Measurement Assurance and Certification Scotland

PERFORMANCE STANDARD MACS-FFA-PS-02

Finfish Aquaculture Sector

Physical and chemical testing

Version 1
March 2022

Every day SEPA works to protect and enhance Scotland's environment, helping communities and businesses thrive within the resources of our planet.

We call this **One Planet Prosperity**

Contents

1 Introduction

Every day SEPA works to protect and enhance Scotland's environment, helping communities and businesses thrive within the resources of our planet. We call this One Planet Prosperity. If everyone in the world lived as we do in Scotland, we would need three planets. There is only one.

We're changing today, creating a world-class 21st Century EPA fit for the challenges of tomorrow by grounding our regulatory activity across whole sectors.

A fundamental principle of our sector approach is that environmental compliance is non-negotiable. In every sector, we will ensure that all regulated businesses fully meet their environmental compliance obligations.

In certain sectors, this means that operators performing authorised activities have an obligation to monitor and report back to us in support of the regulation of those activities. We will determine compliance from the data and evidence submitted to us.

In order to maintain confidence in our regulatory decision making, all operator monitoring data must meet our minimum quality requirements. To help operators meet those requirements, we have established Measurement Assurance and Certification Scotland (MACS) - our quality assurance certification scheme.

MACS comprises a range of performance standards and technical guidance documents, each designed to ensure that operator monitoring data is fit for regulatory assessment. Its remit extends across the entire monitoring process; from planning and scheduling of monitoring activity to sampling, analysis and data reporting.

Where an organisation conforms with the requirements of MACS, the operator monitoring data they produce will be of a standard that meets our minimum quality requirements. To ensure that this remains the case, those organisations will be routinely audited.

Further information on MACS, operator monitoring, and our sector approach may be found on the SEPA website:

www.sepa.org.uk

2 Scope

2.1 This MACS performance standard is applicable to organisations carrying out physical and chemical testing of samples for the assessment of seabed standards relating to marine pen fish farms (MPFFs).

2.2 Sections 5 to 7 and Annexes A, B and C lay out the detailed requirements that those organisations **must** adhere to when carrying out those activities.

2.3 Guidance, which may be applied by an organisation in order to meet certain specific requirements, may be found in complementary technical guide MACS-FFA-TG-01 (ref. 3.1.a).

3 References and bibliography

3.1 Text references

a. MACS Technical Guide MACS-FFA-TG-01, Finfish Aquaculture Sector - Dealing with non-conformance, Scottish Environment Protection Agency, 2022.

3.2 Bibliography

a. BS EN ISO/IEC 17025:2017 - General requirements for the competence of testing and calibration laboratories, ISBN 978 0 539 01414 3.

b. BS ISO 3310-1:2016 - Test sieves. Technical requirements and testing. Test sieves of metal wire cloth, ISBN 978 0 580 83347 2.

c. BS ISO 3310-2:2013 - Test sieves. Technical requirements and testing. Test sieves of perforated metal plate, ISBN 978 0 580 82112 7.

d. MACS Performance Standard MACS-FFA-PS-01, Finfish Aquaculture Sector - Sampling of soft-substrate, Scottish Environment Protection Agency, 2022.

e. MACS Performance Standard MACS-FFA-PS-03, Finfish Aquaculture Sector - Biological testing, Scottish Environment Protection Agency, 2022.

4 Terms and definitions

For the purpose of this MACS performance standard, and unless the context requires otherwise, the following definitions apply:

concession – a written approval, granted to release a non-conforming product or service for use or delivery. For example, a written agreement from SEPA explicitly permitting the submission of data associated with a quality control failure.

direct method – an analytical method where samples are analysed directly with no sample preparation.

interlaboratory comparison – organisation, performance and evaluation of measurements or tests on the same or similar items by two or more laboratories in accordance with predetermined conditions.

metrological traceability – the property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.

operator – an individual or company responsible for the operation of an existing or proposed marine pen fish farm that will be subject to operator monitoring activities.

organisation – an entity performing an activity or activities required under operator monitoring. In the context of this performance standard, this term encompasses an operator, or a body appointed by that operator to undertake testing activity on their behalf.

pre-treatment method – an analytical method where samples undergo some form of sample preparation prior to instrumental analysis (e.g. organic solvent extraction).

proficiency testing – evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons.

recovery – the proportion of the amount of analyte, present in or added to the analytical portion of the test material, which is extracted and presented for measurement.

sample – a volume of water or soft-substrate collected from a sampling station and identified for the assessment or measurement of specific determinand(s).

soft-substrate – areas of sea floor consisting of loose deposited particles including clay, mud, sand and gravel, and shells. Also includes mixed substrata with gravels, small stones and pebbles scattered on a bed of finer material, but excluding cobbles.

sub-sample – a representative portion removed from a sample for separate analysis.

