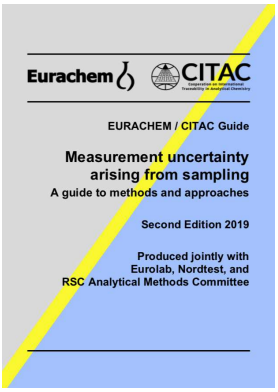
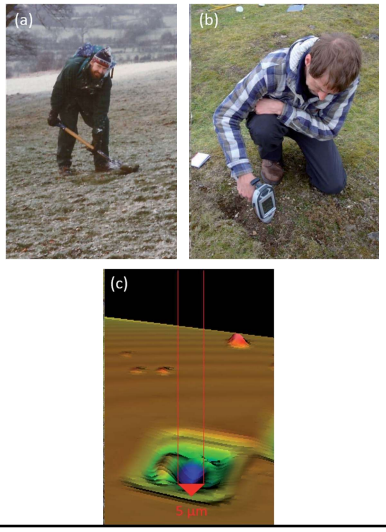
Overview of Eurachem Guide on Measurement Uncertainty Arising from Sampling

Prof. Michael H Ramsey
Chair of Eurachem UfS Working Group
School of Life Sciences,
University of Sussex, Brighton, UK
m.h.ramsey@sussex.ac.uk

ISS Workshop on UfS
9-10th March 2023, Rome
40 min + Q&A

University of Sussex





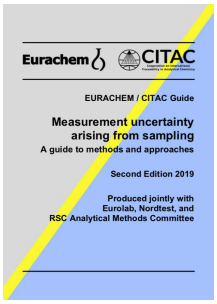
Overview of Talk

- Overview of the Eurachem UfS Guide (New Aspect 2nd Edition)
- Sampling as part of the measurement process
- Uncertainty of measurement values (MU)
  - Why to include the contribution from sampling (UfS)
- Estimation of UfS (& MU) – different approaches
  - Top-down approach – main method used since 1st Edition (2007)
  - Mainly Duplicate Method
    - but also use of Sampling Proficiency Testing
- Benefits of knowing MU (including UfS)
  - Judgement of fitness for purpose (FFP) of measurement values & procedures
    - Enables Validation of Measurement Procedures Including Sampling (VaMPIS)
  - Improved reliability of compliance decisions
- Conclusions

University of Sussex

Overview of the Eurachem UfS Guide

Part: 1. Introduction
2. Fundamental Concepts
   - Sampling as part of the measurement process
   - Uncertainty (U) of measurement values
     - Include contribution from sampling (UfS)
3. Estimation of UfS – different approaches
4. Management Issues
   - Validation and quality control of sampling (Part 4) –→ *Upcoming Supplementary Guide* (SG)
   - Reporting MU
   - Fitness for Purpose of measurements using MU
     - Use in Validation of Measurement Procedures Including Sampling (VaMPIS - SG)
   - Implications for measurement strategies
5. Worked Examples (Appendices A1 – A6)



University of Sussex

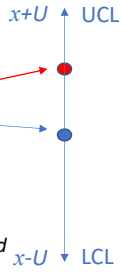
New Aspects of 2nd Edition (2019) Guidance*

1. Unbalanced experimental design (and simplified design) to reduce the cost of estimating UfS - *here*
2. Use of the Uncertainty Factor – *in my second talk*
3. Application of UfS estimation to wider range of measurement types – e.g. *in situ* measurements – *in my third talks*
4. UfS estimation using Sampling Proficiency Testing – *here*

*M H Ramsey, S L R Ellison and P Rostron (eds.) Eurachem/EUROLAB/ CITAC/Nordtest/AMC Guide: *Measurement uncertainty arising from sampling: a guide to methods and approaches.* Second Edition, Eurachem (2019). ISBN (978-0-948926-35-8). http://www.eurachem.org

University of Sussex

What is Measurement Uncertainty (MU)?



