The way forward for Uncertainty from Sampling

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

Overview

• New applications of existing UfS estimation methods
  – In situ – at mm scale (PXRF), and µm scale (SIMS) – sensors in general
    • Passive – Y-ray Spec
  – On site – briefly mentioned here
• Need to further develop methods for estimation of UfS
• Gaining benefits from knowing UfS
  – E.g. improving sampling to reduce UfS to achieve FFP
• External factors affecting take up of UfS estimation
  – Management of the whole measurement process
• Conclusions
UfS estimation for a wider range of measurement types

**In situ measurements in general**

- Taken *in situ* without disturbing or removing the test material
  - Sampling indivisible part of measurement process, e.g.
    - Soils - handheld portable (P)XRF for some metals
    - Minerals – SIMS at micron scale
    - Gases - many sensors – how most measurements are made!
      - e.g. Photolisation detectors for VOCs, (e.g. benzene), IR for CH₄, CO₂
    - Liquids - e.g. pH, UV-Vis for NO₃, TOC, H₂S
    - Clinical - Transcutaneous Bilirubinometer (TcB) – jaundice?

- Less expensive than traditional *ex situ*, so
  - more measurements can be taken
  - giving better coverage of target in space and/or time
    - Even 100% coverage – e.g. groundhog (γ-ray spec)

- Measurements often have larger uncertainty
  - due partially to heterogeneity of analyte concentration (not mixed)
  - detection limits often not as low as for *ex situ* measurements
  - U can be estimated by duplicate method (or SPT)
  - can be shown fit-for-purpose if UfS quantified (example follows)

---

**UfS estimation for *in situ* measurements**

- ‘Sample duplicate’ = duplicated positioning of probe, using same sampling protocol

- ‘Analytical duplicate’ from duplicated measurements made without moving probe

- Systematic component of uncertainty (from bias)
  - can’t be estimated only with matrix-matched CRMs, as:

<table>
<thead>
<tr>
<th>CRM</th>
<th>Dried</th>
<th>Ground</th>
<th>Homogenized</th>
<th>Compacted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test material</td>
<td>Moist</td>
<td>Unground</td>
<td>Heterogeneous</td>
<td>Un-consolidated</td>
</tr>
</tbody>
</table>

- Comparison required, between measurements made *in situ* and those made *ex situ*,
  - ideally with independent analytical method for same measurand
  - on samples taken from same sampling target

- Explain with Example
UfS estimation for *in situ* measurements (at 5mm scale)

- Nature reserve - **West London**
- Ex-Firing range
- Measurements of Pb concentration in topsoil.
- Compare *in situ* PXRF against traditional field sampling
  - with *ex situ* lab (AAS) measurements
- 100 x100m site – 36 sampling targets
- UfS estimated with duplicate method
  - balanced design
  - at 9 sampling targets
- Bias against 6 CRMs, -11% for PXRF, -1% for AAS


‘Bias’ between *in situ* and *ex situ* = systematic component of UfS – using FREML

- Model relationship using FREML* (n =35)
  - Allows for uncertainty on BOTH axes
- Model: In situ Pb = 0.43 (± 0.08) x Ex situ Pb + 77 (± 26)
- ‘Bias’ = -57%
  - caused by soil moisture, material >2mm, surface roughness, and depth difference
- Debate about whether to (1) ‘correct’ *in situ* measurements (or *vice versa*)
  - Perhaps *in situ*Pb is closer to true value being experienced by living organisms*
  - Depends on definition of measurand
- Or (2) include ‘bias’ in estimate of U
- Needs further research

*Functional Relationship Estimation by Maximum Likelihood, AMC Technical Brief Number 10 (2002), software from: https://www.rsc.org/Membership/Networking/InterestGroups/Analytical/AMC/Software/
UfS estimation of \textit{in situ} measurements & FFP

