Schedule for a proficiency test

Once a laboratory has registered the year's pro-gramme the interaction between the laboratory and Eurofins Miljø will be according to the schedule described below.

1   Reminder

Approximately one month before a proficiency test, Eurofins Miljø sends a reminder for all registered participating laboratories. The reminder advises the laboratory of the expected arrival date for samples.

2   Samples

Whenever possible, samples are made up as sample pairs (split-level samples). A sample pair consists of two samples that are identical or almost identical. The difference between the two samples in a sample pair is a small and known (to Eurofins Miljø but not to the participants) difference in con-centration of the parameters to be measured. The purpose of having this small difference is to ensure that the participants do not inadvertently reduce the difference between two results of a duplicate meas-urement, thereby appearing to have repeatability, which is too optimistic. Before statistical treatment of the data the small, added concentration (spike) is removed from the results, leaving data that can be regarded as a duplicate measurement.

3   Shipment of samples

Samples are shipped in a way to allow arrival at the participating laboratories 2-3 days before the proficiency test. The samples are accompanied by an instruction on sample handling, sample identi-fication and where appropriate the procedure for sample preparation from concentrates.

A list of identification numbers used for analytical methods are also sent. Eurofins Miljø requests that the laboratories report the method employed using these identification numbers. The information of the methods used is not important for the proficiency test as such, but when laboratories identify the methods used it becomes possible to extract infor-mation of comparability of methods from the data.

The list of method identification numbers includes methods used by many laboratories and a code for "other methods". If a laboratory uses the number for "other methods" it will aid the development of the system if the precise reference for the method is given. Thereby new methods may be added to the list of methods and the new method may be compared to others at an early stage. Comparison of methods is financed by sources other than the proficiency tests.

4   Analysis

The participants analyse the parameters in the samples according to the registration of the labora-tory. The analysis should be performed using the same procedures as for all other samples received by the laboratory. Calibration, blank correction, and correction for recovery is performed exactly as the laboratory's routine procedures. Occasionally, Eurofins Miljø may include a sample to be used for blank correction with proficiency test samples. If this is the case the instruction accompanying the samples will show this clearly.

5   Reporting of results

The laboratory submits results online by logging on to the Eurofins Proficiency Testing Portal. In order to perform valid statistical data analyses, the laboratories are requested to report results using one significant figure over the number usually employed in similar routine analyses. The results must be reported no later than the date shown in the portal as the deadline. Every participant is assigned a code number, so that all results are processed anonymously. The deadline is normally approximately two weeks after the proficiency testing date. Results that arrive after the deadline are normally not included in the data analysis or reports from Eurofins Miljø.

It is important to everyone - the participants as well as Eurofins Miljø - that the results

- Are reported on time. Late arriving results will mean either delay in statistical data analysis or the late arriving results are not included. This means that the statistical basis for the proficiency test becomes less substantial than intended.

- Are reported using the units specified.

- Are reported showing the method identification number or precise reference if the method is not included in the list of method identification numbers.

It is not possible to correct data that have already been reported.

6 Report from Eurofins Miljø A/S

6.1 Participants' results

Approximately three weeks after the deadline for reporting results, Eurofins Miljø issues a report to each participant. The report contains the laboratory's own submitted results. Participants are requested to check that their results are entered correctly.

The report also includes an assigned value for the concentration of all parameters and an evaluation of the stability and homogeneity of the samples. This evaluation shows the degree to which the samples are considered adequate for evaluating the performance of participants.

The report contains a comparison of the laboratory's data to the average for all participants, giving an assessment of the performance relative to other participants.

The deviation of the laboratory's data from the assigned value and the standard deviation (measure of repeatability) are estimated. The laboratory may use these values for assessment against official requirements or the laboratory's own specifications of quality. In the report, result are compared to quality requirements in Danish legislation.

6.2 Overview of all data and general quality of analyses

The report also includes a table showing the results from all participants for every parameter. Thereby the laboratory may compare its own results to all other data or to results from other laboratories using the same method.

For every parameter a page is included showing the statistical parameters for the general quality of analyses and a plot showing all data graphically. The statistical data analysis is performed according to til DS/ISO 5725:2019, Accuracy (trueness and precision) of measurement methods and results. Section 8 gives a summary of the statistical estimations.

A Youden plot is used to illustrate the data. The plot shows the results of one sample as a function of the result of the other sample in a sample pair. The Youden plot gives an immediate idea of the dominating sources of error in the results. Laboratories having results in the upper left or lower

right hand corner of the diagram have analyses dominated by random error. On the other hand, laboratories having results close to the 45° line shown in the plot, but far away from the assigned value have results dominated by systematic error.

