UPOV

International Union for the Protection of New Varieties of Plants

E

Technical Working Party on Automation and Computer Programs **TWC/36/7**

Thirty-Sixth Session
Hanover, Germany, July 2 to 6, 2018

Original: English
Date: June 24, 2018

RISKS ASSOCIATED WITH ASSESSMENT OF UNIFORMITY BY OFF-TYPES ON THE BASIS OF MORE THAN ONE GROWING CYCLE

Document prepared by experts from Germany and the United Kingdom

INTRODUCTION

1. The report of the Technical Working Party on Automation and Computer Programs (TWC) at its thirty-fifth session (see document TWC/35/21) reads:

> "26. The TWC agreed that the different results obtained using the different approaches for the assessment of off-types on the basis of more than one growing cycle were due in part to the different risks of type I and type II errors associated with each approach. The TWC agreed to invite the experts from Germany, the United Kingdom and other members of the Union to submit papers on the analysis of risks associated with each approach to be considered at its thirty-sixth session."

2. The text below explains the considerations needed when planning off-types assessments over two or more cycles.

THE IMPORTANCE OF CONSIDERING RISKS WHEN ASSESSING UNIFORMITY BY OFF-TYPES

3. Population standards for off-types are given in Test Guidelines, defining the maximum allowable proportions of off-types.

4. The population standard is the maximum level of off-types in all individuals of variety. This is a hypothetical concept since we cannot assess all individuals. Instead, we examine a sample of individuals, with the number depending on circumstances and indicated in the Test Guidelines.

5. As we only look at a relatively small number of individuals, the proportion of off-types we see may not reflect perfectly the proportion in the variety.

6. For example, if the proportion of off-types in the variety is 5% and the sample size is 500 plants, we might find the following numbers of off-types on different occasions (random simulation): 29, 19, 21, 27, 30, 29, 32, 28, 21 or 22.

7. Some of these numbers are more than 5% of 500 plants, others less. If we were to set the maximum number of off-types to 25 (5% of 500), approximately half the time we would expect a variety with 5% off-types to fail this uniformity criterion.

8. We should allow for this natural sampling variability in the way that we assess varieties for off-types. We can do this by understanding the risks associated with the decisions.

TYPE OF RISKS TO CONSIDER

9. From a statistical viewpoint, we should consider two types of risks in particular:

Type I error: The conclusion based on the test statistic, i.e. from the DUS trial, is that a variety is not uniform, when it would be uniform if the trial could be repeated indefinitely.

Type II error: The conclusion based on the test statistic, i.e. from the DUS trial, is that a variety is uniform, when it would be non-uniform if the trial could be repeated indefinitely.

10. We can alter the chances of these two types of errors by the way we assess off-types. For a single cycle assessment, we can do this by changing the maximum number of off-types that is acceptable or by changing the sample size. The probabilities of these types of errors can be calculated using mathematical formulae.

11. In general, the maximum number of off-types is set to ensure that chance of type I errors is sufficiently low. In most cases, we aim to have the chance of type I errors being less than 5%. This equates to an acceptance probability of 95%.

12. Given the sample size, the population standard and the acceptance probability, for a single cycle assessment we can identify the maximum number of off-types. These can be found in tables in document TGP/8/3 "Trial Design and Techniques used in the Examination of Distinctness, Uniformity and Stability" Part II Section 8. For example, the maximum number of off-types with a sample size of 500, a population standard of 5% and an acceptance probability of 95% is 33. Note that 33 is 6.6%, a proportion that is larger than the population standard. So there is a margin of safety for the candidate variety.

13. If the test is set up this way, increasing the sample size reduces the chance of type II errors. For example, with a sample size of 500, a population standard of 5% and an acceptance probability of 95%, the chance of a variety with 10% off-types having an acceptable number of off-types (type II error rate) is just 0.5%. However if the sample size is reduced to 50 plants so that the acceptable number is now 5, then the type II error rate increases to 63%.

14. Guidance on sample sizes can be found in Test Guidelines.

CONSIDERING RISKS WHEN UNIFORMITY BY OFF-TYPES IS ASSESSED ON BASIS OF MORE THAN ONE GROWING CYCLE

15. When examining off-types over two or more cycles, it is necessary to consider which risks are most important when setting thresholds in each cycle. This is particularly true for approaches 1 and 2 as set out in Annex I of document TWP/1/17 Rev., which is reproduced in Annex to this document. There are two key choices:

a) To consider risks in each cycle separately;
b) To consider risks for the complete decision process over multiple cycles.