- Random component of UfS calculated using RANOVA of duplicate measurements -

<table>
<thead>
<tr>
<th>parameter</th>
<th>\textit{Ex situ}</th>
<th>\textit{In situ}</th>
</tr>
</thead>
<tbody>
<tr>
<td>([\text{Pb}]\text{ mg kg}^{-1})</td>
<td>\begin{tabular}{c}749 \end{tabular} &amp; \begin{tabular}{c}1045 \end{tabular}</td>
<td></td>
</tr>
<tr>
<td>(s_{\text{analytical}})</td>
<td>\begin{tabular}{c}14 \end{tabular} &amp; \begin{tabular}{c}61 \end{tabular}</td>
<td></td>
</tr>
<tr>
<td>(s_{\text{sampling}})</td>
<td>\begin{tabular}{c}310 \end{tabular} &amp; \begin{tabular}{c}529 \end{tabular}</td>
<td></td>
</tr>
<tr>
<td>(s_{\text{meas}})</td>
<td>\begin{tabular}{c}311 \end{tabular} &amp; \begin{tabular}{c}532 \end{tabular}</td>
<td></td>
</tr>
<tr>
<td>(U')</td>
<td>\begin{tabular}{c}83% \end{tabular} &amp; \begin{tabular}{c}102% \end{tabular}</td>
<td></td>
</tr>
</tbody>
</table>

- In \textit{In situ} Analysis gives higher \(U\) – but not dominant source
- Sampling is dominant sources of \(U\) (>99\% in both cases)

- One Benefit of knowing UfS is the ability to judge fitness-for-purpose, described in:-
  - UfS Guide Section 16 of, applied here using Optimized Uncertainty method = OCLI
  - i.e. was the sampling (and analysis) good enough?

Judging FFP using Optimized Uncertainty (OCLI) equation

\[ E (L) = C \left[ 1 - \Phi \left( \frac{\epsilon_1}{s_{\text{meas}}} \right) \right] + \frac{D}{s_{\text{meas}}^2} \]

- \(E (L)\) – expectation of financial loss
- \(s_{\text{meas}}\) – measurement uncertainty
- \(\Phi\) – standard normal cumulative distribution function
- \(\epsilon_1\) – error limit = \(|T - c|\)
  - \((T = \text{threshold value, } c = \text{contaminant concentration})\)
- \(D\) – combined optimal cost for sampling and analysis
- \(C\) – consequence costs (e.g. potential losses resulting from misclassification)

<table>
<thead>
<tr>
<th></th>
<th>Cost per measurement (£)</th>
<th>Consequence cost of misclassification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(L_{\text{samp}})</td>
<td>(L_{\text{anal}})</td>
</tr>
<tr>
<td>Ex situ (AAS)</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>In situ (PXRF)</td>
<td>13</td>
<td>10500</td>
</tr>
</tbody>
</table>
Judging FFP – at Acceptable level of Uncertainty?

Cost of lowering $U$ on measurement vs. Uncertainty $mg/kg$ ($s_{meas}$)

- Actual $U$
- Optimal $U$

Cost of misclassification, e.g. unnecessary remediation

General Case

Is in situ or ex situ more cost effective?

For Case Study

Neither in situ or ex situ procedure is currently Fitness-for-Purpose (FFP)

- Both in situ (●) and ex situ (▲) have far larger $U_{\text{meas}}$ ($x5$) than is optimal to achieve FFP
- Reduction of $U_{F}$ needs use of composite measurements within each target
- 4-fold Composite measurement should lower $U_{F} \times 2$, & reduce cost by $\sim x2$
- Using model ($s^2 \sim 1/m$)

- In situ has generally lower overall cost ($x3$) without improvement

Ex situ optimal uncertainty: 68 ug g$^{-1}$

In situ optimal uncertainty: 88 ug g$^{-1}$

Ex situ actual uncertainty: 311 ug g$^{-1}$

In situ actual uncertainty: 532 ug g$^{-1}$

£11,000:

£3,700
**Duplicate Method at micron scale on *in situ***

Estimation of UfS and $U_{\text{heterogeneity}}$ for SIMS measurements

Fragments of NBS 28 Glass - Mean diameter ~ 230μm
Mass estimated as ~ 20μg
Select 100 sample fragments
Use Duplicate method to estimate U

Duplicate measurements - 50 μm apart
- made on each of 100 fragments
- measure at different times in run
- run over 15 hours

Sample mass from crater - ~ 300-350 pg

$^{18}$O/$^{16}$O in NBS 28 Glass CRM
Helmholtz Zentrum, Potsdam, Germany

SIMS = Secondary Ion Mass Spectrometry
Ramsey MH and Wiedenbeck M. (2017) Geostandards and Geoanalytical Research, 42,1,5-24

**Uncertainty estimates for NBS28 at micron scale**

- Use ANOVA to estimate Uncertainty
- expressed U in units of ‘per mil’ ‰ = $1000 \times \frac{s}{x}$
  - Unit widely used in isotopic analysis for repeatability and heterogeneity,

Summary of Uncertainty estimates for NBS28

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Between-dups</th>
<th>Within-dups</th>
<th>Num pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncertainty 1s(‰)*</td>
<td>0.31</td>
<td>0.28</td>
<td>0.14</td>
<td>97</td>
</tr>
</tbody>
</table>

- $U_{\text{anal}} =$ Analytical repeatability estimated from ‘within-duplicate’ = 0.14‰
- $U_{\text{samp}} =$ between-fragments from Heterogeneity ($U_{\text{heter}}$) quantified as 0.28 ‰
  - adds to $U_{\text{anal}}$ to give total repeatability/ $U_{\text{meas}}$ of 0.31‰
  - dominates total measurement variance ($U_{\text{samp}}$ contributes 81%)

- Full measurement uncertainty estimate would require bias against matched CRMs & between-lab variance

Ramsey MH and Wiedenbeck M. (2017) Geostandards and Geoanalytical Research, 42,1,5-24
UfS of passive **in situ** measurements of radioactive decay

**Case Study:** $^{137}$Cs in soil at Dounreay

- Area within nuclear decommissioning site
  - $^{137}$Cs measured with $\gamma$-ray spectrometry (both **in situ** and **ex situ**)
- **Objectives** - *Comparison of in situ against ex situ surveys*

---

**In situ** measurements by $\gamma$-ray spectrometry

- **In situ** NaI detector responds to $\gamma$-rays (e.g. from decay of $^{137}$Cs)
- Mass of soil ‘sample’ 200 - **1000 kg**
- ~**1000 larger mass** than physically extracted (e.g. ~ **0.5 kg**)
  - for **ex situ** measurement by $\gamma$-ray spec.
Measurement Uncertainty at individual sampling targets

U estimated using *duplicate method*.

Sampling and analytical duplicates at 8 sampling targets for *ex situ*, & 12 for *in situ*

Sampling and analytical uncertainty using robust ANOVA.

<table>
<thead>
<tr>
<th>Sampling Target</th>
<th>Analysis 1</th>
<th>Analysis 2</th>
<th>Analysis 1</th>
<th>Analysis 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Sample 2</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ex situ</th>
<th>In situ</th>
</tr>
</thead>
<tbody>
<tr>
<td>U'_anal %</td>
<td>5.1</td>
</tr>
<tr>
<td>U'_samp %</td>
<td>72.5</td>
</tr>
<tr>
<td>U'_meas %</td>
<td>72.6</td>
</tr>
</tbody>
</table>

- *U'_anal* greater for *in situ* - Shorter counting time, environmental conditions.
- *U'_samp* much greater for *ex situ* - Effect of very small sample mass + heterogeneity.
- Combined *U'_measurement* greatest for *ex situ* - exceeds *in situ* by factor of 6 in this area.
  - Bias in measurements could also be included in U, but no CRM was available
  - No significant bias found between measurements made *in situ* and *ex situ*

