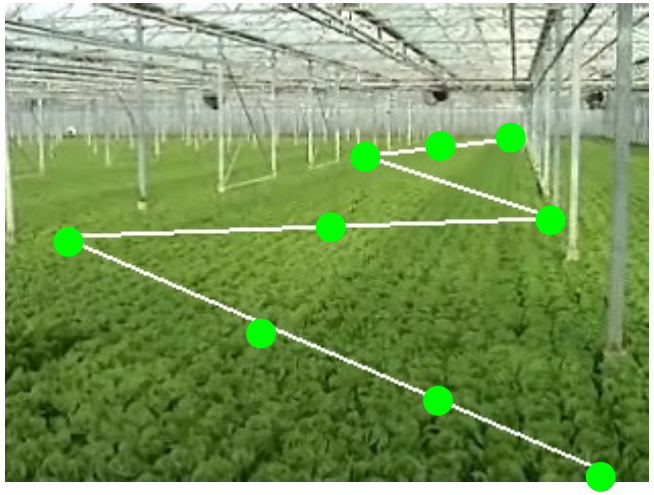
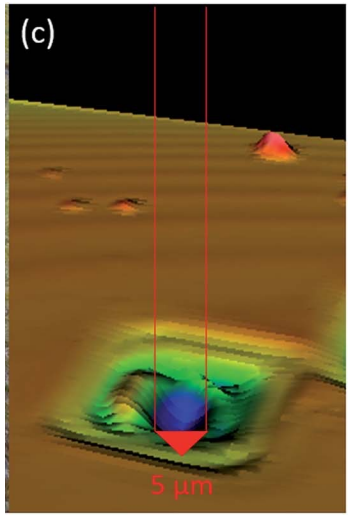
Measurement Uncertainty from Sampling:
an introduction for Early Career Researchers





Prof Michael H Ramsey

School of Life Sciences,

University of Sussex, Brighton, UK

m.h.ramsey@sussex.ac.uk

ECR-SEGH

Uncertainty from sampling

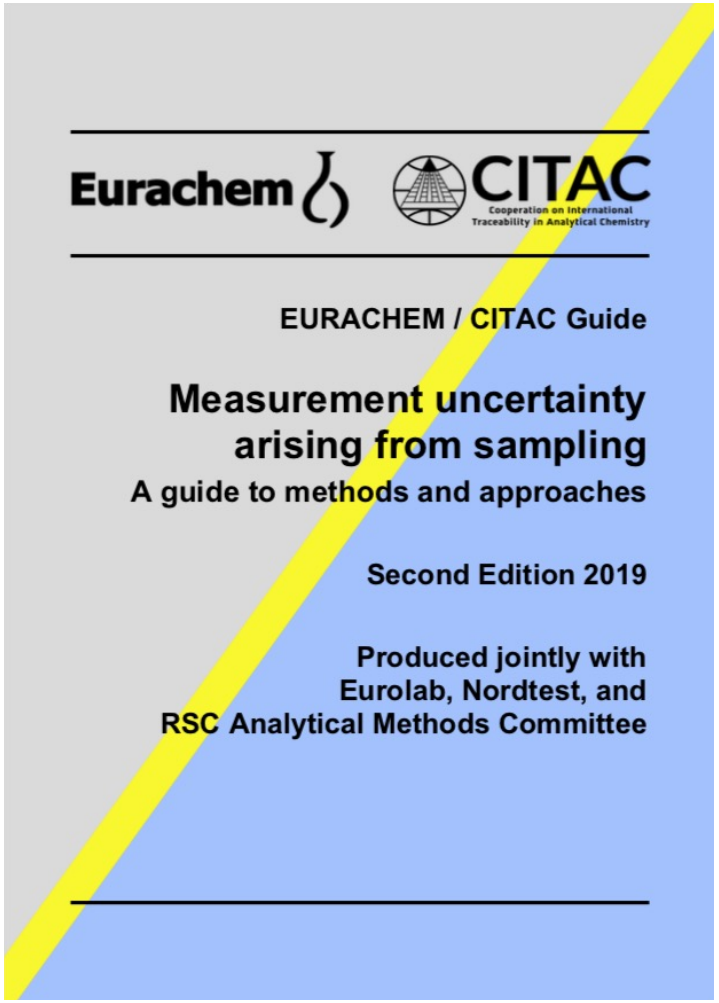
6th May 2022

University of Sussex

Overview

- Uncertainty of measurement (MU) – what is it?
  - Sampling as part of the measurement process
  - Including contribution from sampling (UfS) into MU (not just analytical)
- How to estimate UfS (and MU)
- How to use MU in interpretation of geochemical measurements
- Range of application areas relevant to SEGH
  - soil, sediments, plants (e.g. food), water, gas…..
  - Reference for review of applications of UfS estimation in all these areas
- Worked examples for
  - Nitrate in Lettuce (Normal frequency distribution)
  - Lead in top soil (Log-normal frequency distribution)
  - Application to *in situ* measurements (- no physical sample taken)
    - At macro scale (e.g. PXRF, pH meter) and micro scale (SIMS, SEM-EPMA)
- Conclusions

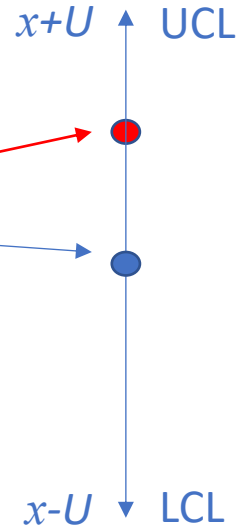
Estimation of UfS & MU – Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide[4]



- Eurachem/EUROLAB UfS Guide[4] describes several methods to estimate UfS
  - six worked examples for quantitative lab measurements made *ex situ* on wide range of analytes, in many different materials (e.g. food, feed, water and soil). Subsequently applied to gases, fuel etc.
- Most widely applicable approach for random components of MU (4 of 6 examples) - is 'Duplicate Method' based on a **balanced design**
  Only needs one sampler
- More sophisticated approach uses multiple samplers
  - e.g. Sampling Proficiency Testing (SPT) results – not covered in this talk

Ramsey M.H., Ellison S. L. R., and Rostron P.(eds.) (2019) Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide: *Measurement uncertainty arising from sampling: a guide to methods and approach,* Second Edition, Eurachem. http://www.eurachem.org/index.php/publications/guides/musamp

Measurement Uncertainty (MU) - including that arising from Sampling (UfS)



- 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[2]
- UfS mainly caused by small-scale heterogeneity of analyte within sampling target, so...
- Research student needs to consider quality of primary sampling
  - as well as quality of instrumental analysis
- Primary metric for expressing quality of a measurement value is its uncertainty
- **It is therefore essential to include UfS - to make a realistic estimate of MU**

[1] Historic definition of MU from ISO 3534-1: 1993 Statistics – Vocabulary and Symbols, International Organization for Standardization, Geneva
[2] Parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand. JCGM 100 (2008) / ISO/IEC Guide 98-3:2008

University of Sussex

Estimation of UfS & MU - Calculations

Analysis of Variance (ANOVA) (in Excel speadsheet RANOVA3*) can quantify 3 components of total variance, s^2_{total}, where *s* is the standard deviation

$$s^2_{total} = s^2_{between-target} + s^2_{sampling} + s^2_{analytical} \qquad (1)$$

- 'between-target' reflects geochemical variation at the larger scale

Standard measurement uncertainty (*u*) arises from combination of sampling and analytical sources:

$$u = s_{meas} = \sqrt{s^2_{sampling} + s^2_{analytical}} \qquad (2)$$

Expanded relative measurement uncertainty with 95% confidence **(*U'*)** for a measurement value (x) given by:

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

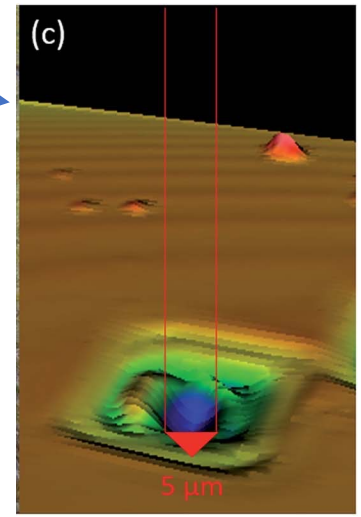
* AMC Software, downloaded gratis from: https://www.rsc.org/Membership/Networking/InterestGroups/Analytical/AMC/Software/

University of Sussex

Sampling as part of the measurement process

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

- 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
- Need to define the Sampling Target:-
  - i.e. 'portion of material, at a particular time, that a sample is intended to represent'
  - e.g. batch of food, area of soil, a crystal etc

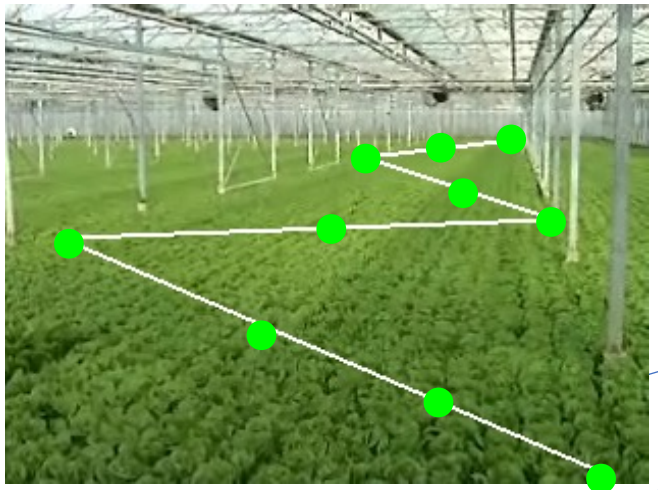


Secondary Ion Mass Spectrometry (SIMS) on quartz, illustrating 5 µm beam scale

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

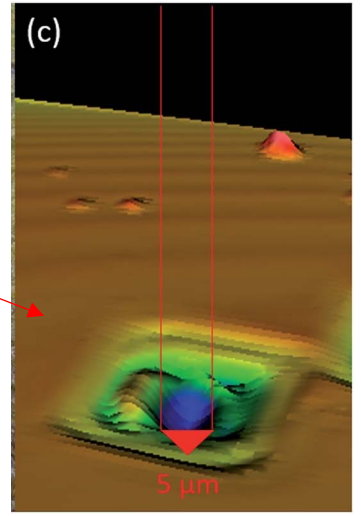
University of Sussex

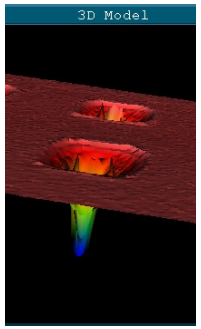
Sampling as part of the measurement process



x10

Process step	Form of material	Description of process step
Sampling	Sampling Target	Collection of a single sample, or several increments combined into composite sample
Physical sample preparation	Primary Sample	Comminution and/or splitting
	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