- Historially: MU (*U*) is *'an estimate attached to a test results (x)….*
  which characterises the range of values within which the true value is asserted to lie' [1]
  - 'True value' equivalent to '**Value of the Measurand**' in more recent definitions
  - *Parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand.* [2]
  - UCL = Upper Confidence Limit, LCL = Lower Confidence Limit.
  - **Confidence Interval** (CI) is between LCL and UCL
- Includes both Random effects (e.g. precision) and Systematic effects (e.g. bias)
- MU arises from all steps in measurement (e.g. sampling & physical sample preparation)
- Key parameter of measurement (and sampling) quality
- Doesn't assume measurements (or sampling) are 'correct' – traditional approach to Sampling Quality

[1] Historic definition of MU from ISO 3534-1: 1993 Statistics – Vocabulary and Symbols, International Organization for Standardization, Geneva
[2] JCGM 100 (2008) / ISO/IEC Guide 98-3:2008

University of Sussex

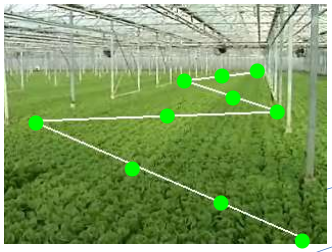
Sampling as part of the measurement process

- Sampling really the first step in the measurement process
- *In situ* measurement techniques reveal this *(Talk #3)*
  - Place the sensor→ make measurement = taking a sample
  - Uncertainty from sampling produces MU in measurement
- Physical sample preparation (in field or lab)
  - e.g. filter, acidify, dry, store, sieve, grind, split
  - is also part of the measurement process
  - and potentially important source of MU
  - include in the validation and QC processes (often omitted by labs)

siebtechnik-tema.com

University of Sussex

Sampling as part of the measurement process

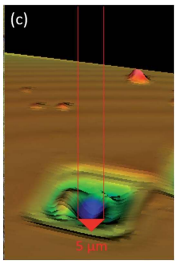


Sampling Target at macro scale

x10

Process step	Form of material	Description of process step
Sampling	Sampling Target	Collection of a single sample, or several increments combined into composite sample
	Primary Sample	Comminution and/or splitting
Physical sample preparation	Sub-sample	Further comminution and/or splitting
	Laboratory sample	Physical preparation, e.g. drying, sieving, milling, splitting, homogenisation
	Test sample	Selection of test portion for chemical treatment preceding chemical analysis
Analysis	Test portion	Chemical treatment leading to analytical determination
	Test solution	Analytical determination of analyte concentration

More careful use of the word 'sample'



Sampling Target at micron scale

University of Sussex

Sampling as part of the measurement process

If the objective is to measure the true value of the analyte concentration (or value of the measurand):-

- in the **sampling target** *(e.g. batch of food, area of soil etc.)*
- Then Sampling is included in measurement process
- U from sampling is part of measurement uncertainty*
  - method validation and QC need to include sampling

If true value (or measurand) defined solely in terms of laboratory sample:-

- sampling is not included

* Most user of analytical measurements assume $x \pm U$ applies to sampling target, not just to lab sample

University of Sussex

Methods for estimating uncertainty of measurement
(including sampling)

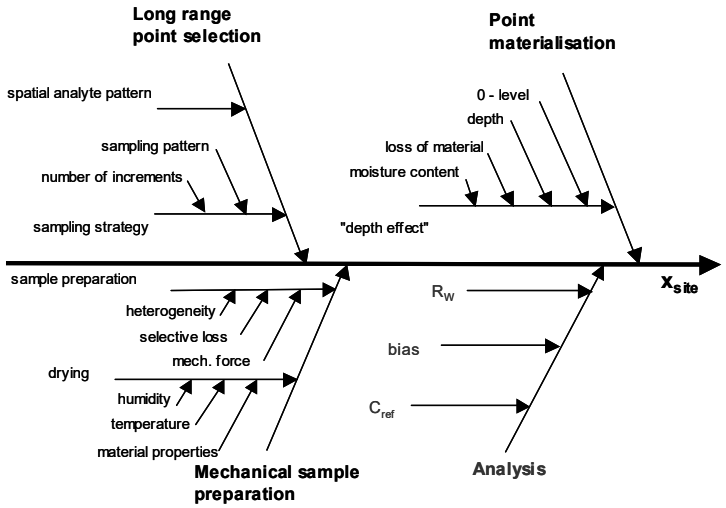
- What are the options?
  - Empirical methods - 'Top down' approach
    - based on replicate measurements (within or between organisations)
    - *applicable to any system*
    - *Examples in the Guide, and this workshop* for food (A1*, A4), soil (A2*) and water(A3)*
  - Modelling methods - 'Bottom up' approach
    - based on identifying, estimating and summing all of the components = 'Budget Modelling Approach' – *Example in Guide for top soil (A6)*
    - sometimes Modelling using Sampling Theory (e.g. Gy's) to estimate components in particulate systems
      - *Example in Guide for animal feed (A5)*

