

# ISTA GMO Proficiency Tests

## Use of presence/absence (qualitative) results for an overall rating from more than one test

By **Christoph Haldemann**, ISTA GMO Task Force Member, Swiss Federal Research Station for Animal Production, Switzerland, e-mail: c\_haldemann@yahoo.de and **Sylvain Grégoire**, ISTA Statistics Committee Chair GEVES, France, e-mail: sylvain.gregoire@geves.fr



Christoph Haldemann



Sylvain Grégoire

### 1. Introduction

The aim of this article is to show how **presence/absence** results of Proficiency Test (PT) with known spiking levels of GM seeds could be managed; using the system which is already in use in ISTA for other types of tests (Purity, Germination, etc.).

Up to now ISTA GMO TF carried out 3 Proficiency Tests; participating laboratories had to detect the presence of GM seeds in samples of conventional seed of corn, (*Zea mays* L.).

In this article both

- a conventional seed sample wrongly reported by the laboratory as containing GM seeds,
  - a sample containing GM seeds wrongly reported by the laboratory as containing no GM,
- are considered as a miss-classification (an error).

With this article we use the number of misclassified samples, and suggest 3 rating systems, one compares with an absolute number of miss-classifications, and 2 others compare with a relative number (a percentage) of miss-classifications.

Actual data from the laboratories which participated in all three ISTA Proficiency have been used. This is a total of 21 laboratories, as not all laboratories participated to the 3 tests.

### 2. Design of the three ISTA Proficiency Tests

The laboratories received samples with corn seeds (*Zea mays* L.) only.

Some samples contained only conventional seeds, for these samples the expected result was "absence of GM seeds".

Other samples of conventional seeds (non GM seeds) were spiked with GM seeds, for these samples the expected result was "presence of GM seeds".

### Proficiency Test 1

Each laboratory received 30 samples containing about 300 seeds.

Composition of samples:

- 12 negative samples
  - 6 positive samples, 1%; Mon810 only (3 seeds)
  - 6 positive samples, 1%; T25 only (3 seeds)
  - 6 positive samples, 1%; Mon810 + T25 (1 seed of Mon810 + 2 seeds of T25)
- 43 laboratories participated.

### Proficiency Test 2

Each laboratory received 10 samples containing about 3000 seeds.

Composition of samples:

- 3 negative samples
  - 3 positive samples, 0.7%; Mon810 only (21 seeds)
  - 4 positive samples, 1.4%; Mon810 only (42 seeds)
- 50 laboratories participated.

### Proficiency Test 3

Each laboratory received 12 maize samples containing about 1500 seeds.

Composition of samples:

- 3 negative samples
  - 3 positive samples, 0.2%; Mon810 + T25 (1 seed of Mon810 + 2 seeds of T25)
  - 3 positive samples, 2%; Mon810 + T25 (10 seed of Mon810 + 20 seeds of T25)
  - 3 positive samples, 4%; Mon810 + T25 (20 seed of Mon810 + 40 seeds of T25)
- 40 laboratories participated.

### 3. The current ISTA system to rate a Proficiency Test, and a run of 6 tests

ISTA Proficiency Tests overall rating procedure is described below (Table 1).

One test rating	One test Score Value	Overall rating on 6 tests	Range on 6 tests <sup>5)</sup>
<sup>1)</sup> A	5 points	<sup>1)</sup> A	28 - 30 points
<sup>2)</sup> B	4 points	<sup>2)</sup> B	21 - 27 points
<sup>3)</sup> C	3 points	<sup>3)</sup> C	16 - 20 points
<sup>4)</sup> BMP	0 points	<sup>4)</sup> BMP	below 16 points

Table 1: ISTA rating system for 6 Proficiency Tests based on the in-round rating values.

Depending on the results obtained by a laboratory on a given proficiency test, the laboratory is rated A, B, C or BMP. There is no official definition of A, B, C and BMP in the case of GM tests, the definitions below are an attempt to show the philosophy of the system which can be applied to many types of tests:

To each rate correspond a number of points, the bigger number the better the results. The decrease from A to C and the "0" for BMP, are another way to understand the meaning of these rates.

The rating for a given Proficiency Test is an indication to the lab on its performance in this test.

For laboratories which have an ISTA Accreditation on a type of test, participation to Proficiency Tests is compulsory. The set of the 6 most recent tests is used in the evaluation process. The sum of points collected on the tests, is compared on the range in last column of Table 1.

<sup>5)</sup> Range for 6 Proficiency Tests

For instance a lab with a total of 23 points on 6 tests will be rated B on the set of 6 tests.

In this article we use exactly this system, but as we have only 3 tests available we use a "Range 2", which is the usual range divided by 2.

