Experimental sampling designs for uncertainty estimation

Both sampling and analysis contribute to measurement uncertainty, sampling uncertainty is actually important, if not even more important than, analytical uncertainty in test result. Therefore, the need for a proper sampling plan cannot be over emphasized.

By the top-down (holistic) approach for the evaluation of measurement uncertainty instead of the GUM method, we need to look at two uncertainty components, namely the *random* and *systematic (bias)* effects.

In order to quantify the variance (random variation) associated with sampling, we have to take a number of samples from the bulk material (commonly known as the population or the target) for laboratory analysis in order to determine the amount of test parameter of interest in replicates. This random part of the uncertainty is expressed by the standard deviation.

Normally in the evaluation of measurement uncertainty from the analytical point of view, we also look at the systematic error (i.e. bias) in the analysis. It is quite simple to do by running the analysis against a certified reference material with known value, comparing the results against a reference analytical method or taking part in a proficiency testing (or inter-laboratory comparison) program.

However, it is not a simple task to the determination of the bias effect due to sampling, as there are few sampling proficiency tests for comparison study exist.

Possible alternative approaches are:

a. When a theoretical value of test parameter has been established (such as from factory production), we can use it as an estimate of the true value;
b. Sampling can be performed on a reference sampling target with known value, if there is one;
c. Inter-sampler comparison can be made when two or more people performing sampling and analysis;

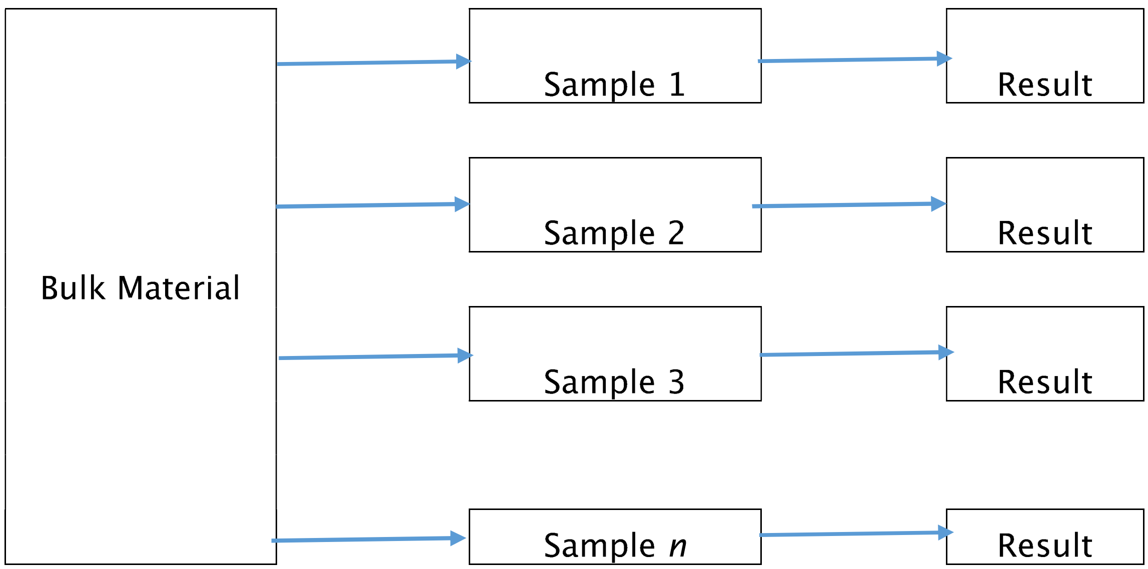
The variances of sampling and analysis bias, if established, should be included as another uncertainty budget to the combined uncertainty of measurement which covers both sampling uncertainty and analytical uncertainty.

In this note, we focus on the random part of measurement uncertainty. There are few approaches to study the randomness of sampling uncertainty.

Sampling variance only

If we take a number of samples from the field or the consignment of material in random order and then analyze these samples one after another, following by computing the mean of the final results. The schematic diagram for this kind of experiment is shown in Figure 1.

Figure 1: Sampling plan by taking several samples for analysis



By this experimental design, we obtain measurement results for all the samples taken from the field and calculate the sample mean $\bar{x}$ and standard deviation, s. The Student's t–equation (1) below provides us the confidence interval say at 95% level of this exercise:

$$\mu = \bar{x} \pm t_{(\alpha=0.05, df=n-1)} \times \frac{s}{\sqrt{n}} \qquad (1)$$

where n is the number of samples analyzed. The confidence interval (also known as standard error of mean SEM) is represented by $t_{\alpha=0.05, df=n-1} \times \frac{s}{\sqrt{n}}$.

