

CCIL Protocol for Estimating Measurement Uncertainty Using QC Data (Type A)

Mark Hugdahl, Technical Manager, ALS Environmental (Vancouver),
& the CCIL Committee on Measurement Uncertainty.
Version 1.0 (June 23, 2003)

Introduction to Recommended Type A Protocol:

This document describes a protocol to estimate the Measurement Uncertainty (MU) of a test method by a Type A statistical evaluation of QC data. It is designed for environmental analytical tests, although it may be appropriate for other related testing disciplines as well. Requisite QC data may have been generated during method validation studies, or may be part of the laboratory's routine QC program.

The Type A approach to uncertainty estimation assumes that statistical analysis of laboratory QC data can provide a reasonable estimate of measurement uncertainty provided several criteria are met:

- The analyses have been carried out over an extended period of time and involve several combinations of operators and equipment. This will ensure that the data has been obtained under internal reproducibility conditions.
- The measurement procedure is stable and has remained in statistical control during the period of the analysis.
- All variables and processes that contribute significantly to the uncertainty of the test have been captured by the selected QC data. Any contributing factors that are not captured by QC must be considered separately.

This protocol is applicable to tests where precision varies in approximate proportion to concentration at levels above approximately ten times the Method Detection Limit (MDL)*. In such cases, uncertainties at levels \gg MDL may be expressed or calculated as a linear function of concentration, using expanded relative standard uncertainties. As concentrations approach zero, relative uncertainties become exponentially greater due to constant contributions to uncertainty represented by the MDL, and sometimes also by factors that contribute to positive method blank values. This model addresses both of these relationships with an equation (Equation 2) that describes the dependence of uncertainty on three factors:

- i) An assumed-constant relative uncertainty associated with higher analyte concentrations (it is recommended that this assumption be verified during method validation studies).
- ii) A constant contribution to uncertainty attributable to S_0 , the standard deviation of the method for concentrations that approach zero (S_0 is related to the MDL).

* For tests that meet this criteria, the protocol is applicable to the estimation of uncertainties at any concentration within the validated range of the method, including concentrations below 10x MDL.

- iii) A constant contribution to uncertainty attributable to the mean long-term method blank value (where significant).

The model is primarily adapted from procedures described in the Eurachem/CITAC Guide "Quantifying Uncertainty in Analytical Measurement, 2nd edition" (Appendix E.4: Documenting uncertainty dependent on analyte level)¹.

An important premise of this procedure is that QC data from laboratory duplicate or split samples, where they are possible and defined, must be considered in order to obtain a reasonable estimate of measurement uncertainty for environmental samples. This is because, for many environmental chemistry tests, the variability due to sub-sampling and sample heterogeneity tends to dominate over all other sources of variability. This is especially true for most analyses of organic compounds, and for tests of soil or sediment that do not involve extensive pre-treatments to guarantee homogenization.

Laboratories can only estimate and report *analytical uncertainty*, which excludes factors outside their control. It is important to note that analytical uncertainty differs from *total measurement uncertainty*, which includes contributions to uncertainty from the sampling component of a test. Michael Ramsey² states that "analytical variance should ideally contribute less than 20% to the measurement variance, if the measurement uncertainty is not to be limited by the analytical component." Since variance is the square of standard deviation, and since uncertainties are directly proportional to standard deviation, it can be shown that analytical uncertainties should ideally be less than $1/\sqrt{5}$ (or 45%) of the data quality objective for total measurement uncertainty. If this is unattainable, the analytical uncertainty will impose limits to total measurement uncertainty, which is likely to be the case for certain difficult and imprecise tests.

Another premise of this procedure is that the potential degradation of analytes that could occur between the times of sampling and analysis are not normally considered, provided that samples are analyzed within referenced holding times, and that they have been appropriately preserved and stored. We assume that there is sufficient scientific basis behind published holding time, storage, and preservation guidelines to prevent significant analyte degradation from occurring when they are met.

The way in which method bias is treated has a direct impact on how measurement uncertainty is estimated. The Eurachem guide contains the following recommendations on bias in section 7.15.2:

Where the bias is significant compared to the combined uncertainty, additional action is required. Appropriate actions might:

- Eliminate or correct for the bias, making due allowance for the uncertainty of the correction.
- Report the observed bias and its uncertainty in addition to the result.

