



EURACHEM / CITAC Guide

Traceability in Chemical Measurement

**A guide to achieving comparable results
in chemical measurement**

2003

EURACHEM/CITAC Guide: Traceability in Chemical Measurement

*A guide to achieving comparable
results in chemical measurement*

Drafting Editors

S L R Ellison (LGC, UK)
B King (UK)
M Rösslein (EMPA, Switzerland)
M Salit (NIST, USA)
A Williams (UK)

Composition of the Working Group

EURACHEM Members

A Williams <i>Chairman</i>	UK
S Ellison <i>Secretary</i>	LGC, Teddington, UK
M Berglund	Institute for Reference Materials and Measurements, Belgium
W Haesselbarth	Bundesanstalt für Materialforschung und Prüfung, Germany
K Hedegaard	EUROM II
R Kaarls	Netherlands Measurement Institute, The Netherlands
R Kaus	EURACHEM Germany
M Rösslein	EMPA St. Gallen, Switzerland
A van der Veen	Netherlands Measurement Institute, The Netherlands
M Walsh	State Laboratory, Ireland
W Wegscheider	University of Leoben, Austria
H van de Wiel	National Institute of Public Health and the Environment, The Netherlands
R Wood	Food Standards Agency, UK

CITAC Members

B King	UK
M Salit	National Institute of Standards and Technology USA
A Squirrel	NATA, Australia

AOAC Representative

A Pohland	AOAC International
-----------	--------------------

EA Representative

W Richter	PTB, Germany
-----------	--------------

Acknowledgements

This document has been produced primarily by a joint Eurachem/CITAC Working Group with the composition shown (right). The editors are grateful to all these individuals and organisations and to others who have contributed comments, advice and assistance.

Production of this Guide was in part supported under contract with the UK Department of Trade and Industry as part of the National Measurement System Valid Analytical Measurement (VAM) Programme.

EURACHEM/CITAC Guide: Traceability in Chemical Measurement

Preface

Measurement underpins a wide range of socio-economic activities, both domestic and international. Every day, thousands of chemical measurements support decisions on food safety, health and environmental protection. The global market, too, needs accurate and reliable measurements so that technical barriers to trade can be minimised. In all these sectors, the concept of “tested once, accepted everywhere” is increasingly important, and the need for reliable measurement results that can be compared across space and time has never been greater. Reliable measurements depend critically on competent staff, validated and tested methods, comprehensive quality systems, and traceability to appropriate measurement references. Recognition of these requirements is underscored by the increasing adoption of standards and measurement quality systems, such as laboratory accreditation against ISO 17025:1999, or the pharmaceutical industry’s GLP and cGMP requirements.

To achieve comparability of results over space and time, it is essential to link all the individual measurement results to some common, stable reference or measurement standard. Results can be compared through their relationship to that reference. This strategy of linking results to a reference is termed “traceability.”

The *International Vocabulary of Basic and General terms in Metrology* (VIM)¹ defines traceability as the:

“property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties.”

This definition implies a need for effort at national and international level to provide widely accepted reference standards, and at the individual laboratory level to demonstrate the necessary links to those standards.

At the national and international level, comparability between national measurement systems is being continually improved by intercomparison of measurement standards at the National Metrology Institute (NMI) level. A multilateral mutual recognition arrangement was signed in 1999 by the member nations of the Meter Convention in response to the need for an open, transparent and comprehensive scheme to give users reliable quantitative information on the comparability of national metrology systems.

Individual measurement and testing laboratories play their part by careful use of appropriate measurement and reference standards for calibration and control of their measurement processes. In an increasingly regulated environment, however, laboratories are under greater pressure to demonstrate that their use of measurement and reference standards is indeed both appropriate and sufficient.

This is particularly true in analytical chemistry. Many of the physical quantities used in routine chemical measurement are underpinned by extensive and effective calibration and traceability systems, making the establishment of traceability for these quantities relatively straightforward. However, the values of chemical quantities involved are typically drawn from a wide range of reference materials and data with varying pedigree and provenance, requiring especial care and judgement in selection of references. Chemical measurements also typically require confirmation

of identity as well as measurement of amount. Another challenge is the measurement of a species in complex matrices, which may influence the apparent value of the measured species. Further, it is not uncommon for useful chemical results to arise from the measurement of operationally defined species, for example “extractable cadmium” (sometimes called “empirical” measurements). In such circumstances, it is not always so straightforward to identify the requirements for traceability, or to demonstrate that the traceability in place is adequate.

The purpose of the present document is accordingly to provide guidance on identifying traceability requirements and establishing traceability of measurement and test results. The document describes a consistent set of principles which laboratories can apply in order to establish traceability for their measurement results, and pays particular attention to the use of appropriate references for chemical quantities.

1. International vocabulary of basic and general terms in metrology. ISO, Geneva (1993)

Contents

EURACHEM/CITAC Guide: Traceability in Chemical Measurement	i
Preface	i
Contents	1
1 Scope and Field of Application	3
2 Introduction	3
3 Principles of traceability	6
3.1 Methods, Measurands and Results	6
3.2 Measurement scales, standards and units	6
3.3 Calibration	7
3.4 Effects on measurement results	7
3.5 Controlling fixed conditions	8
3.6 Controlling variables with calibration standards	9
3.7 Common references allow arbitrary definition	11
3.8 Role of method development	11
3.9 Role of method validation	11
3.10 Traceability and measurement uncertainty	12
4 Traceability: The International Definition	13
5 International System of Quantities and Units (SI)	13
6 Establishing Traceability	15
6.1 Essential activities in establishing traceability	15
6.2 Specifying the measurand and required uncertainty	15
6.3 Choosing a suitable method	16
6.4 Validation	16
6.5 Importance of different influence quantities	18
6.6 Choosing and applying appropriate reference standards	18
6.7 Uncertainty estimation	19
7 Choice of the Reference	20
7.1 Introduction	20

7.2 Physical measurements	20
7.3 Confirmation of identity	21
7.4 Calibration with certified reference materials	21
7.5 Calibration with other materials	22
7.6 Calibration using reference data	23
7.7 Reference materials for Method Development, Validation and Verification.	23
7.8 Assessing the traceability of commercial reference materials	24
 8 Reporting Traceability	 24
 9 Conclusion	 25
 10 Bibliography	 26
 Appendix: Examples of Establishing Traceability	 27
1. Preparation of a calibration standard	28
2. Cadmium Release from Ceramic Ware	32

1 Scope and Field of Application

1.1 This Guide gives detailed guidance for the establishment of measurement traceability in quantitative chemical analysis, based on the definition in the International Vocabulary Of Basic And General Terms in Metrology (VIM) [1]. Though it is aimed principally at testing and measurement laboratories carrying out chemical measurement, the principles are expected to apply from routine analysis to basic research. The document is also intended to assist laboratories in meeting the requirements on traceability of results given in ISO 17025.

1.2 Some common areas in which chemical measurements are needed, and in which the principles of this Guide may be applied, are:

- Quality control and quality assurance in manufacturing industries.
- Measurement and Testing for regulatory compliance.
- Measurement and Testing utilising an agreed method
- Calibration of standards and equipment.
- Measurements associated with the development and certification of reference materials.
- Research and development.

1.3 Though this guide discusses measurement uncertainty and method validation in relation to their role in traceability, a detailed description is not attempted in either case. Readers are referred to the Bibliography for additional guidance.

1.4 Traceability is necessary but not sufficient for reliable results; other measures are necessary. It is accordingly assumed throughout this Guide that, whether carrying out measurements or assessing the performance of the measurement procedure, effective quality assurance and control measures are in place to ensure that the measurement process is stable and in control. Such measures normally include, for example, appropriately qualified staff, proper maintenance of equipment and reagents, use of documented measurement procedures and control charts. Reference 2 provides further information on analytical QA procedures.

2 Introduction

2.1 Good analytical results are essential so that reliable decisions can be made. A key property of good results is comparability; the ability to compare results meaningfully wherever they originate. Comparability is provided by, among other things, traceability to a consistent and agreed set of measurement units and scales. For most chemical measurement results, this is best provided by the SI, the internationally accepted system of units. While it is recognised that other units may be required, this guide will generally assume that measurements will be expressed in, or rely on, SI units.

2.2 Traceability is not a new concept in chemical analysis. Before the advent of automation and instrumental techniques, titrimetry and gravimetry were the workhorses of the chemistry laboratory and even though the average analyst may not explicitly refer to or recognise the significance of uncertainty or traceability, the core elements for their attainment were in place. For example, great care was, and is, paid to the preparation and calibration of volumetric solutions, including their linkage to SI. With more complex measurement methods, it is not

always so straightforward to identify the requirements for traceability, or to demonstrate that the traceability in place is adequate. The purpose of the present document is accordingly to provide guidance on identifying traceability requirements and establishing adequate traceability.

2.3 There has been much discussion, both at workshops and in the literature, about the role of traceability in chemical measurement. This document is based on the following principles, which are fully in line with the VIM definition of traceability:

- Method development establishes a procedure for obtaining an acceptable estimate of the measurand. This procedure includes an equation that describes how to calculate a measurement result from other measured quantities, and specifies the conditions under which this equation is expected to hold.
- Validation demonstrates that this equation and set of conditions is sufficiently complete for the purpose in hand.

Establishing traceability ensures that the values of these measured quantities and the values of the specified conditions are related to appropriate standards. This is achieved by calibration using appropriate measurement standards. Calibration is essential for the critical values in the measurement; for less critical values the required control may be less rigorous.

These basic principles are summarised in Box 1. They are discussed in detail in section 3, and related to the internationally accepted definition of traceability in section 4.

2.4 The document identifies the key elements in establishing traceability as

- i) Specifying the measurand, scope of measurements and the required uncertainty
- ii) Choosing a suitable method of estimating the value, that is, a measurement procedure with associated calculation - an equation - and measurement conditions
- iii) Demonstrating, through validation, that the calculation and measurement conditions include all the “influence quantities” that significantly affect the result, or the value assigned to a standard.
- iv) Identifying the relative importance of each influence quantity
- v) Choosing and applying appropriate reference standards
- vi) Estimating the uncertainty

These activities are discussed individually in sections 6 and 7. Other documents in this series

Box 1. Summary of basic principles

1. We assume that an acceptable estimate y of the measurand value can be obtained from

$$y = f(x_1, x_2, \dots, x_m) \Big|_{x_{m+1}, x_{m+2}, \dots, x_n}$$

that is, y is calculated from $x_1 \dots x_m$ using a relationship f which is valid under measurement conditions specified by $x_{m+1} \dots x_n$.

2. Validation checks that the equation above is sufficient using suitable tests.
3. y is then considered *traceable* to $x_1 \dots x_n$
4. Given that Eq 1 is sufficient, all that is necessary for complete traceability to appropriate references is that all the values x_1 to x_n are themselves traceable or defined values.*

In practice, it is sufficient to ensure that values x_1 to x_n are under sufficient control to provide the required uncertainty in y . For critical quantities, this requires traceable calibration against other reference values. For less critical quantities, less stringent control may be adequate.

*“Defined values”: for example, unit conversion factors, mathematical constants, or the values of constants used to relate some SI units to fundamental constants.

do, however, provide substantial additional guidance. In particular, The EURACHEM/CITAC “Guide to Quality in Analytical Chemistry” [2] describes the implementation of quality systems for chemical measurement. The Eurachem Guide “The fitness for purpose of analytical methods” [3] provides detailed guidance on method validation (item iii) above), while the Eurachem/CITAC guide “Quantifying uncertainty in analytical measurement” [4] describes the evaluation of measurement uncertainty in detail (Item vi). In the present guide, this detail is not repeated, but each of these special topics is discussed briefly to identify their roles in establishing traceability.

3 Principles of traceability

3.1 Methods, Measurands and Results

3.1.1 A *measurand* is a “quantity subjected to measurement”, such as mass, volume or concentration. It is critically important that the quantity to be measured is clearly and unambiguously defined. For example, volume is defined for a specific temperature, and concentration applies to a particular analyte and chemical species. Some measurands are defined in terms of methods used; for example, ‘extractable lead’ would require specification of the extraction conditions^{*}. (Measurands defined in terms of a method are sometimes called ‘empirical’, in comparison with ‘rational’ measurands which can be described without reference to a specific method).

3.1.2 Measurement *methods* are procedures intended to provide *estimates* of the values of measurands. Methods are developed and documented so that they provide reliable estimates, and for the purpose of this document it will be assumed that the method is accepted as providing an adequate estimate for the purpose in hand, and that it incorporates all necessary controls and corrections.

3.1.3 *Results* are values ascribed to measurands following measurement using an appropriate method. Results are accordingly estimates of measurands. Results have properties such as uncertainty, accuracy, and, as will be shown, traceability.

3.2 Measurement scales, standards and units

3.2.1 Meaningful comparisons between measurement results are only possible if the results are expressed in the same *units*. This is actually achieved by quoting measurement results as multiples of a given unit; for example, a mass of 2.1 kilograms has a mass equal to 2.1 times the mass of the international kilogram. The mass of the international kilogram is the ‘unit of mass’. Clearly, in order to express one mass as a multiple of another, the two have to be compared. It is impractical to compare all masses with the international kilogram. This comparison is therefore most commonly indirect, through reference standards, which are in turn calibrated against other standards. This forms a chain of comparisons leading to the relevant primary unit or an accepted ‘realisation’ of a unit. Providing access to consistent units of measurement by means of reference standards is the principal function of traceability; without it, there is no meaningful measurement.

3.2.2 A measurement scale is simply an agreed method of using units of measurement and defining an origin (a ‘zero’ point). Mass, length and concentration are expressed using linear measurement scales with zero at the origin (they are ‘ratio scales’); pH, for example, is a logarithmic scale with a reference at a hydrogen ion activity of 1. When two results are described as being ‘on the same measurement scale’ they are both expressed in the same units and using the same origin.

^{*} Strictly, “extractable lead” is typically an abbreviation for more specific terms such as “*amount concentration* ..” or “*mass fraction*...” of extractable lead.

3.3 Calibration

3.3.1 Section 2.4 stated that calibration is the fundamental process in establishing traceability. It is through calibration that traceability to appropriate reference standards is actually achieved in practice. The following paragraphs review the internationally accepted definition of calibration, and discuss the calibration of parts of a measurement system and of whole systems.

3.3.2 Calibration is described by the VIM as the process of establishing the relationship between values shown by a measuring instrument or system, and the values provided by measurement standards.

3.3.3 Calibration can be (and usually is) applied to parts of a measurement system. In particular, instruments are normally calibrated in isolation, and then used in a larger measurement system. Items such as balances and thermometers are calibrated less frequently because they are relatively stable in the medium term; instruments such as GC or ICP equipment tend to vary much more and are typically calibrated more frequently, often in the same run as a set of test items. For this purpose, one would generally expect to use a pure chemical as the calibration material, though it may be added to a matrix similar to the samples expected in order to reduce matrix effects. Under these circumstances, the reference standard values will appear in the calculation of the result (perhaps indirectly) and it is therefore clear that the result is traceable to these reference values.

3.3.4 In some cases, calibration standards are taken through the complete measurement process. For example, a matrix reference material may be analysed at the same time as the test samples and used to correct the results, or a known amount of material (a 'spike') may be used to estimate *and correct for* the actual analyte recovery during a run. Clearly, if these procedures are employed, either the reference material value or the amount of 'spike' added must appear in the calculation for the result, perhaps via an intermediate 'recovery factor', and the results are accordingly traceable to the value used.

Note: This procedure implicitly assumes that spike recovery is a sufficient correction for analyte losses. See also the note on spiking in section 6.4.1.

3.3.5 One final situation is conceivable, if rare in practice. It may be that during method development and validation, it is decided that a fixed correction should be applied to all future measurements, based on observations of a particular reference material which is not used for regular, day to day calibration. This, too, is calibration, and since the value appears in the calculation of each result, it is meaningful to speak of traceability to the value(s) in question.

Note: This should not be confused with a straightforward bias check using, say, a matrix CRM, which does not generate a correction applied to results.

3.3.6 Note that a calibration remains valid only so long as an instrument remains stable. In practice, this is assured through appropriate QC, and recalibration intervals must reflect the rate of drift.

3.4 Effects on measurement results

3.4.1 Any measurement can be thought of as one or more determinations combined to give a result under specified conditions. For example, analysis of a soil sample for, say, contaminants, typically involves the quantitative determination of the mass of soil taken, and the concentration of analyte in a measured volume of solution containing an extract from the sample. All these parameters are qualified to some extent by the conditions of measurement. Mass is determined by

weighing, strictly *in vacuo* and in a specified gravitational field; volume is typically taken as ‘volume at 20 °C’ and extraction conditions – whether for complete extraction or for a defined partial extraction - are typically defined in terms of time, solvent, and temperature. The mass, concentration and perhaps volume will of course vary from one measurement to the next, as different sized samples are taken – they are the measured values of the ‘variables’ in the calculation of the final result. The extraction and other conditions are usually held close to their nominal values and are not expected to change; they are fixed conditions, and are not generally included in the calculation.

3.4.2 For a given measurement method, if the fixed conditions change, so will the value of the result. For example, if extraction conditions change significantly from those specified in the method, the result will be wrong, just as it would if the mass or concentration values are in error. It follows that both the fixed conditions required for the measurement, and the other measured values obtained and put into the calculation of the result, affect the analytical result. If either fixed conditions or variable measured quantities are incorrect, then so will be the result. These measured values, whether included in the calculation or among specified conditions, are the ‘influence quantities’ for the measurement - all have an influence on the result and all must be controlled. It is simplest to look first at how the fixed measurement conditions are controlled; the control of the variable parameters will be considered later.

3.5 Controlling fixed conditions

3.5.1 If two scientists want to get the same reading for a measurement, the simplest method is to use the same measuring instrument. To continue the soil analysis example with one simple physical aspect of the measurement, if a consistent extraction time is important, then two analysts could simply use the same clock to determine the extraction time. If this is done, it is possible to say that all the results are *traceable* to the time given by the clock; the clock provides the reference standard of time.

3.5.2 This works well, and (at least for a given method) is not even reliant on the clock being correct. As long as the clock is consistent, if everybody involved uses the same clock and times the same interval (i.e. every result is traceable to the clock’s extraction time interval), everybody involved will share a consistent set of conditions, and extraction timing will not cause a spread in the results.

3.5.3 This becomes unworkable very quickly if close control is needed; it is clearly impractical for the same clock to be used by many different scientists at different times and places. What is needed is a collection of clocks which all show the same time. In practice, the simplest way of achieving this is to ensure that all the clocks are themselves compared with a master clock and shown to be indicating the same interval, or corrected so that the correct interval can be deduced from each clock's readings (this is 'calibration' against the master clock). Each analyst using their own clock then generates the same extraction time. Now, it is possible to say not only that each analyst's results are traceable to their own clock's interval, but also that they are all traceable to the master clock. It is this traceability to a single reference standard – the master clock, in the example – that generates consistent measurement in the different laboratories.*

3.5.4 This leads to one key principle;

- traceability to common reference standards allows laboratories to obtain the same set of fixed conditions required for measurements.

This in turn minimises variation due to changes in fixed conditions of measurement.

3.5.5 The issues raised here also apply when measurement conditions are required to vary in a prescribed way. For example when a chromatography column temperature is 'ramped', the times, temperatures and ramp rates all fall into the category of 'conditions of measurement' specified by the chosen method.

3.6 Controlling variables with calibration standards

3.6.1 Very similar principles apply when looking at the measured variables included in the calculation of the result, but the picture is more complex since the values are not supposed to be fixed, but 'consistent' in some way. In particular, each needs to use a consistent measurement scale. This 'consistency' is achieved by using the same calibration standards for successive measurements. The following short discussion develops this concept. For simplicity, only one reference is shown, though of course most measurements rely on several.

3.6.2 Consider two laboratories, A and B, carrying out measurements on samples of broadly the same type (see Figure 1). Each calibrates their equipment using a reference standard with a known nominal concentration (x_1 and x_2 respectively). They calculate their respective results y_1 and y_2 from a calibration equation including the respective values of x . In each case, the result y is a function of the reference value x (usually a simple multiple, assuming a linear response). The reference value x , of course, provides the units of measurement. Where there is such a relationship – one value is calculated from another, reference, value – the calculated value can always be claimed to be traceable to the reference value⁺. Here, y_1 is traceable to x_1 , and y_2 to x_2 , though so far that has very limited value.

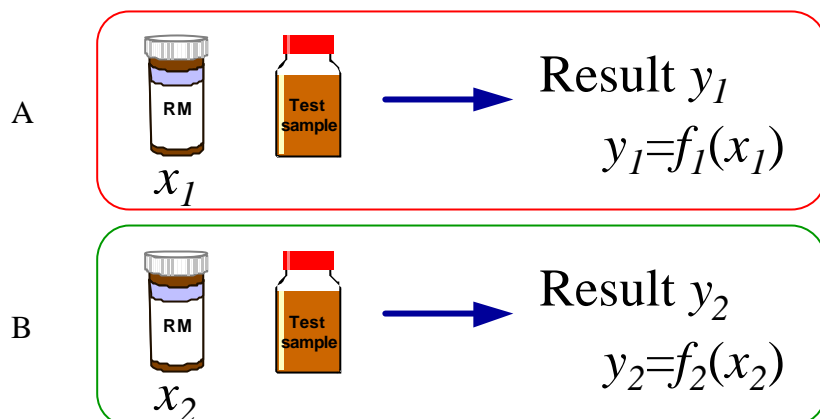
3.6.3 The important question is the relationship between y_1 and y_2 . Though the values of y_1 and y_2 are (usually) different, is the difference genuine, or just due to different references? Clearly, as

* Clearly, in practice, fully calibrated clocks are rarely necessary; simple checks against a time signal are generally adequate for typical time intervals. But the principle is the same; all the clocks are compared with a single reference.

⁺ To support a claim of traceability according to the VIM definition, and to be practically useful, the uncertainty associated with y also needs to be known.

things stand there is no basis for comparing the two results; certainly we cannot write a mathematical equation that would show, for example, y_1 in terms of y_2 .

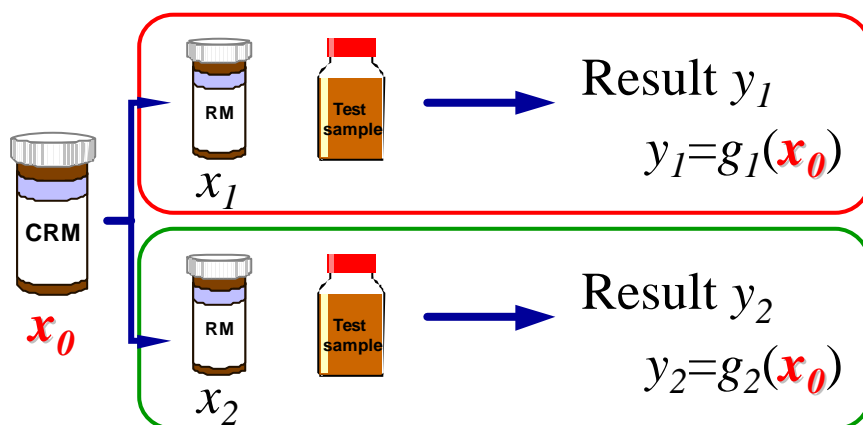
Figure 1



3.6.4 If, however, the two reference standards are both calibrated against some common reference, a comparison becomes meaningful (Figure 2). Now, both results are derived from the same value (x_0). Both will now have the same units of measurement, (the same scale and units as x_0), and direct comparison of the values y_1 and y_2 is now not only possible but also meaningful. By analogy, of course, x_0 could also be derived from a yet higher reference to allow global comparisons.

3.6.5 In this instance, therefore, traceability does not make the results identical; after all, for different samples, they would generally be different. But traceability through calibration permits meaningful comparison by ensuring 'consistency' of measurement units.

Figure 2



3.6.6 This discussion illustrates two further principles:

- When a result is calculated from a reference value, it is traceable to that value.
- Traceability to common references allows meaningful comparisons between results

3.7 Common references allow arbitrary definition

3.7.1 Though the point is abstruse, there is another implication of traceability to a common reference, which is important in metrology. Looking again at figure 2, in principle, it now becomes possible to derive a direct mathematical relationship between y_1 and y_2 in which the value of x_0 is eliminated (at least to first order). For example, in the simple case of linear responses, the ratio y_1/y_2 does not contain x_0 (though in the situation of Figure 1, it would simply include both x_1 and x_2). It follows that if traceability to a common reference is assured, the value of the common reference can, in principle, be defined arbitrarily without affecting the relationship between end results. This is a very useful result; the international kilogram is just such an arbitrary reference, and without traceability to this single artefact, there would be no basis for comparing mass determinations around the globe.

3.8 Role of method development

3.8.1 Method development typically produces a standard operating procedure, incorporating a set of instructions for carrying out a measurement, a set of measurement conditions defining the values of parameters that must be held stable, and an equation from which the result is calculated using the values of the measured parameters. It accordingly provides an equation, which is expected to generate consistent results provided that the specified conditions are correctly set and stable. The implication is that if the values of all these parameters are traceable to stable references, the results will be consistent.

3.8.2 However, this expectation is invariably based on some assumptions; specifically, linearity of response, freedom from overall bias, and absence of other significant effects. If those assumptions are incorrect, for example due to the presence of unsuspected effects, results will be unreliable and often incorrect. Practical experience indicates only too clearly that unknown effects are frequent and often large; such assumptions should therefore not go unchallenged.

3.9 Role of method validation

3.9.1 Method validation, among other important functions concerned with adequacy of performance, is the mechanism used to test these crucial assumptions. It answers the question “are these assumptions valid?” by reviewing the measurement model and making experimental tests of the assumptions, for example by carrying out measurements on appropriate reference materials, or by comparison with the results of independent methods. An overall bias check seeks evidence of significant bias; recovery studies seek evidence of loss of material; linearity checks seek evidence of significant departures from linearity; and ruggedness or similar studies seek evidence for the presence of further effects and so on.

3.9.2 Where an effect is discovered, the method needs to be modified and subjected to further development and validation. Such a modification can take three basic forms:

- elimination of the effect (for example, changing digestion conditions to eliminate precipitation in elemental analysis)
- reducing variation caused by the effect by adding or reducing a control range. For example, it may become necessary to specify a particular operating temperature or range of temperatures to reduce variation.
- correcting for the effect by including it in the calculation of the result.

Notice that the last two actually have the effect of introducing another measurement into the method – that is, another factor requiring traceability.

3.9.3 Where no significant effects are found, the method is considered validated and may be used without modification; the equation, and the specification of measurement conditions, can now be accepted as a sufficient basis for measurement. By implication, of course, the method now explicitly includes all the factors known to require traceability – there are no other known significant effects. If all the identified factors are indeed made traceable to suitable references, the method can be expected to produce consistent results.

3.9.4 The role of validation in establishing traceability is accordingly to test whether the method is sufficiently well defined and incorporates all necessary traceability requirements.

3.10 Traceability and measurement uncertainty

3.10.1 In Figure 2, the two analysts have reached the point where a comparison between their results is at least meaningful. But if they are to decide with any confidence that one sample has a higher level of analyte than another (and not just different results), one more piece of information is essential. The uncertainty of the results is needed.

3.10.2 Uncertainty of measurement is covered in detail in other publications [4, 5] and will not be described in detail here. For the current discussion, the most important points are:

- i) Uncertainty arises, at least in part and sometimes entirely, from inputs to the calculation of the result. Where a reference value is uncertain (and all reference values are uncertain) the uncertainty in the reference value contributes to the uncertainty in the result
- ii) Uncertainty in results therefore arises from the combination of all the uncertainties in reference values and those arising from the measurement procedure, both from random variation and from other causes.

3.10.3 To estimate the uncertainty on a particular result, then, the analyst needs not only the contributions to uncertainty arising from the measurement procedure itself (from precision, operator limitations etc), but also the uncertainty associated with their reference values. It follows that useful measurements with uncertainties can only be provided if all the necessary parameters are traceable to appropriate references *and* the uncertainty on each of those references is known.

4 Traceability: The International Definition

4.1 The previous section has shown that, for consistent and useful measurement results, it is important both that a chain of comparisons to agreed reference standards, and the uncertainties associated with these comparisons, are established. These principles lead directly to the definition of traceability in the International Vocabulary of Basic and General Terms in Metrology (VIM):

Traceability: Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties

NOTES

1 The concept is often expressed by the adjective traceable.

2 The unbroken chain of comparisons is called a traceability chain.

4.2 The definition establishes traceability as a property of measurement results, that is, of values obtained by measurement. Strictly, therefore, the phrase “traceable to a given laboratory” is shorthand for “traceable to a reference value maintained by that laboratory”. Similarly, “traceable to the SI” is shorthand for “traceable to reference values obtained by agreed realisations of the SI units”.

5 International System of Quantities and Units (SI)

5.1 Section 3.2.1 shows that measurements need to be expressed in agreed measurement units. The appropriate system of units for most chemical measurement is the "Système Internationale" (SI). The SI units form a coherent system which is used almost universally in science and very widely in trade.

5.2 The SI defines base units for mass (kilogram, kg), length (metre, m), time (second, s), thermodynamic temperature (Kelvin, K), electric current (amp, A), luminous intensity (candela, cd) and amount of substance (mole, mol). It also defines many derived units in terms of the base units, and a selection of important derived units for chemical measurement is provided in Table 1 (reference 6 provides a more comprehensive set). Note that the mole is the only base unit that requires further qualification; it is essential to specify the *entity* concerned, that is, the specific substance referred to.

5.3 The nature of the substance analysed is, of course, important in all chemical measurements, whether or not they are expressed in moles. In particular, quantities such as mass fraction in chemical measurement are not ‘dimensionless’ in that they invariably refer to the fraction of one substance as a portion of a mixture of other substances. The implication is that for appropriate traceability, each measurement result should be traceable to a reference for the particular substance.

Table 1. Quantities and units in chemical measurement

Quantity	Units
molar fraction	mol/mol, %
mass fraction	kg/kg, %
volume fraction	m ³ /m ³ , %
molar concentration	mol/m ³
mass concentration	kg/m ³
volume concentration	m ³ /m ³
Molality	mol/kg
pH	1 (negative logarithm of hydrogen ion activity)
Enzyme activity	katal (mol s ⁻¹) (SI unit), U (μmol/min)
Purity, an important characteristic for many reference materials and other substances, is generally expressed in terms of one of the fractions or concentrations above	

6 Establishing Traceability

6.1 Essential activities in establishing traceability

6.1.1 The introduction stated a set of activities which are necessary to establish traceability in a working laboratory:

- i) Specifying the measurand, scope of measurements and the required uncertainty
- ii) Choosing a suitable method of estimating the value, that is, a measurement procedure with associated calculation - an equation - and measurement conditions
- iii) Demonstrating, through validation, that the calculation and measurement conditions include all the “influence quantities” that significantly affect the result, or the value assigned to a standard.
- iv) Identifying the relative importance of each influence quantity
- v) Choosing and applying appropriate reference standards
- vi) Estimating the uncertainty

This list does not necessarily imply an order or priority among the activities; they are all important. Some interdependencies will also occasionally result in revisiting prior decisions. The important issue is that they are all carried out adequately for the purpose in hand. For consistency, however, the following paragraphs consider each in turn in the order above.

6.1.2 Note that these steps are sufficient for claiming traceability of results on the assumption that other QA measures, including staff training, measurement quality control etc. are in place.

6.2 Specifying the measurand and required uncertainty

6.2.1 A meaningful measurement requires unambiguous specification of the measurand, or quantity to be measured. For the purposes of this guide, a measurand is usually described adequately in words, but close attention needs to be paid to some specific issues. These are:

- *Identity of the analyte.* Chemical measurement most commonly quantifies particular molecular or elemental species. It will clearly be necessary to take extra care if different forms of a material occur and if the difference is important. For example, different isotopes, isotope mixtures, isotopomers, enantiomers, or crystalline forms may need to be distinguished.
- *Implied measurement conditions.* Most analytical results are expected to be obtained under conditions close to normal ambient temperature, pressure and humidity, and it is common practice to omit these conditions from the stated description of the measurand. In considering traceability, however, it is important to understand exactly what conditions apply, as these form part of the formal definition of the measurand. Where the conditions are not specified, it is normally sufficient to assume that the measurand is defined at NTP (20 centigrade, 1 atmosphere).
- *Recovery correction.* It is most important to state clearly whether the quantity of interest is an amount of material recovered from a substrate, or whether it is the total amount believed to be present. The former is not normally corrected for analytical recovery. The latter may need a recovery correction if recovery is significantly different from 100%. This is important

because recovery correction requires an additional measurement to calculate the correction, and will generally add to the traceability requirements.

- *Specification in terms of a method.* The guidance in this document is unchanged whether a measurand is defined in terms of a procedure or not; definition in terms of a procedure simply leads to a longer list of fixed parameters. Note, however, that defining a measurand in terms of a procedure does not restrict the method used. Though unusual, it is possible in principle to use an entirely different procedure to make the measurements. For example, a purely spectroscopic technique may be used to estimate 'fat content', though 'fat' is most commonly defined in terms of a mass of material extracted under specified conditions. However, it will always be necessary to demonstrate that alternative procedures provide equivalent results.

6.2.2 It is often convenient to consider the required performance of the measurement method at this stage. In considering traceability, the most important concern is the measurement uncertainty required. This is important because:

- a) The uncertainty in a result cannot be better than the uncertainty arising from the measurement standards in use; uncertainty requirements accordingly influence the choice of measurement standards.
- b) For a given measurement technique, achieving a smaller overall uncertainty is likely to require greater control. This will normally increase the number of variables which need to be controlled.

6.3 Choosing a suitable method

6.3.1 Once the measurand is known and understood, a method of measurement is selected, or may be developed especially for the purpose. The choice of method involves a range of factors, including, for example, regulatory requirements for particular methods, customer requirements, cost, experience of different methods, availability of equipment, and criticality of decisions. Choice of method is accordingly a matter of judgement informed by customer needs.

6.3.2 This guide is concerned only with the establishment of traceability for a chosen method. It *assumes* that the choice is the best available in the circumstances. It is for the measurement scientist to decide and, if necessary, demonstrate, whether the method is adequate, and once that is decided, this guide can help show that the results are traceable to appropriate references.

6.3.3 The instructions for the method are expected to include the necessary calculations and to specify any controls required, including but not limited to those required by the definition of the measurand. Typically this will take the form of an equation or set of equations for calculating the measurement result, together with a list of conditions such as times, temperatures, reagent concentrations etc. which must be adhered to. (This equation and set of values is referred to below as the *method specification*, simply for brevity).

6.3.4 The quantities identified in the method specification are all the relevant *influence quantities* for the purpose of establishing traceability, subject to validation as described below.

6.4 Validation

6.4.1 Validation is covered in detail in other sources, and a full discussion is not required here. However, the main requirements relating to traceability need to be considered. First, to fulfil its

role in confirming the adequacy of the method specification, method validation should provide a reasonable test of the measurement equation and conditions. It must be recognised that this cannot be exhaustive, and practical considerations may limit the testing possible. But in an ideal case, validation within a single laboratory will include the following activities for the reasons given:

- Assessment of selectivity and specificity, to ensure that the method responds to the particular species of interest and not to other, similar species.
- A certified reference material check, which demonstrates that the method is not significantly biased by comparison with independently obtained traceable values.
- Reasonable checks on specific, likely effects other than those included in the method specification, which show that no other effects need be included.
- Precision studies over as wide a time interval and set of conditions as reasonably possible, which provide another test for presence of significant unsuspected effects.
- Additional studies on specific and likely sources of bias, including spiking and recovery studies, likely interferences and cross-reactivity studies, which demonstrate, again, that no additional effects are important.
Note: The behaviour of added (“spiked”) analyte may not be equivalent to that of native or “incurred” analyte; spiking may therefore fail to give a true indication of native analyte recovery.
- A linearity check, to demonstrate that the units given may be calculated and quoted as a simple ratio as implied by the normal use of units in measurement results.

Other performance characteristics, such as detection capability, will often be assessed in addition to the above in order to assess fitness for purpose.

6.4.2 Intercomparisons between analysts and different laboratories, or with other methods, can also demonstrate possible deficiencies in the method. If duly treated as tests for additional effects, these, too, will add evidence of the sufficiency of the method specification.

6.4.3 The second important consideration in validation studies is that such references as are used to control, calibrate and test the method during validation are themselves traceable. This is important to ensure that the validation studies are directly relevant to results obtained in routine use.

6.4.4 Validation has been identified (section 3.9, above) as playing a key role in establishing traceability. It is not an optional activity. Even when adopting a standard method which has been validated and thoroughly tested, some level of validation remains necessary. It is not normally necessary to repeat the complete study of all possible or likely effects; the method specification can be taken to be complete without further detailed checking. But analytical methods are complex and consequently prone to human error. It is invariably necessary to at least check that the laboratory can carry out the method correctly (this is often called verification). This is best done with an appropriate certified reference material. Evidence from proficiency tests and other studies may, dependent on the nature of the exercise, also provide evidence of correct operation of a method.

6.5 Importance of different influence quantities

6.5.1 Establishing the relative importance of different influence quantities is crucial in deciding the appropriate degree of control or calibration. It is not always necessary to establish a specific calibration for every quantity.

6.5.2 In general, the importance of different influence quantities is dictated by their quantitative effect on measurement results. Quantities with large effects on the results are likely to be important. A second important issue is the likely effect on the result given the uncertainties or possible gross errors involved. Typically, effects from physical quantities such as time, mass and volume are well controlled and easily measured compared to many chemical effects, particularly at trace levels. Though this situation arises only because a great deal of care has already been paid to physical measurements, it is very likely in practice that an analyst will need to pay far more attention to chemical effects than to intermediate physical measurements.

6.5.3 To decide whether an effect needs to be measured and included in provisions for traceability, it is normally sufficient to consider whether the worst case that might reasonably arise would lead to a significant error in the measurement. If it would not, there is clearly no strong case for additional calibration. For example, ambient temperature in a working laboratory in the UK is extremely unlikely to be outside the range 10-30 centigrade and if such a range is not significant for any measurement within the laboratory, there is no strong case for calibration and control of the room temperature.

6.5.4 A formal uncertainty assessment covering all possible effects (and not just those known to be significant) is clearly an exceptionally powerful tool in deciding the relative importance of different effects. If the uncertainty associated with a particular effect is small compared to the overall uncertainty, further control is unnecessary.

6.5.5 It should be clear that despite the foregoing discussion, environmental and other conditions which are not explicitly stated in the method specification may nonetheless exert some influence on the results. Further, most method development is carried out under relatively restricted environmental conditions and it is rarely possible to test extremes; instead, it is generally assumed that laboratories generally operate in approximately the same conditions as applied in method development. This amounts to an unstated requirement to control environmental or other conditions, and a laboratory will normally be expected to take due care in controlling measurement conditions. In the context of this guide, the most important question is whether such care necessarily extends to traceable measurement and control of conditions. Evaluation of the possible impact should normally follow the principles outlined in paragraph 6.5.3. However, it is common to find that environmental conditions do need some level of control for at least some measurements, and it is accordingly good practice to at least monitor conditions with appropriately checked equipment.

6.6 Choosing and applying appropriate reference standards

6.6.1 To make sure that all the values used in the measurement equation, and all other fixed values used in the measurement are traceable to appropriate references, all that is necessary in practice is to establish procedures for calibration of the equipment measuring or controlling fixed values, and for ensuring the calibration, certification or control of all the references used in the

measurement. Calibration, together with validated methods, is accordingly the key to traceability.

6.6.2 In practice, it is recognised that calibrated and certified reference standards are not always available, but it is always necessary to establish sufficient control by appropriate choice of measurement standards. There are, however, many different types of measurement standard, particularly for chemical measurement, and there are different circumstances for their use. These issues are accordingly discussed in detail in section 7.

6.7 Uncertainty estimation

6.7.1 The requirement for uncertainty information follows from the need to ensure first, that the references used are sufficiently accurate for the purpose, and second, to provide similar information for the result of the measurement. Uncertainty estimation is discussed in detail elsewhere and will not be discussed here. But the minimum required for useful measurements is

- *either* assessing the contribution of each reference value uncertainty to the uncertainty of the measurement result (which may rely on validation to show that changes within the uncertainty make negligible differences to the result)
or if appropriate, complying with the equipment, calibration and control requirements of the standard method (norme) in use.
- assessing the overall uncertainty of the result, including the influence of the references used
- confirming that the overall uncertainty meets end-use requirements

7 Choice of the Reference

7.1 Introduction

7.1.1 Sections 4 and 7 make it clear that appropriate references play a vital role in traceability. The choice of reference is therefore crucial. The following paragraphs consider the choice of reference for:

- Physical measurements made during analytical work
- Confirmation of identity
- Calibrations with certified reference materials
- Calibrations using other materials
- Calibration using reference data
- Method Development, Validation and Verification.

7.1.2 In some circumstances it may not be possible to obtain a suitable certified reference standard. In such cases the limitations on the traceability of the results should be made clear and any adverse effect of this on the applicability of the results should be conveyed to the customer.

7.2 Physical measurements

7.2.1 A large range of physical measurements is common in analytical work. Fortunately, suitable calibration of physical measurement equipment and availability of standards is rarely a major problem in analytical measurement. Equipment and reference standards for mass, length, volume, temperature, time and for electrical measurement normally provide calibration uncertainties well below any level of significance compared to the uncertainties found in analytical measurement. This is, however, entirely dependent upon a long-established and carefully maintained infrastructure, reliant on traceability to national and global references. For all practical analytical work, therefore, reference standards must be chosen to be appropriate to the equipment being calibrated, of sufficiently small uncertainty for the purpose in hand, and their values must be traceable to relevant references. In most cases, this will require a certificate of calibration provided by a competent authority.

7.2.2 Where equipment is calibrated by a third party, and the laboratory does not maintain a calibration standard, the calibration provider must be able to provide a certificate of calibration including uncertainty values. In addition, the laboratory should monitor the continuing performance of the equipment between calibrations, using local, stable check standards to confirm continued operation within calibration uncertainties.

7.3 Confirmation of identity

7.3.1 In most analytical measurements, the identity of the material needs to be confirmed by reference to an authentic sample or reference data^{*}. Identity confirmation by comparison is not generally considered to constitute traceability in the sense defined by the VIM. Nonetheless, due care will always need to be taken in selecting appropriate references for this comparison.

Certified pure materials will often serve for identity confirmation where available in sufficient quantity. Authentic samples from a reputable source are usually adequate substitutes provided that the purity is sufficient to generate an essentially pure response for the analyte of interest.

7.3.2 Comparison with reference data, for example in the form of spectroscopic data, is normally acceptable evidence of identity. In this case, however, it is important to ensure that

- the reference data are obtained under closely similar conditions to those used in the laboratory. For example, infrared spectra of solids should be compared with reference data for solids and not for solutions unless the effects of the change of state are taken into account.
- the reference data and the data on the test items are traceable to appropriate references (for example, wavelength standards) so that direct comparison is possible.

For most instruments currently available, traceability with adequate uncertainty for confirmation of identity is easily achieved via routine calibration and quality control.

7.3.3 Identity can also be established by, for example,

- Knowledge of the synthetic route. For example, if acetic acid is reacted with ethanol then ethyl acetate can be expected.
- Measurement of basic characteristics that allow the identity to be deduced. For example, elemental composition, molecular weight, or presence of specific functional groups, may be established.
- Direct comparison with authentic materials.

7.4 Calibration with certified reference materials

7.4.1 A ‘certified reference material’ is formally defined [1] as a “reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes traceability to an accurate realisation of the unit in which the property is expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence”. The key features which distinguish these materials from other calibration standards are therefore

- Demonstrable traceability to national or international standards.
- A statement of uncertainty.

7.4.2 In many cases it is possible to purchase certified chemical reference standards or calibration solutions and it is convenient and usually cost effective to utilise these. Because the values are traceable to national or international standards and are consequently very reliable, their use is

^{*} Some techniques, such as NMR spectroscopy, may provide sufficiently predictable responses from theory and/or model systems that identity can be confirmed without an authentic sample, but this is not common in general analysis.

recommended where practicable. The supplier should be asked to provide information about the traceability of the value of the reference material supplied.

7.4.3 In some cases, e.g. XRF analyses of alloys, it may be possible to use a suitable certified matrix reference material for calibration. In these cases the result is clearly traceable to the value of the reference material. The points to consider when choosing such a matrix material are the same as those identified in section 7.7. Note, however, that certified matrix reference materials are not generally recommended for calibration; the cost is typically prohibitive, sufficiently close matrix matching is rare in practice, and the uncertainties associated with the values for certified natural matrix materials are often too large for calibration purposes.

7.5 Calibration with other materials

a) Pure materials

7.5.1 In many cases, the measurand is an amount of a chemically distinct substance; an element or single molecular species. Chemists have a long history of isolating and purifying such substances, and it is common to find relevant materials of purity sufficient to serve as reference standards. This follows from an almost unique feature of chemical measurement; 100% purity forms a natural reference value, which cannot be exceeded. Coupled with widely available and excellent reference data for atomic and molecular weight, and often with additional data on physical parameters such as density, a high purity material represents a local, practical realisation of concentration units, through conversion of mass to molar quantity. Calibration with materials of well-established purity is accordingly a valid means of establishing traceability.

7.5.2 Establishing purity relies primarily on appropriate techniques for preparing and purifying a material (which provides a strong expectation of high purity), followed by reasonable efforts to detect significant *impurity*, usually by application of a battery of techniques capable of detecting a wide variety of likely contaminant. The reliability of these processes cannot readily be verified except by long experience and sound professional judgement based on a good understanding of the chemistry involved. Without clear evidence of traceable values of known uncertainty, the adequacy of such a material can only be a matter of care and judgement. Laboratories will, in general, need to be particularly careful to ensure reliable supply, to check materials as required, and to use all reasonable checks to confirm reliability of uncertified pure materials.

7.5.3 Preparation of pure reference materials is sufficiently costly that most working analytical laboratories will not undertake it. Nonetheless, there will be many cases where in-house preparation is the only option, the most common, perhaps, being the need to test for a proprietary material synthesised in house. Such a material should be checked by all available means, typically including (but not limited to) melting point and other thermal properties, spectroscopic evidence of several independent types, moisture determination, non-metal contamination, checks for inorganics (in organic materials), elemental microanalysis, chromatographic examination, and specific checks for any likely impurities.

7.5.4 Finally, even when materials of good purity are available, the continuing need for trace analysis leads to a requirement for low-level solutions of material, and at low concentrations, the analyte content and purity of the material are frequently affected by secondary effects such as container adsorption, contamination, oxidation etc. Considerable care in supplier selection will be necessary, as will care in use and storage, and it is wise to check successive batches of material against one another.

b) Other reference materials

7.5.5 A wide range of other materials and formulations is available for calibration, including, for example, mixed element calibration solutions, alloys, and carefully characterised novel pharmaceutical reference materials. Without formal evidence of traceability and associated uncertainty information, it must be the laboratory's responsibility to demonstrate that the materials are fit for their intended purpose. As in 7.5.4, considerable care in selection is necessary.

7.6 Calibration using reference data

7.6.1 In some situations, reference data are used either to support calibration using a well characterised material, or as calibration factors. Examples might be the use of reference spectroscopic data to calibrate wavelength scales (as in infrared spectroscopy), or the use of reference absorbance data to establish concentrations directly from absorbance measurements. In such cases, the values of the measurement results are traceable to the reference data.

7.6.2 As in section 7.3, it is important to ensure that

- the reference data apply under the conditions used in the measurement (which may involve appropriate conversion methods)
- the reference data are traceable to appropriate references (for example, traceability to wavelength standards is important in using spectroscopic absorbance data)

7.7 Reference materials for Method Development, Validation and Verification.

7.7.1 As is pointed out in sections 4 and 7, reference materials, particularly matrix reference materials, play an important role in method development, validation and verification, and their use for this purpose is strongly recommended. It is, however, important that the material should not only provide traceable reference values, but should also be relevant to the application.

7.7.2 Matrix effects and other factors such as concentration range can be more important than the uncertainty of the certified value. The factors to consider include:

- Measurand (analyte)
- Measurement range (concentration)
- Matrix match and potential interferences
- Sample size
- Homogeneity and stability
- Measurement uncertainty
- Characterisation and certification procedures (measurement and statistical)

7.7.3 More detailed information on the choice and use of reference materials is given in EEE/RM/058. [7]

7.8 Assessing the traceability of commercial reference materials

When choosing a supplier of reference materials the following factors should be taken into account:

- a) Conformance of the production of the reference materials with quality standards such as ISO Guides 34, ISO 17025 or ILAC requirements [8]. Conformance should preferably be demonstrated through third party assessment.
- b) Track record of both the producer and the material. For example, whether the RM in use has been subjected to an interlaboratory comparison, cross-checked by use of different methods, or there is experience of use in a number of laboratories over a period of years.
- c) Availability of a certificate and report conforming to ISO Guide 31.
- d) The validity of the certification and uncertainty data, including conformance of key procedures with ISO Guide 35.

7.8.1 Some or all of these requirements may be specified in the customer and analytical specification, but often it will be necessary for the analyst to use professional judgement. Note that quality does not necessarily equate to small uncertainty and fitness for purpose criteria need to be used.

8 Reporting Traceability

8.1 Evidence of traceability is reported in calibration certificates (for which it is mandatory under ISO 17025) or where required by a customer. Such a report should identify (by reference to other available data if necessary)

- All the chemical calibration standards used
- Where significant, the identity of reference standards used to control the conditions of measurement.

8.2 It is not normal practice to give details of traceability on test reports. However, where it is necessary to report evidence of traceability of results, the report will typically include:

- The identity of calibration standards used
- Where significant, the identity of references used to control the conditions of measurement.

9 Conclusion

9.1 This guide has presented a discussion of the principles underlying the establishment of traceability for a method in use by a calibration, measurement or testing laboratory. The document takes the view, summarised in the Introduction, that:

- Method development establishes an optimised procedure for obtaining an acceptable estimate of the measurand, including the calculation and a set of measurement conditions.
- Validation demonstrates that this calculation and set of conditions is sufficiently complete for the purpose in hand.
- Once these conditions are met, the laboratory need only establish traceability or control for each value in the equation and for each of the specified conditions.
- Traceability, established by calibration using an appropriate measurement standard, is essential for the critical values in the measurement; for less critical values, it is recognised that the required control may be less rigorous.

The detailed discussion of traceability principles, and the activities that are necessary, are developed from this viewpoint to provide a self-consistent and practical approach to establishing and demonstrating the traceability of results.

9.2 It is important, in closing, to note that these simple principles apply well only in the context of a sound quality control and assurance system, and that is an important assumption made in this guide. No amount of attention to traceability, as discussed in this guide, will provide a useful result if the wrong method is chosen, if experience and training are inadequate, or if a method is used well outside its scope. But given good attention to all the other factors necessary for laboratory competence, adherence to this guide will allow a laboratory to declare that its results are fully traceable to appropriate references.

10 Bibliography

- 1 International vocabulary of basic and general terms in metrology. ISO, Geneva (1993)
- 2 Eurachem/CITAC Guide: Guide to Quality in Analytical Chemistry (2002). Available from the Eurachem secretariat, <http://www.eurachem.org/>, or CITAC at <http://www.citac.cc/>
- 3 The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics. Eurachem (1998). Available from <http://www.eurachem.org/>
- 4 A Williams, S L R Ellison, M Roesslein (eds); Eurachem/CITAC Guide: Quantifying Uncertainty in Analytical Measurement, 2nd Edition (2000). ISBN 0 948926 15 5. Available from the Eurachem secretariat (Europe), from LGC, Queens road, Teddington, England (UK) or <http://www.eurachem.org/>
- 5 Guide to the expression of uncertainty in measurement, ISO, Geneva (1993)
- 6 B Taylor. NIST Special Publication 811: Guide for the Use of the International System of Units (SI). (U.S. Government Printing Office, Washington, DC, October 1995).
- 7 EEE/RM/062: The selection and use of reference materials; A basic guide for laboratories and accreditation bodies (2002). Available from the Eurachem secretariat or website (<http://www.eurachem.org/>).
- 8 ILAC Guidelines for the Competence of Reference Material Producers, ILAC G12, (2000) (see www.ILAC.org)

Appendix: Examples of Establishing Traceability

The following examples of traceability are based on those in the EURACHEM/CITAC guide “Quantifying Uncertainty in Analytical Measurement”. This is available from the EURACHEM and CITAC web-sites.

The format of each example follows the list given in section 2.4 and 6.11, which sets out the following activities require to establishing traceability:

- i) Specifying the measurand, scope of measurements and the required uncertainty
- ii) Choosing a suitable method of estimating the value, that is, a measurement procedure with associated calculation - an equation - and measurement conditions
- iii) Demonstrating, through validation, that the calculation and measurement conditions include all the “influence quantities” that significantly affect the result, or the value assigned to a standard.
- iv) Identifying the relative importance of each influence quantity
- v) Choosing and applying appropriate reference standards
- vi) Estimating the uncertainty

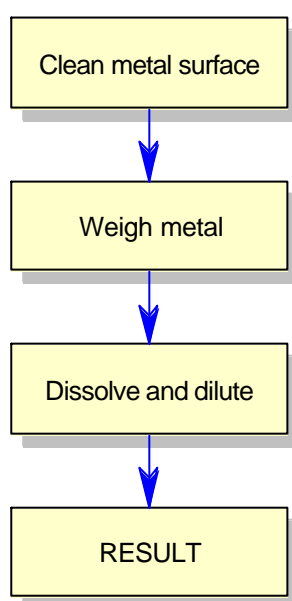
1. Preparation of a calibration standard

Specify the measurand and acceptable uncertainty

A calibration standard is to be prepared, for use within the laboratory, from a high purity metal (Cadmium) with a concentration of $\gg 1000 \text{ mg l}^{-1}$ with a required combined standard uncertainty of 2 mg l^{-1} or smaller. The concentration is defined at 20°C . Because of the small uncertainty required, the use of commercial calibration solutions is not feasible.

