# Comments on the IAM Sampling Paper

## 1 Introduction

We are not comfortable with the argumentation in the IAM Sampling Paper and hence cannot endorse its main conclusions. In these notes, we present the reasons why we reject or disagree with the views and arguments put forward in the IAM Sampling Paper.

## 2 Main criticism: Acceptance sampling vs. Estimation of the total uncertainty from both Analysis and Sampling

We would like to focus on two of the approaches to sampling discussed in the IAM Sampling Paper:

* Acceptance Sampling (AcS)
* Estimation of the total uncertainty from both Analysis and Sampling (UfS)

We take the basic direction proposed in the IAM Sampling Paper regarding these two approaches to be defined in the two recommendations on pp. 11 and 12.

These recommendations read as follows:

IAM Sampling Paper recommendation regarding AcS approach:

*That acceptance plans are only developed for specific system [sic] by those who fully appreciate their significance and associated difficulties and that the consequence of reducing the number of units taken from a lot area [sic] clearly appreciated.*

IAM Sampling Paper recommendations regarding UfS approach:

It is recommended that Codex:

* *Notes the publication of the EURACHEM/EUROLAB/CITAC/Nordtest Guide on the “Estimation of Measurement Uncertainty Arising from Sampling” and the Nordtest handbook.*
* *Discusses the issue of uncertainty and sampling and decides whether it should develop recommendations in the area in the same way that it already has for [Analytical] Measurement Uncertainty.*
* *Discusses whether sampling uncertainty should be taken into account when a lot is assessed for compliance with a Codex specification.*
* *Considers whether it should prepare Guidance for Codex Committee Committees on uncertainty from sampling.*

The reasons why we do not endorse these recommendations are presented in the following two sections (Sections 2.1 and 2.2). In Section 2.3, we also comment on the Strengths & Weaknesses analysis of UfS and AcS found on pages 5 and 6 of the IAM Sampling Paper.

### 2.1 Consumer and producer risks not only relevant for AcS approach but also for UfS approach.

The first issue we would like to discuss is the question of “*the consequence of reducing the number of units from a lot*”, to use the phrasing from the recommendation regarding the AcS approach quoted above. This issue is raised several times in the IAM paper, and concerns the “*high probability of accepting defective items, especially for small batches, that is not appreciated by some regulators*” (discussion of the weaknesses of the AsC approach on page 6).

We agree that there is a high probability of accepting defective items (false negative rate) at low sample sizes, and this point is very aptly illustrated in the table on page 10 of the IAM Sampling Paper (which can be directly constructed on the basis of the information presented in Chart A and Charts B-R from ISO 3951-1).

The gist of the matter, however, seems to us to lie in the fact that in the AcS approach, the false negative and false positive rates constitute the starting point for the determination of the AQL and the sample size. This approach thus compels the user to be aware of the consumer and producer risks (i.e. of false negative and false positive rates). If the desirable risk levels are not attainable due to limitations on sample size occasioned by economical or other constraints, the user will at least remain aware of the ensuing unsatisfactory risk levels.

In our opinion, the emphasis on consumer and producer risks which characterizes the AcS approach is essential for any viable approach to sampling. We are concerned that, in the IAM Sampling Paper, the high “percentage of defectives which will be accepted 10 % of the time” is construed as a weakness of the AcS approach, rather than as an inevitable consequence of the sample size. Indeed, if the consumer and producer risks were computed for the UfS approach, it would be found that they would be essentially the same. The only difference is that, in the UfS approach, it is not required to compute these risks, and the user often remains unaware of them. If, in the framework of the UfS approach, the false negative and false positive rates ensuing from the application of a particular rejection criterion for lots were computed, they would, by and large, be *the same* as in the AcS approach. In other words, what we see as a shortcoming of the UfS approach (i.e. the failure to emphasize the need to compute the false negative and false positive rates) is put forward as a virtue.