As a general metrological principal, known systematic errors that contribute to method bias should be eliminated, or corrections should be made to account for them. However, in environmental chemistry, there are many examples where method biases may be present, but where correction factors cannot reliably be applied, because these biases tend to be inconsistent and poorly characterized. A common form of method bias is due to partial (and variable) losses of analyte incurred during sample processing steps within a test procedure. These and other causes of bias are often dependent on the composition of individual samples.

This protocol assumes that method bias, if present, is not corrected for, but is instead reported as a separate entity, independent from estimates of measurement uncertainty. Note that there is also an uncertainty associated with measures of bias. Where possible, bias should be reported together with its uncertainty.

In situations where bias corrections are applied as a component of a method, an additional component of bias uncertainty must be added to the estimation of measurement uncertainty to account for the error inherent in the correction.

As recommended in the Eurachem/CITAC guide, this protocol uses an expansion factor of 2 to achieve uncertainty estimates that approximate 95% confidence intervals. It is impractical to use more exact t-statistics for the final expansion because the combined standard uncertainty is a combination of several measured values, each with potentially different degrees of freedom. We therefore recommend that all estimates of standard deviation be based on large datasets where possible, with an absolute minimum of 6 degrees of freedom.

Uncertainty values obtained from this procedure must be regarded as **estimates**. Primarily, this is because all environmental samples are different, especially with regard to matrix effects and heterogeneity. These sample specific effects tend to occur infrequently, and only with certain samples; often these effects cannot be predicted or detected, and they can vastly increase the error and uncertainty associated with a given result. **It is our intent with this procedure to arrive at an estimate of a 95% confidence level uncertainty value that can be assumed to apply to 95% (or more) of the samples that a laboratory receives for a given test.** It follows that for samples where undetected matrix effects or interferences occur, or for samples that are atypically heterogeneous, uncertainty estimates may be low. This is related to the rationale for the exclusion of outliers in the evaluation of sample duplicate data in Step 3.

It should also be noted that the uncertainty estimates produced by this protocol are symmetrical, and reflect a “plus or minus” range around a measured value. Symmetric uncertainty intervals work well for uncertainties that are relatively low in comparison to the measured result, but become less meaningful for uncertainties that are large in comparison to the result. If the calculated uncertainty estimate exceeds 100% of the measured result, then a value of zero is included within the realm of possible values, which may not be an accurate representation of the situation. In such situations, asymmetrical uncertainty ranges may be more accurate representations of where the true value lies (i.e. an uncertainty interval of [(measured result * 1/3) to (measured result * 3)] might be reported. It is unlikely that many routine analytical test procedures will be subject to such extreme values of uncertainty. This is a recognized limitation of this protocol, which will not be discussed further.

The next two parts of this document contain a summary of the key equations used with this protocol, and a step-wise description of how the protocol can be applied. A spreadsheet accompanying this document contains a detailed worked-example of this protocol using actual data.

Summary of Type A Uncertainty Determination.

Equation 1 - Standard Uncertainty as a Function of Concentration:

$$u(c) = [\sqrt { s_0^2 + (\Theta c)^2 }] + \text{MBlk}_{\text{LT}}$$

where:

- u(c)** = The standard combined uncertainty at concentration c. Represents 1/2 the range of the approximate 68% confidence interval (i.e. one standard deviation) of $c \pm u(c)$.
The units of u(c) are the same as the units of c.
- c** = Measured concentration of analyte in the sample.
- s₀** = A constant contribution to standard uncertainty represented by the standard deviation at zero concentration, which is related to the method detection limit. The units of s₀ must be the same as the units of c.
- Θ** = Combined relative standard uncertainty, excluding MDL and Method Blank contributions.
- MBlk_{LT}** = Mean positive Long-Term Method Blank value, where significant (i.e. if > 1/3 MDL).
The units of the long-term Method Blanks must be the same as the units of c.