Establish the procedure to prepare the calibration standard

The surface of the high purity metal is cleaned to remove any metal-oxide contamination. Afterwards the metal is weighed and then dissolved in nitric acid in a volumetric flask.



The separate stages are:

- i. The surface of the high purity metal is treated with an acid mixture to remove any metal-oxide contamination. The cleaning method is provided by the manufacturer of the metal and needs to be carried out to obtain the purity quoted on the certificate.
- ii. The volumetric flask (100 ml) is weighed without and with the purified metal inside. The balance used has a resolution of 0.01 mg.
- iii. 1 ml of nitric acid (65% m/m) and 3 ml of ion-free water are added to the flask to dissolve the cadmium (approximately 100 mg, weighed accurately). Afterwards the flask is filled with ion-free water up to the mark and mixed by inverting the flask at least thirty times.
- iv. The concentration is calculated from

$$c_{Cd} = \frac{1000 \cdot m \cdot pur}{V} \quad (\text{mg /l})$$

Where

- c_{Cd} : concentration of the calibration standard (mg/l)
1000 : conversion factor from (ml) to (l)
 m : mass of the high purity metal (mg)
 pur : purity of the metal given as mass fraction()
 V : volume of the liquid of the calibration standard (ml)

Mass, purity and volume are all part of the equation, and are consequently influence quantities and expected to be appropriately controlled. Noting that the specification of the measurand implicitly includes the temperature as a fixed value, it follows that the four values which need to be considered for traceability are mass, purity, volume and temperature.

Validation

Validation is a prerequisite in establishing traceability. For this simple and well-understood procedure, the principal influences are well known. However, an important assumption is the implicit assumption of complete dissolution of the material. To check this in practice, a simple cross-check against an independent preparation is normally sufficient. The validation therefore consists of two major parts. First a calibration solution with a similar combined standard uncertainty has to be obtained. This solution could be either the calibration solution used before in the same laboratory, a solution which has been prepared according to a different procedure, or a solution provided by a national standard program, like an SRM solution from NIST. Second the concentration of the two solutions has to be compared using an analytical technique with measurement capabilities sufficient to detect the kind of gross effect which might arise from incomplete dissolution or reprecipitation. On performing this check, using high performance induced coupled plasma optical emission spectrometry (HP ICP-OES), good agreement is found between observed and expected values. In the light of long experience of dissolution, this is sufficient to confirm the sufficiency of the simple specification.

Identify the relative importance of each influence quantity

Mass, purity and volume are all clearly critical, since they form part of the calculation for the result. The relevant references will accordingly need to be chosen with close attention to their uncertainty. Temperature, however, is not part of the equation, and following sections 6.5.2-3 it is useful to consider whether special attention is required. Section 6.5.3 suggests a 'worst-case' check. The following effects (in mg/l Cd) of different temperature errors were estimated assuming aqueous solution:

Temperature error (°C)	Concentration error (mg/l Cd)
10.0	2.00
5.0	1.00

1.0	0.20
0.1	0.02

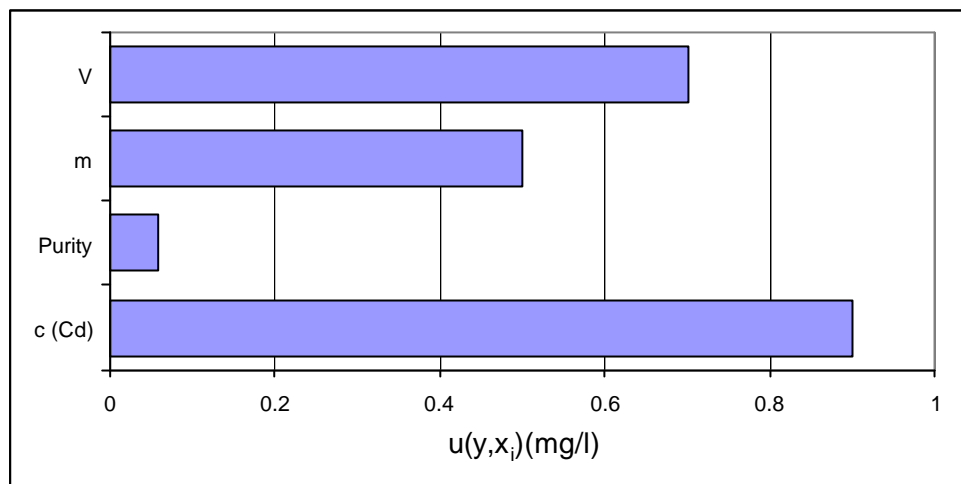
Clearly, the natural temperature range (represented by the 10°C error following the example in section 6.5.3) is likely to be unacceptable. But an error of 5°C leads to an error of only 1 mg/lCd, significantly less than the required uncertainty. This is readily achievable in a routine laboratory with ordinary temperature control. It is likely that no additional measurement or calibration will be required, though as indicated in section 6.5.5, temperature monitoring would be sensible.

Choosing and applying appropriate references.

- The mass m needs to be traceable to reference standards with sufficiently small uncertainty. This is provided routinely by normal calibration procedures for the balance, and confirmed by the associated calibration certificate. Since calibration intervals are relatively long for analytical balances, the linearity is checked on a regular basis with the internal check weights of the balance to stay within the limits given in the manufacturer certificate. Its validity is further reviewed with daily check-weights, which are traceable to national standards and capable of showing significant deviation from nominal values.
- The purity is the certified property of a reference material, as a certified by the supplier, and the uncertainty is demonstrably small enough for the purpose (see the uncertainty figures below). Provided that the metal surface is cleaned according to the instructions given by the supplier, purity can be considered traceable with adequate uncertainty.
- The volume is measured using a flask chosen from a manufacturer who provides information about the traceability of the flask to a national standard, through a calibration certificate. The resulting uncertainty is a substantial contribution, but acceptable. Because glassware can deform slightly over time, and the glassware calibration is a dominant uncertainty source, the volume of the flask is checked regularly by weighing the given volume of water.
- The flask has been calibrated with water at a temperature of 20°C. A check on the laboratory temperature shows effective control within $20 \pm 4^\circ\text{C}$, which is within acceptable limits as expected (see above), so equilibration of solutions at room temperature is sufficient. The laboratory temperature must clearly be monitored using a thermometer with a smaller uncertainty; in practice this can be readily achieved with an ordinary mercury-in-glass thermometer checked against a calibrated thermometer.

Estimating the uncertainty

The estimation of the combined standard uncertainty is described in the first example in the guide "Quantifying Uncertainty in Analytical Measurement" (second edition). The overall uncertainty and major contributions are shown in the figure below. Note that the volume uncertainty includes a temperature uncertainty contribution equivalent to approximately 0.4 mg/l, based on an ambient temperature range of $20 \pm 4^\circ\text{C}$, confirming the acceptability of the ambient temperature control.



2. Cadmium Release from Ceramic Ware

Specify the measurand and acceptable uncertainty

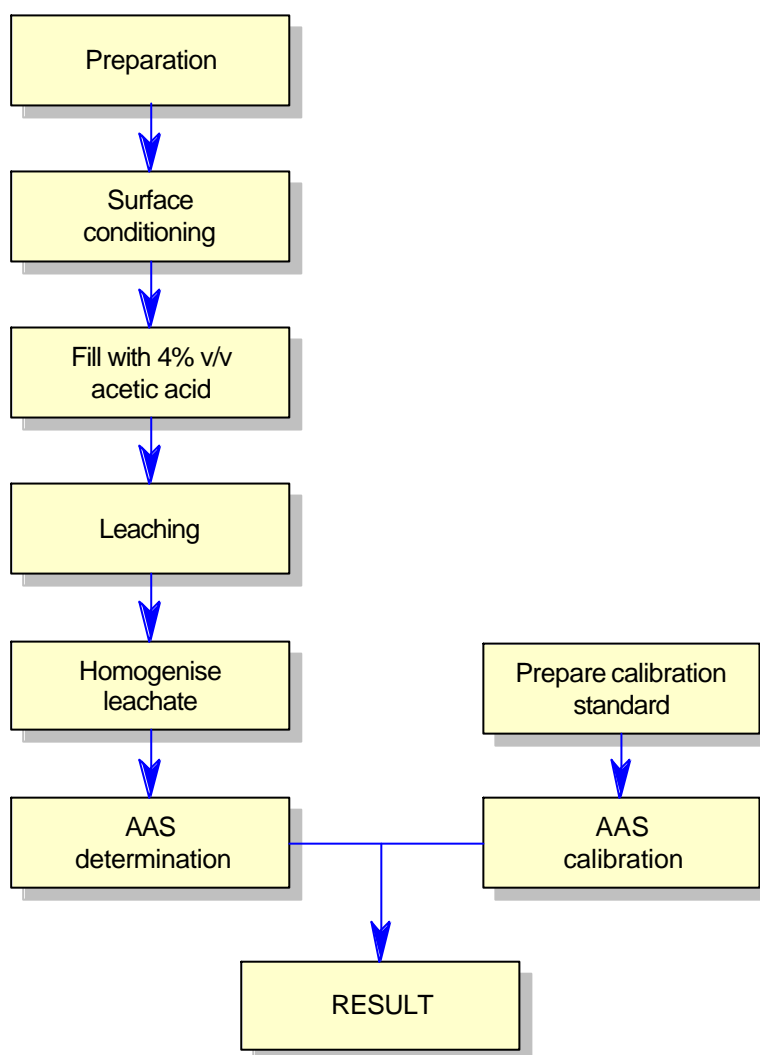
The amount of released cadmium from ceramic ware is determined using atomic absorption spectrometry. The procedure employed is the empirical method BS 6748. The AAS spectrometer needs a detection limit below 0.02 mg/l Cadmium. The acceptable standard uncertainty for this empirical method is 10% (expressed as relative standard deviation).

Establish the procedure to determine the Cadmium release from ceramic ware

The complete procedure is given in British Standard BS 6748:1986 “Limits of metal release from ceramic ware, glass ware, glass ceramic ware and vitreous enamel ware” and this forms the specification for the measurand. Only a general description is given here.

The general procedure is given in the following steps and illustrated schematically:

- i. The sample is conditioned to (22 ± 2) °C. Where appropriate (‘category 1’ articles), the surface area of the article is determined. For the example, a surface of 2.37 dm² was obtained.
- ii. The conditioned sample is filled with 4% v/v acid solution at (22 ± 2) °C to within 1 mm from the overflow point, measured from the upper rim of the sample, or to within 6 mm from the extreme edge of a sample with flat or sloping rim.
- iii. The quantity of 4% v/v acetic acid required or used is recorded to an accuracy of $\pm 2\%$ (in this example, 332 ml acetic acid was used).
- iv. The sample is allowed to stand at 22 ± 2 °C for 24 hours in darkness with due precaution to prevent evaporation loss.
- v. After standing, the solution is stirred sufficiently for homogenisation, and a test portion removed, diluted by a factor d if necessary, and analysed by AA, using appropriate wavelengths and in this example, a least squares calibration curve.
- vi. The result is calculated and reported as the amount of cadmium in the total volume of the extracting solution, expressed in milligrams of cadmium per square decimetre of surface area for category 1 articles, or milligrams of cadmium per litre of the volume for category 2 and 3 articles.



The apparatus and reagent specifications affecting the uncertainty are:

- A freshly prepared solution of 4% v/v glacial acetic acid in water, made up by dilution of 40 ml glacial acetic to 1 l.
- A $(500 \pm 1) \text{ mg l}^{-1}$ standard cadmium solution in 4% v/v acetic acid.
- Laboratory glassware is required to be of at least class B that will not release detectable levels of cadmium in 4% acetic acid during the test procedure.
- The atomic absorption spectrophotometer is required to have detection limits of not greater than 0.02 mg l^{-1} for cadmium.

The amount of cadmium in the total volume of the extracting solution per milligram of cadmium per square decimetre of surface area, determined using the method specified in British Standard BS 6748:1986 (r), expressed in mg/dm^2 , is calculated from^{*}

$$r = \frac{c_0 \cdot V_L}{a_V} \cdot d \quad (\text{mg}/\text{dm}^2)$$

Where

- r : the result; mass of Cd leached per unit area (mg dm^{-2})
- V_L : the volume of the leachate (l)
- a_V : the surface area of the vessel (dm^2)
- d : factor by which the sample was diluted
- c_0 : concentration of cadmium in the extraction solution (mg l^{-1})

With

$$c_0 = \frac{(A_0 - B_0)}{B_1}$$

- A_0 : absorption of the metal in the sample extract
- B_0 : intercept of the calibration curve
- B_1 : slope of the calibration curve

There are four parameters in the equation for the result, and three additional parameters specified in the method to control the leaching process. This gives seven important influence quantities: concentration of cadmium in the extract solution, volume, area, dilution factor, acid concentration, time, and leaching temperature.

Identifying the relative importance of each influence quantity

This standard method gives explicit directions for the control of all the influence quantities, including tolerances on measuring equipment and calibration standards. There are only two noteworthy issues; length-related measurements, and the particular calibration method used for the spectrometer.

Length measurements underpin both the area determination and the volume of leachant used, as the latter is specified by reference to a measurement between the surface of the liquid and the edge of the vessel to be tested. Specifically, British Standard BS 6748:1986 requires the vessel to be filled to within 1 mm from the overflow point, measured from the upper rim of the sample, or to within 6 mm from the extreme edge of a vessel with flat or sloping rim. The requirements

^{*} Note that in reference 4, this equation is expanded to include factors for reagent concentration, time and temperature, simply to make uncertainty estimation explicit. Here, only the calculation used for the result is presented, in accordance with section 6.

themselves are not especially stringent, but still reduce the possible errors in filling to 1-2% for most practical purposes. It is consequently clear that the measurement of the tolerances (1 mm or 6 mm respectively) will have little effect on the test results as long as the requirement is met.

Area measurement will be harder to achieve with sufficient uncertainty, principally because of simple practical difficulties in measuring interior dimensions for even simple shapes. However, while care will be needed in the measurements, control of the ruler or caliper used is a relatively minor problem. Typical requirements will be to measure of the order of 10 cm, and most technical rulers can easily measure this with uncertainties well below 1% (as rsd). While the area measurement is important, therefore, the actual measuring device is unlikely to require close attention.

Though the method specifies the uncertainty for the calibration solution, the exact application of this measurement standard is at the laboratory's discretion. This is considered further under 'validation'.

Validation

This method is an established standard, previously validated, and the list of parameters is accordingly taken as complete. There is, in addition, substantial literature on the process, which confirms that the time, temperature and acid concentration are the sole important parameters in leaching into an unstirred solution.

The standard method does not specify the exact form of the calculation of c_0 , permitting any method with suitable performance. This clearly places the responsibility for choice of AA determination method, and its validation, on the laboratory. The measurement technique is accordingly validated, including a linearity check using serially diluted calibration standards, a precision check, limit of determination (to confirm that the measured value is within the linear range), and a bias check using an independently prepared reference solution. Instrument operating parameters such as pump flow rate were varied to check for significant effects. These measures confirm that, provided the calibration is performed in the same analytical run as the test solutions are measured and the instrument parameters are not changed during the run, there are no additional significant factors. The equation can accordingly be accepted as sufficiently complete, and no additional parameters need be considered.

Choosing the references

The equation of the measurand has the seven parameters concentration of Cadmium in the extract solution, volume, area, dilution factor, correction factor for acid concentration, for time and for temperature. In order to establish traceability of the result is necessary to establish the traceability of these parameters with adequate uncertainty.

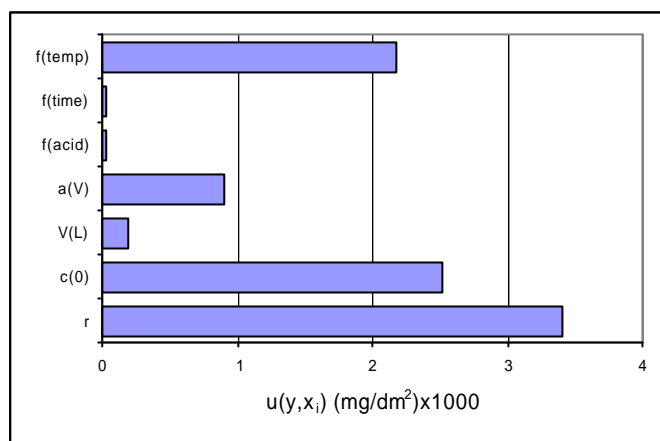
- *Concentration* c_0 , A_0 , B_0 and B_1 relate the concentration of the extraction solution, which has the largest contribution to the overall uncertainty, to the concentration of the calibration solutions, establishing traceability to the calibration solutions. These calibration solutions were obtained by diluting the reference solution of $(500 \pm 1) \text{ mg l}^{-1}$ cadmium in 4% v/v acetic acid. The reference solution is traceable to NIST SRM solution according to the certificate of the manufacturer. NIST has shown its measurement capabilities for determining the concentration of cadmium in solution in a key comparison at CIPM. The dilution steps were done using volumetric glassware, whose manufacturer

specifies the value of the volume and its tolerance. Details were also available about the procedure used to establish traceability to the SI. The calibration solutions were measured using atomic absorption spectrometry and then the absorption values and the concentration of the calibration solution were employed to calculate the intercept (B_0) and slope (B_1) of the calibration curve by least square linear regression. These activities achieve traceability for c_0 .

- V_L is the volume of the leachate. It is measured using a measuring cylinder. The volume determined by the measuring cylinder is adequately controlled by manufacturing tolerances according to the glassware standards referenced, so purchase from a reputable source according to specification is sufficient. As usual, however, a brief check on laboratory glassware on receipt, if only against similar equipment, is prudent to guard against the occasional, if rare, manufacturing error.
- The length measurement is done by placing a mark on the vessel employing a ruler to check the distance of 1 or 6 mm. This is not a critical measurement, so specific calibration of the ruler is unnecessary. As a matter of ordinary prudence, however, the ruler was checked on initial receipt in the laboratory using a calibrated vernier caliper (available for other measurements).
- a_V is the surface area of the vessel. It is measured using a ruler checked as above.
- d is a factor by which the sample was diluted. It is not used in this determination, therefore no traceability statement is needed.
- *Acid concentration.* The British Standard BS 6748:1986 specifies the values for the acid concentration, based on glacial acetic acid of stated purity. Because the influence of changes in acid concentration is small (the uncertainty estimate is based on the manufacturer's purity range), no further measures are required for sufficient traceability to the SI.
- *Time.* The duration of the leaching process is specified in the BS and has to be controlled. Because the time has such a minor influence on the combined standard uncertainty, it is sufficient to control the duration with a normal laboratory clock, which need only be checked occasionally against an appropriate time signal.
- *Temperature.* BS 6748:1986 quotes a temperature range of 22 ± 2 °C. Because the temperature influence is the second largest contribution to the overall uncertainty, it is necessary to control and measure it with a thermometer, which is checked and calibrated by the manufacturer every two years. The manufacturer has accreditation to perform this calibration service. In this way traceability to the SI is established.

Measurement uncertainty evaluation

The measurement uncertainty evaluation is described in EURACHEM/CITAC Guide "Quantifying uncertainty in analytical measurement" second edition page 70 – 78. The uncertainties arising from the different influence quantities are given in the following figure (the time, temperature and acid concentration uncertainty contributions are associated with nominal correction factors, introduced solely to support uncertainty estimation⁴).



EURACHEM / CITAC Guide CG 4

Quantifying Uncertainty in Analytical Measurement

Second Edition

QUAM:2000.1

EURACHEM/CITAC Guide*

Quantifying Uncertainty in Analytical Measurement

Second Edition

Editors

S L R Ellison (LGC, UK)
M Rosslein (EMPA, Switzerland)
A Williams (UK)

Composition of the Working Group

EURACHEM members

A Williams	Chairman	UK
S Ellison	Secretary	LGC, Teddington, UK
M Berglund		Institute for Reference Materials and Measurements, Belgium
W Haesselbarth		Bundesanstalt für Materialforschung und -prüfung, Germany
K Hedegaard		EUROM II
R Kaarls		Netherlands Measurement Institute, The Netherlands
M Månsson		SP Swedish National Testing and Research Institute, Sweden
M Rösslein		EMPA St. Gallen, Switzerland
R Stephany		National Institute of Public Health and the Environment, The Netherlands
A van der Veen		Netherlands Measurement Institute, The Netherlands
W Wegscheider		University of Mining and Metallurgy, Leoben, Austria
H van de Wiel		National Institute of Public Health and the Environment, The Netherlands
R Wood		Food Standards Agency, UK

CITAC Members

Pan Xiu Rong	Director, NRCCRM., China
M Salit	National Institute of Science and Technology USA
A Squirrel	NATA, Australia
K Yasuda	Hitachi Ltd, Japan

AOAC Representatives

R Johnson	Agricultural Analytical Services, Texas State Chemist, USA
Jung-Keun Lee	U.S. F.D.A. Washington
D Mowrey	Eli Lilly & Co., Greenfield, USA

IAEA Representatives

P De Regge	IAEA Vienna
A Fajgelj	IAEA Vienna

EA Representative

D Galsworthy,	UKAS, UK
---------------	----------

Acknowledgements

This document has been produced primarily by a joint EURACHEM/CITAC Working Group with the composition shown (right). The editors are grateful to all these individuals and organisations and to others who have contributed comments, advice and assistance.

Production of this Guide was in part supported under contract with the UK Department of Trade and Industry as part of the National Measurement System Valid Analytical Measurement (VAM) Programme.

*CITAC Reference

This Guide constitutes CITAC Guide number 4

Quantifying Uncertainty in Analytical Measurement
English edition

Second edition 2000
ISBN 0 948926 15 5

Copyright © 2000
Copyright in this document is vested in the
contributing members of the working group.

CONTENTS

FOREWORD TO THE SECOND EDITION	1
1. SCOPE AND FIELD OF APPLICATION	3
2. UNCERTAINTY	4
2.1. DEFINITION OF UNCERTAINTY	4
2.2. UNCERTAINTY SOURCES	4
2.3. UNCERTAINTY COMPONENTS	4
2.4. ERROR AND UNCERTAINTY	5
3. ANALYTICAL MEASUREMENT AND UNCERTAINTY	7
3.1. METHOD VALIDATION	7
3.2. CONDUCT OF EXPERIMENTAL STUDIES OF METHOD PERFORMANCE	8
3.3. TRACEABILITY	9
4. THE PROCESS OF MEASUREMENT UNCERTAINTY ESTIMATION	11
5. STEP 1. SPECIFICATION OF THE MEASURAND	13
6. STEP 2. IDENTIFYING UNCERTAINTY SOURCES	14
7. STEP 3. QUANTIFYING UNCERTAINTY	16
7.1. INTRODUCTION	16
7.2. UNCERTAINTY EVALUATION PROCEDURE	16
7.3. RELEVANCE OF PRIOR STUDIES	17
7.4. EVALUATING UNCERTAINTY BY QUANTIFICATION OF INDIVIDUAL COMPONENTS	17
7.5. CLOSELY MATCHED CERTIFIED REFERENCE MATERIALS	17
7.6. USING PRIOR COLLABORATIVE METHOD DEVELOPMENT AND VALIDATION STUDY DATA	17
7.7. USING IN-HOUSE DEVELOPMENT AND VALIDATION STUDIES	18
7.8. EVALUATION OF UNCERTAINTY FOR EMPIRICAL METHODS	20
7.9. EVALUATION OF UNCERTAINTY FOR AD-HOC METHODS	20
7.10. QUANTIFICATION OF INDIVIDUAL COMPONENTS	21
7.11. EXPERIMENTAL ESTIMATION OF INDIVIDUAL UNCERTAINTY CONTRIBUTIONS	21
7.12. ESTIMATION BASED ON OTHER RESULTS OR DATA	22
7.13. MODELLING FROM THEORETICAL PRINCIPLES	22
7.14. ESTIMATION BASED ON JUDGEMENT	22
7.15. SIGNIFICANCE OF BIAS	24
8. STEP 4. CALCULATING THE COMBINED UNCERTAINTY	25
8.1. STANDARD UNCERTAINTIES	25
8.2. COMBINED STANDARD UNCERTAINTY	25
8.3. EXPANDED UNCERTAINTY	27

Quantifying Uncertainty	Contents
9. REPORTING UNCERTAINTY	29
9.1. GENERAL	29
9.2. INFORMATION REQUIRED	29
9.3. REPORTING STANDARD UNCERTAINTY	29
9.4. REPORTING EXPANDED UNCERTAINTY	29
9.5. NUMERICAL EXPRESSION OF RESULTS	30
9.6. COMPLIANCE AGAINST LIMITS	30
APPENDIX A. EXAMPLES	32
INTRODUCTION	32
EXAMPLE A1: PREPARATION OF A CALIBRATION STANDARD	34
EXAMPLE A2: STANDARDISING A SODIUM HYDROXIDE SOLUTION	40
EXAMPLE A3: AN ACID/BASE TITRATION	50
EXAMPLE A4: UNCERTAINTY ESTIMATION FROM IN-HOUSE VALIDATION STUDIES. DETERMINATION OF ORGANOPHOSPHORUS PESTICIDES IN BREAD.	59
EXAMPLE A5: DETERMINATION OF CADMIUM RELEASE FROM CERAMIC WARE BY ATOMIC ABSORPTION SPECTROMETRY	70
EXAMPLE A6: THE DETERMINATION OF CRUDE FIBRE IN ANIMAL FEEDING STUFFS	79
EXAMPLE A7: DETERMINATION OF THE AMOUNT OF LEAD IN WATER USING DOUBLE ISOTOPE DILUTION AND INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY	87
APPENDIX B. DEFINITIONS	95
APPENDIX C. UNCERTAINTIES IN ANALYTICAL PROCESSES	99
APPENDIX D. ANALYSING UNCERTAINTY SOURCES	100
D.1 INTRODUCTION	100
D.2 PRINCIPLES OF APPROACH	100
D.3 CAUSE AND EFFECT ANALYSIS	100
D.4 EXAMPLE	101
APPENDIX E. USEFUL STATISTICAL PROCEDURES	102
E.1 DISTRIBUTION FUNCTIONS	102
E.2 SPREADSHEET METHOD FOR UNCERTAINTY CALCULATION	104
E.3 UNCERTAINTIES FROM LINEAR LEAST SQUARES CALIBRATION	106
E.4: DOCUMENTING UNCERTAINTY DEPENDENT ON ANALYTE LEVEL	108
APPENDIX F. MEASUREMENT UNCERTAINTY AT THE LIMIT OF DETECTION/ LIMIT OF DETERMINATION	112
F1. INTRODUCTION	112
F2. OBSERVATIONS AND ESTIMATES	112
F3. INTERPRETED RESULTS AND COMPLIANCE STATEMENTS	113
APPENDIX G. COMMON SOURCES AND VALUES OF UNCERTAINTY	114
APPENDIX H. BIBLIOGRAPHY	120

Foreword to the Second Edition

Many important decisions are based on the results of chemical quantitative analysis; the results are used, for example, to estimate yields, to check materials against specifications or statutory limits, or to estimate monetary value. Whenever decisions are based on analytical results, it is important to have some indication of the quality of the results, that is, the extent to which they can be relied on for the purpose in hand. Users of the results of chemical analysis, particularly in those areas concerned with international trade, are coming under increasing pressure to eliminate the replication of effort frequently expended in obtaining them. Confidence in data obtained outside the user's own organisation is a prerequisite to meeting this objective. In some sectors of analytical chemistry it is now a formal (frequently legislative) requirement for laboratories to introduce quality assurance measures to ensure that they are capable of and are providing data of the required quality. Such measures include: the use of validated methods of analysis; the use of defined internal quality control procedures; participation in proficiency testing schemes; accreditation based on ISO 17025 [H.1], and establishing traceability of the results of the measurements

In analytical chemistry, there has been great emphasis on the precision of results obtained using a specified method, rather than on their traceability to a defined standard or SI unit. This has led the use of "official methods" to fulfil legislative and trading requirements. However as there is now a formal requirement to establish the confidence of results it is essential that a measurement result is traceable to a defined reference such as a SI unit, reference material or, where applicable, a defined or empirical (sec. 5.2.) method. Internal quality control procedures, proficiency testing and accreditation can be an aid in establishing evidence of traceability to a given standard.

As a consequence of these requirements, chemists are, for their part, coming under increasing pressure to demonstrate the quality of their results, and in particular to demonstrate their fitness for purpose by giving a measure of the confidence that can be placed on the result. This is expected to include the degree to which a result would be expected to agree with other results, normally irrespective of the analytical methods used. One useful measure of this is measurement uncertainty.

Although the concept of measurement uncertainty has been recognised by chemists for many years, it was the publication in 1993 of the "Guide to the Expression of Uncertainty in Measurement" [H.2] by ISO in collaboration with BIPM, IEC, IFCC, IUPAC, IUPAP and OIML, which formally established general rules for evaluating and expressing uncertainty in measurement across a broad spectrum of measurements. This EURACHEM document shows how the concepts in the ISO Guide may be applied in chemical measurement. It first introduces the concept of uncertainty and the distinction between uncertainty and error. This is followed by a description of the steps involved in the evaluation of uncertainty with the processes illustrated by worked examples in Appendix A.

The evaluation of uncertainty requires the analyst to look closely at all the possible sources of uncertainty. However, although a detailed study of this kind may require a considerable effort, it is essential that the effort expended should not be disproportionate. In practice a preliminary study will quickly identify the most significant sources of uncertainty and, as the examples show, the value obtained for the combined uncertainty is almost entirely controlled by the major contributions. A good estimate of uncertainty can be made by concentrating effort on the largest contributions. Further, once evaluated for a given method applied in a particular laboratory (i.e. a particular measurement procedure), the uncertainty estimate obtained may be reliably applied to subsequent results obtained by the method in the same laboratory, provided that this is justified by the relevant quality control data. No further effort should be necessary unless the procedure itself or the equipment used is changed, in which case the uncertainty estimate would be reviewed as part of the normal re-validation.

The first edition of the EURACHEM Guide for "Quantifying Uncertainty in Analytical Measurement" [H.3] was published in 1995 based on the ISO Guide.

This second edition of the EURACHEM Guide has been prepared in the light of practical experience of uncertainty estimation in chemistry laboratories and the even greater awareness of the need to introduce formal quality assurance procedures by laboratories. The second edition stresses that the procedures introduced by a laboratory to estimate its measurement uncertainty should be integrated with existing quality assurance measures, since these measures frequently provide much of the information required to evaluate the measurement uncertainty. The guide therefore provides explicitly for the use of validation and related data in the construction of uncertainty estimates in full compliance with formal ISO Guide principles. The approach is also consistent with the requirements of ISO 17025:1999 [H.1]

NOTE Worked examples are given in Appendix A. A numbered list of definitions is given at Appendix B. The convention is adopted of printing defined terms in bold face upon their first occurrence in the text, with a reference to Appendix B enclosed in square brackets. The definitions are, in the main, taken from the International vocabulary of basic and general standard terms in Metrology (VIM) [H.4], the Guide [H.2] and ISO 3534 (Statistics - Vocabulary and symbols) [H.5]. Appendix C shows, in general terms, the overall structure of a chemical analysis leading to a measurement result. Appendix D describes a general procedure which can be used to identify uncertainty components and plan further experiments as required; Appendix E describes some statistical operations used in uncertainty estimation in analytical chemistry. Appendix F discusses measurement uncertainty near detection limits. Appendix G lists many common uncertainty sources and methods of estimating the value of the uncertainties. A bibliography is provided at Appendix H.

1. Scope and Field of Application

1.1. This Guide gives detailed guidance for the evaluation and expression of uncertainty in quantitative chemical analysis, based on the approach taken in the ISO “Guide to the Expression of Uncertainty in Measurement” [H.2]. It is applicable at all levels of accuracy and in all fields - from routine analysis to basic research and to empirical and rational methods (see section 5.3.). Some common areas in which chemical measurements are needed, and in which the principles of this Guide may be applied, are:

- Quality control and quality assurance in manufacturing industries.
- Testing for regulatory compliance.
- Testing utilising an agreed method.
- Calibration of standards and equipment.
- Measurements associated with the development and certification of reference materials.
- Research and development.

1.2. Note that additional guidance will be required in some cases. In particular, reference material value assignment using consensus methods (including multiple measurement methods) is not covered, and the use of uncertainty estimates in compliance statements and the expression and use of uncertainty at low levels may require additional guidance. Uncertainties associated with sampling operations are not explicitly treated.

1.3. Since formal quality assurance measures have been introduced by laboratories in a number of sectors this second EURACHEM Guide is now able to illustrate how data from the following procedures may be used for the estimation of measurement uncertainty:

- Evaluation of the effect of the identified sources of uncertainty on the analytical result for a single method implemented as a defined **measurement procedure [B.8]** in a single laboratory .
- Results from defined internal quality control procedures in a single laboratory.
- Results from collaborative trials used to validate methods of analysis in a number of competent laboratories.
- Results from proficiency test schemes used to assess the analytical competency of laboratories.

1.4. It is assumed throughout this Guide that, whether carrying out measurements or assessing the performance of the measurement procedure, effective quality assurance and control measures are in place to ensure that the measurement process is stable and in control. Such measures normally include, for example, appropriately qualified staff, proper maintenance and calibration of equipment and reagents, use of appropriate reference standards, documented measurement procedures and use of appropriate check standards and control charts. Reference [H.6] provides further information on analytical QA procedures.

NOTE: This paragraph implies that all analytical methods are assumed in this guide to be implemented via fully documented procedures. Any general reference to analytical methods accordingly implies the presence of such a procedure. Strictly, measurement uncertainty can only be applied to the results of such a procedure and not to a more general **method of measurement [B.9]**.

2. Uncertainty

2.1. Definition of uncertainty

2.1.1. The definition of the term uncertainty (of measurement) used in this protocol and taken from the current version adopted for the International Vocabulary of Basic and General Terms in Metrology [H.4] is:

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

Note 1 The parameter may be, for example, a **standard deviation** [B.23] (or a given multiple of it), or the width of a confidence interval.

NOTE 2 Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterised by standard deviations. The other components, which also can be characterised by standard deviations, are evaluated from assumed probability distributions based on experience or other information. The ISO Guide refers to these different cases as Type A and Type B estimations respectively.

2.1.2. In many cases in chemical analysis, the **measurand** [B.6] will be the concentration* of an analyte. However chemical analysis is used to measure other quantities, *e.g.* colour, pH, *etc.*, and therefore the general term "measurand" will be used.

2.1.3. The definition of uncertainty given above focuses on the range of values that the analyst believes could reasonably be attributed to the measurand.

2.1.4. In general use, the word *uncertainty* relates to the general concept of *doubt*. In this guide, the

word *uncertainty*, without adjectives, refers either to a parameter associated with the definition above, or to the limited knowledge about a particular value. *Uncertainty of measurement* does not imply doubt about the validity of a measurement; on the contrary, knowledge of the uncertainty implies increased confidence in the validity of a measurement result.

2.2. Uncertainty sources

2.2.1. In practice the uncertainty on the result may arise from many possible sources, including examples such as incomplete definition, sampling, matrix effects and interferences, environmental conditions, uncertainties of masses and volumetric equipment, reference values, approximations and assumptions incorporated in the measurement method and procedure, and random variation (a fuller description of uncertainty sources is given in section 6.7.)

2.3. Uncertainty components

2.3.1. In estimating the overall uncertainty, it may be necessary to take each source of uncertainty and treat it separately to obtain the contribution from that source. Each of the separate contributions to uncertainty is referred to as an uncertainty component. When expressed as a standard deviation, an uncertainty component is known as a **standard uncertainty** [B.13]. If there is correlation between any components then this has to be taken into account by determining the covariance. However, it is often possible to evaluate the combined effect of several components. This may reduce the overall effort involved and, where components whose contribution is evaluated together are correlated, there may be no additional need to take account of the correlation.

2.3.2. For a measurement result y , the total uncertainty, termed **combined standard uncertainty** [B.14] and denoted by $u_c(y)$, is an estimated standard deviation equal to the positive square root of the total variance obtained by combining all the uncertainty components, however evaluated, using the law of propagation of uncertainty (see section 8.).

* In this guide, the unqualified term "concentration" applies to any of the particular quantities *mass* concentration, *amount* concentration, *number* concentration or *volume* concentration unless units are quoted (*e.g.* a concentration quoted in mg l^{-1} is evidently a mass concentration). Note also that many other quantities used to express composition, such as mass fraction, substance content and mole fraction, can be directly related to concentration.

2.3.3. For most purposes in analytical chemistry, an **expanded uncertainty [B.15]** U , should be used. The expanded uncertainty provides an interval within which the value of the measurand is believed to lie with a higher level of confidence. U is obtained by multiplying $u_c(y)$, the combined standard uncertainty, by a **coverage factor [B.16]** k . The choice of the factor k is based on the level of confidence desired. For an approximate level of confidence of 95%, k is 2.

NOTE The coverage factor k should always be stated so that the combined standard uncertainty of the measured quantity can be recovered for use in calculating the combined standard uncertainty of other measurement results that may depend on that quantity.

2.4. Error and uncertainty

2.4.1. It is important to distinguish between error and uncertainty. **Error [B.19]** is defined as the difference between an individual result and the **true value [B.3]** of the measurand. As such, error is a single value. In principle, the value of a known error can be applied as a correction to the result.

NOTE Error is an idealised concept and errors cannot be known exactly.

2.4.2. Uncertainty, on the other hand, takes the form of a range, and, if estimated for an analytical procedure and defined sample type, may apply to all determinations so described. In general, the value of the uncertainty cannot be used to correct a measurement result.

2.4.3. To illustrate further the difference, the result of an analysis after correction may by chance be very close to the value of the measurand, and hence have a negligible error. However, the uncertainty may still be very large, simply because the analyst is very unsure of how close that result is to the value.

2.4.4. The uncertainty of the result of a measurement should never be interpreted as representing the error itself, nor the error remaining after correction.

2.4.5. An error is regarded as having two components, namely, a random component and a systematic component.

2.4.6. Random error [B.20] typically arises from unpredictable variations of influence quantities. These random effects give rise to variations in repeated observations of the measurand. The

random error of an analytical result cannot be compensated for, but it can usually be reduced by increasing the number of observations.

NOTE 1 The experimental standard deviation of the **arithmetic mean [B.22]** or average of a series of observations is *not* the random error of the mean, although it is so referred to in some publications on uncertainty. It is instead a measure of the uncertainty of the mean due to some random effects. The exact value of the random error in the mean arising from these effects cannot be known.

2.4.7. Systematic error [B.21] is defined as a component of error which, in the course of a number of analyses of the same measurand, remains constant or varies in a predictable way. It is independent of the number of measurements made and cannot therefore be reduced by increasing the number of analyses under constant measurement conditions.

2.4.8. Constant systematic errors, such as failing to make an allowance for a reagent blank in an assay, or inaccuracies in a multi-point instrument calibration, are constant for a given level of the measurement value but may vary with the level of the measurement value.

2.4.9. Effects which change systematically in magnitude during a series of analyses, caused, for example by inadequate control of experimental conditions, give rise to systematic errors that are not constant.

EXAMPLES:

1. A gradual increase in the temperature of a set of samples during a chemical analysis can lead to progressive changes in the result.
2. Sensors and probes that exhibit ageing effects over the time-scale of an experiment can also introduce non-constant systematic errors.

2.4.10. The result of a measurement should be corrected for all recognised significant systematic effects.

NOTE Measuring instruments and systems are often adjusted or calibrated using measurement standards and reference materials to correct for systematic effects. The uncertainties associated with these standards and materials and the uncertainty in the correction must still be taken into account.

2.4.11. A further type of error is a spurious error, or blunder. Errors of this type invalidate a measurement and typically arise through human failure or instrument malfunction. Transposing

digits in a number while recording data, an air bubble lodged in a spectrophotometer flow-through cell, or accidental cross-contamination of test items are common examples of this type of error.

2.4.12. Measurements for which errors such as these have been detected should be rejected and no attempt should be made to incorporate the errors into any statistical analysis. However, errors such as digit transposition can be corrected (exactly), particularly if they occur in the leading digits.

2.4.13. Spurious errors are not always obvious and, where a sufficient number of replicate measurements is available, it is usually appropriate to apply an outlier test to check for the presence of suspect members in the data set. Any positive result obtained from such a test should be considered with care and, where possible, referred back to the originator for confirmation. It is generally not wise to reject a value on purely statistical grounds.

2.4.14. Uncertainties estimated using this guide are not intended to allow for the possibility of spurious errors/blunders.

3. Analytical Measurement and Uncertainty

3.1. Method validation

3.1.1. In practice, the fitness for purpose of analytical methods applied for routine testing is most commonly assessed through method validation studies [H.7]. Such studies produce data on overall performance and on individual influence factors which can be applied to the estimation of uncertainty associated with the results of the method in normal use.

3.1.2. Method validation studies rely on the determination of overall method performance parameters. These are obtained during method development and interlaboratory study or following in-house validation protocols. Individual sources of error or uncertainty are typically investigated only when significant compared to the overall precision measures in use. The emphasis is primarily on identifying and removing (rather than correcting for) significant effects. This leads to a situation in which the majority of potentially significant influence factors have been identified, checked for significance compared to overall precision, and shown to be negligible. Under these circumstances, the data available to analysts consists primarily of overall performance figures, together with evidence of insignificance of most effects and some measurements of any remaining significant effects.

3.1.3. Validation studies for quantitative analytical methods typically determine some or all of the following parameters:

Precision. The principal precision measures include repeatability standard deviation s_r , reproducibility standard deviation s_R , (ISO 3534-1) and intermediate precision, sometimes denoted s_{zi} , with i denoting the number of factors varied (ISO 5725-3:1994). The repeatability s_r indicates the variability observed within a laboratory, over a short time, using a single operator, item of equipment *etc.* s_r may be estimated within a laboratory or by inter-laboratory study. Interlaboratory reproducibility standard deviation s_R for a particular method may only be estimated directly by interlaboratory study; it shows the variability obtained when different laboratories analyse the same sample. Intermediate precision relates to the variation in results observed when

one or more factors, such as time, equipment and operator, are varied within a laboratory; different figures are obtained depending on which factors are held constant. Intermediate precision estimates are most commonly determined within laboratories but may also be determined by interlaboratory study. The observed precision of an analytical procedure is an essential component of overall uncertainty, whether determined by combination of individual variances or by study of the complete method in operation.

Bias. The bias of an analytical method is usually determined by study of relevant reference materials or by spiking studies. The determination of overall bias with respect to appropriate reference values is important in establishing **traceability [B.12]** to recognised standards (see section 3.2). Bias may be expressed as analytical recovery (value observed divided by value expected). Bias should be shown to be negligible or corrected for, but in either case the uncertainty associated with the determination of the bias remains an essential component of overall uncertainty.

Linearity. Linearity is an important property of methods used to make measurements at a range of concentrations. The linearity of the response to pure standards and to realistic samples may be determined. Linearity is not generally quantified, but is checked for by inspection or using significance tests for non-linearity. Significant non-linearity is usually corrected for by use of non-linear calibration functions or eliminated by choice of more restricted operating range. Any remaining deviations from linearity are normally sufficiently accounted for by overall precision estimates covering several concentrations, or within any uncertainties associated with calibration (Appendix E.3).

Detection limit. During method validation, the detection limit is normally determined only to establish the lower end of the practical operating range of a method. Though uncertainties near the detection limit may require careful consideration and special treatment (Appendix F), the detection limit, however determined, is not of direct relevance to uncertainty estimation.

Robustness or ruggedness. Many method development or validation protocols require that sensitivity to particular parameters be investigated directly. This is usually done by a preliminary ‘ruggedness test’, in which the effect of one or more parameter changes is observed. If significant (compared to the precision of the ruggedness test) a more detailed study is carried out to measure the size of the effect, and a permitted operating interval chosen accordingly. Ruggedness test data can therefore provide information on the effect of important parameters.

Selectivity/specificity. Though loosely defined, both terms relate to the degree to which a method responds uniquely to the required analyte. Typical selectivity studies investigate the effects of likely interferences, usually by adding the potential interferent to both blank and fortified samples and observing the response. The results are normally used to demonstrate that the practical effects are not significant. However, since the studies measure changes in response directly, it is possible to use the data to estimate the uncertainty associated with potential interferences, given knowledge of the range of interferent concentrations.

3.2. Conduct of experimental studies of method performance

3.2.1. The detailed design and execution of method validation and method performance studies is covered extensively elsewhere [H.7] and will not be repeated here. However, the main principles as they affect the relevance of a study applied to uncertainty estimation are pertinent and are considered below.

3.2.2. *Representativeness* is essential. That is, studies should, as far as possible, be conducted to provide a realistic survey of the number and range of effects operating during normal use of the method, as well as covering the concentration ranges and sample types within the scope of the method. Where a factor has been representatively varied during the course of a precision experiment, for example, the effects of that factor appear directly in the observed variance and need no additional study unless further method optimisation is desirable.

3.2.3. In this context, *representative variation* means that an influence parameter must take a distribution of values appropriate to the uncertainty in the parameter in question. For continuous parameters, this may be a permitted range or stated uncertainty; for discontinuous

factors such as sample matrix, this range corresponds to the variety of types permitted or encountered in normal use of the method. Note that representativeness extends not only to the range of values, but to their distribution.

3.2.4. In selecting factors for variation, it is important to ensure that the larger effects are varied where possible. For example, where day to day variation (perhaps arising from recalibration effects) is substantial compared to repeatability, two determinations on each of five days will provide a better estimate of intermediate precision than five determinations on each of two days. Ten single determinations on separate days will be better still, subject to sufficient control, though this will provide no additional information on within-day repeatability.

3.2.5. It is generally simpler to treat data obtained from random selection than from systematic variation. For example, experiments performed at random times over a sufficient period will usually include representative ambient temperature effects, while experiments performed systematically at 24-hour intervals may be subject to bias due to regular ambient temperature variation during the working day. The former experiment needs only evaluate the overall standard deviation; in the latter, systematic variation of ambient temperature is required, followed by adjustment to allow for the actual distribution of temperatures. Random variation is, however, less efficient. A small number of systematic studies can quickly establish the size of an effect, whereas it will typically take well over 30 determinations to establish an uncertainty contribution to better than about 20% relative accuracy. Where possible, therefore, it is often preferable to investigate small numbers of major effects systematically.

3.2.6. Where factors are known or suspected to interact, it is important to ensure that the effect of interaction is accounted for. This may be achieved either by ensuring random selection from different levels of interacting parameters, or by careful systematic design to obtain both variance and covariance information.

3.2.7. In carrying out studies of overall bias, it is important that the reference materials and values are relevant to the materials under routine test.

3.2.8. Any study undertaken to investigate and test for the significance of an effect should have sufficient power to detect such effects before they become practically significant.

3.3. Traceability

3.3.1. It is important to be able to compare results from different laboratories, or from the same laboratory at different times, with confidence. This is achieved by ensuring that all laboratories are using the same measurement scale, or the same ‘reference points’. In many cases this is achieved by establishing a chain of calibrations leading to primary national or international standards, ideally (for long-term consistency) the Systeme Internationale (SI) units of measurement. A familiar example is the case of analytical balances; each balance is calibrated using reference masses which are themselves checked (ultimately) against national standards and so on to the primary reference kilogram. This unbroken chain of comparisons leading to a known reference value provides ‘traceability’ to a common reference point, ensuring that different operators are using the same units of measurement. In routine measurement, the consistency of measurements between one laboratory (or time) and another is greatly aided by establishing traceability for all relevant intermediate measurements used to obtain or control a measurement result. Traceability is therefore an important concept in all branches of measurement.

3.3.2. Traceability is formally defined [H.4] as:

“The property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties.”

The reference to uncertainty arises because the agreement between laboratories is limited, in part, by uncertainties incurred in each laboratory’s traceability chain. Traceability is accordingly intimately linked to uncertainty. Traceability provides the means of placing all related measurements on a consistent measurement scale, while uncertainty characterises the ‘strength’ of the links in the chain and the agreement to be expected between laboratories making similar measurements.

3.3.3. In general, the uncertainty on a result which is traceable to a particular reference, will be the uncertainty on that reference together with the uncertainty on making the measurement relative to that reference.

3.3.4. Traceability of the result of the complete analytical procedure should be established by a combination of the following procedures:

1. Use of traceable standards to calibrate the measuring equipment
2. By using, or by comparison to the results of, a primary method
3. By using a pure substance RM.
4. By using an appropriate matrix Certified Reference Material (CRM)
5. By using an accepted, closely defined procedure.

Each procedure is discussed in turn below.

3.3.5. Calibration of measuring equipment

In all cases, the calibration of the measuring equipment used must be traceable to appropriate standards. The quantification stage of the analytical procedure is often calibrated using a pure substance reference material, whose value is traceable to the SI. This practice provides traceability of the results to SI for this part of the procedure. However, it is also necessary to establish traceability for the results of operations prior to the quantification stage, such as extraction and sample clean up, using additional procedures.

3.3.6. Measurements using Primary Methods

A primary method is currently described as follows:

“A primary method of measurement is a method having the highest metrological qualities, whose operation is completely described and understood in terms of SI units and whose results are accepted without reference to a standard of the same quantity.”

The result of a primary method is normally traceable directly to the SI, and is of the smallest achievable uncertainty with respect to this reference. Primary methods are normally implemented only by National Measurement Institutes and are rarely applied to routine testing or calibration. Where applicable, traceability to the results of a primary method is achieved by direct comparison of measurement results between the primary method and test or calibration method.

3.3.7. Measurements using a pure substance Reference Material (RM).

Traceability can be demonstrated by measurement of a sample composed of, or

containing, a known quantity of a pure substance RM. This may be achieved, for example, by spiking or by standard additions. However, it is always necessary to evaluate the difference in response of the measurement system to the standard used and the sample under test. Unfortunately, for many chemical analyses and in the particular case of spiking or standard additions, both the correction for the difference in response and its uncertainty may be large. Thus, although the traceability of the result to SI units can in principle be established, in practice, in all but the most simple cases, the uncertainty on the result may be unacceptably large or even unquantifiable. If the uncertainty is unquantifiable then traceability has not been established

3.3.8. Measurement on a Certified Reference Material (CRM)

Traceability may be demonstrated through comparison of measurement results on a certified matrix CRM with the certified value(s). This procedure can reduce the uncertainty compared to the use of a pure substance RM where there is a suitable matrix CRM available. If the value of the CRM is traceable to SI, then these measurements provide traceability to SI units and the evaluation of the uncertainty utilising

reference materials is discussed in 7.5. However, even in this case, the uncertainty on the result may be unacceptably large or even unquantifiable, particularly if there is not a good match between the composition of the sample and the reference material.

3.3.9. Measurement using an accepted procedure.

Adequate comparability can often only be achieved through use of a closely defined and generally accepted procedure. The procedure will normally be defined in terms of input parameters; for example a specified set of extraction times, particle sizes *etc.* The results of applying such a procedure are considered traceable when the values of these input parameters are traceable to stated references in the usual way. The uncertainty on the results arises both from uncertainties in the specified input parameters and from the effects of incomplete specification and variability in execution (see section 7.8.1.). Where the results of an alternative method or procedure are expected to be comparable to the results of such an accepted procedure, traceability to the accepted values is achieved by comparing the results obtained by accepted and alternative procedures.

4. The Process of Measurement Uncertainty Estimation

4.1. Uncertainty estimation is simple in principle. The following paragraphs summarise the tasks that need to be performed in order to obtain an estimate of the uncertainty associated with a measurement result. Subsequent chapters provide additional guidance applicable in different circumstances, particularly relating to the use of data from method validation studies and the use of formal uncertainty propagation principles. The steps involved are:

Step 1. Specify measurand

Write down a clear statement of what is being measured, including the relationship between the measurand and the input quantities (*e.g.* measured quantities, constants, calibration standard values *etc.*) upon which it depends. Where possible, include corrections for known systematic effects. The specification information should be given in the relevant Standard Operating Procedure (SOP) or other method description.

Step 2. Identify uncertainty sources

List the possible sources of uncertainty. This will include sources that contribute to the uncertainty on the parameters in the relationship specified in Step 1, but may include other sources and must include sources arising from chemical assumptions. A general procedure for forming a structured list is suggested at Appendix D.

Step 3. Quantify uncertainty components

Measure or estimate the size of the uncertainty component associated with each potential source of uncertainty identified. It is often possible to estimate or determine a single contribution to uncertainty associated with a number of separate sources. It is also important to consider whether available data accounts sufficiently for all sources of uncertainty, and plan additional experiments and studies carefully to ensure that all sources of uncertainty are adequately accounted for.

