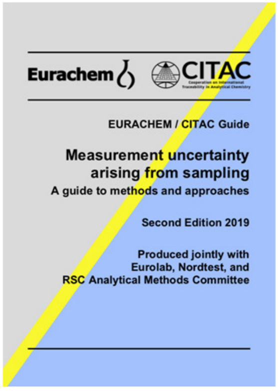
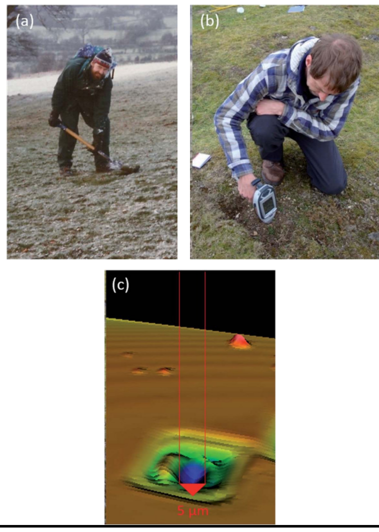
Overview of Uncertainty from Sampling and the Eurachem UfS Guide (2019)

Michael H Ramsey
School of Life Sciences,
University of Sussex, Brighton, UK

Eurachem/Eurolab Workshop,
Uncertainty from sampling and analysis for accredited laboratories
November 2019, Berlin

University of Sussex





Overview

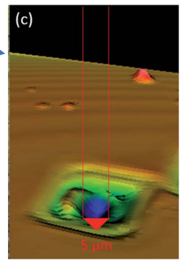
- Sampling as part of the measurement process
- Uncertainty (U) in measurement and sampling -
  - *key parameter of measurement (and sampling) quality*
  - *Brief overview of UfS estimation*
- Rationale for revision of UfS Guidance (2019)
- Focus on <u>New Aspects</u> to Guidance
  1. Uncertainty Factor – *explained in my later talk*
  2. Unbalanced experimental design to reduce the cost of estimating UfS - *in later talk (by Peter Rostron)*
  3. Application of UfS estimation to wider range of measurement types – e.g. *in situ*, passive and micro-scale measurements – *in other talks*
  4. UfS estimation using Sampling Proficiency Testing - *here*
- Conclusions

University of Sussex

Sampling as part of the measurement process

- Sampling is really the first step in the measurement process (traditional sampling at the macro scale, e.g. soil)
- *In situ* measurement techniques – sampling integral
  - Place the sensor→ make measurement
    - taking a ‘beam’ sample at micro scale (e.g. mm or μm)
  - Uncertainty in sampling produces U in measurement value
- 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 U
- Include both in validation and QC processes

hand-held portable X-ray Fluorescence (pXRF) on soil at 5 mm scale

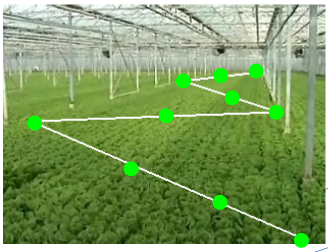


Secondary Ion Mass Spectrometry (SIMS) on quartz, illustrating 5 μm beam scale

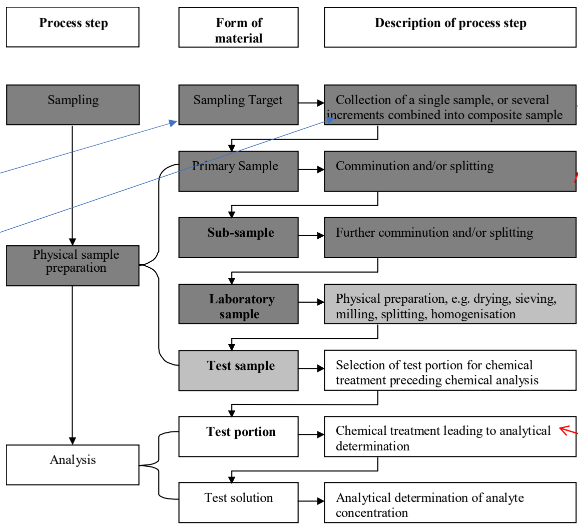
Analytical Methods Committee (2018) AMC Technical Brief No 84. Beam sampling: taking samples at the micro-scale, ***Analytical Methods*, 10, 1100-1102**