Equation 2 - Expanded 95% Uncertainty as a Function of Concentration:

$$U(c) = 2 * [\sqrt { s_0^2 + (\Theta c)^2 }] + \text{MBlk}_{\text{LT}}$$

where:

- U(c)** = The expanded uncertainty at concentration c. Represents one half the range of the approximate 95% confidence interval (i.e. two standard deviations) of $c \pm U(c)$. The units of u(c) are the same as the units of c.
- c** = Measured concentration of analyte in the sample.
- s₀** = A constant contribution to standard uncertainty represented by the standard deviation at zero concentration, which is related to the method detection limit.
- Θ** = Combined relative standard uncertainty, excluding MDL and Method Blank contributions.
- MBlk_{LT}** = Mean positive long-term Method Blank value, where significant (i.e. if > 1/3 MDL).
The units of the long-term Method Blanks must be the same as the units of c.
Note that the Method Blank term is not expanded.

Equation 3 - Determination of Θ :

$$\Theta = \sqrt{[RSD_{DUP}^2 + RSD_{LT}^2 + \dots + RSD_n^2]}$$

where:

- Θ = Combined relative standard uncertainty, excluding MDL and Method Blank contributions.
- RSD_{DUP} = Relative standard deviation of single measurements from within sets of duplicate pairs (applicable at concentrations $> \sim 10x$ MDL).
- RSD_{LT} = Relative standard deviation contribution of long-term components that are not captured by duplicates (at concentrations $> \sim 10x$ MDL).
- RSD_n = Relative standard deviation contribution for any additional uncertainty component(s) (at concentrations $> \sim 10x$ MDL).

Equation 4 - Determination of RSD_{DUP} :

$$RSD_{DUP} = \sqrt{\{[\sum ((a_i - b_i) / \bar{x}_i)^2] / 2N\}}$$

where:

- RSD_{DUP} = Relative Standard Deviation of single measurements from within sets of duplicate pairs.
- $a_i - b_i$ = The difference of the two duplicate values for sample i. Whether or not the absolute value of the difference is used here doesn't affect the outcome.
- \bar{x}_i = The mean value for of the set of duplicate pair results for sample i.
- N = The number of duplicate pairs used in the evaluation.

Note 1: Duplicate pairs where the mean value is below ~ 10 times the MDL must be excluded from the data used to determine this value. The resulting relative standard deviation represents the total relative uncertainty at levels greater than approximately 10 times the MDL.

Note 2: RSD_{DUP} is in decimal-fraction format, not a percentage value (i.e. an RSD of 0.10 is equivalent to 10% RSD).

Equation 5 – The Standard Grubbs Test for Identification of Outliers:

$$T = |(X_n - \bar{X})| / s$$

where:

s = Calculated standard deviation of dataset, including the suspected outlier.

\bar{X} = mean value of dataset.

X_n = value of suspected outlier.

For normally distributed datasets, the Grubbs test can be used to determine whether suspected outliers can be reasonably removed, at a selected risk of false rejection³. Using the Grubbs test, one essentially calculates how much a suspected outlier differs from the population mean, measured in units of standard deviation. For a suspected outlier, calculate a value for T using either Equation 5 or Equation 6 as appropriate.

The standard Grubbs test (Equation 5) is directly applicable to data that is comprised of replicates of a single measurement or entity (i.e. groups of reference material or spike data). A modified version of this test is presented in Equation 6 for normalized duplicate pair data.

Compare the calculated T value to the appropriate critical value in Table 1 based on the number of datapoints in the set and the risk that can be tolerated for false rejection. We recommend the use of a 5% risk of false rejection.

Table 1: Critical Values for Grubbs' Test for $n \geq 7$.

Number of Data Points	Risk of False Rejection			
	0.1%	0.5%	1%	5%
7	2.201	2.139	2.097	1.938
8	2.358	2.274	2.221	2.032
9	2.492	2.387	2.323	2.110
10	2.606	2.482	2.410	2.176
15	2.997	2.806	2.705	2.409
20	3.230	3.001	2.884	2.557
25	3.389	3.135	3.009	2.663
50	3.789	3.483	3.336	2.956
100	4.084	3.754	3.600	3.207

Where the calculated T value for a result exceeds the appropriate value in the table (interpolate if necessary), the result is a probable outlier from the data population, and it may be reasonable to remove it. Wherever possible, attempts should be made to correct and/or explain the anomalous result before resorting to exclusion based solely on the Grubbs test. Table 1 is taken from Taylor³.