2 further empirical examples in Nordtest Guide i.e. Fe in iron ore, Conductivity in wastewater - (variography)*

** B. Magnusson, M. Krysell, E. Sahlin and T. Näykki, Uncertainty from sampling, Nordtest Report TR 604 (2nd) 2020, ISBN 978-91-89167-31-5. Available from www.nordtest.info*

US University of Sussex

Budget Modelling Approach
to estimating U - Cause & effect diagram



R_w is within-laboratory reproducibility

US University of Sussex

Budget Modelling Approach *to estimating MU*

Summation of all individual components of uncertainty

-e.g. applied to concentration of Cd and P in field of arable top soils

$$\bar{x}_{site} = \bar{x}_{anal} \times f_{b-loc} \times f_{strat} \times f_{depth} \times f_{prep} \times f_{dry}$$

- $\bar{x}_{site}$ = measurement result
- $\bar{x}_{anal}$ = mean from the analysis of test samples
- $f_{\text{b-loc}}$ = correction factor for deviation "between locations"
- f_{strat} = correction factor for bias due to sampling strategy
- f_{depth} = correction factor for the "depth effect"
- f_{prep} = correction factor for errors during mechanical sample preparation
- f_{dry} = correction factor for deviation of moisture content

$$u_{site} = \sqrt{u^2_{anly} + u^2_{b\text{-}loc} + u^2_{strat} + u^2_{depth} + u^2_{prep} + u^2_{dry}}$$

Challenging to estimate realistic UfS for many applications - Applied in Example A6

University of Sussex

Statistical model
for ***Empirical*** estimation of uncertainty - One Sampling Target

$$x = X_{true} + \varepsilon_{sampling} + \varepsilon_{analytical}$$

x = ***measured*** **value** of the analyte concentration in one sampling target

X_{true} = ***true*** **value** of the analyte concentration in the sampling target

$\varepsilon_{sampling} + \varepsilon_{analytical}$ = effects on measured concentration from sampling and analysis

Variance (standard deviation squared) **of measurement value** = σ^2_{meas}

$$\sigma^2_{meas} = \sigma^2_{sampling} + \sigma^2_{analytical}$$

$\sigma^2_{sampling}$ is the between-sample variance on one target (largely due to analyte heterogeneity)

$\sigma^2_{analytical}$ is the between-analysis variance on one sample (as Repeatability)

For ***estimates*** of variance, we have:

$$s^2_{meas} = s^2_{sampling} + s^2_{analytical}$$

University of Sussex

Statistical model

for ***Empirical*** estimation of uncertainty - Multiple Sampling Targets

Multiple sampling targets (n > 8) are needed for realistic estimate of MU & UfS

$$x = X_{true} + \varepsilon_{target} + \varepsilon_{sampling} + \varepsilon_{analytical}$$

ε_{target} represents the variation of concentration between the targets and has variance $\sigma^2_{between-target}$.