5 Resource requirements

5.1 Personnel

5.1.1 Organisations must ensure that all personnel performing activities relating to the requirements of this performance standard have been deemed competent in, and are authorised to undertake, those activities.

5.1.2 Organisations must document and implement procedure(s) for:

- determining competence requirements;
- the training and supervision of personnel;
- assessing the initial competence of personnel;
- ongoing monitoring of the competence of personnel; and
- the authorisation of personnel.

5.2 Testing facilities

5.2.1 Testing facilities must be suitable for the activities being undertaken and must not affect the validity of reported result(s).

5.2.2 Measures put in place to ensure the suitability of those testing facilities must be documented, implemented, monitored and periodically reviewed.

5.3 Equipment

5.3.1 Equipment must be suitable for the activities being undertaken and must not affect the validity of reported result(s).

5.3.2 Organisations must document the equipment necessary for the correct performance of their testing activities.

5.3.3 Organisations must have documented procedure(s) in place for the handling, transport, storage, use and planned maintenance of equipment in order to ensure its proper functioning and to prevent its contamination and deterioration.

5.3.4 Organisations must verify that all equipment is functioning properly before placing or returning it into use.

5.3.5 Measuring equipment must be calibrated when:

- the measurement accuracy or uncertainty will affect the validity of reported result(s), and/or
- calibration is required to establish the metrological traceability of reported result(s).

5.3.6 For all calibrated measuring equipment, organisations must implement an ongoing calibration programme to maintain confidence in the calibration status of that equipment.

5.3.7 All measuring equipment requiring calibration must be clearly labelled, such that the user of the equipment can readily identify its calibration status.

5.4 Control of documents

5.4.1 Organisations must implement a management system for the control of documents.

5.4.2 Such a system must ensure that all documents are:

- uniquely identified;
- suitably marked to indicate their current revision status;
- approved by authorised personnel prior to their issue; and
- periodically reviewed and updated where necessary.

5.4.3 Where obsolete documents are retained for any purpose, they must be clearly marked so as to prevent their unintended use.

5.4.4 Copies of any documents relating to the requirements of this performance standard must be provided to SEPA upon request.

5.5 Control of records

5.5.1 Organisations must establish and retain records to demonstrate fulfilment of the requirements of this MACS performance standard.

5.5.2 Such records must be retained for a minimum period of five years.

5.5.3 Copies of these records must be provided to SEPA upon request.

6 Testing requirements

6.1 Test method selection

6.1.1 Organisations must ensure that analytical methodologies employed are fit for purpose and appropriate for the determinand, sample matrix and concentration range to be determined.

6.1.2 Analysis methods must be able to provide absolute test result values for all determinands.

NOTE: The only exception to this requirement is the use of a '<' qualifier when submitting a test result determined at less than the stated method detection limit.

6.1.3 Analysis methods must have fully documented analytical procedure(s). Copies of those procedure(s) must be provided to SEPA upon request.

6.2 Method validation

6.2.1 Analysis methods must be fully validated using appropriate matrices prior to being used for the generation of data for submission to SEPA.

6.2.2 For all methods, organisations must demonstrate that the analytical performance measured during method validation meets the targets detailed in Annex A of this performance standard.

6.2.3 For all method validation exercises, organisations must keep records of the following information:

- the validation procedure employed;
- the analytical results obtained during the validation exercise;
- a determination of the validated method's analytical performance; and
- a statement detailing the fitness for intended use of the validated method.

Copies of those records must be provided to SEPA upon request.

6.2.4 **Validation procedure**

a. As far as is reasonably practicable, any validation exercise must encompass the whole analysis method.

b. Validation must be performed over multiple analytical batches.

c. The total number of validation batches must be sufficient to allow analytical performance to be determined with at least ten degrees of freedom.

NOTE: This requirement applies to both within-batch and between-batch estimates of degrees of freedom. See Annex C for further detail.

d. As a minimum, each validation batch must contain duplicates of each chosen test sample type.

e. The following performance characteristics must be assessed as part of a validation exercise. Assessments must be carried out in line with the protocols detailed in Annexes B and C:

- Precision (%RSD).
- Bias.
- Method Detection Limit (MDL).

f. Analytical performance must be calculated using final sample concentrations, i.e. corrected for sample weight, blank levels, and recovery (where applicable).

NOTE: Correction for recovery is required for pre-treatment methods where there is known to be significant loss of analyte during sample preparation. Depending on the requirements of an individual method, the recovery factor used to perform this correction may be calculated on either a historic or per-batch basis.

g. Each validation test sample type will return individual estimates of %RSD and Bias. For each performance characteristic, the overall method performance for a determinand must be quoted as the largest estimate taken from all relevant test sample types.

h. No changes shall be made to the analysis method once a validation exercise has commenced. If circumstances indicate that significant changes are required to the method employed, then the validation exercise must be repeated.

