SAMPLING COLUMN

Sampling quality criteria (SQC)

Kim H. Esbensen[a] and Claas Wagner[b, *]
[a]KHE Consulting, www.kheconsult.com
[b]Sampling Consultant—Specialist in Feed, Food and Fuel QA/QC. E-mail: cw@wagnerconsultants.com

Formulating proper Sampling Quality Criteria (SQC) is the initial step in a scientific approach to representative sampling: this activity can be characterised as "a framework for planning and managing sampling and analytical operations consistent with the overall project objectives". It includes establishment of concise sampling objectives, precise outlining of the decision unit (DU) and deciding on the level of confidence wanted regarding the kind(s) of decisions to be made based on the analytical results. Once defined, these criteria serve as input to the Theory of Sampling for developing a representative sampling protocol.

Sampling quality criteria (SQC)

The first component in SQC is defining the *analyte(s)* to be involved, the concentration level of interest of the analyte(s) and how inference(s) will be made from the analytical data to the decision unit.

The second component of the SQC concerns definition and establishment of the physical *decision unit (DU)*, also known as the "lot" in the Theory of Sampling (TOS). The decision unit establishes the spatial and temporal boundary conditions of the sample collection process.

The third SQC component, the confidence level, establishes the desired probability that a correct inference (decision) *can* be made. The confidence level should typically correlate to the potential consequences of an incorrect decision, e.g. health, economic, societal consequences. The magnitude of the total combined errors in the sampling, sample processing and analytical protocols constitute the unavoidable basic risk level involved, and determines the likelihood of an incorrect decision. Thus, controlling errors to a greater extent increases the probability of a correct decision. The required confidence level also directly affects the sampling effort and quality control measures.

Establishing proper SQC is nothing but very carefully thinking through the *why*, *what* and *how* regarding the use of the final analytical results; surprisingly often this prerequisite does not get the full attention it deserves.

First SQC component— definition of analyte(s)

The first SQC component addresses definition of the analyte(s), the expected concentration level of interest, as well as how the analytical data will be used; i.e. how *inference* is made with respect to the decision unit. This prerequisite sets the scope and limits the selection possibilities for sampling tools, sampling containers, sample handling procedures, sample preservation, laboratory preparation equipment, sample mass reduction procedures etc. For all sampling operations it is critical to secure analyte integrity from the primary sample all the way through to the final analysis.[1]

Second SQC component— delineating the decision unit (DU)

The decision unit (*aka* the lot) defines the target (material, form, size, conditions) from which the primary sample is to be collected and importantly sets the *scale of decision-making*. This scale can be based on volume, mass, package size or any other definable criteria relevant for the project. As defined in a previous sampling column, a pre-requirement for the decision unit is physical accessibility (Fundamental Sampling Principle).[2] If a certain part of the decision unit is not accessible, the section either **must** be made accessible for the purpose of sampling or the (limited) target is

It so easy to do it **WRONG!** And so easy to do it **RIGHT!**

Model photo: with permission

disqualified from being a proper decision unit from which defensible and reliable decisions can be made. Sampling of a decision unit under limiting constraints, be these practical, logistical, economic etc., will unavoidably forfeit any chance of extracting representative samples; only worthless specimens will ensue.

Third SQC component—inference and confidence

Three basic types of inference can be applied from the final analytical data, i.e. from the concentration of the analytes of interest: *judgement*, *direct inference* and *statistical analysis*.

Judgement is never an acceptable methodology, irrespective of whatever personal experience is involved. Judgemental inference is not discussed further here.

Direct inference is the simplest proper inference approach in which the analytical result from a single primary (composite) sample is used as a reliable, and defensible, estimate of the average concentration of the analyte in the decision unit/lot. This approach requires no statistical analysis as long as the principles of the Theory of Sampling are obeyed. This is the tacitly assumed sampling situation in many "straightforward" situations.

In comparison to direct inference, *statistical analysis* involves multiple primary samples. In such case, the upper confidence limit of the mean is compared to a specification limit or a statistical comparison of one decision unit to another decision unit (e.g. reference decision unit). In "acceptance sampling", e.g. in releasing a batch of drug dosages/tablets, it is the lower limit of the confidence interval that sets the operative threshold.

There is no rule for setting the confidence level. It is a function of the consequences of an undesired incorrect decision. Typically the larger the consequences of an incorrect decision the higher the desired confidence needs to be. It is **not** a good policy to always set the confidence at the same level (e.g., 95%) if it is not known *why* this level is actually used. Never mind that this "usually" is the level encountered; there

are enough "template statistics" governing complex problems, projects etc. that it may just be that careful consideration (proper SQC deliberations) reveal that some other level of confidence is more appropriate.

In order to calculate a confidence level for statistical analysis, an estimate of the global estimation error (*GEE*) is required, as defined in previous sampling columns:

$$GEE = Total\ Sampling\ Error + Total\ Analytical\ Error$$

For either case of direct inference or statistical analysis, the estimate of the *GEE* (all bias sampling errors need to have been eliminated first, TOS) can be determined from a properly conducted replication experiment, see previous *SE* column).[3]

If the sampling bias errors have not been properly eliminated, the estimate of *GEE* will forever be of varying magnitude, see previous sampling column,[2] with the consequence that the premise for a constant confidence level is broken.

Figure 1 graphically depicts the relationships between confidence, error and representativeness. Applying the Theory of Sampling, including the defined sampling quality criteria, to any sampling protocol ensures representativeness, because the primary goal of the Theory of Sampling is to minimise the total sampling error (*TSE*), which in turn increases the confidence of the inference made from the analytical data to the decision unit.