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

University of Sussex

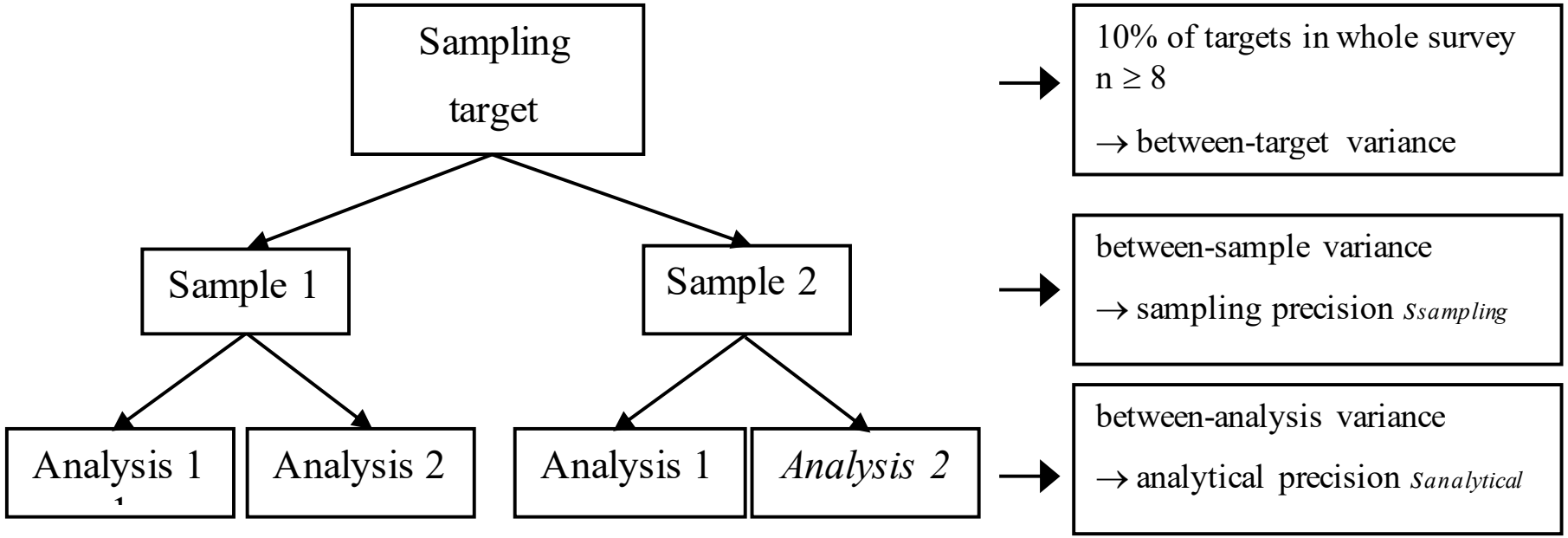
Example A1 from Eurachem UfS Guide:
Nitrate Concentration in Lettuce

- Nitrate a potential risk to human health
- EU threshold 4500 mg kg^{-1} for batch concentration
  - Batch = bay of ~20,000 lettuces = sampling target
- Current sampling protocol specifies taking 10 heads to make a single composite sample from each batch
- What is the uncertainty on measurements?

- Is that amount of U acceptable?
  - can be answered using the Optimised Uncertainty approach
  - not discussed further here, details in UfS Guide Section 16 and..
    - Lyn, J.A., Ramsey, M.H., and Wood, R. (2002) Optimised uncertainty in food analysis: application and comparison between four contrasting 'analyte-commodity' combinations, Analyst, 127, 1252 – 1260.

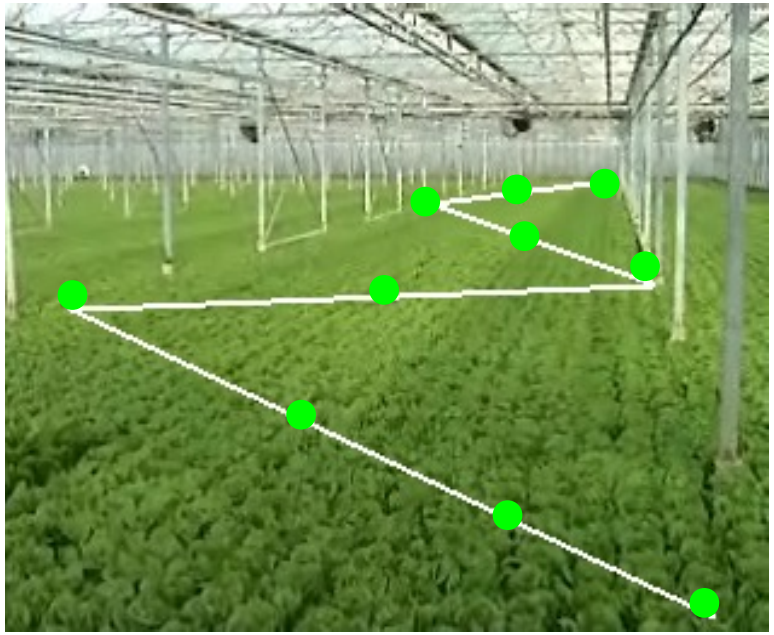
Estimating U with Duplicate Method - using the Balanced Design

Figure 1: A balanced design

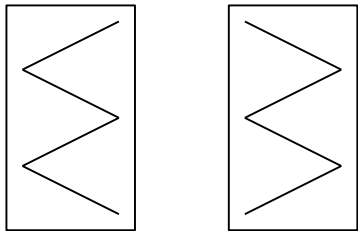


Precision estimated as repeatability

‘W’ Sampling Design for Lettuce



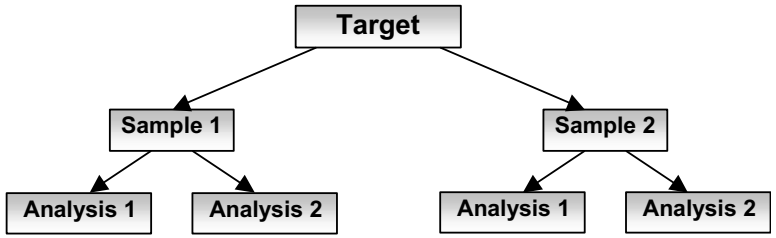
Duplicate sample is equally likely interpretation of ‘W’ design



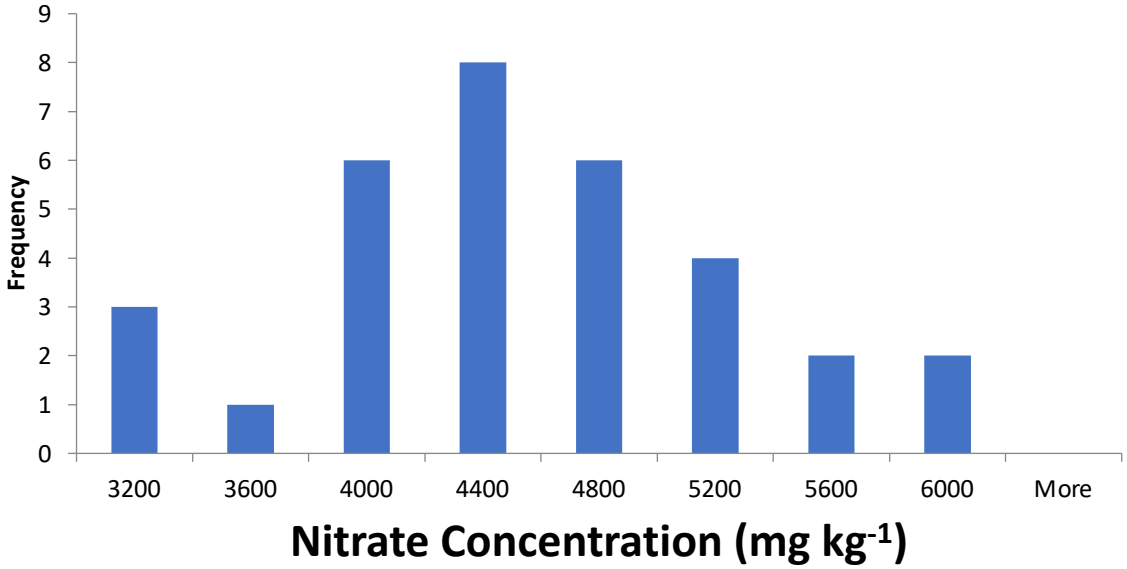
Sampling of Lettuce for Nitrate

University of Sussex

Nitrate conc. in Duplicate Samples



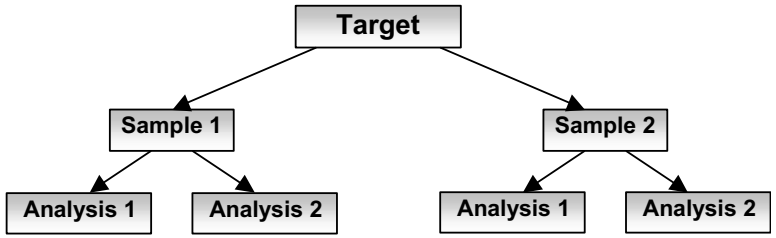
Sample Target	S1A1	S1A2	S2A1	S2A2
A	3898	4139	4466	4693
B	3910	3993	4201	4126
C	5708	5903	4061	3782
D	5028	4754	5450	5416
E	4640	4401	4248	4191
F	5182	5023	4662	4839
G	3028	3224	3023	2901
H	3966	4283	4131	3788



Approximately Normal Distribution
With< 10% outliers
Had to be sure with so few observations

University of Sussex

Nitrate conc. in Duplicate Samples



Sample Target	S1A1	S1A2	S2A1	S2A2
A	3898	4139	4466	4693
B	3910	3993	4201	4126
C	5708	5903	4061	3782
D	5028	4754	5450	5416
E	4640	4401	4248	4191
F	5182	5023	4662	4839
G	3028	3224	3023	2901
H	3966	4283	4131	3788

- Most analytical duplicates agree < x0.1 (approx)
- Sampling duplicates agree only < x0.2 (approx)
- C = Outlying target? S2 outlying analysis?