As only a single measurement is done on each sample, we have no idea of its analytical variance. In other words, we assume the analytical variance to be zero or negligible which is not always true.

Hence this approach provides us only the extent of sampling random error and not really sampling uncertainty which covers also its true value with a certain degree of confidence.

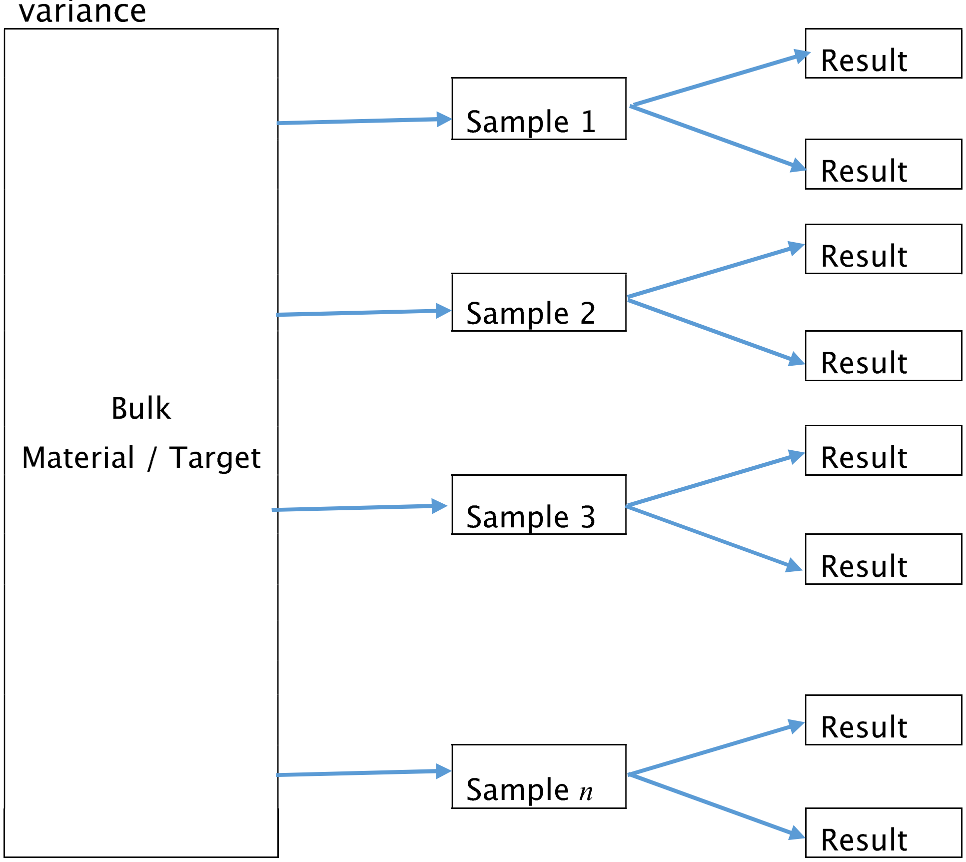
Sampling and analytical variances

The approach described below quantify the variance associated with sampling and analysis.

This is a case where we want to estimate the separate variances associated with sampling and testing. We take a number of samples from a field randomly and send the samples to the laboratory, instructing the laboratory personnel to carry out *duplicate* analysis on each of the samples.

Figure 2 below shows a schematic drawing of having a simple design for the determination of sampling and analytical variance.

Figure 2: An simple split experimental design for sampling and analytical variance



For the above simple balanced design, we can use the one-way analysis of variance (ANOVA) technique to estimate the mean squares (variances) of *between samples* σ_s and *within sample (or analysis),* σ_a. The total standard deviation for a combined operation of sampling and analysis is going to be the square root of sum of these two variances.

