RC-6a
Quality Assurance and the Evaluation of Uncertainties in Environmental Measurements
Dra. Lourdes Romero González
Contents

1 - Introduction

2 - Measurement Uncertainty
   Concepts and definitions
   Sources of uncertainty
   Uncertainty components
   Process of evaluating uncertainties
   Reporting uncertainty

3 - Close to Detection/Decision Limits

4 - Final recommendations

5 - Bibliography
Uncertainty has implications for decision purposes.
Radioactivity environmental monitoring (REM) programmes provide relevant information on radioactivity levels in all compartments of the biosphere. Compliance against Regulatory Limits involves a large number of results being compared to basic standards.

To make decisions on the potential risk to humans or the environment, the reliability of the assessment obtained from these programs requires that laboratories producing the analytical data be able to provide results of the required quality. The confidence to be placed in results is possible only if a quantitative and reliable expression of their relative quality: the associated uncertainty is assessed.

Environmental measurements are performed at levels where the radionuclide of interest cannot be distinguished from natural background levels. The relative uncertainty associated with the result tends to increase. Uncertainty has implications for decision purposes.
Introduction

The user of the data reported by a laboratory should be aware that the information provided by measurement can rarely be assumed to be complete.

Laboratories are increasingly requested to demonstrate not only the quality of their results, but their fitness for purpose by giving a measure of the confidence that can be placed on the result.

Over the years different uncertainty evaluation procedures have been developed.

1993 ISO Guide to the Expression of Uncertainty in Measurement (GUM)
- established general rules for evaluating and expressing uncertainty in measurement
- promotes the achievement of international harmonization for stating formally measurement results
- making possible international comparability

1995 EURACHEM Quantifying uncertainty in analytical measurement
application of the concepts in the GUM to analytical measurement

many of the opinions and recommendations here expressed are the outcome of multiple review and discussions within the Working Group

GTINC
Spanish WG for the study of Uncertainties in REM
sponsored by the CSN (Regulatory Body)
from which the author is the coordinator

Over the years different uncertainty evaluation procedures have been developed.

Criteria and Recommendations for Radioactive Determinations
in compliance with MARLAP, ISO 11843 and ISO 11929 series
any measurement result is in general a point estimate of the measured quantity (measurand).

**Error of the measurement**
the difference between the measured result and the actual value of the measurand.

**Uncertainty of measurement** (VIM)
“A parameter associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand”

**Accuracy**
the closeness of the agreement between the result of a measurement and a true value of the measurand
(a measurement is accurate if its error is small)

**Precision** (not defined in VIM)
we take as “a quantitative indication of the variability of a series of repeatable measurement results” (Lira, L., 2002)
**Measurement uncertainty**

**Concepts and definitions**

The measurement model:

- **Input quantity** $X_i$
- **Input estimate** $x_i$
- **Output quantity** $Y$
- **Output estimate** $y_i$

The measurand $Y$ (output quantity) depends upon a number $N$, of input quantities, $(X_1, X_2, ..., X_N)$:

$$Y = f(X_1, X_2, ..., X_N)$$

When measuring, we get an estimate of $Y$ (output estimate $y$) obtained from above using input estimates $x_i$:

$$y_i = f(x_1, x_2, ..., x_N)$$

The **Standard uncertainty of measurement** $u_c(y)$ associated with the output estimate (measurement result $y$) is the standard deviation of the measurand $Y$.

- To be determined from the input estimates $x_i$, and their associated standard uncertainties $u(x_i)$.

The **Combined Standard uncertainty** $u_c(y)$ (total uncertainty of $y$) is an estimated standard deviation obtained by combining all the uncertainty components $u(x_i)$ evaluated using the "Law of propagation of uncertainty".

When the input quantities $X_i$, are uncorrelated, $u_c(y)$, is given by:

$$u_c^2(y) = \sum_{i=1}^{N} \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i)$$

$u(x_i)$ evaluated by using a TypeA or TypeB methods.

$u_c(y)$ is an estimated standard deviation obtained by combining all the uncertainty components $u(x_i)$.

- When the input quantities $X_i$, are correlated to some degree, the covariance also has to be considered.

The **Expanded uncertainty** $U$ provides an interval within which the value of the measurand $Y$ is believed to lie with a higher level of confidence.