Youden plot,
NH4-N, µg/L N sample pair AB
Spike value 0 of B

Laboratories omitted from the above Youden plot : 6 17
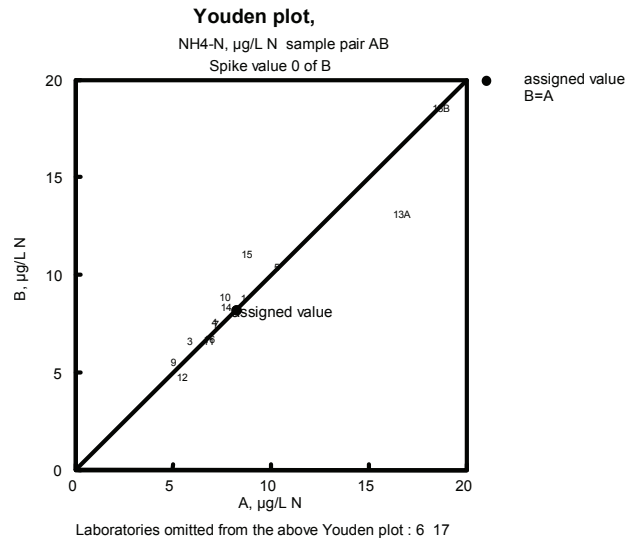
An example of a Youden plot is shown above.

6.3 Assigned value

Eurofins Miljø assigns a value for each parameter. The assigned value is Eurofins Miljø best estimate of the true concentration of the sample. This is the value used to estimate the accuracy of participants' results. The principles used to assign values to the samples are described below.

Assigned values in synthetic samples are calculated from the weight of chemicals and water used to prepare the samples.

Assigned values in natural samples are consensus median values either from all participants or from a limited number of participants who have been selected based on good accuracy in two or more previous proficiency tests of the same parameter. For certain parameters, such as for example biochemical oxygen demand, the Danish Ministry of Environment, regulation on quality requirements for environmental measurements, specifies the method. In such cases only results from participants using the specified method are included in the assigned value.

Eurofins Miljø report to participants describes how the assigned value was determined for all parameters.

On rare occasions the structure of the data may be such that an estimated assigned value would be extremely uncertain. In such cases no assigned value is estimated. This situation may arise in case

several methods giving incomparable results are used, in case of few participants or large variability.

7 Quality assurance of the proficiency test

Eurofins Miljø gives high priority to quality assurance and documentation of the quality of proficiency tests in order that participants, accreditation bodies and other users of the proficiency tests may be confident that the results from a participant adequately reflect the participant's performance.
Therefore, Eurofins Miljø has an extensive quality system for performance of proficiency tests and a well-validated software programme for statistical data analysis. Eurofins Miljø constantly looks for improvement by standardising and simplifying routines whenever possible.

Eurofins Miljø also performs measurements and tests to control and document the performance of the proficiency test.

The documentation for preparation and control of proficiency test samples is presented in a report: "Quality Documentation". This report is available to all interested parties on the Eurofins Miljø webpage (www.eurofins.dk/proficiencytest) and Eurofins Proficiency Testing Portal approximately six weeks after the deadline for submitting participants' results.

7.1 Sample preparation
Sample preparation is documented in work sheets on which every step in the preparation process is documented and signed by the staff member who performed the step in question as well as a control. Sample preparation is summarised in a table including the concentrations calculated from formulation. The table is included in Eurofins Miljø Quality Documentation report.

7.2 Homogeneity and stability
All samples distributed in proficiency tests are checked by a programme of analysis to ensure that samples are homogeneous and stable. The analyses include those parameters in each proficiency test posing the greatest risk for instability or contamination that may result in heterogeneity. The programme of analyses takes into account that the statistical tests performed must have optimal power. This reduces the risk of erroneously accepting a batch of samples of inferior quality. The programme is based on Eurofins Miljø many year of experience with the proficiency tests offered.

All measurement of the control programme as well as the statistical analyses used to conclude on homogeneity or stability and the conclusions of the tests are documented in Eurofins Miljø Quality Documentation report.

Homogenety and stability are evaluated using the target standard deviation for reproducibility $\hat{\sigma}$ we use 0.5 times * the required maximum deviation from the assigned value stipulated in the Danish EPA regulation on quality reguirements for environmental measurements.

7.3 Spike
In most cases samples are prepared as split-level samples, that is sample pairs where one may have received a spike of a small amount of the parameter in question. When calculating repeatability from sample pairs, after removing by calculation the small spike added, it is necessary that the true spike is identical to the spike calculated.