In short: we would have expected the IAM Sampling Paper to

* emphasize the importance of computing consumer and producer risks
* point out that the computation of these risks plays a central role in the AsC approach
* require that the UfS approach be expanded to include the computation of these risks

### 2.2 Is sampling uncertainty not taken into account in the AcS approach?

The second issue we would like to discuss is addressed a few times in the IAM Sampling Paper: namely, the question of whether the sampling uncertainty is taken into account in the AcS approach. In the IAM Sampling Paper, it is stated a few times that this is not the case, and that only analytical uncertainty is taken into account in the AcS approach. See for example the **Conclusions – comparison of approaches** Section on page 7:

“*1. Only in the UfS approach does the information from the validation step (e.g. on the portion of the measurement uncertainty from sampling and sample preparation) get reported to the user of the measurement results. (e.g. 15 ± 10 ng g-1, rather than just the analytical portion 15 ± 1 ng g-1)*

*2. The differences in terminology of the three approaches reflect deeper distinctions. For example, the ‘variability’ due to sampling in AcS produces ‘uncertainty’ in the measurement (of concentration) that is not reported to the user (i.e. producer, consumer or regulator)*

*3. The more realistic estimate of measurement uncertainty given by the UfS approach is essential to making reliable decisions and classifications on the acceptability of material for its intended purpose (e.g. safety of food for consumption). The methodology for using this uncertainty information in enforcement decisions is not yet agreed internationally, for example in deciding the acceptable levels of false positive (producer’s or seller’s risk) and false negative (consumer’s or buyer’s risk) classifications. However, the UfS approach will enable this methodology to be applied not just at the validation stage, but also in routine operation.*

*4. Both ARS and AcS consider sampling variability in the design of the initial sampling protocol, but don’t consider or express it as part of the measurement process. This has the advantage of apparent simplicity, but misleads the decision maker on the reliability of the classification decision. However, the AcS approach uses the equivalent of the UfS information, in moving the effective threshold value (e.g. to AQL)”*

In short, the IAM Sampling Paper’s position seems to be that the only approach in which information about sampling uncertainty is adequately taken into consideration, quantified and reported to the user is the UfS approach. This position seems to us to be inaccurate.

Indeed, the very producer and consumer risks (or, more generally, the OC curves) quantify and codify the information about sampling uncertainty in a manner which is relevant for the user. We would argue that providing a sampling variance in the framework of a variance decomposition (i.e. the UfS approach) does not constitute information which can easily be interpreted by the user. On the other hand, a diagram showing that a lot with $p$ % defective items will yield an acceptance of the lot with such and such probability is precisely the information which the user needs.

In the Annex we provide further details on the treatment of sampling uncertainty in the AcS method (ISO 3951 approach).

### 2.3 Comments on the analysis of Strengths & Weaknesses for UfS and AcS

In Sections 2.1 and 2.2 we pointed out that:

1. High consumer or producer risks are not a flaw of the AcS method, but rather an inevitable consequence of fit-for-purpose decisions (e.g. economic constraints on the sample sizes) and of the assumption of a normal distribution.
2. These high risks would also be found to attend decisions taken in accordance with the UfS approach. However, in the UfS approach, it is not required that these risks be computed.
3. These consumer and producer risks (or, more generally, the OC curves) constitute an estimate of sampling uncertainty, expressed in a manner which is relevant to the decision-maker.

We will now review how these three points apply to the Strengths & Weaknesses analysis of UfS and AcS found on pages 5 and 6 of the IAM Sampling Paper. In addition we will point out what, in our opinion, seem to be logical inconsistencies.

|  |  |
| --- | --- |
| Strengths of the UfS approach according to IAM Sampling Paper (pages 5, 6) | Our comments |
| Gives a realistic estimate of the measurement uncertainty, which will make decisions on batch acceptance/rejection more reliable. | This depends on the way in which the uncertainty ranges are computed. In our opinion, this cannot be construed as an inherent strength of the UfS approach (see Section 3.3 below). |
| Enables the fitness-for-purpose of the measurements (& sampling) to be judged in terms of minimizing the overall costs of both measurement and incorrect regulatory decisions. | This depends on the way in which the uncertainty ranges are computed. Minimizing the overall costs of both measurement and incorrect regulatory decisions can also be achieved by using the AcS approach. We fail to see how this point constitutes a strength of the UfS approach as compared to the AcS approach (see point 1 above). |
| Inclusion of sampling quality control monitors on-going performance of samplers in routine application of the protocol, not just at validation. | Sampling quality control is also an issue of the AcS approach.  |