>4500 Theshold?

- Range of conc. between batches x1.6 (approx)
- Contrast between-target (i.e. geochemical) is evident e.g. F is high and G is low – is level of Uncertainty OK?

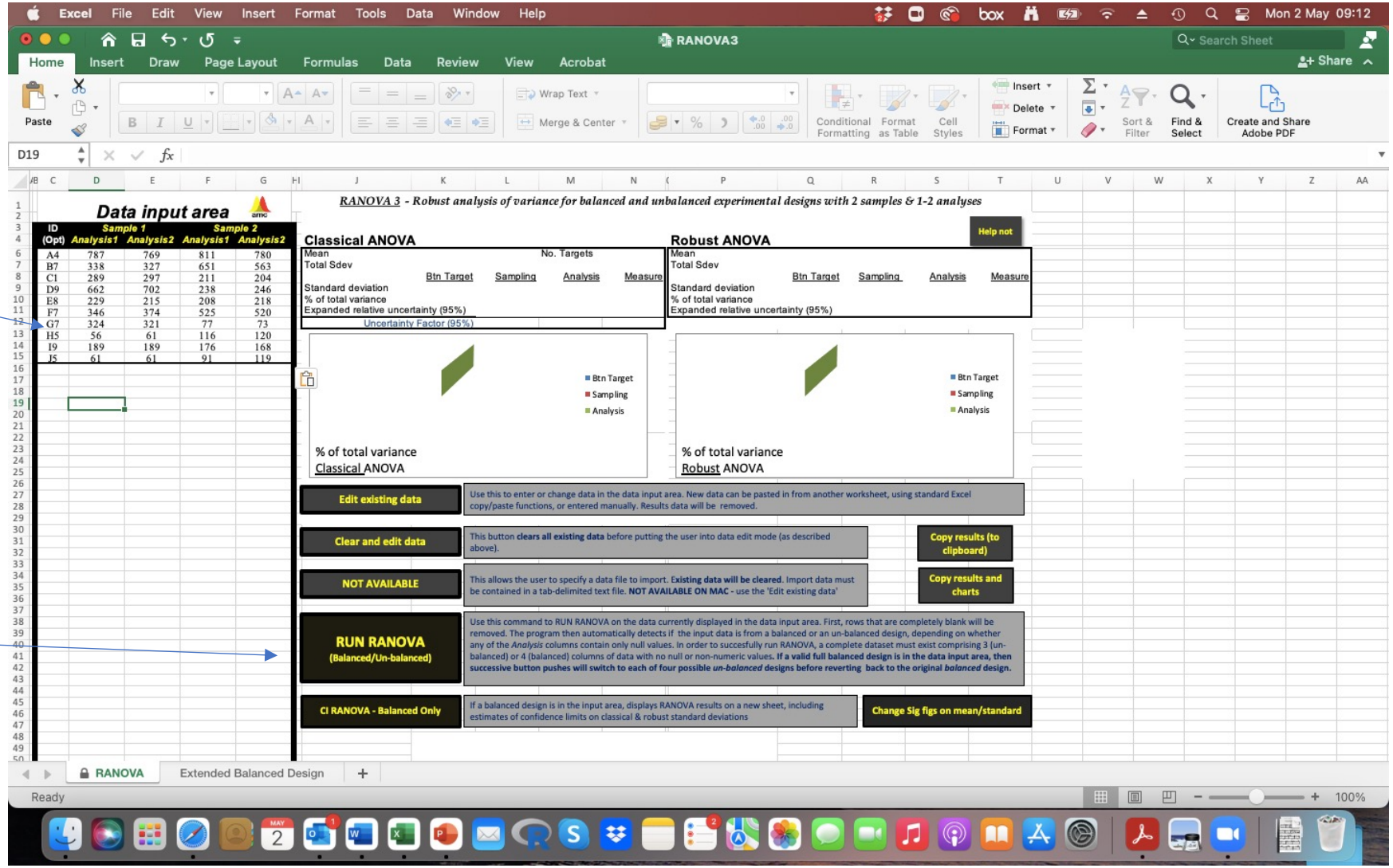
<4500?

University of Sussex

RANOVA3 output for Nitrate in Lettuce Example A1

Cut and paste raw measurement values into RANOVA3 within Excel

Click on RUN RANOVA to run



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

University of Sussex

RANOVA3 output for Nitrate in Lettuce Example A1

Classical ANOVA

Mean	*4345.6*	No. Targets		*8*
Total Sdev	*774.53*			
	Btn Target	Sampling	Analysis	Measure
Standard deviation	*556.28*	*518.16*	*148.18*	*538.93*
% of total variance	*51.58*	*44.76*	*3.66*	*48.42*
Expanded relative uncertainty (95%)		*23.85*	*6.82*	*24.80*
Uncertainty Factor (95%)		*1.2432*	*1.0738*	*1.2574*

Robust ANOVA

Mean	*4408.3*			
Total Sdev	*670.58*			
	Btn Target	Sampling	Analysis	Measure
Standard deviation	*565.4*	*319.05*	*167.94*	*360.55*
% of total variance	*71.09*	*22.64*	*6.27*	*28.91*
Expanded relative uncertainty (95%)		*14.47*	*7.62*	*16.36*

- **Classical** ANOVA
  - *Assumes Normal distribtion*
- *U'* = 24.8% -not reliable
- Uncertainty Factor 1.26 - not relevant as not a log-normal distribution

- Histogram suggests Normal distribution with < 10% of outlying values, (analytical, sampling and between-target), so Robust ANOVA needed

- **Robust** ANOVA
  - *accommodates < 10% outlying values*
- ***U'* = 16.4%** (s_{meas} = 360 mg kg^{-1})

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

University of Sussex

Geochemical interpretation – are lettuces safe to eat?

Sample Target	Nitrate Conc in S1A1 mg/kg	Deterministic Classification C<4500
A	3898	Y
B	3910	Y
C	5708	N
D	5028	N
E	4640	N
F	5182	N
G	3028	Y
H	3966	Y
Batches Accepted		4

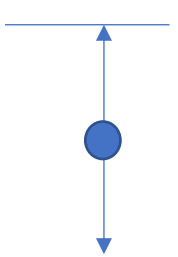
T = 4500 mg kg^{-1}

$C < T$

- **Deterministic classification** of the contamination
  - against the threshold of T = 4500 mg/kg
- shows four batches (A, B, G & H) are below threshold (C < T)

Using MU to improve geochemical interpretation – are lettuces safe to eat?

Sample Target	Nitrate Conc in S1A1 mg/kg	Deterministic Classification C<4500	10-head MU= 16.4%	C+U10	Probabilistic Classification C+U10 < 4500
A	3898	Y	639.3	4537	N
B	3910	Y	641.2	4551	N
C	5708	N	936.1	6644	N
D	5028	N	824.6	5853	N
E	4640	N	761	5401	N
F	5182	N	849.8	6032	N
G	3028	Y	496.6	3525	Y
H	3966	Y	650.4	4616	Y
Batches Accepted		**4**			**2**



T = 4500 mg kg^{-1}

$C + U < T$

- MU can be used to make a **probabilistic classification** of the contamination
  - against the threshold of T = 4500 mg/kg
- Reveals that two batches (A & B) may be false negatives (C + U > T)

University of Sussex

Example A2: Estimation of UfS in Soil - using Duplicate Method

Scenario:

- Former landfill, in West London
- 9 hectare = 90 000 m^2
- Potential housing development
- measurand → Pb conc. in each sampling target

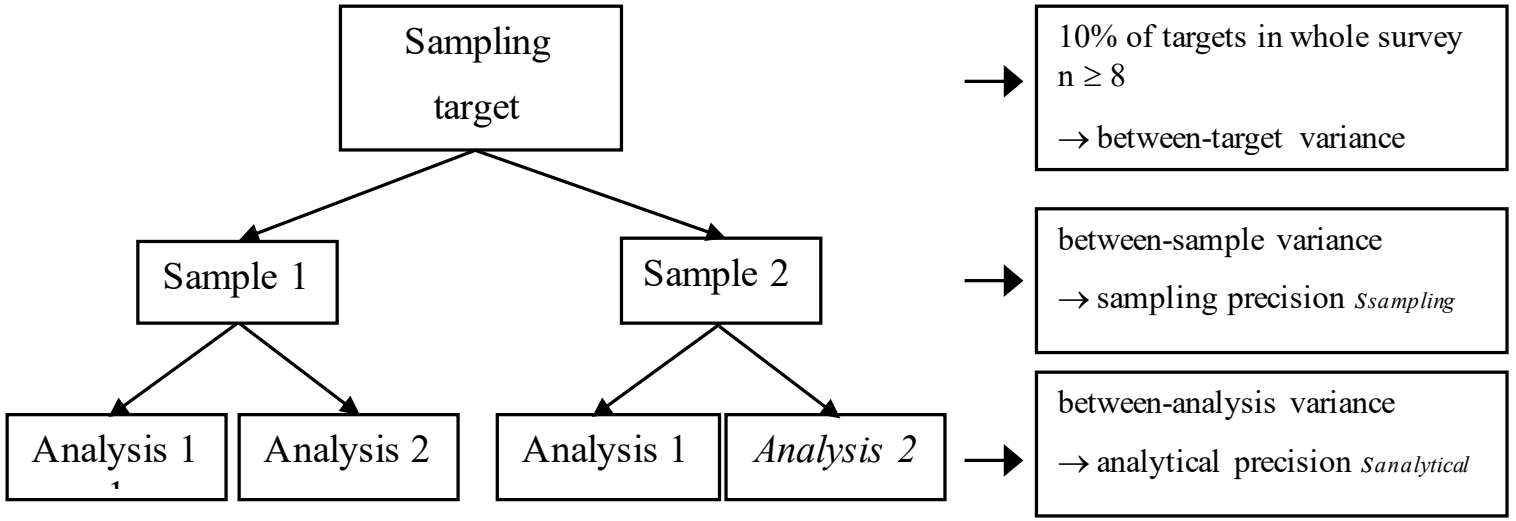
Area of investigation:

- 300 m x 300 m area → depth of 0.15 m
- 100 sampling targets in a regular grid (10 x 10)
- 100 primary samples (taken with soil auger)
  - each intended to represent a 30 m x 30 m target