Equation 6 – The Grubbs Test to Identify Outliers of Normalized Duplicate Pairs:

$$T = |RD| / \sqrt{2} \cdot RSD_{DUP}$$

where:

RD = Relative Difference for each duplicate pair, as a decimal fraction [$RD = (a_i - b_i) / X_{mean}$].

RSD_{DUP} = Relative Standard Deviation of single measurements from within sets of duplicate pairs, as determined using Equation 4 (expressed as a decimal fraction).

Note: The relative standard deviation of a set of signed RPD values for duplicate pair data is equal to $\sqrt{2} \cdot RSD_{DUP}$, provided that the assignments of a_i and b_i are random (i.e. half the RPD's are positive and half are negative).

For duplicate pair data, the Grubbs test can be applied to normalized Relative Percent Difference (RPD) values, using Equation 6. In this application the “number of data points” column in Table 1 refers to the number of duplicate pairs used in the estimation of RSD_{DUP} .

Equations 5 and 6 are identical, except that Equation 6 uses compares normalized Relative Difference (RD) values to normalized Relative Standard Deviation (RSD) values. Note that for groups of Relative Difference values derived from duplicate pair datasets, RD is equivalent to $(X_n - X_{mean})$, because by definition, the mean RD is zero (or approaches zero). If the assignments of a_i and b_i are random (i.e. half the RD values are positive, and half are negative), then the mean RD value will approach zero. Using Equations 4 and 6, it doesn't matter whether the assignments of a_i and b_i are random in practice; there is no impact on the results if absolute value RD values are used.

The $\sqrt{2}$ factor is necessary because the RSD_{DUP} value calculated by Equation 4 represents the relative standard deviation of single measurements, and for this outlier test we need the RSD applicable to the difference of two measurements.

Sequential Description of the Type A Protocol.

Step 1: Estimate s_0 .

Estimate S_0 by the following two procedures:

- i) Prepare a series of low level method spikes (minimum $n = 7$) at between approximately one to five times the anticipated MDL level, or refer to existing method validation data if this data is already available. Calculate the standard deviation of the data series. This is a component of the protocol for determination of the MDL as described by the US EPA⁴.
- ii) Evaluate a series of method blank data that encompasses a long time period (ideally one year or more). Ensure that raw results are used wherever possible (even when less than detection limit). Calculate the standard deviation of the long-term method blank. This step may not be relevant for tests where contamination is rarely encountered and/or where numerical values are rarely obtained for method blanks. The Grubbs test (Equation 5) may be applied to this data if necessary to identify potential outliers.

Select the largest of the two values calculated above as S_0 for purposes of uncertainty determinations.

Step 2: Evaluate the Long-Term Method Blank Value.

- i) Using the method blank data from Step 1 part (ii), calculate the mean value of the long-term method blank to at least 2 significant digits. This step may not be relevant for tests where contamination is rarely encountered and/or where numerical values are rarely obtained for method blanks. The Grubbs test (Equation 5) may be applied to this data if necessary to identify potential outliers.
- ii) Compare the calculated mean value of the long-term method blank with the value for S_0 determined in Step 1. If the long-term method blank value is less than 1/5 the value of S_0 , it is insignificant, and can be ignored ($S_0/5$ is 10% of the 95% confidence expanded uncertainty contribution of S_0).
- iii) If the long-term method blank is significant in comparison with S_0 , consider whether blank subtraction can reasonably be applied to the method. Blank subtraction is recommended only in situations where the long-term blank is well characterized and has been demonstrated to be in statistical control.

Step 3: Estimate RSD_{DUP}

Duplicate data is often the best data available for estimating some of the most significant contributors to uncertainty for environmental analytical tests. Most importantly, duplicate data includes variance

due to sample heterogeneity and sub-sampling. For many tests in environmental chemistry, this represents the single greatest source of measurement uncertainty.