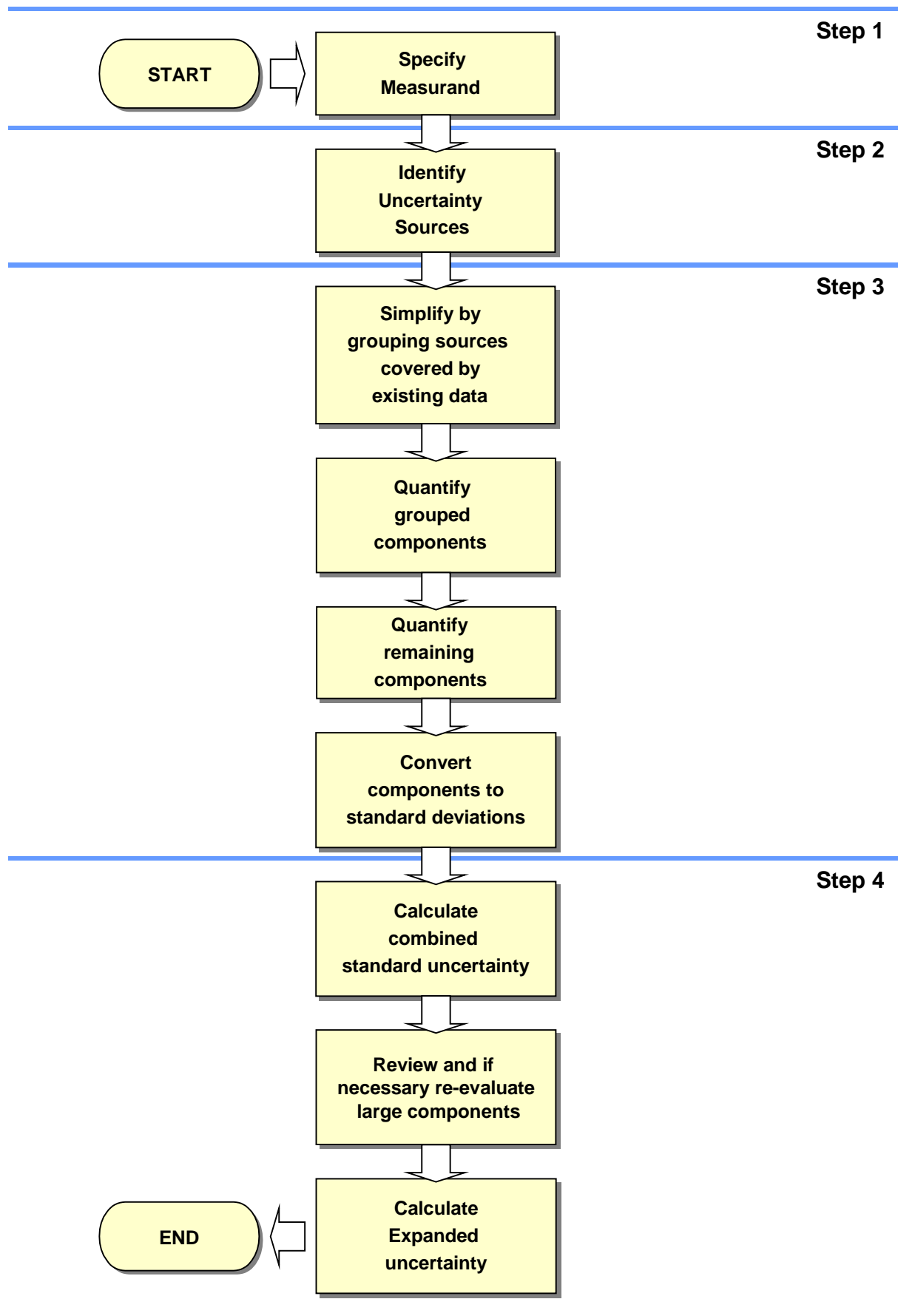
Step 4. Calculate combined uncertainty

The information obtained in step 3 will consist of a number of quantified contributions to overall uncertainty, whether associated with individual sources or with the combined effects of several sources. The contributions have to be expressed as standard deviations, and combined according to the appropriate rules, to give a combined standard uncertainty. The appropriate coverage factor should be applied to give an expanded uncertainty.

Figure 1 shows the process schematically.

4.2. The following chapters provide guidance on the execution of all the steps listed above and shows how the procedure may be simplified depending on the information that is available about the combined effect of a number of sources.

Figure 1: The Uncertainty Estimation Process



5. Step 1. Specification of the Measurand

5.1. In the context of uncertainty estimation, “specification of the measurand” requires both a clear and unambiguous statement of what is being measured, and a quantitative expression relating the value of the measurand to the parameters on which it depends. These parameters may be other measurands, quantities which are not directly measured, or constants. It should also be clear whether a sampling step is included within the procedure or not. If it is, estimation of uncertainties associated with the sampling procedure need to be considered. All of this information should be in the Standard Operating Procedure (SOP).

5.2. In analytical measurement, it is particularly important to distinguish between measurements intended to produce results which are independent of the method used, and those which are not so intended. The latter are often referred to as *empirical methods*. The following examples may clarify the point further.

EXAMPLES:

1. Methods for the determination of the amount of nickel present in an alloy are normally expected to yield the same result, in the same units, usually expressed as a mass or mole fraction. In principle, any systematic effect due to method bias or matrix would need to be corrected for, though it is more usual to ensure that any such effect is small. Results would not normally need to quote the particular method used, except for information. The method is not empirical.

2. Determinations of “extractable fat” may differ substantially, depending on the extraction

conditions specified. Since “extractable fat” is entirely dependent on choice of conditions, the method used is *empirical*. It is not meaningful to consider correction for bias intrinsic to the method, since the measurand is defined by the method used. Results are generally reported with reference to the method, uncorrected for any bias intrinsic to the method. The method is considered empirical.

3. In circumstances where variations in the substrate, or matrix, have large and unpredictable effects, a procedure is often developed with the sole aim of achieving comparability between laboratories measuring the same material. The procedure may then be adopted as a local, national or international standard method on which trading or other decisions are taken, with no intent to obtain an absolute measure of the true amount of analyte present. Corrections for method bias or matrix effect are ignored by convention (whether or not they have been minimised in method development). Results are normally reported uncorrected for matrix or method bias. The method is considered to be empirical.

5.3. The distinction between empirical and non-empirical (sometimes called *rational*) methods is important because it affects the estimation of uncertainty. In examples 2 and 3 above, because of the conventions employed, uncertainties associated with some quite large effects are not relevant in normal use. Due consideration should accordingly be given to whether the results are expected to be dependent upon, or independent of, the method in use and only those effects relevant to the result as reported should be included in the uncertainty estimate.

6. Step 2. Identifying Uncertainty Sources

6.1. A comprehensive list of relevant sources of uncertainty should be assembled. At this stage, it is not necessary to be concerned about the quantification of individual components; the aim is to be completely clear about what should be considered. In Step 3, the best way of treating each source will be considered.

6.2. In forming the required list of uncertainty sources it is usually convenient to start with the basic expression used to calculate the measurand from intermediate values. All the parameters in this expression may have an uncertainty associated with their value and are therefore potential uncertainty sources. In addition there may be other parameters that do not appear explicitly in the expression used to calculate the value of the measurand, but which nevertheless affect the measurement results, e.g. extraction time or temperature. These are also potential sources of uncertainty. All these different sources should be included. Additional information is given in Appendix C (Uncertainties in Analytical Processes).

6.3. The cause and effect diagram described in Appendix D is a very convenient way of listing the uncertainty sources, showing how they relate to each other and indicating their influence on the uncertainty of the result. It also helps to avoid double counting of sources. Although the list of uncertainty sources can be prepared in other ways, the cause and effect diagram is used in the following chapters and in all of the examples in Appendix A. Additional information is given in Appendix D (Analysing uncertainty sources).

6.4. Once the list of uncertainty sources is assembled, their effects on the result can, in principle, be represented by a formal measurement model, in which each effect is associated with a parameter or variable in an equation. The equation then forms a complete model of the measurement process in terms of all the individual factors affecting the result. This function may be very complicated and it may not be possible to write it down explicitly. Where possible, however, this should be done, as the form of the expression will generally determine the method of combining individual uncertainty contributions.

6.5. It may additionally be useful to consider a measurement procedure as a series of discrete operations (sometimes termed *unit operations*), each of which may be assessed separately to obtain estimates of uncertainty associated with them. This is a particularly useful approach where similar measurement procedures share common unit operations. The separate uncertainties for each operation then form contributions to the overall uncertainty.

6.6. In practice, it is more usual in analytical measurement to consider uncertainties associated with elements of overall method performance, such as observable precision and bias measured with respect to appropriate reference materials. These contributions generally form the dominant contributions to the uncertainty estimate, and are best modelled as separate effects on the result. It is then necessary to evaluate other possible contributions only to check their significance, quantifying only those that are significant. Further guidance on this approach, which applies particularly to the use of method validation data, is given in section 7.2.1.

6.7. Typical sources of uncertainty are

- Sampling

Where in-house or field sampling form part of the specified procedure, effects such as random variations between different samples and any potential for bias in the sampling procedure form components of uncertainty affecting the final result.

- Storage Conditions

Where test items are stored for any period prior to analysis, the storage conditions may affect the results. The duration of storage as well as conditions during storage should therefore be considered as uncertainty sources.

- Instrument effects

Instrument effects may include, for example, the limits of accuracy on the calibration of an analytical balance; a temperature controller that may maintain a mean temperature which differs (within specification) from its

Quantifying Uncertainty

indicated set-point; an auto-analyser that could be subject to carry-over effects.

- Reagent purity

The concentration of a volumetric solution will not be known exactly even if the parent material has been assayed, since some uncertainty related to the assaying procedure remains. Many organic dyestuffs, for instance, are not 100% pure and can contain isomers and inorganic salts. The purity of such substances is usually stated by manufacturers as being *not less than* a specified level. Any assumptions about the degree of purity will introduce an element of uncertainty.

- Assumed stoichiometry

Where an analytical process is assumed to follow a particular reaction stoichiometry, it may be necessary to allow for departures from the expected stoichiometry, or for incomplete reaction or side reactions.

- Measurement conditions

For example, volumetric glassware may be used at an ambient temperature different from that at which it was calibrated. Gross temperature effects should be corrected for, but any uncertainty in the temperature of liquid and glass should be considered. Similarly, humidity may be important where materials are sensitive to possible changes in humidity.

- Sample effects

The recovery of an analyte from a complex matrix, or an instrument response, may be affected by composition of the matrix. Analyte speciation may further compound this effect.

Step 2. Identifying Uncertainty Sources

The stability of a sample/analyte may change during analysis because of a changing thermal regime or photolytic effect.

When a 'spike' is used to estimate recovery, the recovery of the analyte from the sample may differ from the recovery of the spike, introducing an uncertainty which needs to be evaluated.

- Computational effects

Selection of the calibration model, *e.g.* using a straight line calibration on a curved response, leads to poorer fit and higher uncertainty.

Truncation and round off can lead to inaccuracies in the final result. Since these are rarely predictable, an uncertainty allowance may be necessary.

- Blank Correction

There will be an uncertainty on both the value and the appropriateness of the blank correction. This is particularly important in trace analysis.

- Operator effects

Possibility of reading a meter or scale consistently high or low.

Possibility of making a slightly different interpretation of the method.

- Random effects

Random effects contribute to the uncertainty in all determinations. This entry should be included in the list as a matter of course.

NOTE: These sources are not necessarily independent.

7. Step 3. Quantifying Uncertainty

7.1. Introduction

7.1.1. Having identified the uncertainty sources as explained in Step 2 (Chapter 6), the next step is to quantify the uncertainty arising from these sources. This can be done by

- evaluating the uncertainty arising from each individual source and then combining them as described in Chapter 8. Examples A1 to A3 illustrate the use of this procedure.

or

- by determining directly the combined contribution to the uncertainty on the result from some or all of these sources using method performance data. Examples A4 to A6 represent applications of this procedure.

In practice, a combination of these is usually necessary and convenient.

7.1.2. Whichever of these approaches is used, most of the information needed to evaluate the uncertainty is likely to be already available from the results of validation studies, from QA/QC data and from other experimental work that has been carried out to check the performance of the method. However, data may not be available to evaluate the uncertainty from all of the sources and it may be necessary to carry out further work as described in sections 7.10. to 7.14.

7.2. Uncertainty evaluation procedure

7.2.1. The procedure used for estimating the overall uncertainty depends on the data available about the method performance. The stages involved in developing the procedure are

- **Reconcile the information requirements with the available data**

First, the list of uncertainty sources should be examined to see which sources of uncertainty are accounted for by the available data, whether by explicit study of the particular contribution or by implicit variation within the course of whole-method experiments. These sources should be checked against the list prepared in Step 2 and any remaining sources should be listed to provide an auditable record of which contributions to the uncertainty have been included.

- **Plan to obtain the further data required**

For sources of uncertainty not adequately covered by existing data, either seek additional information from the literature or standing data (certificates, equipment specifications *etc.*), or plan experiments to obtain the required additional data. Additional experiments may take the form of specific studies of a single contribution to uncertainty, or the usual method performance studies conducted to ensure representative variation of important factors.

7.2.2. It is important to recognise that not all of the components will make a significant contribution to the combined uncertainty; indeed, in practice it is likely that only a small number will. Unless there is a large number of them, components that are less than one third of the largest need not be evaluated in detail. A preliminary estimate of the contribution of each component or combination of components to the uncertainty should be made and those that are not significant eliminated.

7.2.3. The following sections provide guidance on the procedures to be adopted, depending on the data available and on the additional information required. Section 7.3. presents requirements for the use of prior experimental study data, including validation data. Section 7.4. briefly discusses evaluation of uncertainty solely from individual sources of uncertainty. This may be necessary for all, or for very few of the sources identified, depending on the data available, and is consequently also considered in later sections. Sections 7.5. to 7.9. describe the evaluation of uncertainty in a range of circumstances. Section 7.5. applies when using closely matched reference materials. Section 7.6. covers the use of collaborative study data and 7.7. the use of in-house validation data. 7.8. describes special considerations for empirical methods and 7.9. covers ad-hoc methods. Methods for quantifying individual components of uncertainty, including experimental studies, documentary and other data, modelling, and professional judgement are covered in more detail in sections 7.10. to 7.14. Section 7.15. covers the treatment of known bias in uncertainty estimation.

7.3. Relevance of prior studies

7.3.1. When uncertainty estimates are based at least partly on prior studies of method performance, it is necessary to demonstrate the validity of applying prior study results. Typically, this will consist of:

- Demonstration that a comparable precision to that obtained previously can be achieved.
- Demonstration that the use of the bias data obtained previously is justified, typically through determination of bias on relevant reference materials (see, for example, ISO Guide 33 [H.8]), by appropriate spiking studies, or by satisfactory performance on relevant proficiency schemes or other laboratory intercomparisons.
- Continued performance within statistical control as shown by regular QC sample results and the implementation of effective analytical quality assurance procedures.

7.3.2. Where the conditions above are met, and the method is operated within its scope and field of application, it is normally acceptable to apply the data from prior studies (including validation studies) directly to uncertainty estimates in the laboratory in question.

7.4. Evaluating uncertainty by quantification of individual components

7.4.1. In some cases, particularly when little or no method performance data is available, the most suitable procedure may be to evaluate each uncertainty component separately.

7.4.2. The general procedure used in combining individual components is to prepare a detailed quantitative model of the experimental procedure (cf. sections 5. and 6., especially 6.4.), assess the standard uncertainties associated with the individual input parameters, and combine them using the law of propagation of uncertainties as described in Section 8.

7.4.3. In the interests of clarity, detailed guidance on the assessment of individual contributions by experimental and other means is deferred to sections 7.10. to 7.14. Examples A1 to A3 in Appendix A provide detailed illustrations of the procedure. Extensive guidance on the application of this procedure is also given in the ISO *Guide* [H.2].

7.5. Closely matched certified reference materials

- **7.5.1.** Measurements on certified reference materials are normally carried out as part of method validation or re-validation, effectively constituting a calibration of the whole measurement procedure against a traceable reference. Because this procedure provides information on the combined effect of many of the potential sources of uncertainty, it provides very good data for the assessment of uncertainty. Further details are given in section 7.7.4.

NOTE: ISO Guide 33 [H.8] gives a useful account of the use of reference materials in checking method performance.

7.6. Uncertainty estimation using prior collaborative method development and validation study data

7.6.1. A collaborative study carried out to validate a published method, for example according to the AOAC/IUPAC protocol [H.9] or ISO 5725 standard [H.10], is a valuable source of data to support an uncertainty estimate. The data typically include estimates of reproducibility standard deviation, s_R , for several levels of response, a linear estimate of the dependence of s_R on level of response, and may include an estimate of bias based on CRM studies. How this data can be utilised depends on the factors taken into account when the study was carried out. During the ‘reconciliation’ stage indicated above (section 7.2.), it is necessary to identify any sources of uncertainty that are not covered by the collaborative study data. The sources which may need particular consideration are:

- **Sampling.** Collaborative studies rarely include a sampling step. If the method used in-house involves sub-sampling, or the measurand (see Specification) is estimating a bulk property from a small sample, then the effects of sampling should be investigated and their effects included.
- **Pre-treatment.** In most studies, samples are homogenised, and may additionally be stabilised, before distribution. It may be necessary to investigate and add the effects of the particular pre-treatment procedures applied in-house.
- **Method bias.** Method bias is often examined prior to or during interlaboratory study, where possible by comparison with reference

methods or materials. Where the bias itself, the uncertainty in the reference values used, and the precision associated with the bias check, are all small compared to s_R , no additional allowance need be made for bias uncertainty. Otherwise, it will be necessary to make additional allowances.

- **Variation in conditions.** Laboratories participating in a study may tend towards the means of allowed ranges of experimental conditions, resulting in an underestimate of the range of results possible within the method definition. Where such effects have been investigated and shown to be insignificant across their full permitted range, however, no further allowance is required.
- **Changes in sample matrix.** The uncertainty arising from matrix compositions or levels of interferences outside the range covered by the study will need to be considered.

7.6.2. Each significant source of uncertainty not covered by the collaborative study data should be evaluated in the form of a standard uncertainty and combined with the reproducibility standard deviation s_R in the usual way (section 8.)

7.6.3. For methods operating within their defined scope, when the reconciliation stage shows that all the identified sources have been included in the validation study or when the contributions from any remaining sources such as those discussed in section 7.6.1. have been shown to be negligible, then the reproducibility standard deviation s_R , adjusted for concentration if necessary, may be used as the combined standard uncertainty.

7.6.4. The use of this procedure is shown in example A6 (Appendix A)

7.7. Uncertainty estimation using in-house development and validation studies

7.7.1. In-house development and validation studies consist chiefly of the determination of the method performance parameters indicated in section 3.1.3. Uncertainty estimation from these parameters utilises:

- The best available estimate of overall precision.
- The best available estimate(s) of overall bias and its uncertainty.

- Quantification of any uncertainties associated with effects incompletely accounted for in the above overall performance studies.

Precision study

7.7.2. The precision should be estimated as far as possible over an extended time period, and chosen to allow natural variation of all factors affecting the result. This can be obtained from

- The standard deviation of results for a typical sample analysed several times over a period of time, using different analysts and equipment where possible (the results of measurements on QC check samples can provide this information).
- The standard deviation obtained from replicate analyses performed on each of several samples.

NOTE: Replicates should be performed at materially different times to obtain estimates of intermediate precision; within-batch replication provides estimates of repeatability only.

- From formal multi-factor experimental designs, analysed by ANOVA to provide separate variance estimates for each factor.

7.7.3. Note that precision frequently varies significantly with the level of response. For example, the observed standard deviation often increases significantly and systematically with analyte concentration. In such cases, the uncertainty estimate should be adjusted to allow for the precision applicable to the particular result. Appendix E.4 gives additional guidance on handling level-dependent contributions to uncertainty.

Bias study

7.7.4. Overall bias is best estimated by repeated analysis of a relevant CRM, using the complete measurement procedure. Where this is done, and the bias found to be insignificant, the uncertainty associated with the bias is simply the combination of the standard uncertainty on the CRM value with the standard deviation associated with the bias.

NOTE: Bias estimated in this way combines bias in laboratory performance with any bias intrinsic to the method in use. Special considerations may apply where the method in use is empirical; see section 7.8.1.

- When the reference material is only approximately representative of the test

materials, additional factors should be considered, including (as appropriate) differences in composition and homogeneity; reference materials are frequently more homogeneous than test samples. Estimates based on professional judgement should be used, if necessary, to assign these uncertainties (see section 7.14.).

- Any effects following from different concentrations of analyte; for example, it is not uncommon to find that extraction losses differ between high and low levels of analyte.

7.7.5. Bias for a method under study can also be determined by comparison of the results with those of a reference method. If the results show that the bias is not statistically significant, the standard uncertainty is that for the reference method (if applicable; see section 7.8.1.), combined with the standard uncertainty associated with the measured difference between methods. The latter contribution to uncertainty is given by the standard deviation term used in the significance test applied to decide whether the difference is statistically significant, as explained in the example below.

EXAMPLE

A method (method 1) for determining the concentration of Selenium is compared with a reference method (method 2). The results (in mg kg⁻¹) from each method are as follows:

	\bar{x}	s	n
Method 1	5.40	1.47	5
Method 2	4.76	2.75	5

The standard deviations are pooled to give a pooled standard deviation s_c

$$s_c = \sqrt{\frac{1.47^2 \times (5-1) + 2.75^2 \times (5-1)}{5+5-2}} = 2.205$$

and a corresponding value of t :

$$t = \frac{(5.40 - 4.76)}{2.205 \sqrt{\left(\frac{1}{5} + \frac{1}{5}\right)}} = \frac{0.64}{1.4} = 0.46$$

t_{crit} is 2.3 for 8 degrees of freedom, so there is no significant difference between the means of the results given by the two methods. However, the difference (0.64) is compared with a standard deviation term of 1.4 above. This value of 1.4 is the standard deviation associated with the difference, and accordingly represents the relevant contribution to uncertainty associated with the measured bias.

7.7.6. Overall bias can also be estimated by the addition of analyte to a previously studied material. The same considerations apply as for the study of reference materials (above). In addition, the differential behaviour of added material and material native to the sample should be considered and due allowance made. Such an allowance can be made on the basis of:

- Studies of the distribution of the bias observed for a range of matrices and levels of added analyte.
- Comparison of result observed in a reference material with the recovery of added analyte in the same reference material.
- Judgement on the basis of specific materials with known extreme behaviour. For example, oyster tissue, a common marine tissue reference, is well known for a tendency to co-precipitate some elements with calcium salts on digestion, and may provide an estimate of 'worst case' recovery on which an uncertainty estimate can be based (e.g. By treating the worst case as an extreme of a rectangular or triangular distribution).
- Judgement on the basis of prior experience.

7.7.7. Bias may also be estimated by comparison of the particular method with a value determined by the method of standard additions, in which known quantities of the analyte are added to the test material, and the correct analyte concentration inferred by extrapolation. The uncertainty associated with the bias is then normally dominated by the uncertainties associated with the extrapolation, combined (where appropriate) with any significant contributions from the preparation and addition of stock solution.

NOTE: To be directly relevant, the additions should be made to the original sample, rather than a prepared extract.

7.7.8. It is a general requirement of the ISO *Guide* that corrections should be applied for all recognised and significant systematic effects. Where a correction is applied to allow for a significant overall bias, the uncertainty associated with the bias is estimated as paragraph 7.7.5. described in the case of insignificant bias

7.7.9. Where the bias is significant, but is nonetheless neglected for practical purposes, additional action is necessary (see section 7.15.).

Additional factors

7.7.10. The effects of any remaining factors should be estimated separately, either by experimental variation or by prediction from established theory. The uncertainty associated with such factors should be estimated, recorded and combined with other contributions in the normal way.

7.7.11. Where the effect of these remaining factors is demonstrated to be negligible compared to the precision of the study (i.e. statistically insignificant), it is recommended that an uncertainty contribution equal to the standard deviation associated with the relevant significance test be associated with that factor.

EXAMPLE

The effect of a permitted 1-hour extraction time variation is investigated by a *t*-test on five determinations each on the same sample, for the normal extraction time and a time reduced by 1 hour. The means and standard deviations (in mg l⁻¹) were: Standard time: mean 1.8, standard deviation 0.21; alternate time: mean 1.7, standard deviation 0.17. A *t*-test uses the pooled variance of

$$\frac{(5-1) \times 0.21^2 + (5-1) \times 0.17^2}{(5-1) + (5-1)} = 0.037$$

to obtain

$$t = \frac{(1.8 - 1.7)}{\sqrt{0.037 \times \left(\frac{1}{5} + \frac{1}{5} \right)}} = 0.82$$

This is not significant compared to $t_{\text{crit}} = 2.3$. But note that the difference (0.1) is compared with a calculated standard deviation term of $\sqrt{0.037 \times (1/5 + 1/5)} = 0.12$. This value is the contribution to uncertainty associated with the effect of permitted variation in extraction time.

7.7.12. Where an effect is detected and is statistically significant, but remains sufficiently small to neglect in practice, the provisions of section 7.15. apply.

7.8. Evaluation of uncertainty for empirical methods

7.8.1. An 'empirical method' is a method agreed upon for the purposes of comparative measurement within a particular field of application where the measurand characteristically depends upon the method in use. The method accordingly defines the measurand. Examples include methods for

leachable metals in ceramics and dietary fibre in foodstuffs (see also section 5.2. and example A5)

7.8.2. Where such a method is in use within its defined field of application, the bias associated with the method is defined as zero. In such circumstances, bias estimation need relate only to the laboratory performance and should not additionally account for bias intrinsic to the method. This has the following implications.

7.8.3. Reference material investigations, whether to demonstrate negligible bias or to measure bias, should be conducted using reference materials certified using the particular method, or for which a value obtained with the particular method is available for comparison.

7.8.4. Where reference materials so characterised are unavailable, overall control of bias is associated with the control of method parameters affecting the result; typically such factors as times, temperatures, masses, volumes *etc.* The uncertainty associated with these input factors must accordingly be assessed and either shown to be negligible or quantified (see example A6).

7.8.5. Empirical methods are normally subjected to collaborative studies and hence the uncertainty can be evaluated as described in section 7.6.

7.9. Evaluation of uncertainty for ad-hoc methods

7.9.1. Ad-hoc methods are methods established to carry out exploratory studies in the short term, or for a short run of test materials. Such methods are typically based on standard or well-established methods within the laboratory, but are adapted substantially (for example to study a different analyte) and will not generally justify formal validation studies for the particular material in question.

7.9.2. Since limited effort will be available to establish the relevant uncertainty contributions, it is necessary to rely largely on the known performance of related systems or blocks within these systems. Uncertainty estimation should accordingly be based on known performance on a related system or systems. This performance information should be supported by any study necessary to establish the relevance of the information. The following recommendations assume that such a related system is available and has been examined sufficiently to obtain a reliable uncertainty estimate, or that the method consists of blocks from other methods and that the uncertainty in these blocks has been

established previously.

7.9.3. As a minimum, it is essential that an estimate of overall bias and an indication of precision be available for the method in question. Bias will ideally be measured against a reference material, but will in practice more commonly be assessed from spike recovery. The considerations of section 7.7.4. then apply, except that spike recoveries should be compared with those observed on the related system to establish the relevance of the prior studies to the ad-hoc method in question. The overall bias observed for the ad-hoc method, on the materials under test, should be comparable to that observed for the related system, within the requirements of the study.

7.9.4. A minimum precision experiment consists of a duplicate analysis. It is, however, recommended that as many replicates as practical are performed. The precision should be compared with that for the related system; the standard deviation for the ad-hoc method should be comparable.

NOTE: It recommended that the comparison be based on inspection. Statistical significance tests (e.g. an F-test) will generally be unreliable with small numbers of replicates and will tend to lead to the conclusion that there is 'no significant difference' simply because of the low power of the test.

7.9.5. Where the above conditions are met unequivocally, the uncertainty estimate for the related system may be applied directly to results obtained by the ad-hoc method, making any adjustments appropriate for concentration dependence and other known factors.

7.10. Quantification of individual components

7.10.1. It is nearly always necessary to consider some sources of uncertainty individually. In some cases, this is only necessary for a small number of sources; in others, particularly when little or no method performance data is available, every source may need separate study (see examples 1,2 and 3 in Appendix A for illustrations). There are several general methods for establishing individual uncertainty components:

- Experimental variation of input variables
- From standing data such as measurement and calibration certificates
- By modelling from theoretical principles

- Using judgement based on experience or informed by modelling of assumptions

These different methods are discussed briefly below.

7.11. Experimental estimation of individual uncertainty contributions

7.11.1. It is often possible and practical to obtain estimates of uncertainty contributions from experimental studies specific to individual parameters.

7.11.2. The standard uncertainty arising from random effects is often measured from repeatability experiments and is quantified in terms of the standard deviation of the measured values. In practice, no more than about fifteen replicates need normally be considered, unless a high precision is required.

7.11.3. Other typical experiments include:

- Study of the effect of a variation of a single parameter on the result. This is particularly appropriate in the case of continuous, controllable parameters, independent of other effects, such as time or temperature. The rate of change of the result with the change in the parameter can be obtained from the experimental data. This is then combined directly with the uncertainty in the parameter to obtain the relevant uncertainty contribution.

NOTE: The change in parameter should be sufficient to change the result substantially compared to the precision available in the study (e.g. by five times the standard deviation of replicate measurements)

- Robustness studies, systematically examining the significance of moderate changes in parameters. This is particularly appropriate for rapid identification of significant effects, and commonly used for method optimisation. The method can be applied in the case of discrete effects, such as change of matrix, or small equipment configuration changes, which have unpredictable effects on the result. Where a factor is found to be significant, it is normally necessary to investigate further. Where insignificant, the associated uncertainty is (at least for initial estimation) that obtained from the robustness study.
- Systematic multifactor experimental designs intended to estimate factor effects and interactions. Such studies are particularly

useful where a categorical variable is involved. A categorical variable is one in which the value of the variable is unrelated to the size of the effect; laboratory numbers in a study, analyst names, or sample types are examples of categorical variables. For example, the effect of changes in matrix type (within a stated method scope) could be estimated from recovery studies carried out in a replicated multiple-matrix study. An analysis of variance would then provide within- and between-matrix components of variance for observed analytical recovery. The between-matrix component of variance would provide a standard uncertainty associated with matrix variation.

7.12. Estimation based on other results or data

7.12.1. It is often possible to estimate some of the standard uncertainties using whatever relevant information is available about the uncertainty on the quantity concerned. The following paragraphs suggest some sources of information.

7.12.2. Proficiency testing (PT) schemes. A laboratory's results from participation in PT schemes can be used as a check on the evaluated uncertainty, since the uncertainty should be compatible with the spread of results obtained by that laboratory over a number of proficiency test rounds. Further, in the special case where

- the compositions of samples used in the scheme cover the full range analysed routinely
- the assigned values in each round are traceable to appropriate reference values, and
- the uncertainty on the assigned value is small compared to the observed spread of results

then the dispersion of the differences between the reported values and the assigned values obtained in repeated rounds provides a basis for a good estimate of the uncertainty arising from those parts of the measurement procedure within the scope of the scheme. For example, for a scheme operating with similar materials and analyte levels, the standard deviation of differences would give the standard uncertainty. Of course, systematic deviation from traceable assigned values and any other sources of uncertainty (such as those noted in section 7.6.1.) must also be taken into account.

7.12.3. Quality Assurance (QA) data. As noted previously it is necessary to ensure that the quality criteria set out in standard operating procedures are achieved, and that measurements on QA samples show that the criteria continue to be met. Where reference materials are used in QA checks, section 7.5. shows how the data can be used to evaluate uncertainty. Where any other stable material is used, the QA data provides an estimate of intermediate precision (Section 7.7.2.). QA data also forms a continuing check on the value quoted for the uncertainty. Clearly, the combined uncertainty arising from random effects cannot be less than the standard deviation of the QA measurements.

7.12.4. Suppliers' information. For many sources of uncertainty, calibration certificates or suppliers catalogues provide information. For example, the tolerance of volumetric glassware may be obtained from the manufacturer's catalogue or a calibration certificate relating to a particular item in advance of its use.

7.13. Modelling from theoretical principles

7.13.1. In many cases, well-established physical theory provides good models for effects on the result. For example, temperature effects on volumes and densities are well understood. In such cases, uncertainties can be calculated or estimated from the form of the relationship using the uncertainty propagation methods described in section 8.

7.13.2. In other circumstances, it may be necessary to use approximate theoretical models combined with experimental data. For example, where an analytical measurement depends on a timed derivatisation reaction, it may be necessary to assess uncertainties associated with timing. This might be done by simple variation of elapsed time. However, it may be better to establish an approximate rate model from brief experimental studies of the derivatisation kinetics near the concentrations of interest, and assess the uncertainty from the predicted rate of change at a given time.

7.14. Estimation based on judgement

7.14.1. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement method and procedure used. The quality and utility of the

uncertainty quoted for the result of a measurement therefore ultimately depends on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.

7.14.2. Most distributions of data can be interpreted in the sense that it is less likely to observe data in the margins of the distribution than in the centre. The quantification of these distributions and their associated standard deviations is done through repeated measurements.

7.14.3. However, other assessments of intervals may be required in cases when repeated measurements cannot be performed or do not provide a meaningful measure of a particular uncertainty component.

7.14.4. There are numerous instances in analytical chemistry when the latter prevails, and judgement is required. For example:

- An assessment of recovery and its associated uncertainty cannot be made for every single sample. Instead, an assessment is made for classes of samples (*e.g.* grouped by type of matrix), and the results applied to all samples of similar type. The degree of similarity is itself an unknown, thus this inference (from type of matrix to a specific sample) is associated with an extra element of uncertainty that has no frequentistic interpretation.
- The model of the measurement as defined by the specification of the analytical procedure is used for converting the measured quantity to the value of the measurand (analytical result). This model is - like all models in science - subject to uncertainty. It is only assumed that nature behaves according to the specific model, but this can never be known with ultimate certainty.
- The use of reference materials is highly encouraged, but there remains uncertainty regarding not only the true value, but also regarding the relevance of a particular reference material for the analysis of a specific sample. A judgement is required of the extent to which a proclaimed standard substance reasonably resembles the nature of the samples in a particular situation.
- Another source of uncertainty arises when the measurand is insufficiently defined by the procedure. Consider the determination of

"permanganate oxidizable substances" that are undoubtedly different whether one analyses ground water or municipal waste water. Not only factors such as oxidation temperature, but also chemical effects such as matrix composition or interference, may have an influence on this specification.

- A common practice in analytical chemistry calls for spiking with a single substance, such as a close structural analogue or isotopomer, from which either the recovery of the respective native substance or even that of a whole class of compounds is judged. Clearly, the associated uncertainty is experimentally assessable provided the analyst is prepared to study the recovery at all concentration levels and ratios of measurands to the spike, and all "relevant" matrices. But frequently this experimentation is avoided and substituted by judgements on
 - the concentration dependence of recoveries of measurand,
 - the concentration dependence of recoveries of spike,
 - the dependence of recoveries on (sub)type of matrix,
 - the identity of binding modes of native and spiked substances.

7.14.5. Judgement of this type is not based on immediate experimental results, but rather on a subjective (personal) probability, an expression which here can be used synonymously with "degree of belief", "intuitive probability" and "credibility" [H.11]. It is also assumed that a degree of belief is not based on a snap judgement, but on a well considered mature judgement of probability.

7.14.6. Although it is recognised that subjective probabilities vary from one person to another, and even from time to time for a single person, they are not arbitrary as they are influenced by common sense, expert knowledge, and by earlier experiments and observations.

7.14.7. This may appear to be a disadvantage, but need not lead in practice to worse estimates than those from repeated measurements. This applies particularly if the true, real-life, variability in experimental conditions cannot be simulated and the resulting variability in data thus does not give a realistic picture.

7.14.8. A typical problem of this nature arises if long-term variability needs to be assessed when

no collaborative study data are available. A scientist who dismisses the option of substituting subjective probability for an actually measured one (when the latter is not available) is likely to ignore important contributions to combined uncertainty, thus being ultimately less objective than one who relies on subjective probabilities.

7.14.9. For the purpose of estimation of combined uncertainties two features of degree of belief estimations are essential:

- degree of belief is regarded as interval valued which is to say that a lower and an upper bound similar to a classical probability distribution is provided,
- the same computational rules apply in combining 'degree of belief' contributions of uncertainty to a combined uncertainty as for standard deviations derived by other methods.

7.15. Significance of bias

7.15.1. It is a general requirement of the *ISO Guide* that corrections should be applied for all recognised and significant systematic effects.

7.15.2. In deciding whether a known bias can reasonably be neglected, the following approach is recommended:

- i) Estimate the combined uncertainty without considering the relevant bias.
- ii) Compare the bias with the combined uncertainty.
- iii) Where the bias is not significant compared to the combined uncertainty, the bias may be neglected.
- iv) Where the bias is significant compared to the combined uncertainty, additional action is required. Appropriate actions might:
 - Eliminate or correct for the bias, making due allowance for the uncertainty of the correction.
 - Report the observed bias and its uncertainty in addition to the result.

NOTE: Where a known bias is uncorrected by convention, the method should be considered empirical (see section 7.8).

8. Step 4. Calculating the Combined Uncertainty

8.1. Standard uncertainties

8.1.1. Before combination, all uncertainty contributions must be expressed as standard uncertainties, that is, as standard deviations. This may involve conversion from some other measure of dispersion. The following rules give some guidance for converting an uncertainty component to a standard deviation.

8.1.2. Where the uncertainty component was evaluated experimentally from the dispersion of repeated measurements, it can readily be expressed as a standard deviation. For the contribution to uncertainty in single measurements, the standard uncertainty is simply the observed standard deviation; for results subjected to averaging, the **standard deviation of the mean [B.24]** is used.

8.1.3. Where an uncertainty estimate is derived from previous results and data, it may already be expressed as a standard deviation. However where a confidence interval is given with a level of confidence, (in the form $\pm a$ at $p\%$) then divide the value a by the appropriate percentage point of the Normal distribution for the level of confidence given to calculate the standard deviation.

EXAMPLE

A specification states that a balance reading is within ± 0.2 mg with 95% confidence. From standard tables of percentage points on the normal distribution, a 95% confidence interval is calculated using a value of 1.96σ . Using this figure gives a standard uncertainty of $(0.2/1.96) \approx 0.1$.

8.1.4. If limits of $\pm a$ are given without a confidence level and there is reason to expect that extreme values are likely, it is normally appropriate to assume a rectangular distribution, with a standard deviation of $a/\sqrt{3}$ (see Appendix E).

EXAMPLE

A 10 ml Grade A volumetric flask is certified to within ± 0.2 ml. The standard uncertainty is $0.2/\sqrt{3} \approx 0.12$ ml.

8.1.5. If limits of $\pm a$ are given without a confidence level, but there is reason to expect that extreme values are unlikely, it is normally

appropriate to assume a triangular distribution, with a standard deviation of $a/\sqrt{6}$ (see Appendix E).

EXAMPLE

A 10 ml Grade A volumetric flask is certified to within ± 0.2 ml, but routine in-house checks show that extreme values are rare. The standard uncertainty is $0.2/\sqrt{6} \approx 0.08$ ml.

8.1.6. Where an estimate is to be made on the basis of judgement, it may be possible to estimate the component directly as a standard deviation. If this is not possible then an estimate should be made of the maximum deviation which could reasonably occur in practice (excluding simple mistakes). If a smaller value is considered substantially more likely, this estimate should be treated as descriptive of a triangular distribution. If there are no grounds for believing that a small error is more likely than a large error, the estimate should be treated as characterising a rectangular distribution.

8.1.7. Conversion factors for the most commonly used distribution functions are given in Appendix E.1.

8.2. Combined standard uncertainty

8.2.1. Following the estimation of individual or groups of components of uncertainty and expressing them as standard uncertainties, the next stage is to calculate the combined standard uncertainty using one of the procedures described below.

8.2.2. The general relationship between the combined standard uncertainty $u_c(y)$ of a value y and the uncertainty of the independent parameters x_1, x_2, \dots, x_n on which it depends is

$$u_c(y(x_1, x_2, \dots)) = \sqrt{\sum_{i=1, n} c_i^2 u(x_i)^2} = \sqrt{\sum_{i=1, n} u(y, x_i)^2}^*$$

where $y(x_1, x_2, \dots)$ is a function of several parameters x_1, x_2, \dots , c_i is a sensitivity coefficient evaluated as $c_i = \partial y / \partial x_i$, the partial differential of y with respect to x_i and $u(y, x_i)$ denotes the uncertainty in y arising from the uncertainty in x_i . Each variable's contribution $u(y, x_i)$ is just the

* The ISO *Guide* uses the shorter form $u_i(y)$ instead of $u(y, x_i)$

square of the associated uncertainty expressed as a standard deviation multiplied by the square of the relevant sensitivity coefficient. These sensitivity coefficients describe how the value of y varies with changes in the parameters x_1, x_2 etc.

NOTE: Sensitivity coefficients may also be evaluated directly by experiment; this is particularly valuable where no reliable mathematical description of the relationship exists.

8.2.3. Where variables are not independent, the relationship is more complex:

$$u(y(x_{i,j}, \dots)) = \sqrt{\sum_{i=1,n} c_i^2 u(x_i)^2 + \sum_{\substack{i,k=1,n \\ i \neq k}} c_i c_k \cdot u(x_i, x_k)}$$

where $u(x_i, x_k)$ is the covariance between x_i and x_k and c_i and c_k are the sensitivity coefficients as described and evaluated in 8.2.2. The covariance is related to the correlation coefficient r_{ik} by

$$u(x_i, x_k) = u(x_i) \cdot u(x_k) \cdot r_{ik}$$

where $-1 \leq r_{ik} \leq 1$.

8.2.4. These general procedures apply whether the uncertainties are related to single parameters, grouped parameters or to the method as a whole. However, when an uncertainty contribution is associated with the whole procedure, it is usually expressed as an effect on the final result. In such cases, or when the uncertainty on a parameter is expressed directly in terms of its effect on y , the sensitivity coefficient $\partial y / \partial x_i$ is equal to 1.0.

EXAMPLE

A result of 22 mg l⁻¹ shows a measured standard deviation of 4.1 mg l⁻¹. The standard uncertainty $u(y)$ associated with precision under these conditions is 4.1 mg l⁻¹. The implicit model for the measurement, neglecting other factors for clarity, is

$$y = (\text{Calculated result}) + \varepsilon$$

where ε represents the effect of random variation under the conditions of measurement. $\partial y / \partial \varepsilon$ is accordingly 1.0

8.2.5. Except for the case above, when the sensitivity coefficient is equal to one, and for the special cases given in Rule 1 and Rule 2 below, the general procedure, requiring the generation of partial differentials or the numerical equivalent must be employed. Appendix E gives details of a numerical method, suggested by Kragten [H.12], which makes effective use of spreadsheet software to provide a combined standard uncertainty from input standard uncertainties and a known measurement model. It is recommended

that this method, or another appropriate computer-based method, be used for all but the simplest cases.

8.2.6. In some cases, the expressions for combining uncertainties reduce to much simpler forms. Two simple rules for combining standard uncertainties are given here.

Rule 1

For models involving only a sum or difference of quantities, e.g. $y = (p + q + r + \dots)$, the combined standard uncertainty $u_c(y)$ is given by

$$u_c(y(p, q, \dots)) = \sqrt{u(p)^2 + u(q)^2 + \dots}$$

Rule 2

For models involving only a product or quotient, e.g. $y = (p \times q \times r \times \dots)$ or $y = p / (q \times r \times \dots)$, the combined standard uncertainty $u_c(y)$ is given by

$$u_c(y) = y \sqrt{\left(\frac{u(p)}{p}\right)^2 + \left(\frac{u(q)}{q}\right)^2 + \dots}$$

where $(u(p)/p)$ etc. are the uncertainties in the parameters, expressed as relative standard deviations.

NOTE Subtraction is treated in the same manner as addition, and division in the same way as multiplication.

8.2.7. For the purposes of combining uncertainty components, it is most convenient to break the original mathematical model down to expressions which consist solely of operations covered by one of the rules above. For example, the expression

$$(o + p) / (q + r)$$

should be broken down to the two elements $(o+p)$ and $(q+r)$. The interim uncertainties for each of these can then be calculated using rule 1 above; these interim uncertainties can then be combined using rule 2 to give the combined standard uncertainty.

8.2.8. The following examples illustrate the use of the above rules:

EXAMPLE 1

$y = (p - q + r)$ The values are $p=5.02$, $q=6.45$ and $r=9.04$ with standard uncertainties $u(p)=0.13$, $u(q)=0.05$ and $u(r)=0.22$.

$$y = 5.02 - 6.45 + 9.04 = 7.61$$

$$u(y) = \sqrt{0.13^2 + 0.05^2 + 0.22^2} = 0.26$$

EXAMPLE 2

$y = (op/qr)$. The values are $o=2.46$, $p=4.32$, $q=6.38$ and $r=2.99$, with standard uncertainties of $u(o)=0.02$, $u(p)=0.13$, $u(q)=0.11$ and $u(r)=0.07$.

$$y = (2.46 \times 4.32) / (6.38 \times 2.99) = 0.56$$

$$u(y) = 0.56 \times \sqrt{\left(\frac{0.02}{2.46}\right)^2 + \left(\frac{0.13}{4.32}\right)^2 + \left(\frac{0.11}{6.38}\right)^2 + \left(\frac{0.07}{2.99}\right)^2}$$

$$\Rightarrow u(y) = 0.56 \times 0.043 = 0.024$$

8.2.9. There are many instances in which the magnitudes of components of uncertainty vary with the level of analyte. For example, uncertainties in recovery may be smaller for high levels of material, or spectroscopic signals may vary randomly on a scale approximately proportional to intensity (constant coefficient of variation). In such cases, it is important to take account of the changes in the combined standard uncertainty with level of analyte. Approaches include:

- Restricting the specified procedure or uncertainty estimate to a small range of analyte concentrations.
- Providing an uncertainty estimate in the form of a relative standard deviation.
- Explicitly calculating the dependence and recalculating the uncertainty for a given result.

Appendix E4 gives additional information on these approaches.

8.3. Expanded uncertainty

8.3.1. The final stage is to multiply the combined standard uncertainty by the chosen coverage factor in order to obtain an expanded uncertainty. The expanded uncertainty is required to provide an interval which may be expected to encompass a large fraction of the distribution of values which could reasonably be attributed to the measurand.

8.3.2. In choosing a value for the coverage factor k , a number of issues should be considered. These include:

- The level of confidence required
- Any knowledge of the underlying distributions

- Any knowledge of the number of values used to estimate random effects (see 8.3.3 below).

8.3.3. For most purposes it is recommended that k is set to 2. However, this value of k may be insufficient where the combined uncertainty is based on statistical observations with relatively few degrees of freedom (less than about six). The choice of k then depends on the effective number of degrees of freedom.

8.3.4. Where the combined standard uncertainty is dominated by a single contribution with fewer than six degrees of freedom, it is recommended that k be set equal to the two-tailed value of Student's t for the number of degrees of freedom associated with that contribution, and for the level of confidence required (normally 95%). Table 1 (page 28) gives a short list of values for t .

EXAMPLE:

A combined standard uncertainty for a weighing operation is formed from contributions $u_{cal}=0.01$ mg arising from calibration uncertainty and $s_{obs}=0.08$ mg based on the standard deviation of five repeated observations. The combined standard uncertainty u_c is equal to $\sqrt{0.01^2 + 0.08^2} = 0.081$ mg. This is clearly dominated by the repeatability contribution s_{obs} , which is based on five observations, giving $5-1=4$ degrees of freedom. k is accordingly based on Student's t . The two-tailed value of t for four degrees of freedom and 95% confidence is, from tables, 2.8; k is accordingly set to 2.8 and the expanded uncertainty $U=2.8 \times 0.081=0.23$ mg.

8.3.5. The *Guide* [H.2] gives additional guidance on choosing k where a small number of measurements is used to estimate large random effects, and should be referred to when estimating degrees of freedom where several contributions are significant.

8.3.6. Where the distributions concerned are normal, a coverage factor of 2 (or chosen according to paragraphs 8.3.3.-8.3.5. using a level of confidence of 95%) gives an interval containing approximately 95% of the distribution of values. It is not recommended that this interval is taken to imply a 95% confidence interval without a knowledge of the distribution concerned.

Table 1: Student's t for 95% confidence (2-tailed)

Degrees of freedom ν	t
1	12.7
2	4.3
3	3.2
4	2.8
5	2.6
6	2.5

9. Reporting Uncertainty

9.1. General

9.1.1. The information necessary to report the result of a measurement depends on its intended use. The guiding principles are:

- present sufficient information to allow the result to be re-evaluated if new information or data become available
- it is preferable to err on the side of providing too much information rather than too little.

9.1.2. When the details of a measurement, including how the uncertainty was determined, depend on references to published documentation, it is imperative that the documentation to hand is kept up to date and consistent with the methods in use.

9.2. Information required

9.2.1. A complete report of a measurement result should include or refer to documentation containing,

- a description of the methods used to calculate the measurement result and its uncertainty from the experimental observations and input data
- the values and sources of all corrections and constants used in both the calculation and the uncertainty analysis
- a list of all the components of uncertainty with full documentation on how each was evaluated

9.2.2. The data and analysis should be presented in such a way that its important steps can be readily followed and the calculation of the result repeated if necessary.

9.2.3. Where a detailed report including intermediate input values is required, the report should

- give the value of each input value, its standard uncertainty and a description of how each was obtained
- give the relationship between the result and the input values and any partial derivatives, covariances or correlation coefficients used to account for correlation effects

- state the estimated number of degrees of freedom for the standard uncertainty of each input value (methods for estimating degrees of freedom are given in the ISO Guide [H.2]).

NOTE: Where the functional relationship is extremely complex or does not exist explicitly (for example, it may only exist as a computer program), the relationship may be described in general terms or by citation of appropriate references. In such cases, it must be clear how the result and its uncertainty were obtained.

9.2.4. When reporting the results of routine analysis, it may be sufficient to state only the value of the expanded uncertainty and the value of k .

9.3. Reporting standard uncertainty

9.3.1. When uncertainty is expressed as the combined standard uncertainty u_c (that is, as a single standard deviation), the following form is recommended:

"(Result): x (units) [with a] standard uncertainty of u_c (units) [where standard uncertainty is as defined in the International Vocabulary of Basic and General terms in Metrology, 2nd ed., ISO 1993 and corresponds to one standard deviation.]"

NOTE The use of the symbol \pm is not recommended when using standard uncertainty as the symbol is commonly associated with intervals corresponding to high levels of confidence.

Terms in parentheses [] may be omitted or abbreviated as appropriate.

EXAMPLE:

Total nitrogen: 3.52 %w/w

Standard uncertainty: 0.07 %w/w *

*Standard uncertainty corresponds to one standard deviation.

9.4. Reporting expanded uncertainty

9.4.1. Unless otherwise required, the result x should be stated together with the expanded uncertainty U calculated using a coverage factor

$k=2$ (or as described in section 8.3.3.). The following form is recommended:

"(Result): $(x \pm U)$ (units)"

[where] the reported uncertainty is [an expanded uncertainty as defined in the International Vocabulary of Basic and General terms in metrology, 2nd ed., ISO 1993,] calculated using a coverage factor of 2, [which gives a level of confidence of approximately 95%]"

Terms in parentheses [] may be omitted or abbreviated as appropriate. The coverage factor should, of course, be adjusted to show the value actually used.

EXAMPLE:

Total nitrogen: (3.52 ± 0.14) % w/w *

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

9.5. Numerical expression of results

9.5.1. The numerical values of the result and its uncertainty should not be given with an excessive number of digits. Whether expanded uncertainty U or a standard uncertainty u is given, it is seldom necessary to give more than two significant digits for the uncertainty. Results should be rounded to be consistent with the uncertainty given.

9.6. Compliance against limits

9.6.1. Regulatory compliance often requires that a measurand, such as the concentration of a toxic substance, be shown to be within particular limits. Measurement uncertainty clearly has implications for interpretation of analytical results in this context. In particular:

- 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. The following paragraphs give examples of common practice.

9.6.2. Assuming that limits were set with no allowance for uncertainty, four situations are apparent for the case of compliance with an upper limit (see Figure 2):

- i) The result exceeds the limit value plus the expanded uncertainty.
- ii) The result exceeds the limiting value by less than the expanded uncertainty.
- iii) The result is below the limiting value by less than the expanded uncertainty
- iv) The result is less than the limiting value minus the expanded uncertainty.

Case i) is normally interpreted as demonstrating clear non-compliance. Case iv) is normally interpreted as demonstrating compliance. Cases ii) and iii) will normally require individual consideration in the light of any agreements with the user of the data. Analogous arguments apply in the case of compliance with a lower limit.

9.6.3. Where it is known or believed that limits have been set with some allowance for uncertainty, a judgement of compliance can reasonably be made only with knowledge of that allowance. An exception arises where compliance is set against a stated method operating in defined circumstances. Implicit in such a requirement is the assumption that the uncertainty, or at least reproducibility, of the stated method is small enough to ignore for practical purposes. In such a case, provided that appropriate quality control is in place, compliance is normally reported only on the value of the particular result. This will normally be stated in any standard taking this approach.

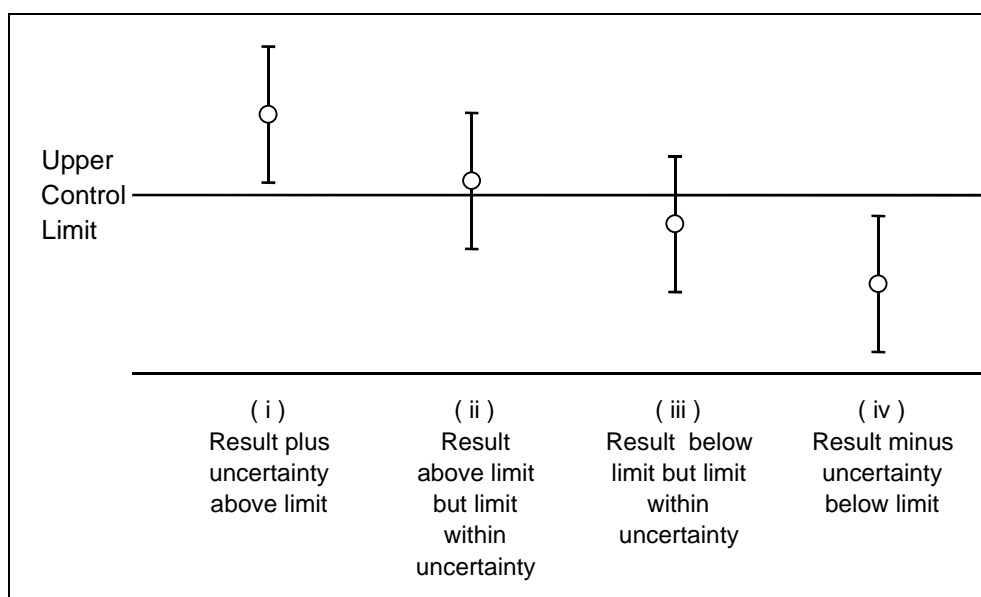


Figure 2: Uncertainty and compliance limits

Appendix A. Examples

Introduction

General introduction

These examples illustrate how the techniques for evaluating uncertainty, described in sections 5-7, can be applied to some typical chemical analyses. They all follow the procedure shown in the flow diagram (Figure 1 on page 12). The uncertainty sources are identified and set out in a cause and effect diagram (see appendix D). This helps to avoid double counting of sources and also assists in the grouping together of components whose combined effect can be evaluated. Examples 1-6 illustrate the use of the spreadsheet method of Appendix E.2 for calculating the combined uncertainties from the calculated contributions $u(y, x_i)$.*

Each of examples 1-6 has an introductory summary. This gives an outline of the analytical method, a table of the uncertainty sources and their respective contributions, a graphical comparison of the different contributions, and the combined uncertainty.