Example A2 from Eurachem UfS Guide (2019)

Application of Duplicate Method to estimate UfS

Figure 1: A balanced design



- Duplicate samples taken at 10/100 sampling targets (i.e. 10%)
  - randomly selected.
  - Duplicate sampling point 3 m from the original sampling point
    - within the sampling location,
    - in a random direction
    - within the sampling target

University of Sussex

Application of Duplicate method to estimate UfS

- Aims of design of duplicate taking to reflect:-
  - ambiguity in the sampling protocol
    - how differently could it be interpreted by a different samplers?
  - uncertainty in locating sampling location within sampling target
    - e.g. survey error by using tape and compass (or by GPS now)
  - effect of small-scale heterogeneity within each sampling target on measured concentration
    - e.g. at 10% of grid spacing distance, 3m for 30m

Sample prep and analysis in the lab

- Soil samples dried, sieved (<2 mm), ground (<100 μm)
- Test portions of 0.25g digested in nitric/perchloric acid
- Pb concentration measured with ICP-AES, under full AQC
- 6 soil CRMs measured to estimate analytical bias over range of concentration
- corrected for reagent blank concentrations where statistically different to zero
- Raw measurements for use for estimation of uncertainty were:
  - **untruncated** – e.g. 0.0124 mg/kg, <u>not</u> < 0.1 or < detection limit
  - **unrounded** – e.g. 2.64862 mg/kg, <u>not</u> 3 mg/kg

NIST – Wiley online

Results as Spatial Map of Measured Pb concentration

Row	A	B	C	D	E	F	G	H	I	J
1	474	287	250	338	212	458	713	125	77	168
2	378	3590	260	152	197	711	165	69	206	126
3	327	197	240	159	327	264	105	137	131	102
4	787	207	197	87	254	1840	78	102	71	107
5	395	165	188	344	314	302	284	89	87	83
6	453	371	155	462	258	245	237	173	152	83
7	72	470	194	83	162	441	199	326	290	164
8	71	101	108	521	218	327	540	132	258	246
9	72	188	104	463	482	228	135	285	181	146
10	89	366	495	779	60	206	56	135	137	149

Argyraki (1997)

- Measured Pb concentration ranges from 56 to 3590 mg kg^{-1}
- Straddles then UK threshold of > 500 mg Pb kg^{-1} for action required (further risk assessment) – 8% of site
- Gives **Deterministic Map of the contamination (ignores MU) – 92% uncontaminated**

University of Sussex

Spatial Map of Duplicated Sampling Targets

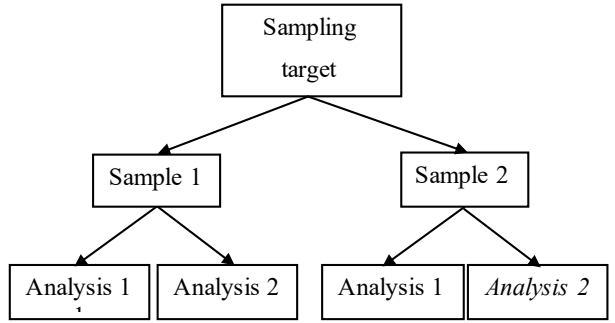
Row	A	B	C	D	E	F	G	H	I	J
1	474	287	250	338	212	458	713	125	77	168
2	378	3590	260	152	197	711	165	69	206	126
3	327	197	240	159	327	264	105	137	131	102
4	787	207	197	87	254	1840	78	102	71	107
5	395	165	188	344	314	302	284	89	87	83
6	453	371	155	462	258	245	237	173	152	83
7	72	470	194	83	162	441	199	326	290	164
8	71	101	108	521	218	327	540	132	258	246
9	72	188	104	463	482	228	135	285	181	146
10	89	366	495	779	60	206	56	135	137	149

Argyraki (1997)

- Uncertainty of measurements estimated by taking of Duplicate Samples at 10% of sampling targets
- - at random selected positions

University of Sussex

Measurements from balanced design for UfS estimation



- Large differences between some sample duplicates (e.g. D9) = high level of UfS
- Good agreement between analytical duplicates (< 10 % difference)

Target #	S1A1	S1A2	S2A1	S2A2
A4	787	769	811	780
B7	338	327	651	563
C1	289	297	211	204
D9	662	702	238	246
E8	229	215	208	218
F7	346	374	525	520
G7	324	321	77	73
H5	56	61	116	120
I9	189	189	176	168
J5	61	61	91	119

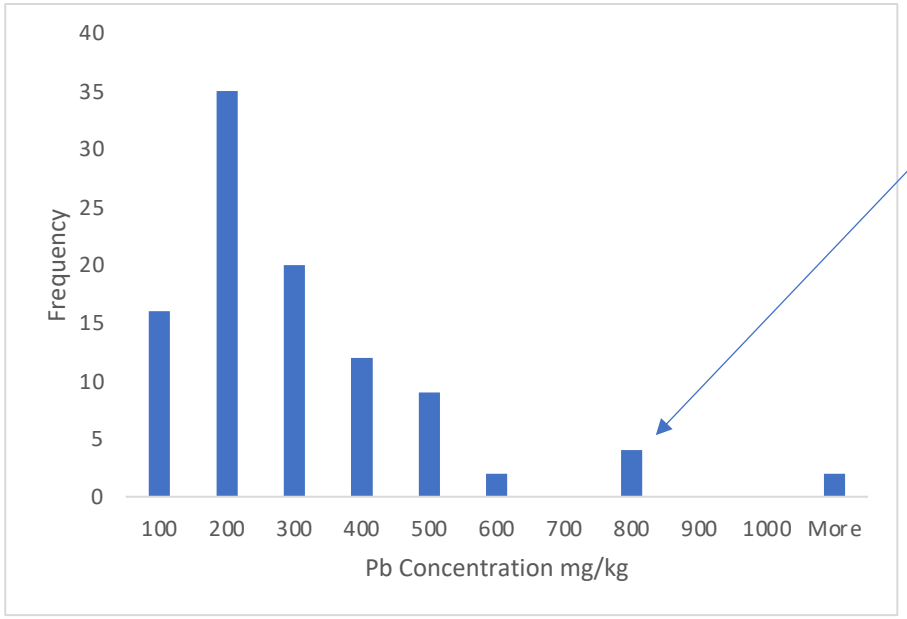
mg kg^{-1}

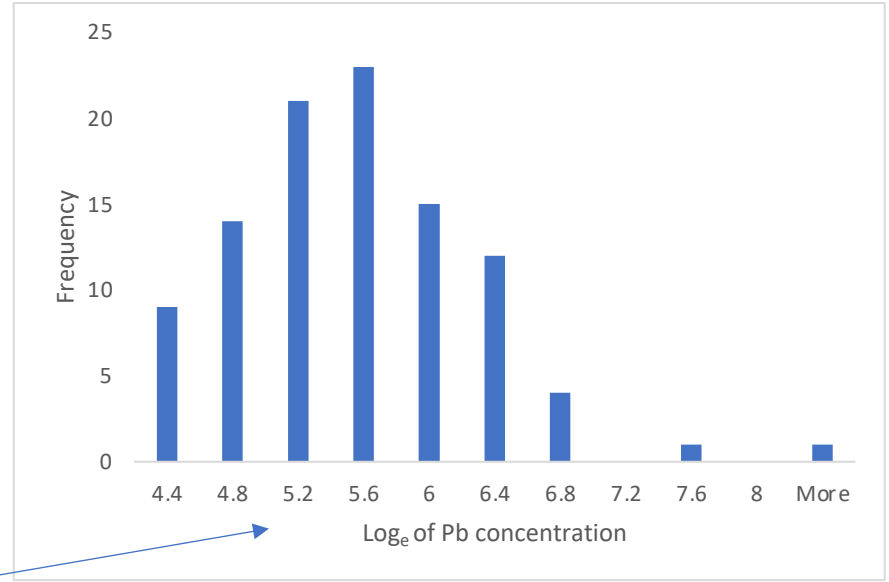
- Needs inspection of frequency distribution to select the best approach to UfS estimation

University of Sussex

Judge Frequency Distribution using Histograms

- Frequency distribution of Pb concentration across the site = long range heterogeneity
- Distribution of Pb measurements on 100 sampling targets is positively skewed = approximately log-normal
- Log-transformation necessary to remove skew





- Distribution closer to Normal after log$_e$ transformation
  - Needed for use of ANOVA

University of Sussex

Need for log-transformation?

- Classical analysis of variance (ANOVA) assumes approximately normal distributions
- Robust ANOVA can accommodate up to 10% outlying values,
  - but not more, and not heavy skew
- However, once transformed, measurement values (and ANOVA results) are no longer given in input units of concentration (e.g. mass fraction, mg kg $^{-1}$)

Measurement values of Pb concentration

In mg kg^{-1}

Target #	S1A1	S1A2	S2A1	S2A2
A4	787	769	811	780
B7	338	327	651	563
C1	289	297	211	204
D9	662	702	238	246
E8	229	215	208	218
F7	346	374	525	520
G7	324	321	77	73
H5	56	61	116	120
I9	189	189	176	168
J5	61	61	91	119

$\log_e$-transformed

Target #	S1A1	S1A2	S2A1	S2A2
A4	6.67	6.65	6.70	6.66
B7	5.82	5.79	6.48	6.33
C1	5.67	5.69	5.35	5.32
D9	6.50	6.55	5.47	5.51
E8	5.43	5.37	5.34	5.38
F7	5.85	5.92	6.26	6.25
G7	5.78	5.77	4.34	4.29
H5	4.03	4.11	4.75	4.79
I9	5.24	5.24	5.17	5.12
J5	4.11	4.11	4.51	4.78

- Need a different way to express MU in this case = Uncertainty factor
  - s_G = standard deviation of the $\log_e$-transformed values (= $s(\log_e(x))$)

$$^{F}U = \exp(2s_G)$$

RANOVA3 output for Soil Example A2

Classical ANOVA

Mean	***317.8***		No. Targets	***10***
Total Sdev	***240.19***			
	Btn Target	Sampling	Analysis	Measure
Standard deviation	***197.55***	***135.43***	***17.99***	***136.62***
% of total variance	***67.65***	***31.79***	***0.56***	***32.35***
Expanded relative uncertainty (95%)		***85.23***	***11.32***	***85.98***
Uncertainty Factor (95%)		***2.6032***	***1.12***	***2.6207***