In addition, duplicate data includes the variance of method bias (if any), and most sources of within-batch or within-run variance. It excludes random error associated with the preparation of calibration standards, along with some other components of long-term variance.

The procedure below is based on the Canadian Association of Environmental Analytical Laboratories' (CAEAL) Policy on Uncertainty of Measurement in Environmental Testing.⁵ The calculations specified for Relative Standard Deviation are strongly recommended, unless it can be shown that standard deviation does not vary with concentration. The calculation of RSD_{DUP} is referenced by CAEAL to the Nordic Committee on Food Analysis' (NMKL) "Estimation and Expression of Measurement Uncertainty in Chemical Analysis."⁶

- i) Obtain a large population of duplicate data from the method. If available, compile 100 or more data pairs. A time-span of one year (longer if necessary) is recommended to capture seasonal variations.
- ii) Filter out all data where the mean value is less than 10x the estimated MDL (Note: MDL is equal to approximately $1.64 S_0 + MB_{LongTerm}$). This is necessary because the CAEAL/NMKL method of calculating RSD_{DUP} does not consider variance due to the MDL independently from other sources of variance. Typically, at mean values $>10x MDL^\dagger$, the uncertainty due to the MDL is $< 1/3$ of the total variability, and so will not affect the independent measurement of RSD_{DUP} .
- iii) Remove any anomalous datapoints that can be identified as errors.
- iv) Calculate the following for each duplicate data pair:
 - Mean of 2 results.
 - Difference of 2 results.
 - Relative Difference as a decimal value (Difference/Mean).
 - Square of Relative Difference.
- v) Calculate a first-pass value for RSD_{DUP} using the CAEAL/NMKL approach (Equation 4). Note that RSD_{DUP} refers to the RSD applicable to single measurements from within duplicate pairs.

This approach assumes that RSD_{DUP} is constant over the concentration range of the method at levels above approximately $10 \cdot MDL$. Where sufficient data is available, this assumption should be verified for sub-sets of method/analyte pairs by comparing RSD values for low, medium, and high concentration ranges.

[†] The value of 10x MDL is given as a general guideline. The point at which the contribution to uncertainty associated with the MDL becomes insignificant varies with the magnitude of the relative uncertainty. For example, if the relative uncertainty at $\gg MDL$ is 30%, then at 10x MDL, the uncertainty attributed to the MDL $\approx 1/(10 \cdot 0.3) = 1/3$ of the total uncertainty, and is therefore insignificant. If the relative uncertainty at $\gg MDL$ is only 10%, then the MDL does not become $< 1/3$ of the total uncertainty until 30x MDL ($1/30 \cdot 0.1 = 1/3$). The rough value of 10x MDL should be applicable in most cases, but higher cut-off values may be warranted for tests with very low uncertainties.

- vi) If desired, use Equation 6 to identify results that are statistical outliers. Where practical, investigate outliers in order to attempt to determine whether an error may have occurred. Some professional judgement may be necessary to determine whether outliers should be excluded from calculations. Some questions to consider are:
- Is it likely that the outlier is the result of an error (whether or not an error can be identified)?
 - Is the outlying duplicate pair likely caused by an atypical heterogeneous sample? For purposes of this protocol, we are defining atypical as representative of less than 5% of the total samples submitted for a test.
- vii) If outliers are excluded, recalculate RSD_{DUP} using Equation 4 with the outliers removed. Multiple passes of the outlier checker are acceptable, but should be used with caution.

Step 4: Determine RSD_{LT} (Long-Term Uncertainty Component)

The objective of this step is to estimate the relative standard uncertainty attributable to long-term variance that is not captured by duplicate data (e.g. long-term variability associated with calibration standards, multiple analysts, seasonal variations, etc.). Long-term Method QC data such as Reference Materials or Method Spikes can be used to evaluate this component of uncertainty.

Because Long-term QC data includes many of the components of variance that are present within duplicate data, it may sometimes be advisable to subtract the within-batch uncertainty associated with the same type of QC sample in order to avoid double-counting (see Step 4B). If the within-batch uncertainty is $< 1/3$ of the uncertainty due to duplicates, then the double-counting is insignificant, and the subtraction is not necessary.