16. The advantage of following a) is that it is easier to work out the maximum number of off-types required for each cycle to achieve the required acceptance probability; tables in TGP/8/3 will give these. For b), tables have not yet been produced.

17. However, b) has the advantage that the selected acceptance probability will be achieved for the whole tests which will have the effect of reducing the chance of type II errors. The choice should lead to greater consistency in standards for off-type assessment members, independently of whether off-types are assessed in one cycle or more than one cycle (and independent of approach). Risks are balanced appropriately for the overall decision on uniformity.

18. For example, let consider approach 1 with 50 plants in each cycle, a 1% population standard with acceptable probability of 95%. [In each approach 1, the two cycles are assessed separately. If there is inconsistency in the verdicts for the two cycles, a third cycle is assessed]. If the risks are considered separately for each cycle then the maximum number of off-types in each cycle would be set at 2. If the overall risks are considered, the maximum number is 1. The overall chance of declaring a variety with 5% off-types as uniform is 56% when the maximum number is 2 and 19% when it is 1. So setting the number of off-types allowed based on the acceptance probability for a single cycle results in a much higher overall chance of type II errors.

ILLUSTRATION OF THE CALCULATION OF OVERALL RISKS

19. Figure 1 shows different cases for the three approaches and how different type I and type II errors are allocated to produce overall errors.

20. On left hand side of the Figure in the first column 1 the year is year (or cycle) is indicated. To remember the content of the three approaches, information (such as the number of plants) is included in the figure.

21. Seven different cases have defined for each approach to illustrate the different type I and type II errors. For example, in case A to case F, the number of plants is equal to 50 and for case G the number is 100.

Approach 1:

Case A: The variety is uniform in the first and in the second year. The stagewise type **II** errors are β_1 and $\beta_{2.}$ The decision is that the variety is overall uniform. The corresponding overall type **II** error for case A is $\beta_{TA.}$

In the draft guidance found in document TWP/1/17 Rev. (Annex I, Page 2), Examples 1 and 2 (Off-types in first cycle=1, second=1; and First=2, Second=2, respectively) fulfill the conditions of case A.

Case B: The variety is uniform in the first year and in second year non-uniform. The stagewise type II and type **I** errors are β_1 and $\alpha_{2.}$The procedure needs a third year. As the variety is uniform in the third year (type II error is β_3) then the decision is that the variety is overall uniform. The corresponding overall type II error for case B is $\beta_{TB.}$

Examples 3, 4 and 5 (First=0, Second=3; First=1, Second=3; First=1, Second=4, respectively) could fulfill the conditions of case B.

Case C: The variety is uniform in the first year and second year non-uniform. The stagewise type II and type **I** errors are β_1 and $\alpha_{2.}$The procedure needs a third year. As the variety is non-uniform in the third year (type I error is α_3) then the decision is that the variety is overall non-uniform. The corresponding type I error for case C is $\alpha_{TC.}$

Examples 3, 4 and 5 (First=0, Second=3; First=1, Second=3; First=1, Second=4, respectively) could fulfill the conditions of case C.

Case D: The variety is non-uniform in the first year and uniform in the second year. Case G conditions have not been fulfilled. The stagewise type **I** and type **II** errors are α_1 and $\beta_{2.}$The procedure needs a third year. As the variety is uniform in the third year (type II error is β_3) then the decision is that the variety is overall uniform. The corresponding type **II** error for case D is $\beta_{TD.}$

There is no example in document TWP/1/17 Rev. for this case.

Case E: The variety is non-uniform in the first year and uniform in the second year. Case G conditions have not been fulfilled. The stagewise type I and type II errors are α_1 and $\beta_{2.}$The procedure needs a third year. As the variety is non-uniform in the third year (type **I** error is α_3) then the decision is that the variety is overall non-uniform. The corresponding type I error for case E is $\alpha_{TE.}$

There is no example in document TWP/1/17 Rev. for this case.