To test this, the average difference between results of the two samples in a sample pair is calculated for each participant after removing the added spike. A t-test at 95% confidence level is performed to see if the calculated difference deviates significantly from the expected value of zero. If the difference found by participants differs significantly from zero it is concluded that participants do not measure the spike calculated.

In such cases the sample preparation is scrutinised to detect any possible error in preparation. The spike measured by Eurofins Miljøs control programme is calculated and checked, and the participants' results are scrutinised to ensure that obvious outliers do not skew the average difference. On this basis it is decided if the spike calculated is maintained. A recommendation is given in case it is deemed necessary to increase the range of requirements for repeatability due to the uncertainty of the size of the spike. All data, conclusions and evaluations are documented in the report: Quality Documentation.

7.4 Assigned value
A reliable assignment of the concentration of parameters in proficiency test samples is essential for assessment of the performance of participants. The overall average of all participants' results is compared to the assigned value using a t-test at the 95% confidence level. If the test is significant it is concluded that the participants do not measure an average value, which is identical to the assigned value.

In case the test shows significance the results of the participants is scrutinised because a significant difference in most cases is due to outliers or outlying methods. If this is the case, the outlying results are excluded manually and the test repeated.

If the difference cannot be attributed to outlying data, the sample preparation and Eurofins Miljøs control measurements are scrutinised in an attempt to identify the cause of the deviation. On this basis it is concluded if the assigned value is to be main-

tained or changed. A recommendation is given in case it is deemed necessary to increase the range of requirements for accuracy due to the uncertainty of the assigned value. All data, conclusions and evaluations are documented in the report: Quality Documentation.

8 Statistical data analysis

Statistical data analysis is performed according to til DS/ISO 5725:2019, Accuracy (trueness and precision) of measurement methods and results.

8.1 Outlier tests

Cochran's test is performed to identify any data set with an unusually large standard deviation. This is followed by Grubbs' test to identify any data set with an average result, which deviates from the overall average for all participants. Any data set which deviates significantly at the 99% confidence level (outliers) in these tests, are excluded from further data analysis. Data sets, which deviate only at the 95% confidence level (stragglers) are retained.

An outlier is a result, which deviates significantly from the main body of data, either because of unusually large standard deviation (Cochran) or because of a deviating average (Grubbs). An outlier test does <u>not</u> take into consideration requirements for quality.

8.2 General quality of analyses

The general quality of analyses is found by analysis of variance (ANOVA) of the results from all data sets. This analysis yields the between laboratories variance, between samples variance and residual variance. The repeatability standard deviation (s_r) and reproducibility standard deviation (s_R) are estimated as follows:

s_r^2 = residual variance
s_L^2 = ½ * (between laboratories variance - residual variance)
$s_R^2 = s_L^2 + s_r^2$

Furthermore the overall average and median for all participants' data are estimated. Comparison of the average and the median to the assigned value yields an estimate of the overall trueness of the results from all participants.

Statistical analyses for proficiency tests having very few participants may be by robust statistics.

9 Symbols and abbreviations

Tables

<	"Less than" - results are not included in calculations
U, UL, **	Manually excluded results
UC, C**	Cochran outlier. Results are not included in statistical data analysis
UG, G**	Grubbs outlier. Results are not included in statistical data analysis

Graphical presentations

<	"Less than" - results are not included in the graphical presentations
U, UL, **	Manually excluded results are not included in the graphical presentations
•	Assigned value (= "true" value)

Statistical symbols

µ	Assigned value
p	Number of laboratories included in statistical data analysis. This corresponds to the total number of reported results minus manually excluded results (**) minus Cochran outliers (UC) and minus Grubbs outliers (UG)
n	Number of results included in statistical data analysis
m	Average of all not-excluded results
M	Median
d	The average difference between the two results of a sample pair after subtraction of the added spike
t	Test value in Student's t-test
p	The level of probability in a statistical test
s	Standard deviation
F	Test value in an F-test
s_i	Standard deviation for participant "i", corresponding to the variability by repeated measurement of a sample
s_r	Standard deviation within a laboratory
s_r^2	Repeatability variance
s_L	Standard deviation between laboratories
s_L^2	Laboratory variance
s_R	Reproducibility standard deviation
s_R^2	Reproducibility variance $s_R^2 = s_r^2 + s_L^2$
CV_r	Coefficient of variation within a laboratory $(s_r \cdot 100)/µ$
CV_R	Total coefficient of variation $(s_R \cdot 100)/µ$