|  |  |
| --- | --- |
| Weaknesses of the AcS approach according to IAM Sampling Paper (page 6) | Our comments |
| The AcS approach underestimates the overall uncertainty of the measurement (excludes contribution from sampling), which will affect the reliability of decisions on batch acceptance/rejection. | * This is not correct: The AcS approach does not exclude the contribution from sampling (see point 3 above)
* This is also a direct contradiction with point 2 under Strengths (“Makes empirical estimates of variability arising from sampling,…)
 |
| Does not give the information on sampling variability (and hence larger measurement uncertainty) to the decision maker. | The AcS approach allows for information about sampling variability to be given to the decision maker in the form of consumer and producer risks (see point 3 above) |
| No way of checking on the quality of the actual implementation of the sampling protocol in routine operation. | We fail to see how this can be construed as a weakness of the AcS approach: The same could be said of any approach, and it depends on the way the approach is implemented. |
| Hard to devise correct protocol for heterogeneous material sampled in situ (e.g. un-mixed nuts in a container, or contaminated land). | We fail to see how this can be construed as a weakness of the AcS approach: The same could be said of any approach, and it depends on the way the approach is implemented. |
| Does not include potential financial losses that may arise from decision errors (caused by uncertainty) in calculation of final sampling protocol. | We fail to see how this can be construed as a weakness of the AcS approach: The same could be said of any approach, and it depends on the way the approach is implemented. |
| Gives rise to a high probability of accepting defective items, especially for small batches, that is not appreciated by some regulators. | The same can also be said of the UfS approach (see point 2 above) |

Finally, we are concerned that the IAM Sampling Paper addresses only 2 weaknesses altogether

Almost no weaknesses of the UfS approach are addressed. Are there really no serious weaknesses – apart from the fact that the “*methodology for including uncertainty from sampling in decision-making process (is) not yet agreed*”? It hardly seems tenable that there are no other serious weaknesses.

## Further comments

### 3.1 The question of a sampling bias (Key questions 5+7+8, page 1 of IAM Sampling Paper)

In our opinion, the concept “Sampling bias” can be somewhat misleading (see for example key questions (7) and (8)). Indeed, the presence or absence of a bias depends not only on the sampling procedure but also on the choice of statistical method applied in the evaluation of the data. For instance, if a stratified sampling procedure has been followed (as is frequently the case), the arithmetical mean will be considerably biased whereas the appropriate stratified mean will not. This shows that the presence or absence of a bias is not a property of the sampling procedure alone but rather of the combination of sampling procedure and statistical method.

For this reason, the answer to key question (5) should be: Yes, sampling must be considered as the first step in the measurement process rather than as a separate process.

### 3.2 The question of a representative versus appropriate samples (Key question 12, page 2 of IAM Sampling Paper)

In our opinion, the formulation of key question (12) is too vague: in particular, it seems to us the concepts “representative” and “appropriate” are not sufficiently well defined to allow concrete measures to be taken in terms of establishing useful practical guidance. The question of whether to use a representative or a pragmatic sampling (see page 2) suffers, in our opinion, from the same ill focus.

### 3.3 The issue of large sampling uncertainty

It should be noted that, in cases where the sampling uncertainty is large (say, relative standard deviations larger than 30 %), classical uncertainty ranges can be considerably misleading when the assumption of symmetrical ranges is erroneously made. Thus, in these cases, the UfS approach is fraught with substantial difficulties. By construction, the AcS approach is not affected by such problems.

## 4 Conclusion

We identified a large number of weaknesses, inconsistencies and inaccuracies in the arguments of the IAM paper. Apart from that we disagree with the views and arguments put forward in the IAM Sampling Paper.

We are concerned that the IAM Sampling Paper’s recommendation for the AcS approach fails to lay sufficient emphasis on the most important aspect of Acceptance Sampling: the fact that in this approach, producer and consumer risks – and thus the attendant risks of financial losses and food safety - play a central role. We are concerned that the IAM Sampling Paper does not recognize that sampling uncertainty is implicitly taken into account in the AcS approach, and that it is reported to the user in a very practical and relevant way; namely, as the producer and consumer risks, or more generally, as the OC curves. Finally, we are concerned that CCMAS shall decide on a “methodology for including uncertainty from sampling in decision-making process (which has) not yet (been) agreed” (cf. IAM Sampling Paper, p6).

Our intention is by no means to blindly embrace the AcS approach. In particular, we recognize the ISO 2859 and ISO 3951 standards are written in a manner which makes them relatively difficult to apply. Both standards use complicated terminology, fail to provide proper explanation of statistical procedures and use complicated tables and charts for decision-making. Therefore we suggest to develop more appropriate terminology, produce explanations for statistical concepts, simplify outdated procedures of decision-making and provide additional tools if needed. In the Annex some ideas are presented which could be used to support this work.

Currently both the ISO 2859 and the ISO 3951 standards are under revision, and we encourage member states to support this revision – which should include a thorough examination and description of uncertainty components. We also suggest preparing documents which describe more specifically and on a scientifically sound basis the differences and similarities between the different approaches in order to achieve an impartial appraisal of their respective advantages and shortcomings.