NOTE: Before data can be submitted from a revalidated method SEPA must be informed and the validation data submitted to allow analytical performance data to be reviewed.

6.3 Ensuring the validity of results

6.3.1 Variance of replicates

Where replicate samples have been collected from a sampling station, all replicates for an individual analysis must be prepared and analysed in the same analytical batch.

6.3.2 Internal analytical quality control

a. Organisations must ensure that their analysis methods are:

- free from the effects of interferences and contamination; and
- statistically under control and continuing to meet performance targets.

b. The objectives required by 6.3.2 a. above must be achieved through appropriate implementation of the following checks and measures:

- System suitability check(s).
- Analysis of blank sample(s).
- Use of laboratory control sample(s) and statistical control charts.

c. Organisations must document and implement procedure(s) that define the loss of analytical quality control and specify the actions that must be taken in such circumstances.

6.3.3 Proficiency testing scheme participation

a. Organisations must demonstrate the ongoing performance of their analysis methods by participation in an appropriate external proficiency testing (PT) scheme.

b. Where no appropriate external PT scheme is available, organisations must demonstrate ongoing performance by other means (e.g. interlaboratory comparisons other than proficiency testing, use of certified reference materials, replicate testing, intralaboratory comparisons).

c. Organisations must document and implement procedure(s) which provide for review, investigation, and implementation of corrective action when the results submitted for a PT sample are deemed unsatisfactory or questionable by the scheme administrator.

d. Details of an organisation's PT scheme programme, including records of their PT performance, must be provided to SEPA upon request.

7 Control of non-conforming work

7.1 Organisations must have documented procedure(s) which are implemented when any aspect of their testing activity does not conform with the requirements of this performance standard.

7.2 As a minimum, these procedure(s) must provide for incidences of non-conforming work to be recorded, investigated, and evaluated for their significance; and require that a determination is made as to whether the results of that work remain valid.

7.3 Where such an evaluation indicates that a non-conformance could recur, or that there is doubt around the conformity of an activity with either the organisation's own procedure(s) or the requirements of this performance standard, then appropriate corrective action must be implemented.

7.4 SEPA may accept submission of analytical results associated with testing that has not been undertaken according to an organisation's own procedure(s) or the requirements of this performance standard. In each case, a concession to report the affected results must be requested from SEPA.

7.5 Concession requests must include a full assessment of the circumstances of the non-conformance and its potential impacts, and justification as to how the submitted data remains fit for its intended purpose. Where it is not possible to provide a suitable justification, then the non-conforming data will not be accepted by SEPA.

NOTE: For additional guidance on dealing with non-conformance, please refer to complementary technical guide MACS-FFA-TG-01 (ref. 3.1.a).

8 MACS document review and control

8.1 All MACS documentation will be subject to periodic review and may occasionally be amended. For the latest versions of all MACS performance standards, please refer to the SEPA website:

www.sepa.org.uk

Annex A

Performance characteristics

Performance characteristic targets for all tests included in this performance standard are detailed in Table A1.

This list is not exhaustive; parameters and targets will be amended as regulatory and/or statutory environmental monitoring requirement changes dictate.

Table A1 – Physical and chemical parameters

Determinand	Units	MDL[1]	Precision[2]	%Bias
Emamectin benzoate	ng/kg	10% EQS[3]	25	50
Particle Size Analysis (PSA)	%	N/A	10	N/A[4]
Total Organic Carbon (TOC)	%	0.1	10	20

1. For further detail on Method Detection Limit, consult Annex B.
2. Expressed as %RSD.
3. MDL assessment must use a target value of one tenth of the Environmental Quality Standard (EQS) required by the operator's CAR authorisation. EmBz MDL targets may therefore vary by operator.
4. Rather than assessing bias, organisations must demonstrate that the instrumentation used is operating within manufacturer's tolerances.

Annex B

Assessment of method detection limit

MACS requires the adoption of a common approach to the assessment of method detection limit (MDL) in order to ensure that all operator data can be evaluated in a consistent and comparable fashion. The following protocols must be applied when undertaking MDL assessment during a method validation exercise for a physical or chemical test.

B.1 MDL must be determined with a minimum of ten degrees of freedom, using the within-batch performance data generated by the analysis of replicate MDL test sample types during method validation.

B.2 **MDL test sample type**

Wherever possible, the MDL validation test sample type must be prepared using real sample matrix. Choice of appropriate MDL test sample is dependent on the type of analytical method to be employed:

a. For methods which **are** capable of returning numeric values at levels below the instrument detection limit (i.e. negative values), no determinand(s) of interest should be present in the chosen test sample type.