Examples 1-3 illustrate the evaluation of the uncertainty by the quantification of the uncertainty arising from each source separately. Each gives a detailed analysis of the uncertainty associated with the measurement of volumes using volumetric glassware and masses from difference weighings. The detail is for illustrative purposes, and should not be taken as a general recommendation as to the level of detail required or the approach taken. For many analyses, the uncertainty associated with these operations will not be significant and such a detailed evaluation will not be necessary. It would be sufficient to use typical values for these operations with due allowance being made for the actual values of the masses and volumes involved.

Example A1

Example A1 deals with the very simple case of the preparation of a calibration standard of cadmium in HNO_3 for AAS. Its purpose is to

show how to evaluate the components of uncertainty arising from the basic operations of volume measurement and weighing and how these components are combined to determine the overall uncertainty.

Example A2

This deals with the preparation of a standardised solution of sodium hydroxide (NaOH) which is standardised against the titrimetric standard potassium hydrogen phthalate (KHP). It includes the evaluation of uncertainty on simple volume measurements and weighings, as described in example A1, but also examines the uncertainty associated with the titrimetric determination.

Example A3

Example A3 expands on example A2 by including the titration of an HCl against the prepared NaOH solution.

Example A4

This illustrates the use of in house validation data, as described in section 7.7., and shows how the data can be used to evaluate the uncertainty arising from combined effect of a number of sources. It also shows how to evaluate the uncertainty associated with method bias.

Example A5

This shows how to evaluate the uncertainty on results obtained using a standard or “empirical” method to measure the amount of heavy metals leached from ceramic ware using a defined procedure, as described in section 7.2.-7.8. Its purpose is to show how, in the absence of collaborative trial data or ruggedness testing results, it is necessary to consider the uncertainty arising from the range of the parameters (e.g. temperature, etching time and acid strength) allowed in the method definition. This process is considerably simplified when collaborative study data is available, as is shown in the next example.

Example A6

The sixth example is based on an uncertainty estimate for a crude (dietary) fibre determination.

* Section 8.2.2. explains the theory behind the calculated contributions $u(y, x_i)$.

Since the analyte is defined only in terms of the standard method, the method is empirical. In this case, collaborative study data, in-house QA checks and literature study data were available, permitting the approach described in section 7.6. The in-house studies verify that the method is performing as expected on the basis of the collaborative study. The example shows how the use of collaborative study data backed up by in-house method performance checks can substantially reduce the number of different contributions required to form an uncertainty estimate under these circumstances.

Example A7

This gives a detailed description of the evaluation of uncertainty on the measurement of the lead content of a water sample using IDMS. In addition to identifying the possible sources of uncertainty and quantifying them by statistical means the examples shows how it is also necessary to include the evaluation of components based on judgement as described in section 7.14. Use of judgement is a special case of Type B evaluation as described in the ISO Guide [H.2].

Example A1: Preparation of a Calibration Standard

Summary

Goal

A calibration standard is prepared from a high purity metal (cadmium) with a concentration of ca.1000 mg l⁻¹.

Measurement procedure

The surface of the high purity metal is cleaned to remove any metal-oxide contamination. Afterwards the metal is weighed and then dissolved in nitric acid in a volumetric flask. The stages in the procedure are shown in the following flow chart.

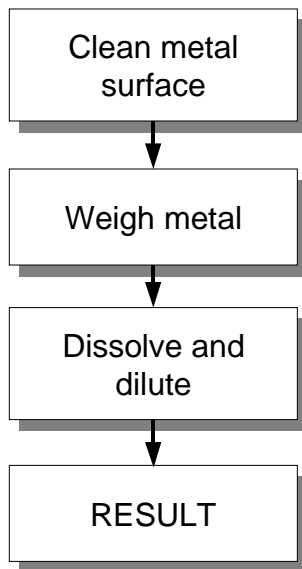


Figure A1. 1: Preparation of cadmium standard

Measurand

$$c_{Cd} = \frac{1000 \cdot m \cdot P}{V} [\text{mg l}^{-1}]$$

where

c_{Cd} :concentration of the calibration standard [mg l⁻¹]

1000 :conversion factor from [ml] to [l]

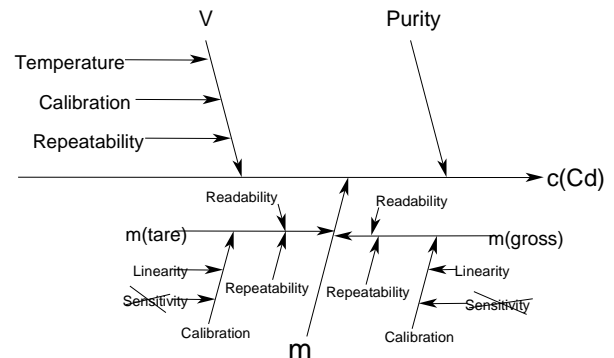
m :mass of the high purity metal [mg]

P :purity of the metal given as mass fraction

V :volume of the liquid of the calibration standard [ml]

Identification of the uncertainty sources:

The relevant uncertainty sources are shown in the cause and effect diagram below:



Quantification of the uncertainty components

The values and their uncertainties are shown in the Table below.

Combined Standard Uncertainty

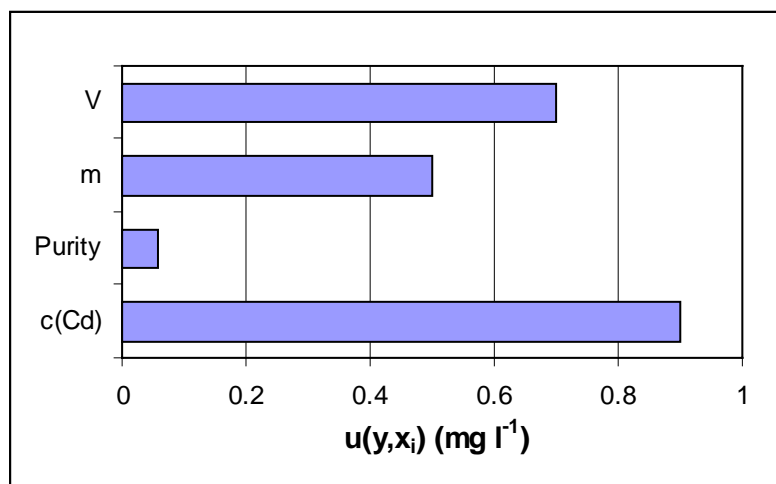
The combined standard uncertainty for the preparation of a 1002.7 mg l⁻¹ Cd calibration standard is 0.9 mg l⁻¹

The different contributions are shown diagrammatically in Figure A1.2.

Table A1.1: Values and uncertainties

	Description	Value	Standard uncertainty	Relative standard uncertainty $u(x)/x$
P	Purity of the metal	0.9999	0.000058	0.000058
m	Mass of the metal	100.28 mg	0.05 mg	0.0005
V	Volume of the flask	100.0 ml	0.07 ml	0.0007
c_{Cd}	concentration of the calibration standard	1002.7 mg l ⁻¹	0.9 mg l ⁻¹	0.0009

Figure A1.2: Uncertainty contributions in cadmium standard preparation



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A1.3

Example A1: Preparation of a calibration standard. Detailed discussion

A1.1 Introduction

This first introductory example discusses the preparation of a calibration standard for atomic absorption spectroscopy (AAS) from the corresponding high purity metal (in this example $\approx 1000 \text{ mg l}^{-1}$ Cd in dilute HNO_3). Even though the example does not represent an entire analytical measurement, the use of calibration standards is part of nearly every determination, because modern routine analytical measurements are relative measurements, which need a reference standard to provide traceability to the SI.

A1.2 Step 1: Specification

The goal of this first step is to write down a clear statement of what is being measured. This specification includes a description of the preparation of the calibration standard and the mathematical relationship between the measurand and the parameters upon which it depends.

Procedure

The specific information on how to prepare a calibration standard is normally given in a Standard Operating Procedure (SOP). The preparation consists of the following stages

The separate stages are:

- i) The surface of the high purity metal is treated with an acid mixture to remove any metal-oxide contamination. The cleaning method is provided by the manufacturer of the metal and needs to be carried out to obtain the purity quoted on the certificate.
- ii) The volumetric flask (100 ml) is weighed without and with the purified metal inside. The balance used has a resolution of 0.01 mg.
- iii) 1 ml of nitric acid (65% m/m) and 3 ml of ion-free water are added to the flask to dissolve the cadmium (approximately 100 mg, weighed accurately). Afterwards the flask is filled with ion-free water up to the mark and mixed by inverting the flask at least thirty times.

Calculation:

The measurand in this example is the concentration of the calibration standard solution, which depends upon the weighing of the high purity metal (Cd), its purity and the volume of the liquid in which it is dissolved. The concentration is given by

$$c_{Cd} = \frac{1000 \cdot m \cdot P}{V} \text{ mg l}^{-1}$$

where

c_{Cd} :concentration of the calibration standard [mg l^{-1}]

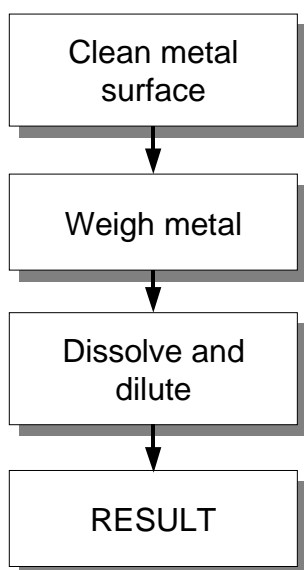
1000 :conversion factor from [ml] to [l]

m :mass of the high purity metal [mg]

P :purity of the metal given as mass fraction

V :volume of the liquid of the calibration standard [ml]

Figure A1.3: Preparation of cadmium standard



A1.3 Step 2: Identifying and analysing uncertainty sources

The aim of this second step is to list all the uncertainty sources for each of the parameters which affect the value of the measurand.

Purity

The purity of the metal (Cd) is quoted in the supplier's certificate as $99.99 \pm 0.01\%$. P is therefore 0.9999 ± 0.0001 . These values depend on the effectiveness of the surface cleaning of the high purity metal. If the manufacturer's procedure is strictly followed, no additional uncertainty due to the contamination of the surface with metal-oxide needs to be added to the value given in the certificate. There is no information available that 100% of the metal dissolves. Therefore one has to check with a repeated preparation experiment that this contribution can be neglected.

Mass m

The second stage of the preparation involves weighing the high purity metal. A 100 ml quantity of a 1000 mg l^{-1} cadmium solution is to be prepared.

The relevant mass of cadmium is determined by a tared weighing, giving $m = 0.10028 \text{ g}$

The manufacturer's literature identifies three uncertainty sources for the tared weighing: the repeatability; the readability (digital resolution) of the balance scale; and the contribution due to the uncertainty in the calibration function of the scale. This calibration function has two potential uncertainty sources, identified as the sensitivity of the balance and its linearity. The sensitivity can be neglected because the mass by difference is done on the same balance over a very narrow range.

NOTE: Buoyancy correction is not considered because all weighing results are quoted on the conventional basis for weighing in air [H.19]. The remaining uncertainties are too small to consider. Note 1 in Appendix G refers.

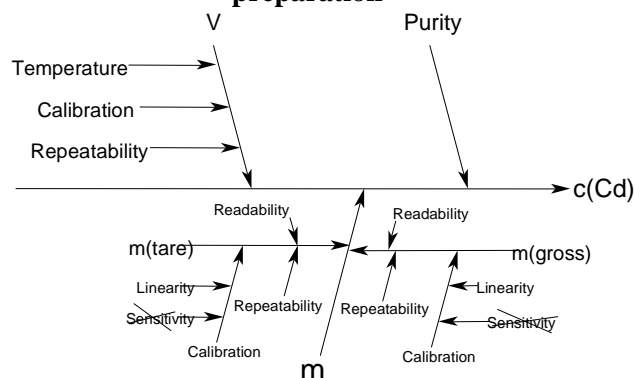
Volume V

The volume of the solution contained in the volumetric flask is subject to three major sources of uncertainty:

- The uncertainty in the certified internal volume of the flask.
- Variation in filling the flask to the mark.
- The flask and solution temperatures differing from the temperature at which the volume of the flask was calibrated.

The different effects and their influences are shown as a cause and effect diagram in Figure A1.4 (see Appendix D for description).

Figure A1.4: Uncertainties in Cd Standard preparation



A1.4 Step 3: Quantifying the uncertainty components

In step 3 the size of each identified potential source of uncertainty is either directly measured, estimated using previous experimental results or derived from theoretical analysis.

Purity

The purity of the cadmium is given on the certificate as 0.9999 ± 0.0001 . Because there is no additional information about the uncertainty value, a rectangular distribution is assumed. To obtain the standard uncertainty $u(P)$ the value of 0.0001 has to be divided by $\sqrt{3}$ (see Appendix E1.1)

$$u(P) = \frac{0.0001}{\sqrt{3}} = 0.000058$$

Mass m

The uncertainty associated with the mass of the cadmium is estimated, using the data from the calibration certificate and the manufacturer's recommendations on uncertainty estimation, as 0.05 mg. This estimate takes into account the three contributions identified earlier (Section A1.3).

NOTE: Detailed calculations for uncertainties in mass can be very intricate, and it is important to refer to manufacturer's literature where mass uncertainties are dominant. In this example, the calculations are omitted for clarity.

Volume V

The volume has three major influences; calibration, repeatability and temperature effects.

- i) *Calibration*: The manufacturer quotes a volume for the flask of 100 ml ± 0.1 ml measured at a temperature of 20 °C. The value of the uncertainty is given without a confidence level or distribution information, so an assumption is necessary. Here, the standard uncertainty is calculated assuming a triangular distribution.

$$\frac{0.1 \text{ ml}}{\sqrt{6}} = 0.04 \text{ ml}$$

NOTE: A triangular distribution was chosen, because in an effective production process, the nominal value is more likely than extremes. The resulting distribution is better represented by a triangular distribution than a rectangular one.

- ii) *Repeatability*: The uncertainty due to variations in filling can be estimated from a repeatability experiment on a typical example of the flask used. A series of ten fill and weigh experiments on a typical 100 ml flask gave a standard deviation of 0.02 ml. This can be used directly as a standard uncertainty.

- iii) *Temperature*: According to the manufacturer the flask has been calibrated at a temperature of 20 °C, whereas the laboratory temperature varies between the limits of ± 4 °C. The uncertainty from this effect can be calculated from the estimate of the temperature range and the coefficient of the volume expansion. The volume expansion of the liquid is considerably larger than that of the flask, so only the former needs to be considered. The coefficient of volume expansion for water is $2.1 \times 10^{-4} \text{ } ^\circ\text{C}^{-1}$,

which leads to a volume variation of

$$\pm (100 \times 4 \times 2.1 \times 10^{-4}) = \pm 0.084 \text{ ml}$$

The standard uncertainty is calculated using the assumption of a rectangular distribution for the temperature variation i.e.

$$\frac{0.084 \text{ ml}}{\sqrt{3}} = 0.05 \text{ ml}$$

The three contributions are combined to give the standard uncertainty $u(V)$ of the volume V

$$u(V) = \sqrt{0.04^2 + 0.02^2 + 0.05^2} = 0.07 \text{ ml}$$

A1.5 Step 4: Calculating the combined standard uncertainty

c_{Cd} is given by

$$c_{Cd} = \frac{1000 \cdot m \cdot P}{V} \quad [\text{mg l}^{-1}]$$

The intermediate values, their standard uncertainties and their relative standard uncertainties are summarised overleaf (Table A1.2)

Using those values, the concentration of the calibration standard is

$$c_{Cd} = \frac{1000 \times 100.28 \times 0.9999}{100.0} = 1002.7 \text{ mg l}^{-1}$$

For this simple multiplicative expression, the uncertainties associated with each component are combined as follows.

$$\begin{aligned} \frac{u_c(c_{Cd})}{c_{Cd}} &= \sqrt{\left(\frac{u(P)}{P}\right)^2 + \left(\frac{u(m)}{m}\right)^2 + \left(\frac{u(V)}{V}\right)^2} \\ &= \sqrt{0.000058^2 + 0.0005^2 + 0.0007^2} \\ &= 0.0009 \end{aligned}$$

$$\begin{aligned} u_c(c_{Cd}) &= c_{Cd} \times 0.0009 = 1002.7 \text{ mg l}^{-1} \times 0.0009 \\ &= 0.9 \text{ mg l}^{-1} \end{aligned}$$

It is preferable to derive the combined standard uncertainty ($u_c(c_{Cd})$) using the spreadsheet method given in Appendix E, since this can be utilised even for complex expressions. The completed spreadsheet is shown in Table A1.3.

The contributions of the different parameters are shown in Figure A1.5. The contribution of the uncertainty on the volume of the flask is the

Table A1.2: Values and Uncertainties

Description	Value x	$u(x)$	$u(x)/x$
Purity of the metal P	0.9999	0.000058	0.000058
Mass of the metal m (mg)	100.28	0.05 mg	0.0005
Volume of the flask V (ml)	100.0	0.07 ml	0.0007

largest and that from the weighing procedure is similar. The uncertainty on the purity of the cadmium has virtually no influence on the overall uncertainty.

The expanded uncertainty $U(c_{Cd})$ is obtained by multiplying the combined standard uncertainty with a coverage factor of 2, giving

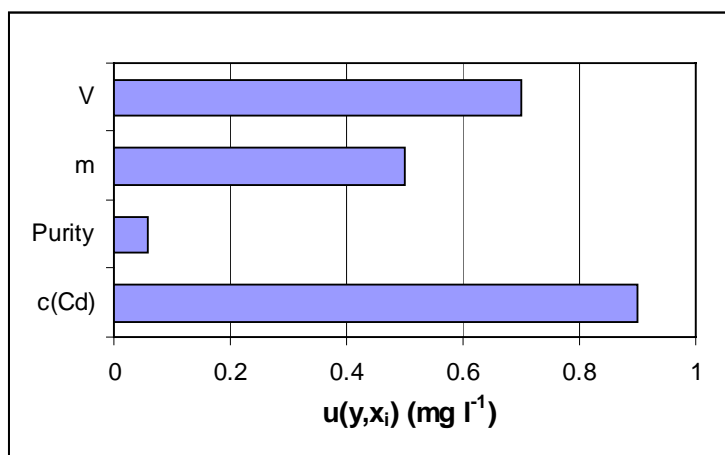
$$U(c_{Cd}) = 2 \times 0.9 \text{ mg l}^{-1} = 1.8 \text{ mg l}^{-1}$$

Table A1.3: Spreadsheet calculation of uncertainty

	A	B	C	D	E
1			P	m	V
2		Value	0.9999	100.28	100.00
3		Uncertainty	0.000058	0.05	0.07
4					
5	P	0.9999	0.999958	0.9999	0.9999
6	m	100.28	100.28	100.33	100.28
7	V	100.0	100.00	100.00	100.07
8					
9	c(Cd)	1002.69972	1002.75788	1003.19966	1001.99832
10	$u(y, x_i)$		0.05816	0.49995	-0.70140
11	$u(y)^2, u(y, x_i)^2$	0.74529	0.00338	0.24995	0.49196
12					
13	$u(c(Cd))$	0.9			

The values of the parameters are entered in the second row from C2 to E2. Their standard uncertainties are in the row below (C3-E3). The spreadsheet copies the values from C2-E2 into the second column from B5 to B7. The result $c(Cd)$ using these values is given in B9. The C5 shows the value of P from C2 plus its uncertainty given in C3. The result of the calculation using the values C5-C7 is given in C9. The columns D and E follow a similar procedure. The values shown in the row 10 (C10-E10) are the differences of the row (C9-E9) minus the value given in B9. In row 11 (C11-E11) the values of row 10 (C10-E10) are squared and summed to give the value shown in B11. B13 gives the combined standard uncertainty, which is the square root of B11.

Figure A1.5: Uncertainty contributions in cadmium standard preparation



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A1.3

Example A2: Standardising a Sodium Hydroxide Solution

Summary

Goal

A solution of sodium hydroxide (NaOH) is standardised against the titrimetric standard potassium hydrogen phthalate (KHP).

Measurement procedure

The titrimetric standard (KHP) is dried and weighed. After the preparation of the NaOH solution the sample of the titrimetric standard (KHP) is dissolved and then titrated using the NaOH solution. The stages in the procedure are shown in the flow chart Figure A2.1.

Measurand:

$$c_{\text{NaOH}} = \frac{1000 \cdot m_{\text{KHP}} \cdot P_{\text{KHP}}}{M_{\text{KHP}} \cdot V_T} \quad [\text{mol l}^{-1}]$$

where

c_{NaOH} :concentration of the NaOH solution
[mol l⁻¹]

1000 :conversion factor [ml] to [l]

m_{KHP} :mass of the titrimetric standard KHP [g]

P_{KHP} :purity of the titrimetric standard given as
:mass fraction

M_{KHP} :molar mass of KHP [g mol⁻¹]

V_T :titration volume of NaOH solution [ml]

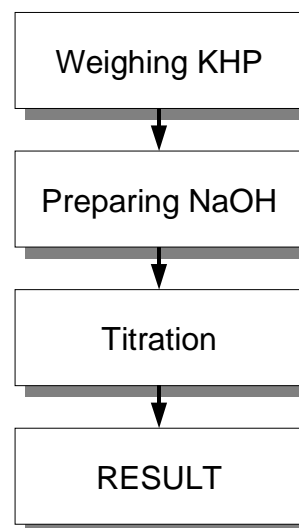


Figure A2.1: Standardising NaOH

Identification of the uncertainty sources:

The relevant uncertainty sources are shown as a cause and effect diagram in Figure A2.2.

Quantification of the uncertainty components

The different uncertainty contributions are given in Table A2.1, and shown diagrammatically in Figure A2.3. The combined standard uncertainty for the 0.10214 mol l⁻¹ NaOH solution is 0.00010 mol l⁻¹.

Table A2.1: Values and uncertainties in NaOH standardisation

	Description	Value x	Standard uncertainty u	Relative standard uncertainty $u(x)/x$
rep	Repeatability	1.0	0.0005	0.0005
m_{KHP}	Mass of KHP	0.3888 g	0.00013 g	0.00033
P_{KHP}	Purity of KHP	1.0	0.00029	0.00029
M_{KHP}	Molar mass of KHP	204.2212 g mol ⁻¹	0.0038 g mol ⁻¹	0.000019
V_T	Volume of NaOH for KHP titration	18.64 ml	0.013 ml	0.0007
c_{NaOH}	NaOH solution	0.10214 mol l ⁻¹	0.00010 mol l ⁻¹	0.00097

Figure A2.2: Cause and effect diagram for titration

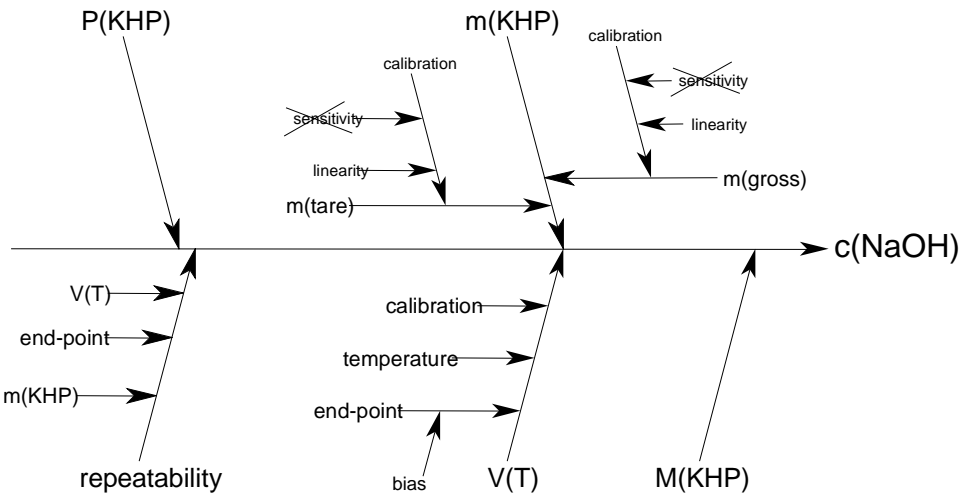
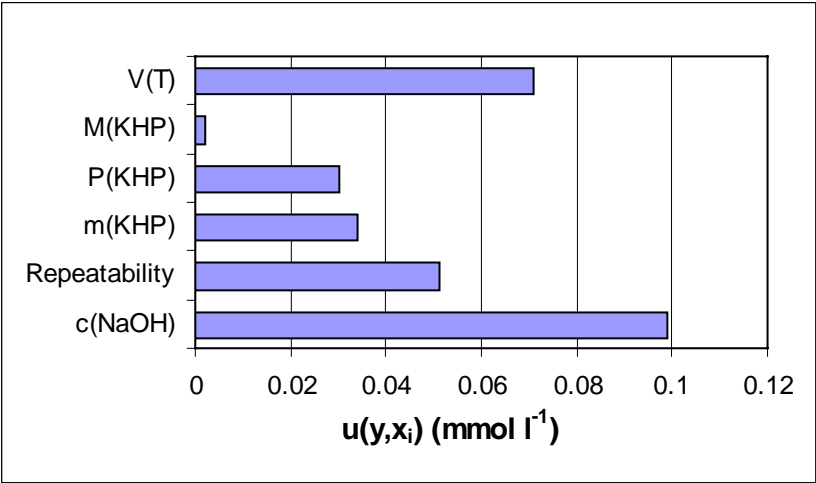


Figure A2.3: Contributions to Titration uncertainty



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A2.3

Example A2: Standardising a sodium hydroxide solution. Detailed discussion

A2.1 Introduction

This second introductory example discusses an experiment to determine the concentration of a solution of sodium hydroxide (NaOH). The NaOH is titrated against the titrimetric standard potassium hydrogen phthalate (KHP). It is assumed that the NaOH concentration is known to be of the order of 0.1 mol l^{-1} . The end-point of the titration is determined by an automatic titration system using a combined pH-electrode to measure the shape of the pH-curve. The functional composition of the titrimetric standard potassium hydrogen phthalate (KHP), which is the number of free protons in relation to the overall number of molecules, provides traceability of the concentration of the NaOH solution to the SI system.

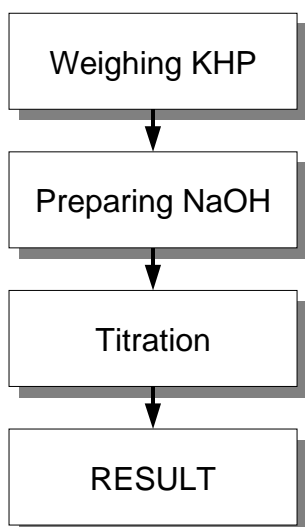
A2.2 Step 1: Specification

The aim of the first step is to describe the measurement procedure. This description consists of a listing of the measurement steps and a mathematical statement of the measurand and the parameters upon which it depends.

Procedure:

The measurement sequence to standardise the NaOH solution has the following stages.

Figure A2.4: Standardisation of a solution of sodium hydroxide



The separate stages are:

- i) The primary standard potassium hydrogen phthalate (KHP) is dried according to the supplier's instructions. The instructions are given in the supplier's catalogue, which also states the purity of the titrimetric standard and its uncertainty. A titration volume of approximately 19 ml of 0.1 mol l^{-1} solution of NaOH entails weighing out an amount as close as possible to

$$\frac{204.2212 \times 0.1 \times 19}{1000 \times 1.0} = 0.388 \text{ g}$$

The weighing is carried out on a balance with a last digit of 0.1 mg.

- ii) A 0.1 mol l^{-1} solution of sodium hydroxide is prepared. In order to prepare 1 l of solution, it is necessary to weigh out $\approx 4 \text{ g}$ NaOH. However, since the concentration of the NaOH solution is to be determined by assay against the primary standard KHP and not by direct calculation, no information on the uncertainty sources connected with the molecular weight or the mass of NaOH taken is required.
- iii) The weighed quantity of the titrimetric standard KHP is dissolved with $\approx 50 \text{ ml}$ of ion-free water and then titrated using the NaOH solution. An automatic titration system controls the addition of NaOH and records the pH-curve. It also determines the end-point of the titration from the shape of the recorded curve.

Calculation:

The measurand is the concentration of the NaOH solution, which depends on the mass of KHP, its purity, its molecular weight and the volume of NaOH at the end-point of the titration

$$c_{\text{NaOH}} = \frac{1000 \cdot m_{\text{KHP}} \cdot P_{\text{KHP}}}{M_{\text{KHP}} \cdot V_T} \quad [\text{mol l}^{-1}]$$

where

c_{NaOH} : concentration of the NaOH solution $[\text{mol l}^{-1}]$

1000 : conversion factor [ml] to [l]

m_{KHP} : mass of the titrimetric standard KHP [g]

- P_{KHP} :purity of the titrimetric standard given as mass fraction
- M_{KHP} :molar mass of KHP [g mol^{-1}]
- V_T :titration volume of NaOH solution [ml]

A2.3 Step 2: Identifying and analysing uncertainty sources

The aim of this step is to identify all major uncertainty sources and to understand their effect on the measurand and its uncertainty. This has been shown to be one of the most difficult step in evaluating the uncertainty of analytical measurements, because there is a risk of

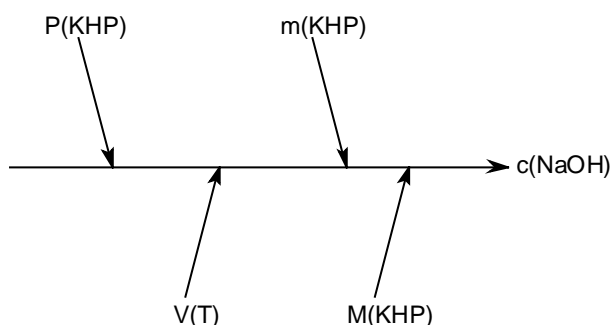


Figure A2.5: First step in setting up a cause and effect diagram

neglecting uncertainty sources on the one hand and an the other of double-counting them. The use of a cause and effect diagram (Appendix D) is one possible way to help prevent this happening. The first step in preparing the diagram is to draw the four parameters of the equation of the measurand as the main branches.

Afterwards, each step of the method is considered

and any further influence quantity is added as a factor to the diagram working outwards from the main effect. This is carried out for each branch until effects become sufficiently remote, that is, until effects on the result are negligible.

Mass m_{KHP}

Approximately 388 mg of KHP are weighed to standardise the NaOH solution. The weighing procedure is a weight by difference. This means that a branch for the determination of the tare (m_{tare}) and another branch for the gross weight (m_{gross}) have to be drawn in the cause and effect diagram. Each of the two weighings is subject to run to run variability and the uncertainty of the calibration of the balance. The calibration itself has two possible uncertainty sources: the sensitivity and the linearity of the calibration function. If the weighing is done on the same scale and over a small range of weight then the sensitivity contribution can be neglected.

All these uncertainty sources are added into the cause and effect diagram (see Figure A2.6).

Purity P_{KHP}

The purity of KHP is quoted in the supplier's catalogue to be within the limits of 99.95% and 100.05%. P_{KHP} is therefore 1.0000 ± 0.0005 . There is no other uncertainty source if the drying procedure was performed according to the suppliers specification.

Molar mass M_{KHP}

Potassium hydrogen phthalate (KHP) has the

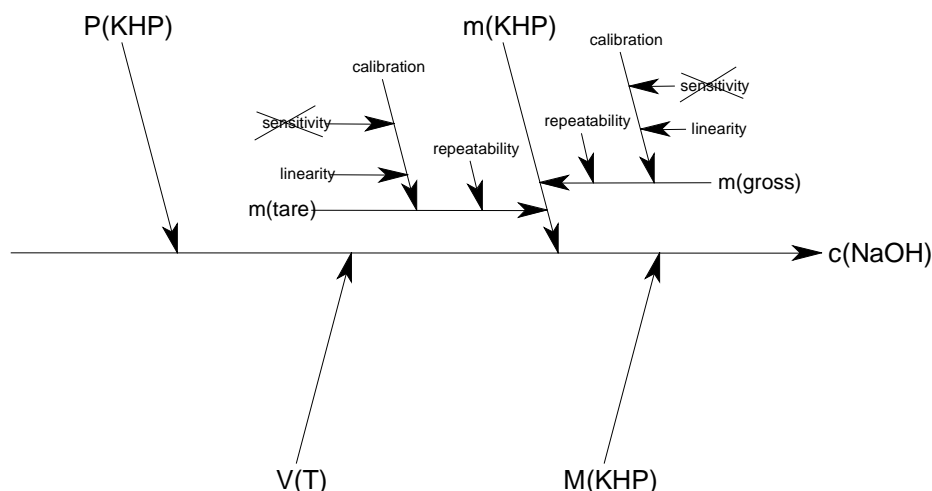


Figure A2.6: Cause and effect diagram with added uncertainty sources for the weighing procedure

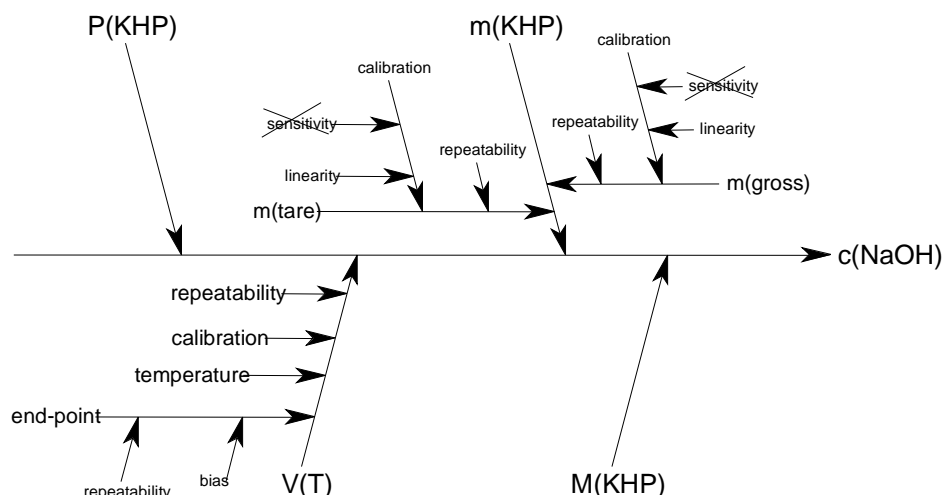


Figure A2.7: Cause and effect diagram (all sources)

empirical formula



The uncertainty in the molar mass of the compound can be determined by combining the uncertainty in the atomic weights of its constituent elements. A table of atomic weights including uncertainty estimates is published biennially by IUPAC in the Journal of Pure and Applied Chemistry. The molar mass can be calculated directly from these; the cause and effect diagram (Figure A2.7) omits the individual atomic masses for clarity

Volume V_T

The titration is accomplished using a 20 ml piston burette. The delivered volume of NaOH from the piston burette is subject to the same three uncertainty sources as the filling of the volumetric flask in the previous example. These uncertainty sources are the repeatability of the delivered volume, the uncertainty of the calibration of that volume and the uncertainty resulting from the difference between the temperature in the laboratory and that of the calibration of the piston burette. In addition there is the contribution of the end-point detection, which has two uncertainty sources.

1. The repeatability of the end-point detection, which is independent of the repeatability of the volume delivery.
2. The possibility of a systematic difference between the determined end-point and the equivalence point (bias), due to carbonate

absorption during the titration and inaccuracy in the mathematical evaluation of the end-point from the titration curve.

These items are included in the cause and effect diagram shown in Figure A2.7.

A2.4 Step 3: Quantifying uncertainty components

In step 3, the uncertainty from each source identified in step 2 has to be quantified and then converted to a standard uncertainty. All experiments always include at least the repeatability of the volume delivery of the piston burette and the repeatability of the weighing operation. Therefore it is reasonable to combine all the repeatability contributions into one contribution for the overall experiment and to use the values from the method validation to quantify its size, leading to the revised cause and effect diagram in Figure A2.8.

The method validation shows a repeatability for the titration experiment of 0.05%. This value can be directly used for the calculation of the combined standard uncertainty.

Mass m_{KHP}

The relevant weighings are:

container and KHP:	60.5450 g(observed)
container less KHP:	60.1562 g(observed)
KHP	0.3888 g(calculated)

Because of the combined repeatability term identified above, there is no need to take into account the weighing repeatability. Any

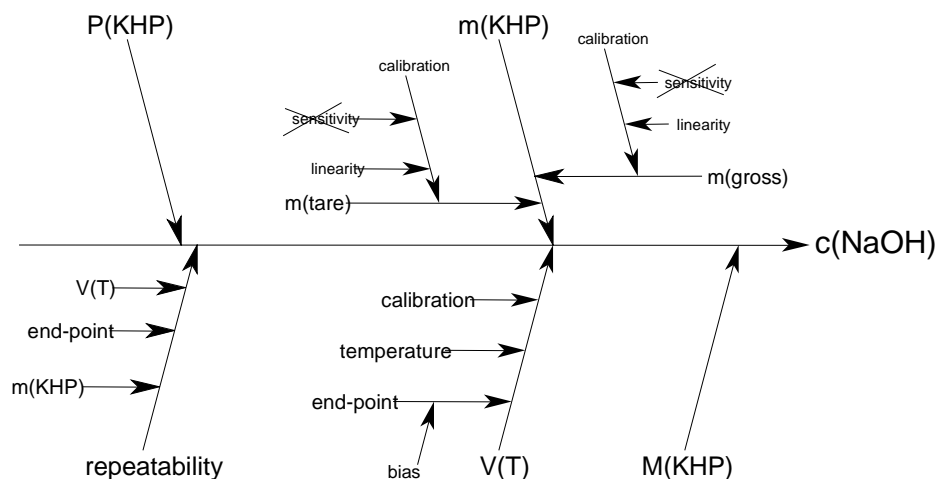


Figure A2.8: Cause and effect diagram (Repeatabilities combined)

systematic offset across the scale will also cancel. The uncertainty therefore arises solely from the balance linearity uncertainty.

Linearity: The calibration certificate of the balance quotes ± 0.15 mg for the linearity. This value is the maximum difference between the actual mass on the pan and the reading of the scale. The balance manufacture's own uncertainty evaluation recommends the use of a rectangular distribution to convert the linearity contribution to a standard uncertainty.

The balance linearity contribution is accordingly

$$\frac{0.15 \text{ mg}}{\sqrt{3}} = 0.09 \text{ mg}$$

This contribution has to be counted twice, once for the tare and once for the gross weight, because each is an independent observation and the linearity effects are not correlated.

This gives for the standard uncertainty $u(m_{KHP})$ of the mass m_{KHP} , a value of

$$u(m_{KHP}) = \sqrt{2 \times (0.09^2)}$$

$$\Rightarrow u(m_{KHP}) = 0.13 \text{ mg}$$

NOTE 1: Buoyancy correction is not considered because all weighing results are quoted on the conventional basis for weighing in air [H.19]. The remaining uncertainties are too small to consider. Note 1 in Appendix G refers.

NOTE 2: There are other difficulties when weighing a titrimetric standard. A temperature difference of only 1 °C between the standard and the

balance causes a drift in the same order of magnitude as the repeatability contribution. The titrimetric standard has been completely dried, but the weighing procedure is carried out at a humidity of around 50 % relative humidity, so adsorption of some moisture is expected.

Purity P_{KHP}

P_{KHP} is 1.0000 ± 0.0005 . The supplier gives no further information concerning the uncertainty in the catalogue. Therefore this uncertainty is taken as having a rectangular distribution, so the standard uncertainty $u(P_{KHP})$ is $0.0005/\sqrt{3} = 0.00029$.

Molar mass M_{KHP}

From the latest IUPAC table, the atomic weights and listed uncertainties for the constituent elements of KHP ($C_8H_5O_4K$) are:

Element	Atomic weight	Quoted uncertainty	Standard uncertainty
C	12.0107	± 0.0008	0.00046
H	1.00794	± 0.00007	0.000040
O	15.9994	± 0.0003	0.00017
K	39.0983	± 0.0001	0.000058

For each element, the standard uncertainty is found by treating the IUPAC quoted uncertainty as forming the bounds of a rectangular distribution. The corresponding standard uncertainty is therefore obtained by dividing those values by $\sqrt{3}$.

The separate element contributions to the molar mass, together with the uncertainty contribution for each, are:

	Calculation	Result	Standard uncertainty
C ₈	8×12.0107	96.0856	0.0037
H ₅	5×1.00794	5.0397	0.00020
O ₄	4×15.9994	63.9976	0.00068
K	1×39.0983	39.0983	0.000058

The uncertainty in each of these values is calculated by multiplying the standard uncertainty in the previous table by the number of atoms.

This gives a molar mass for KHP of

$$M_{KHP} = 96.0856 + 5.0397 + 63.9976 + 39.0983 \\ = 204.2212 \text{ g mol}^{-1}$$

As this expression is a sum of independent values, the standard uncertainty $u(M_{KHP})$ is a simple square root of the sum of the squares of the contributions:

$$u(M_{KHP}) = \sqrt{0.0037^2 + 0.0002^2 + 0.00068^2 + 0.000058^2} \\ \Rightarrow u(M_{KHP}) = 0.0038 \text{ g mol}^{-1}$$

NOTE: Since the element contributions to M_{KHP} are simply the sum of the single atom contributions, it might be expected from the general rule for combining uncertainty contributions that the uncertainty for each element contribution would be calculated from the sum of squares of the single atom contributions, that is, for carbon, $u(M_C) = \sqrt{8 \times 0.00037^2} = 0.001$. Recall, however, that this rule applies only to independent contributions, that is, contributions from separate determinations of the value. In this case, the total is obtained by multiplying a single value by 8. Notice that the contributions from different elements are independent, and will therefore combine in the usual way.

Volume V_T

1. *Repeatability of the volume delivery:* As before, the repeatability has already been taken into account via the combined repeatability term for the experiment.

2. *Calibration:* The limits of accuracy of the delivered volume are indicated by the manufacturer as a \pm figure. For a 20 ml piston burette this number is typically ± 0.03 ml. Assuming a triangular distribution gives a standard uncertainty of $0.03/\sqrt{6} = 0.012$ ml.

Note: The ISO Guide (F.2.3.3) recommends adoption of a triangular distribution if there are reasons to expect values in the centre of the range being more likely than those near the bounds. For the glassware in examples A1 and A2, a triangular distribution has been assumed (see the discussion under Volume uncertainties in example A1).

3. *Temperature:* The uncertainty due to the lack of temperature control is calculated in the same way as in the previous example, but this time taking a possible temperature variation of ± 3 °C (with a 95% confidence). Again using the coefficient of volume expansion for water as $2.1 \times 10^{-4} \text{ °C}^{-1}$ gives a value of

$$\frac{19 \times 2.1 \times 10^{-4} \times 3}{1.96} = 0.006 \text{ ml}$$

Thus the standard uncertainty due to incomplete temperature control is 0.006 ml.

NOTE: When dealing with uncertainties arising from incomplete control of environmental factors such as temperature, it is essential to take account of any correlation in the effects on different intermediate values. In this example, the dominant effect on the solution temperature is taken as the differential heating effects of different solutes, that is, the solutions are not equilibrated to ambient temperature. Temperature effects on each solution concentration at STP are therefore uncorrelated in this example, and are consequently treated as independent uncertainty contributions.

4. *Bias of the end-point detection:* The titration is performed under a layer of Argon to exclude any bias due to the absorption of CO₂ in the titration solution. This approach follows the principle that it is better to prevent any bias than to correct for it. There are no other indications that the end-point determined from the shape of the pH-curve does not correspond to the equivalence-point, because a strong acid is titrated with a strong base. Therefore it is

Table A2.2: Values and uncertainties for titration

	Description	Value x	Standard uncertainty $u(x)$	Relative standard uncertainty $u(x)/x$
rep	Repeatability	1.0	0.0005	0.0005
m_{KHP}	Weight of KHP	0.3888 g	0.00013 g	0.00033
P_{KHP}	Purity of KHP	1.0	0.00029	0.00029
M_{KHP}	Molar mass of KHP	204.2212 g mol ⁻¹	0.0038 g mol ⁻¹	0.000019
V_T	Volume of NaOH for KHP titration	18.64 ml	0.013 ml	0.0007

assumed that the bias of the end-point detection and its uncertainty are negligible.

V_T is found to be 18.64 ml and combining the remaining contributions to the uncertainty $u(V_T)$ of the volume V_T gives a value of

$$u(V_T) = \sqrt{0.012^2 + 0.006^2}$$

$$\Rightarrow u(V_T) = 0.013 \text{ ml}$$

A2.5 Step 4: Calculating the combined standard uncertainty

c_{NaOH} is given by

$$c_{NaOH} = \frac{1000 \cdot m_{KHP} \cdot P_{KHP}}{M_{KHP} \cdot V_T} \quad [\text{mol l}^{-1}]$$

The values of the parameters in this equation, their standard uncertainties and their relative standard uncertainties are collected in Table A2.2

Using the values given above:

$$c_{NaOH} = \frac{1000 \times 0.3888 \times 1.0}{204.2212 \times 18.64} = 0.10214 \text{ mol l}^{-1}$$

For a multiplicative expression (as above) the standard uncertainties are used as follows:

$$\frac{u_c(c_{NaOH})}{c_{NaOH}} = \sqrt{\left(\frac{u(rep)}{rep}\right)^2 + \left(\frac{u(m_{KHP})}{m_{KHP}}\right)^2 + \left(\frac{u(P_{KHP})}{P_{KHP}}\right)^2 + \left(\frac{u(M_{KHP})}{M_{KHP}}\right)^2 + \left(\frac{u(V_T)}{V_T}\right)^2}$$

$$\Rightarrow \frac{u_c(c_{NaOH})}{c_{NaOH}} = \sqrt{0.0005^2 + 0.00033^2 + 0.00029^2 + 0.000019^2 + 0.0007^2}$$

$$= 0.00097$$

$$\Rightarrow u_c(c_{NaOH}) = c_{NaOH} \times 0.00097 = 0.00010 \text{ mol l}^{-1}$$

Spreadsheet software is used to simplify the above calculation of the combined standard uncertainty (see Appendix E.2). The spreadsheet filled in with the appropriate values is shown as Table A2.3, which appears with additional explanation.

It is instructive to examine the relative contributions of the different parameters. The contributions can easily be visualised using a histogram. Figure A2.9 shows the calculated values $|u(y, x_i)|$ from Table A2.3.

The contribution of the uncertainty of the titration volume V_T is by far the largest followed by the repeatability. The weighing procedure and the purity of the titrimetric standard show the same order of magnitude, whereas the uncertainty in the molar mass is again nearly an order of magnitude smaller.

Figure A2.9: Uncertainty contributions in NaOH standardisation

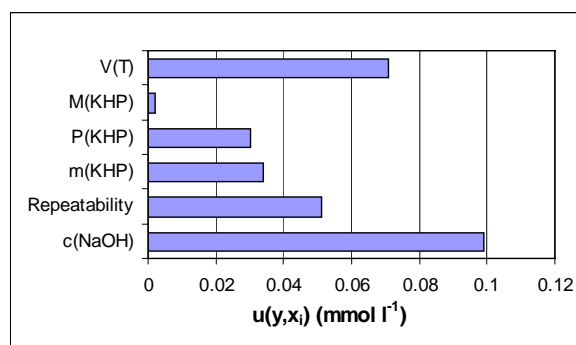


Table A2.3: Spreadsheet calculation of titration uncertainty

	A	B	C	D	E	F	G
1			Rep	m(KHP)	P(KHP)	M(KHP)	V(T)
2		Value	1.0	0.3888	1.0	204.2212	18.64
3		Uncertainty	0.0005	0.00013	0.00029	0.0038	0.013
4							
5	rep	1.0	1.0005	1.0	1.0	1.0	1.0
6	m(KHP)	0.3888	0.3888	0.38893	0.3888	0.3888	0.3888
7	P(KHP)	1.0	1.0	1.0	1.00029	1.0	1.0
8	M(KHP)	204.2212	204.2212	204.2212	204.2212	204.2250	204.2212
9	V(T)	18.64	18.64	18.64	18.64	18.64	18.653
10							
11	c(NaOH)	0.102136	0.102187	0.102170	0.102166	0.102134	0.102065
12	$u(y, x_i)$		0.000051	0.000034	0.000030	-0.000002	-0.000071
13	$u(y)^2, u(y, x_i)^2$	9.72E-9	2.62E-9	1.16E-9	9E-10	4E-12	5.041E-9
14							
15	$u(c(\text{NaOH}))$	0.000099					

The values of the parameters are given in the second row from C2 to G2. Their standard uncertainties are entered in the row below (C3-G3). The spreadsheet copies the values from C2-G2 into the second column from B5 to B9. The result ($c(\text{NaOH})$) using these values is given in B11. C5 shows the value of the repeatability from C2 plus its uncertainty given in C3. The result of the calculation using the values C5-C9 is given in C11. The columns D and G follow a similar procedure. The values shown in the row 12 (C12-G12) are the differences of the row (C11-G11) minus the value given in B11. In row 13 (C13-G13) the values of row 12 (C12-G12) are squared and summed to give the value shown in B13. B15 gives the combined standard uncertainty, which is the square root of B13.

A2.6 Step 5: Re-evaluate the significant components

The contribution of $V(T)$ is the largest one. The volume of NaOH for titration of KHP ($V(T)$) itself is affected by four influence quantities: the repeatability of the volume delivery, the calibration of the piston burette, the difference between the operation and calibration temperature of the burette and the repeatability of the end-point detection. Checking the size of each contribution, the calibration is by far the largest. Therefore this contribution needs to be investigated more thoroughly.

The standard uncertainty of the calibration of $V(T)$ was calculated from the data given by the manufacturer assuming a triangular distribution. The influence of the choice of the shape of the distribution is shown in Table A2.4.

According to the ISO Guide 4.3.9 Note 1:

“For a normal distribution with expectation μ and standard deviation σ , the interval $\mu \pm 3\sigma$ encompasses approximately 99.73 percent of the distribution. Thus, if the upper and lower bounds a_+ and a_- define 99.73 percent limits rather than 100 percent limits, and X_i can be assumed to be approximately normally distributed rather than there being no specific knowledge about X_i [between the bounds], then $u^2(x_i) = a^2/9$. By comparison, the variance of a symmetric rectangular distribution of the half-width a is $a^2/3$... and that of a symmetric triangular distribution of the half-width a is $a^2/6$... The magnitudes of the variances of the three distributions are surprisingly similar in view of the differences in the assumptions upon which they are based.”

Thus the choice of the distribution function of this influence quantity has little effect on the value of the combined standard uncertainty ($u_c(c_{\text{NaOH}})$) and it is adequate to assume that it is triangular.

The expanded uncertainty $U(c_{NaOH})$ is obtained by multiplying the combined standard uncertainty by a coverage factor of 2.

$$U(c_{NaOH}) = 0.00010 \times 2 = 0.0002 \text{ mol l}^{-1}$$

Thus the concentration of the NaOH solution is **$(0.1021 \pm 0.0002) \text{ mol l}^{-1}$** .

Table A2.4: Effect of different distribution assumptions

Distribution	factor	$u(V(T;cal))$ (ml)	$u(V(T))$ (ml)	$u_c(c_{NaOH})$
Rectangular	$\sqrt{3}$	0.017	0.019	$0.00011 \text{ mol l}^{-1}$
Triangular	$\sqrt{6}$	0.012	0.015	$0.00009 \text{ mol l}^{-1}$
Normal ^{Note 1}	$\sqrt{9}$	0.010	0.013	$0.000085 \text{ mol l}^{-1}$

Note 1: The factor of $\sqrt{9}$ arises from the factor of 3 in Note 1 of ISO Guide 4.3.9 (see page 48 for details).

Example A3: An Acid/Base Titration

Summary

Goal

A solution of hydrochloric acid (HCl) is standardised against a solution of sodium hydroxide (NaOH) with known content.

Measurement procedure

A solution of hydrochloric acid (HCl) is titrated against a solution of sodium hydroxide (NaOH), which has been standardised against the titrimetric standard potassium hydrogen phthalate (KHP), to determine its concentration. The stages of the procedure are shown in Figure A3.1.

Measurand:

$$c_{\text{HCl}} = \frac{1000 \cdot m_{\text{KHP}} \cdot P_{\text{KHP}} \cdot V_{\text{T2}}}{V_{\text{T1}} \cdot M_{\text{KHP}} \cdot V_{\text{HCl}}} \quad [\text{mol l}^{-1}]$$

where the symbols are as given in Table A3.1 and the value of 1000 is a conversion factor from ml to litres.

Identification of the uncertainty sources:

The relevant uncertainty sources are shown in Figure A3.2.

Quantification of the uncertainty components

The final uncertainty is estimated as $0.00016 \text{ mol l}^{-1}$. Table A3.1 summarises the values and their uncertainties; Figure A3.3 shows the values diagrammatically.

