOCATION

Approach 1: Third growing cycle in the case of inconsistent results

A variety is considered uniform if it is within the uniformity standard in both of the two growing cycles.

A variety is considered non-uniform if it fails to meet the uniformity standard in both of the two growing cycles.

If at the end of the two growing cycles the variety is within the uniformity standard in one growing cycle but is not within the uniformity standard in the other growing cycle, then uniformity is assessed in a third growing cycle after consultation with the applicant. If in the third growing cycle the variety is within the uniformity standard, the variety is considered uniform. If in the third growing cycle the variety fails to meet the uniformity standard, the variety is considered non-uniform.

Care is needed when considering results that were very different in each of the growing cycles, such as when a type of off-type was observed at a high level in one growing cycle and was absent in another growing cycle.

A variety may be rejected after a single growing cycle, under certain circumstances.

Approach 2: Combining the results of two growing cycles

A variety is considered uniform if it is within the uniformity standard in both of the two growing cycles.

A variety is considered non-uniform if it fails to meet the uniformity standard in both of the two growing cycles.

If at the end of the two growing cycles the variety is within the uniformity standard in one growing cycle but is not within the uniformity standard in the other growing cycle, a variety is considered uniform if the total number of off-types at the end of the two growing cycles does not exceed the number of allowed off-types for the combined sample.

Care is needed when considering results that were very different in each of the growing cycles, such as when a type of off-type was observed at a high level in one growing cycle and was absent in another growing cycle.

A variety may be rejected after a single growing cycle, under certain circumstances.

Example:

Population Standard = 1%
Acceptance Probability ≥ 95%

Sample Size in each of growing cycles 1 and 2 = 50 Maximum number of Off-Types = 2
Sample Size in growing cycles 1 and 2 combined = 100 Maximum number of Off-Types = 3

	Growing cycle		Decision	
	First	Second	Approach 1	Approach 2
Number of Off-Types	2	2	uniform	uniform
	0	3	third growing cycle	uniform
	1	3	third growing cycle	non-uniform
	0	10*	third growing cycle*	non-uniform*
	10**	0	third growing cycle**	non-uniform**

* Care is needed when considering results that were very different in each of the growing cycles, such as when a type of off-type was observed at a high level in one growing cycle and was absent in another growing cycle.

** A variety may be rejected after a single growing cycle, under certain circumstances.

[Annex II follows]

ANNEX II

SITUATION B: TWO GROWING LOCATIONS IN THE SAME YEAR

Approach 1: Third growing cycle for inconsistent results

A variety is considered uniform if it is within the uniformity standard in both of the growing locations.

A variety is considered non-uniform if it fails to meet the uniformity standard in both of the growing locations.

If the variety is within the uniformity standard in one growing location but is not within the uniformity standard in the other growing location, then

Alternative (a) the trial is repeated at both locations for a second year;

Alternative (b) the trial is repeated at the Leading station (location)

Care is needed when considering results that were very different in each of the growing cycles, such as when a type of off-type was observed at a high level in one growing cycle and was absent in another growing cycle.

A variety may be rejected after a single trial in one growing location, under certain circumstances.

Approach 2: Combining the results of two locations

A variety is considered uniform if it is within the uniformity standard in both locations.

A variety is considered non-uniform if it fails to meet the uniformity standard in both locations.

If the variety is within the uniformity standard in one growing location but is not within the uniformity standard in the other growing location, a variety is considered within the uniformity standard if the number of off-type plants or parts of plants does not exceed the allowed number of off-types for the combined sample (two locations).

Care is needed when considering results that were very different in each of the growing cycles, such as when a type of off-type was observed at a high level in one growing cycle and was absent in another growing cycle.

A variety may be rejected after a single trial in one growing location, under certain circumstances.