   MDL must be determined using measurements obtained from a blank real sample matrix.

b. For methods which **are not** capable of returning numeric values at levels below the instrument detection limit, a measurable amount of the determinand(s) of interest should be present in the chosen test sample type.

   MDL must be determined using measurements obtained from a blank real sample matrix, spiked with determinand(s) of interest at a level approximately two to five times the instrument detection limit.

   Alternatively, a real sample matrix may be used with sufficiently low levels of determinand(s) of interest naturally present. There would be no requirement to spike this sample matrix.

NOTE 1: In certain circumstances it may not be possible to find a suitable real matrix for MDL assessment, e.g. where there is potential for the presence of significant natural levels of the determinand(s) of interest. In these situations, use of ideal matrix is acceptable.

NOTE 2: During validation, the MDL test sample type must **not** be used as the blank or process blank.

B.3 MDL calculation

B.3.1 Theory

For the purposes of this MACS performance standard, MDL is defined by the equation:

$$\boldsymbol{MDL = 4.65 \times s_w}$$

where:

- s_w is the pooled within-batch standard deviation of the MDL test sample type.

$$\boldsymbol{s_w = \sqrt{M_0}}$$

where:

- M_0 is the within-batch mean square (also known as the pooled estimate of within-batch variance).

$$\boldsymbol{M_0 = \sum_{i=1}^{m} \frac{{s_i}^2}{m}}$$

where:

- s_i is the standard deviation of an individual batch.
- m is the total number of batches.

NOTE 1: Before accepting the calculated method detection limit, it must be ensured that s_w is calculated from data consisting of final sample concentrations i.e. recovery corrected (where applicable). Data used in MDL calculations must **not** be blank corrected.

NOTE 2: Quoted MDL values must always be reported in the same units as the determinand represented. The calculated MDL value for a determinand may be rounded up for convenience and ease of use.

B.3.2 Worked example

Table B1 – '11×2' MDL test sample results (ng/kg)

Batch No.	Replicate 1 (1)	Replicate 2 (1)	Within-batch st. dev (s_i)	Within-batch variance (s_i^2)
1	22.92	23.38	0.32527	0.1058
2	21.81	22.27	0.32527	0.1058
3	23.99	23.68	0.21920	0.04805
4	23.23	22.77	0.32527	0.1058
5	23.39	22.78	0.43134	0.18605
6	22.57	22.90	0.23335	0.05445
7	22.25	21.22	0.72832	0.53045
8	22.70	22.15	0.38891	0.15125
9	22.83	22.48	0.24749	0.06125
10	22.74	23.78	0.73539	0.5408
11	23.39	21.87	1.07480	1.1552

1. Final sample concentrations. Recovery corrected; not blank corrected.

Applying the theory previously outlined in B.3.1 to the example test sample data from Table B1, above, produces the following results:

$$M_0 = \sum_{i=1}^{m} \frac{{s_i}^2}{m} = \frac{3.0449}{11} = 0.2768$$

$$s_w = \sqrt{M_0} = \sqrt{0.2768} = 0.5261$$

$$MDL = 4.65 \times s_w = 4.65 \times 0.5261 = 2.45 \text{ ng/kg}$$

Annex C

Assessment of precision and bias

C.1 Assessment of precision

Two separate comparisons must be made as part of the overall precision assessment:

- Comparison of within-batch and between-batch variance.
- Comparison of measured and target precision (%RSD).

The outcome of both comparisons must be acceptable in order for measured precision to be considered satisfactory.

C.1.1 Comparison of within-batch and between-batch variance

This comparison assesses whether a significant difference exists between observed within-batch and between-batch variances for each validation test sample type.

a. In practice, this first requires the calculation of the within-batch and between-batch mean squares, M_0 and M_1 respectively:

$$\boldsymbol{M_O = \sum_{i=1}^{m} \frac{s_i^2}{m}}$$

where:

- s_i is the standard deviation of an individual batch.
- m is the total number of batches.

$$\boldsymbol{M_1 = n.s_{bm}^2}$$

where:

- s_{bm} is the standard deviation of the batch means.
- n is the number of replicates in each batch.

b. A two-tailed F-test at the 95% confidence interval (see NOTE 2, below) is then applied to determine whether there is a statistically significant difference between the calculated variances:

$$\boldsymbol{F_{(obs)} = \frac{\sigma_1}{\sigma_2}}$$

where:

- σ_1 and σ_2 are the within-batch and between-batch mean squares (M_0 and M_1 respectively), assigned by $\sigma_1 > \sigma_2$ (see NOTE 1, below).

NOTE 1: In a two-tailed F-test the highest variance must always be used as the numerator when calculating the observed F value ($F_{(obs)}$) in order to ensure a result greater than one.