Case F: The variety is non-uniform in the first and in the second year. Case G conditions have not been fulfilled. The stagewise type I errors are α_1 and $\alpha_{2.}$ The decision is that the variety is overall non-uniform. The corresponding overall type I error for case F is $\alpha_{TF.}$

There is no example in document TWP/1/17 Rev. for this case.

Case G (Exception): The number of off-types in the first exceeds the allowed number of off-types for two growing cycles combined (e.g. based on 100 plants instead of 50) the variety is non-uniform. The stagewise type I error for case G is $\alpha_{S1.}$

Example 6 (First=4) fulfills the conditions of case G.

Overall for approach 1: The uniformity decision of approach 1 has an overall type II error $\beta_{A1,}$ which is a combination of the different type II errors β_{TA}, β_{TB} and $\beta_{TD.}$

The non-uniformity decision of approach 1 has an overall type I error αA1, which is a combination of the different type I errors αS1, αTC, αTE and αTF.

Approach 2:

Case A: The variety is uniform in the first and in the second year. The stagewise type II errors are β_1 and $\beta_{2.}$ The decision is that the variety is overall uniform. The corresponding overall type II error for case A is $\beta_{TA.}$

Case A of Approach 2 is the same as case A of approach 1.

Examples 1 and 2 (First=1, Second=1; First=2, Second=2, respectively) fulfill the conditions of case A.

Case B: The variety is uniform in the first year and in second year not. The stagewise type II and type I errors are β_1 and $\alpha_{2.}$The procedure continues by using the sum of number of off-types of year 1 and 2. As the variety is uniform by using the sum then the decision is that the variety is overall uniform. The corresponding overall type II error for case B is $\beta_{SB.}$

Example 3 (First=0, Second=3) fulfills the conditions of case B.

Case C: The variety is uniform in the first year and second year not. The stagewise type II and type I errors are β_1 and $\alpha_{2.}$The procedure continues by using the sum of number of off-types of year 1 and 2. As the variety is non-uniform by using the sum then the decision is that the variety is overall non-uniform. The corresponding type I error for case C is $\alpha_{SC.}$

Examples 4 and 5 (First=1, Second=3; First=1, Second=4, respectively) fulfill the conditions of case C.

Case D: The variety is non-uniform in the first year and uniform second year. Case G conditions have not been fulfilled. The stagewise type I and type II errors are are α_1 and β_2. The procedure continues by using the sum of number of off-types of year 1 and 2. As the variety is uniform by using the sum then the decision is that the variety is overall uniform. The corresponding type II error for case D is $\beta_{SD.}$

There is no example in document TWP/1/17 Rev. for N1=N2=50.

Case E: The variety is non-uniform in the first year and uniform second year. Case G conditions have not been fulfilled. The stagewise type I and type II errors are α_1 and $\beta_{2.}$The procedure continues by using the sum of number of off-types of year 1 and 2. As the variety is non-uniform by using the sum then the decision is that the variety is overall non-uniform. The corresponding type I error for case E is $\alpha_{SE.}$

There is no example in document TWP/1/17 Rev. for N1=N2=50.

Case F: The variety is non-uniform in the first and in the second year. The stagewise type I errors are α_1 and $\alpha_{2.}$ The procedure can stopped and the decision is that the variety is overall non-uniform. The corresponding overall type I error for case F is $\alpha_{TF.}$

There is no example in document TWP/1/17 Rev. for N1=N2=50.

Case G (Exception): The number of off-types in the first year exceeds the allowed number of off-types for two growing cycles (based on 100 plants, for example) and so the variety is non-uniform. The stagewise type I error for case G is $\alpha_{S1.}$

Example 6 (First=4) fulfills the conditions of case G.

Overall for approach 2: The uniformity decision of approach 2 has an overall type II error $\beta_{A2,}$ which is a combination of the different type II errors β_{TA}, β_{SB} and $\beta_{SD.}$

The non-uniformity decision of approach 2 has an overall type I error αA2, which is a combination of the different type I errors αS1, αSC, αSE and αTF.

Approach 3:

Case A: The variety is uniform based on the sum of the first and in the second year. The decision is that the variety is overall uniform. The corresponding overall type II error for case A is $\beta_{S2.}$

Examples 1 and 3 (First=1, Second=1; First=0, Second=3, respectively) fulfill the conditions of case A.