Variance of measurement value = σ^2_{meas} $\quad \sigma^2_{meas} = \sigma^2_{sampling} + \sigma^2_{analytical}$

$$\sigma^2_{total} = \sigma^2_{between-target} + \sigma^2_{sampling} + \sigma^2_{analytical}$$

For our estimates of variances, we have:-

$$s^2_{total} = s^2_{between-target} + s^2_{sampling} + s^2_{analytical}$$

s^2_{meas}

University of Sussex

Separate the component variances using ANOVA (Analysis of Variance)

- Typically on results of the Duplicate Method – Balanced Design

  – e.g. using RANOVA3 (free from AMC*)
- Two options:
  - **Classical ANOVA** – assumes frequency distributions are Normal
    - but we often have a small proportion of outlying values (analytical, sampling or targets) – so we often need:-
  - **Robust ANOVA** – 'accommodates' up to 10% outlying values

* https://www.rsc.org/membership-and-community/connect-with-others/join-scientific-networks/subject-communities/analytical-science-community/amc/software/

University of Sussex

How MU is expressed & reported

- MU usually expressed using standard deviation (*s*), e.g.:-

1. Standard uncertainty (*u*)

$u = s_{meas}$ (often = $s_{analytical}$)

2. Expanded uncertainty (*U*)

$U = ks_{meas} = 2s_{meas}$

with coverage factor (*k*) of 2 for 95% confidence

3. Expanded relative uncertainty (*U'*)

$$U' = 100\frac{2s_{meas}}{x}\%$$

– for measurement value (x)

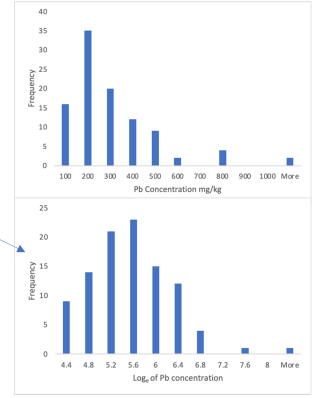
MU can also be expressed as a Confidence Interval, e.g. $= x \pm U$

MU expressed as Uncertainty Factor

4. Uncertainty Factor (^{F}U)

$$^{F}U = \exp(2s_{G,meas})$$

– $s_{G.meas}$ is SD of $\log_e$–transformed measurement values [1]



Confidence Interval $= x$ ×/ ^{F}U

×/ called *'times over'*

- *Discussed in 2nd Talk*

[1] What is the uncertainty factor? Eurachem-AMC Information Leaflet, May 2021
https://www.eurachem.org/index.php/publications/leaflets/uncertainty-factor

Four empirical methods for estimating uncertainty
including that from sampling

Method #	Method description	Samplers (People)	Protocols/ Procedures	Component estimated			
				Sampling Precision	Sampling Bias	Anal. Precision	Anal. Bias
1	Duplicates	single	single	Yes	No	Yes[3]	No[1]
2	Multiple protocols	single	multiple	between protocols		Yes[3]	No [1]
3	CTS	multiple	single	between samplers		Yes	Yes [2]
4	SPT	multiple	multiple	between protocols +between samplers		Yes	Yes [2]

Example of SPT later

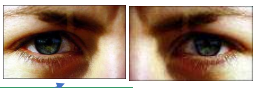
CTS = Collaborative Trial in Sampling , and SPT = Sampling Proficiency Test.

Simplest Empirical method is 'Duplicate Method' (#1) – *applied in Examples A1, A2, A3, A4*

1 estimate analytical bias using CRM, 2 Analytical bias partially or completely included where multiple labs involved 3 Repeatability conditions

US University of Sussex

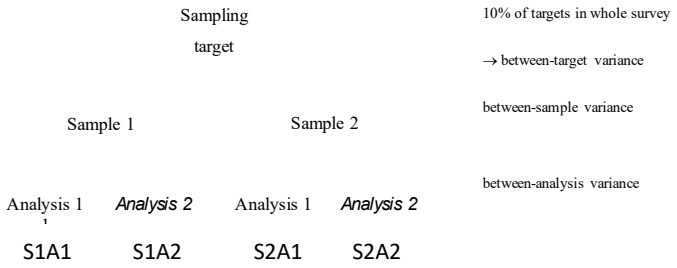
Duplicate Method of UfS Estimation – General Principles