NOTE 2: Use of a two-tailed F-test requires that the significance level is halved when determining the critical value of F ($F_{(crit)}$) i.e. for this performance standard α = 0.025.

c. In determining the critical value of F ($F_{(crit)}$), degrees of freedom for each variance are to be calculated as follows:

within-batch (M_0): $\boldsymbol{df = m(n-1)}$

between-batch (M_1): $\boldsymbol{df = m-1}$

where:

- m is the total number of batches.
- n is the number of replicates in each batch.

d. There are three possible outcomes:

i. No significant difference exists between M_0 and M_1 (i.e. $F_{(obs)} \leq F_{(crit)}$) - this is considered a pass.

ii. M_1 is significantly greater than M_0 (i.e. $F_{(obs)} > F_{(crit)}$; and between-batch variance > within-batch variance) - this is a common situation in many methods and may also be considered a pass, providing the target %RSD is also met (see C.5.2).

iii. M_0 is significantly greater than M_1 (i.e. $F_{(obs)} > F_{(crit)}$; and within-batch variance > between-batch variance) - this is considered a failure and is indicative of a potential problem with the method. The laboratory must investigate, assess, and perform additional method development and/or repeat the validation exercise as required.

NOTE: It is recognised that in exceptional circumstance M_0 may be significantly greater than M_1, but method performance cannot be further improved by additional development (e.g. when total standard deviation (s_t) is very low). In such instances, the laboratory may conform with the requirements of this performance standard provided that both the target %RSD is met, and they are able to justify their acceptance of the validation data to SEPA.

C.1.2 **Comparison of measured and target precision (%RSD)**

This comparison assesses whether the measured precision, expressed as percent relative standard deviation (%RSD), meets the required target precision (%RSD) detailed in Annex A.

C.1.2.1 **Calculation of measured precision (%RSD)**

a. By manipulating the mean square values obtained from ANOVA (see C.1.1) using the calculation detailed below, an estimate of total standard deviation (s_t) will be made for each validation test sample type:

$$\boldsymbol{s_t = \sqrt{\frac{(M_1 + (n-1)M_0)}{n}}}$$

where:

- M_0 is the within-batch mean square.
- M_1 is the between-batch mean square.
- n is the number of replicates in each batch.

b. The measured %RSD of each test sample type may then be calculated as follows:

$$\boldsymbol{\%RSD = \frac{s_t}{\bar{x}} \times 100}$$

where:

- s_t is the total standard deviation.
- $\bar{x}$ is the mean of results.

c. The measured %RSD for each validation test sample type must then be assessed against the appropriate target %RSD detailed in Annex A.

If the measured value is less than or equal to the target value, the required precision has been achieved, performance is considered satisfactory, and no further action is required.

If the measured value is greater than the target value, it is still possible to comply with the requirements of this performance standard if statistical significance testing indicates that the exceedance is not significant (see C.1.2.2).

C.1.2.2 **Significance testing of precision (%RSD)**

a. A one-tailed F-test at the 95% confidence interval (α = 0.05) is applied to determine whether the difference between the measured precision (%RSD) and the target precision (%RSD) is statistically significant:

$$\boldsymbol{F_{(obs)} = \frac{{s_t}^2}{{Z_p}^2}}$$

where:

- s_t is the measured total standard deviation.
- Z_p is the target standard deviation.

b. The target standard deviation (Z_p) can be calculated from both the MACS target %RSD and the operator's target MDL.

The value used when determining the observed F value ($F_{(obs)}$) will be whichever of the two calculated Z_p values below is the greater:

$$\boldsymbol{Z_p = \bar{x} \times \frac{target\ \%RSD}{100}}$$

or

$$\boldsymbol{Z_p = \frac{target\ MDL}{4}}$$

where:

- $\bar{x}$ is the mean of results.

c. In determining the critical value of F ($F_{(crit)}$), an estimated number of degrees of freedom for s_t are to be calculated as follows, with the final value rounded to the nearest whole number:

$$\boldsymbol{df = \frac{m(m-1)(M_1 + (n-1)M_0)^2}{mM_1{}^2 + (m-1)(n-1)M_0{}^2}}$$

where:

- M_0 is the within-batch mean square.
- M_1 is the between-batch mean square.
- m is the total number of batches.
- n is the number of replicates in each batch.

Degrees of freedom for Z_p are infinite, although for calculation purposes a value of $\geq 10^{10}$ is considered sufficient for the requirements of this performance standard.

d. There are two possible outcomes:

i. The measured precision **is not** significantly greater than the target precision (i.e. $F_{(obs)} \leq F_{(crit)}$) - this is considered a pass; the required precision has been achieved and performance is considered satisfactory.

ii. The measured precision **is** significantly greater than the target precision (i.e. $F_{(obs)} > F_{(crit)}$) - this is considered a failure; the required precision has not been achieved and performance is not considered satisfactory.