Case F: The variety is non-uniform based on the sum of the first and in the second year. The decision is that the variety is overall non-uniform. The corresponding overall type I error for case F is $\alpha_{S2.}$ Examples 2 and 3 (First=2, Second=2; First=1, Second=3, respectively) fulfill the conditions of case F.

Comparison of these cases for the different approaches is difficult at least for approach 3. But, at least for approach 2 and 3, case B and case C for approach 2 might be included as case A for approach 3. Case D and case E for approach 2 might be included as case F for approach 3.

Case G (Exception): The number of off-types in the first year exceeds the allowed number of off-types for two growing cycles (e.g. based 100 plants) and so the variety is non-uniform. The stagewise type I error for case G is $\alpha_{S1.}$

Example 6 (First=4) fulfills the conditions of case G.

Overall for approach 3: The uniformity decision of approach 3 has an overall type II error $\beta_{A3,}$ which equals to $\beta_{S2.}$

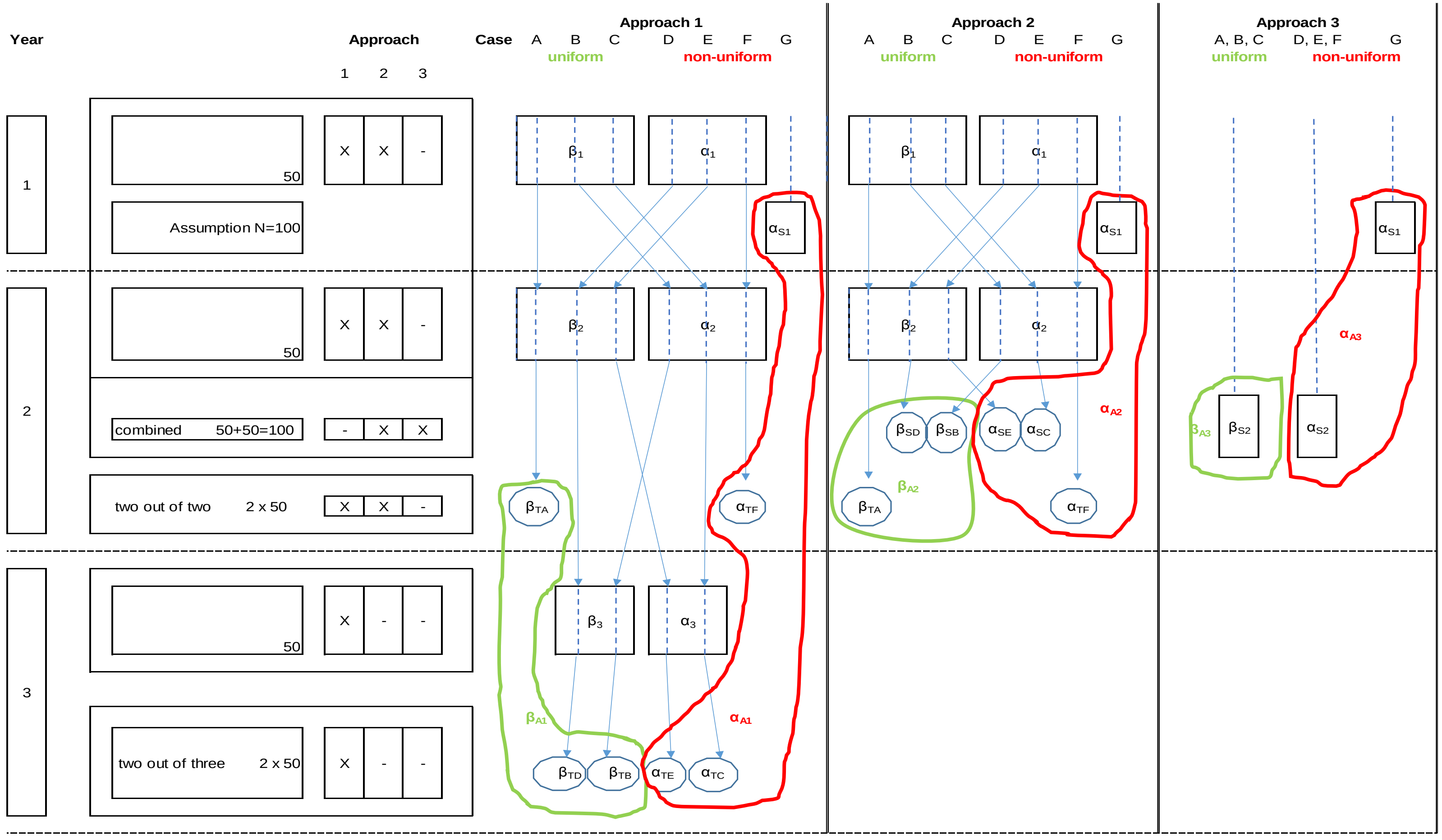
The non-uniformity decision of approach 3 has an overall type I error αA3, which is a combination of the different type I errors αS1 and αS2.