Figure A3.1: Titration procedure

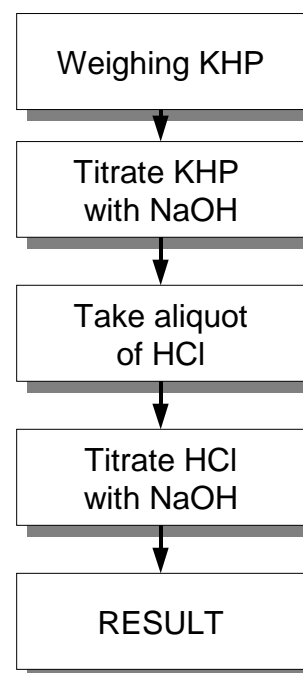


Figure A3.2: Cause and Effect diagram for acid-base titration

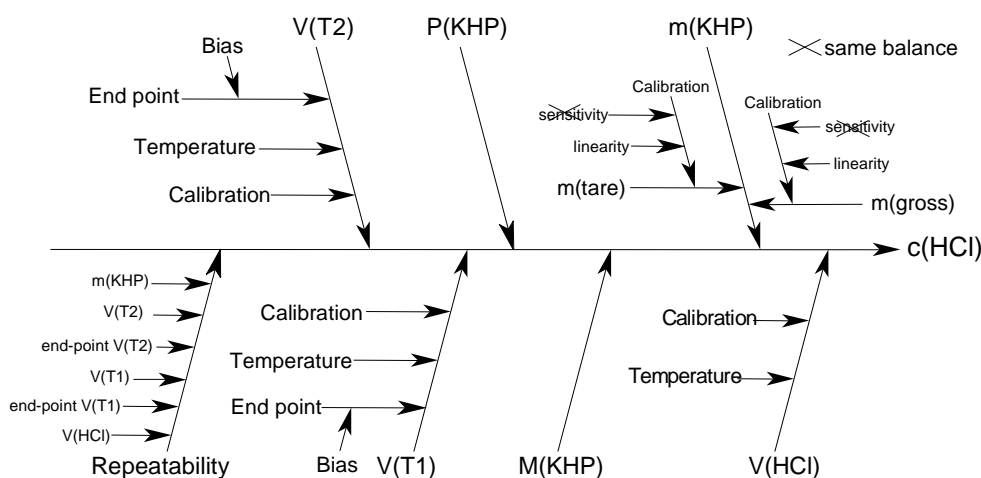
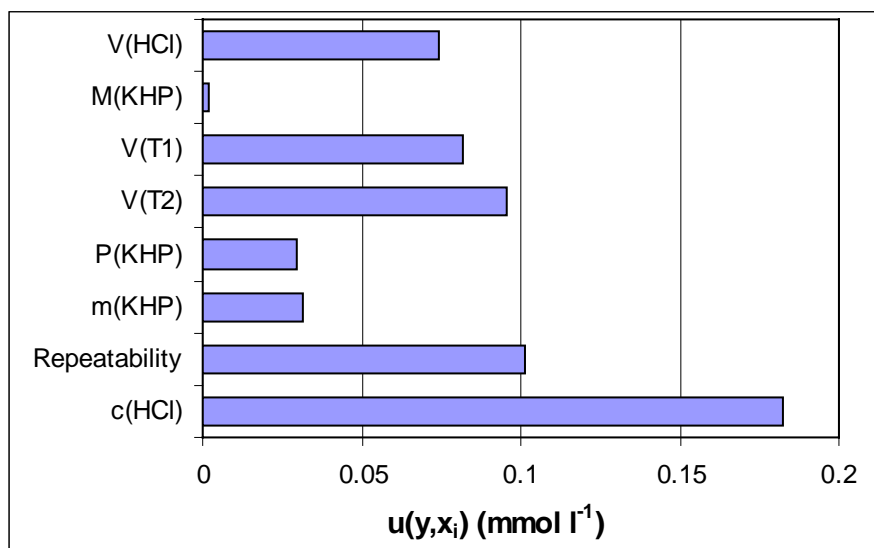


Table A3.1: Acid-base Titration values and uncertainties

	Description	Value x	Standard uncertainty $u(x)$	Relative standard uncertainty $u(x)/x$
rep	Repeatability	1	0.001	0.001
m_{KHP}	Weight of KHP	0.3888 g	0.00013 g	0.00033
P_{KHP}	Purity of KHP	1.0	0.00029	0.00029
V_{T2}	Volume of NaOH for HCl titration	14.89 ml	0.015 ml	0.0010
V_{T1}	Volume of NaOH for KHP titration	18.64 ml	0.016 ml	0.00086
M_{KHP}	Molar mass of KHP	204.2212 g mol ⁻¹	0.0038 g mol ⁻¹	0.000019
V_{HCl}	HCl aliquot for NaOH titration	15 ml	0.011 ml	0.00073
c_{HCl}	HCl solution concentration	0.10139 mol l ⁻¹	0.00016 mol l ⁻¹	0.0016

Figure A3.3: Uncertainty contributions in acid-base titration



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A3.3.

Example A3: An acid/base titration. Detailed discussion

A3.1 Introduction

This example discusses a sequence of experiments to determine the concentration of a solution of hydrochloric acid (HCl). In addition, a number of special aspects of the titration technique are highlighted. The HCl is titrated against solution of sodium hydroxide (NaOH), which was freshly standardised with potassium hydrogen phthalate (KHP). As in the previous example (A2) it is assumed that the HCl concentration is known to be of the order of 0.1 mol l^{-1} and that the end-point of the titration is determined by an automatic titration system using the shape of the pH-curve. This evaluation gives the measurement uncertainty in terms of the SI units of measurement.

A3.2 Step 1: Specification

A detailed description of the measurement procedure is given in the first step. It comprises a listing of the measurement steps and a mathematical statement of the measurand.

Procedure

The determination of the concentration of the HCl solution consists of the following stages (See also Figure A3.4):

- i) The titrimetric standard potassium hydrogen phthalate (KHP) is dried to ensure the purity quoted in the supplier's certificate. Approximately 0.388 g of the dried standard is then weighed to achieve a titration volume of 19 ml NaOH.
- ii) The KHP titrimetric standard is dissolved with $\approx 50 \text{ ml}$ of ion free water and then titrated using the NaOH solution. A titration system controls automatically the addition of NaOH and samples the pH-curve. The end-point is evaluated from the shape of the recorded curve.
- iii) 15 ml of the HCl solution is transferred by means of a volumetric pipette. The HCl solution is diluted with de-ionised water to give $\approx 50 \text{ ml}$ solution in the titration vessel.
- iv) The same automatic titrator performs the measurement of HCl solution.

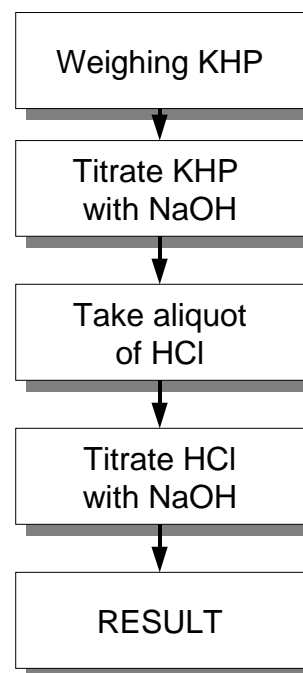


Figure A3.4: Determination of the concentration of a HCl solution

Calculation:

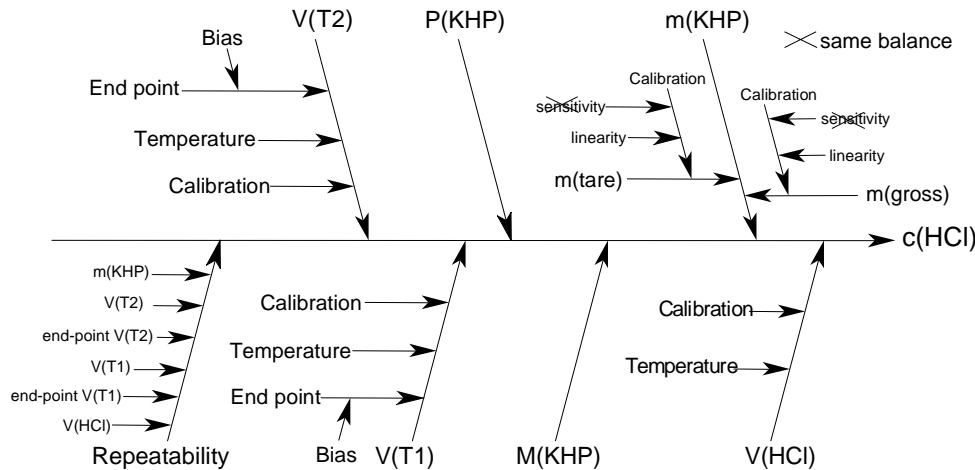
The measurand is the concentration of the HCl solution, c_{HCl} . It depends on the mass of KHP, its purity, its molecular weight, the volumes of NaOH at the end-point of the two titrations and the aliquot of HCl.:

$$c_{\text{HCl}} = \frac{1000 \cdot m_{\text{KHP}} \cdot P_{\text{KHP}} \cdot V_{T2}}{V_{T1} \cdot M_{\text{KHP}} \cdot V_{\text{HCl}}} \quad [\text{mol l}^{-1}]$$

where

- c_{HCl} :concentration of the HCl solution $[\text{mol l}^{-1}]$
 1000 :conversion factor [ml] to [l]
 m_{KHP} :mass of KHP taken [g]
 P_{KHP} :purity of KHP given as mass fraction
 V_{T2} :volume of NaOH solution to titrate HCl [ml]
 V_{T1} :volume of NaOH solution to titrate KHP [ml]
 M_{KHP} : molar mass of KHP $[\text{g mol}^{-1}]$
 V_{HCl} :volume of HCl titrated with NaOH solution [ml]

Figure A3.5: Final cause and effect diagram



A3.3 Step 2: Identifying and analysing uncertainty sources

The different uncertainty sources and their influence on the measurand are best analysed by visualising them first in a cause and effect diagram (Figure A3.5).

Because a repeatability estimate is available from validation studies for the procedure as a whole, there is no need to consider all the repeatability contributions individually. They are therefore grouped into one contribution (shown in the revised cause and effect diagram in Figure A3.5).

The influences on the parameters V_{T2} , V_{T1} , m_{KHP} , P_{KHP} and M_{KHP} have been discussed extensively in the previous example, therefore only the new influence quantities of V_{HCl} will be dealt with in more detail in this section.

Volume V_{HCl}

15 ml of the investigated HCl solution is to be transferred by means of a volumetric pipette. The delivered volume of the HCl from the pipette is subject to the same three sources of uncertainty as all the volumetric measuring devices.

1. The variability or repeatability of the delivered volume
2. The uncertainty in the stated volume of the pipette
3. The solution temperature differing from the calibration temperature of the pipette.

A3.4 Step 3: Quantifying uncertainty components

The goal of this step is to quantify each uncertainty source analysed in step 2. The

quantification of the branches or rather of the different components was described in detail in the previous two examples. Therefore only a summary for each of the different contributions will be given.

repeatability

The method validation shows a repeatability for the determination of 0.1% (as %rsd). This value can be used directly for the calculation of the combined standard uncertainty associated with the different repeatability terms.

Mass m_{KHP}

Calibration/linearity: The balance manufacturer quotes ± 0.15 mg for the linearity contribution. This value represents the maximum difference between the actual mass on the pan and the reading of the scale. The linearity contribution is assumed to show a rectangular distribution and is converted to a standard uncertainty:

$$\frac{0.15}{\sqrt{3}} = 0.087 \text{ mg}$$

The contribution for the linearity has to be accounted for twice, once for the tare and once for the gross mass, leading to an uncertainty $u(m_{KHP})$ of

$$u(m_{KHP}) = \sqrt{2 \times (0.087)^2} \\ \Rightarrow u(m_{KHP}) = 0.12 \text{ mg}$$

NOTE 1: The contribution is applied twice because no assumptions are made about the form of the non-linearity. The non-linearity is accordingly treated as a systematic effect on each weighing, which varies randomly in magnitude across the measurement range.

NOTE 2: Buoyancy correction is not considered because all weighing results are quoted on the conventional basis for weighing in air [H.19]. The remaining uncertainties are too small to consider. Note 1 in Appendix G refers.

P(KHP)

$P(\text{KHP})$ is given in the supplier's certificate as 100% $\pm 0.05\%$. The quoted uncertainty is taken as a rectangular distribution, so the standard uncertainty $u(P_{\text{KHP}})$ is

$$u(P_{\text{KHP}}) = \frac{0.0005}{\sqrt{3}} = 0.00029.$$

V(T2)

- Calibration:* Figure given by the manufacturer (± 0.03 ml) and approximated to a triangular distribution $0.03/\sqrt{6} = 0.012$ ml.
- Temperature:* The possible temperature variation is within the limits of $\pm 4^\circ\text{C}$ and approximated to a rectangular distribution $15 \times 2.1 \times 10^{-4} \times 4/\sqrt{3} = 0.007$ ml.
- Bias of the end-point detection:* A bias between the determined end-point and the equivalence-point due to atmospheric CO_2 can be prevented by performing the titration under Argon. No uncertainty allowance is made.

V_{T2} is found to be 14.89 ml and combining the two contributions to the uncertainty $u(V_{T2})$ of the volume V_{T2} gives a value of

$$u(V_{T2}) = \sqrt{0.012^2 + 0.007^2} \\ \Rightarrow u(V_{T2}) = 0.014 \text{ ml}$$

Volume V_{T1}

All contributions except the one for the temperature are the same as for V_{T2}

- Calibration:* $0.03/\sqrt{6} = 0.012$ ml
- Temperature:* The approximate volume for the titration of 0.3888 g KHP is 19 ml NaOH, therefore its uncertainty contribution is $19 \times 2.1 \times 10^{-4} \times 4/\sqrt{3} = 0.009$ ml.
- Bias:* Negligible

V_{T1} is found to be 18.64 ml with a standard uncertainty $u(V_{T1})$ of

$$u(V_{T1}) = \sqrt{0.012^2 + 0.009^2} \\ \Rightarrow u(V_{T1}) = 0.015 \text{ ml}$$

Molar mass M_{KHP}

Atomic weights and listed uncertainties (from current IUPAC tables) for the constituent elements of KHP ($\text{C}_8\text{H}_5\text{O}_4\text{K}$) are:

Element	Atomic weight	Quoted uncertainty	Standard uncertainty
C	12.0107	± 0.0008	0.00046
H	1.00794	± 0.00007	0.000040
O	15.9994	± 0.0003	0.00017
K	39.0983	± 0.0001	0.000058

For each element, the standard uncertainty is found by treating the IUPAC quoted uncertainty as forming the bounds of a rectangular distribution. The corresponding standard uncertainty is therefore obtained by dividing those values by $\sqrt{3}$.

The molar mass M_{KHP} for KHP and its uncertainty $u(M_{\text{KHP}})$ are, respectively:

$$M_{\text{KHP}} = 8 \times 12.0107 + 5 \times 1.00794 + 4 \times 15.9994 \\ + 39.0983 \\ = 204.2212 \text{ g mol}^{-1}$$

$$u(M_{\text{KHP}}) = \sqrt{(8 \times 0.00046)^2 + (5 \times 0.00004)^2 \\ + (4 \times 0.00017)^2 + 0.000058^2}$$

$$\Rightarrow u(F_{\text{KHP}}) = 0.0038 \text{ g mol}^{-1}$$

NOTE: The single atom contributions are not independent. The uncertainty for the atom contribution is therefore calculated by multiplying the standard uncertainty of the atomic weight by the number of atoms.

Volume V_{HCl}

- Calibration:* Uncertainty stated by the manufacturer for a 15 ml pipette as ± 0.02 ml and approximated with a triangular distribution: $0.02/\sqrt{6} = 0.008$ ml.
- Temperature:* The temperature of the laboratory is within the limits of $\pm 4^\circ\text{C}$. Using a rectangular temperature distribution gives a standard uncertainty of $15 \times 2.1 \times 10^{-4} \times 4/\sqrt{3} = 0.007$ ml.

Combining these contributions gives

$$u(V_{\text{HCl}}) = \sqrt{0.0037^2 + 0.008^2 + 0.007^2} \\ \Rightarrow u(V_{\text{HCl}}) = 0.011 \text{ ml}$$

Table A3.2: Acid-base Titration values and uncertainties (2-step procedure)

	Description	Value x	Standard Uncertainty $u(x)$	Relative standard uncertainty $u(x)/x$
rep	Repeatability	1	0.001	0.001
m_{KHP}	Mass of KHP	0.3888 g	0.00012 g	0.00031
P_{KHP}	Purity of KHP	1.0	0.00029	0.00029
V_{T2}	Volume of NaOH for HCl titration	14.89 ml	0.014 ml	0.00094
V_{T1}	Volume of NaOH for KHP titration	18.64 ml	0.015 ml	0.00080
M_{KHP}	Molar mass of KHP	204.2212 g mol ⁻¹	0.0038 g mol ⁻¹	0.000019
V_{HCl}	HCl aliquot for NaOH titration	15 ml	0.011 ml	0.00073

A3.5 Step 4: Calculating the combined standard uncertainty

c_{HCl} is given by

$$c_{HCl} = \frac{1000 \cdot m_{KHP} \cdot P_{KHP} \cdot V_{T2}}{V_{T1} \cdot M_{KHP} \cdot V_{HCl}}$$

NOTE: The repeatability estimate is, in this example, treated as a relative effect; the complete model equation is therefore

$$c_{HCl} = \frac{1000 \cdot m_{KHP} \cdot P_{KHP} \cdot V_{T2}}{V_{T1} \cdot M_{KHP} \cdot V_{HCl}} \times rep$$

All the intermediate values of the two step experiment and their standard uncertainties are collected in Table A3.2. Using these values:

$$c_{HCl} = \frac{1000 \times 0.3888 \times 1.0 \times 14.89}{18.64 \times 204.2212 \times 15} \times 1 = 0.10139 \text{ mol l}^{-1}$$

The uncertainties associated with each component are combined accordingly:

$$\begin{aligned} \frac{u_c(c_{HCl})}{c_{HCl}} &= \sqrt{\left(\frac{u(m_{KHP})}{m_{KHP}}\right)^2 + \left(\frac{u(P_{KHP})}{P_{KHP}}\right)^2 + \left(\frac{u(V_{T2})}{V_{T2}}\right)^2 + \left(\frac{u(V_{T1})}{V_{T1}}\right)^2 + \left(\frac{u(M_{KHP})}{M_{KHP}}\right)^2 + \left(\frac{u(V_{HCl})}{V_{HCl}}\right)^2 + u(rep)^2} \\ &= \sqrt{0.00031^2 + 0.00029^2 + 0.00094^2 + 0.00080^2 + 0.000019^2 + 0.00073^2 + 0.001^2} \\ &= 0.0018 \end{aligned}$$

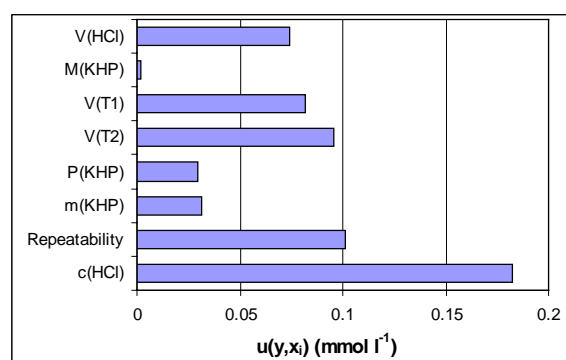
$$\Rightarrow u_c(c_{HCl}) = c_{HCl} \times 0.0018 = 0.00018 \text{ mol l}^{-1}$$

A spreadsheet method (see Appendix E) can be used to simplify the above calculation of the

combined standard uncertainty. The spreadsheet filled in with the appropriate values is shown in Table A3.3, with an explanation.

The sizes of the different contributions can be compared using a histogram. Figure A3.6 shows the values of the contributions $|u(y, x_i)|$ from Table A3.3.

Figure A3.6: Uncertainties in acid-base titration



The expanded uncertainty $U(c_{HCl})$ is calculated by multiplying the combined standard uncertainty by a coverage factor of 2:

$$U(c_{HCl}) = 0.00018 \times 2 = 0.0004 \text{ mol l}^{-1}$$

The concentration of the HCl solution is

$$(0.1014 \pm 0.0004) \text{ mol l}^{-1}$$

Table A3.3: Acid-base Titration – spreadsheet calculation of uncertainty

	A	B	C	D	E	F	G	H	I
1			rep	m(KHP)	P(KHP)	V(T2)	V(T1)	M(KHP)	V(HCl)
2		value	1.0	0.3888	1.0	14.89	18.64	204.2212	15
3		uncertainty	0.001	0.00012	0.00029	0.014	0.015	0.0038	0.011
4									
5	rep	1.0	1.001	1.0	1.0	1.0	1.0	1.0	1.0
6	m(KHP)	0.3888	0.3888	0.38892	0.3888	0.3888	0.3888	0.3888	0.3888
7	P(KHP)	1.0	1.0	1.0	1.00029	1.0	1.0	1.0	1.0
8	V(T2)	14.89	14.89	14.89	14.89	14.904	14.89	14.89	14.89
9	V(T1)	18.64	18.64	18.64	18.64	18.64	18.655	18.64	18.64
10	M(KHP)	204.2212	204.2212	204.2212	204.2212	204.2212	204.2212	204.2250	204.2212
11	V(HCl)	15	15	15	15	15	15	15	15.011
12									
13	c(HCl)	0.101387	0.101489	0.101418	0.101417	0.101482	0.101306	0.101385	0.101313
14	$u(y, x_i)$		0.000101	0.000031	0.000029	0.000095	-0.000082	-0.0000019	-0.000074
15	$u(y)^2, u(y, x_i)^2$	3.34E-8	1.03E-8	9.79E-10	8.64E-10	9.09E-9	6.65E-9	3.56E-12	5.52E-9
16									
17	$u(c(HCl))$	0.00018							

The values of the parameters are given in the second row from C2 to I2. Their standard uncertainties are entered in the row below (C3-I3). The spreadsheet copies the values from C2-I2 into the second column from B5 to B11. The result (c(HCl)) using these values is given in B13. The C5 shows the value of the repeatability from C2 plus its uncertainty given in C3. The result of the calculation using the values C5-C11 is given in C13. The columns D to I follow a similar procedure. The values shown in the row 14 (C14-I14) are the differences of the row (C13-H13) minus the value given in B13. In row 15 (C15-I15) the values of row 14 (C14-I14) are squared and summed to give the value shown in B15. B17 gives the combined standard uncertainty, which is the square root of B15.

A3.6 Special aspects of the titration example

Three special aspects of the titration experiment will be dealt with in this second part of the example. It is interesting to see what effect changes in the experimental set up or in the implementation of the titration would have on the final result and its combined standard uncertainty.

Influence of a mean room temperature of 25°C

For routine analysis, analytical chemists rarely correct for the systematic effect of the temperature in the laboratory on the volume. This question considers the uncertainty introduced by the corrections required.

The volumetric measuring devices are calibrated at a temperature of 20°C. But rarely does any analytical laboratory have a temperature controller to keep the room temperature that level. For illustration, consider correction for a mean room temperature of 25°C.

The final analytical result is calculated using the corrected volumes and not the calibrated volumes at 20°C. A volume is corrected for the temperature effect according to

$$V' = V[1 - \alpha(T - 20)]$$

where

V' : actual volume at the mean temperature T

V : volume calibrated at 20°C

α : expansion coefficient of an aqueous solution [°C⁻¹]

T : observed temperature in the laboratory [°C]

The equation of the measurand has to be rewritten:

$$c_{HCl} = \frac{1000 \cdot m_{KHP} \cdot P_{KHP}}{M_{KHP}} \cdot \frac{V'_{T2}}{V'_{T1} \cdot V'_{HCl}}$$

Including the temperature correction terms gives:

$$c_{HCl} = \frac{1000 \cdot m_{KHP} \cdot P_{KHP}}{M_{KHP}} \cdot \frac{V'_{T2}}{V'_{T1} \cdot V'_{HCl}} \\ = \left(\frac{1000 \cdot m_{KHP} \cdot P_{KHP}}{M_{KHP}} \right) \\ \times \left(\frac{V_{T2} [1 - \alpha(T - 20)]}{V_{T1} [1 - \alpha(T - 20)] \cdot V_{HCl} [1 - \alpha(T - 20)]} \right)$$

This expression can be simplified by assuming that the mean temperature T and the expansion coefficient of an aqueous solution α are the same for all three volumes

$$c_{HCl} = \left(\frac{1000 \cdot m_{KHP} \cdot P_{KHP}}{M_{KHP}} \right) \\ \times \left(\frac{V_{T2}}{V_{T1} \cdot V_{HCl} \cdot [1 - \alpha(T - 20)]} \right)$$

This gives a slightly different result for the HCl concentration at 20°C:

$$c_{HCl} = \frac{1000 \times 0.3888 \times 1.0 \times 14.89}{204.2236 \times 18.64 \times 15 \times [1 - 2.1 \times 10^{-4} (25 - 20)]} \\ = 0.10149 \text{ mol l}^{-1}$$

The figure is still within the range given by the combined standard uncertainty of the result at a mean temperature of 20°C, so the result is not significantly affected. Nor does the change affect the evaluation of the combined standard uncertainty, because a temperature variation of $\pm 4^\circ\text{C}$ at the mean room temperature of 25°C is still assumed.

Visual end-point detection

A bias is introduced if the indicator phenolphthalein is used for visual end-point detection, instead of an automatic titration system extracting the equivalence-point from the pH curve. The change of colour from transparent to red/purple occurs between pH 8.2 and 9.8 leading to an excess volume, introducing a bias compared to the end-point detection employing a pH meter. Investigations have shown that the excess volume is around 0.05 ml with a standard uncertainty for the visual detection of the end-point of approximately 0.03 ml. The bias arising from the excess volume has to be considered in the calculation of the final result. The actual volume for the visual end-point detection is given by

$$V_{T1;Ind} = V_{T1} + V_{Excess}$$

where

$V_{T1;Ind}$: volume from a visual end-point detection

V_{T1} : volume at the equivalence-point

V_{Excess} : excess volume needed to change the colour of phenolphthalein

The volume correction quoted above leads to the following changes in the equation of the measurand

$$c_{HCl} = \frac{1000 \cdot m_{KHP} \cdot P_{KHP} \cdot (V_{T2;Ind} - V_{Excess})}{M_{KHP} \cdot (V_{T1;Ind} - V_{Excess}) \cdot V_{HCl}}$$

The standard uncertainties $u(V_{T2})$ and $u(V_{T1})$ have to be recalculated using the standard uncertainty of the visual end-point detection as the uncertainty component of the repeatability of the end-point detection.

$$u(V_{T1}) = u(V_{T1;Ind} - V_{Excess}) \\ = \sqrt{0.004^2 + 0.012^2 + 0.009^2 + 0.03^2} \\ = 0.034 \text{ ml}$$

$$u(V_{T2}) = u(V_{T2;Ind} - V_{Excess}) \\ = \sqrt{0.004^2 + 0.012^2 + 0.007^2 + 0.03^2} \\ = 0.033 \text{ ml}$$

The combined standard uncertainty

$$u_c(c_{HCl}) = 0.0003 \text{ mol l}^{-1}$$

is considerable larger than before.

Triple determination to obtain the final result

The two step experiment is performed three times to obtain the final result. The triple determination is expected to reduce the contribution from repeatability, and hence reduce the overall uncertainty.

As shown in the first part of this example, all the run to run variations are combined to one single component, which represents the overall experimental repeatability as shown in the in the cause and effect diagram (Figure A3.5).

The uncertainty components are quantified in the following way:

Mass m_{KHP}

$$\text{Linearity: } 0.15/\sqrt{3} = 0.087 \text{ mg}$$

$$\Rightarrow u(m_{KHP}) = \sqrt{2 \times 0.87^2} = 0.12 \text{ mg}$$

Purity P_{KHP}

$$\text{Purity: } 0.0005/\sqrt{3} = 0.00029$$

Volume V_{T2} calibration: $0.03/\sqrt{6} = 0.012 \text{ ml}$

temperature:

$$15 \times 2.1 \times 10^{-4} \times 4/\sqrt{3} = 0.007 \text{ ml}$$

$$\Rightarrow u(V_{T2}) = \sqrt{0.012^2 + 0.007^2} = 0.014 \text{ ml}$$

Repeatability

The quality log of the triple determination shows a mean long term standard deviation of the experiment of 0.001 (as RSD). It is not recommended to use the actual standard deviation obtained from the three determinations because this value has itself an uncertainty of 52%. The standard deviation of 0.001 is divided by the square root of $\sqrt{3}$ to obtain the standard uncertainty of the triple determination (three independent measurements)

$$Rep = 0.001/\sqrt{3} = 0.00058 \text{ (as RSD)}$$

Volume V_{HCl} calibration: $0.02/\sqrt{6} = 0.008 \text{ ml}$ temperature: $15 \times 2.1 \times 10^{-4} \times 4/\sqrt{3} = 0.007 \text{ ml}$

$$\Rightarrow u(V_{HCl}) = \sqrt{0.008^2 + 0.007^2} = 0.01 \text{ ml}$$

Molar mass M_{KHP}

$$u(M_{KHP}) = 0.0038 \text{ g mol}^{-1}$$

Volume V_{T1} calibration: $0.03/\sqrt{6} = 0.12 \text{ ml}$

temperature:

$$19 \times 2.1 \times 10^{-4} \times 4/\sqrt{3} = 0.009 \text{ ml}$$

$$\Rightarrow u(V_{T1}) = \sqrt{0.012^2 + 0.009^2} = 0.015 \text{ ml}$$

All the values of the uncertainty components are summarised in Table A3.4. The combined standard uncertainty is $0.00016 \text{ mol l}^{-1}$, which is a very modest reduction due to the triple determination. The comparison of the uncertainty contributions in the histogram, shown in Figure A3.7, highlights some of the reasons for that result. Though the repeatability contribution is much reduced, the volumetric uncertainty contributions remain, limiting the improvement.

Figure A3.7: Replicated Acid-base Titration values and uncertainties

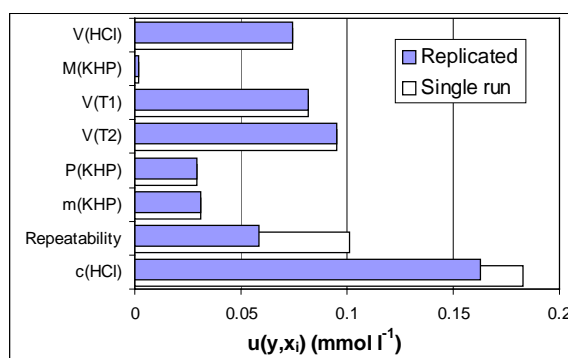


Table A3.4: Replicated Acid-base Titration values and uncertainties

	Description	Value x	Standard Uncertainty $u(x)$	Relative Standard Uncertainty $u(x)/x$
Rep	Repeatability of the determination	1.0	0.00058	0.00058
m_{KHP}	Mass of KHP	0.3888 g	0.00013 g	0.00033
P_{KHP}	Purity of KHP	1.0	0.00029	0.00029
V_{T2}	Volume of NaOH for HCl titration	14.90 ml	0.014 ml	0.00094
V_{T1}	Volume of NaOH for KHP titration	18.65 ml	0.015 ml	0.0008
M_{KHP}	Molar mass of KHP	204.2212 g mol ⁻¹	0.0038 g mol ⁻¹	0.000019
V_{HCl}	HCl aliquot for NaOH titration	15 ml	0.01 ml	0.00067

Example A4: Uncertainty Estimation from In-House Validation Studies. Determination of Organophosphorus Pesticides in Bread.

Summary

Goal

The amount of an organophosphorus pesticide residue in bread is determined employing an extraction and a GC procedure.

Measurement procedure

The stages needed to determine the amount of organophosphorus pesticide residue are shown in Figure A4.1

Measurand:

$$P_{op} = \frac{I_{op} \cdot c_{ref} \cdot V_{op}}{I_{ref} \cdot Rec \cdot m_{sample}} \cdot F_{hom} \text{ mg kg}^{-1}$$

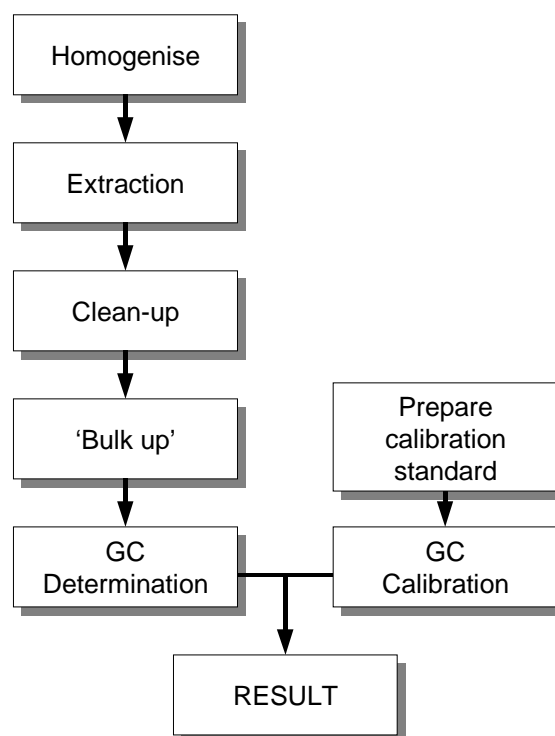
where

- P_{op} :Level of pesticide in the sample [mg kg⁻¹]
 I_{op} :Peak intensity of the sample extract
 c_{ref} :Mass concentration of the reference standard [µg ml⁻¹]
 V_{op} :Final volume of the extract [ml]
 I_{ref} :Peak intensity of the reference standard
 Rec :Recovery
 m_{sample} :Mass of the investigated sub-sample [g]
 F_{hom} :Correction factor for sample inhomogeneity

Identification of the uncertainty sources:

The relevant uncertainty sources are shown in the cause and effect diagram in Figure A4.2.

Figure A4.1: Organophosphorus pesticides analysis



Quantification of the uncertainty components:

Based on in-house validation data, the three major contributions are listed in Table A4.1 and shown diagrammatically in Figure A4.3 (values are from Table A4.5).

Table A4.1: Uncertainties in pesticide analysis

Description	Value x	Standard uncertainty $u(x)$	Relative standard uncertainty $u(x)/x$	Comments
Repeatability(1)	1.0	0.27	0.27	Based on duplicate tests of different types of samples
Bias (Rec) (2)	0.9	0.043	0.048	Spiked samples
Other sources (3) (Homogeneity)	1.0	0.2	0.2	Estimation based on model assumptions
P_{op}	--	--	0.34	Relative standard uncertainty

Figure A4.2: Uncertainty sources in pesticide analysis

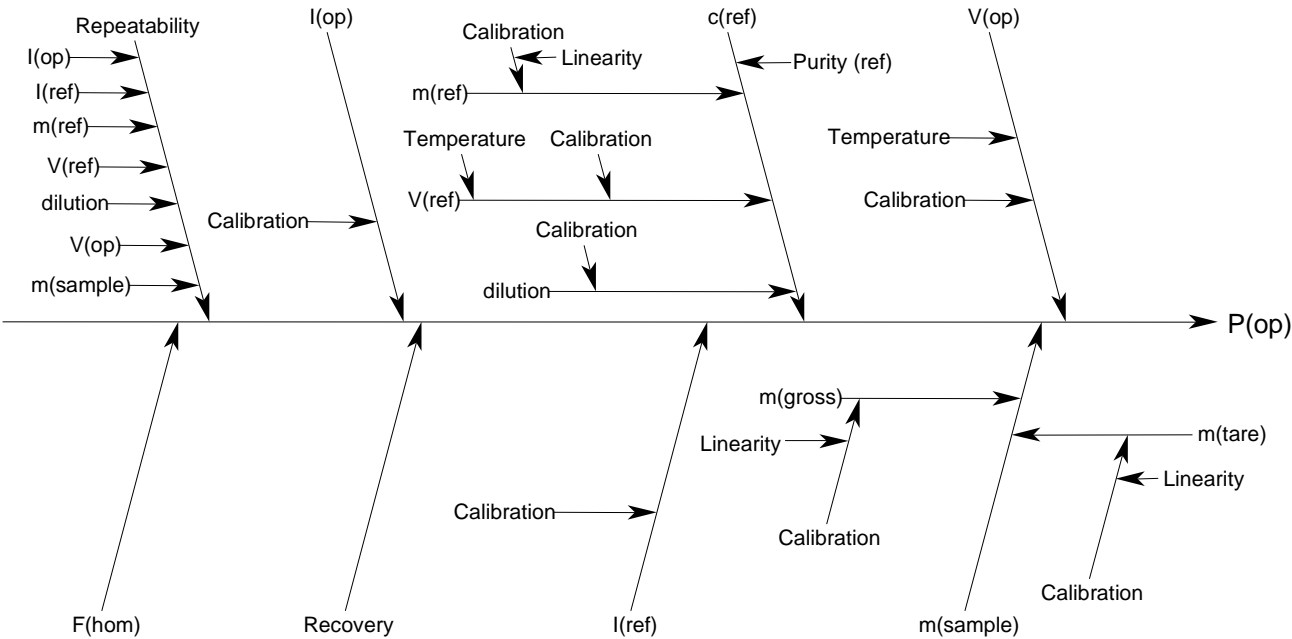
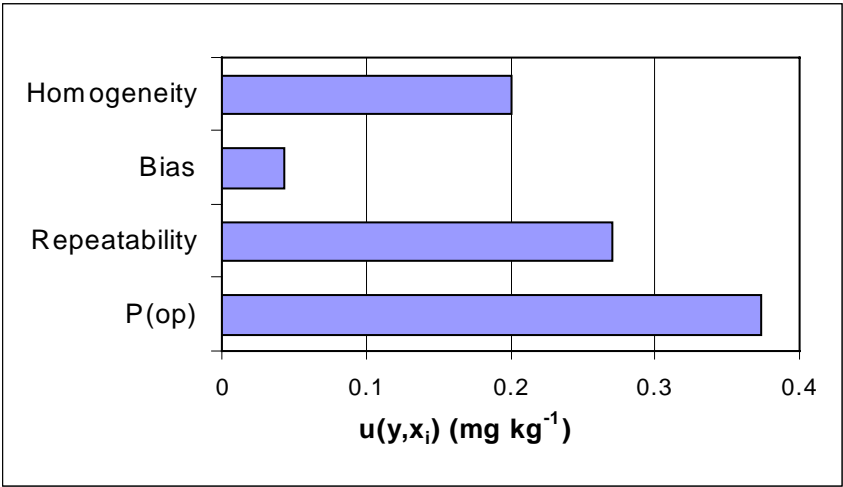


Figure A4.3: Uncertainties in pesticide analysis



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A4.5

Example A4: Determination of organophosphorus pesticides in bread. Detailed discussion.

A4.1 Introduction

This example illustrates the way in which in-house validation data can be used to quantify the measurement uncertainty. The aim of the measurement is to determine the amount of an organophosphorus pesticides residue in bread. The validation scheme and experiments establish traceability by measurements on spiked samples. It is assumed the uncertainty due to any difference in response of the measurement to the spike and the analyte in the sample is small compared with the total uncertainty on the result.

A4.2 Step 1: Specification

The specification of the measurand for more extensive analytical methods is best done by a comprehensive description of the different stages of the analytical method and by providing the equation of the measurand.

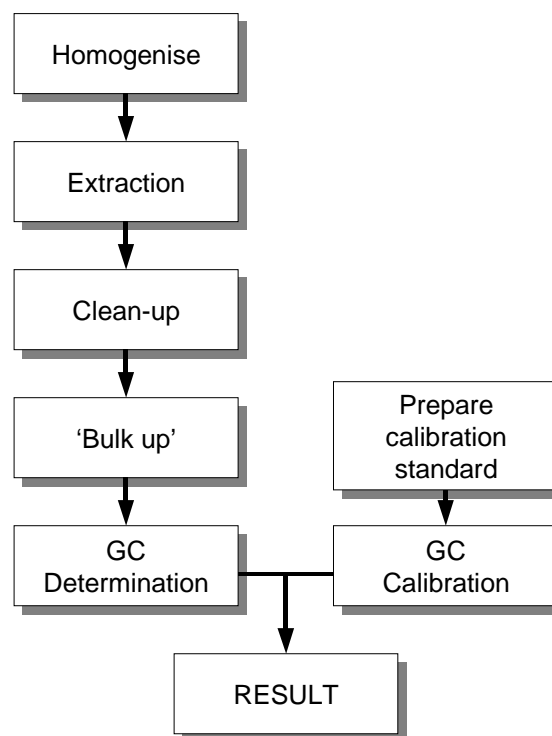
Procedure

The measurement procedure is illustrated schematically in Figure A4.4. The separate stages are:

- i) Homogenisation: The complete sample is divided into small (approx. 2 cm) fragments, a random selection is made of about 15 of these, and the sub-sample homogenised. Where extreme inhomogeneity is suspected proportional sampling is used before blending.
- ii) Weighing of sub-sampling for analysis gives mass m_{sample}
- iii) Extraction: Quantitative extraction of the analyte with organic solvent, decanting and drying through a sodium sulphate columns, and concentration of the extract using a Kuderna-Danish apparatus.
- iv) Liquid-liquid extraction:
- v) Acetonitrile/hexane liquid partition, washing the acetonitrile extract with hexane, drying the hexane layer through sodium sulphate column.

- vi) Concentration of the washed extract by gas blown-down of extract to near dryness.
- vii) Dilution to standard volume V_{op} (approx. 2 ml) in a 10 ml graduated tube.
- viii) Measurement: Injection and GC measurement of 5 μl of sample extract to give the peak intensity I_{op} .
- ix) Preparation of an approximately 5 $\mu\text{g ml}^{-1}$ standard (actual mass concentration c_{ref}).
- x) GC calibration using the prepared standard and injection and GC measurement of 5 μl of the standard to give a reference peak intensity I_{ref} .

Figure A4.4: Organophosphorus pesticides analysis



Calculation

The mass concentration c_{op} in the final sample is given by

$$c_{op} = c_{ref} \cdot \frac{I_{op}}{I_{ref}} \quad \mu\text{g ml}^{-1}$$

and the estimate P_{op} of the level of pesticide in the bulk sample (in mg kg^{-1}) is given by

$$P_{op} = \frac{c_{op} \cdot V_{op}}{Rec \cdot m_{sample}} \quad \text{mg kg}^{-1}$$

or, substituting for c_{op} ,

$$P_{op} = \frac{I_{op} \cdot c_{ref} \cdot V_{op}}{I_{ref} \cdot Rec \cdot m_{sample}} \quad \text{mg kg}^{-1}$$

where

- P_{op} :Level of pesticide in the sample [mg kg^{-1}]
 I_{op} :Peak intensity of the sample extract
 c_{ref} :Mass concentration of the reference standard [$\mu\text{g ml}^{-1}$]
 V_{op} :Final volume of the extract [ml]
 I_{ref} :Peak intensity of the reference standard
 Rec :Recovery
 m_{sample} :Mass of the investigated sub-sample [g]

Scope

The analytical method is applicable to a small range of chemically similar pesticides at levels between 0.01 and 2 mg kg^{-1} with different kinds of bread as matrix.

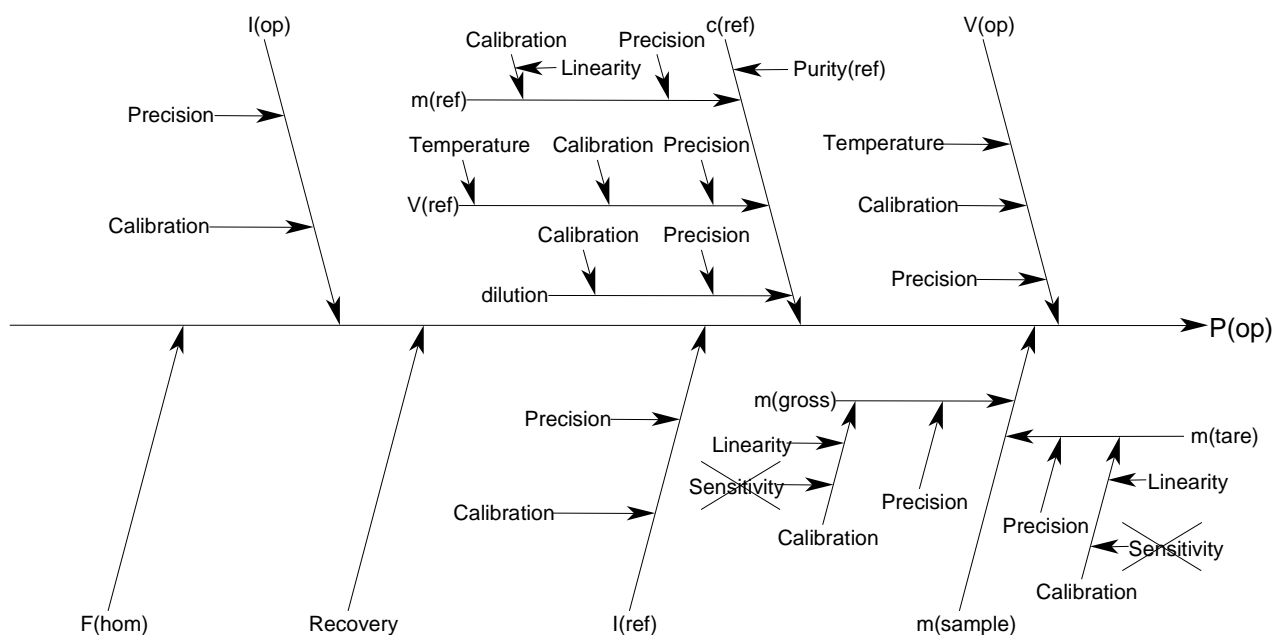
A4.3 Step 2: Identifying and analysing uncertainty sources

The identification of all relevant uncertainty sources for such a complex analytical procedure is best done by drafting a cause and effect diagram. The parameters in the equation of the measurand are represented by the main branches of the diagram. Further factors are added to the diagram, considering each step in the analytical procedure (A4.2), until the contributory factors become sufficiently remote.

The sample inhomogeneity is not a parameter in the original equation of the measurand, but it appears to be a significant effect in the analytical procedure. A new branch, $F(\text{hom})$, representing the sample inhomogeneity is accordingly added to the cause and effect diagram (Figure A4.5).

Finally, the uncertainty branch due to the inhomogeneity of the sample has to be included in the calculation of the measurand. To show the effect of uncertainties arising from that source clearly, it is useful to write

Figure A4.5: Cause and effect diagram with added main branch for sample inhomogeneity



$$P_{op} = F_{hom} \cdot \frac{I_{op} \cdot c_{ref} \cdot V_{op}}{I_{ref} \cdot Rec \cdot m_{sample}} \quad [\text{mg kg}^{-1}]$$

where F_{hom} is a correction factor assumed to be unity in the original calculation. This makes it clear that the uncertainties in the correction factor must be included in the estimation of the overall uncertainty. The final expression also shows how the uncertainty will apply.

NOTE: Correction factors: This approach is quite general, and may be very valuable in highlighting hidden assumptions. In principle, every measurement has associated with it such correction factors, which are normally assumed unity. For example, the uncertainty in c_{op} can be expressed as a standard uncertainty for c_{op} , or as the standard uncertainty which represents the uncertainty in a correction factor. In the latter case, the value is identically the uncertainty for c_{op} expressed as a relative standard deviation.

A4.4 Step 3: Quantifying uncertainty components

In accordance with section 7.7., the quantification of the different uncertainty components utilises data from the in-house development and validation studies:

- The best available estimate of the overall run to run variation of the analytical process.

- The best possible estimation of the overall bias (Rec) and its uncertainty.
- Quantification of any uncertainties associated with effects incompletely accounted for the overall performance studies.

Some rearrangement the cause and effect diagram is useful to make the relationship and coverage of these input data clearer (Figure A4.6).

NOTE: In normal use, samples are run in small batches, each batch including a calibration set, a recovery check sample to control bias and random duplicate to check precision. Corrective action is taken if these checks show significant departures from the performance found during validation. This basic QC fulfils the main requirements for use of the validation data in uncertainty estimation for routine testing.

Having inserted the extra effect 'Repeatability' into the cause and effect diagram, the implied model for calculating P_{op} becomes

$$P_{op} = F_{hom} \cdot \frac{I_{op} \cdot c_{ref} \cdot V_{op}}{I_{ref} \cdot Rec \cdot m_{sample}} \cdot F_{Rep} \quad \text{mg kg}^{-1}$$

Eq. A4.1

That is, the repeatability is treated as a multiplicative factor F_{Rep} like the homogeneity. This form is chosen for convenience in calculation, as will be seen below.

Figure A4.6: Cause and effect diagram after rearrangement to accommodate the data of the validation study

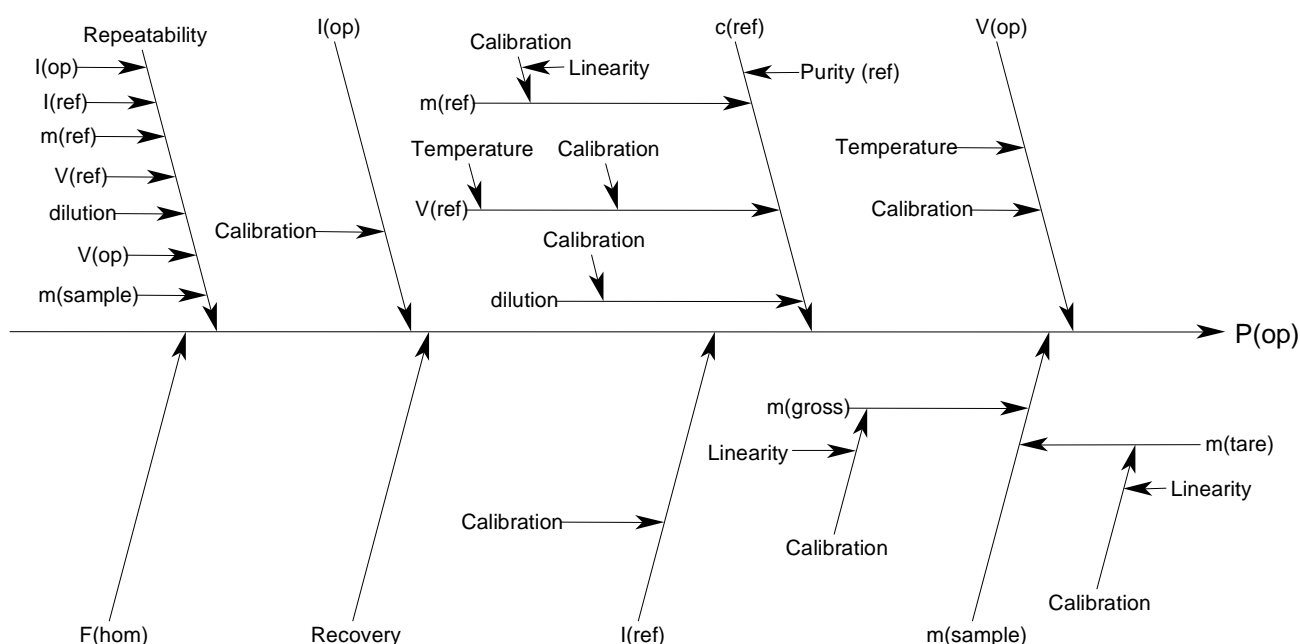


Table A4.2: Results of duplicate pesticide analysis

Residue	D1 [mg kg ⁻¹]	D2 [mg kg ⁻¹]	Mean [mg kg ⁻¹]	Difference D1-D2	Difference/ mean
Malathion	1.30	1.30	1.30	0.00	0.000
Malathion	1.30	0.90	1.10	0.40	0.364
Malathion	0.57	0.53	0.55	0.04	0.073
Malathion	0.16	0.26	0.21	-0.10	-0.476
Malathion	0.65	0.58	0.62	0.07	0.114
Pirimiphos Methyl	0.04	0.04	0.04	0.00	0.000
Chlorpyrifos Methyl	0.08	0.09	0.085	-0.01	-0.118
Pirimiphos Methyl	0.02	0.02	0.02	0.00	0.000
Chlorpyrifos Methyl	0.01	0.02	0.015	-0.01	-0.667
Pirimiphos Methyl	0.02	0.01	0.015	0.01	0.667
Chlorpyrifos Methyl	0.03	0.02	0.025	0.01	0.400
Chlorpyrifos Methyl	0.04	0.06	0.05	-0.02	-0.400
Pirimiphos Methyl	0.07	0.08	0.75	-0.10	-0.133
Chlorpyrifos Methyl	0.01	0.01	0.10	0.00	0.000
Pirimiphos Methyl	0.06	0.03	0.045	0.03	0.667

The evaluation of the different effects is now considered.

1. Precision study

The overall run to run variation (precision) of the analytical procedure was performed with a number of duplicate tests (same homogenised sample, complete extraction/determination procedure) for typical organophosphorus pesticides found in different bread samples. The results are collected in Table A4.2.

The normalised difference data (the difference divided by the mean) provides a measure of the overall run to run variability. To obtain the estimated relative standard uncertainty for single determinations, the standard deviation of the normalised differences is taken and divided by $\sqrt{2}$ to correct from a standard deviation for pairwise differences to the standard uncertainty for the single values. This gives a value for the standard uncertainty due to run to run variation of the overall analytical process, including run to run recovery variation but excluding homogeneity effects, of $0.382/\sqrt{2} = 0.27$

NOTE: At first sight, it may seem that duplicate tests provide insufficient degrees of freedom. But it

is not the goal to obtain very accurate numbers for the precision of the analytical process for one specific pesticide in one special kind of bread. It is more important in this study to test a wide variety of different materials and sample levels, giving a representative selection of typical organophosphorus pesticides. This is done in the most efficient way by duplicate tests on many materials, providing (for the repeatability estimate) approximately one degree of freedom for each material studied in duplicate.

2. Bias study

The bias of the analytical procedure was investigated during the in-house validation study using spiked samples (homogenised samples were split and one portion spiked). Table A4.3 collects the results of a long term study of spiked samples of various types.

The relevant line (marked with grey colour) is the "bread" entry line, which shows a mean recovery for forty-two samples of 90%, with a standard deviation (s) of 28%. The standard uncertainty was calculated as the standard deviation of the mean $u(\overline{Rec}) = 0.28/\sqrt{42} = 0.0432$.