- Duplication is most cost-effective form of replication
  - Apply to both duplicate samples and duplicate chemical analyses
    - using two-stage nested experimental design (balanced or unbalanced)
  - But can have large confidence interval of resulting estimates of MU
  - Unless it is applied to at least 8 sampling targets (ideally more, e.g. 20)
- Realistic taking of duplicate samples is crucial
  - Not just the splitting of a single sample
- Take duplicate samples independently by fresh interpretation of the sampling procedure
  - How far away (in space or time) might duplicate sample be taken? Reflect..
    - ambiguity in sampling procedure
    - spatial uncertainty in the surveying device in use
    - Examples below for Nitrate in lettuce (A1), Pb in soil (A2)

US University of Sussex

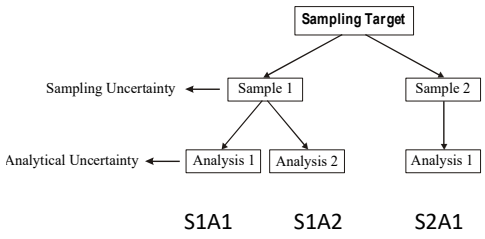
Estimation of MU (including UfS)
Using Duplicate Method – **Full Balanced Design**

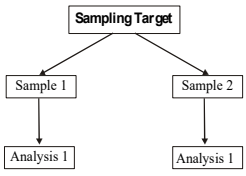


- Usually uses this full balanced experimental design
- Sampling Target = Portion of material, at a particular time, that the sample is intended to represent
- Only requires one 'sampler' (or measurement scientist)
  - Can be improved using multiple 'samplers' - using SPT results *(see later slide, and UfS Guide)*
- Explain Duplicate Method for Case Studies – followed by ANOVA
  - Applicable to both *ex situ* and *in situ* measurement methods

University of Sussex

Alternative designs for UfS estimation
using Duplicate Method at reduced cost





- **Unbalance design**
- Saves 33% of extra cost for UfS estimation
- Larger CI on U_{anal} – not limiting factor
- *Allowed for in RANOVA2 & 3*

- **Simplified balanced design**
- Saves 66% of extra cost, but..
- Only gives estimate of U_{meas}
- Needs external estimate of U_{anal}
- UfS=U_{samp} calculate by subtraction

Appendix D.2 of Guide

University of Sussex

Benefits of knowing MU (including UfS)

1. Validation of measurement procedures including sampling (VaMPIS)
   - Using judgement of fitness for purpose (FFP) of measurement values & procedures
2. Improved reliability of compliance decisions with more realistic MU
   - Explain with Examples

University of Sussex

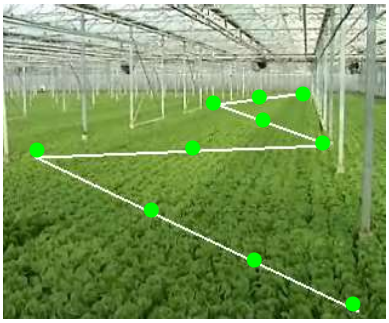
Example A1 from Eurachem UfS Guide:
Nitrate Concentration in Lettuce + Validation

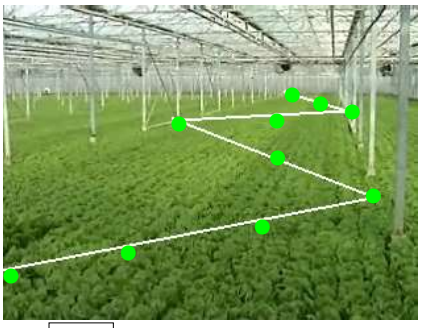
- EU threshold 4500 mg kg^{-1} for nitrate concentration of Sampling Target[1]
  - i.e. ~ 12,000 – 20,000 heads in each bay/batch/target – *full details in next talk*
- Current EU sampling procedure[2] specifies taking 10 heads (increments)
  - to make a single **composite sample** from each Sampling Target
- Analytical procedure/method (HPLC[3]) already validated using Collaborative Trial[4]
  - $U_{analysis}$ around 6% at that validation (RSDR = ~ 3%)
- Need to validate the whole measurement procedure (including sampling & sample prep)
- MU is key metric that affects compliance decisions
  - MU is affected by (and reflects) all of metrics for the measurement procedure
    - precision, bias, LOD, working range, selectivity, sensitivity, ruggedness
  - how much MU from the sampling (UfS)?
- Judge FFP of measurement procedure by the MU - is it close to Target MU?