C.2 Assessment of bias (systematic error)

An assessment of bias, or systematic error, need only be made if the assessment of precision (see C.1) has proved acceptable.

C.2.1 Comparison of measured and target bias

This comparison assesses whether the measured bias, expressed as a percentage (%Bias), meets the required target %Bias detailed in Annex A.

Assessment of measured %Bias for a method is based on the difference of the actual mean of results from a 'true' or expected concentration. This assessment may be made using data generated from the analysis of reference materials or from the results of spiked/unspiked sample matrix pairs.

C.2.1.1 Calculation of measured bias

a. The theory behind the calculation of measured %Bias is identical regardless of whether the analysis of reference materials or the results of spiked/unspiked pairs are used:

$$\%\boldsymbol{Bias} = \frac{(\bar{\boldsymbol{x}} - \boldsymbol{E})}{\boldsymbol{E}} \times \mathbf{100}$$

where:

- $\bar{x}$ is the mean of results.
- E is the expected, or 'true' concentration.

b. It is important to note that the expected concentration (E) used in the calculation above is defined differently depending on which experimental approach is used.

- Where analysis of reference materials has been used to generate the result mean, the expected concentration is the accepted reference value of the material(s) used.

- Where spiked/unspiked pairs have been used to generate the result mean, the expected concentration is the expected final result based on the amount of material added to the spiked sample.

c. The measured %Bias for each validation test sample type must then be assessed against the appropriate target %Bias detailed in Annex A.

If the measured value is less than or equal to the target value, the required bias has been achieved, performance is considered satisfactory, and no further action is required.

If the measured value is greater than the target value, it is still possible to comply with the requirements of this performance standard if statistical significance testing indicates that the exceedance is not significant (see C.2.2).

C.2.2 Significance testing of bias

a. A one-tailed t-test at the 95% confidence interval (α = 0.05) is applied to determine whether the difference between the measured bias (expressed as a concentration) and the target bias is statistically significant:

$$\boldsymbol{t_{(obs)} = \frac{|(|measured\ bias| - Z_b)|}{SE}}$$

where:

- Z_b is the target bias (expressed as a concentration).
- SE is the standard error of the batch means, calculated as:

$$\boldsymbol{SE = \frac{s_{bm}}{\sqrt{m}}}$$

where:

- s_{bm} is the standard deviation of the batch means.
- m is the total number of batches.

NOTE: The symbol $|measured\ bias|$ signifies the value of measured bias regardless of sign. Likewise, the symbol $|(|measured\ bias| - Z_b)|$ signifies the value of $(|measured\ bias| - Z_b)$ regardless of sign.

b. The target bias (Z_b) can be calculated from both the MACS target %Bias and the operator's target MDL.

The value used when determining the observed t value ($t_{(obs)}$) will be whichever of the two calculated Z_b values below is the greater:

$$\boldsymbol{Z_b = E \times \frac{target\ \%Bias}{100}}$$

or

$$\boldsymbol{Z_b = \frac{target\ MDL}{2}}$$

where:

- E is the expected, or 'true' concentration.

c. In determining the critical value of t ($t_{(crit)}$), degrees of freedom are to be calculated as follows:

$$\boldsymbol{df = m - 1}$$

where:

- m is the total number of batches.

d. There are two possible outcomes:

i. The measured bias **is not** significantly different from the target bias (i.e. $t_{(obs)} \leq t_{(crit)}$) - this is considered a pass; the required bias has been achieved and performance is considered satisfactory.

ii. The measured bias **is** significantly different from the target bias (i.e. $t_{(obs)} > t_{(crit)}$) - this is considered a failure; the required bias has not been achieved and performance is not considered satisfactory.

C.3 Worked example

C.3.1 The following example is presented to demonstrate the application of the theory, statistical tests and assessments described above.

It considers a hypothetical 11×2 validation exercise of a pre-treatment method determinand with the following minimum performance criteria:

- Precision (%RSD) target: 25%
- Bias target: ±50%
- Required MDL: 0.5 ng/kg

NOTE 1: The test sample results used in Table C1 have been generated manually for illustrative purposes only, and do not represent real analytical validation data.

NOTE 2: Although not explicitly shown, results are intended to be representative of final sample concentrations (i.e. blank and recovery corrected). In situations where the 'Unspiked' sample matrix result is greater than that of the 'Spiked' sample matrix, the calculated 'Spiked minus Unspiked' result is defaulted to zero.