Robust ANOVA

Mean	***297.31***			
Total Sdev	***218.49***			
	Btn Target	Sampling	Analysis	Measure
Standard deviation	***179.67***	***123.81***	***11.144***	***124.31***
% of total variance	***67.63***	***32.11***	***0.26***	***32.37***
Expanded relative uncertainty (95%)		***83.29***	***7.50***	***83.63***

- Software RANOVA3* (in Excel) performs:-
- Classical ANOVA gives poor estimate of *U'* = 85.98%,
- but also estimate of ${}^{F}U$ as 2.62
  - after $\log_e$-transformation within RANOVA3

Robust U as 83.63% (for comparison)
Histogram suggests > 10% of outlying values,
so direct classical, and robust estimate are not very reliable
So log-transformation before classical ANOVA is likely to be a better option

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

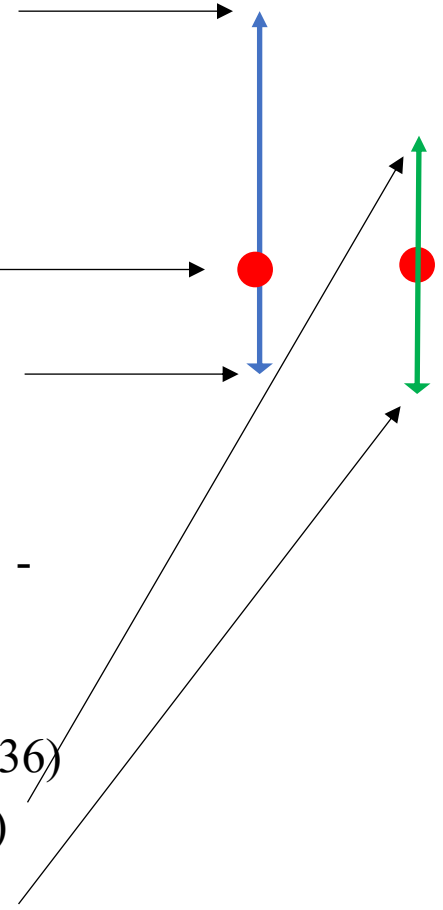
Confidence Limits on Measurement Value

- *For* ^{F}U = 2.62, for a typical Pb measurement value of 300 mg kg^{-1}

Upper confidence limit (UCL) = 784 mg kg^{-1} (300 x 2.62)

Measurement value of 300 mg kg^{-1}

- Lower confidence limit (LCL) = 115 mg kg^{-1} (300 / 2.62)
- Asymmetric confidence limits around the measured value
- -185 and +484 mg kg^{-1} (away from 300)
- Reflects skew in frequency distribution of the uncertainty as seen in histograms



-

- Not seen in symmetrical confidence limits from robust U' = 83.6% = 251 (300 * 0.836)
- = +/- 251 mg kg^{-1}

UCL = 551 (300 + 251)

LCL = 49 (300 - 251)

-
  - calculated without log-transformation.

Inclusion of analytical bias in ${}^{F}U_{meas}$ estimate

- Analytical bias - modelled as Linear functional relationship fitted between measured values on certified values of 6 CRMs (using FREML*)

  - 3.41 % ± 1.34 %

- Systematic component of relative expanded uncertainty:

$$u_{systematic}' = \sqrt{-3.41^2 + 1.34^2}\ \% = 3.72\ \%$$

$$s'_{systematic} = 0.0372\ \text{mg kg}^{-1}$$

NIST – Wiley online

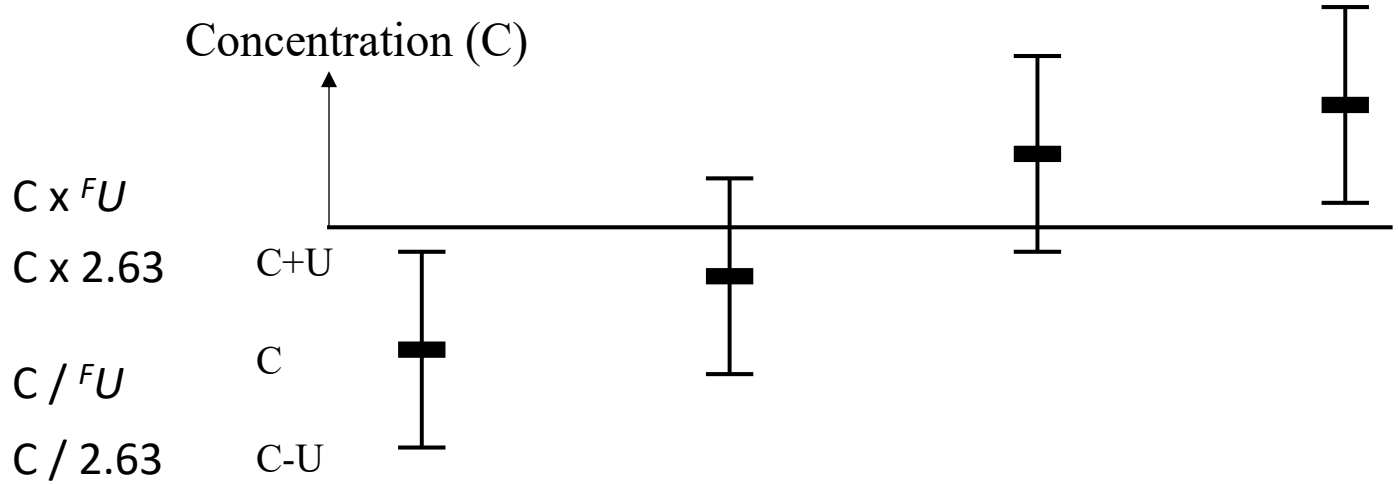
- Currently no consensus on how to combine systematic and random components of uncertainty.
- One method is to add them by the sum of their squares (extending previous equation):
- $${}^{F}u_{\text{meas}} = exp\sqrt{s^2_{\text{G,samp}} + \left(s'_{\text{anal}}\right)^2 + (s'_{systematic})^2}$$
- $$\exp\sqrt{0.4784^2 + 0.0566^2 + 0.0372^2} = 1.621 \quad \text{(up from 1.619)}$$

- ${}^{F}U = ({}^{F}u)^2 = 2.628$ **=2.63** (up from 2.621) – **Analytical bias has almost no effect on MU in this case**

*Functional Relationship Estimation by Maximum Likelihood, AMC Technical Brief Number 10 (2002), software from:- https://www.rsc.org/Membership/Networking/InterestGroups/Analytical/AMC/Software/

University of Sussex

Effect of MU on geochemical interpretation



500 – then UK limit for risk assessment

Uncontaminated | Possibly Contaminated | Probably Contaminated | Contaminated - Probabilistic Classification

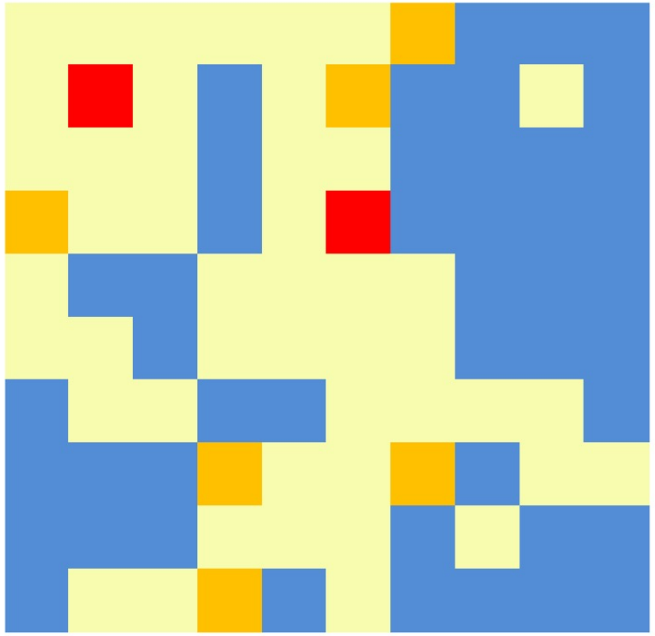
Case Study	<190	190-500	500-1318	>1318 mg/kg
Map Symbol				

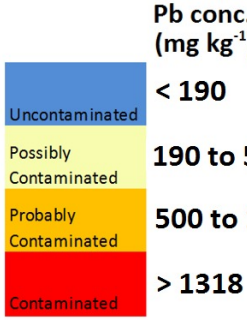
Uncertainty Factor ^{F}U = 2.63, so 190 = 500 / 2.63 1318 = 500 x 2.63

Probabilistic Geochemical Mapping using MU

Example for Pb at Hounslow Site

Row	A	B	C	D	E	F	G	H	I	J
1	474	287	250	338	212	458	713	125	77	168
2	378	3590	260	152	197	711	165	69	206	126
3	327	197	240	159	327	264	105	137	131	102
4	787	207	197	87	254	1840	78	102	71	107
5	395	165	188	344	314	302	284	89	87	83
6	453	371	155	462	258	245	237	173	152	83
7	72	470	194	83	162	441	199	326	290	164
8	71	101	108	521	218	327	540	132	258	246
9	72	188	104	463	482	228	135	285	181	146
10	89	366	495	779	60	206	56	135	137	149