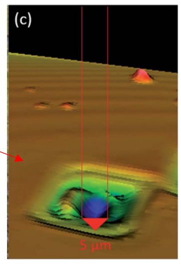
University of Sussex

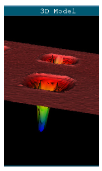
Sampling as part of the measurement process



x10







Primary sample
= Test portion mass from SIMS crater ~ 300-350 pg

University of Sussex

Sampling as part of the measurement process

If objective is to measure true value of analyte concentration (or measurand)
- in **sampling target** *(e.g. batch of food, area of soil, a crystal etc.)*
- Sampling is included in measurement process
- UfS part of measurement uncertainty (& method validation and QC)

If measurand (or true value) defined solely in terms of laboratory sample
- Primary sampling is not included

- Most <u>users</u> of analytical measurements assume x ± U apply to target
  - not just to lab sample

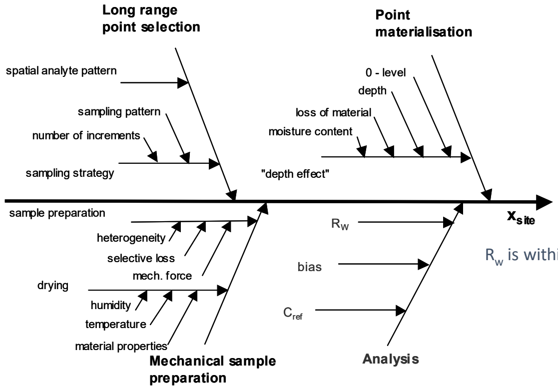
US University of Sussex

Methods for estimating uncertainty of measurement *(including sampling)*

- What are the options?
  1. Empirical methods - 'Top down' approach
     - based on replicate measurements (within or between organisations)
     - *applicable to any system*
     - *Examples in the Guide – for food (A1, A4), soil (A2) and water(A3)*
  2. 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)*
       - (Kurfurst *et al*, 2004, Accred Qual Assur., 9, 64-75)
     - sometimes Modelling using Sampling Theory (e.g. Gy's) to estimate components in particulate systems
       - (Minkkinen 2004, Chemometrics and Intelligent Lab. Systems, 74, 85-94)
       - *Example in Guide for animal feed (A5)*

US University of Sussex

Budget Modelling Approach *to estimating U* - *Cause & effect diagram*



R_w is within-laboratory reproducibility

Example A6 for top soil in UfS Guide

University of Sussex

Budget Modelling Approach *to estimating U*

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

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

Explained by Ulrich Kurfürst in Example A6

University of Sussex

Modelling using Sampling Theory

Sampling theory of Gy defines 8 sampling errors

- includes 'fundamental sampling error' described by:-

$$\sigma_r^2 = Cd^3\left(\frac{1}{M_S} - \frac{1}{M_L}\right)$$

$\sigma_r = \frac{\sigma_a}{a_L}$ = Relative standard deviation of the fundamental sampling error = $u'_{sampling}$

σ_a= absolute standard deviation (in concentration units)

a_L = average concentration of the lot

d = characteristic particle size = 95 % upper limit of the size distribution

M_S = Sample size

M_L = Lot size

Explained by Pentti Minkkinen in Example A5 for animal feed

General relationship: **$s^2 \propto 1/m$** – useful in modifying UfS estimated by any method

University of Sussex

Empirical estimation of uncertainty: Statistical model

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

x = *measured* value of the analyte concentration in the 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 of measurement = $s^2_{meas} = s^2_{sampling} + s^2_{analytical}$

- includes between-organisational effects (e.g. sampling & analytical bias)

Standard uncertainty = $u_{meas} = s_{meas}$

Relative expanded uncertainty as % (for 95% confidence) = $U' = 100 \times 2s_{meas}/\bar{x}$

University of Sussex

Empirical methods for estimating uncertainty of measurement *(including sampling)*