1. Commission Regulation (EC) No 563/2002 of 2 April 2002 amending Regulation (EC) No 466/2001
2. European Directive 79/700/EEC. OJ L 207, 15.8.1979, p26.
3. BS EN 12014-2:1997, Foodstuffs. Determination of nitrate and/or nitrite content. General considerations
4. Farrington et al.,(2006), Journal of the Association of Public Analysts (Online), 34, 1-11

University of Sussex

UfS estimation for Lettuce using Duplicated 'W' Sampling Design





Sampling target

Sample 1 Sample 2

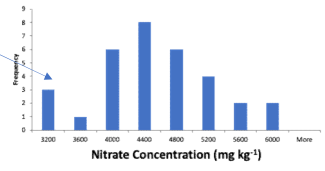
Analysis 1 Analysis 2 Analysis 1 *Analysis 2*

Duplicate sample is equally likely interpretation of 'W' design

US University of Sussex

Estimating UfS (and MU) with Duplicate Method

- Use Duplicate Method – *as described earlier in talk*
- Selected 8 typical sampling targets (bays of ~20,000 lettuce) – sampled in duplicate
  - see measurement values in talk on Example A2
- Estimated UfS and MU using Analysis of Variance (ANOVA) – program RANOVA3* *(output in later talk)*
- Selected Robust ANOVA - as it *accommodates < 10% outlying values*
- $U' = 16.4\%$ ($u = s_{meas} = 360$ mg kg^{-1}) – *as repeatability*
  - MU dominated by UfS (78% of MU)
  - UfS mainly caused by nitrate heterogeneity within the sampling targets
- $U'_{anal} = 7.6\%$ – *as repeatability*
  - Very similar to MU = 6 % reported at separate validation of the analytical procedure[4]



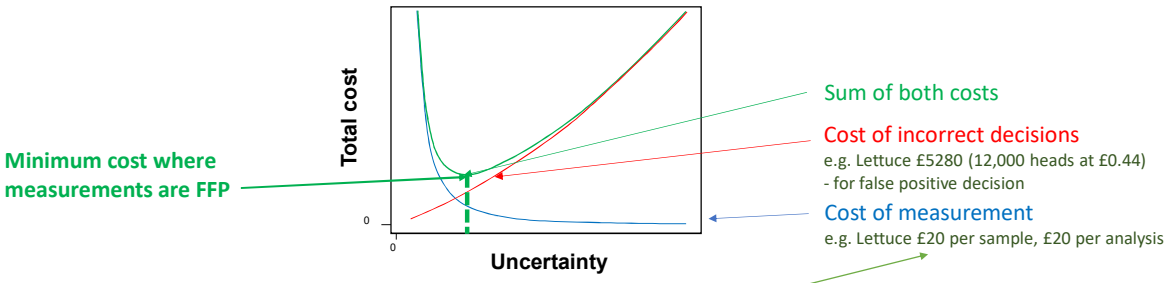
* http://www.rsc.org/Membership/Networking/InterestGroups/Analytical/AMC/Software/

Lyn, J.A., Palestra, I.M., Ramsey, M.H., Damant, A.P. and Wood, R. (2007) Modifying uncertainty from sampling to achieve fitness for purpose: a case study on nitrate in lettuce *Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement*, 12, 67-74

US University of Sussex

FFP at Target MU – e.g. set at minimum overall cost

- **Validation by judging Fitness for Purpose** (FFP) - *Section 16 of UfS Guide*
- Target MU - can be Option (1) set externally (e.g. arbitrary 20%, 16% < 20% so FFP), or Option (2)…
- At **Optimal MU*** that minimises the overall cost (including the consequences of incorrect decisions)



- By knowing UfS, can judge how Target MU (however set) can be achieved most cost-effectively by:
  - Spending more (or less) on **chemical analysis** (e.g. more precise technique), or
  - Spending more (or less) on **sampling** (e.g. taking more increments)