Table C1 – '11×2' validation results

Batch	Replicate	CRM	Unspiked Sample Matrix	10% Method Range		90% Method Range	
				Spiked Sample Matrix	Spiked minus Unspiked	Spiked Sample Matrix	Spiked minus Unspiked
1	1	43.231	4.133	7.148	3.015	106.532	102.399
	2	43.556	4.550	6.216	1.666	99.024	94.474
batch mean	$\bar{x}_i$	43.3935	4.3415	6.6820	2.3405	102.7780	98.4365
within batch st dev	s_i	0.22981	-	-	0.9539	-	5.6038
within batch variance	s_i^2	0.0528	-	-	0.9099	-	31.4028
2	1	43.086	4.688	6.638	1.950	102.494	97.806
	2	39.914	4.376	8.489	4.113	100.745	96.369
batch mean	$\bar{x}_i$	41.5000	4.5320	7.5635	3.0315	101.6195	97.0875
within batch st dev	s_i	2.24294	-	-	1.5295	-	1.0161
within batch variance	s_i^2	5.0308	-	-	2.3393	-	1.0325
3	1	46.674	4.560	6.204	1.644	105.252	100.692
	2	45.165	4.417	6.100	1.683	106.048	101.631
batch mean	$\bar{x}_i$	45.9195	4.4885	6.1520	1.6635	105.6500	101.1615
within batch st dev	s_i	1.06702	-	-	0.0276	-	0.6640
within batch variance	s_i^2	1.1385	-	-	0.0008	-	0.4409
4	1	45.585	4.770	5.520	0.750	103.164	98.394
	2	37.062	4.564	6.331	1.767	104.587	100.023
batch mean	$\bar{x}_i$	41.3235	4.6670	5.9255	1.2585	103.8755	99.2085
within batch st dev	s_i	6.02667	-	-	0.7191	-	1.1519
within batch variance	s_i^2	36.3208	-	-	0.5171	-	1.3268
5	1	44.693	5.189	5.641	0.452	104.353	99.164
	2	45.247	5.882	5.470	0.000	99.958	94.076
batch mean	$\bar{x}_i$	44.9700	5.5355	5.5555	0.2260	102.1555	96.6200
within batch st dev	s_i	0.39174	-	-	0.3196	-	3.5978
within batch variance	s_i^2	0.1535	-	-	0.1022	-	12.9439
6	1	50.017	5.055	5.742	0.687	104.130	99.075
	2	46.385	5.720	5.136	0.000	101.544	95.824
batch mean	$\bar{x}_i$	48.2010	5.3875	5.4390	0.3435	102.8370	97.4495
within batch st dev	s_i	2.56821	-	-	0.4858	-	2.2988
within batch variance	s_i^2	6.5957	-	-	0.2360	-	5.2845
7	1	46.369	4.239	7.153	2.914	102.721	98.482
	2	44.948	4.678	6.638	1.960	104.978	100.300
batch mean	$\bar{x}_i$	45.6585	4.4585	6.8955	2.4370	103.8495	99.3910
within batch st dev	s_i	1.00480	-	-	0.6746	-	1.2855
within batch variance	s_i^2	1.0096	-	-	0.4551	-	1.6526
8	1	42.043	5.271	6.383	1.112	104.735	99.464
	2	42.905	5.310	5.604	0.294	99.948	94.638
batch mean	$\bar{x}_i$	42.4740	5.2905	5.9935	0.7030	102.3415	97.0510
within batch st dev	s_i	0.60953	-	-	0.5784	-	3.4125
within batch variance	s_i^2	0.3715	-	-	0.3346	-	11.6451
9	1	50.800	4.501	5.783	1.282	104.087	99.586
	2	49.954	5.149	5.017	0.000	96.457	91.308
batch mean	$\bar{x}_i$	50.3770	4.8250	5.4000	0.6410	100.2720	95.4470
within batch st dev	s_i	0.59821	-	-	0.9065	-	5.8534
within batch variance	s_i^2	0.3579	-	-	0.8218	-	34.2626
10	1	47.608	4.802	7.066	2.264	100.738	95.936
	2	46.678	4.920	5.832	0.912	98.436	93.516
batch mean	$\bar{x}_i$	47.1430	4.8610	6.4490	1.5880	99.5870	94.7260
within batch st dev	s_i	0.65761	-	-	0.9560	-	1.7112
within batch variance	s_i^2	0.4325	-	-	0.9140	-	2.9282
11	1	45.255	5.172	5.614	0.442	98.134	92.962
	2	41.990	5.277	6.700	1.423	98.164	92.887
batch mean	$\bar{x}_i$	43.6225	5.2245	6.1570	0.9325	98.1490	92.9245
within batch st dev	s_i	2.30870	-	-	0.6937	-	0.0530
within batch variance	s_i^2	5.3301	-	-	0.4812	-	0.0028
expected concentration (wt/wt)	E	50	-	-	10	-	90
mean	$\bar{x}$	44.962	4.874	6.201	1.379	102.101	97.228
%bias		-10.076	-	-	-86.214	-	8.031
within-batch mean square	M_0	5.1631	-	-	0.6465	-	9.3566
between-batch mean square	M_1	16.3282	-	-	1.7033	-	10.7911
st dev of batch means	s_{bm}	2.8573	-	-	0.9228	-	2.3228
standard error of batch means	SE	0.8615	-	-	0.2782	-	0.7004
total st dev	s_t	3.2780	-	-	1.0839	-	3.1739