Nested (multiple – target) balanced design of sampling

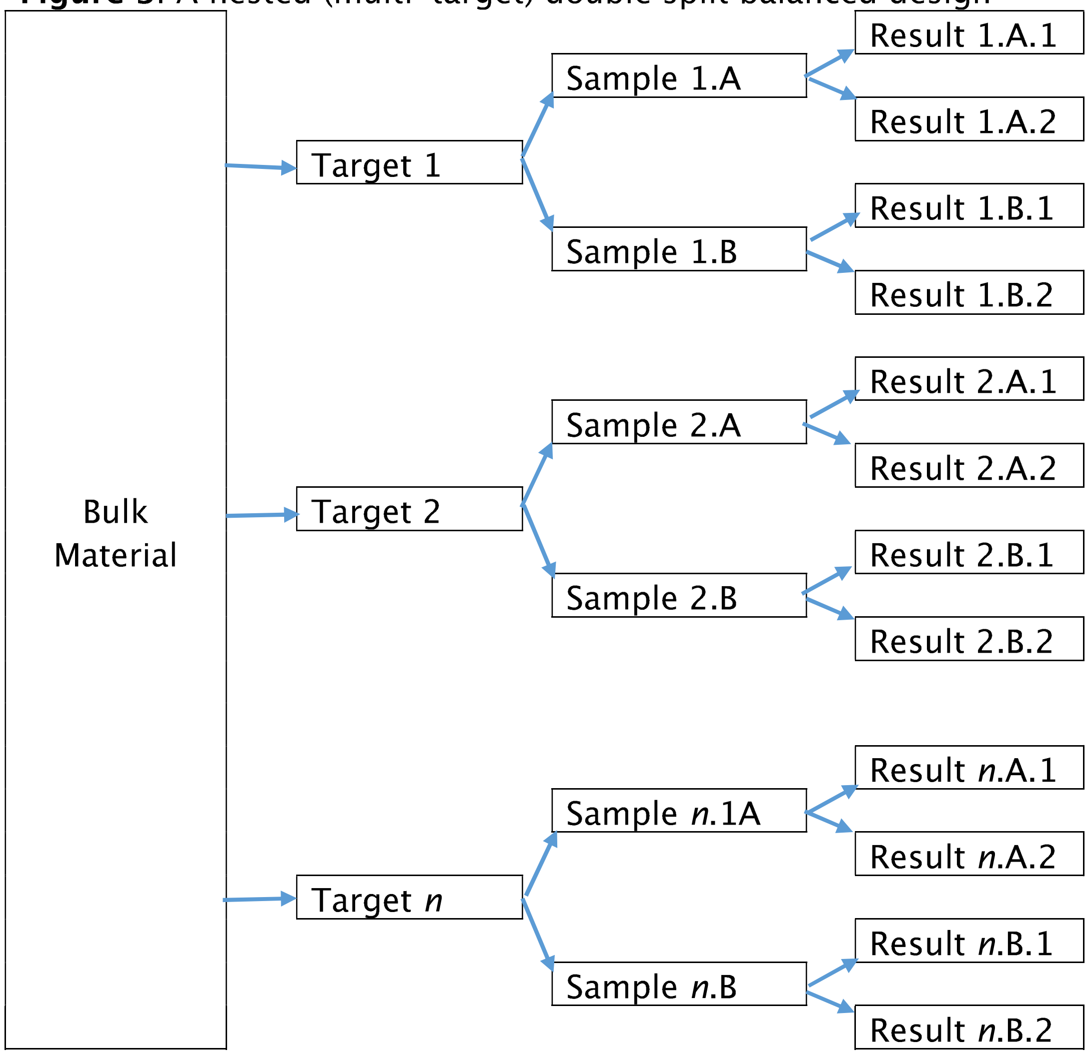
Figure 2 above shows taking multi-samples from a target of the bulk material with duplicate analysis. This simple design, however, is based on the assumption that the particular target under study is typical of all targets in the bulk material. It has disregarded the possibility that targets may vary in their degree of heterogeneity, and therefore in the value of σ_s.

A greatly preferable estimate, characterizing a whole class of material, is by taking duplicate samples from a succession of different targets of the bulk material. See Figure 3.

This double split procedure is also quite straight forward though a greater time may be required to accumulate and analyze the results using the ANOVA technique. In this manner, we will obtain estimates of sampling standard deviation, in addition to the between-target and analytical variations.

In fact, this double spilt method is also applicable in the conduct of *quality control of sampling* which ensures the sampling protocol is not out of control. We shall discuss this topic in a later article.

Figure 3: A nested (multi-target) double split balanced design



References

1. Eurachem/CITAC Guide (2007) "*Measurement Uncertainty arising from Sampling*"
2. Nordtest Technical Report TR 604 (2007) "*Uncertainty from Sampling – A Nordtest Handbook for Sampling Planners on Sample Quality Assurance and Uncertainty Estimation*"
3. Michael Thompson & Philip Lowthian (2011) "*Notes on Statistics and Data Quality for Analytical Chemists*", Imperial College Press