$U$ is obtained by:

$$U = k \cdot u_c(y)$$

- The choice of the **coverage factor** $k$ is based on the level of confidence desired.
- for an approximate level of confidence of 95%, the value of $k$ is 2.
Some sources are common to any analytical process:
- incomplete definition of the measurand
- sampling, sub-sampling, storage conditions
- matrix effects/interferences, environmental conditions
- masses and volumetric equipment, reference values
- approximations included in the measurement method
- digital displays and rounding, ...

Radioactive determinations
- many analytical techniques are used before measuring
- measurement involves sophisticated instrumentation.
- specific sources due to the random nature of radioactive decay and radiation counting

Sampling
- other possible causes
  - Radioactive standards
  - Radionuclide half-life
  - Counting efficiency
  - Background
  - Radioactive decay
  - Source geometry & placement
  - Variable instrument backgrounds and efficiencies
  - Time measurements in decay & ingrowth calculations
  - Instrument dead-time corrections
  - Approximation errors in mathematical models
  - Published values for half-lives & radiation emission probabilities

Measurement uncertainty Sources
- The counting uncertainty is the predominant source of uncertainty at the low activity levels encountered in environmental samples
In estimating overall uncertainty, it may be necessary to treat each source separately to obtain its contribution.

- Each of the separate contributions to uncertainty (input estimates) is an *uncertainty component*.
  - when expressed as a standard deviation, is the *standard uncertainty* $u(x_i)$.

Components are grouped into two categories according to the way in which their numerical value is estimated: **Type A or a Type B method of evaluation**.

- **“Type A”**: Uncertainty that is evaluated from the statistical distribution of series of measurements.
  - can be characterised by standard deviations, $s_i$:
    - the associated number of degrees of freedom is $v_i$.
    - and the standard uncertainty $u_i = s_i$.

- **“Type B”**: Uncertainty evaluated by means *other than the statistical analysis* of a series of observations.
  - The standard uncertainty is evaluated by scientific judgement based on all available information on the possible variability of the input quantity: assumed probability distributions based on experience or other information, represented by $u_j$.
  - $u_j$ can be characterised by a corresponding standard deviation:
    - (since the quantity $u_j$ like a standard deviation, the standard uncertainty is $u_j$).
In estimating overall uncertainty, it may be necessary to treat each source separately to obtain its contribution.

- Each of the separate contributions to uncertainty (input estimates) is an uncertainty component when expressed as a standard deviation, is
  - **Type A**: Uncertainty that is evaluated from the statistical distribution of series of measurements. The standard uncertainty is evaluated by
    - a corresponding standard deviation: the associated number of degrees of freedom is \( \nu_j \) and the standard uncertainty \( u_i = s_i \).
  - **Type B**: Uncertainty evaluated by means other than the statistical analysis of a series of observations. The standard uncertainty is eval-
    - uated by scientific judgement based on all available information on the possible variability of the input quantity: assumed prob-
    - ability distributions based on experience or other information, represented by \( u_j \).
    - \( u_j \) can be characterised by a corresponding standard deviation, the standard uncertainty \( u(x) = s \).

Components are grouped into two categories according to the way in which their numerical value is estimated: Type A or a Type B method of evaluation.

### Frequent distributions used for Type B evaluation method

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Form</th>
<th>Use when:</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectangular</td>
<td>( u(x) = \frac{a}{\sqrt{3}} )</td>
<td>A certificate or other specification gives limits without specifying a level of confidence (eg 25ml ± 0.05ml)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>An estimate is made in the form of a maximum range (±a) with no knowledge of the shape of the distribution.</td>
<td></td>
</tr>
<tr>
<td>Triangular</td>
<td>( u(x) = \frac{a}{\sqrt{6}} )</td>
<td>The available information concerning x is less limited than for a rectangular distribution. Values close to x are more likely than near the bounds.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>An estimate is made in the form of a maximum range (±a) described by a symmetric distribution.</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>( u(x) = s )</td>
<td>An estimate is made from repeated observations of a randomly varying process.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( u(x) = s )</td>
<td>An uncertainty is given in the form of a standard deviation ( s ) or ( \sigma ), a relative standard deviation ( s / \Sigma ), or a coefficient of variance ( CV% ) without specifying the distribution.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( u(x) = \frac{CV \times X}{100} )</td>
<td>An uncertainty is given in the form of a 95% (or other) confidence interval I without specifying the distribution.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( u(x) = \frac{I}{2} ) (for I at 95%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Process of evaluating uncertainty

Specify Measurand

Relation between measurand and input quantities: \( Y = f(X_1, X_2, \ldots, X_n) \)

Identify Uncertainty Sources

List all possible sources; parameters, processes, assumptions...