A significance test is used to determine whether the mean recovery is significantly different from

Table A4.3: Results of pesticide recovery studies

Substrate	Residue Type	Conc. [mg kg ⁻¹]	N ¹⁾	Mean ²⁾ [%]	s ²⁾ [%]
Waste Oil	PCB	10.0	8	84	9
Butter	OC	0.65	33	109	12
Compound Animal Feed I	OC	0.325	100	90	9
Animal & Vegetable Fats I	OC	0.33	34	102	24
Brassicas 1987	OC	0.32	32	104	18
Bread	OP	0.13	42	90	28
Rusks	OP	0.13	30	84	27
Meat & Bone Feeds	OC	0.325	8	95	12
Maize Gluten Feeds	OC	0.325	9	92	9
Rape Feed I	OC	0.325	11	89	13
Wheat Feed I	OC	0.325	25	88	9
Soya Feed I	OC	0.325	13	85	19
Barley Feed I	OC	0.325	9	84	22

(1) The number of experiments carried out

(2) The mean and sample standard deviation s are given as percentage recoveries.

1.0. The test statistic t is calculated using the following equation

$$t = \frac{|1 - \overline{Rec}|}{u(\overline{Rec})} = \frac{(1 - 0.9)}{0.0432} = 2.315$$

This value is compared with the 2-tailed critical value t_{crit} for $n-1$ degrees of freedom at 95% confidence (where n is the number of results used to estimate \overline{Rec}). If t is greater or equal than the critical value t_{crit} than \overline{Rec} is significantly different from 1.

$$t = 2.31 \geq t_{crit;41} \cong 2.021$$

In this example a correction factor ($1/\overline{Rec}$) is being applied and therefore \overline{Rec} is explicitly included in the calculation of the result.

3. Other sources of uncertainty

The cause and effect diagram in Figure A4.7 shows which other sources of uncertainty are (1) adequately covered by the precision data, (2) covered by the recovery data or (3) have to be further examined and eventually considered in the calculation of the measurement uncertainty.

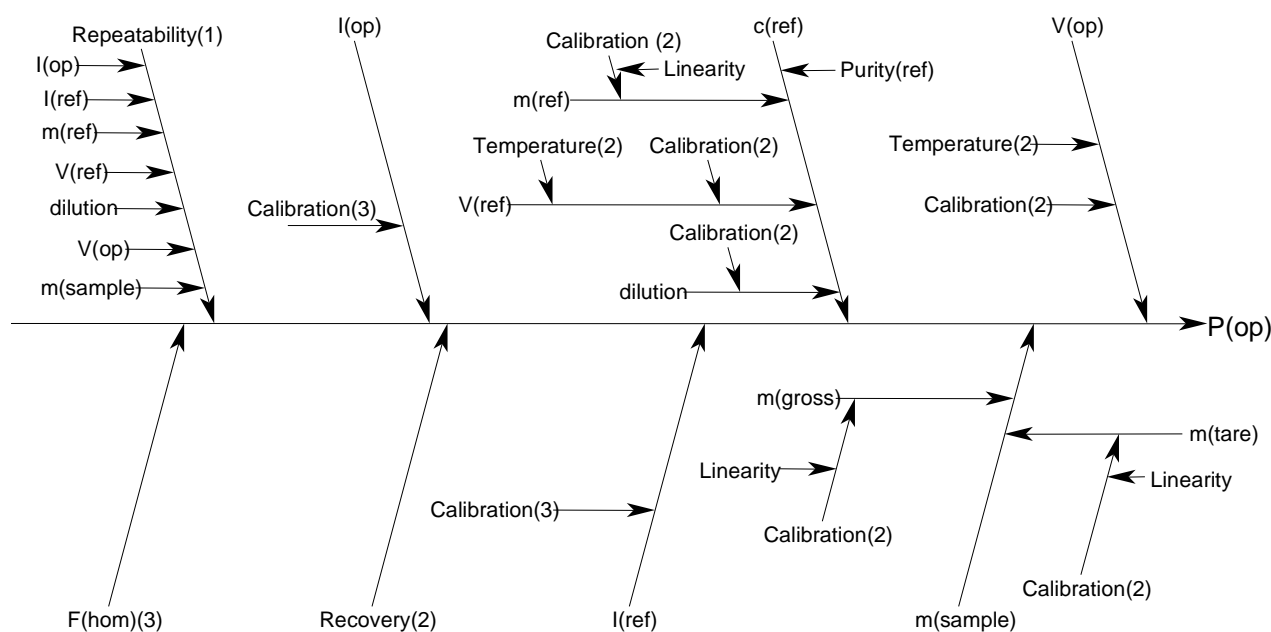
All balances and the important volumetric measuring devices are under regular control. Precision and recovery studies take into account

the influence of the calibration of the different volumetric measuring devices because during the investigation various volumetric flasks and pipettes have been used. The extensive variability studies, which lasted for more than half a year, also cover influences of the environmental temperature on the result. This leaves only the reference material purity, possible nonlinearity in GC response (represented by the 'calibration' terms for I_{ref} and I_{op} in the diagram), and the sample homogeneity as additional components requiring study.

The purity of the reference standard is given by the manufacturer as 99.53% $\pm 0.06\%$. The purity is potential an additional uncertainty source with a standard uncertainty of $0.0006/\sqrt{3} = 0.00035$ (Rectangular distribution). But the contribution is so small (compared, for example, to the precision estimate) that it is clearly safe to neglect this contribution.

Linearity of response to the relevant organophosphorus pesticides within the given concentration range is established during validation studies. In addition, with multi-level studies of the kind indicated in Table A4.2 and Table A4.3, nonlinearity would contribute to the observed precision. No additional allowance is

Figure A4.7: Evaluation of other sources of uncertainty



- (1) Repeatability (F_{Rep} in equation A4.1) considered during the variability investigation of the analytical procedure.
- (2) Considered during the bias study of the analytical procedure.
- (3) To be considered during the evaluation of the other sources of uncertainty.

required. The in-house validation study has proven that this is not the case.

The homogeneity of the bread sub-sample is the last remaining other uncertainty source. No literature data were available on the distribution of trace organic components in bread products, despite an extensive literature search (at first sight this is surprising, but most food analysts attempt homogenisation rather than evaluate inhomogeneity separately). Nor was it practical to measure homogeneity directly. The contribution has therefore been estimated on the basis of the sampling method used.

To aid the estimation, a number of feasible pesticide residue distribution scenarios were considered, and a simple binomial statistical distribution used to calculate the standard uncertainty for the total included in the analysed sample (see section A4.6). The scenarios, and the calculated relative standard uncertainties in the amount of pesticide in the final sample, were:

- Scenario (a) Residue distributed on the top surface only: 0.58.
- Scenario (b) Residue distributed evenly over the surface only: 0.20.
- Scenario (c) Residue distributed evenly

through the sample, but reduced in concentration by evaporative loss or decomposition close to the surface: 0.05-0.10 (depending on the "surface layer" thickness).

Scenario (a) is specifically catered for by proportional sampling or complete homogenisation: It would arise in the case of decorative additions (whole grains) added to one surface. Scenario (b) is therefore considered the likely worst case. Scenario (c) is considered the most probable, but cannot be readily distinguished from (b). On this basis, the value of 0.20 was chosen.

NOTE: For more details on modelling inhomogeneity see the last section of this example.

A4.5 Step 4: Calculating the combined standard uncertainty

During the in-house validation study of the analytical procedure the repeatability, the bias and all other feasible uncertainty sources had been thoroughly investigated. Their values and uncertainties are collected in Table A4.4.

The relative values are combined because the model (equation A4.1) is entirely multiplicative:

Table A4.4: Uncertainties in pesticide analysis

Description	Value x	Standard uncertainty $u(x)$	Relative standard uncertainty $u(x)$	Remark
Repeatability(1)	1.0	0.27	0.27	Duplicate tests of different types of samples
Bias (<i>Rec</i>) (2)	0.9	0.043	0.048	Spiked samples
Other sources (3) (Homogeneity)	1.0	0.2	0.2	Estimations founded on model assumptions
P_{op}	--	--	0.34	Relative standard uncertainty

$$\frac{u_c(P_{op})}{P_{op}} = \sqrt{0.27^2 + 0.048^2 + 0.2^2} = 0.34$$

$$\Rightarrow u_c(P_{op}) = 0.34 \times P_{op}$$

The spreadsheet for this case (Table A4.5) takes the form shown in Table A4.5. Note that the spreadsheet calculates an absolute value uncertainty (0.377) for a nominal corrected result of 1.1111, giving a value of $0.373/1.11=0.34$.

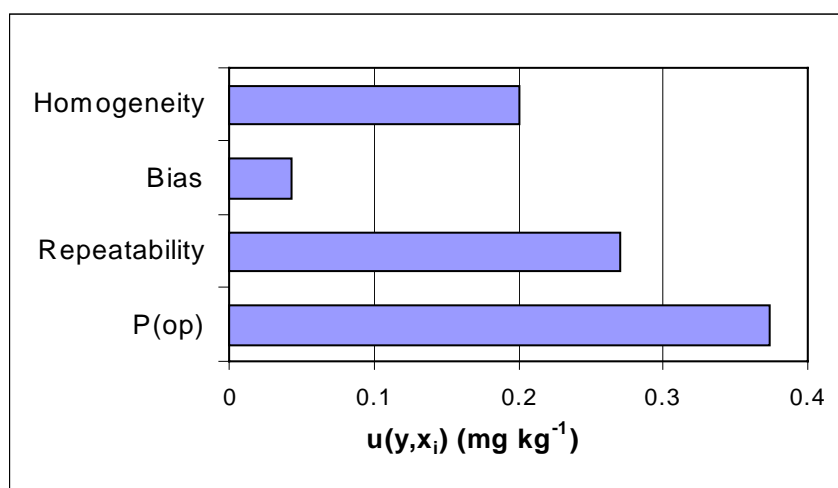
The relative sizes of the three different contributions can be compared by employing a histogram. Figure A4.8 shows the values $|u(y, x_i)|$ taken from Table A4.5.

The repeatability is the largest contribution to the measurement uncertainty. Since this component is derived from the overall variability in the method, further experiments would be needed to show where improvements could be made. For example, the uncertainty could be reduced significantly by homogenising the whole loaf before taking a sample.

The expanded uncertainty $U(P_{op})$ is calculated by multiplying the combined standard uncertainty with a coverage factor of 2 to give:

$$U(P_{op}) = 0.34 \times P_{op} \times 2 = 0.68 \times P_{op}$$

Figure A4.8: Uncertainties in pesticide analysis



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A4.5

Table A4.5: Uncertainties in pesticide analysis

	A	B	C	D	E
1			Repeatability	Bias	Homogeneity
2		value	1.0	0.9	1.0
3		uncertainty	0.27	0.043	0.2
4					
5	Repeatability	1.0	1.27	1.0	1.0
6	Bias	0.9	0.9	0.943	0.9
7	Homogeneity	1.0	1.0	1.0	1.2
8					
9	P_{op}	1.1111	1.4111	1.0604	1.333
10	$u(y, x_i)$		0.30	-0.0507	0.222
11	$u(y)^2, u(y, x_i)^2$	0.1420	0.09	0.00257	0.04938
12					
13	$u(P_{op})$	0.377	(0.377/1.111 = 0.34 as a relative standard uncertainty)		

The values of the parameters are entered in the second row from C2 to E2. Their standard uncertainties are in the row below (C3:E3). The spreadsheet copies the values from C2-E2 into the second column from B5 to B7. The result using these values is given in B9 (=B5×B7/B6, based on equation A4.1). C5 shows the value of the repeatability from C2 plus its uncertainty given in C3. The result of the calculation using the values C5:C7 is given in C9. The columns D and E follow a similar procedure. The values shown in the row 10 (C10:E10) are the differences of the row (C9:E9) minus the value given in B9. In row 11 (C11:E11) the values of row 10 (C10:E10) are squared and summed to give the value shown in B11. B13 gives the combined standard uncertainty, which is the square root of B11.

A4.6 Special aspect: Modelling inhomogeneity for organophosphorus pesticide uncertainty

Assuming that all of the material of interest in a sample can be extracted for analysis irrespective of its state, the worst case for inhomogeneity is the situation where some part or parts of a sample contain all of the substance of interest. A more general, but closely related, case is that in which two levels, say L_1 and L_2 of the material are present in different parts of the whole sample. The effect of such inhomogeneity in the case of random sub-sampling can be estimated using binomial statistics. The values required are the mean μ and the standard deviation σ of the amount of material in n equal portions selected randomly after separation.

These values are given by

$$\mu = n \cdot (p_1 l_1 + p_2 l_2) \Rightarrow$$

$$\mu = np_1 \cdot (l_1 - l_2) + nl_2 \quad [1]$$

$$\sigma^2 = np_1 \cdot (1 - p_1) \cdot (l_1 - l_2)^2 \quad [2]$$

where l_1 and l_2 are the amount of substance in portions from regions in the sample containing total fraction L_1 and L_2 respectively, of the total amount X , and p_1 and p_2 are the probabilities of selecting portions from those regions (n must be small compared to the total number of portions from which the selection is made).

The figures shown above were calculated as follows, assuming that a typical sample loaf is approximately $12 \times 12 \times 24$ cm, using a portion size of $2 \times 2 \times 2$ cm (total of 432 portions) and assuming 15 such portions are selected at random and homogenised.

Scenario (a)

The material is confined to a single large face (the top) of the sample. L_2 is therefore zero as is l_2 ; and $L_1=1$. Each portion including part of the top surface will contain an amount l_1 of the material. For the dimensions given, clearly one in six (2/12) of the portions meets this criterion, p_1 is

therefore 1/6, or 0.167, and l_1 is $X/72$ (*i.e.* there are 72 "top" portions).

This gives

$$\mu = 15 \times 0.167 \times l_1 = 2.5 l_1$$

$$\sigma^2 = 15 \times 0.167 \times (1 - 0.17) \times l_1^2 = 2.08 l_1^2$$

$$\Rightarrow \sigma = \sqrt{2.08 l_1^2} = 1.44 l_1$$

$$\Rightarrow RSD = \frac{\sigma}{\mu} = 0.58$$

NOTE: To calculate the level X in the entire sample, μ is multiplied back up by $432/15$, giving a mean estimate of X of

$$X = \frac{432}{15} \times 2.5 \times l_1 = 72 \times \frac{X}{72} = X$$

This result is typical of random sampling; the expectation value of the mean is exactly the mean value of the population. For random sampling, there is thus no contribution to overall uncertainty other than the run to run variability, expressed as σ or RSD here.

Scenario (b)

The material is distributed evenly over the whole surface. Following similar arguments and assuming that all surface portions contain the same amount l_1 of material, l_2 is again zero, and p_1 is, using the dimensions above, given by

$$p_1 = \frac{(12 \times 12 \times 24) - (8 \times 8 \times 20)}{(12 \times 12 \times 24)} = 0.63$$

i.e. p_1 is that fraction of sample in the "outer" 2 cm. Using the same assumptions then $l_1 = X/272$.

NOTE: The change in value from scenario (a)

This gives:

$$\mu = 15 \times 0.63 \times l_1 = 9.5 l_1$$

$$\sigma^2 = 15 \times 0.63 \times (1 - 0.63) \times l_1^2 = 3.5 l_1^2$$

$$\Rightarrow \sigma = \sqrt{3.5 l_1^2} = 1.87 l_1$$

$$\Rightarrow RSD = \frac{\sigma}{\mu} = 0.2$$

Scenario (c)

The amount of material near the surface is reduced to zero by evaporative or other loss. This

case can be examined most simply by considering it as the inverse of scenario (b), with $p_1=0.37$ and l_1 equal to $X/160$. This gives

$$\mu = 15 \times 0.37 \times l_1 = 5.6 l_1$$

$$\sigma^2 = 15 \times 0.37 \times (1 - 0.37) \times l_1^2 = 3.5 l_1^2$$

$$\Rightarrow \sigma = \sqrt{3.5 \times l_1^2} = 1.87 l_1$$

$$\Rightarrow RSD = \frac{\sigma}{\mu} = 0.33$$

However, if the loss extends to a depth less than the size of the portion removed, as would be expected, each portion contains some material l_1 and l_2 would therefore both be non-zero. Taking the case where all outer portions contain 50% "centre" and 50% "outer" parts of the sample

$$l_1 = 2 \times l_2 \Rightarrow l_1 = X/296$$

$$\begin{aligned} \mu &= 15 \times 0.37 \times (l_1 - l_2) + 15 \times l_2 \\ &= 15 \times 0.37 \times l_2 + 15 \times l_2 = 20.6 l_2 \end{aligned}$$

$$\sigma^2 = 15 \times 0.37 \times (1 - 0.37) \times (l_1 - l_2)^2 = 3.5 l_2^2$$

giving an RSD of $1.87/20.6 = 0.09$

In the current model, this corresponds to a depth of 1 cm through which material is lost. Examination of typical bread samples shows crust thickness typically of 1 cm or less, and taking this to be the depth to which the material of interest is lost (crust formation itself inhibits lost below this depth), it follows that realistic variants on scenario (c) will give values of σ/μ not above 0.09.

NOTE: In this case, the reduction in uncertainty arises because the inhomogeneity is on a smaller scale than the portion taken for homogenisation. In general, this will lead to a reduced contribution to uncertainty. It follows that no additional modelling need be done for cases where larger numbers of small inclusions (such as grains incorporated in the bulk of a loaf) contain disproportionate amounts of the material of interest. Provided that the probability of such an inclusion being incorporated into the portions taken for homogenisation is large enough, the contribution to uncertainty will not exceed any already calculated in the scenarios above.

Example A5: Determination of Cadmium Release from Ceramic Ware by Atomic Absorption Spectrometry

Summary

Goal

The amount of released cadmium from ceramic ware is determined using atomic absorption spectrometry. The procedure employed is the empirical method BS 6748.

Measurement procedure

The different stages in determining the amount of cadmium released from ceramic ware are given in the flow chart (Figure A5.1).

Measurand:

$$r = \frac{c_0 \cdot V_L}{a_V} \cdot d \cdot f_{acid} \cdot f_{time} \cdot f_{temp} \quad \text{mg dm}^{-2}$$

The variables are described in Table A5.1.

Identification of the uncertainty sources:

The relevant uncertainty sources are shown in the cause and effect diagram at Figure A5.2.

Quantification of the uncertainty sources:

The sizes of the different contributions are given in Table A5.1 and shown diagrammatically in Figure A5.2

Figure A5.1: Extractable metal procedure

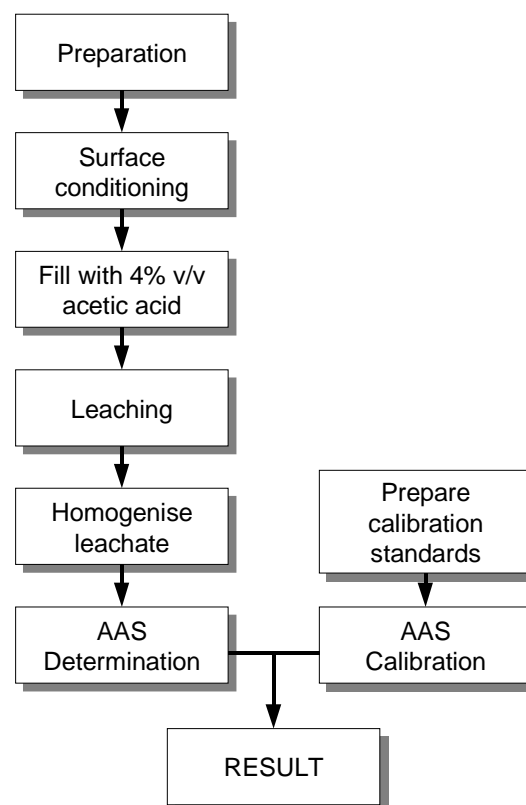


Table A5.1: Uncertainties in extractable cadmium determination

	Description	Value x	Standard uncertainty $u(x)$	Relative standard uncertainty $u(x)/x$
c_0	Content of cadmium in the extraction solution	0.26 mg l ⁻¹	0.018 mg l ⁻¹	0.069
d	Dilution factor (if used)	1.0 ^{Note 1}	0 ^{Note 1}	0 ^{Note 1}
V_L	Volume of the leachate	0.332 l	0.0018 l	0.0054
a_V	Surface area of the vessel	2.37 dm ²	0.06 dm ²	0.025
f_{acid}	Influence of the acid concentration	1.0	0.0008	0.0008
f_{time}	Influence of the duration	1.0	0.001	0.001
f_{temp}	Influence of temperature	1.0	0.06	0.06
r	Mass of cadmium leached per unit area	0.036 mg dm ⁻²	0.0033 mg dm ⁻²	0.09

Note 1: No dilution was applied in the present example; d is accordingly exactly 1.0

Figure A5.2: Uncertainty sources in leachable cadmium determination

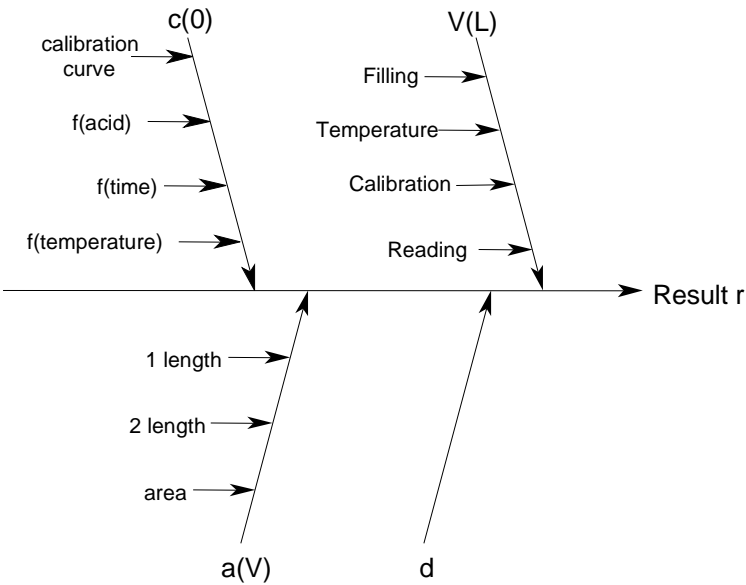
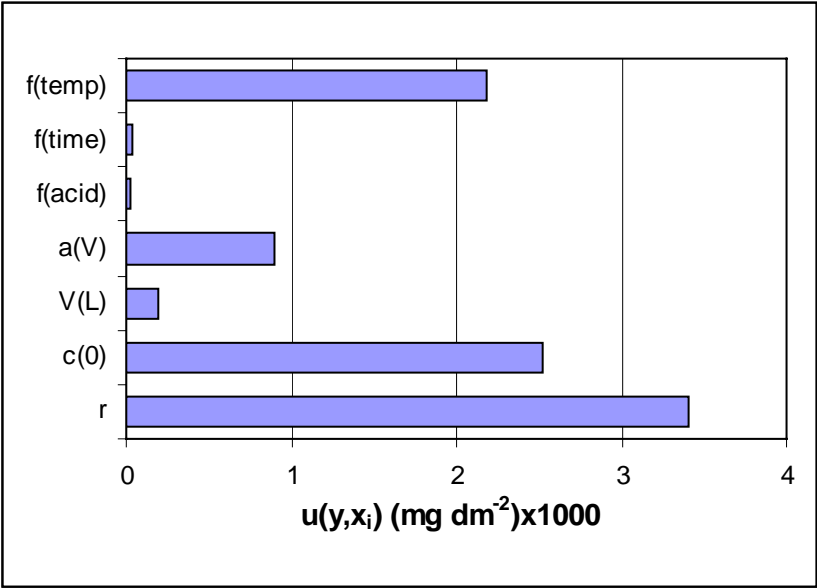


Figure A5.3: Uncertainties in leachable Cd determination



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A5.4

Example A5: Determination of cadmium release from ceramic ware by atomic absorption spectrometry. Detailed discussion.

A5.1 Introduction

This example demonstrates the uncertainty evaluation of an empirical method; in this case (BS 6748), the determination of metal release from ceramic ware, glassware, glass-ceramic ware and vitreous enamel ware. The test is used to determine by atomic absorption spectroscopy (AAS) the amount of lead or cadmium leached from the surface of ceramic ware by a 4% (v/v) aqueous solution of acetic acid. The results obtained with this analytical method are only expected to be comparable with other results obtained by the same method.

A5.2 Step 1: Specification

The complete procedure is given in British Standard BS 6748:1986 “Limits of metal release from ceramic ware, glass ware, glass ceramic ware and vitreous enamel ware” and this forms the specification for the measurand. Only a general description is given here (right).

A5.2.1 Apparatus and Reagent specifications

The reagent specifications affecting the uncertainty study are:

- A freshly prepared solution of 4% v/v glacial acetic acid in water, made up by dilution of 40 ml glacial acetic to 1 l.
- A (1000 ± 1) mg l⁻¹ standard lead solution in 4% (v/v) acetic acid.
- A (500 ± 0.5) mg l⁻¹ standard cadmium solution in 4% (v/v) acetic acid.

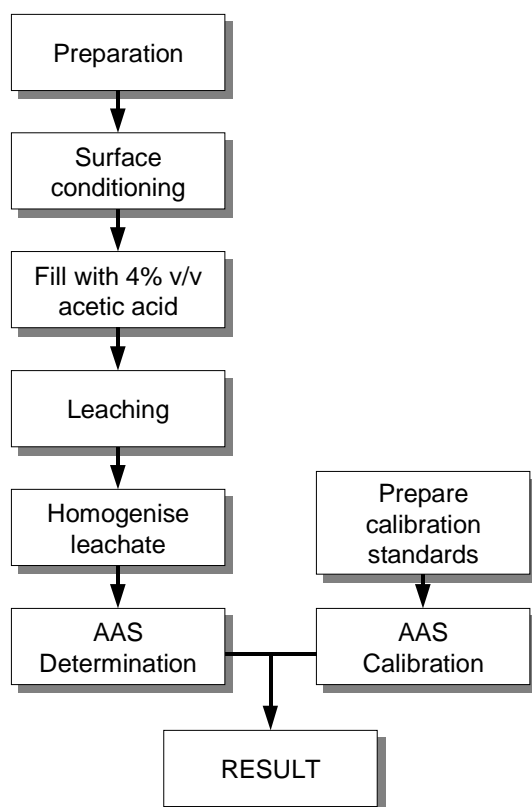
Laboratory glassware is required to be of at least class B and incapable of releasing detectable levels of lead or cadmium in 4% acetic acid during the test procedure. The atomic absorption spectrophotometer is required to have detection limits of at most 0.2 mg l⁻¹ for lead and 0.02 mg l⁻¹ for cadmium.

A5.2.2 Procedure

The general procedure is illustrated schematically in Figure A5.4. The specifications affecting the uncertainty estimation are:

- The sample is conditioned to (22 ± 2) °C. Where appropriate (‘category 1’ articles), the surface area of the article is determined. For this example, a surface area of 2.37 dm² was obtained (Table A5.1 and Table A5.3 include the experimental values for the example).
- The conditioned sample is filled with 4% v/v acid solution at (22 ± 2) °C to within 1 mm from the overflow point, measured from the upper rim of the sample, or to within 6 mm from the extreme edge of a sample with a flat or sloping rim.
- The quantity of 4% v/v acetic acid required or used is recorded to an accuracy of $\pm 2\%$ (in this example, 332 ml acetic acid was used).
- The sample is allowed to stand at (22 ± 2) °C for 24 hours (in darkness if cadmium is determined) with due precaution to prevent evaporation loss.
- After standing, the solution is stirred sufficiently for homogenisation, and a test portion removed, diluted by a factor d if necessary, and analysed by AA, using

Figure A5.4: Extractable metal procedure



appropriate wavelengths and, in this example, a least squares calibration curve.

- vi) The result is calculated (see below) and reported as the amount of lead and/or cadmium in the total volume of the extracting solution, expressed in milligrams of lead or cadmium per square decimetre of surface area for category 1 articles or milligrams of lead or cadmium per litre of the volume for category 2 and 3 articles.

NOTE: Complete copies of BS 6748:1986 can be obtained by post from BSI customer services, 389 Chiswick High Road, London W4 4AL England ☎ +44 (0) 208 996 9001.

A5.3 Step 2: Identity and analysing uncertainty sources

Step 1 describes an 'empirical method'. If such a method is used within its defined field of application, the bias of the method is defined as zero. Therefore bias estimation relates to the laboratory performance and not to the bias intrinsic to the method. Because no reference material certified for this standardised method is available, overall control of bias is related to the control of method parameters influencing the result. Such influence quantities are time, temperature, mass and volumes, *etc.*

The concentration c_0 of lead or cadmium in the acetic acid after dilution is determined by atomic absorption spectrometry and calculated using

$$c_0 = \frac{(A_0 - B_0)}{B_1} \quad \text{mg l}^{-1}$$

where

c_0 :concentration of lead or cadmium in the extraction solution [mg l⁻¹]

A_0 :absorbance of the metal in the sample extract

B_0 :intercept of the calibration curve

B_1 :slope of the calibration curve

For vessels that can be filled, the result r' is then

$$r' = c_0 \cdot d$$

where d is the dilution factor employed. Otherwise, the empirical method calls for the result to be expressed as mass r of lead or cadmium leached per unit area. r is given by

$$r = \frac{c_0 \cdot V_L}{a_v} \cdot d = \frac{V_L \cdot (A_0 - B_0)}{a_v \cdot B_1} \cdot d \quad \text{mg dm}^{-2}$$

where the additional parameters are

r :mass of Cd or Pb leached per unit area [mg dm⁻²]

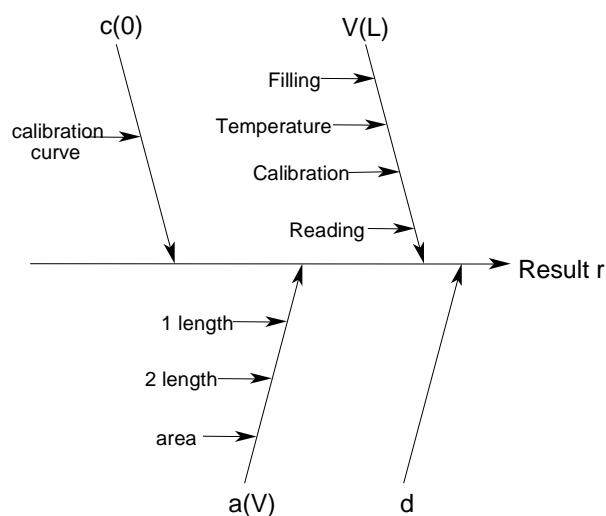
V_L :the volume of the leachate [l]

a_v :the surface area of the vessel [dm²]

d :factor by which the sample was diluted

The first part of the above equation of the measurand is used to draft the basic cause and effect diagram (Figure A5.5).

Figure A5.5:Initial cause and effect diagram



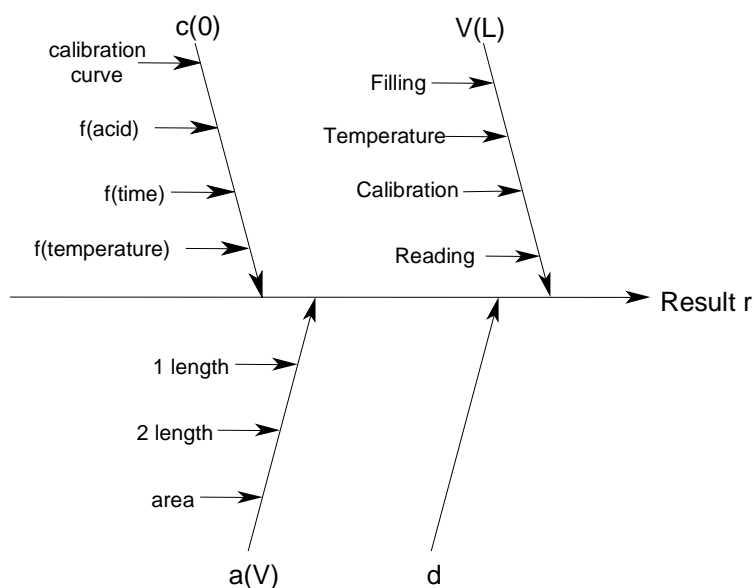
There is no reference material certified for this empirical method with which to assess the laboratory performance. All the feasible influence quantities, such as temperature, time of the leaching process and acid concentration therefore have to be considered. To accommodate the additional influence quantities the equation is expanded by the respective correction factors leading to

$$r = \frac{c_0 \cdot V_L}{a_v} \cdot d \cdot f_{acid} \cdot f_{time} \cdot f_{temp}$$

These additional factors are also included in the revised cause and effect diagram (Figure A5.6). They are shown there as effects on c_0 .

NOTE: The latitude in temperature permitted by the standard is a case of an uncertainty arising as a result of incomplete specification of the measurand. Taking the effect of temperature into account allows estimation of the range of results which could be reported whilst complying with the empirical method as well as is practically possible. Note particularly that variations in the result caused by different operating temperatures within the range

Figure A5.6: Cause and effect diagram with added hidden assumptions (correction factors)



cannot reasonably be described as bias as they represent results obtained in accordance with the specification.

$$\frac{2.1 \times 10^{-4} \times 332 \times 2}{\sqrt{3}} = 0.08 \text{ ml}$$

A5.4 Step 3: Quantifying uncertainty sources

The aim of this step is to quantify the uncertainty arising from each of the previously identified sources. This can be done either by using experimental data or from well based assumptions.

Dilution factor d

For the current example, no dilution of the leaching solution is necessary, therefore no uncertainty contribution has to be accounted for.

Volume V_L

Filling: The empirical method requires the vessel to be filled 'to within 1 mm from the brim'. For a typical drinking or kitchen utensil, 1 mm will represent about 1% of the height of the vessel. The vessel will therefore be $99.5 \pm 0.5\%$ filled (i.e. V_L will be approximately 0.995 ± 0.005 of the vessel's volume).

Temperature: The temperature of the acetic acid has to be $22 \pm 2^\circ\text{C}$. This temperature range leads to an uncertainty in the determined volume, due to a considerable larger volume expansion of the liquid compared with the vessel. The standard uncertainty of a volume of 332 ml, assuming a rectangular temperature distribution, is

Reading: The volume V_L used is to be recorded to within 2%, in practice, use of a measuring cylinder allows an inaccuracy of about 1% (i.e. $0.01V_L$). The standard uncertainty is calculated assuming a triangular distribution.

Calibration: The volume is calibrated according to the manufacturer's specification within the range of ± 2.5 ml for a 500 ml measuring cylinder. The standard uncertainty is obtained assuming a triangular distribution.

For this example a volume of 332 ml is used and the four uncertainty components are combined accordingly

$$u(V_L) = \sqrt{\left(\frac{0.005 \times 332}{\sqrt{6}}\right)^2 + (0.08)^2 + \left(\frac{0.01 \times 332}{\sqrt{6}}\right)^2 + \left(\frac{2.5}{\sqrt{6}}\right)^2} = 1.83 \text{ ml}$$

Cadmium concentration c_0

The amount of leached cadmium is calculated using a manually prepared calibration curve. For this purpose five calibration standards, with a concentration 0.1 mg l^{-1} , 0.3 mg l^{-1} , 0.5 mg l^{-1} , 0.7 mg l^{-1} and 0.9 mg l^{-1} , were prepared from a $500 \pm 0.5 \text{ mg l}^{-1}$ cadmium reference standard. The linear least squares fitting procedure used

assumes that the uncertainties of the values of the abscissa are considerably smaller than the uncertainty on the values of the ordinate. Therefore the usual uncertainty calculation procedures for c_0 only reflect the uncertainty in the absorbance and not the uncertainty of the calibration standards, nor the inevitable correlations induced by successive dilution from the same stock. In this case, however, the uncertainty of the calibration standards is sufficiently small to be neglected.

The five calibration standards were measured three times each, providing the results in Table A5.2.

The calibration curve is given by

$$A_j = c_i \cdot B_1 + B_0$$

where

A_j : j^{th} measurement of the absorbance of the i^{th} calibration standard

c_i : concentration of the i^{th} calibration standard

B_1 : slope

B_0 : intercept

and the results of the linear least square fit are

	Value	Standard deviation
B_1	0.2410	0.0050
B_0	0.0087	0.0029

with a correlation coefficient r of 0.997. The

Table A5.2: Calibration results

Concentration [mg l ⁻¹]	1	2	3
0.1	0.028	0.029	0.029
0.3	0.084	0.083	0.081
0.5	0.135	0.131	0.133
0.7	0.180	0.181	0.183
0.9	0.215	0.230	0.216

fitted line is shown in Figure A5.7. The residual standard deviation S is 0.005486.

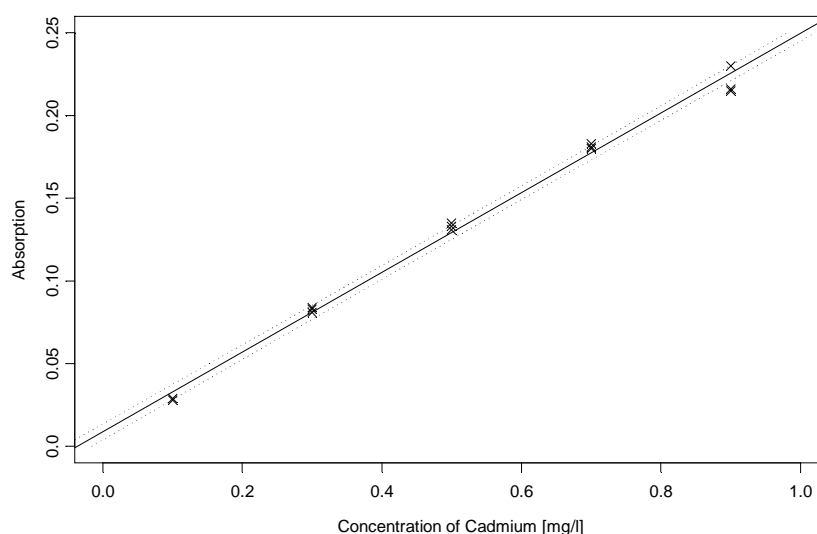
The actual leach solution was measured twice, leading to a concentration c_0 of 0.26 mg l⁻¹. The calculation of the uncertainty $u(c_0)$ associated with the linear least square fitting procedure is described in detail in Appendix E3. Therefore only a short description of the different calculation steps is given here.

$u(c_0)$ is given by

$$\begin{aligned}
 u(c_0) &= \frac{S}{B_1} \sqrt{\frac{1}{p} + \frac{1}{n} + \frac{(c_0 - \bar{c})^2}{S_{xx}}} \\
 &= \frac{0.005486}{0.241} \sqrt{\frac{1}{2} + \frac{1}{15} + \frac{(0.26 - 0.5)^2}{1.2}} \\
 \Rightarrow u(c_0) &= 0.018 \text{ mg l}^{-1}
 \end{aligned}$$

with the residual standard deviation S given by

Figure A5.7: Linear least square fit and uncertainty interval for duplicate determinations



$$S = \sqrt{\frac{\sum_{j=1}^n [A_j - (B_0 + B_1 \cdot c_j)]^2}{n-2}} = 0.005486$$

and

$$S_{xx} = \sum_{j=1}^n (c_j - \bar{c})^2 = 1.2$$

where

- B_1 :slope
 p :number of measurements to determine c_0
 n :number of measurements for the calibration
 c_0 :determined cadmium concentration of the leached solution
 \bar{c} :mean value of the different calibration standards (n number of measurements)
 i :index for the number of calibration standards
 j :index for the number of measurements to obtain the calibration curve

Area a_v

Length measurement: The total surface area of the sample vessel was calculated, from measured dimensions, to be 2.37 dm². Since the item is approximately cylindrical but not perfectly regular, measurements are estimated to be within 2 mm at 95% confidence. Typical dimensions are between 1.0 dm and 2.0 dm leading to an estimated dimensional measurement uncertainty of 1 mm (after dividing the 95% figure by 1.96). Area measurements typically require two length measurements, height and width respectively (i.e. 1.45 dm and 1.64 dm)

Area: Since the item has not a perfect geometric shape, there is also an uncertainty in any area calculation; in this example, this is estimated to contribute an additional 5% at 95% confidence.

The uncertainty contribution of the length measurement and area itself are combined in the usual way.

$$u(a_v) = \sqrt{0.01^2 + 0.01^2 + \left(\frac{0.05 \times 2.37}{1.96}\right)^2}$$

$$\Rightarrow u(a_v) = 0.06 \text{ dm}^2$$

Temperature effect f_{temp}

A number of studies of the effect of temperature on metal release from ceramic ware have been undertaken⁽¹⁻⁵⁾. In general, the temperature effect

is substantial and a near-exponential increase in metal release with temperature is observed until limiting values are reached. Only one study¹ has given an indication of effects in the range of 20-25°C. From the graphical information presented the change in metal release with temperature near 25°C is approximately linear, with a gradient of approximately 5% °C⁻¹. For the ±2°C range allowed by the empirical method this leads to a factor f_{temp} of 1±0.1. Converting this to a standard uncertainty gives, assuming a rectangular distribution:

$$u(f_{temp}) = 0.1/\sqrt{3} = 0.06$$

Time effect f_{time}

For a relatively slow process such as leaching, the amount leached will be approximately proportional to time for small changes in the time. Krinitz and Franco¹ found a mean change in concentration over the last six hours of leaching of approximately 1.8 mg l⁻¹ in 86 mg l⁻¹, that is, about 0.3%/h. For a time of (24±0.5)h c_0 will therefore need correction by a factor f_{time} of 1±(0.5×0.003) = 1±0.0015. This is a rectangular distribution leading to the standard uncertainty

$$u(f_{time}) = 0.0015/\sqrt{3} \approx 0.001.$$

Acid concentration f_{acid}

One study of the effect of acid concentration on lead release showed that changing concentration from 4 to 5% v/v increased the lead released from a particular ceramic batch from 92.9 to 101.9 mg l⁻¹, i.e. a change in f_{acid} of (101.9 – 92.9)/92.9 = 0.097 or close to 0.1. Another study, using a hot leach method, showed a comparable change (50% change in lead extracted on a change of from 2 to 6% v/v)³. Assuming this effect as approximately linear with acid concentration gives an estimated change in f_{acid} of approximately 0.1 per % v/v change in acid concentration. In a separate experiment the concentration and its standard uncertainty have been established using titration with a standardised NaOH titre (3.996% v/v $u = 0.008\%$ v/v). Taking the uncertainty of 0.008% v/v on the acid concentration suggests an uncertainty for f_{acid} of 0.008×0.1 = 0.0008. As the uncertainty on the acid concentration is already expressed as a standard uncertainty, this value can be used directly as the uncertainty associated with f_{acid} .

NOTE: In principle, the uncertainty value would need correcting for the assumption that the single study above is sufficiently representative of all ceramics. The present value does, however, give a reasonable estimate of the magnitude of the uncertainty.

A5.5 Step 4: Calculating the combined standard uncertainty

The amount of leached cadmium per unit area, assuming no dilution, is given by

$$r = \frac{c_0 \cdot V_L}{a_V} \cdot f_{acid} \cdot f_{time} \cdot f_{temp} \quad \text{mg dm}^{-2}$$

The intermediate values and their standard uncertainties are collected in Table A5.3. Employing those values

$$r = \frac{0.26 \times 0.332}{2.37} \times 1.0 \times 1.0 \times 1.0 = 0.036 \text{ mg dm}^{-2}$$

In order to calculate the combined standard uncertainty of a multiplicative expression (as above) the standard uncertainties of each component are used as follows:

$$\begin{aligned} \frac{u_c(r)}{r} &= \sqrt{\left(\frac{u(c_0)}{c_0}\right)^2 + \left(\frac{u(V_L)}{V_L}\right)^2 + \left(\frac{u(a_V)}{a_V}\right)^2 + \left(\frac{u(f_{acid})}{f_{acid}}\right)^2 + \left(\frac{u(f_{time})}{f_{time}}\right)^2 + \left(\frac{u(f_{temp})}{f_{temp}}\right)^2} \\ &= \sqrt{0.069^2 + 0.0054^2 + 0.025^2 + 0.0008^2 + 0.001^2 + 0.06^2} = 0.095 \\ \Rightarrow u_c(r) &= 0.095r = 0.0034 \text{ mg dm}^{-2} \end{aligned}$$

The simpler spreadsheet approach to calculate the combined standard uncertainty is shown in Table

A5.4. A description of the method is given in Appendix E.

The contributions of the different parameters and influence quantities to the measurement uncertainty are illustrated in Figure A5.8, comparing the size of each of the contributions (C13:H13 in Table A5.4) with the combined uncertainty (B16).

The expanded uncertainty $U(r)$ is obtained by applying a coverage factor of 2

$$U_r = 0.0034 \times 2 = 0.007 \text{ mg dm}^{-2}$$

Thus the amount of released cadmium measured according to BS 6748:1986

$$(0.036 \pm 0.007) \text{ mg dm}^{-2}$$

where the stated uncertainty is calculated using a coverage factor of 2.

A5.6 References for Example 5

1. B. Krinitz, V. Franco, J. AOAC **56** 869-875 (1973)
2. B. Krinitz, J. AOAC **61**, 1124-1129 (1978)
3. J. H. Gould, S. W. Butler, K. W. Boyer, E. A. Stelle, J. AOAC **66**, 610-619 (1983)
4. T. D. Seht, S. Sircar, M. Z. Hasan, Bull. Environ. Contam. Toxicol. **10**, 51-56 (1973)
5. J. H. Gould, S. W. Butler, E. A. Steele, J. AOAC **66**, 1112-1116 (1983)

Table A5.3: Intermediate values and uncertainties for leachable cadmium analysis

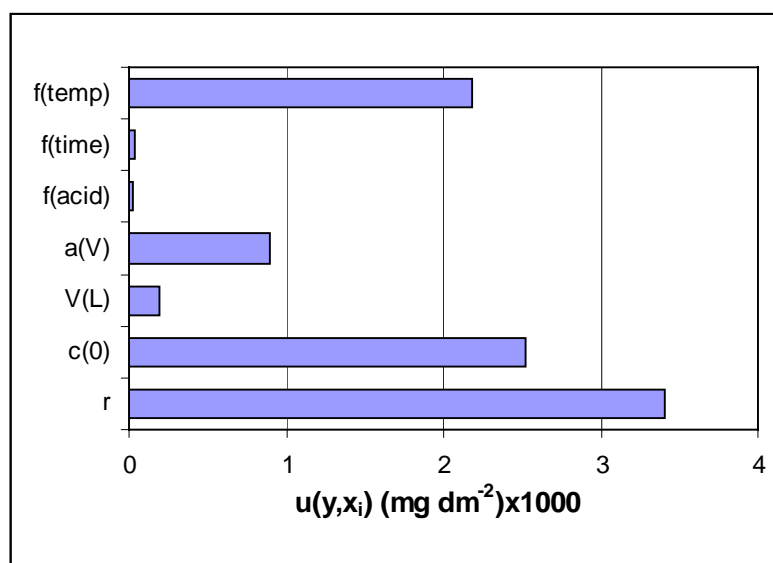
	Description	Value	Standard uncertainty $u(x)$	Relative standard uncertainty $u(x)/x$
c_0	Content of cadmium in the extraction solution	0.26 mg l ⁻¹	0.018 mg l ⁻¹	0.069
V_L	Volume of the leachate	0.332 l	0.0018 l	0.0054
a_V	Surface area of the vessel	2.37 dm ²	0.06 dm ²	0.025
f_{acid}	Influence of the acid concentration	1.0	0.0008	0.0008
f_{time}	Influence of the duration	1.0	0.001	0.001
f_{temp}	Influence of temperature	1.0	0.06	0.06

Table A5.4: Spreadsheet calculation of uncertainty for leachable cadmium analysis

	A	B	C	D	E	F	G	H
1			c_0	V_L	a_V	f_{acid}	f_{time}	f_{temp}
2		value	0.26	0.332	2.37	1.0	1.0	1.0
3		uncertainty	0.018	0.0018	0.06	0.0008	0.001	0.06
4								
5	c_0	0.26	0.278	0.26	0.26	0.26	0.26	0.26
6	V_L	0.332	0.332	0.3338	0.332	0.332	0.332	0.332
7	a_V	2.37	2.37	2.37	2.43	2.37	2.37	2.37
8	f_{acid}	1.0	1.0	1.0	1.0	1.0008	1.0	1.0
9	f_{time}	1.0	1.0	1.0	1.0	1.0	1.001	1.0
10	f_{temp}	1.0	1.0	1.0	1.0	1.0	1.0	1.06
11								
12	r	0.036422	0.038943	0.036619	0.035523	0.036451	0.036458	0.038607
13	$u(y, x_i)$		0.002521	0.000197	-0.000899	0.000029	0.000036	0.002185
14	$u(y)^2$, $u(y, x_i)^2$	1.199 E-5	6.36 E-6	3.90 E-8	8.09 E-7	8.49 E-10	1.33 E-9	4.78 E-6
15								
16	$u_c(r)$	0.0034						

The values of the parameters are entered in the second row from C2 to H2, and their standard uncertainties in the row below (C3:H3). The spreadsheet copies the values from C2:H2 into the second column (B5:B10). The result (r) using these values is given in B12. C5 shows the value of c_0 from C2 plus its uncertainty given in C3. The result of the calculation using the values C5:C10 is given in C12. The columns D and H follow a similar procedure. Row 13 (C13:H13) shows the differences of the row (C12:H12) minus the value given in B12. In row 14 (C14:H14) the values of row 13 (C13:H13) are squared and summed to give the value shown in B14. B16 gives the combined standard uncertainty, which is the square root of B14.

Figure A5.8: Uncertainties in leachable Cd determination



The values of $u(y, x_i) = (\partial y / \partial x_i) \cdot u(x_i)$ are taken from Table A5.4

Example A6: The Determination of Crude Fibre in Animal Feeding Stuffs

Summary

Goal

The determination of crude fibre by a regulatory standard method.

Measurement procedure

The measurement procedure is a standardised procedure involving the general steps outlined in Figure A6.1. These are repeated for a blank sample to obtain a blank correction.

Measurand

The fibre content as a percentage of the sample by weight, C_{fibre} , is given by:

$$C_{fibre} = \frac{(b - c) \times 100}{a}$$

Where:

- a is the mass (g) of the sample.
(Approximately 1 g)
- b is the loss of mass (g) after ashing during the determination;
- c is the loss of mass (g) after ashing during the blank test.

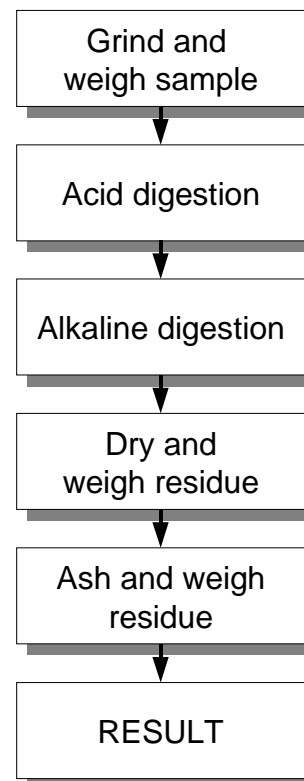
Identification of uncertainty sources

A full cause and effect diagram is provided as Figure A6.9.

Quantification of uncertainty components

Laboratory experiments showed that the method was performing in house in a manner that fully justified adoption of collaborative study

Figure A6.1: Fibre determination.



reproducibility data. No other contributions were significant in general. At low levels it was necessary to add an allowance for the specific drying procedure used. Typical resulting uncertainty estimates are tabulated below (as standard uncertainties) (Table A6.1).

Table A6.1: Combined standard uncertainties

Fibre content (%w/w)	Standard uncertainty $u_c(C_{fibre})$ (%w/w)	Relative Standard uncertainty $u_c(C_{fibre}) / C_{fibre}$
2.5	$\sqrt{0.29^2 + 0.115^2} = 0.31$	0.12
5	0.4	0.08
10	0.6	0.06

Example A6: The determination of crude fibre in animal feeding stuffs.

Detailed discussion

A6.1 Introduction

Crude fibre is defined in the method scope as the amount of fat-free organic substances which are insoluble in acid and alkaline media. The procedure is standardised and its results used directly. Changes in the procedure change the measurand; this is accordingly an example of an empirical method.

Collaborative trial data (repeatability and reproducibility) were available for this statutory method. The precision experiments described were planned as part of the in-house evaluation of the method performance. There is no suitable reference material (i.e. certified by the same method) available for this method.

A6.2 Step 1: Specification

The specification of the measurand for more extensive analytical methods is best done by a comprehensive description of the different stages of the analytical method and by providing the equation of the measurand.

Procedure

The procedure, a complex digestion, filtration, drying, ashing and weighing procedure, which is also repeated for a blank crucible, is summarised in Figure A6.2. The aim is to digest most components, leaving behind all the undigested material. The organic material is ashed, leaving an inorganic residue. The difference between the dry organic/inorganic residue weight and the ashed residue weight is the “fibre content”. The main stages are:

- i) Grind the sample to pass through a 1mm sieve
- ii) Weigh 1g of the sample into a weighed crucible
- iii) Add a set of acid digestion reagents at stated concentrations and volumes. Boil for a stated, standardised time, filter and wash the residue.
- iv) Add standard alkali digestion reagents and boil for the required time, filter, wash and rinse with acetone.

- v) Dry to constant weight at a standardised temperature (“constant weight” is not defined within the published method; nor are other drying conditions such as air circulation or dispersion of the residue).
- vi) Record the dry residue weight.
- vii) Ash at a stated temperature to “constant weight” (in practice realised by ashing for a set time decided after in house studies).
- viii) Weigh the ashed residue and calculate the fibre content by difference, after subtracting the residue weight found for the blank crucible.