UfS estimation for *On site* measurements

- ‘On site’ measurements - when a sample taken from original location, usually prepared and homogenised, measured close to its original location.
- Situation is intermediate in complexity between traditional *ex situ* measurements, and *in situ* measurements

Example: Determination of total petrol hydrocarbons (TPH) in stockpiled soil*

- *ex situ* measurement made by on-site method
  - SiteLAB_UVF 3100 fluorescence spectrometer
- compared against those made in remote laboratory (GC-FID, with more QC)
- Duplicate method applied to both methods in usual way (+CRMs for bias)
  - on site [TPH] 3 x higher than off site measurement, maybe due to:
  - loss of volatile TPH in samples taken off site
  - differences between the definitions of measurand for the two TPH analytical methods

MOU23 Ditch this slide - just put key points elsewhere
(conclusions?)
Microsoft Office User, 07/11/2019
General directions for UfS estimation and reduction (1)

1. **Encourage use of Uncertainty Factor**
   - as a better way to express U in appropriate circumstances

2. **Make available Confidence Limits on estimates of UfS** (and heterogeneity)
   - Enables rigorous comparison of UfS values estimated by different methods and for different analytes

3. **Improve ways of Modifying UfS** in order to achieve Fitness-for-Purpose (FFP)
   - Why some systems behave in predictable ways ($s^2 \propto 1/m$), and others don’t
   - Improve the modelling of UfS for such systems

4. **Encourage and develop Sampling QC procedures** (Section 13 of Guide) to check:
   - whether the conditions present at validation are still present
   - whether initial estimates of U are still applicable
     - Especially where subsequent targets very different (e.g. contaminated land)
     - Perhaps add in separate estimate of heterogeneity for each site?

---

General directions for UfS estimation and reduction (2)

5. **Compile databases of UfS/UoM estimates across a whole sector** to see:
   A. If there are typical values that could be used for prediction of UfS (e.g. by regulators)
      - As Ellison et al. (2017)*, did for food sector
      - ~27 different food products (from field/store/factory/retail)
      - ~75 different analytes
      - Found Horwitz-like relationship
      - Provides estimate of UfS to within ~ an order of magnitude.

   B. Get more evidence to test whether UfS increases as a function of concentration in other sectors*

   ![Graph showing relationship between log(UfS) and log(concentration)](image)

---

MOU13 This slide is general wish list - move later? Or combine with later lists?
Microsoft Office User, 02/10/2019
External factors affecting take up of UfS estimation

- Awareness of UfS - improve with 2019 UfS Guide etc.
- Regulatory and accreditation requirements to estimate UfS
  - improving with sampling in ISO/IEC 17025:2017 clause 7.6.1 to evaluate UfS
- Cost of estimating UfS
  - reduce cost of estimation with unbalanced or simple design
  - emphasise reduction in overall cost by avoiding adverse effects of UfS (e.g. loss of product)
- Including UfS in conformity assessment and compliance decisions
  - find better ways
    - e.g. Food sector currently excludes UfS
    - Assumes samples are ‘representative’, therefore UfS is ‘zero’!
- Management of samplers/sampling process
  - need to integrate sampling into whole measurement process
    - Not administer sampling as a separate process
    - Educate samplers in the measurement process

Conclusions

- Increase range of sectors and situations where UfS is being estimated, e.g.
  - in situ measurements (e.g. sensors) makes UfS estimation more applicable
    - needs more development (e.g. correct for ‘bias’?)
  - Passive measurements - in situ can be better than ex situ
    - due to much larger (x 1000) sample mass, and much lower cost (so better coverage)
    - at micro-scale (EPMA, SIMS, LA-ICPMS etc.)
    - for heterogeneity estimation (e.g. for U_{\text{HET}} candidate CRMs)
    - Microbiology and other new sectors
- Realise benefits of knowing UfS, e.g.
  - Enables critical assessment of FFP of all sampling and analytical methods
  - measurements with higher U can be shown to be fit for some purposes
    - e.g. In situ and sensor measurements
- Encourage integrated management of field sampling as part of measurement process