Quantify Uncertainty Components

\( u(x_i) \) evaluated by using a Type A or Type B methods

Convert Components to Standard Deviations

\( u_c(y) \) by using the Law of propagation of uncertainty

Calculate Combined Standard Uncertainty

Review

re-evaluate large Components

¿?

Apply appropriate coverage factor \( k \)

Calculate Expanded Uncertainty

Expression of Results

\(( y \pm U )\) (stating the units)
Reporting Uncertainty

The information necessary to report the result of a measurement depends on its intended use

Guiding principle:
- to present sufficient information to allow the result to be re-evaluated (if new information become available),
- to better enable statistical analyses and to observe trends in the data

References recommend that the result should be reported as expanded uncertainty $U$:

**Result:** $(y \pm U)$ *(stating the units)*

$k$ must always be reported and the confidence level associated to the $y \pm U$ interval

Example: The activity concentration of a radionuclide $(A)$ in a water sample,

$$A = (0.85 \pm 0.13) \text{ Bq/m}^3$$

- The reported uncertainty is an expanded uncertainty calculated using a coverage factor of 2 which gives a level of confidence of approx. 95%

**Rounding**

The number of significant figures that should be reported depends on the uncertainty of the result

- Round the uncertainty (standard or expanded) to either 1 or 2 significant figures and report both the measured value and the uncertainty to the resulting number of decimal places
- **should only be applied to final results**

Intermediate results shall be carried through all steps with additional figures to prevent unnecessary round off errors


**recommended by References**

Uncertainties should be rounded to 2 figures, when possible
Reporting Uncertainty

In radioactive environmental measurements it is possible to calculate results that are less than zero although negative radioactivity is physically impossible.

Negative values may occur when the measured result is less than a pre-established average background level for the particular system and procedure used.

Censoring of results is *not* recommended, these values should be reported to better enable statistical analyses and to observe trends in the data.

All results, whether positive, negative, or zero, should be reported as obtained, together with their uncertainties [references]


Compliance against regulatory limits in REM involves that large numbers of results from environmental radioactive determinations be compared to basic standards or to be within specific limits.
(The uncertainty associated to the result has obviously implications for interpretation of analytical data)

According to section 9.6 of EURACHEM:
“The uncertainty in the analytical result may need to be taken into account when assessing compliance. The LIMITS may have been set with some allowance for measurement uncertainties. Consideration should be given to both factors in any assessment”
Environmental measurements are frequently performed at levels where the radionuclide of interest cannot be distinguished from natural background levels. The relative uncertainty associated with the result tends to increase to the point where the (symmetric) uncertainty interval includes zero.

This region is typically associated with the practical Limit of Detection for a given method.

"Conflicting region" exists some confusion due to:

- the difficulty of establishing a Decision Threshold
- the numerous criteria, terminology and formulation since the first articles on making a detection decision for radioactive measurements were published

All methodologies involve the principles of statistical hypothesis testing.

Resume latest harmonized international criteria, terminology and definitions based on ISO 11843 & 11929 series.
Environmental measurements are frequently performed at levels where the radionuclide of interest cannot be distinguished from natural background levels. The relative uncertainty associated with the result tends to increase to the point where the (symmetric) uncertainty interval includes zero. Does the sample contain a positive amount of the radionuclide?

"Conflicting region" exists due to:
- the difficulty of establishing a Decision Threshold
- the numerous criteria, terminology and formulation since the first articles on making a detection decision for radioactive measurements were published

This region is typically associated with the practical Limit of Detection for a given method.

All methodologies involve the principles of statistical hypothesis testing.

Resume latest harmonized international criteria, terminology and definitions based on ISO 11843 & 11929 series.


Close to Detection/Decision Levels

Environmental measurements are frequently performed at levels where the radionuclide of interest cannot be distinguished from natural background levels. The relative uncertainty associated with the result tends to increase to the point where the (symmetric) uncertainty interval includes zero. This region is typically associated with the practical Limit of Detection for a given method.