Example:

Population Standard = 1%
Acceptance Probability ≥ 95%

Sample Size in each of growing locations 1 and 2 = 50 Maximum number of Off-Types = 2
Sample Size in growing locations 1 and 2 combined = 100 Maximum number of Off-Types = 3

	Growing location		Decision	
	First	Second	Approach 1	Approach 2
Number of Off-Types	2	2	uniform	uniform
	0	3	repeat trial	uniform
	1	3	repeat trial	non-uniform
	0	10*	repeat trial*	non-uniform*
	10**	0	repeat trial**	non-uniform**

* Care is needed when considering results that were very different in each of the growing cycles, such as when a type of off-type was observed at a high level in one growing cycle and was absent in another growing cycle.

** A variety may be rejected after a single growing cycle, under certain circumstances.

[Annex III follows]

ANNEX III

SITUATION C: MORE THAN ONE SAMPLE OR SUBSAMPLE FOR A CHARACTERISTIC IN THE SAME GROWING CYCLE

Approach: Additional growing cycle in the case of inconsistent results

A variety is considered to be uniform for a characteristic if it is within the uniformity standard for the characteristic in all samples or subsamples.

A variety is considered non-uniform if it fails to meet the uniformity standard for the characteristic in all samples or subsamples.

In the case where a variety is within the uniformity standard for the characteristic in one sample or subsample (e.g. main trial) and not in another sample or subsample (e.g. ear-row plot), both samples or subsamples are examined in a further growing cycle.

Note: to consider whether Situation C is already covered by document TGP/10 "Examining Uniformity", Section 6 "Combining all observations on a variety".

Extract from document TGP/10 "Examining Uniformity"

"SECTION 6: Combining all observations on a variety

[…]

"6.3 Off-types only: characteristics observed on different samples

"In many cases, uniformity is assessed by observations on different samples of plants or parts of plants. For example, for uniformity in wheat (see UPOV Test Guidelines for Wheat: TG/3), some characteristics are observed on a sample of 2,000 plants, whilst some other characteristics are observed on a sample of 100 parts of plants taken from 100 plants. Off-type plants observed in the plot of 2,000 plants can be excluded from further observations. For the plant parts taken from 100 plants, it is not normally possible to trace back the plant part to the original plant in the plot. Therefore, the sample of 100 plant parts needs to be considered to be independent from the 2,000 plants. Another independent sample of the variety is observed for seed characteristics. In such cases, a uniformity assessment should be carried out on all the independent samples, using the appropriate population standard. A variety should be considered to be uniform if the uniformity requirements are fulfilled in all samples."

[Annex IV follows]

ANNEX IV

SITUATION D: ASSESSING SUB-SAMPLES WITHIN A SINGLE TEST/TRIAL

Approach: Use of sub-sample as a first step of assessment

A variety is considered uniform if no off-types are observed in a sub-sample.

A variety is considered non–uniform if the number of off-types in the sub-sample exceeds the accepted number of off-types for the whole sample.

If the number of off-types is 1 or more, but below the accepted number of off-types for the whole sample, the whole sample is assessed.

Example:

In a sample size of 100 plants, the acceptable number of off-types is 3 (based on a population standard of 1% and an acceptance probability of at least 95%).

In a subsample of 20 plants used in the context of the sample size of 100 plants above:

A variety is considered uniform if no off-types are observed in the sub-sample.

A variety is considered non–uniform if the number of off-types in the sub-sample exceeds 3.

If the number of off-types is 1 to 3, the whole sample of 100 plants is assessed.

(The background to this example is provided in Annex V to this document)

[Annex V follows]

ANNEX V

MEMORANDUM FROM THE EXPERT FROM GERMANY

Use of a stepwise approach in the off-type procedure within the same growing cycle

Background

The TC noted that the TWC had proposed that the statistical basis for the acceptable number of off-types in situation D be considered further at its session in 2014, as set out in document TC/50/12, paragraphs 34 and 35.

The following text is based on document TWC/31/22 Annex V.

The method of uniformity assessment on the basis of off-types (off-types procedure) was described in document TGP/8. Paragraph 8.1.7 provides guidance on the use of the off-types procedure on more than one single test, including a combined test, a two-stage test and sequential tests.

A combined test is described as follows:
Make a decision after two (or three) years based on the total number of plants examined and the total number of off-types recorded.

A two-stage test is described as follows:
Use the result of the first year to see if the data suggests a clear decision (reject or accept). If the decision is not clear then proceed with the second year and decide after the second year.