This procedure assumes RSD_{LT} is constant throughout the concentration range of the method. If multi-level spike or RM data is available, this assumption should be verified for at least a sub-set of method/analyte pairs. Ideally, RSD_{LT} should be evaluated across the concentration range for which the method is validated.

Note that Method Spikes fully satisfy the objective of this step only when the spiking solution used is independent of calibration standards. If Spikes are prepared from calibration standards, then the long-term random error associated with calibration standards must be assessed independently (e.g. with RSD of 2nd source Control Standards) and added as a separate uncertainty term.

Step 4A: Determine RSD of Multi-Batch Method QC

- i) Obtain a large population of Method Spike or Reference Material data, which spans a long time period. A time-span of one year (longer if necessary) is recommended to capture seasonal variations. All data should ideally be $>10x$ MDL to minimize the contribution of s_0 to the measurement of RSD_{LT} .

- ii) Remove any anomalous datapoints that can be identified as errors.
- iii) Determine the RSD (first-pass) of the Multi-Batch Spike or RM data.
- iv) If necessary, use the Grubbs outlier test as described in Equation 5 to identify any anomalous data that cannot definitively be shown to be in error. A risk level for false rejection of 5% is recommended.
- v) Re-calculate the RSD if outliers were removed. It is acceptable to repeat the outlier removal process if necessary, but caution is recommended to avoid the rejection of false outliers.

Step 4B: Determine RSD of Single-Batch Method QC (Optional)

If the RSD determined above (Step 4A) is $\gg 1/3$ of RSD_{DUP} , then it may be worthwhile to evaluate the single-batch RSD for the same type of Method QC. If the single-batch RSD is known, it is possible to subtract it (as a variance) from the RSD of multi-batch, as determined in Step 4A, in order to correct for potential double counting.

Before proceeding with this procedure, ensure that the value obtained for RSD_{DUP} fully represents all components of single-batch variability. Some factors to consider include:

- Are duplicates for this test always taken side-by-side through the process, or are they randomized? If they are not randomized, some aspects of single-batch variability may be under-estimated.
- Are duplicate data pre-screened in any way before data entry (i.e. are "bad duplicates" typically re-analyzed until they meet specifications). Use of pre-screened data under-estimates the true value of RSD_{DUP} , because samples that are not duplicated will be subject to a lesser degree of scrutiny.

If either of the above questions indicate that duplicate data may not fully capture typical single-batch variability, then the subtraction of single-batch variance from long-term variance is not recommended.

Next, one might anticipate whether the single-batch relative standard deviation will likely be less than $1/3$ of the long-term relative standard deviation. If this is the case, then the subtraction will be insignificant, and the effort required to perform this step would not be justified.

In those cases where the correction is warranted and desired:

- i) Obtain a reasonable population of Method Spike or Reference Material data from within single analytical batches or runs. A minimum of 7 samples per batch is recommended. It is recommended to evaluate multiple concentration levels that bracket the validated range of the method. Treat all datasets from different batches separately.
- ii) Determine the RSD of each set of within-batch QC. Remove outliers if necessary using the Grubbs test (Equation 5, 5% risk of false rejection recommended).
- iii) If multiple sets of data were evaluated, use an average value for Single-Batch RSD.

Step 4C: Determine RSD_{LT}

- i) If Step 4B was omitted, then use the RSD of Multi-Batch Method QC (from Step 4A) as RSD_{LT} .
- ii) If Step 4B was followed, then calculate RSD_{LT} as follows:

$$RSD_{LT} = \sqrt{(RSD_{Multi-Batch}^2 - RSD_{Single-Batch}^2)}$$

Step 5: Evaluate Other Sources of Relative Uncertainty (if any).

If any other sources of relative uncertainty have been identified for a method that are not covered by RSD_{DUP} and RSD_{LT} , then determine a mechanism to evaluate and add this contribution in a similar fashion (as described in Equation 3).

Step 6: Combine All Calculated Factors into the Uncertainty Function.