"Conflicting region" exists due to confusion caused by:
- the difficulty of establishing a Decision Threshold
- the numerous criteria, terminology and formulation

The application of more advanced statistical techniques, i.e.: Bayesian inference, can be found in:
- Incorporates scientific hypothesis in the analysis (by means of "a priori distribution"), resulting in a distribution of likely outcomes

Resume latest harmonized international criteria, terminology and definitions based on ISO 11843 & 11929 series:

- Limits for Qualitative Detection and Quantitative Determination. Analytical Chemistry 40:587-593 (1968)

All methodologies incorporate scientific hypothesis in the analysis, resulting in a distribution of likely outcomes.
Terminology and definitions adopted by ISO are presented for the basic concepts and criteria (the classical names assigned by Currie are indicated in brackets). Complete definitions, terminology, equations, and explanations on when/how to apply, are described in ISO 11843 and ISO 11929. Symbols utilized for the same terminology (same physical concepts) are different in ISO 11843 & ISO 11929.

Close to Detection/Decision Levels

Decision Threshold, \( R_n^* \)
(Critical Level (Lc) Currie’s)
“allows a decision to be made for each measurement with a given probability of error as to whether the registered pulses include a contribution by the sample”

\[
R_n^* = K_{1-\alpha} \cdot \sigma(R_n=0)
\]

\[
R_n = R_s - R_0
\]

\[
N_n = N_s - N_0
\]

Decision Threshold

Statistical concept: the lowest usable action level

Results of \( R_n \) are compared with \( R_n^* \)

“Critical value of a statistical test for the decision between the hypothesis \( \rho_s = \rho_0 \) and the alternative hypothesis \( \rho_s > \rho_0 \)”

\[
\rho_s = \text{Expectation value of } R_s \quad \text{(gross effect counting rate quotient of the number pulses } N_s \text{ counted during the preselected duration of measurement } t_s; \quad R_s = \frac{N_s}{t_s}
\]

\[
\rho_0 = \text{Expectation value of } R_0 \quad \text{(background effect counting rate, quotient of the pulses } N_n \text{ counted during the preselected duration of measurement } t_0; \quad R_0 = \frac{N_0}{t_0}
\]

\[
R_n = \text{net effect counting rate, } R_n = R_s - R_0
\]

\[
\rho_n = \text{Expectation value of } R_n
\]

\( R_n^* \) is a value chosen so that results above it are unlikely to be false positive, with a probability \( \alpha \) fixed a priori.

A smaller value of \( \alpha \) makes type I errors (false +) less likely, but also type II errors (false -) more likely (sample~blank).
Close to Detection/Decision Levels

Frequently the results of radioactive determinations must meet certain *reference-guide values* established by the user of the results.

Monitoring of the Environmental Radioactivity:

- A minimum value for the so-called Detection Limit for a method is required by the Regulatory body.
- The EU REM sparse network (implemented within the Member States to obtain data on *actual levels* of radioactivity) requires that laboratories provide data with the highest achievable accuracy and high sensitivity measurements (to allow comparison of data sets for extended time periods).

An important performance characteristic of a radioanalytical procedure is the *Detection Capability or Detection Limit*.
Close to Detection/Decision Levels

Decision Threshold, $R_n^*$
(Critical Level (Lc) Currie’s)
“allows a decision to be made for each measurement with a given probability of error as to whether the registered pulses include a contribution by the sample”

Detection Limit, $\rho_n^*$

the lowest useable action level

$R_n^*$
$\rho_n^*$

allows a decision to be made as to whether a MEASURING METHOD satisfies certain requirements and is consequently suitable for the given purpose of measurement

$\rho_n = R_n^* + K_{1-\beta} \cdot \sigma(R_n = \rho_n^*)$

Guideline Value
“Value constituted by requirements on measuring procedures arising for scientific, legal or other reasons which are specified, (i.e.: activity, specific activity, dose rate, ...)”

In the REM programmes this value is fixed by the Regulatory Body for specific activity in the different types of samples

$\rho_s = \text{Expectation value of } R_s = N_s / t_s \quad R_n = \text{net effect counting rate, } R_n = R_s - R_0$

$\rho_0 = \text{Expectation value of } R_0 = N_0 / t_0 \quad \rho_n = \text{Expectation value of } R_n$
Close to Detection/Decision

Decision Threshold, $R_n^*$

(Critical Level (Lc) Currie’s)
"allows a decision to be made for each measurement with a given probability of error as to whether the registered pulses include a contribution by the sample"

Detection Treshold

Detection Limit

$\rho_n^*$

$R_n^*$

the lowest useable action level

Detection Limit, $\rho_n^*$

(Detection Capability) (Ld Currie’s)
"specifies the minimum sample contribution which can be detected with a given probability of error using the measuring procedure in question"

$\rho_n = R_n^* + K_{1-\beta} \cdot \sigma (R_n = \rho_n^*)$

Guideline Value

"Value constituted by requirements on measuring procedures arising for scientific, legal or other reasons which are specified, (i.e.: activity, specific activity, dose rate, ...)"