This range 2 shall not be used in the current ISTA system, but will allow us to describe the mechanism of the proposal (as soon as 6 tests are available, the current ISTA system can be applied).

rating	score	Overall rating	Range <sup>5)</sup>	Range 2 <sup>6)</sup>
<sup>1)</sup> A	5	<sup>1)</sup> A	28 - 30	<b>14 - 15</b>
<sup>2)</sup> B	4	<sup>2)</sup> B	21 - 27	<b>11 - 13</b>
<sup>3)</sup> C	3	<sup>3)</sup> C	16 - 20	<b>8 - 10</b>
<sup>4)</sup> BMP	0	<sup>4)</sup> BMP	below 16	<b>below 8</b>

<sup>1)</sup> A No problem has been detected in this test.

<sup>2)</sup> B There are small problems, but no specific look or action is suggested to the participant.

<sup>3)</sup> C Problems, ISTA indicates there might be things to consider by the laboratory to explain or correct things.

<sup>4)</sup> BMP Below Minimum of Performance, ISTA indicates by a letter that the results were poor and the laboratory has to explain and correct things.

<sup>6)</sup> Range 2 is for Proficiency Tests in this article

**Table 2:** Range 2, used to show the mechanism with only three tests available.

In the current ISTA system, the mechanism to derive a A,B,C or BMP from the raw results of the test may differ depending on the type of test. The computations are for instance not the same to check Germination, and for the retrieval of spiked seeds of other species.

For presence/absence of GM seed, there is not yet an official computation system in place in ISTA. The following proposals show possibilities that have been considered by the Proficiency Test sub-group of the ISTA GMO TF in conjunction with the ISTA Statistics Committee.

#### 4. Proposal of a rating system for presence/absence of GM seeds in conventional seeds

The Proficiency Tests carried out by ISTA on detection of GM seeds in conventional seeds are prepared with great efforts to ensure that the expected result (presence or absence) of each sample is very sure.

This can be compared to the spiking of seeds from other species in seeds from a species, as carried out for other seeds determination. This is different from the situation of germination, where the number of normal seeds in a given sample can not be known in advance.

The proposals are based on counting the "mistakes". Two types of mistakes can occur:

- a conventional seed sample is wrongly reported by the laboratory as containing GM seeds,
- a sample containing GM seeds is wrongly reported by the laboratory as containing no GM seeds.

The total number of miss-classified samples is used as a basis for the rating of a given Proficiency Test.

#### Rating calculated with an absolute number of misclassifications (Rating System 1).

In this system the number of misclassifications is used, and a table gives the correspondence to the rate as shown below in Table 3.

rate	Number of misclassified samples
A	0 errors
B	1 or 2 errors
C	3 errors
BMP	more than 3 errors

**Table 3:** Rating system 1

#### Rating calculated with relative (percentage) errors (Rating System 2 and Rating System 3).

The principle for the calculation remains the same, counting the number of samples misclassified. But instead of working with an absolute number of errors, the errors are computed in percent of all samples received for each given Proficiency Test, and compare to the right column (Percentage of misclassified samples) shown in Table 4 and Table 5.

Two proposals are described in Table 4 and Table 5.

rate	Number of misclassified samples
A	0% - 5%
B	>5% - 10%
C	>10% - 20%
BMP	>20%

**Table 4:** Rating system 2

rate	Number of misclassified samples
A	0% - 6%
B	>6% - 20%
C	>20% - 30%
BMP	>30%

**Table 5:** Rating system 3

*ISTA GMO Proficiency Tests is continued on page 10*

### 5. Results

The number of misclassifications (Y axis) for each laboratory is summarised in Graph 1 for the 21 laboratories (X axis) which carried out the 3 Proficiency Tests.

See Graph 1

13 laboratories made no misclassification. 8 laboratories had some misclassifications in one, two, or the three Proficiency Tests. PT1, PT2, PT3 refers respectively to the first, second and third ISTA Proficiency Test.

The 3 proposed rating systems have been applied to each of the 3 Proficiency Tests separately. Then the sum of points has been compared to Range 2 in Table 2, to obtain an overall rate on 3 tests. The 4 overall rates (A,B,C and BMP on X axis) is shown in graph 2, where the Y axis is the number of labs (out of 21) in each rate category for each rating system (black for rating system 1, red for rating system 2, green for rating system 3).

See Graph 2

16 to 18 labs obtained an overall A. With the range applied, and the current ISTA system, an overall A can be obtained even if not all individual tests are A.

In this article a 14 can be obtained with 2 A and 1 B resulting in an overall A. (A is obtained if the sum of points is 14 or 15)

### 6. Discussion

The three proposed rating systems give rather similar outcomes.

The figure in Graph 2 is rather similar to graphs shown at Budapest Congress on other types of tests.