Measurand

The fibre content as a percentage of the sample by weight, C_{fibre} , is given by:

$$C_{fibre} = \frac{(b - c) \times 100}{a}$$

Where:

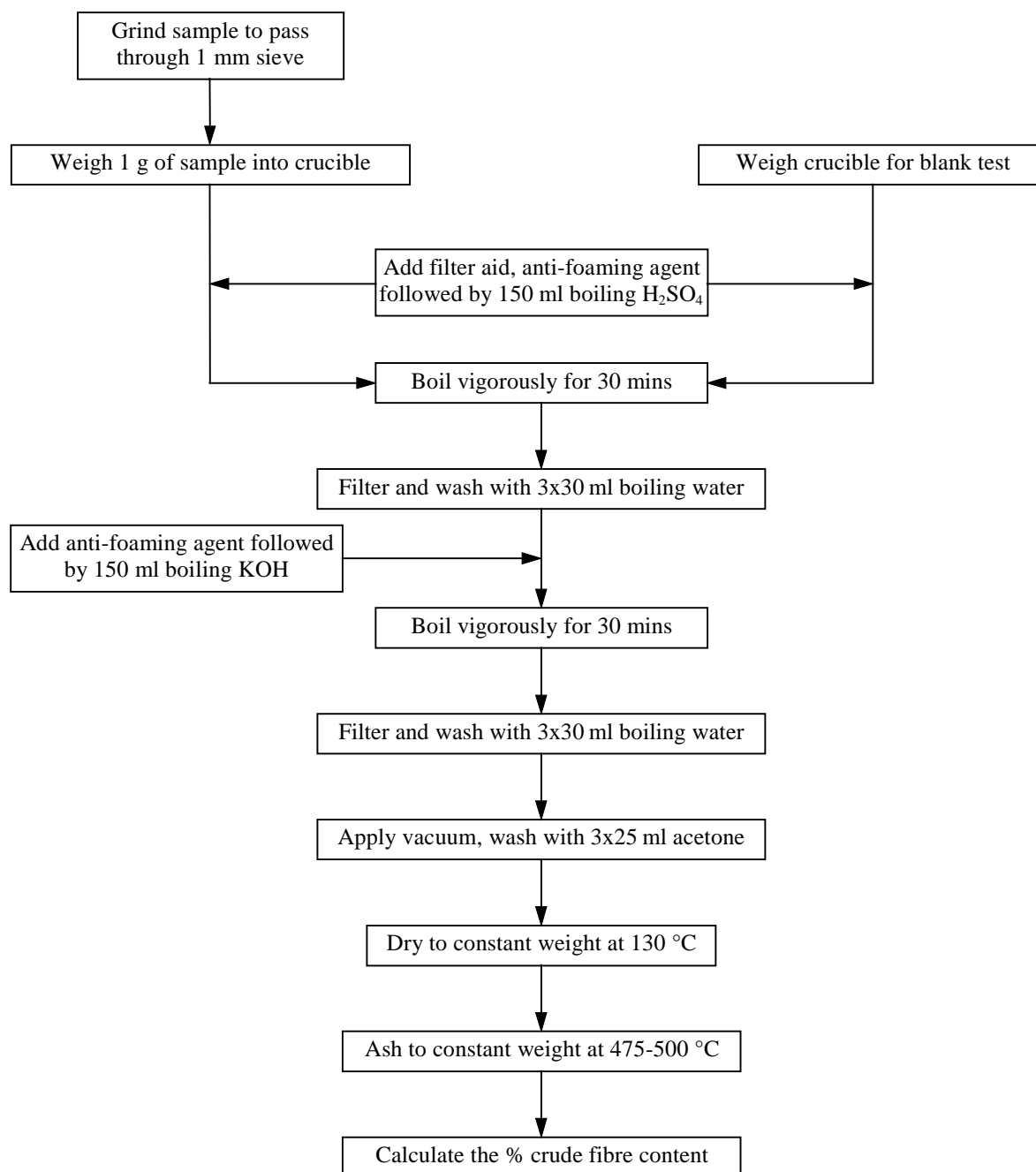
- a* is the mass (g) of the sample. Approximately 1 g of sample is taken for analysis.
- b* is the loss of mass (g) after ashing during the determination.
- c* is the loss of mass (g) after ashing during the blank test.

A6.3 Step 2: Identifying and analysing uncertainty sources

A range of sources of uncertainty was identified. These are shown in the cause and effect diagram for the method (see Figure A6.9). This diagram was simplified to remove duplication following the procedures in Appendix D; this, together with removal of insignificant components, leads to the simplified cause and effect diagram in Figure A6.10.

Since prior collaborative and in-house study data were available for the method, the use of these data is closely related to the evaluation of different contributions to uncertainty and is accordingly discussed further below.

Figure A6.2: Flow diagram illustrating the stages in the regulatory method for the determination of fibre in animal feeding stuffs



A6.4 Step 3: Quantifying uncertainty components

Collaborative trial results

The method has been the subject of a collaborative trial. Five different feeding stuffs representing typical fibre and fat concentrations were analysed in the trial. Participants in the trial carried out all stages of the method, including grinding of the samples. The repeatability and

reproducibility estimates obtained from the trial are presented in Table A6.2.

As part of the in-house evaluation of the method, experiments were planned to evaluate the repeatability (within batch precision) for feeding stuffs with fibre concentrations similar to those of the samples analysed in the collaborative trial. The results are summarised in Table A6.2. Each estimate of in-house repeatability is based on 5 replicates.

Table A6.2: Summary of results from collaborative trial of the method and in-house repeatability check

Sample	Fibre content (% w/w)			
	Collaborative trial results			In-house repeatability standard deviation
	Mean	Reproducibility standard deviation (s_R)	Repeatability standard deviation (s_r)	
A	2.3	0.293	0.198	0.193
B	12.1	0.563	0.358	0.312
C	5.4	0.390	0.264	0.259
D	3.4	0.347	0.232	0.213
E	10.1	0.575	0.391	0.327

The estimates of repeatability obtained in-house were comparable to those obtained from the collaborative trial. This indicates that the method precision in this particular laboratory is similar to that of the laboratories which took part in the collaborative trial. It is therefore acceptable to use the reproducibility standard deviation from the collaborative trial in the uncertainty budget for the method. To complete the uncertainty budget we need to consider whether there are any other effects not covered by the collaborative trial which need to be addressed. The collaborative trial covered different sample matrices and the pre-treatment of samples, as the participants were supplied with samples which required grinding prior to analysis. The uncertainties associated with matrix effects and sample pre-treatment do not therefore require any additional consideration. Other parameters which affect the result relate to the extraction and drying conditions used in the method. These were investigated separately to ensure the laboratory bias was under control (i.e., small compared to the reproducibility standard deviation). The parameters considered are discussed below.

Loss of mass on ashing

As there is no appropriate reference material for this method, in-house bias has to be assessed by considering the uncertainties associated with individual stages of the method. Several factors will contribute to the uncertainty associated with the loss of mass after ashing:

- acid concentration;

- alkali concentration;
- acid digestion time;
- alkali digestion time;
- drying temperature and time;
- ashing temperature and time.

Reagent concentrations and digestion times

The effects of acid concentration, alkali concentration, acid digestion time and alkali digestion time have been studied in previously published papers. In these studies, the effect of changes in the parameter on the result of the analysis was evaluated. For each parameter the sensitivity coefficient (i.e., the rate of change in the final result with changes in the parameter) and the uncertainty in the parameter were calculated.

The uncertainties given in Table A6.3 are small compared to the reproducibility figures presented in Table A6.2. For example, the reproducibility standard deviation for a sample containing 2.3 % w/w fibre is 0.293 % w/w. The uncertainty associated with variations in the acid digestion time is estimated as 0.021 % w/w (i.e., 2.3×0.009). We can therefore safely neglect the uncertainties associated with variations in these method parameters.

Drying temperature and time

No prior data were available. The method states that the sample should be dried at 130 °C to “constant weight”. In this case the sample is dried for 3 hours at 130 °C and then weighed. It is then dried for a further hour and re-weighed. Constant

Table A6.3: Uncertainties associated with method parameters

Parameter	Sensitivity coefficient ^{Note 1}	Uncertainty in parameter	Uncertainty in final result as RSD ^{Note 4}
acid concentration	0.23 (mol l ⁻¹) ⁻¹	0.0013 mol l ⁻¹ ^{Note 2}	0.00030
alkali concentration	0.21 (mol l ⁻¹) ⁻¹	0.0023 mol l ⁻¹ ^{Note 2}	0.00048
acid digestion time	0.0031 min ⁻¹	2.89 mins ^{Note 3}	0.0090
alkali digestion time	0.0025 min ⁻¹	2.89 mins ^{Note 3}	0.0072

Note 1. The sensitivity coefficients were estimated by plotting the normalised change in fibre content against reagent strength or digestion time. Linear regression was then used to calculate the rate of change of the result of the analysis with changes in the parameter.

Note 2. The standard uncertainties in the concentrations of the acid and alkali solutions were calculated from estimates of the precision and trueness of the volumetric glassware used in their preparation, temperature effects etc. See examples A1-A3 for further examples of calculating uncertainties for the concentrations of solutions.

Note 3. The method specifies a digestion time of 30 minutes. The digestion time is controlled to within ± 5 minutes. This is a rectangular distribution which is converted to a standard uncertainty by dividing by $\sqrt{3}$.

Note 4. The uncertainty in the final result, as a relative standard deviation, is calculated by multiplying the sensitivity coefficient by the uncertainty in the parameter.

weight is defined in this laboratory as a change of less than 2 mg between successive weighings. In an in-house study, replicate samples of four feeding stuffs were dried at 110, 130 and 150 °C and weighed after 3 and 4 hours drying time. In the majority of cases, the weight change between 3 and 4 hours was less than 2 mg. This was therefore taken as the worst case estimate of the uncertainty in the weight change on drying. The range ± 2 mg describes a rectangular distribution, which is converted to a standard uncertainty by dividing by $\sqrt{3}$. The uncertainty in the weight recorded after drying to constant weight is therefore 0.00115 g. The method specifies a sample weight of 1 g. For a 1 g sample, the uncertainty in drying to constant weight corresponds to a standard uncertainty of 0.115 % w/w in the fibre content. This source of uncertainty is independent of the fibre content of the sample. There will therefore be a fixed contribution of 0.115 % w/w to the uncertainty budget for each sample, regardless of the concentration of fibre in the sample. At all fibre concentrations, this uncertainty is smaller than the reproducibility standard deviation, and for all but the lowest fibre concentrations is less than 1/3 of the s_R value. Again, this source of uncertainty can usually be neglected. However for low fibre concentrations, this uncertainty is more than 1/3 of the s_R value so an additional term should be

included in the uncertainty budget (see Table A6.4).

Ashing temperature and time

The method requires the sample to be ashed at 475 to 500 °C for at least 30 mins. A published study on the effect of ashing conditions involved determining fibre content at a number of different ashing temperature/time combinations, ranging from 450 °C for 30 minutes to 650 °C for 3 hours. No significant difference was observed between the fibre contents obtained under the different conditions. The effect on the final result of small variations in ashing temperature and time can therefore be assumed to be negligible.

Loss of mass after blank ashing

No experimental data were available for this parameter. However, as discussed above, the effects of variations in this parameter are likely to be small.

A6.5 Step 4: Calculating the combined standard uncertainty

This is an example of an empirical method for which collaborative trial data were available. The in-house repeatability was evaluated and found to be comparable to that predicted by the collaborative trial. It is therefore appropriate to use the s_R values from the collaborative trial. The discussion presented in Step 3 leads to the

conclusion that, with the exception of the effect of drying conditions at low fibre concentrations, the other sources of uncertainty identified are all small in comparison to s_R . In cases such as this, the uncertainty estimate can be based on the reproducibility standard deviation, s_R , obtained from the collaborative trial. For samples with a fibre content of 2.5 % w/w, an additional term has been included to take account of the uncertainty associated with the drying conditions.

Standard uncertainty

Typical standard uncertainties for a range of fibre concentrations are given in the Table A6.4 below.

Expanded uncertainty

Typical expanded uncertainties are given in Table A6.5 below. These were calculated using a coverage factor k of 2, which gives a level of confidence of approximately 95%.

Table A6.4: Combined standard uncertainties

Fibre content (% w/w)	Standard uncertainty $u_c(C_{fibre})$ (%w/w)	Relative standard uncertainty $u_c(C_{fibre}) / C_{fibre}$
2.5	$\sqrt{0.29^2 + 0.115^2} = 0.31$	0.12
5	0.4	0.08
10	0.6	0.06

Table A6.5: Expanded uncertainties

Fibre content (% w/w)	Expanded uncertainty $U(C_{fibre})$ (% w/w)	Expanded uncertainty (% of fibre content)
2.5	0.62	25
5	0.8	16
10	0.12	12

Figure A6.9: Cause and effect diagram for the determination of fibre in animal feeding stuffs

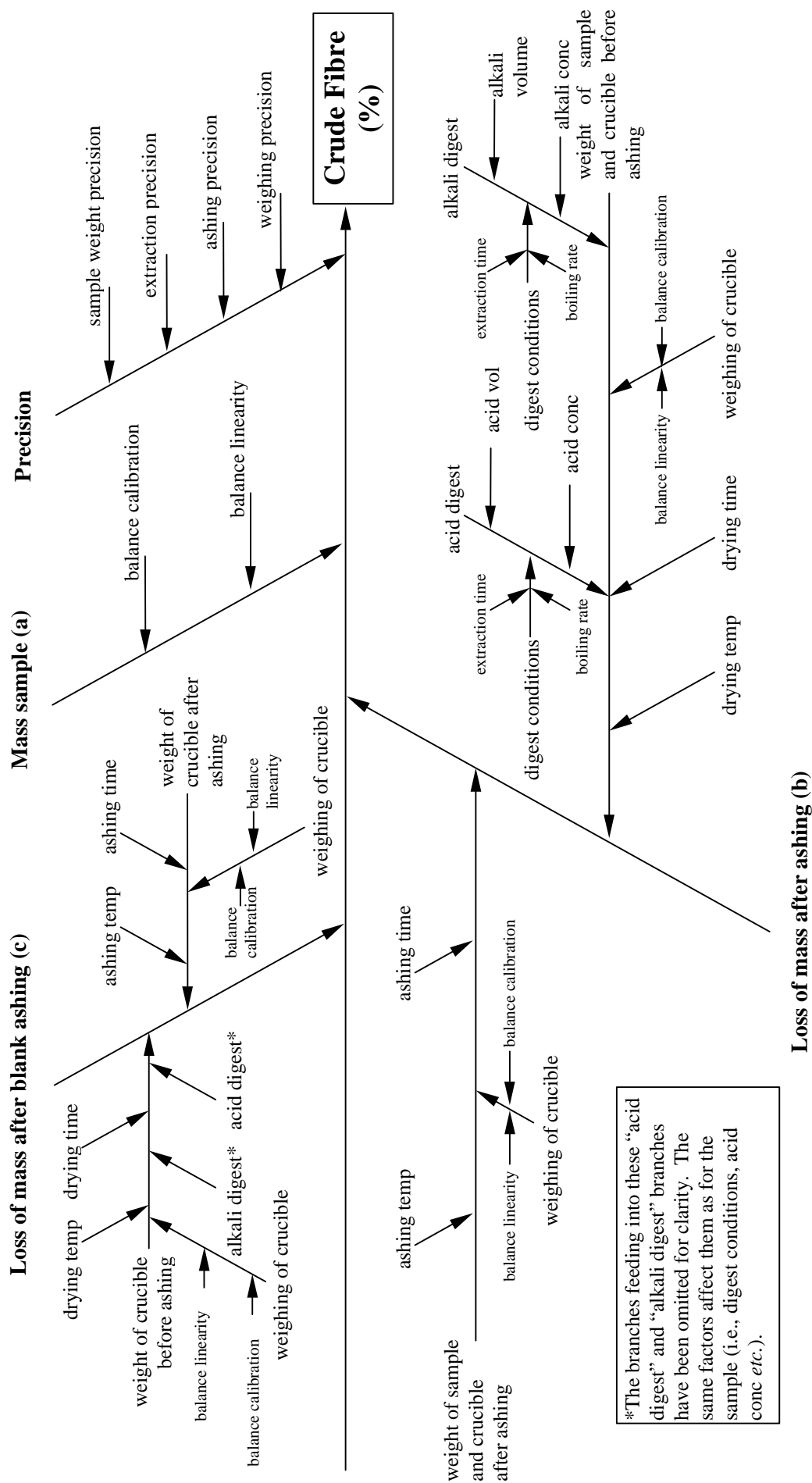
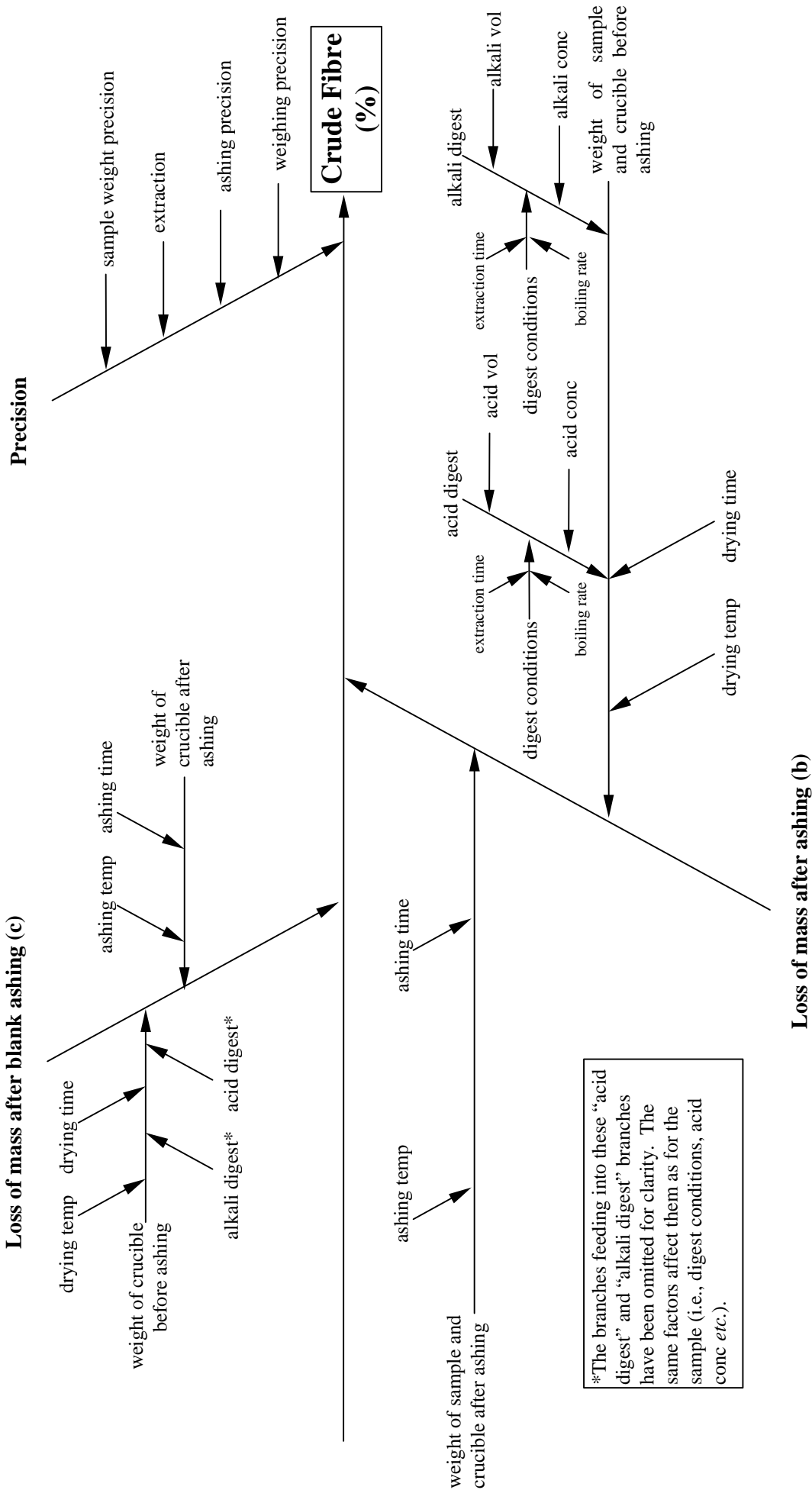


Figure A6.10: Simplified cause and effect diagram



Example A7: Determination of the Amount of Lead in Water Using Double Isotope Dilution and Inductively Coupled Plasma Mass Spectrometry

A7.1 Introduction

This example illustrates how the uncertainty concept can be applied to a measurement of the amount content of lead in a water sample using Isotope Dilution Mass Spectrometry (IDMS) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS).

General introduction to Double IDMS

IDMS is one of the techniques that is recognised by the Comité consultatif pour la quantité de matière (CCQM) to have the potential to be a primary method of measurement, and therefore a well defined expression which describes how the measurand is calculated is available. In the simplest case of isotope dilution using a certified spike, which is an enriched isotopic reference material, isotope ratios in the spike, the sample and a blend b of known masses of sample and spike are measured. The element amount content c_x in the sample is given by:

$$c_x = c_y \cdot \frac{m_y}{m_x} \cdot \frac{K_{y1} \cdot R_{y1} - K_b \cdot R_b}{K_b \cdot R_b - K_{x1} \cdot R_{x1}} \cdot \frac{\sum_i (K_{xi} \cdot R_{xi})}{\sum_i (K_{yi} \cdot R_{yi})} \quad (1)$$

where c_x and c_y are element amount content in the sample and the spike respectively (the symbol c is used here instead of k for amount content¹ to avoid confusion with K -factors and coverage factors k). m_x and m_y are mass of sample and spike respectively. R_x , R_y and R_b are the isotope amount ratios. The indexes x , y and b represent the sample, the spike and the blend respectively. One isotope, usually the most abundant in the sample, is selected and all isotope amount ratios are expressed relative to it. A particular pair of isotopes, the reference isotope and preferably the most abundant isotope in the spike, is then selected as monitor ratio, e.g. $n(^{208}\text{Pb})/n(^{206}\text{Pb})$. R_{xi} and R_{yi} are all the possible isotope amount ratios in the sample and the spike respectively. For the reference isotope, this ratio is unity. K_{xi} , K_{yi} and K_b are the correction factors for mass discrimination, for a particular isotope amount ratio, in sample, spike and blend respectively. The K -factors are measured using a certified isotopic reference material according to equation (2).

$$K = K_0 + K_{\text{bias}}; \text{ where } K_0 = \frac{R_{\text{certified}}}{R_{\text{observed}}} \quad (2)$$

where K_0 is the mass discrimination correction factor at time 0, K_{bias} is a bias factor coming into effect as soon as the K -factor is applied to correct a ratio measured at a different time during the measurement. The K_{bias} also includes other possible sources of bias such as multiplier dead time correction, matrix effects *etc.* $R_{\text{certified}}$ is the certified isotope amount ratio taken from the certificate of an isotopic reference material and R_{observed} is the observed value of this isotopic reference material. In IDMS experiments, using Inductively Coupled Plasma Mass Spectrometry (ICP-MS), mass fractionation will vary with time which requires that all isotope amount ratios in equation (1) need to be individually corrected for mass discrimination.

Certified material enriched in a specific isotope is often unavailable. To overcome this problem, 'double' IDMS is frequently used. The procedure uses a less well characterised, isotopically enriched spiking material in conjunction with a certified material (denoted z) of natural isotopic composition. The certified, natural composition material acts as the primary assay standard. Two blends are used; blend b is a blend between sample and enriched spike, as in equation (1). To perform double IDMS a second blend, b' is prepared from the primary assay standard with amount content c_z , and the enriched material y . This gives a similar expression to equation (1):

$$c_z = c_y \cdot \frac{m'_y}{m_z} \cdot \frac{K'_{y1} \cdot R_{y1} - K'_b \cdot R'_b}{K'_b \cdot R'_b - K'_{z1} \cdot R_{z1}} \cdot \frac{\sum_i (K_{zi} \cdot R_{zi})}{\sum_i (K_{yi} \cdot R_{yi})} \quad (3)$$

where c_z is the element amount content of the primary assay standard solution and m_z the mass of the primary assay standard when preparing the new blend. m'_y is the mass of the enriched spike solution, K'_b , R'_b , K'_{z1} and R_{z1} are the K -factor and the ratio for the new blend and the assay standard respectively. The index z represents the assay

Table A7.1. Summary of IDMS parameters

Parameter	Description	Parameter	Description
m_x	mass of sample in blend b [g]	m_y	mass of enriched spike in blend b [g]
m'_y	mass of enriched spike in blend b' [g]	m_z	mass of primary assay standard in blend b' [g]
c_x	amount content of the sample x [mol g ⁻¹ or μmol g ⁻¹] ^{Note 1}	c_z	amount content of the primary assay standard z [mol g ⁻¹ or μmol g ⁻¹] ^{Note 1}
c_y	amount content of the spike y [mol g ⁻¹ or μmol g ⁻¹] ^{Note 1}	c_{blank}	observed amount content in procedure blank [mol g ⁻¹ or μmol g ⁻¹] ^{Note 1}
R_b	measured ratio of blend b, $n(^{208}\text{Pb})/n(^{206}\text{Pb})$	K_b	mass bias correction of R_b
R'_b	measured ratio of blend b', $n(^{208}\text{Pb})/n(^{206}\text{Pb})$	K'_b	mass bias correction of R'_b
R_{y1}	measured ratio of enriched isotope to reference isotope in the enriched spike	K_{y1}	mass bias correction of R_{y1}
R_{zi}	all ratios in the primary assay standard, R_{z1}, R_{z2} etc.	K_{zi}	mass bias correction factors for R_{zi}
R_{xi}	all ratios in the sample	K_{xi}	mass bias correction factors for R_{xi}
R_{x1}	measured ratio of enriched isotope to reference isotope in the sample x	R_{z1}	as R_{x1} but in the primary assay standard

Note 1: Units for amount content are always specified in the text.

standard. Dividing equation (1) with equation (3) gives

$$\frac{c_x}{c_z} = \frac{c_y \cdot \frac{m_y}{m_x} \cdot \frac{K_{y1} \cdot R_{y1} - K_b \cdot R_b}{K_b \cdot R_b - K_{x1} \cdot R_{x1}} \cdot \frac{\sum_i (K_{xi} \cdot R_{xi})}{\sum_i (K_{yi} \cdot R_{yi})}}{c_y \cdot \frac{m'_y}{m_z} \cdot \frac{K_{y1} \cdot R_{y1} - K'_b \cdot R'_b}{K'_b \cdot R'_b - K_{z1} \cdot R_{z1}} \cdot \frac{\sum_i (K_{zi} \cdot R_{zi})}{\sum_i (K_{yi} \cdot R_{yi})}} \quad (4)$$

Simplifying this equation and introducing a procedure blank, c_{blank} , we get:

$$c_x = c_z \cdot \frac{m_y}{m_x} \cdot \frac{m_z}{m'_y} \cdot \frac{K_{y1} \cdot R_{y1} - K_b \cdot R_b}{K_b \cdot R_b - K_{x1} \cdot R_{x1}} \times \frac{K'_b \cdot R'_b - K_{z1} \cdot R_{z1}}{K_{y1} \cdot R_{y1} - K'_b \cdot R'_b} \cdot \frac{\sum_i (K_{xi} \cdot R_{xi})}{\sum_i (K_{zi} \cdot R_{zi})} - c_{\text{blank}} \quad (5)$$

This is the final equation, from which c_y has been eliminated. In this measurement the number index on the amount ratios, R , represents the following actual isotope amount ratios:

$$R_1 = n(^{208}\text{Pb})/n(^{206}\text{Pb}) \quad R_2 = n(^{206}\text{Pb})/n(^{206}\text{Pb})$$

$$R_3 = n(^{207}\text{Pb})/n(^{206}\text{Pb}) \quad R_4 = n(^{204}\text{Pb})/n(^{206}\text{Pb})$$

For reference, the parameters are summarised in Table A7.1.

A7.2 Step 1: Specification

The general procedure for the measurements is shown in Table A7.2. The calculations and measurements involved are described below.

Calculation procedure for the amount content c_x

For this determination of lead in water, four blends each of b', (assay + spike), and b, (sample + spike), were prepared. This gives a total of 4 values for c_x . One of these determinations will be described in detail following Table A7.2, steps 1 to 4. The reported value for c_x will be the average of the four replicates.

Table A7.2. General procedure

Step	Description
1	Preparing the primary assay standard
2	Preparation of blends: b' and b
3	Measurement of isotope ratios
4	Calculation of the amount content of Pb in the sample, c_x
5	Estimating the uncertainty in c_x

Calculation of the Molar Mass

Due to natural variations in the isotopic composition of certain elements, e.g. Pb, the molar mass, M , for the primary assay standard has to be determined since this will affect the amount content c_z . Note that this is not the case when c_z is expressed in mol g^{-1} . The molar mass, $M(\text{E})$, for an element E, is numerically equal to the atomic weight of element E, $A_r(\text{E})$. The atomic weight can be calculated according to the general expression:

$$A_r(\text{E}) = \frac{\sum_{i=1}^p R_i \cdot M(^i\text{E})}{\sum_{i=1}^p R_i} \quad (6)$$

where the values R_i are all true isotope amount ratios for the element E and $M(^i\text{E})$ are the tabulated nuclide masses.

Note that the isotope amount ratios in equation (6) have to be absolute ratios, that is, they have to be corrected for mass discrimination. With the use of proper indexes, this gives equation (7). For the calculation, nuclide masses, $M(^i\text{E})$, were taken from literature values², while Ratios, R_{zi} , and K_0 -factors, $K_0(zi)$, were measured (see Table A7.8). These values give

$$M(\text{Pb, Assay 1}) = \frac{\sum_{i=1}^p K_{zi} \cdot R_{zi} \cdot M_z(^i\text{E})}{\sum_{i=1}^p K_{zi} \cdot R_{zi}} \quad (7)$$

$$= 207.21034 \text{ g mol}^{-1}$$

Measurement of K -factors and isotope amount ratios

To correct for mass discrimination, a correction factor, K , is used as specified in equation (2). The K_0 -factor can be calculated using a reference

material certified for isotopic composition. In this case, the isotopically certified reference material NIST SRM 981 was used to monitor a possible change in the K_0 -factor. The K_0 -factor is measured before and after the ratio it will correct. A typical sample sequence is: 1. (blank), 2. (NIST SRM 981), 3. (blank), 4. (blend 1), 5. (blank), 6. (NIST SRM 981), 7. (blank), 8. (sample), etc.

The blank measurements are not only used for blank correction, they are also used for monitoring the number of counts for the blank. No new measurement run was started until the blank count rate was stable and back to a normal level. Note that sample, blends, spike and assay standard were diluted to an appropriate amount content prior to the measurements. The results of ratio measurements, calculated K_0 -factors and K_{bias} are summarised in Table A7.8.

Preparing the primary assay standard and calculating the amount content, c_z

Two primary assay standards were produced, each from a different piece of metallic lead with a chemical purity of $w=99.999\%$. The two pieces came from the same batch of high purity lead. The pieces were dissolved in about 10 ml of 1:3 w/w HNO_3 :water under gentle heating and then further diluted. Two blends were prepared from each of these two assay standards. The values from one of the assays is described hereafter.

0.36544 g lead, m_1 , was dissolved and diluted in aqueous HNO_3 (0.5 mol l^{-1}) to a total of $d_1=196.14 \text{ g}$. This solution is named *Assay 1*. A more diluted solution was needed and $m_2=1.0292 \text{ g}$ of *Assay 1*, was diluted in aqueous HNO_3 (0.5 mol l^{-1}) to a total mass of $d_2=99.931 \text{ g}$. This solution is named *Assay 2*. The amount content of Pb in *Assay 2*, c_z , is then calculated according to equation (8)

$$c_z = \frac{m_2}{d_2} \cdot \frac{m_1 \cdot w}{d_1} \cdot \frac{1}{M(\text{Pb, Assay1})} \quad (8)$$

$$= 9.2605 \times 10^{-8} \text{ mol g}^{-1} = 0.092605 \mu\text{mol g}^{-1}$$

Preparation of the blends

The mass fraction of the spike is known to be roughly $20 \mu\text{g Pb per g solution}$ and the mass fraction of Pb in the sample is also known to be in this range. Table A7.3 shows the weighing data for the two blends used in this example.

Table A7.3

Blend	b		b'	
	Spike	Sample	Spike	Assay 2
Solutions used				
Parameter	m_y	m_x	m'_y	m_z
Mass (g)	1.1360	1.0440	1.0654	1.1029

Measurement of the procedure blank c_{Blank}

In this case, the procedure blank was measured using external calibration. A more exhaustive procedure would be to add an enriched spike to a blank and process it in the same way as the samples. In this example, only high purity reagents were used, which would lead to extreme ratios in the blends and consequent poor reliability for the enriched spiking procedure. The externally calibrated procedure blank was measured four times, and c_{Blank} found to be $4.5 \times 10^{-7} \mu\text{mol g}^{-1}$, with standard uncertainty $4.0 \times 10^{-7} \mu\text{mol g}^{-1}$ evaluated as type A.

Calculation of the unknown amount content c_x

Inserting the measured and calculated data (Table A7.8) into equation (5) gives $c_x = 0.053738 \mu\text{mol g}^{-1}$. The results from all four replicates are given in Table A7.4.

A7.3 Steps 2 and 3: Identifying and quantifying uncertainty sources

Strategy for the uncertainty calculation

If equations (2), (7) and (8) were to be included in the final IDMS equation (5), the sheer number of parameters would make the equation almost impossible to handle. To keep it simpler, K_0 -factors and amount content of the standard assay solution and their associated uncertainties are treated separately and then introduced into the IDMS equation (5). In this case it will not affect the final combined uncertainty of c_x , and it is advisable to simplify for practical reasons.

For calculating the combined standard uncertainty, $u_c(c_x)$, the values from one of the measurements, as described in A7.2, will be used. The combined uncertainty of c_x will be calculated using the spreadsheet method described in Appendix E.

Table A7.4

	$c_x (\mu\text{mol g}^{-1})$
Replicate 1 (our example)	0.053738
Replicate 2	0.053621
Replicate 3	0.053610
Replicate 4	0.053822
Average	0.05370
Experimental standard deviation (s)	0.0001

Uncertainty on the K -factors

i) Uncertainty on K_0

K is calculated according to equation (2) and using the values of K_{x1} as an example gives for K_0 :

$$K_0(x1) = \frac{R_{\text{certified}}}{R_{\text{observed}}} = \frac{2.1681}{2.1699} = 0.9992 \quad (9)$$

To calculate the uncertainty on K_0 we first look at the certificate where the certified ratio, 2.1681, has a stated uncertainty of 0.0008 based on a 95% confidence interval. To convert an uncertainty based on a 95% confidence interval to standard uncertainty we divide by 2. This gives a standard uncertainty of $u(R_{\text{certified}}) = 0.0004$. The observed amount ratio, $R_{\text{observed}} = n(^{208}\text{Pb})/n(^{206}\text{Pb})$, has a standard uncertainty of 0.0025 (as rsd). For the K -factor, the combined uncertainty can be calculated as:

$$\frac{u_c(K_0(x1))}{K_0(x1)} = \sqrt{\left(\frac{0.0004}{2.1681}\right)^2 + (0.0025)^2} \quad (10)$$

$$= 0.002507$$

This clearly points out that the uncertainty contributions from the certified ratios are negligible. Henceforth, the uncertainties on the measured ratios, R_{observed} , will be used for the uncertainties on K_0 .

Uncertainty on K_{bias}

This bias factor is introduced to account for possible deviations in the value of the mass discrimination factor. As can be seen in the cause and effect diagram above, and in equation (2), there is a bias associated with every K -factor. The values of these biases are in our case not known, and a value of 0 is applied. An uncertainty is, of

course, associated with every bias and this has to be taken into consideration when calculating the final uncertainty. In principle, a bias would be applied as in equation (11), using an excerpt from equation (5) and the parameters K_{y1} and R_{y1} to demonstrate this principle.

$$c_x = \dots \cdot \frac{(K_0(y1) + K_{\text{bias}}(y1)) \cdot R_{y1} - \dots}{\dots} \cdot \dots \quad (11)$$

The values of all biases, $K_{\text{bias}}(y_i, x_i, z_i)$, are (0 ± 0.001). This estimation is based on a long experience of lead IDMS measurements. All $K_{\text{bias}}(y_i, x_i, z_i)$ parameters are not included in detail in Table A7.5, Table A7.8 or in equation 5, but they are used in all uncertainty calculations.

Uncertainty of the weighed masses

In this case, a dedicated mass metrology lab performed the weighings. The procedure applied was a bracketing technique using calibrated weights and a comparator. The bracketing technique was repeated at least six times for every sample mass determination. Buoyancy correction was applied. Stoichiometry and impurity corrections were not applied in this case. The uncertainties from the weighing certificates were treated as standard uncertainties and are given in

Table A7.5

	Value	Standard Uncertainty	Type ^{Note 1}
$K_{\text{bias}}(z_i)$	0	0.001	B
R_{z1}	2.1429	0.0054	A
$K_0(z1)$	0.9989	0.0025	A
$K_0(z3)$	0.9993	0.0035	A
$K_0(z4)$	1.0002	0.0060	A
R_{z2}	1	0	A
R_{z3}	0.9147	0.0032	A
R_{z4}	0.05870	0.00035	A
M_1	207.976636	0.000003	B
M_2	205.974449	0.000003	B
M_3	206.975880	0.000003	B
M_4	203.973028	0.000003	B

Note 1. Type A (statistical evaluation) or Type B (other)

Table A7.8.

Uncertainty in the amount content of the Standard Assay Solution, c_z

i) Uncertainty in the atomic weight of Pb

First, the combined uncertainty of the molar mass of the assay solution, *Assay 1*, will be calculated. The values in Table A7.5 are known or have been measured:

According to equation (7), the calculation of the molar mass takes this form:

$$M(\text{Pb}, \text{Assay1}) = \frac{K_{z1} \cdot R_{z1} \cdot M_1 + R_{z2} \cdot M_2 + K_{z3} \cdot R_{z3} \cdot M_3 + K_{z4} \cdot R_{z4} \cdot M_4}{K_{z1} \cdot R_{z1} + K_{z2} \cdot R_{z2} + K_{z3} \cdot R_{z3} + K_{z4} \cdot R_{z4}} \quad (12)$$

To calculate the combined standard uncertainty of the molar mass of Pb in the standard assay solution, the spreadsheet model described in Appendix E was used. There were eight measurements of every ratio and K_0 . This gave a molar mass $M(\text{Pb}, \text{Assay 1}) = 207.2103 \text{ g mol}^{-1}$, with uncertainty $0.0010 \text{ g mol}^{-1}$ calculated using the spreadsheet method.

ii) Calculation of the combined standard uncertainty in determining c_z

To calculate the uncertainty on the amount content of Pb in the standard assay solution, c_z the data from A7.2 and equation (8) are used. The uncertainties were taken from the weighing certificates, see A7.3. All parameters used in equation (8) are given with their uncertainties in Table A7.6.

The amount content, c_z , was calculated using equation (8). Following Appendix D.5 the combined standard uncertainty in c_z , is calculated to be $u_c(c_z) = 0.000028$. This gives $c_z = 0.092606 \text{ } \mu\text{mol g}^{-1}$ with a standard uncertainty of $0.000028 \text{ } \mu\text{mol g}^{-1}$ (0.03% as %rsd).

To calculate $u_c(c_x)$, for replicate 1, the spreadsheet model was applied (Appendix E). The uncertainty budget for replicate 1 will be representative for the measurement. Due to the number of parameters in equation (5), the spreadsheet will not be displayed. The value of the parameters and their uncertainties as well as the combined uncertainty of c_x can be seen in Table A7.8.

Table A7.6

	Value	Uncertainty
Mass of lead piece, m_1 (g)	0.36544	0.00005
Total mass first dilution, d_1 (g)	196.14	0.03
Aliquot of first dilution, m_2 (g)	1.0292	0.0002
Total mass of second dilution, d_2 (g)	99.931	0.01
Purity of the metallic lead piece, w (mass fraction)	0.99999	0.000005
Molar mass of Pb in the Assay Material, M (g mol ⁻¹)	207.2104	0.0010

A7.4 Step 4: Calculating the combined standard uncertainty

The average and the experimental standard deviation of the four replicates are displayed in Table A7.7. The numbers are taken from Table A7.4 and Table A7.8.

Table A7.7

Replicate 1		Mean of replicates 1-4		
c_x	0.05374	c_x	0.05370	$\mu\text{mol g}^{-1}$
$u_c(c_x)$	0.00018	s	0.00010 ^{Note 1}	$\mu\text{mol g}^{-1}$

Note 1. This is the experimental standard uncertainty and not the standard deviation of the mean.

In IDMS, and in many non-routine analyses, a complete statistical control of the measurement procedure would require limitless resources and

time. A good way then to check if some source of uncertainty has been forgotten is to compare the uncertainties from the type A evaluations with the experimental standard deviation of the four replicates. If the experimental standard deviation is higher than the contributions from the uncertainty sources evaluated as type A, it could indicate that the measurement process is not fully understood. As an approximation, using data from Table 8, the sum of the type A evaluated experimental uncertainties can be calculated by taking 92.2% of the total experimental uncertainty, which is $0.00041 \mu\text{mol g}^{-1}$. This value is then clearly higher than the experimental standard deviation of $0.00010 \mu\text{mol g}^{-1}$, see Table A7.7. This indicates that the experimental standard deviation is covered by the contributions from the type A evaluated uncertainties and that no further type A evaluated uncertainty contribution, due to the preparation of the blends, needs to be considered. There could however be a bias associated with the preparations of the blends. In this example, a possible bias in the preparation of the blends is judged to be insignificant in comparison to the major sources of uncertainty.

The amount content of lead in the water sample is then:

$$c_x = (0.05370 \pm 0.00036) \mu\text{mol g}^{-1}$$

The result is presented with an expanded uncertainty using a coverage factor of 2.

References for Example 7

1. T. Cvitaš, *Metrologia*, 1996, **33**, 35-39
2. G. Audi and A.H. Wapstra, *Nuclear Physics*, A565 (1993)

Table A7.8

parameter	uncertainty evaluation	value	experimental uncertainty (Note 1)	contribution to total $u_c(\%)$	final uncertainty (Note 2)	contribution to total $u_c(\%)$
ΣK_{bias}	B	0	0.001 ^{Note 3}	7.2	0.001 ^{Note 3}	37.6
c_z	B	0.092605	0.000028	0.2	0.000028	0.8
$K_0(\mathbf{b})$	A	0.9987	0.0025	14.4	0.00088	9.5
$K_0(\mathbf{b}')$	A	0.9983	0.0025	18.3	0.00088	11.9
$K_0(\mathbf{x}1)$	A	0.9992	0.0025	4.3	0.00088	2.8
$K_0(\mathbf{x}3)$	A	1.0004	0.0035	1	0.0012	0.6
$K_0(\mathbf{x}4)$	A	1.001	0.006	0	0.0021	0
$K_0(\mathbf{y}1)$	A	0.9999	0.0025	0	0.00088	0
$K_0(\mathbf{z}1)$	A	0.9989	0.0025	6.6	0.00088	4.3
$K_0(\mathbf{z}3)$	A	0.9993	0.0035	1	0.0012	0.6
$K_0(\mathbf{z}4)$	A	1.0002	0.006	0	0.0021	0
m_x	B	1.0440	0.0002	0.1	0.0002	0.3
m_{y1}	B	1.1360	0.0002	0.1	0.0002	0.3
m_{y2}	B	1.0654	0.0002	0.1	0.0002	0.3
m_z	B	1.1029	0.0002	0.1	0.0002	0.3
R_b	A	0.29360	0.00073	14.2	0.00026 ^{Note 4}	9.5
R'_b	A	0.5050	0.0013	19.3	0.00046	12.7
R_{x1}	A	2.1402	0.0054	4.4	0.0019	2.9
R_{x2}	Cons.	1	0		0	
R_{x3}	A	0.9142	0.0032	1	0.0011	0.6
R_{x4}	A	0.05901	0.00035	0	0.00012	0
R_{y1}	A	0.00064	0.00004	0	0.000014	0
R_{z1}	A	2.1429	0.0054	6.7	0.0019	4.4
R_{z2}	Cons.	1	0		0	
R_{z3}	A	0.9147	0.0032	1	0.0011	0.6
R_{z4}	A	0.05870	0.00035	0	0.00012	0
c_{Blank}	A	4.5×10^{-7}	4.0×10^{-7}	0	2.0×10^{-7}	0
c_x		0.05374	0.00041		0.00018	
			$\Sigma A_{\text{contrib.}} =$	92.2	$\Sigma A_{\text{contrib.}} =$	60.4
			$\Sigma B_{\text{contrib.}} =$	7.8	$\Sigma B_{\text{contrib.}} =$	39.6

Notes overleaf

Notes to Table A7.8

Note 1. The experimental uncertainty is calculated without taking the number of measurements on each parameter into account.

Note 2. In the final uncertainty the number of measurements has been taken into account. In this case all type A evaluated parameters have been measured 8 times. Their standard uncertainties have been divided by $\sqrt{8}$.

Note 3. This value is for one single K_{bias} . The parameter ΣK_{bias} is used instead of listing all $K_{\text{bias}}(z_i, x_i, y_i)$, which all have the same value (0 ± 0.001).

Note 4. R_p has been measured 8 times per blend giving a total of 32 observations. When there is no blend to blend variation, as in this example, all these 32 observations could be accounted for by implementing all four blend replicates in the model. This can be very time consuming and since, in this case, it does not affect the uncertainty noticeably, it is not done.

Appendix B. Definitions

General

B.1 Accuracy of measurement

The closeness of the agreement between the result of a measurement and a *true value* of the measurand [H.4].

NOTE 1 "Accuracy" is a qualitative concept.

NOTE 2 The term "precision" should not be used for "accuracy".

B.2 Precision

The closeness of agreement between independent test results obtained under stipulated conditions [H.5].

NOTE 1 Precision depends only on the distribution of random errors and does not relate to the true value or the specified value.

NOTE 2 The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a larger standard deviation.

NOTE 3 "Independent test results" means results obtained in a manner not influenced by any previous result on the same or similar test object. Quantitative measures of precision depend critically on the stipulated conditions. Repeatability and reproducibility conditions are particular sets of extreme stipulated conditions.

B.3 True value

Value consistent with the definition of a given particular quantity [H.4].

NOTE 1 This is a value that would be obtained by a perfect measurement.

NOTE 2 True values are by nature indeterminate.

NOTE 3 The indefinite article "a" rather than the definite article "the" is used in

conjunction with "true value" because there may be many values consistent with the definition of a given particular quantity.

B.4 Conventional true value

Value attributed to a particular quantity and accepted, sometimes by convention, as having an uncertainty appropriate for a given purpose [H.4].

EXAMPLES

a) At a given location, the value assigned to the quantity realised by a reference standard may be taken as a conventional true value.

b) The CODATA (1986) recommended value for the Avogadro constant, N_A : $6.0221367 \times 10^{23} \text{ mol}^{-1}$

NOTE 1 "Conventional true value" is sometimes called *assigned value*, *best estimate* of the value, *conventional value* or *reference value*.

NOTE 2 Frequently, a number of results of measurements of a quantity is used to establish a conventional true value.

B.5 Influence quantity

A quantity that is not the measurand but that affects the result of the measurement [H.4].

EXAMPLES

1. Temperature of a micrometer used to measure length;

2. Frequency in the measurement of an alternating electric potential difference;

3. Bilirubin concentration in the measurement of haemoglobin concentration in human blood plasma.

Measurement

B.6 Measurand

Particular quantity subject to measurement [H.4].

NOTE The specification of a measurand may require statements about quantities such as time, temperature and pressure..

B.7 Measurement

Set of operations having the object of determining a value of a quantity [H.4].

B.8 Measurement procedure

Set of operations, described specifically, used in the performance of measurements according to a given method [H.4].

NOTE A measurement procedure is usually recorded in a document that is sometimes itself called a "measurement procedure" (or a *measurement method*) and is usually in sufficient detail to enable an operator to carry out a measurement without additional information.

B.9 Method of measurement

A logical sequence of operations, described generically, used in the performance of measurements [H.4].

NOTE Methods of measurement may be qualified in various ways such as:

- substitution method
- differential method
- null method

B.10 Result of a measurement

Value attributed to a measurand, obtained by measurement [H.4].

NOTE 1 When the term "result of a measurement" is used, it should be made clear whether it refers to:

- The indication.
- The uncorrected result.
- The corrected result.

and whether several values are averaged.

NOTE 2 A complete statement of the result of a measurement includes information about the uncertainty of measurement.

Uncertainty

B.11 Uncertainty (of measurement)

Parameter associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand [H.4].

NOTE 1 The parameter may be, for example, a standard deviation (or a given multiple of it), or the width of a confidence interval.

NOTE 2 Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of a series of measurements and can be characterised by experimental standard deviations. The other components, which can also be characterised by standard deviations, are evaluated from assumed probability distributions based on experience or other information.

NOTE 3 It is understood that the result of the measurement is the best estimate of the value of the measurand and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion.

B.12 Traceability

The property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties [H.4].

B.13 Standard uncertainty

$u(x_i)$ Uncertainty of the result x_i of a measurement expressed as a standard deviation [H.2].

B.14 Combined standard uncertainty

$u_c(y)$ Standard uncertainty of the result y of a measurement when the result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being

the variances or covariances of these other quantities weighted according to how the measurement result varies with these quantities [H.2].

B.15 Expanded uncertainty

U Quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand [H.2].

NOTE 1 The fraction may be regarded as the coverage probability or level of confidence of the interval.

NOTE 2 To associate a specific level of confidence with the interval defined by the expanded uncertainty requires explicit or implicit assumptions regarding the probability distribution characterised by the measurement result and its combined standard uncertainty. The level of confidence that may be attributed to this interval can be known only to the extent to which such assumptions can be justified.

NOTE 3 An expanded uncertainty *U* is calculated from a combined standard uncertainty u_c and a coverage factor *k* using

$$U = k \times u_c$$

B.16 Coverage factor

k Numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty [H.2].

NOTE A coverage factor is typically in the range 2 to 3.

B.17 Type A evaluation (of uncertainty)

Method of evaluation of uncertainty by the statistical analysis of series of observations [H.2].

B.18 Type B evaluation (of uncertainty)

Method of evaluation of uncertainty by means other than the statistical analysis of series of observations [H.2]

Error

B.19 Error (of measurement)

The result of a measurement minus a true value of the measurand [H.4].

NOTE 1 Since a true value cannot be determined, in practice a conventional true value is used.

B.20 Random error

Result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions [H.4].

NOTE 1 Random error is equal to error minus systematic error.

NOTE 2 Because only a finite number of measurements can be made, it is possible to determine only an estimate of random error.

B.21 Systematic error

Mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand [H.4].

NOTE 1: Systematic error is equal to error minus random error.

NOTE 2: Like true value, systematic error and its causes cannot be known.

Statistical terms

B.22 Arithmetic mean

\bar{x} Arithmetic mean value of a sample of *n* results.

$$\bar{x} = \frac{\sum_{i=1, n} x_i}{n}$$

B.23 Sample Standard Deviation

s An estimate of the population standard deviation σ from a sample of *n* results.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

B.24 Standard deviation of the mean

$s_{\bar{x}}$ The standard deviation of the mean \bar{x} of n values taken from a population is given by

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

The terms "standard error" and "standard error of the mean" have also been used to describe the same quantity.

B.25 Relative Standard Deviation (RSD)

RSD An estimate of the standard deviation of a population from a sample of n results divided by the mean of that sample. Often known as coefficient of variation (CV). Also frequently stated as a percentage.

$$\mathbf{RSD} = \frac{s}{\bar{x}}$$

Appendix C. Uncertainties in Analytical Processes

C.1 In order to identify the possible sources of uncertainty in an analytical procedure it is helpful to break down the analysis into a set of generic steps:

1. **Sampling**
2. **Sample preparation**
3. **Presentation of Certified Reference Materials to the measuring system**
4. **Calibration of Instrument**
5. **Analysis (data acquisition)**
6. **Data processing**
7. **Presentation of results**
8. **Interpretation of results**

C.2 These steps can be further broken down by contributions to the uncertainty for each. The following list, though not necessarily comprehensive, provides guidance on factors which should be considered.

1. **Sampling**
 - Homogeneity.
 - Effects of specific sampling strategy (e.g. random, stratified random, proportional *etc.*)
 - Effects of movement of bulk medium (particularly density selection)
 - Physical state of bulk (solid, liquid, gas)
 - Temperature and pressure effects.
 - Does sampling process affect composition? E.g. differential adsorption in sampling system.
2. **Sample preparation**
 - Homogenisation and/or sub-sampling effects.
 - Drying.
 - Milling.
 - Dissolution.
 - Extraction.
 - Contamination.

- Derivatisation (chemical effects)
- Dilution errors.
- (Pre-)Concentration.
- Control of speciation effects.

3. **Presentation of Certified Reference Materials to the measuring system**
 - Uncertainty for CRM.
 - CRM match to sample
4. **Calibration of instrument**
 - Instrument calibration errors using a Certified Reference Material.
 - Reference material and its uncertainty.
 - Sample match to calibrant
 - Instrument precision
5. **Analysis**
 - Carry-over in auto analysers.
 - Operator effects, e.g. colour blindness, parallax, other systematic errors.
 - Interferences from the matrix, reagents or other analytes.
 - Reagent purity.
 - Instrument parameter settings, e.g. integration parameters
 - Run-to-run precision
6. **Data Processing**
 - Averaging.
 - Control of rounding and truncating.
 - Statistics.
 - Processing algorithms (model fitting, e.g. linear least squares).
7. **Presentation of Results**
 - Final result.
 - Estimate of uncertainty.
 - Confidence level.
8. **Interpretation of Results**
 - Against limits/bounds.
 - Regulatory compliance.
 - Fitness for purpose.

Appendix D. Analysing Uncertainty Sources

D.1 Introduction

It is commonly necessary to develop and record a list of sources of uncertainty relevant to an analytical method. It is often useful to structure this process, both to ensure comprehensive coverage and to avoid over-counting. The following procedure (based on a previously published method [H.14]), provides one possible means of developing a suitable, structured analysis of uncertainty contributions.