Uncontaminated
Possibly Contaminated
Probably Contaminated
Contaminated

Deterministic Map
- *Ignores MU*
- **92% 'uncontaminated'**

Probabilistic Map
- *Allows for MU*
- **46% 'uncontaminated'**

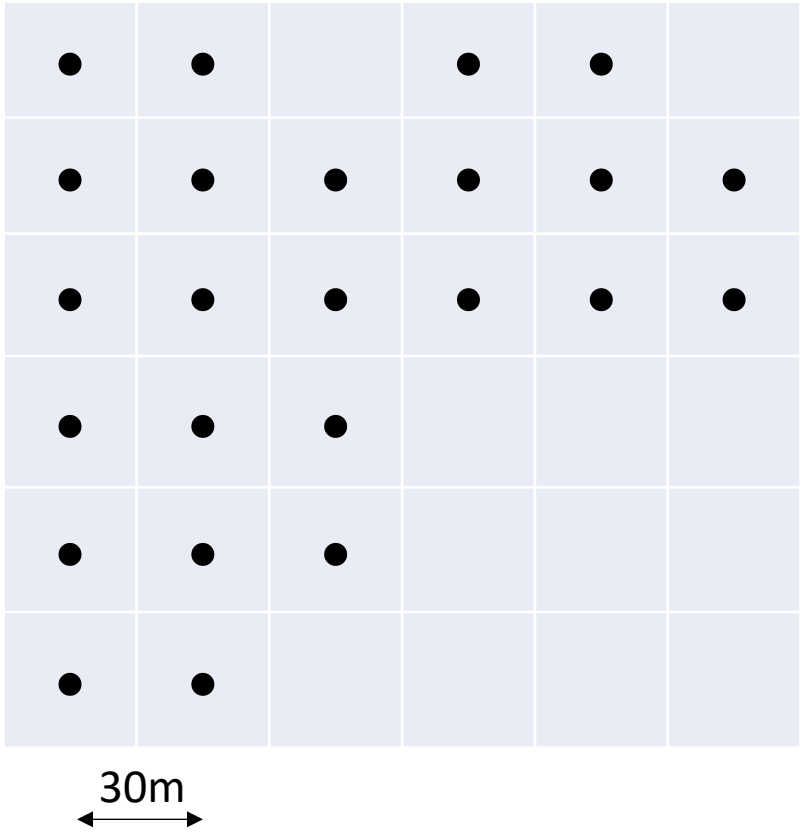
Bettencourt da Silva, R., Argyraki, A., Borges, C., Ramsey, M.H. (2022) Spatial modelling of concentration in topsoil using random and systematic uncertainty components. Analytical Letters 210574656 https://www.tandfonline.com/doi/full/10.1080/00032719.2022.2050383

University of Sussex

Case Study 3: Estimation of UfS & MU
for measurements made *in situ*

Site of a medieval Pb smelter at Wirksworth, Derbyshire, UK

Hand-held portable x-ray fluorescence spectrometer (PXRF) used to measure Pb concentration [Pb] in topsoil *in situ*



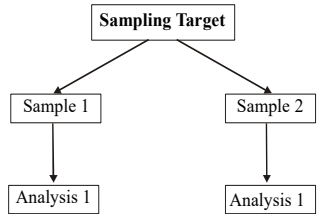
Grid of 24 sampling targets used to survey [Pb] across site

Details in Ramsey M.H. (2020) Measurement Uncertainty from Sampling: Implication for Testing, Diagnostics and Inspection. Presentation to 17th IMEKO TC 10 and EUROLAB Virtual Conference *"Global Trends in Testing, Diagnostics & Inspection for 2030"* October 20-22, 2020. https://www.imeko.org/publications/tc10-2020/IMEKO-TC10-2020-042.pdf

University of Sussex

Case Study: Estimation of UfS & MU for measurements made *in situ* (Method)

- Duplicate Method used to estimate random components of MU of *in situ* measurements
  - as repeatability
- Equivalent of 'duplicate samples' are taken by placing the *in situ* measurement device twice, reflecting two independent interpretations of measurement protocol.
- In this study PXRF duplicates were 2m apart, in a randomly chosen direction, to reflect uncertainty in location



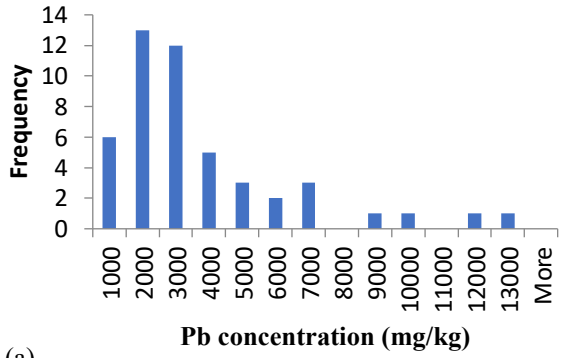
- These two sampling points are both equally likely interpretations of the protocol given that particular surveying technology
- Simplified design used for speed (no analytical duplicates)

University of Sussex

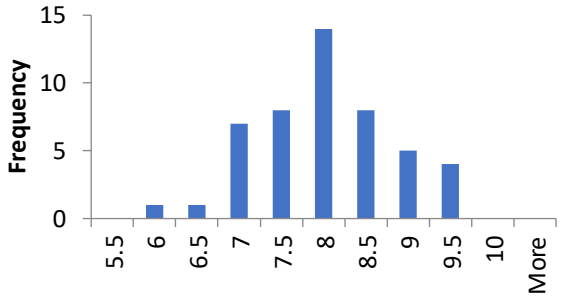
Duplicated PXRF measurements – for random component of UfS (Results)

Target Number	S1Pb mg/kg	S2Pb mg/kg
1	1005	1633
2	4631	3723
3	1415	2264
4	865	1350
5	2899	2216
6	721	1758
7	2122	1014
8	1321	1043
9	3348	3904
10	11543	5570
11	2904	2833
12	2617	2762
13	976	786
14	6127	3874
15	331	576
16	12878	8948
17	3246	4332
18	9006	6098
19	1936	1989
20	5811	6289
21	4611	2880
22	1326	1442
23	1215	2713
24	2070	2305

- Duplicated 'samples' show quite large variation (from small scale heterogeneity)
- Again, distribution is log-normal (made normal by log-transformation)



(a)



(b)

- So use Classical ANOVA in RANOVA3, gave uncertainty factor ${}^{F}U = 1.85$
- External estimate of PXRF alone $U'_{analysis} = 3\%$.
  - Made using additional *ex situ* PXRF measurements (made in lab on prepared versions of removed samples from same 24 targets), in fully balanced experimental design (i.e. with duplicated analyses)
- Similar to value reported by PXRF instrument
- Estimate of $U'_{analysis}$ has little effect on value of MU
- **Actual MU** ${}^{F}U = 1.85$ **is much higher** than when UfS included (3%)

University of Sussex

Estimation of UfS (and MU) for measurements made *in situ* (Method)

- Systematic component of MU of *in situ* measurements from analytical bias
  - estimated by measurements made on matrix-matched CRMs (e.g. NIST 2710), but….
  - CRMs are homogeneous, fine grained, and dry
  - unlike most test materials in real world (field soils heterogeneous, coarse grained and wet)
- To overcome this mis-match, compare *in situ* against *ex situ* measurements
  - made for same analyte on same sampling targets.
- Need to also match value of the 'measurand', which is effectively the true value that is being estimated
  - i.e. total Pb concentration in dry soil
- Therefore, in Case Study, also removed *ex situ* samples taken at same locations as where *in situ* measurement made
  - with full balanced design
  - for all 24 sampling targets, but 8 targets would be OK for routine investigation
  - then analysed by ICP-AES (traceable to CRMs)
  - after drying, sieving, grinding and acid digestion in a remote laboratory (i.e. *ex situ*).

‘Bias’ of *in situ* PXRF against *ex situ* ICP-AES measurements

PXRF-in situ	ICP-ex situ
Target Av. [Pb] mg/kg	Target Av. [Pb] mg/kg
1319	7340
4177	8815
1840	1522
1108	1290
2558	9340
1240	3080
1568	4180
1182	1926
3626	3670
8557	6718
2869	5630
2690	3630
881	6880
5001	9370
454	1522
10913	21877
3789	5230
7552	18784
1963	2800
6050	10584
3746	7316
1384	2235
1964	3860
2188	5210

- Systematic component of MU estimated as bias…
  - by comparing average value of both *in situ* PXRF measurements
  - against *ex situ* ICP-AES measurement
- Relationship modelled as a function of concentration using FREML
  - functional relationship estimation by maximum likelihood
- In FREML uncertainty of both variables properly taken into account.
  - Also possible to use ordinary least-squares regression, but this can only allow for uncertainty in y-axis (e.g. PXRF) and ignores uncertainty for x-axis (e.g. ICP-AES)

$$\text{Model} \rightarrow [Pb]_{in\ situ} = b(1) \times [Pb]_{ex\ situ} + b(0)$$

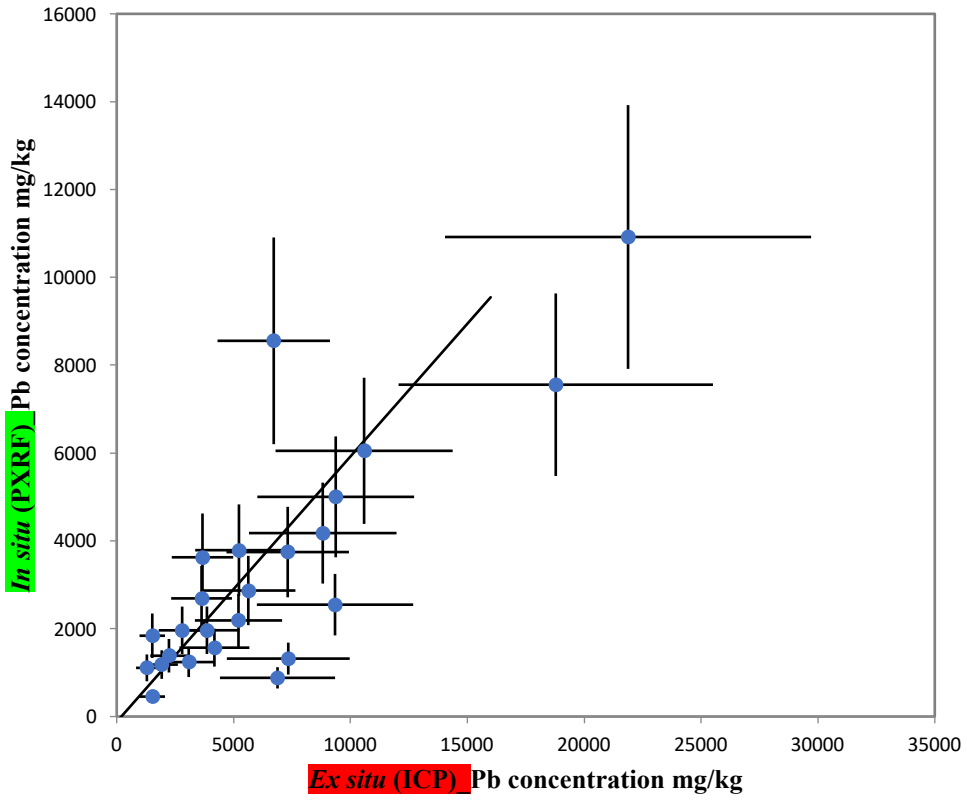
- Slope coefficient of linear model (b(1)) → rotational component of bias
- Intercept coefficient b(0) → translational component

