Data handling in molecular geography

ULRICH M. HANKE
SOIL SCIENCE & BIOGEOCHEMISTRY
"All measurements, however careful and scientific, are subject to some
uncertainties. Error analysis is the study and evaluation of these uncertainties.”

### Origins of errors

<table>
<thead>
<tr>
<th>Inverse relationship</th>
<th>Decreasing sample amount</th>
<th>Increasing potential errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical sample mass (g)</td>
<td>10^1 to 10^3</td>
<td>10^{-3} to 10^2</td>
</tr>
<tr>
<td>Work step</td>
<td>Sampling</td>
<td>Preparation</td>
</tr>
<tr>
<td>Sources of errors</td>
<td>Sampling devices</td>
<td>Lab tools</td>
</tr>
</tbody>
</table>
Two classes of errors

Instrumentation

- Depending on the sensitivity of the analytical device
- Quantification of uncertainty via calibration procedures

Sampling and analytical uncertainty

- Depending on the sample size and number of sample transfer
- Quantification of uncertainty via external standards
Organic geochemistry using isotopes

Molecular investigations in complex matrices (often << 1 g kg$^{-1}$)

1. Extraction of compound classes
2. Identification of molecular structures using chromatography
3. Isolation and concentration of target compounds
4. Measurement of the isotopic composition using mass spectrometry
5. Data handling
Model of constant contamination

Statistical approximation of theoretical and observed frequencies

**Key features**

- Every measurement campaign is affected by the same error
- Extraction of error allows for the correction of data

Error propagation yields:

\[
\sigma^2_{F^{14}C} = \left[ \sigma_{m_c} \left( \frac{F^{14}C_m \cdot m_m - F^{14}C_c \cdot m_c}{(m_m - m_c)^2} \right) \right]^2 + \left[ \sigma_{F^{14}C_m} \left( \frac{m_m}{m_m - m_c} \right) \right]^2 + \left[ \sigma_{F^{14}C_c} \left( \frac{-m_c}{m_m - m_c} \right) \right]^2
\]
Sensitivity test by means of (AMS) accelerating mass spectrometry

Fossil standards (n = 25)
$\chi^2_{\text{CHAR}}$: 2.0 (n=5)
Modern contamination: 1.4±0.5 μg C

Modern standards (n = 17)
$\chi^2_{\text{CHAR}}$: 0.2 (n=5)
Fossil contamination: 1.8±0.5μg C
Summing up

- All physical measurements are subject to errors (even if they are negligible)

- Uncertainties increase with decreasing sample size (Gt → ng)

- Errors originate from sampling, laboratory handling and instrumentation

- Standard materials, reference samples and calibration procedures enable high precision

- ‘Error analysis‘ allows for extraction of errors and subsequently correction of data