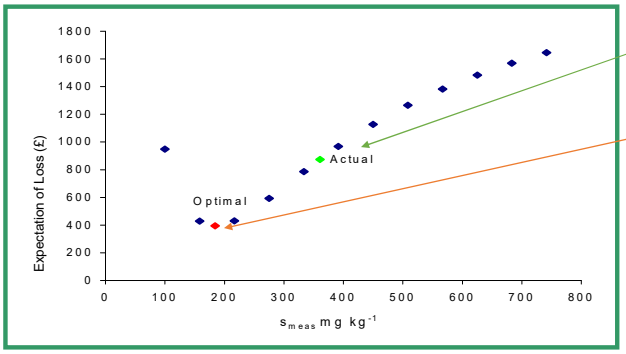
Optimal MU* method explained more in 3rd Talk

University of Sussex

Judge FFP - level of Uncertainty

- For lettuce example estimate MU (s_{meas}) using Duplicate Method
- Calculate Target MU using optimised uncertainty (OU) method*
- **Measurement Procedure is judged as NOT FFP**

Cost ↑



Uncertainty→

Actual MU (360 mg kg^{-1}) i.e. *U'* = 16.4%
- and consequent cost (£800 per target)
is much higher than…

Optimal MU value (184 mg kg^{-1}) i.e. *U'* = 8.3%
At minimum cost (£400)

To achieve FFP - we need to reduce the MU by factor of 2
UfS accounts for 78% of MU (from ANOVA)
- So reducing UfS is most cost-effective

Sampling Theory predicts we can reduce UfS x2 by increasing sample mass by factor of 4 (= 2^2)

So take composite sample with 40 heads instead of 10 heads – to make FFP

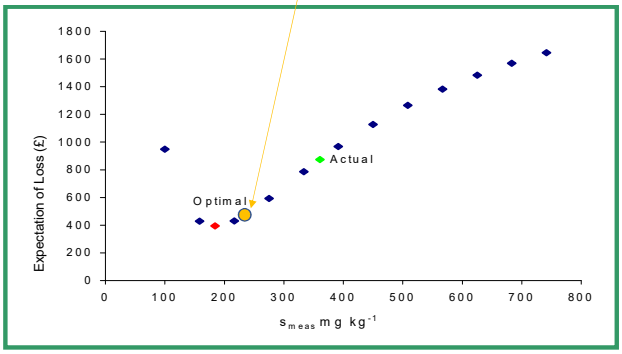
* In upcoming SG-VaMPIS – not in UfS Guide

University of Sussex

Reducing the Uncertainty – to achieve FFP

- **Increasing number of increments from 10 to 40 heads**
- Reduces s_{samp} from 319 to 177 mg kg^{-1} - by a factor of x 1.8 (similar to model prediction of x2)
- Reduces MU (s_{meas}) from 360 to 244 mg kg^{-1}. (*U'* from 16.4 % to 11.1%)
- Close to the optimal value (184 mg kg^{-1}) at similar Cost (~£500, down from £800 per target)
- A**chieves Fitness-for-Purpose (FFP)** = MU that minimises to overall financial loss

Cost ↑

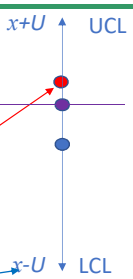


Uncertainty→

Lyn, J.A., Palestra, I.M., Ramsey, M.H., Damant, A.P. and Wood, R. (2007) Modifying uncertainty from sampling to achieve fitness for purpose: a case study on nitrate in lettuce *Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement*, 12, 67-74

University of Sussex

Making compliance decisions more reliable – using UfS



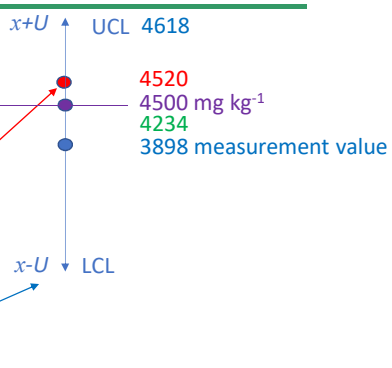
- Measurement value (*x*)
- and its MU ($U_{analytical}$)
  = range within which true value lies [1]
- Threshold value for compliance
- Appears that true value cannot exceed the Threshold
- = measurement indicates **Compliant** target/batch
- Use MU that includes UfS
- True value of one target/batch
- Over Threshold, therefore measurement indicates **Non-Compliant** target/batch
- Such Non-Compliance appears impossible with MU based only upon $U_{analytical}$
- **Only by including UfS within MU can non-compliant batches be rejected reliably**
- **It is therefore essential to include UfS - to make a realistic estimate of MU**
  - Analytical MU alone may not include the true value