University of Sussex

‘Bias’ of *in situ* PXRF against *ex situ* ICP-AES measurements (2)

Equation describing relationship,
showing both coefficients and their standard errors (in parentheses):

$[Pb]_{in\ situ} = 0.60\ (\pm 0.09) \times [Pb]_{ex\ situ} - 120\ (\pm 288)$



- Estimated rotational bias of *in situ* PXRF measurements
  - compared against the *ex situ* ICP measurements
- calculated from slope coefficient, is **- 40%** (± 9%)
  - i.e. 100 x (1 - 0.60).
- No translational bias detected, as..
  - intercept coefficient = -120 mk/kg (± 288)
  - not statistically different from zero
- Possible causes of measurement bias identified as:
  - soil moisture
  - material/particles > 2mm diameter
  - surface roughness in the PXRF ‘undisturbed sample’
  - depth difference between undisturbed sample for *in situ* PXRF (~1mm)
  - but removed *ex situ* field sample for ICP-AES (150 mm)

University of Sussex

Treatment of Systematic component of MU for *in situ* measurements

- Issue needs further discussion by users of *in situ* measurements in general, to reach a consensus, as identified [12], in brief…

Option 1 - 'correct' *in situ* measurements ($[Pb]_{PXRF,\ corr}$) to agree with *ex situ* values by applying a rearrangement of the bias model

  - omitting the non-significant intercept for the Case Study

$$[Pb]_{PXRF,corr} = \frac{[Pb]_{PXRF,raw} - b(0)}{b(1)} = \frac{[Pb]_{PXRF,raw}}{0.60}$$

- Uncertainty of this correction ($s'_{bias} = 0.09$, as <0.2) can be combined into $s_{G,meas}$ using an approximation [9]

$$s_{G,meas} = \sqrt{s_{G,meas}^2 + \left(s'_{bias}\right)^2}$$

- Expanded uncertainty factor ${}^{F}U = 1.88$ (up from 1.85)

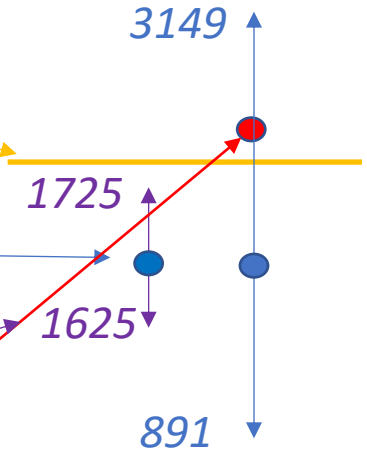
$${}^{F}U = \exp\left(2s_{G,meas}\right)$$

Option 2 is not to correct, but to add the entire bias, and its uncertainty, to MU

[12] Ramsey M.H. (2020) Challenges for the estimation of uncertainty of measurements made in situ. Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement. DOI: 10.1007/s00769-020-01446-4

Benefits of MU - Regulatory compliance – Case Study

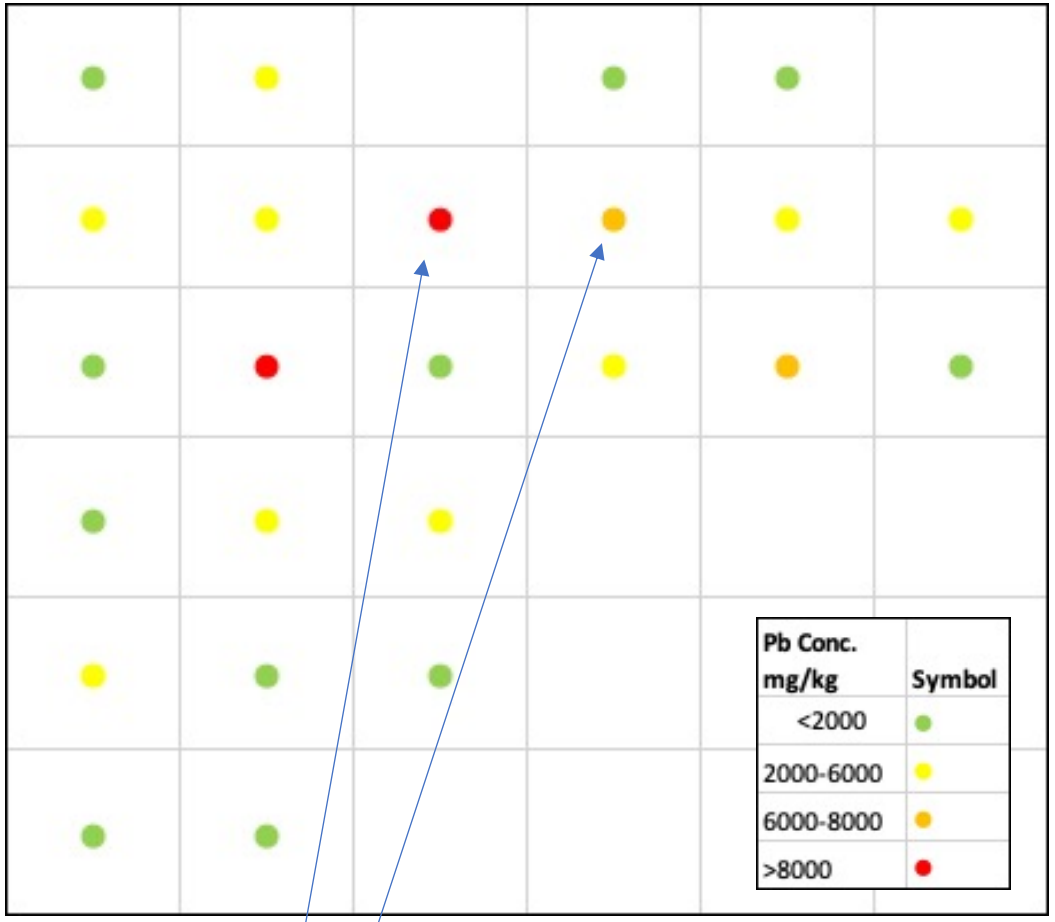
- In case study, one UK regulatory threshold for Pb in soil was 2000 mg/kg
- First PXRF measurement value on Target 1 is 1005 mg/kg
- 'Correction of bias' using Equation (6) (Option 1) gives 1675 mg/kg



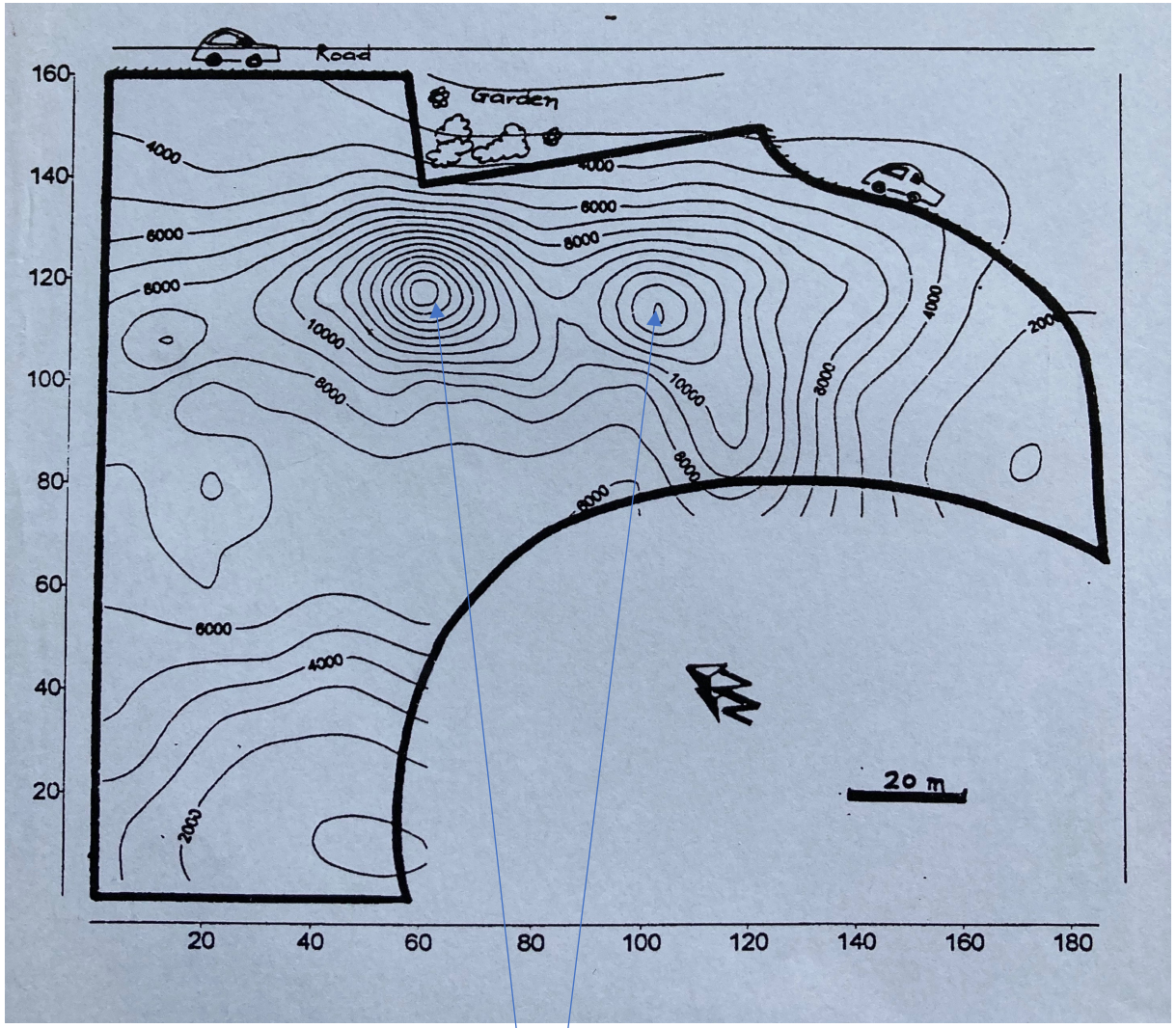
- If MU based on $U'_{analysis}$ of 3%, gives true value between 1625 and 1725 mg/kg (i.e. under threshold of 2000) - excludes a false positive classification
- MU (including UfS as $^{F}U = 1.88$, bias corrected) gives true value between 891 to 3149 mg/kg,
  - indicates possibility true value of [Pb] over threshold of 2000 mg/kg
- Evidence that ignoring UfS can cause financial loss, e.g. :-
  - False negatives can cause litigation
  - False positive can cause unnecessary remediation