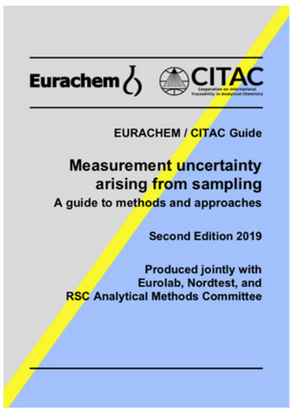
Method #	Method description	Samplers (persons)	Protocols	Component estimated		
				Sampling		Analytic
				Precision	Bias	Precision
1	Duplicates	Single	Single	Yes	No	Yes[3]
2	Protocols	Single	Multiple	Between protocols		Yes[3]
3	CTS	Multiple	Single	Between samplers		Yes
4	SPT	Multiple	Multiple	Between protocols +between samplers		Yes

- [1]Analytical bias information may be obtained by including certified reference materials in the analytical run (see Example A2, Appendix A).
- [2]Analytical bias is partially or completely included in collaborative exercises where multiple laboratories are involved.
- [3]In these approaches, precision is estimated under repeatability conditions

- Examples of Method #1 (*see later talk*) and Method #4 (using SPT) *discussed here*

Rationale for revision of UfS Guidance

- Second Edition of the Eurachem UfS Guide*
  - initiated to update, explain and integrate several recent research developments
  - whilst leaving most of the text unchanged
- Retains the same basic approach and structure as First Edition of 2007
  - based on concept that primary sampling as first part of measurement process
  - two main approaches to estimating UfS
- Six worked examples retained
  - demonstrate both approaches
  - across a range of application sectors, including food, animal feed, soil and water.
  - two of the examples partially updated to illustrate some of research developments
- Four main aspects of new developments included in Second Edition:-



*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). Available from http://www.eurachem.org

1. Using Uncertainty Factor to express MU and UfS

- Uncertainty Factor (^{F}U) is an alternative way to express measurement uncertainty.
- Upper and lower confidence limits (UCL & LCL) of a measurement value expressed by:-
  - multiplying and dividing the measurement value by the uncertainty factor
  - e.g. For measurement value of 5 mg/kg,

    $^{F}U = 2.5$ UCL = 5 * 2.5 = 12.5 mg/kg, LCL = 5/2.5 = 2 mg/kg

  Contrasts against traditional approach of adding and subtracting the uncertainty.
  - e.g. For measurement value of 5 mg/kg.

    $U = 2.5$ mg/kg UCL = 5 + 2.5 = 7.5 mg/kg, LCL = 5 – 2.5 = 2.5 mg/kg

US University of Sussex

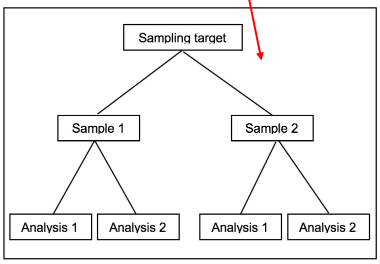
New Aspect 1. Using Uncertainty Factor to express MU and UfS

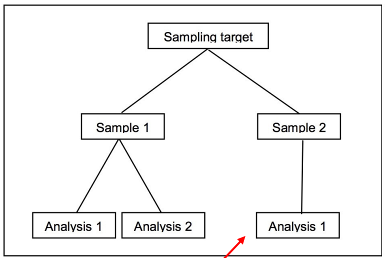
- Uncertainty Factor (^{F}U) is approach more accurate when:-
  - relative expanded uncertainty value is large (e.g. >20%),
  - frequency distribution of uncertainty is approximately log-normal, rather than normal.
- Both conditions often apply to measurement uncertainty that arises from sampling process,
  - particularly when spatial distribution of analyte in test material is substantially heterogeneous.
- Also for some purely analytical systems (e.g. GMO in soya, with $u' = 70\%$)
- Guide also explains how measurement uncertainty can be calculated
  - by adding the component arising from sampling, expressed as an uncertainty factor (^{F}U)
  - with that arising from chemical analysis, expressed as relative uncertainty (U')
- *More details of these issues in my later talk*

US University of Sussex

New Aspect 2. Unbalanced design to reduce cost of estimating UfS

- First edition used 'duplicate method' with a balanced experimental design for estimation of:-
  - measurement uncertainty as a whole
  - and its two components in sampling and analytical steps
  - using balanced design (analytical duplicates on both of the two sample duplicates)