This column is but an introduction serving to raise the awareness of the need for proper SQC; the reader is

referred to more in-depth treatments in the references,[4–9] with a natural starting point in Reference 1.

Perspectives

Industry is critically dependent on the highest data quality (data relevance, data reliability, data representativity). Representative data cannot be acquired without a *sampling process* initiating the full "lot-to-analysis-to-decision" pathway. In this endeavour the critical determinant is the potential *sampling bias,* an inconstant deviation between an analytical sampling result and the true average concentration of the lot, product or process, which must be *eliminated* from all sampling processes at all stages to ensure data validity. The sampling bias is fundamentally different from the well-known analytical bias, which can be eliminated by a statistical bias correction. However, this approach is impossible for the sampling bias, which, in addition, is orders of magnitude larger in most cases.

Sampling does not necessarily only refer to physical sample extraction, but also to on-line process and/or product measurements through the use of Process Analytical Technological (PAT) sensors. PAT is a framework originally developed by the US Food and Drug Administration to design, analyse and control pharmaceutical manufacturing processes through (continuous) measurement of critical process parameters, which is now being applied to many other manufacturing and processing industry sectors. In case the highly advanced measurement sensor technologies extract information from a wrongly defined decision unit or biased sample,
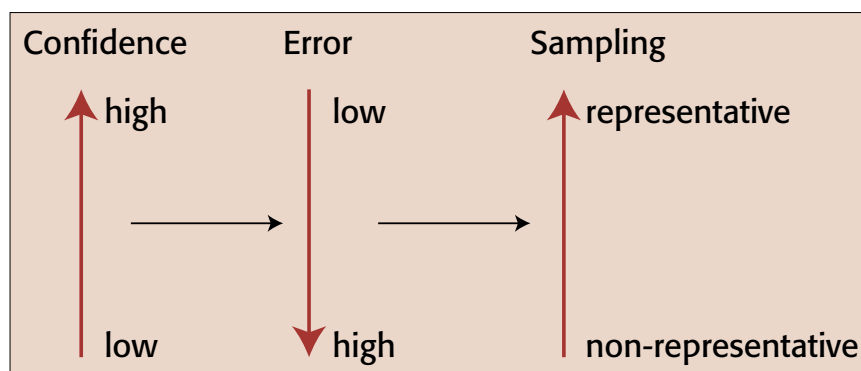

Figure 1. Relationship between confidence level, error magnitude and representativeness.

SAMPLING COLUMN

the results of this advanced process control technique will unavoidably lead to biased results. The main challenge is therefore to ensure that highly precise measurements also reflect the true target value of the decision unit (requiring elimination of the sampling bias), which currently most of the times **remains unaddressed**. **A fully comprehensive SQC is a missing link in this context.**

In all sampling situations (whether sampled and analysed physically, or measured by the use of PAT), *representativeness* (of samples and signals) is the prime objective without which the derived analytical results and the decisions based on these data are invalid. Representativeness implies both elimination of the fatal sampling bias, as well as high reproducibility of the sampling process. Depending on the required use of data (used "as is" for monitoring or aggregate data for higher-level decision-making), there will always be a problem-dependent "decision unit" defining the target from which a sample is extracted or for which a signal is measured. **But critical DUs do not always conveniently suggest themselves; problem-based due diligence is required!**

Conclusions

Defining sample quality criteria must be the initial step in the development of any sampling protocol. SQC defines the analytes, the decision unit and addresses the required inference and its confidence level. The precise definition of the analyte(s) and the decision unit is currently one of the weaker elements, and not always sufficiently addressed in sampling protocols, with the potential consequence of using improper sampling protocols, ultimately with the end result of invalid inferences. Above all: never resort to using grab sampling!

References

1. C. Ramsey and C. Wagner, "Sample quality criteria", *J. AOAC Int.* **98,** 265 (2015). doi: http://dx.doi.org/10.5740/jaoacint.14-247
2. K.H. Esbensen and C. Wagner, "Composite sampling I: the Fundamental Sampling Principle", *Spectrosc. Europe* **27(5),** 18 (2015). http://bit.ly/1OyT9r0
3. K.H. Esbensen and C. Wagner, "Sampling quality assessment: the replication experiment", *Spectrosc. Europe* **28(1),** 20 (2016). http://bit.ly/1RQ11XJ
4. K.H. Esbensen, *DS 3077 Representative Sampling—Horizontal Standard*. Danish Standards, http://www.ds.dk (2013).
5. *GOODSamples: Guidance on Obtaining Defensible Samples*. Association of American Feed Control Officials, Champaign, IL (2015). http://www.aafco.org/Publications/GOODSamples
6. F.F. Pitard, *Pierre Gy's Sampling Theory and Sampling Practice*, 2nd Edn. CRC Press, Boca Raton, FL (1993).
7. C.A. Ramsey and A.D. Hewitt, "A methodology for assessing sample representativeness", *Environ. Forensics* **6,** 71 (2005). doi: http://dx.doi.org/10.1080/15275920590913877
8. C. Ramsey, "Considerations for inference to decision units", *J. AOAC Int.* **98,** 288 (2015). doi: http://dx.doi.org/10.5740/jaoacint.14-292
9. C. Wagner and K.H. Esbensen, "Theory of Sampling: four critical success factors before analysis", *J. AOAC Int.* **98,** 275 (2015). doi: http://dx.doi.org/10.5740/jaoacint.14-236

WHAT did we tell you about grab sampling?