University of Sussex

Case Study: Estimation of UfS & MU for measurements made *in situ*



Map of 24 Pb concentrations (mg/kg) measured by PXRF showing two 'hot spots' of high [Pb]

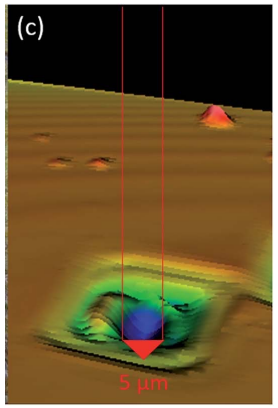


Map of modelled Pb concentration (mg/kg) based on data from several surveys, showing two clear peaks due to Pb smelters

Argyraki A (1997) Estimation of measurement uncertainty in the sampling of contaminated land. PhD Thesis, Imperial College, University of London

Conclusions

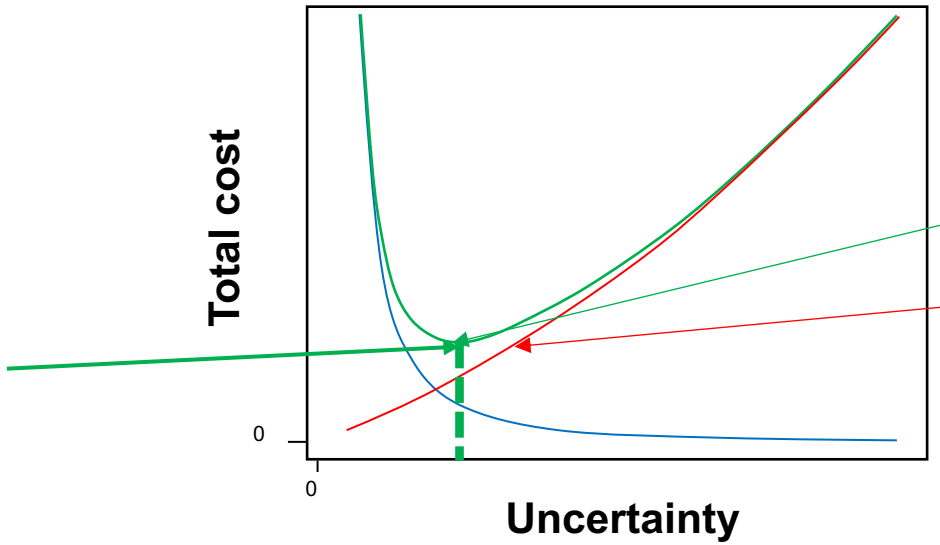
- Knowing uncertainty of measurements (MU) is crucial for their reliable geochemical interpretation
- Estimating MU, including the contribution from sampling, can be done for *ex situ* (i.e. lab) measurements with the Duplicate Method
  - Take ~10% of your field samples in duplicate
  - Use ANOVA (e.g.RANOVA3) to calculate MU and its components (e.g. UfS)
  - Also analyse reference materials (CRMs) to estimate analytical bias (add into MU if significant)
  - Applicable to any sampling medium: soil, sediment, herbage, waters, gases etc.
  - Reviewed by Argyraki A (2019) Applications of UfS estimation across a range of sectors. Presentation at Eurachem/Eurolab Workshop Uncertainty from sampling and analysis for accredited laboratories, BAM, Berlin November 2019 https://www.eurachem.org/images/stories/workshops/2019_11_MU/pdf/P1-08_Application_review_Argyraki.pdf
- Can also be applied to *in situ* measurements, such as using PXRF
    - At any scale; macro or micro-scale (e.g. SIMS on mineral grains)
      - Ramsey and Wiedenbeck (2017) Geostandards and Geoanalytical Research, 42,1,5-24
- Use MU values to improve your geochemical interpretation
  - E.g. are concentration levels different from those at (1) another site, (2) regulatory limits?
  - Include MU in probabilistic risk assessment – e.g. in geochemical mapping



University of Sussex

Optimal level of MU - at minimum overall cost

- Can be used to judge Fitness for Purpose (FFP) of the measurements
- Optimal level of MU (= Target MU) – can be set at…
- At MU that minimises the overall cost (including the consequences of incorrect decisions)



Minimum cost where measurements are FFP

Sum of both costs

Cost of incorrect decisions

e.g. Lettuce £5280 (12,000 heads at £0.44)

Cost of measurement

e.g. Lettuce £20 per sample, £20 per analysis - for false negative decision

- By knowing UfS, can judge how Target MU 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)

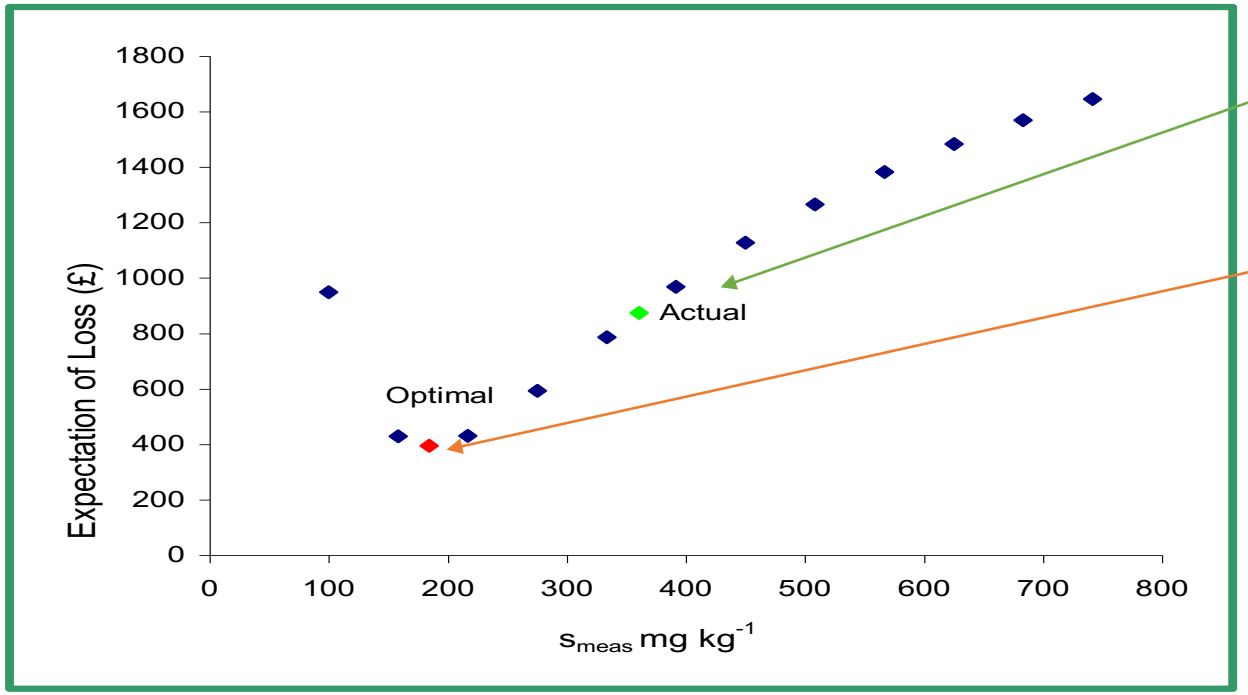
- Lyn, J.A., Ramsey, M.H., and Wood, R. (2002) Optimised uncertainty in food analysis: application and comparison between four contrasting 'analyte-commodity' combinations, Analyst, 127, 1252 – 1260.

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)

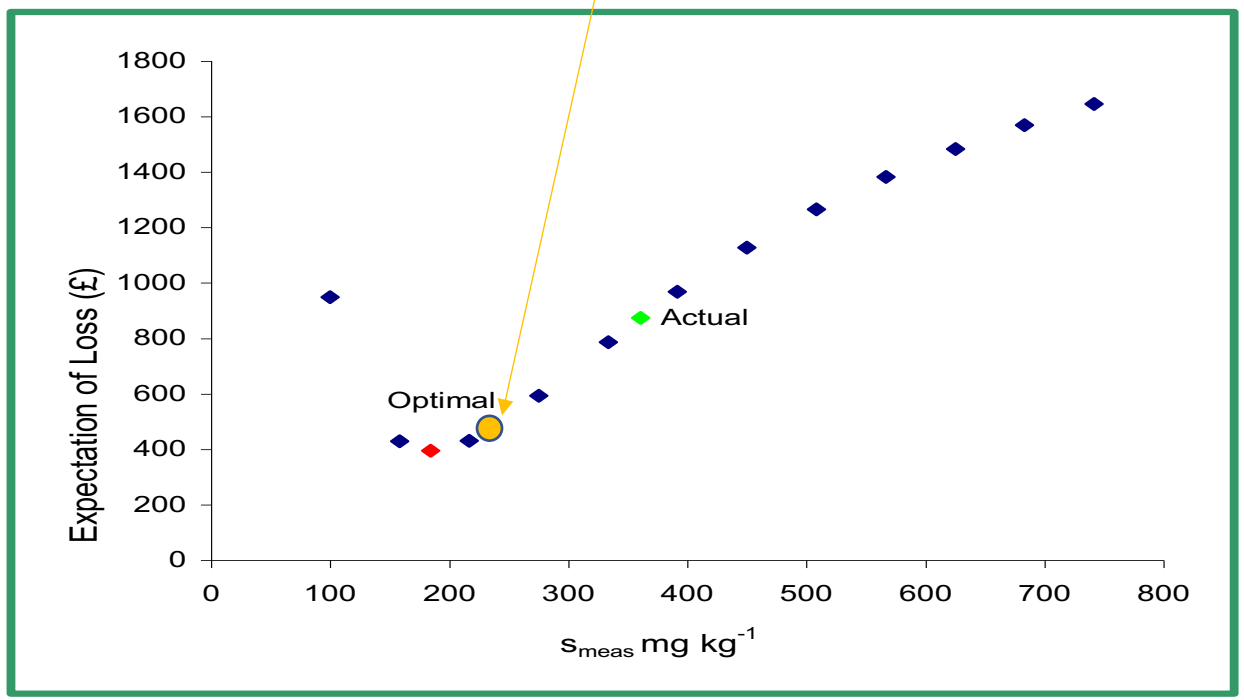
So take composite sample with 40 heads instead of 10 heads

University of Sussex

Reducing the Uncertainty – by taking more increments

- 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