- Second edition stresses the advantage of using unbalanced design,
  - analytical dreduces the extra cost uplicate on only one of the two sample duplicates.
  - of estimating the uncertainty by 33%.
  - *details in later talk by Peter Rostron*

University of Sussex

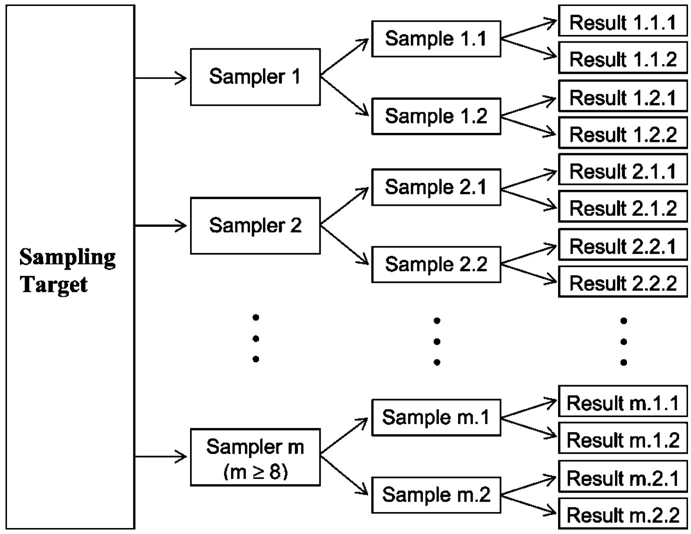
New Aspect 4. UfS estimation using Sampling Proficiency Testing

- First Edition of UfS Guide this approach was discussed in theory
- Second Edition now refers to the first practical example
- Multiple samplers each apply whatever sampling protocol they consider appropriate
  - to achieve the same stated objective for the same sampling target
- Using a balanced design across all of the different samplers (next slide)
- Then possible to include 'between-sampler' bias in estimate of UfS
  - in addition to the components that were previously included
- General principles of SPTs*
- First practical SPT used to estimate UfS**
  - concerned measurement of moisture content of 20 ton batch of fresh butter

-***Proficiency testing of sampling.** Technical Brief 78, July 2017, Anal. Methods, 2017, 9, 4110, DOI: 10.1039/c7ay90092a
-https://pubs.rsc.org/en/content/articlehtml/2017/ay/c7ay90092a
**M H Ramsey. B Geelhoed, A P Damant, R Wood (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

University of Sussex

Experimental Design of SPTs



- Each sampler takes two samples,
  - both analysed twice
  - In balanced design
- Allows effects of analysis to be removed from those of sampling
- Multiple targets measured
  - in different rounds of the SPT

University of Sussex

σ_{pt}

Scoring an SPT – *in general*

$$z = (x - x_{pt}) / \sigma_{pt}$$

x = submitted result

e.g. participants estimated mean concentration of the sampling target

x_{pt} is assigned value

- independent of the result, either…
  - by expert/prior measurement
  - by spiking
  - by consensus

σ_{pt} is fitness for purpose (FFP) criterion

- independent of the result, either…
  - from external FFP requirement, or
  - from internal information (e.g. $s_{within\text{-}participant}$)

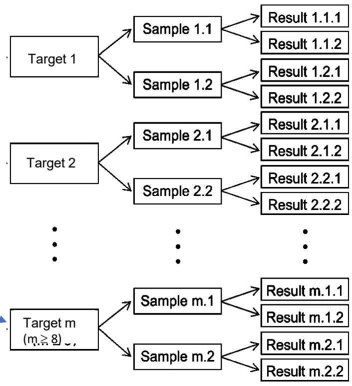
University of Sussex

Uncertainty estimation from SPT

- Use ANOVA to separate the component variances (s^2)

$$s^2_{total} = s^2_{between-sampler} + s^2_{within-sampler} + s^2_{analytical}$$