A sequential test is a multi-stage test where decision rules can be defined dependently or independently on results of the test.

A specific example for the use of a two-step test is provided in document TWC/29/09 “Assessing uniformity by off-types on the basis of more than one sample or sub-sample”, Annex I, page 13. The following specific approach is applied by several European examination offices in wheat and barley for the assessment of uniformity in case of characteristics observed on a sample size of 100 plants or parts of plants. The population standard is fixed at 1% with an acceptance probability of 95% for each decision.

In the first step 20 plants or parts of plants are observed.

- If there are no off-type plants in 20 plants then the variety does not exceed the number of allowed off-types for this characteristic for this growing cycle
- If there are more than 3 off-type plants then the variety exceeds the number of allowed off-types for this characteristic for this growing cycle.
- If there are 1, 2 or 3 off-type plants then the second step follows.

In the second step further 80 plants or part of plants are observed.

- If there are 3 or less off-type plants in 100 (20 of step 1 + 80 of step 2) plants then the variety does not exceed the number of allowed off-types for this characteristic for this growing cycle
- If there are more than 3 off-type plants in 100 (20 of step 1 + 80 of step 2) plants then the variety exceeds the number of allowed off-types for this characteristic for this growing cycle

The decision rule is defined as follows:

> A variety is considered to be within the uniformity standard in a given growing cycle if the number of off-types in all samples does not exceed the number of allowed off-types in either of the samples.
>
> A variety is considered to be uniform if it is within the uniformity standard in both of the two growing cycles.

If at the end of two growing cycles the variety is within the uniformity standard in one cycle but not in the other cycle, the test will be continued for a third growing cycle. If at the end of the third growing cycle the variety is within the uniformity standard, the variety is considered to be uniform. If at the end of the third growing cycle the variety fails to meet the uniformity standard, the variety is considered to be non-uniform.

This so-called stepwise approach can be used for a number of characteristics in each growing cycle in which the steps are represented by the samples (20 plants or 20 + 80 plants).

In order to compare different tests and decision rules it is useful to compare appropriate type-I and type-II errors. Basic ideas were described by experts from France in document TWC/13/17 “Sequential analysis”.

It should be reminded that a decision for uniformity of a variety has always a so-called type-II error (acceptance of null hypothesis) whereas a decision for non-uniformity of a variety has a so-called type-I error (rejection of null hypothesis). The following applies for the described cereal example:

It is assumed that the population standard is 1% and the acceptance probability is 95%. The number of allowed off-types within 100 plants is 3. All risks are evaluated on the basis of the binomial distribution. To compute the type-II error the population standard for the non-uniform varieties is assumed to be 2% (two times 1%).

For 20 plants the number of allowed off-types is normally 1. However, for 20 plants in the first step of the two-step procedure (see above) it is defined that no off-type is allowed. If we have 20 plants in the first step and no off-type the type-II error (beta risk) is 66.8%. This is high but comparable with 2 off-types in 100 plants. In this case the type-II error is 67.7%.

Looking on the type-I error, the actual error is 1.7 % in case of 20 plants and 1 off- type, and 1.8 % in case of 3 off-types in 100 plants. A decision on the basis of a sample of 20 plants is only taken if there are no off-types which represents a smaller error than the decision taken on 100 plants.

In case of 20 plants in the first step and more than 3 off-types the type-I error (alpha risk) is almost zero. It is very small and smaller than in case of more than 3 off-types in 100 plants. In this case the type-I error (alpha risk) is 0.3 %.

If in the first step there are 1, 2 or 3 off-types in 20 plants the next 80 plants are assessed and the decision is taken on the basis of 100 plants. In that case the type-I and type-II errors have to be evaluated using special formulas or using the software from France (http://www.seedtest.org/en/stats_tool_box_content---1--1143.html).

So it is possible to compare all the other situations.

The open question is: Do we need a statement regarding to the different types of errors for each decision or do we need a statement over all decisions. This should be discussed with statistical experts.
Because of underlying nonlinear formulas there are no general rules for all practical cases usable in DUS tests.

[End of Annex V and of document]