[1] Historic definition of MU from ISO 3534-1: 1993 Statistics – Vocabulary and Symbols, International Organization for Standardization, Geneva

University of Sussex

Lettuce Example of Compliance decision using UfS

Nitrate in Lettuce (Example A1 – Target A)



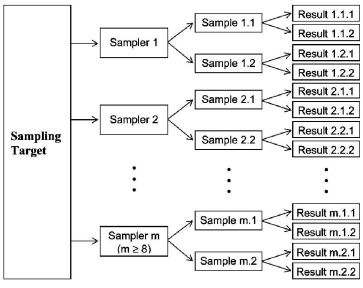
- Measurement value *(x)* 3898 for Target A (S1A1)
- and its MU ($U_{analytical}$) u_{anal} = 168, U = 336 mg kg^{-1}
  = range within which true value lies [1]
- Threshold value for compliance 4500 mg kg^{-1}
- Appears that true value cannot exceed the Threshold
- = measurement indicates Target A is **Compliant**
- Use MU that includes UfS u_{meas} = 360, U = 720 mg kg^{-1}
- True value possible for Target A 4520
- Over Threshold, therefore measurement indicates Target A is **Non-Compliant**
- Non-Compliance of Target A appears impossible with MU based only upon $U_{analytical}$
- **Only by including UfS within MU can non-compliant batch (Target A) be rejected reliably**

[1] Historic definition of MU from ISO 3534-1: 1993 Statistics – Vocabulary and Symbols, International Organization for Standardization, Geneva

University of Sussex

Estimate of Uncertainty using SPT - including Between-Sampler Bias

- Example using Sampling PT for moisture in butter*



* Ramsey M.H. Geelhoed B, Damant, A.P., Wood, R. (2011) Improved evaluation of measurement uncertainty from sampling by inclusion of between-sampler bias using sampling proficiency testing. Analyst, 136 (7), 1313 – 1321. DOI:10.1039/C0AN00705F.

ANOVA: U' as % on concentration of moisture in butter (200 tons)

≈ Duplicate Method (single sampler) gives U' = **0.39 %**

SPT (multiple samplers, n=9) gives U' = 0.87%

- *U' larger* x 2.2* - includes bias between-samplers

Remove two samplers with potentially non-proficient z-scores (RSz > 3)

SPT (n=7) gives U' = 0.69%

- **U' still larger x 1.8**
- **a more reliable estimate of Uncertainty**
- Ideally apply over multiple rounds of SPT, if targets comparable
- e.g. 16 rounds, stack-gas measurement SPT [Coleman et al ,2013, Accred Qual Assur 18:517–524]
- Multiple samplers (e.g. in CTS) better for Validation of Sampling
- More expensive than Duplicate Method, but sometimes justified

University of Sussex

Conclusions

- Eurachem UfS Guide explains importance of UfS (& MU), and how to estimate it
- Including sampling within the measurement process:
  - Is essential for making reliable estimates of MU (including UfS)
    - E.g. For Compliance Decisions: e.g. are concentration levels above from regulatory limits?
    - Conforms to ISO/IEC 17025:2017
  - Being able to judge FFP, and hence validate the whole measurement process
  - Hence rigorous Validation of the whole Measurement Process (Including Sampling)
    - Upcoming Supplementary Guidance on VaMPIS
- UfS (and hence MU) can be estimated with Duplicate Method *(most practical)*
  - Applicable to any sampling medium: soil, sediment, herbage, waters, gases etc.
  - Also applicable to *in situ* measurements (such as PXRF) (*Talk 3*)
  - Sampling PT (or CT) results can be used to also include between-sampler bias within MU
- *Questions?*