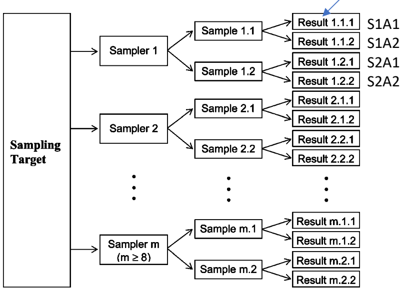
- Use s^2_{total} as estimate of the measurement uncertainty
- Extra component is between-sampler effect
  - includes sampling bias – between-samplers
    - comes out as random effect in this model
  - not estimated by the **'Duplicate method'** usually used
    - by <u>single</u> sampler on 'm' different targets (m ≥ 8)
    - using balanced design
- UfS from Duplicate Method – for single sampler
  - estimated for butter using balanced design on single sampler
  - averaged across all samplers



University of Sussex

SPT on % Moisture in Fresh Butter

- Target: 20 tons of fresh butter in boxes – 6-fold composite sample
- Standard experimental design of SPT
- 9 samplers (A – I) operating independently
- % Moisture determined gravimetrically – by centralized analysis
  - Hence low analytical uncertainty (U' = 0.35%)



Sampler #	S1A1	S1A2	S2A1	S2A2
A	15.4741	15.4155	15.4972	15.4796
B	15.3655	15.3257	15.3653	15.3373
C	15.4417	15.4069	15.4552	15.4518
D	15.4161	15.4134	15.4486	15.4143
E	15.4085	15.3675	15.4392	15.406
F	15.4148	15.3876	15.4176	15.3473
G	15.4959	15.4757	15.4853	15.5185
H	15.3673	15.3732	15.372	15.3427
I	15.3214	15.2779	15.3424	15.3721

University of Sussex

Z- scores for SPT on Moisture in Fresh Butter

- z- scores calculated for both measurements (m.1.1 & m.1.2)
  - for two samples (m.1 and m.2) taken by each sampler (m)

Sampler	First z-score m.1.1	Second z-score m.1.2	Third z-score m.2.1	Fourth z-score m.2.2	Rescaled sum of z-scores = $\Sigma z/\sqrt{n}$ = $\Sigma z/\sqrt{4}$
A	1.36	0.18	1.83	1.47	2.42
B	-0.83	-1.64	-0.84	-1.40	-2.35
C	0.71	0.00	0.98	0.91	1.30
D	0.19	0.14	0.85	0.15	0.66
E	0.04	-0.79	0.66	-0.01	-0.06
F	0.16	-0.39	0.22	-1.20	-0.60
G	1.80	1.39	1.59	**2.26**	**3.52**
H	-0.80	-0.68	-0.70	-1.29	-1.73
I	-1.72	**-2.60**	-1.30	-0.70	**-3.16**

- Two samplers had potentially non-proficient RSz-scores* (> 3)
- Samplers; I (z = -3.2) & G (z = +3.5) – in opposite directions
- Video evidence suggests might be related to angle of sampling device

*AMC, Proficiency testing of analytical laboratories: organization and statistical assessment, Analyst, 1992, 117, 97–117.

University of Sussex

Estimate of Uncertainty using SPT - including Between-Sampler Bias

- Example using SPT for moisture in butter

ANOVA: U' as % on concentration of moisture in butter

≈ 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)

Samplers; I (z = -3.2) & G (z = +3.5)

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

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.

*Whether U′ values are significantly different – talk by Peter Rostron in 'Methods' Session

University of Sussex

Example of SPT using *in situ* measurements
NPL's Gas Measurement PT scheme

- Measurement of combustion gases
- From stack simulator facility at NPL (for sampling component)
  - and directly from cylinders (for analytical component)
- Concentrations and conditions are typical of plant
  - falling under Industrial Emissions Directive (IED)

Test	**NOx (ppm)**	**VOC (ppm)**	**SO_2 (ppm)**	**CO (ppm)**	**O_2 (%)**	**H_2O (%)**
Max Range	270	10	200	100	17	20

http://www.stack-pt-schemes.net/?page=stack_gases

- Results from 16 rounds published*
- (S)PT score used in accreditation
- UfS not yet estimates
- Combination of sampling and *in situ* analytical steps (i.e. measurement)
  - Better described as **Measurement Proficiency Test (MPT)** ?