total number of batches	m	11
no. of replicates in each batch	n	2

C.3.2 Applying the protocols for assessment of precision and bias previously outlined in C.1 and C.2 to the corrected test sample results presented in Table C1 produces the statistical summary in Table C2, below:

Table C2 – Summary statistics

		CRM	Spiked minus Unpiked		
			10% Method Range	90% Method Range	
between-batch mean square	M_1	16.3282	1.7033	10.7911	Precision - ANOVA
within-batch mean square	M_0	5.1631	0.6465	9.3566	
between-batch degrees of freedom	d.f. (M_1)	11	11	11	
within-batch degrees of freedom	d.f. (M_0)	10	10	10	
observed F value	$F_{(obs)}$	3.162	2.635	1.153	
critical F value	$F_{(crit)}$	3.665	3.665	3.665	
significant?		N.S.	N.S.	N.S.	
assessment		PASS	PASS	PASS	
mean	$\bar{x}$	44.9620	1.3786	97.2275	Precision - %RSD
total st dev	s_t	3.278	1.084	3.174	
measured relative st dev	%RSD	7.29	78.62	3.26	
st dev from target %RSD		N/A	0.345	N/A	
st dev from target MDL		N/A	0.125	N/A	
target st dev	Z_p	N/A	0.345	N/A	
estimated degrees of freedom (s_t)	d.f. (s_t)	N/A	17	N/A	
observed F value	$F_{(obs)}$	N/A	9.891	N/A	
critical F value	$F_{(crit)}$	N/A	1.644	N/A	
significant?		N/A	*	N/A	
assessment		PASS	FAIL	PASS	
measured %bias		-10.08	-86.21	8.03	%Bias
measured bias (conc.)		N/A	-8.621	N/A	
bias (conc.) from target bias		N/A	5.000	N/A	
bias (conc.) from target MDL		N/A	0.250	N/A	
target bias (conc.)	Z_b	N/A	5.000	N/A	
degrees of freedom	d.f.	N/A	10	N/A	
observed t value	$t_{(obs)}$	N/A	13.015	N/A	
critical t value	$t_{(crit)}$	N/A	1.812	N/A	
significant?		N/A	*	N/A	
assessment		PASS	FAIL	PASS	

N/A	significance testing not applicable
N.S.	not significant
*	significant at the 0.05 level

In this example, the summary statistics are interpreted as follows:

a. **'*Precision - ANOVA*' assessment**

Comparison of within-batch and between-batch variance is acceptable for the CRM, 10% Spiked minus Unspiked results and 90% Spiked minus Unspiked results.

b. **'*Precision - %RSD*' assessment**

The required target precision is met for the CRM (7.59%) and the 90% Spiked minus Unspiked results (3.26%), so significance testing is not necessary.

The measured %RSD of the 10% Spiked minus Unspiked results does not meet the required target and is found to be significantly different once an F-test is performed (i.e. $F_{(obs)} > F_{(crit)}$). As a result, performance is not considered acceptable for this test type.

c. **'*%Bias*' assessment**

The required target bias is met for the CRM (-10.08%) and the 90% Spiked minus Unspiked results (8.03%). Significance testing is not necessary.

The measured %Bias of the 10% Spiked minus Unspiked results (-86.21%) does not meet the required target and is found to be significantly different once a t-test is performed (i.e. $t_{(obs)} > t_{(crit)}$). As a result, performance is not considered acceptable for this test type.

NOTE: Had the 10% Spiked minus Unspiked results in this example been generated from a real validation, assessment of bias would not be required as the precision assessment has already been deemed unsatisfactory (see C.2). Bias assessment has been performed in this case for indicative purposes only.

For information on accessing this document in an alternative format or language please either contact SEPA by telephone on 03000 99 66 99 or by email to equalities@sepa.org.uk

If you are a user of British Sign Language (BSL) the Contact Scotland BSL service gives you access to an online interpreter enabling you to communicate with us using sign language.

http://contactscotland-bsl.org/

www.sepa.org.uk

03000 99 66 99

Strathallan House, Castle Business Park, Stirling, FK9 4TZ