Uses

- DECISION THRESHOLD $R_n^*$ is to be compared with the MEASURED VALUES
- DETECTION LIMIT $\rho_n^*$ is to be compared with the GUIDELINE VALUE

$\rho_s = \text{Expectation value of } R_s = N_s / t_s$

$R_n = \text{net effect counting rate, } R_n = R_s - R_0$

$\rho_0 = \text{Expectation value of } R_0 = N_0 / t_0$

$\rho_n = \text{Expectation value of } R_n$
Close to Detection/Decision Levels

To stress the significance of producing reliable measurements together with adequate uncertainty evaluation and having the “Conflicting Region” well characterized.

Areas of application of the Detection Capability for a radiochemical procedure

- Fulfilling Safeguard agreements (Treaty on the Non-Proliferation of Nuclear Weapons, NPT)
- Verification activities include monitoring systems to detect the flow of nuclear material past key points (detection of very small amounts of specific radionuclides), to ensure peaceful nuclear activities
- Exemption levels, Clearance of materials,
- Cleanup of contaminated areas,
- Bioassay excreta radioanalyses (internal dosimetry)
- Waste management, ...

Expense and consequences of making incorrect decisions in REM programs

- Reporting false positive in environmental samples, can produce unnecessary costly cleanup, unnecessarily alarm public, spend money on re-sampling, analyses and further investigations
- Reporting a false negative, the consequences could affect directly the population,
  - not protective actions of public and environment would be taken
  - if later discovered can destroy trust and communication > political consequences
Final Recommendations

**Documentation**

**THE VALUE (AND ITS UNCERTAINTY) SHOULD ALWAYS BE REPORTED** if it does not exceed the *Decision Threshold*, the comment “no detected” should be added.

For established sample contributions, in addition to the measured value, confidence intervals and the confidence level shall be reported.

A report on measurements shall be accompanied by details on the probabilities of error, the DECISION THRESHOLD and the DETECTION LIMIT.

**censoring data** means CHANGING measured results from numbers to some other form that cannot be averaged or analyzed numerically.

\[
\text{Result} \leq \text{P}_{\text{th}}(Ld)
\]
Measurement uncertainty

Laboratory measurements always involve uncertainty. Every measured value obtained by a radioanalytical procedure should be accompanied by an explicit uncertainty estimate.

Uncertainty must be considered when:
- analytical results are used as part of a basis for making decisions
- comparing data against Regulatory Limits
- comparing data among results of laboratories from other countries

All results, (positive, negative, or zero) should be reported, together with their uncertainties.
- The coverage factor and approximate probability should be stated with the expanded uncertainty.

Assessment of measured results

**MEASUREMENT RESULT** shall be compared with the **DECISION THRESHOLD**.
- If a result is greater than the **Decision Threshold** $R_n^*$, it is assumed to be a real sample contribution.

Assessment of measuring procedure

the determined **DETECTION LIMIT** shall be compared with the **GUIDELINE VALUE**.
- If the **DETECTION LIMIT** $\rho_n^*$ is greater than the **GUIDELINE VALUE**, the procedure is not suitable for the purpose of the measurement.
Final Recommendations

At the environmental radioactivity levels, the relative uncertainty associated with the measurement result tends to increase:

- Uncertainties should be correctly assessed
- Detection/decision levels must be carefully characterized

*further harmonization of criteria and terminology is needed*

The radiological protection of the environment and the population requires from all states to have laboratories with Internationally comparable Quality levels

*Adequate management of any eventual situation of nuclear emergency can only be assured on the basis of reliable and traceable measurements to international standards*

*To conclude*

- Efforts should be made by the scientific community to have all involved laboratories in closer collaboration for an international harmonization of criteria and terminology and to diffuse this information among the **users of the results**
- To study the use of proper statistics for decision making at the “conflicting region” application of Bayesian Statistics
Bibliography