Calculate Θ , using Equation 3:

$$\Theta = \sqrt{[RSD_{DUP}^2 + RSD_{LT}^2 + \dots + RSD_n^2]}$$

Calculate Uncertainty (95%), using Equation 2:

$$U(c) = 2 * [\sqrt{\{s_0^2 + (\Theta c)^2\}}] + MBk_{LT}$$

Modifications to the Type A Protocol for Special Cases.

Special Case 1: Reporting of Averaged Results

The uncertainty associated with averaged results is lower than the uncertainty of a single result. A more representative (and usually smaller) value for uncertainty can be obtained if several replicate measurements are taken for a given sample.

For results subjected to averaging, the standard deviation of the mean is represented as:

$$s_{\text{mean}} = s / \sqrt{n}$$

In the above equation, s is the standard deviation *applicable to single measurements* (note that the special case does not apply if s is also determined by averaged results). In the case where an averaged result is reported, the reported uncertainty may be estimated in the same manner as for single results, except that an alternative value for RSD_{DUP} may be determined directly from the standard deviation of the replicate measurements of the sample, as indicated:

$$RSD_{\text{DUP}} = s / [\text{Mean} * \sqrt{n}]$$

Where:

- s** = standard deviation of the replicated results.
- Mean** = average value of the replicated results.
- n** = number of replicate measurements taken.

Then, as for single measurements:

$$\Theta = \sqrt{ [RSD_{\text{DUP}}^2 + RSD_{\text{LT}}^2 + \dots + RSD_n^2]}$$

$$U(c) = 2 * [\sqrt{ \{ s_0^2 + (\Theta c)^2 \} }] + MBk_{\text{LT}}$$

Special Case 2: Tests Where Duplicates are not possible.

For methods where entire samples are analyzed as submitted, lab duplicate cannot be analyzed. Sub-sampling is not a component of measurement uncertainty for these methods. The analytical component of these methods is therefore not influenced by sample heterogeneity.

In such cases, the variability associated with duplicates is undefined, and this component can be omitted from estimates of measurement uncertainty. Where duplicates are omitted from the estimation process, it is not suitable to subtract within-batch variance from long-term variance as described in steps 4B and 4C.

Where a test does not include a sub-sampling component, and where duplicates are not possible, measurement uncertainty can be determined using long-term Reference Material or Method Spike data. Ideally, this data should bracket the validated concentration range of the method.

As in the general, case, if long-term Method Spikes are used for this purpose, they must be prepared independently from calibration standards. Otherwise, the random error associated with calibration standards must be assessed independently (e.g. with RSD of Control Standards), and added as a separate uncertainty term. If additional sources of uncertainty are unaccounted for, then they must also be added.

$$\Theta = \sqrt{[\text{RSD}_{\text{LT}}^2 + \dots + \text{RSD}_n^2]}$$
$$U(c) = 2 * [\sqrt{ \{ s_0^2 + (\Theta c)^2 \} }] + \text{MBIk}_{\text{LT}}$$

References:

- ¹ Eurachem/CITAC Guide, "Quantifying Uncertainty in Analytical Measurement," Second Edition (QUAM:2000.P1). Especially Appendix E.4, Statistical Procedures, documenting uncertainty dependent on analyte level. Editors SLR Ellison, M Rosslein, A Williams.
- ² Ramsey, Michael, 1998. "Sampling as a source of measurement uncertainty: techniques for quantification and comparison with analytical sources", Journal of Analytical Atomic Spectrometry, Feb. 1998, Vol. 13 (97-104).
- ³ Taylor, John Keenan, 1987. "Quality Assurance of Chemical Measurements," Lewis Publishers, pages 37 and 271.
- ⁴ US EPA 40 CFR Part 136, Appendix B, Rev 1.11, "Definition and Procedure for the Determination of the Method Detection Limit".
- ⁵ "CAEAL Policy on Uncertainty of Measurement in Environmental Testing," Revision 1.4. Treatment of duplicates is discussed under "Sample Duplicate Insertion".
- ⁶ NMKL Procedure No. 5 (1997), "Estimation and expression of measurement uncertainty in chemical analysis," Nordic Committee on Food Analysis (NMKL).