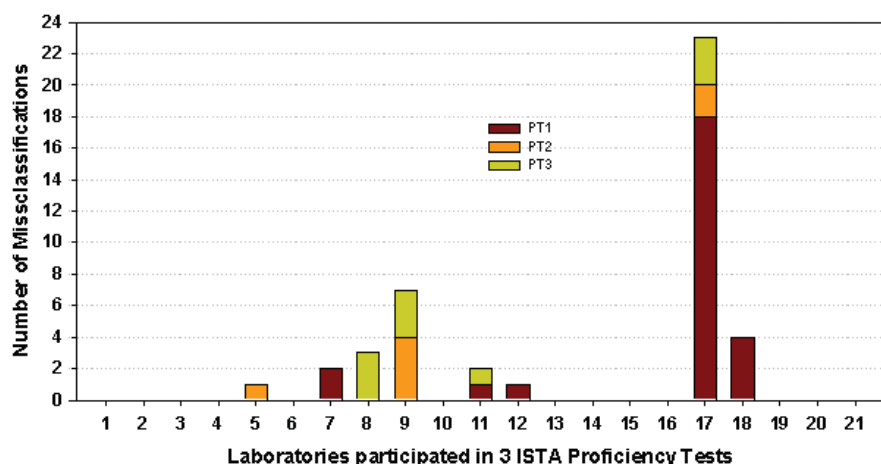
If we look at the most crucial score class BMP (Below Minimum Performance), one or two labs are pointed out, depending on the rating system. It make sense that a lab with mistakes in all the three tests and 23 misclassifications in total is ranked as BMP.

Only a few laboratories had problems to detect the presence or absence of GM seeds in samples of conventional seed of corn (*Zea mays* L.). Some participating laboratories are already very well experienced, while some other laboratories may have participated to the tests with yet few or not enough experience, as a way for them to check their progress.

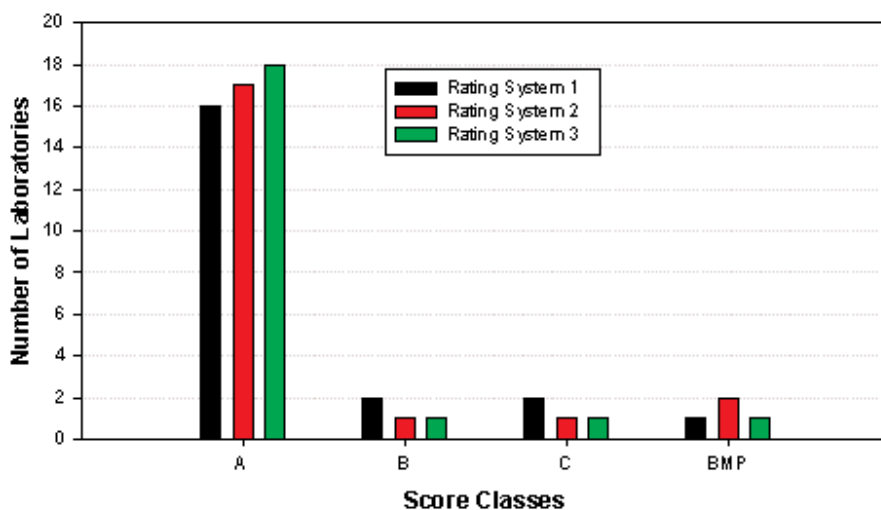
The values used to derive a rate from misclassifications (Number of misclassified samples in Table 3, respectively Percentage of misclassified samples in Table 4 and 5) are possibilities among hundreds. The same idea can be kept, and the numbers changed if necessary in the future. These values have been selected not only using the 3 ISTA Proficiency Tests, but also from discussions with laboratory experts.

These proposals are under evaluation by ISTA.

The same question of rating procedure is also under development for the rating of results expressed as a quantity of GM material (% seeds, % DNA copies, etc.) The general idea has been shown at Budapest Congress; a short description of the proposals will be published in a next issue. ■



Graph 1: Total errors for PT1, PT2 and PT3 (30 + 10 + 12 = 52 samples)



Graph 2: Comparison of the 3 Rating Systems for the laboratories participated in all 3 ISTA Proficiency Tests (21 laboratories)

## Announcement 4<sup>th</sup> ISTA Proficiency Test on GMO Testing

After performing three rounds on *Zea mays* L. there will be a change in the species for the 4<sup>th</sup> round to Soybean (*Glycine max* (L.) Merr).

Laboratories interested in participating should please contact the ISTA Secretariat:

E-mail: [ista.office@ista.ch](mailto:ista.office@ista.ch)

Fax: +41 1 838 60 01

More details can be found in the Announcement posted on the ISTA Website at [www.seedtest.org](http://www.seedtest.org)

Please find more details about the Proficiency Test on page 48.