*MD. Coleman et al, State of UK emissions monitoring of stacks and flues: an evaluation of proficiency testing data for SO2, NO and particulates, Accred Qual Assur (2013) 18:517–524, DOI 10.1007/s00769-013-1011-xhttp://www.npl.co.uk/measurement-services/environmental-monitoring/stack-proficiency-testing-scheme

University of Sussex

Review of published SPTs since 1995

Target Medium Synthetic target? (if in bold)		Analytes Central Lab? In situ test?	No. of Participants	No. of rounds	Ref	Comment
Soil	Top-Soil	Pb & Cu	9	1	[3]	First reported SPT realisation.
	Top-Soil	Pb & Cu	9	2	[6]	Improved performance in second round (s down 36% to 20%)
	Top-Soil	Ba	9	1	[7]	Spatial resolved SPT to locate 'hot spot'. Synthetic Target
Air	**Workplace**	Hydrocarbons Aromatic & Chlorinated	38	3	[8]	Not called an 'SPT', no z-scores. 3 sampling exercise each with sets of 4-8 runs. Has evidence for improved performance between rounds.
	Landfill-gas	*CH_4, CO_2, O_2*	9	1	[9]	Temporal variability monitored and corrected by reference borehole
	Stack gas	*Gases* & particulates	4	?	[10]	4 sampler at a time, use in-house-constructed emission simulation apparatus. Adapted homogeneity test for ISO 13528 compliance
	Stack gas	*Gases*	15	16	[11]	Run annually. 3 participants at a time, repeat over 4 to 8 days of testing. In-house constructed emission simulation rig allowing the testing of hot and wet test atmospheres. Homogeneity to EN15259.
Food (bulk)	Wheat	N, Mo, Pb	5	1	[12]	All participants used same protocol.
	Coffee	H_2O	8	1	[12]	All participants used same protocol.
	Lettuce 'in field'	NO_3^-	16	1	[13]	First example of virtual SPT, with participants from 16 different countries.
	Butter	H_2O	9	1	[5]	Results used to estimate U, including between-sampler bias.
	Apple Juice	Patulin	9	1	[14]	Not all SPT design criteria were met:- Limited access to the sampling target, ownership of the target, and this being the same as sampler's employer, limited independence of the samplers.
Water	Water-waste	COD, TOC, *pH,Temp*	16-20	3	[15]	*In situ* measurement of pH &Temp. Own labs used as well as central analysis. .

Wide range of sampling media, and analytes

Some targets spiked to derive assigned value

Some analyses done:
- *In situ*
- centrally by one lab

First example of virtual SPT,

Number of participants generally >8, up to 38

No. of Rounds
1 = feasibility assessment, multiple for ongoing SPT

Improved performance detected between rounds

University of Sussex

New #3. Application of UfS estimation to wider range of situations

- These include measurements made:
- (a) *in situ* (e.g. by field sensors without removing a sample)
- (b) on site (e.g. in a field laboratory on a removed sample)
  - Intermediate in complexity between *ex situ* and *in situ*
- (c) passive measurements of radioactive decay, and
- (d) at the microscopic scale (e.g. *In situ* PXRF in mm scale and SIMS at micron scale).
- *Discussed in later talk*

University of Sussex

Conclusions

- UfS needs to be included in estimates of Measurement Uncertainty – in most situations
- There are two main approaches to estimation UfS
  - empirical approach generally more flexible* than modelling approach
- Second edition of the Eurachem UfS Guide has several new aspects *(some discussed in later talks)*
- Use of Sampling Proficiency Testing data can help identify unsuspected* between-sampler bias

University of Sussex

Acknowledgements

Composition of UfS Working Group

Eurachem members

- Mike Ramsey (Chair)
- Steve Ellison (Secretary)
- Paolo de Zorzi (ISPRA, Italy)
- Pentti Minkkinen
- Eskil Sahlin (RISE. Sweden)
- Alex Williams

EUROLAB members

- Irache Visiers (Applus)
- Rudiger Kaus (EUROLAB, DE)

CITAC. Members

- Ilya Kuselman (Israel)
- Jorge Eduardo S Sarkis (INER, Brazil)

Nordtest representative

- Bertil Magnusson (Sweden)

RSC AMC representatives

- Roger Wood
- Peter Rostron

Additional members of RSC/AMC Subcommittee:-

- Bob Barnes
- Mike Thompson

Funding from Analytical Methods Trust (AMT)

US University of Sussex