RECOMMENDATION

22. Whichever approach is used, when uniformity is assessed over two or more cycles, the thresholds for the acceptable number of off-types should be based on the overall type I error (1 minus the acceptance probability). This will lead to greater harmonisation of standards, whatever approach is used.

23. Tables or software could be developed to support this.

Figure 1



Legend:

α_1	type I error for the first year	stagewise
α_2	type I error for the second year	stagewise
α_3	type I error for the third year	stagewise
β_1	type II error for the first year	stagewise
…		
α_{A1}	Type I error for approach 1	overall
β_{A1}	Type II error for approach 1	overall
…		

T stands for two out of two or three years
S1 stands for the assumption to have 100 plants instead of 50 for example
SD stand for Sum over two years but with looking on year1 and year2for case D
S2 stands for Sum over two years without looking on year 1 or year 2
A,B,C,D,E,F,G are different cases

[Annex follows]

ANNEX I OF DOCUMENT TWP/1/17 REV.

ASSESSING UNIFORMITY BY OFF-TYPES ON BASIS OF MORE THAN ONE GROWING CYCLE

Two independent growing cycles could take place in a single location in different years, or in different locations in the same year, according to document TGP/8 Part I, Sections 1.2 and 1.3.

The following guidance is not intended to be used for the assessment of uniformity by off-types on the same plants in two growing cycles. Results from growing cycles using different lots of plant material should not be combined.

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. 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. It is important to identify whether differences in number of off-types between growing cycles were due to biological reasons or sampling variation. Furthermore, ~~on the basis of a clear lack of uniformity, a~~ if a variety exceeds in the first growing cycle the allowed number of off-types in two growing cycles, the variety may be rejected after a single growing cycle.

Approach 2: Combining the results of two growing cycles 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, 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 sample size of growing cycles 1 and 2 combined.

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 statistical test for consistency should be applied when appropriate. It is important to identify whether differences in number of off-types between growing cycles were due to biological reasons or sampling variation. Furthermore, ~~on the basis of a clear lack of uniformity, a~~ if a variety exceeds in the first growing cycle the allowed number of off-types in two growing cycles, the variety may be rejected after a single growing cycle.

Approach 3: Combining the results of two growing cycles

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.

A variety is considered non-uniform if the total number of off-types at the end of the two growing cycles exceeds the number of allowed off-types for the combined sample.

A variety may be rejected after a single growing cycle, if the number of off-types exceeds the number of allowed off-types for the combined sample (over two cycles).

Care is needed when considering results that are very different in each of the growing cycles, such as when a type of off-type is observed at a high level in one growing cycle and is absent in another growing cycle. A statistical test for consistency ~~is possible~~ should be applied when appropriate. It is important to identify whether differences in number of off-types between growing cycles were due to biological reasons or sampling variation.

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	Approach 3
Number of Off-Types	1	1	uniform	uniform	uniform
	2	2	uniform	uniform	non-uniform
	0	3*	third growing cycle*	uniform*	uniform*
	1	3*	third growing cycle*	non-uniform*	non-uniform*
	1	4*	third growing cycle*	non-uniform*	non-uniform*
	4**	1*	third growing cycle*	non-uniform*	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 statistical test for consistency should be applied when appropriate. It is important to identify whether differences in number of off-types between growing cycles were due to biological reasons or sampling variation.

** If a variety exceeds in the first growing cycle the allowed number of off-types in two growing cycles, the variety may be rejected after a single growing cycle.

[End of Annex and of document]