D.2 Principles of approach

D.2.1 The strategy has two stages:

- Identifying the effects on a result
In practice, the necessary structured analysis is effected using a *cause and effect diagram* (sometimes known as an Ishikawa or ‘fishbone’ diagram) [H.15].
- Simplifying and resolving duplication
The initial list is refined to simplify presentation and ensure that effects are not unnecessarily duplicated.

D.3 Cause and effect analysis

D.3.1 The principles of constructing a cause and effect diagram are described fully elsewhere. The procedure employed is as follows:

1. Write the complete equation for the result. The parameters in the equation form the main branches of the diagram. It is almost always necessary to add a main branch representing a nominal correction for overall bias, usually as recovery, and this is accordingly recommended at this stage if appropriate.
2. Consider each step of the method and add any further factors to the diagram, working outwards from the main effects. Examples include environmental and matrix effects.
3. For each branch, add contributory factors until effects become sufficiently remote, that is, until effects on the result are negligible.
4. Resolve duplications and re-arrange to clarify contributions and group related causes. It is

convenient to group precision terms at this stage on a separate precision branch.

D.3.2 The final stage of the cause and effect analysis requires further elucidation. Duplications arise naturally in detailing contributions separately for every input parameter. For example, a run-to-run variability element is always present, at least nominally, for any influence factor; these effects contribute to any overall variance observed for the method as a whole and should not be added in separately if already so accounted for. Similarly, it is common to find the same instrument used to weigh materials, leading to over-counting of its calibration uncertainties. These considerations lead to the following additional rules for refinement of the diagram (though they apply equally well to any structured list of effects):

- Cancelling effects: remove both. For example, in a weight by difference, two weights are determined, both subject to the balance ‘zero bias’. The zero bias will cancel out of the weight by difference, and can be removed from the branches corresponding to the separate weighings.
- Similar effect, same time: combine into a single input. For example, run-to-run variation on many inputs can be combined into an overall run-to-run precision ‘branch’. Some caution is required; specifically, variability in operations carried out individually for every determination can be combined, whereas variability in operations carried out on complete batches (such as instrument calibration) will only be observable in between-batch measures of precision.
- Different instances: re-label. It is common to find similarly named effects which actually refer to different instances of similar measurements. These must be clearly distinguished before proceeding.

D.3.3 This form of analysis does not lead to uniquely structured lists. In the present example, temperature may be seen as either a direct effect on the density to be measured, or as an effect on the measured mass of material contained in a

density bottle; either could form the initial structure. In practice this does not affect the utility of the method. Provided that all significant effects appear once, somewhere in the list, the overall methodology remains effective.

D.3.4 Once the cause-and-effect analysis is complete, it may be appropriate to return to the original equation for the result and add any new terms (such as temperature) to the equation.

D.4 Example

D.4.1 The procedure is illustrated by reference to a simplified direct density measurement. Consider the case of direct determination of the density $d(\text{EtOH})$ of ethanol by weighing a known volume V in a suitable volumetric vessel of tare weight m_{tare} and gross weight including ethanol m_{gross} . The density is calculated from

$$d(\text{EtOH}) = (m_{\text{gross}} - m_{\text{tare}}) / V$$

For clarity, only three effects will be considered: Equipment calibration, Temperature, and the precision of each determination. Figures D1-D3 illustrate the process graphically.

D.4.2 A cause and effect diagram consists of a hierarchical structure culminating in a single outcome. For the present purpose, this outcome is a particular analytical result (' $d(\text{EtOH})$ ' in Figure D1). The 'branches' leading to the outcome are the contributory effects, which include both the results of particular intermediate measurements and other factors, such as environmental or matrix effects. Each branch may in turn have further contributory effects. These 'effects' comprise all factors affecting the result, whether variable or constant; uncertainties in any of these effects will clearly contribute to uncertainty in the result.

D.4.3 Figure D1 shows a possible diagram obtained directly from application of steps 1-3. The main branches are the parameters in the equation, and effects on each are represented by subsidiary branches. Note that there are two 'temperature' effects, three 'precision' effects and three 'calibration' effects.

D.4.4 Figure D2 shows precision and temperature effects each grouped together following the second rule (same effect/time); temperature may be treated as a single effect on density, while the individual variations in each determination contribute to variation observed in replication of the entire method.

D.4.5 The calibration bias on the two weighings cancels, and can be removed (Figure D3) following the first refinement rule (cancellation).

D.4.6 Finally, the remaining 'calibration' branches would need to be distinguished as two (different) contributions owing to possible non-linearity of balance response, together with the calibration uncertainty associated with the volumetric determination.

Figure D1: Initial list

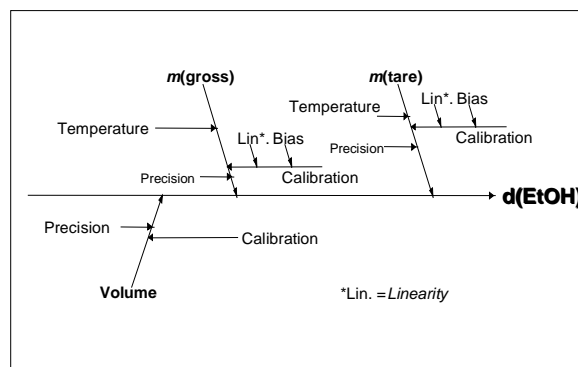


Figure D2: Combination of similar effects

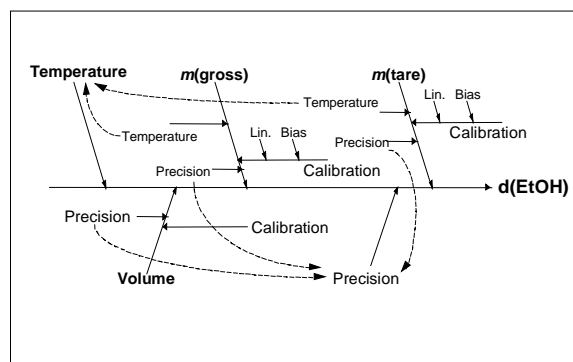
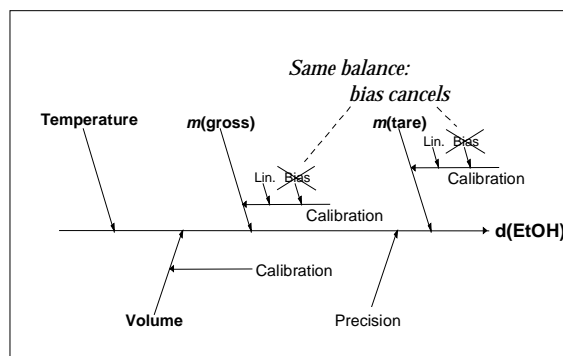


Figure D3: Cancellation



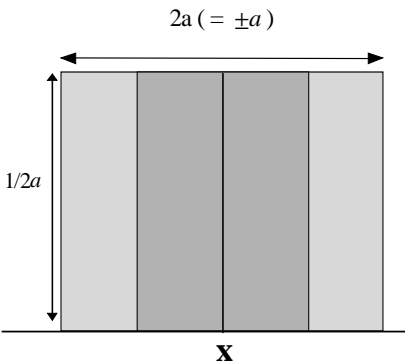
Appendix E. Useful Statistical Procedures

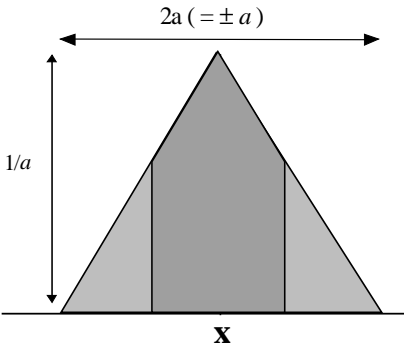
E.1 Distribution functions

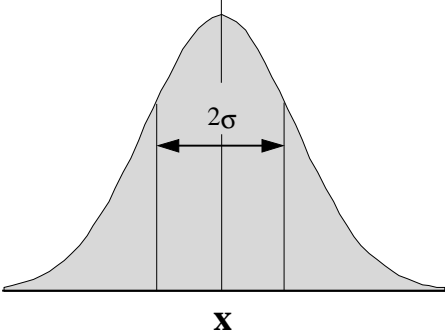
The following table shows how to calculate a standard uncertainty from the parameters of the two most important distribution functions, and gives an indication of the circumstances in which each should be used.

EXAMPLE

A chemist estimates a contributory factor as not less than 7 or more than 10, but feels that the value could be anywhere in between, with no idea of whether any part of the range is more likely than another. This is a description of a rectangular distribution function with a range $2a=3$ (semi range of $a=1.5$). Using the function below for a rectangular distribution, an estimate of the standard uncertainty can be calculated. Using the above range, $a=1.5$, results in a standard uncertainty of $(1.5/\sqrt{3}) = 0.87$.

Rectangular distribution		
Form	Use when:	Uncertainty
	<ul style="list-style-type: none"> A certificate or other specification gives limits without specifying a level of confidence (e.g. 25ml \pm 0.05ml) An estimate is made in the form of a maximum range ($\pm a$) with no knowledge of the shape of the distribution. 	$u(x) = \frac{a}{\sqrt{3}}$

Triangular distribution		
Form	Use when:	Uncertainty
	<ul style="list-style-type: none"> The available information concerning x is less limited than for a rectangular distribution. Values close to x are more likely than near the bounds. An estimate is made in the form of a maximum range ($\pm a$) described by a symmetric distribution. 	$u(x) = \frac{a}{\sqrt{6}}$

Normal distribution		
Form	Use when:	Uncertainty
 <p style="text-align: center;">X</p>	<ul style="list-style-type: none"> ▪ An estimate is made from repeated observations of a randomly varying process. • An uncertainty is given in the form of a standard deviation s, a relative standard deviation s/\bar{x}, or a coefficient of variance CV% without specifying the distribution. • An uncertainty is given in the form of a 95% (or other) confidence interval $x \pm c$ without specifying the distribution. 	<p>$u(x) = s$</p> <p>$u(x) = s$</p> <p>$u(x) = x \cdot (s / \bar{x})$</p> <p>$u(x) = \frac{CV\%}{100} \cdot x$</p> <p>$u(x) = c / 2$ (for c at 95%)</p> <p>$u(x) = c / 3$ (for c at 99.7%)</p>

E.2 Spreadsheet method for uncertainty calculation

E.2.1 Spreadsheet software can be used to simplify the calculations shown in Section 8. The procedure takes advantage of an approximate numerical method of differentiation, and requires knowledge only of the calculation used to derive the final result (including any necessary correction factors or influences) and of the numerical values of the parameters and their uncertainties. The description here follows that of Kragten [H.12].

E.2.2 In the expression for $u(y(x_1, x_2, \dots, x_n))$

$$\sqrt{\sum_{i=1,n} \left(\frac{\partial y}{\partial x_i} \cdot u(x_i) \right)^2 + \sum_{i,k=1,n} \left(\frac{\partial y}{\partial x_i} \cdot \frac{\partial y}{\partial x_k} \cdot u(x_i, x_k) \right)}$$

provided that either $y(x_1, x_2, \dots, x_n)$ is linear in x_i or $u(x_i)$ is small compared to x_i , the partial differentials $(\partial y / \partial x_i)$ can be approximated by:

$$\frac{\partial y}{\partial x_i} \approx \frac{y(x_i + u(x_i)) - y(x_i)}{u(x_i)}$$

Multiplying by $u(x_i)$ to obtain the uncertainty $u(y, x_i)$ in y due to the uncertainty in x_i gives

$$u(y, x_i) \approx y(x_1, x_2, \dots, (x_i + u(x_i)), \dots, x_n) - y(x_1, x_2, \dots, x_i, \dots, x_n)$$

Thus $u(y, x_i)$ is just the difference between the values of y calculated for $[x_i + u(x_i)]$ and x_i respectively.

E.2.3 The assumption of linearity or small values of $u(x_i)/x_i$ will not be closely met in all cases. Nonetheless, the method does provide acceptable accuracy for practical purposes when considered against the necessary approximations made in estimating the values of $u(x_i)$. Reference H.12 discusses the point more fully and suggests methods of checking the validity of the assumption.

E.2.4 The basic spreadsheet is set up as follows, assuming that the result y is a function of the four parameters p, q, r , and s :

- i) Enter the values of p, q , etc. and the formula for calculating y in column A of the spreadsheet. Copy column A across the following columns once for every variable in y (see Figure E2.1). It is convenient to place the values of the uncertainties $u(p), u(q)$ and so on in row 1 as shown.
- ii) Add $u(p)$ to p in cell B3, $u(q)$ to q in cell C4 etc., as in Figure E2.2. On recalculating the spreadsheet, cell B8 then becomes

$f(p+u(p), q, r, \dots)$ (denoted by $f(p', q, r, \dots)$ in Figures E2.2 and E2.3), cell C8 becomes $f(p, q+u(q), r, \dots)$ etc.

- iii) In row 9 enter row 8 minus A8 (for example, cell B9 becomes B8-A8). This gives the values of $u(y, p)$ as

$$u(y, p) = f(p+u(p), q, r, \dots) - f(p, q, r, \dots) \text{ etc.}$$

- iv) To obtain the standard uncertainty on y , these individual contributions are squared, added together and then the square root taken, by entering $u(y, p)^2$ in row 10 (Figure E2.3) and putting the square root of their sum in A10. That is, cell A10 is set to the formula

$$\text{SQRT}(\text{SUM}(\text{B10}+\text{C10}+\text{D10}+\text{E10}))$$

which gives the standard uncertainty on y .

E.2.5 The contents of the cells B10, C10 etc. show the squared contributions $u(y, x_i)^2 = (c_i u(x_i))^2$ of the individual uncertainty components to the uncertainty on y and hence it is easy to see which components are significant.

E.2.6 It is straightforward to allow updated calculations as individual parameter values change or uncertainties are refined. In step i) above, rather than copying column A directly to columns B-E, copy the values p to s by reference, that is, cells B3 to E3 all reference A3, B4 to E4 reference A4 etc. The horizontal arrows in Figure E2.1 show the referencing for row 3. Note that cells B8 to E8 should still reference the values in columns B to E respectively, as shown for column B by the vertical arrows in Figure E2.1. In step ii) above, add the references to row 1 by reference (as shown by the arrows in Figure E2.1). For example, cell B3 becomes A3+B1, cell C4 becomes A4+C1 etc. Changes to either parameters or uncertainties will then be reflected immediately in the overall result at A8 and the combined standard uncertainty at A10.

E.2.7 If any of the variables are correlated, the necessary additional term is added to the SUM in A10. For example, if p and q are correlated, with a correlation coefficient $r(p, q)$, then the extra term $2 \times r(p, q) \times u(y, p) \times u(y, q)$ is added to the calculated sum before taking the square root. Correlation can therefore easily be included by adding suitable extra terms to the spreadsheet.

Figure E2.1

	A	B	C	D	E
1		$u(p)$	$u(q)$	$u(r)$	$u(s)$
2					
3	p	p	p	p	p
4	q	q	q	q	q
5	r	r	r	r	r
6	s	s	s	s	s
7					
8	$y=f(p,q,...)$	$y=f(p,q,...)$	$y=f(p,q,...)$	$y=f(p,q,...)$	$y=f(p,q,...)$
9					
10					
11					

Figure E2.2

	A	B	C	D	E
1		$u(p)$	$u(q)$	$u(r)$	$u(s)$
2					
3	p	$p+u(p)$	p	p	p
4	q	q	$q+u(q)$	q	q
5	r	r	r	$r+u(r)$	r
6	s	s	s	s	$s+u(s)$
7					
8	$y=f(p,q,...)$	$y=f(p',...)$	$y=f(..q',...)$	$y=f(..r',...)$	$y=f(..s',...)$
9		$u(y,p)$	$u(y,q)$	$u(y,r)$	$u(y,s)$
10					
11					

Figure E2.3

	A	B	C	D	E
1		$u(p)$	$u(q)$	$u(r)$	$u(s)$
2					
3	p	$p+u(p)$	p	p	p
4	q	q	$q+u(q)$	q	q
5	r	r	r	$r+u(r)$	r
6	s	s	s	s	$s+u(s)$
7					
8	$y=f(p,q,...)$	$y=f(p',...)$	$y=f(..q',...)$	$y=f(..r',...)$	$y=f(..s',...)$
9		$u(y,p)$	$u(y,q)$	$u(y,r)$	$u(y,s)$
10	$u(y)$	$u(y,p)^2$	$u(y,q)^2$	$u(y,r)^2$	$u(y,s)^2$
11					

E.3 Uncertainties from linear least squares calibration

E.3.1 An analytical method or instrument is often calibrated by observing the responses, y , to different levels of the analyte, x . In most cases this relationship is taken to be linear viz:

$$y = b_0 + b_1 x \quad \text{Eq. E3.1}$$

This calibration line is then used to obtain the concentration x_{pred} of the analyte from a sample which produces an observed response y_{obs} from

$$x_{pred} = (y_{obs} - b_0)/b_1 \quad \text{Eq. E3.2}$$

It is usual to determine the constants b_1 and b_0 by weighted or un-weighted least squares regression on a set of n pairs of values (x_i, y_i) .

E.3.2 There are four main sources of uncertainty to consider in arriving at an uncertainty on the estimated concentration x_{pred} :

- Random variations in measurement of y , affecting both the reference responses y_i and the measured response y_{obs} .
- Random effects resulting in errors in the assigned reference values x_i .
- Values of x_i and y_i may be subject to a constant unknown offset, for example arising when the values of x are obtained from serial dilution of a stock solution
- The assumption of linearity may not be valid

Of these, the most significant for normal practice are random variations in y , and methods of estimating uncertainty for this source are detailed here. The remaining sources are also considered briefly to give an indication of methods available.

E.3.3 The uncertainty $u(x_{pred}, y)$ in a predicted value x_{pred} due to variability in y can be estimated in several ways:

From calculated variance and covariance.

If the values of b_1 and b_0 , their variances $\text{var}(b_1)$, $\text{var}(b_0)$ and their covariance, $\text{covar}(b_1, b_0)$, are determined by the method of least squares, the variance on x , $\text{var}(x)$, obtained using the formula in Chapter 8. and differentiating the normal equations, is given by

$$\text{var}(x_{pred}) = \frac{\text{var}(y_{obs}) + x_{pred}^2 \cdot \text{var}(b_1) + 2 \cdot x_{pred} \cdot \text{covar}(b_0, b_1) + \text{var}(b_0)}{b_1^2} \quad \text{Eq. E3.3}$$

and the corresponding uncertainty $u(x_{pred}, y)$ is $\sqrt{\text{var}(x_{pred})}$.

From the calibration data.

The above formula for $\text{var}(x_{pred})$ can be written in terms of the set of n data points, (x_i, y_i) , used to determine the calibration function:

$$\text{var}(x_{pred}) = \text{var}(y_{obs}) / b_1^2 + \frac{S^2}{b_1^2} \cdot \left(\frac{1}{\sum w_i} + \frac{(x_{pred} - \bar{x})^2}{(\sum w_i x_i^2) - (\sum w_i x_i)^2 / \sum w_i} \right) \quad \text{Eq. E3.4}$$

where $S^2 = \frac{\sum w_i (y_i - y_{fi})^2}{(n - 2)}$, $(y_i - y_{fi})$ is the

residual for the i^{th} point, n is the number of data points in the calibration, b_1 the calculated best fit gradient, w_i the weight assigned to y_i and $(x_{pred} - \bar{x})$ the difference between x_{pred} and the mean \bar{x} of the n values x_1, x_2, \dots

For unweighted data and where $\text{var}(y_{obs})$ is based on p measurements, equation E3.4 becomes

$$\text{var}(x_{pred}) = \frac{S^2}{b_1^2} \cdot \left(\frac{1}{p} + \frac{1}{n} + \frac{(x_{pred} - \bar{x})^2}{(\sum x_i^2) - (\sum x_i)^2 / n} \right) \quad \text{Eq. E3.5}$$

This is the formula which is used in example 5 with $S_{xx} = \left[\sum (x_i^2) - \left(\sum x_i \right)^2 / n \right] = \sum (x_i - \bar{x})^2$.

From information given by software used to derive calibration curves.

Some software gives the value of S , variously described for example as RMS error or residual standard error. This can then be used in equation E3.4 or E3.5. However some software may also give the standard deviation $s(y_c)$ on a value of y calculated from the fitted line for some new value of x and this can be used to calculate $\text{var}(x_{pred})$ since, for $p=1$

$$s(y_c) = S \sqrt{1 + \frac{1}{n} + \frac{(x_{pred} - \bar{x})^2}{(\sum x_i^2) - (\sum x_i)^2 / n}}$$

giving, on comparison with equation E3.5,

$$\text{var}(x_{\text{pred}}) = [s(y_c) / b_1]^2 \quad \text{Eq. E3.6}$$

E.3.4 The reference values x_i may each have uncertainties which propagate through to the final result. In practice, uncertainties in these values are usually small compared to uncertainties in the system responses y_i , and may be ignored. An approximate estimate of the uncertainty $u(x_{\text{pred}}, x_i)$ in a predicted value x_{pred} due to uncertainty in a particular reference value x_i is

$$u(x_{\text{pred}}, x_i) \approx u(x_i)/n \quad \text{Eq. E3.7}$$

where n is the number of x_i values used in the calibration. This expression can be used to check the significance of $u(x_{\text{pred}}, x_i)$.

E.3.5 The uncertainty arising from the assumption of a linear relationship between y and x is not normally large enough to require an additional estimate. Providing the residuals show that there is no significant systematic deviation from this assumed relationship, the uncertainty arising from this assumption (in addition to that covered by the resulting increase in y variance) can be taken to be negligible. If the residuals show a systematic trend then it may be necessary to include higher

terms in the calibration function. Methods of calculating $\text{var}(x)$ in these cases are given in standard texts. It is also possible to make a judgement based on the size of the systematic trend.

E.3.6 The values of x and y may be subject to a constant unknown offset (e.g. arising when the values of x are obtained from serial dilution of a stock solution which has an uncertainty on its certified value). If the standard uncertainties on y and x from these effects are $u(y, \text{const})$ and $u(x, \text{const})$, then the uncertainty on the interpolated value x_{pred} is given by:

$$u(x_{\text{pred}})^2 = u(x, \text{const})^2 +$$

$$(u(y, \text{const})/b_1)^2 + \text{var}(x) \quad \text{Eq. E3.8}$$

E.3.7 The four uncertainty components described in E.3.2 can be calculated using equations Eq. E3.3 to Eq. E3.8. The overall uncertainty arising from calculation from a linear calibration can then be calculated by combining these four components in the normal way.

E.4: Documenting uncertainty dependent on analyte level

E.4.1 Introduction

E.4.1.1 It is often observed in chemical measurement that, over a large range of analyte levels, dominant contributions to the overall uncertainty vary approximately proportionately to the level of analyte, that is $u(x) \propto x$. In such cases it is often sensible to quote uncertainties as relative standard deviations or, for example, coefficient of variation (%CV).

E.4.1.2 Where the uncertainty is unaffected by level, for example at low levels, or where a relatively narrow range of analyte level is involved, it is generally most sensible to quote an absolute value for the uncertainty.

E.4.1.3 In some cases, both constant and proportional effects are important. This section sets out a general approach to recording uncertainty information where variation of uncertainty with analyte level is an issue and reporting as a simple coefficient of variation is inadequate.

E.4.2 Basis of approach

E.4.2.1 To allow for both proportionality of uncertainty and the possibility of an essentially constant value with level, the following general expression is used:

$$u(x) = \sqrt{s_0^2 + (x \cdot s_1)^2} \quad [1]$$

where

$u(x)$ is the combined standard uncertainty in the result x (that is, the uncertainty expressed as a standard deviation)

s_0 represents a constant contribution to the overall uncertainty

s_1 is a proportionality constant.

The expression is based on the normal method of combining of two contributions to overall uncertainty, assuming one contribution (s_0) is constant and one (xs_1) proportional to the result. Figure E.4.1 shows the form of this expression.

NOTE: The approach above is practical only where it is possible to calculate a large number of values. Where experimental study is employed, it will not often be possible to establish the relevant parabolic relationship. In such circumstances, an adequate

approximation can be obtained by simple linear regression through four or more combined uncertainties obtained at different analyte concentrations. This procedure is consistent with that employed in studies of reproducibility and repeatability according to ISO 5725:1994. The relevant expression is then $u(x) \approx s'_0 + x \cdot s'_1$

E.4.2.2 The figure can be divided into approximate regions (**A** to **C** on the figure):

- A:** The uncertainty is dominated by the term s_0 , and is approximately constant and close to s_0 .
- B:** Both terms contribute significantly; the resulting uncertainty is significantly higher than either s_0 or xs_1 , and some curvature is visible.
- C:** The term xs_1 dominates; the uncertainty rises approximately linearly with increasing x and is close to xs_1 .

E.4.2.3 Note that in many experimental cases the complete form of the curve will not be apparent. Very often, the whole reporting range of analyte level permitted by the scope of the method falls within a single chart region; the result is a number of special cases dealt with in more detail below.

E.4.3 Documenting level-dependent uncertainty data

E.4.3.1 In general, uncertainties can be documented in the form of a value for each of s_0 and s_1 . The values can be used to provide an uncertainty estimate across the scope of the method. This is particularly valuable when calculations for well characterised methods are implemented on computer systems, where the general form of the equation can be implemented independently of the values of the parameters (one of which may be zero - see below). It is accordingly recommended that, except in the special cases outlined below or where the dependence is strong but not linear*, uncertainties

* An important example of non-linear dependence is the effect of instrument noise on absorbance measurement at high absorbances near the upper limit of the instrument capability. This is particularly pronounced where absorbance is calculated from transmittance (as in infrared spectroscopy). Under these circumstances, baseline noise causes very large uncertainties in high absorbance figures, and the

are documented in the form of values for a constant term represented by s_0 and a variable term represented by s_I .

E.4.4. Special cases

E.4.4.1. Uncertainty not dependent on level of analyte (s_0 dominant)

The uncertainty will generally be effectively independent of observed analyte concentration when:

- The result is close to zero (for example, within the stated detection limit for the method). Region **A** in Figure E.4.1
- The possible range of results (stated in the method scope or in a statement of scope for the uncertainty estimate) is small compared to the observed level.

Under these circumstances, the value of s_I can be recorded as zero. s_0 is normally the calculated standard uncertainty.

E.4.4.2. Uncertainty entirely dependent on analyte (s_I dominant)

Where the result is far from zero (for example, above a 'limit of determination') and there is clear evidence that the uncertainty changes proportionally with the level of analyte permitted within the scope of the method, the term xs_I dominates (see Region **C** in Figure E.4.1). Under these circumstances, and where the method scope does not include levels of analyte near zero, s_0 may reasonably be recorded as zero and s_I is simply the uncertainty expressed as a relative standard deviation.

E.4.4.3. Intermediate dependence

In intermediate cases, and in particular where the situation corresponds to region **B** in Figure E.4.1, two approaches can be taken:

a) Applying variable dependence

The more general approach is to determine, record and use both s_0 and s_I . Uncertainty

uncertainty rises much faster than a simple linear estimate would predict. The usual approach is to reduce the absorbance, typically by dilution, to bring the absorbance figures well within the working range; the linear model used here will then normally be adequate. Other examples include the 'sigmoidal' response of some immunoassay methods.

estimates, when required, can then be produced on the basis of the reported result. This remains the recommended approach where practical.

NOTE: See the note to section E.4.2.

b) Applying a fixed approximation

An alternative which may be used in general testing and where

- the dependence is not strong (that is, evidence for proportionality is weak)
- or
- the range of results expected is moderate

leading in either case to uncertainties which do not vary by more than about 15% from an average uncertainty estimate, it will often be reasonable to calculate and quote a fixed value of uncertainty for general use, based on the mean value of results expected. That is,

either

a mean or typical value for x is used to calculate a fixed uncertainty estimate, and this is used in place of individually calculated estimates

or

a single standard deviation has been obtained, based on studies of materials covering the full range of analyte levels permitted (within the scope of the uncertainty estimate), and there is little evidence to justify an assumption of proportionality. This should generally be treated as a case of zero dependence, and the relevant standard deviation recorded as s_0 .

E.4.5. Determining s_0 and s_I

E.4.5.1. In the special cases in which one term dominates, it will normally be sufficient to use the uncertainty as standard deviation or relative standard deviation respectively as values of s_0 and s_I . Where the dependence is less obvious, however, it may be necessary to determine s_0 and s_I indirectly from a series of estimates of uncertainty at different analyte levels.

E.4.5.2. Given a calculation of combined uncertainty from the various components, some of which depend on analyte level while others do not, it will normally be possible to investigate the dependence of overall uncertainty on analyte level by simulation. The procedure is as follows:

- 1: Calculate (or obtain experimentally) uncertainties $u(x_i)$ for at least ten levels x_i of analyte, covering the full range permitted.
2. Plot $u(x_i)^2$ against x_i^2
3. By linear regression, obtain estimates of m and c for the line $u(x)^2 = mx^2 + c$
4. Calculate s_0 and s_I from $s_0 = \sqrt{c}$, $s_I = \sqrt{m}$
5. Record s_0 and s_I

E.4.6. Reporting

E.4.6.1. The approach outlined here permits estimation of a standard uncertainty for any single result. In principle, where uncertainty information is to be reported, it will be in the form of

$$[\text{result}] \pm [\text{uncertainty}]$$

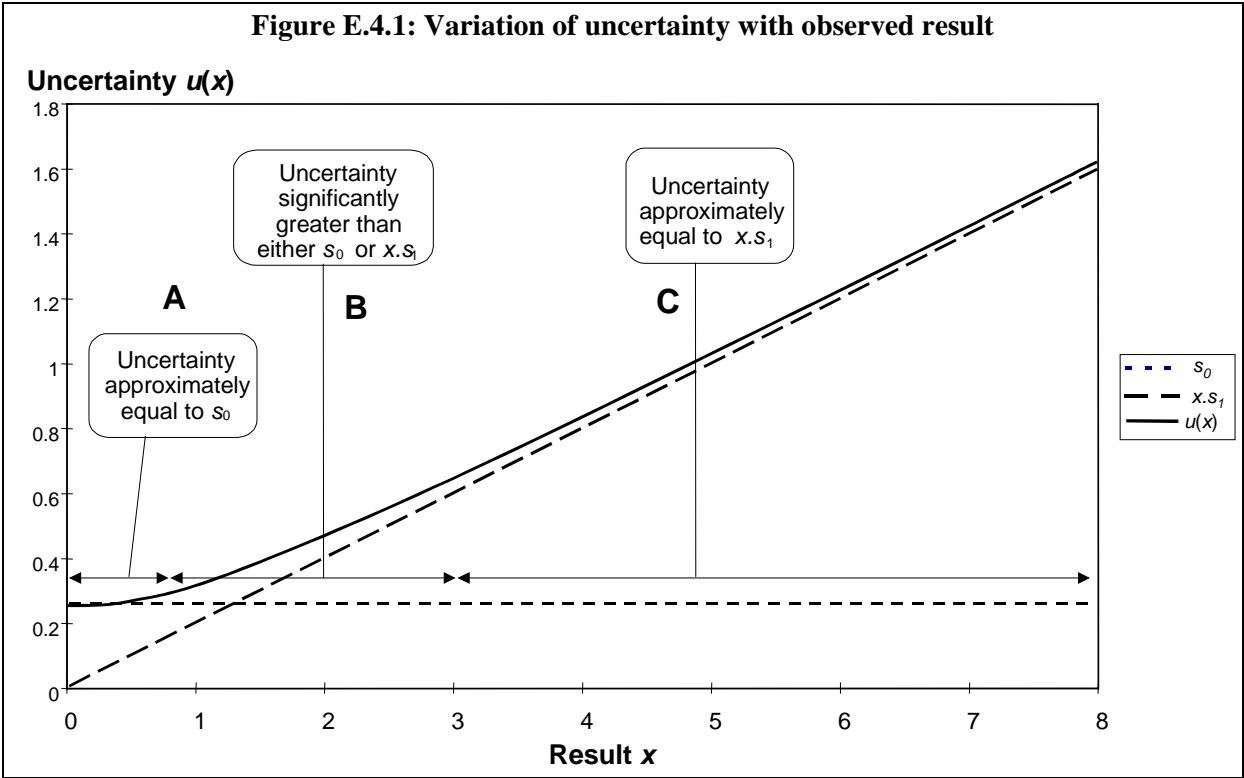
where the uncertainty as standard deviation is calculated as above, and if necessary expanded (usually by a factor of two) to give increased confidence. Where a number of results are reported together, however, it may be possible, and is perfectly acceptable, to give an estimate of uncertainty applicable to all results reported.

E.4.6.2. Table E.4.1 gives some examples. The uncertainty figures for a list of different analytes may usefully be tabulated following similar principles.

NOTE: Where a 'detection limit' or 'reporting limit' is used to give results in the form "<x" or "nd", it will normally be necessary to quote the limits used in addition to the uncertainties applicable to results above reporting limits.

Table E.4.1: Summarising uncertainty for several samples

Situation	Dominant term	Reporting example(s)
Uncertainty essentially constant across all results	s_0 or fixed approximation (sections E.4.4.1. or E.4.4.3.a)	Standard deviation: expanded uncertainty; 95% confidence interval
Uncertainty generally proportional to level	xs_I (see section E.4.4.2.)	relative standard deviation; coefficient of variance (%CV)
Mixture of proportionality and lower limiting value for uncertainty	Intermediate case (section E.4.4.3.)	quote %CV or rsd together with lower limit as standard deviation.



Appendix F. Measurement Uncertainty at the Limit of Detection/Limit of Determination

F.1. Introduction

F.1.1. At low concentrations, an increasing variety of effects becomes important, including, for example,

- the presence of noise or unstable baseline,
- the contribution of interferences to the (gross) signal,
- the influence of any analytical blank used, and
- losses during extraction, isolation or clean-up.

Because of such effects, as analyte concentrations drop, the relative uncertainty associated with the result tends to increase, first to a substantial fraction of the result and finally 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.

NOTE: The terminology and conventions associated with measuring and reporting low levels of analyte have been widely discussed elsewhere (See Bibliography [H.16, H.17, H.18] for examples and definitions). Here, the term ‘limit of detection’ only implies a level at which detection becomes problematic, and is not associated with any specific definition.

F.1.2. It is widely accepted that the most important use of the ‘limit of detection’ is to show where method performance becomes insufficient for acceptable quantitation, so that improvements can be made. Ideally, therefore, quantitative measurements should not be made in this region. Nonetheless, so many materials are important at very low levels that it is inevitable that measurements must be made, and results reported, in this region.

F.1.3. The ISO Guide on Measurement Uncertainty [H.2] does not give explicit instructions for the estimation of uncertainty when the results are small and the uncertainties large compared to the results. Indeed, the basic form of the ‘law of propagation of uncertainties’, described in chapter 8 of this guide, may cease to apply accurately in this region; one assumption on which the calculation is based is that the

uncertainty is small relative to the value of the measurand. An additional, if philosophical, difficulty follows from the definition of uncertainty given by the ISO Guide: though negative observations are quite possible, and even common in this region, an implied dispersion including values below zero cannot be “... reasonably ascribed to the value of the measurand” when the measurand is a concentration, because concentrations themselves cannot be negative.

F.1.4. These difficulties do not preclude the application of the methods outlined in this guide, but some caution is required in interpretation and reporting the results of measurement uncertainty estimation in this region. The purpose of the present Appendix is to provide limited guidance to supplement that already available from other sources.

NOTE: Similar considerations may apply to other regions; for example, mole or mass fractions close to 100% may lead to similar difficulties.

F.2. Observations and estimates

F.2.1. A fundamental principle of measurement science is that *results are estimates of true values*. Analytical results, for example, are available initially in units of the observed signal, e.g. mV, absorbance units *etc.* For communication to a wider audience, particularly to the customers of a laboratory or to other authorities, the raw data need to be converted to a chemical quantity, such as concentration or amount of substance. This conversion typically requires a calibration procedure (which may include, for example, corrections for observed and well characterised losses). Whatever the conversion, however, the figure generated remains an observation, or signal. If the experiment is properly carried out, this observation remains the ‘best estimate’ of the value of the measurand.

F.2.2. Observations are not often constrained by the same fundamental limits that apply to real concentrations. For example, it is perfectly sensible to report an ‘observed concentration’,

that is, an estimate, below zero. It is equally sensible to speak of a dispersion of possible *observations* which extends into the same region. For example, when performing an unbiased measurement on a sample with no analyte present, one *should* see about half of the observations falling below zero. In other words, reports like

$$\text{observed concentration} = 2.4 \pm 8 \text{ mg l}^{-1}$$

$$\text{observed concentration} = -4.2 \pm 8 \text{ mg l}^{-1}$$

are not only possible; they should be seen as valid statements.

F.2.3. The methods of uncertainty estimation described in this guide apply well to the estimation of uncertainties on observations. It follows that while reporting observations and their associated uncertainties to an informed audience, there is no barrier to, or contradiction in, reporting the best estimate and its associated uncertainty even where the result implies an impossible physical situation. Indeed, in some circumstances (for example, when reporting a value for an analytical blank which will subsequently be used to correct other results) it is absolutely essential to report the observation and its uncertainty (however large).

F.2.4. This remains true wherever the end use of the result is in doubt. Since only the observation and its associated uncertainty can be used directly (for example, in further calculations, in trend analysis or for re-interpretation), the uncensored observation should always be available.

F.2.5. The ideal is accordingly to report valid observations and their associated uncertainty regardless of the values.

F.3. Interpreted results and compliance statements

F.3.1. Despite the foregoing, it must be accepted that many reports of analysis and statements of compliance include some interpretation for the end user's benefit. Typically, such an interpretation would include any relevant inference about the levels of analyte which could reasonably be present in a material. Such an interpretation is an inference about the real world, and consequently would be expected (by the end user) to conform to real limits. So, too, would any associated estimate of uncertainty in 'real' values.

F.3.2. Under such circumstances, where the end use is well understood, and where the end user cannot realistically be informed of the nature of measurement observations, the general guidance provided elsewhere (for example in references H.16, H.17, H.18) on the reporting of low level results may reasonably apply.

F.3.3. One further caution is, however, pertinent. Much of the literature on capabilities of detection relies heavily on the statistics of repeated observations. It should be clear to readers of the current guide that observed variation is only rarely a good guide to the full uncertainty of results. Just as with results in any other region, careful consideration should accordingly be given to all the uncertainties affecting a given result before reporting the values.

Appendix G. Common Sources and Values of Uncertainty

The following tables summarise some typical examples of uncertainty components. The tables give:

- The particular measurand or experimental procedure (determining mass, volume *etc*)
- The main components and sources of uncertainty in each case
- A suggested method of determining the uncertainty arising from each source.
- An example of a typical case

The tables are intended only to summarise the examples and to indicate general methods of estimating uncertainties in analysis. They are not intended to be comprehensive, nor should the values given be used directly without independent justification. The values may, however, help in deciding whether a particular component is significant.

Determination	Uncertainty Components	Cause	Method of determination	Typical values	
				Example	Value
Mass	Balance calibration uncertainty	Limited accuracy in calibration	Stated on calibration certificate, converted to standard deviation	4-figure balance	0.5 mg
	Linearity		i) Experiment, with range of certified weights ii) Manufacturer's specification		ca. 0.5x last significant digit
	Readability	Limited resolution on display or scale	From last significant digit		0.5x last significant digit/ $\sqrt{3}$
	Daily drift	Various, including temperature	Standard deviation of long term check weighings. Calculate as RSD if necessary.		ca. 0.5x last significant digit.
	Run to run variation	Various	Standard deviation of successive sample or check weighings		ca. 0.5x last significant digit.
	Density effects (<i>conventional</i> basis) ^{Note 1}	Calibration weight/sample density mismatch causes a difference in the effect of atmospheric buoyancy	Calculated from known or assumed densities and typical atmospheric conditions	Steel, Nickel Aluminium Organic solids Water Hydrocarbons	1 ppm 20 ppm 50-100 ppm 65 ppm 90 ppm
	Density effects (<i>in vacuo</i> basis) ^{Note 1}	As above.	Calculate atmospheric buoyancy effect and subtract buoyancy effect on calibration weight.	100 g water 10 g Nickel	+0.1g (effect) <1 mg (effect)

Note 1. For fundamental constants or SI unit definitions, mass determinations by weighing are usually corrected to the weight in vacuum. In most other practical situations, weight is quoted on a *conventional* basis as defined by OIML [H.18]. The convention is to quote weights at an air density of 1.2 kg m^{-3} and a sample density of 8000 kg m^{-3} , which corresponds to weighing steel at sea level in normal atmospheric conditions. The buoyancy correction to conventional mass is zero when the sample density is 8000 kg m^{-3} or the air density is 1.2 kg m^{-3} . Since the air density is usually very close to the latter value, correction to conventional weight can normally be neglected. The standard uncertainty values given for density-related effects on a conventional weight basis in the table above are sufficient for preliminary estimates for weighing on a conventional basis without buoyancy correction at sea level. Mass determined on the conventional basis may, however, differ from the 'true mass' (*in vacuo*) by 0.1% or more (see the effects in the bottom line of the table above).

Determination	Uncertainty Components	Cause	Method of determination	Typical values	
				Example	Value
Volume (liquid)	Calibration uncertainty	Limited accuracy in calibration	Stated on manufacturer's specification, converted to standard deviation. For ASTM class A glassware of volume V, the limit is approximately $V^{0.6}/200$	10 ml (Grade A)	$0.02 / \sqrt{3} = 0.01 \text{ ml}^*$
	Temperature	Temperature variation from the calibration temperature causes a difference in the volume at the standard temperature.	$\Delta T \cdot \alpha / (2\sqrt{3})$ gives the relative standard deviation, where ΔT is the possible temperature range and α the coefficient of volume expansion of the liquid. α is approximately $2 \times 10^{-4} \text{ K}^{-1}$ for water and $1 \times 10^{-3} \text{ K}^{-1}$ for organic liquids.	100 ml water	0.03 ml for operating within 3°C of the stated operating temperature
	Run to run variation	Various	Standard deviation of successive check deliveries (found by weighing)	25 ml pipette	Replicate fill/weight: $s = 0.0092 \text{ ml}$

* Assuming rectangular distribution

Determination	Uncertainty Components	Cause	Method of determination	Typical values	
				Example	Value
Reference material concentration	Purity	Impurities reduce the amount of reference material present. Reactive impurities may interfere with the measurement.	Stated on manufacturer's certificate. Reference certificates usually give unqualified limits; these should accordingly be treated as rectangular distributions and divided by $\sqrt{3}$. Note: where the nature of the impurities is not stated, additional allowance or checks may need to be made to establish limits for interference etc.	Reference potassium hydrogen phthalate certified as 99.9 $\pm 0.1\%$	$0.1/\sqrt{3} = 0.06\%$
	Concentration (certified)	Certified uncertainty in reference material concentration.	Stated on manufacturer's certificate. Reference certificates usually give unqualified limits; these should accordingly be treated as rectangular distributions and divided by $\sqrt{3}$.	Cadmium acetate in 4% acetic acid. Certified as $(1000 \pm 2) \text{ mg l}^{-1}$.	$2/\sqrt{3} = 1.2 \text{ mg l}^{-1}$ (0.0012 as RSD)*
	Concentration (made up from certified material)	Combination of uncertainties in reference values and intermediate steps	Combine values for prior steps as RSD throughout.	Cadmium acetate after three dilutions from 1000 mg l^{-1} to 0.5 mg l^{-1}	$\sqrt{0.0012^2 + 0.0017^2 + 0.0021^2 + 0.0017^2} = 0.0034$ as RSD

* Assuming rectangular distribution

Determination	Uncertainty Components	Cause	Method of determination	Typical values	
				Example	Value
Absorbance	Instrument calibration Note: this component relates to absorbance reading versus reference absorbance, not to the calibration of concentration against absorbance reading	Limited accuracy in calibration.	Stated on calibration certificate as limits, converted to standard deviation		
	Run to run variation	Various	Standard deviation of replicate determinations, or QA performance.	Mean of 7 absorbance readings with $s=1.63$	$1.63/\sqrt{7} = 0.62$
Sampling	Homogeneity	Sub-sampling from inhomogeneous material will not generally represent the bulk exactly. Note: random sampling will generally result in zero bias. It may be necessary to check that sampling is actually random.	i) Standard deviation of separate sub-sample results (if the inhomogeneity is large relative to analytical accuracy). ii) Standard deviation estimated from known or assumed population parameters.	Sampling from bread of assumed two-valued inhomogeneity (See Example A4)	For 15 portions from 72 contaminated and 360 uncontaminated bulk portions: $RSD = 0.58$

Determination	Uncertainty Components	Cause	Method of determination	Typical values	
				Example	Value
Extraction recovery	Mean recovery	Extraction is rarely complete and may add or include interferences.	Recovery calculated as percentage recovery from comparable reference material or representative spiking. Uncertainty obtained from standard deviation of mean of recovery experiments. Note: recovery may also be calculated directly from previously measured partition coefficients.	Recovery of pesticide from bread; 42 experiments, mean 90%, s=28% (See Example A4)	$28/\sqrt{42}=4.3\%$ (0.048 as RSD)
	Run to run variation in recovery	Various	Standard deviation of replicate experiments.	Recovery of pesticides from bread from paired replicate data. (See Example A4)	0.31 as RSD.

Appendix H. Bibliography

- H.1. ISO/IEC 17025:1999. General Requirements for the Competence of Calibration and Testing Laboratories. ISO, Geneva (1999).
- H.2. Guide To The Expression Of Uncertainty In Measurement. ISO, Geneva (1993). (ISBN 92-67-10188-9) (Reprinted 1995)
- H.3. EURACHEM, Quantifying Uncertainty in Analytical Measurement. Laboratory of the Government Chemist, London (1995). ISBN 0-948926-08-2
- H.4. International Vocabulary of basic and general terms in Metrology. ISO, Geneva, (1993). (ISBN 92-67-10175-1)
- H.5. ISO 3534:1993. Statistics - Vocabulary and Symbols. ISO, Geneva, Switzerland (1993).
- H.6. Analytical Methods Committee, *Analyst* (London). **120** 29-34 (1995).
- H.7. EURACHEM, The Fitness for Purpose of Analytical Methods. (1998) (ISBN 0-948926-12-0)
- H.8. ISO/IEC Guide 33:1989, "Uses of Certified Reference Materials". ISO, Geneva (1989).
- H.9. International Union of Pure and Applied Chemistry. Pure Appl. Chem., **67**, 331-343, (1995).
- H.10. ISO 5725:1994 (Parts 1-4 and 6). "Accuracy (trueness and precision) of measurement methods and results". ISO, Geneva (1994). See also ISO 5725-5:1998 for alternative methods of estimating precision.
- H.11. I. J. Good, "Degree of Belief", in Encyclopaedia of Statistical Sciences, Vol. 2, Wiley, New York (1982).
- H.12. J. Kragten, "Calculating standard deviations and confidence intervals with a universally applicable spreadsheet technique", *Analyst*, **119**, 2161-2166 (1994).
- H.13. British Standard BS 6748:1986. Limits of metal release from ceramic ware, glassware, glass ceramic ware and vitreous enamel ware.
- H.14. S. L. R. Ellison, V. J. Barwick. Accred. Qual. Assur. **3** 101-105 (1998).
- H.15. ISO 9004-4:1993, Total Quality Management. Part 2. Guidelines for quality improvement. ISO, Geneva (1993).
- H.16. H. Kaiser, *Anal. Chem.* **42** 24A (1970).
- H.17. L.A. Currie, *Anal. Chem.* **40** 583 (1968).
- H.18. IUPAC, Limit of Detection, *Spectrochim. Acta* **33B** 242 (1978).
- H.19. OIML Recommendations IR 33 Conventional value of the result of weighing in air.

Quality Assurance for Research and Development and Non-routine Analysis



Co-Operation on International Traceability in Analytical Chemistry

This document has been produced primarily by a joint EURACHEM / CITAC Working Group, the membership of which is

Prof C Adams, Unilever, UK
Prof K Cammann, ICBFhM, Germany
ir HA Deckers, RvA, Netherlands
Prof Z Dobkowski, Ind. Chem. Res. Inst., Poland
Mr D Holcombe, LGC, UK
Dr PD LaFleur, Kodak, USA
Dr P Radvila, EMPA, Switzerland
Dr C Rohrer, Lenzing AG, Austria
Dr W Steck, BASF AG, Germany
ir P Vermaercke, S.C.K., Belgium

English Edition

Second Internet Version, November 1998

First Edition October 1998

Quality Assurance for Research and Development and Non-routine Analysis

This document has been produced primarily by a joint EURACHEM / CITAC Working Group, the membership of which is listed in Annex A. The secretary would also like to thank all of those individuals and organisations who have contributed comments, advice and background documentation.

Production of this Guide was in part supported under contract with the UK Department of Trade and Industry as part of the National Measurement System Valid Analytical Measurement (VAM) Programme.

Comments are invited on this document and should be sent to:

Mr David Holcombe

Drafting Secretary for EURACHEM / CITAC R&D Working Group

LGC, Queens Rd, Teddington, Middlesex, TW11 0LY, United Kingdom

☎: Int + 44 181 943 7613, 📠: Int + 44 181 943 2767, ✉: dgh@lgc.co.uk

English Edition 1.0, 1998

ISBN: 0 948926 11 2

Ruling language

The text may be freely translated into other languages, but where such action results in a dispute over interpretation, the guidance given in this English version is taken as being the definitive version.

Copyright of text

Copyright of the guidance presented in this guide is the property of the organisations represented by the working group members as listed in Annex A. All rights are reserved. This guidance may be downloaded for personal and non-commercial use, but no part of the guide may be distributed, published or stored in any retrieval system, in any media, by persons other than EURACHEM or CITAC members, without the written consent of the EURACHEM and CITAC Secretariats. Similarly, enquiries regarding the translation, production and distribution of new editions of this guide should be directed to the EURACHEM and CITAC Secretariats

This edition is Copyright © LGC (Teddington) Ltd, 1998

CONTENTS

Section	Title	Page
1	Aims and objectives	1
	Who this guide is for	1
	Using this Guide	1
	Emphasis of guidance	1
	Customers	2
2	Introduction	2
	What is Research and Development	2
	Importance of QA	3
	What needs to be controlled in R&D	3
3	Definitions	4
4	Principles of making Valid Analytical Measurements	5
5	Organisational quality elements	6
	Administrative and technical planning of work	6
	Quality management, corporate and local	7
	Record keeping and document control	7
	Staff qualifications, training and supervision of staff	8
	Equipment and computer controlled equipment	10
	Monitoring the quality system	10
	Subcontracting	12
6	Technical quality elements	12
	Unit operations	12
	Technical capability of laboratory	13
	Methodology	13
	Reagents, reference materials and calibrants	14
	Calibration and traceability	14
	Instrument performance	15
	Use of statistics	16
	Technical requirement related to particular unit processes (Sampling, isolation of analyte, measurement, validation, measurement uncertainty)	18
7	Analytical task quality elements	21
	Preparation and planning before starting work	21
	While the work is in progress	25
	When the work is complete	26
8	External verification	28
	Formal assessment against conventional quality assurance standards	29
	Benchmarking	30
	Visiting groups / Peer review	32
	Ranking of organisations	33
	External quality assessment procedures	34
	Conclusions	34
9	Bibliography and references	35
Annex A	- EURACHEM / CITAC Working Group	42
Annex B	- Flowchart showing lifecycle of an R&D project	44
Annex C	- Questionnaire for Analytical Work	45

Annex D	- Concepts for the accreditation of R&D tests by type.	46
Annex E	- R&D to develop analytical instrumentation	49

1. AIMS AND OBJECTIVES

1.1 *Who this guide is for*

- 1.1.1 This guide is intended to be used by managers and analytical staff, both in industry and the academic world, involved in the planning, performance and management of non-routine measurements in analytical science and associated research and development. Those responsible for the evaluation of the quality of such work will also find the guide useful. It provides principles from which assessing organisations such as accreditation or certification bodies could specify assessment criteria.

1.2 *Using this guide*

- 1.2.1 This guide aims to state and promote quality assurance (QA) *good practice*, or at least practice that meets the professional standards of the peer group. Many of these practices have already been stated in an earlier CITAC guide (CG1)^[1], which provides advice for mainly routine analysis, and an earlier EURACHEM / WELAC guide ^[2], which advises on the interpretation of EN 45001 and ISO Guide 25 for chemistry laboratories. Predictably there is likely to be a high degree of overlap between what is good practice in a routine situation and what is good practice in a non-routine situation. To avoid duplication those practices are only repeated below where it has been considered appropriate that further clarification is necessary for non-routine purposes. Where the guidance has **not** been restated, reference to the relevant part of the CITAC guide has been stated instead. Thus this guide should be used in conjunction with CG1.

1.3 *Emphasis of guidance*

- 1.3.1 There is still much discussion as to how applicable the various established quality standards/protocols, such as ISO Guide 25 ^[3], EN 45001 ^[4], ISO 9000 ^[5], and OECD Principles of Good Laboratory Practice (GLP) ^[6], are to non-routine work. GLP is study based, and the studies often involve non-routine or developmental work. R&D is compatible with the design element of ISO 9001. However it is widely argued that non-routine work does not fit easily into a highly documented and formalised quality system. For this reason the guidance is directed towards good practice rather than compliance with formal standards. The two approaches are not necessarily at odds with one another, but compliance may occasionally place requirements which are considered to be over and above what is considered to be best practice. Conversely no single quality standard necessarily covers all the elements of activity which might be considered relevant as best practice. The aim is to produce guidelines for analysts, their customers, and their managers, and not a quality manual template for an organisation. Note also that external verification, such as can be provided against a formal quality standard, is not mandatory, even though it may be desirable in some cases.

- 1.3.2 It is anticipated that once this guide is published it may be possible for accreditation bodies and other authoritative organisations to adapt the text for compliance purposes, for example to the published standards/protocols mentioned in §1.3.1 above.

1.4 **Customers**

- 1.4.1 Non-routine work regulated by this guidance may be performed for a number of different types of customer, such as:
- other departments within the same organisation which lack the specialist skills the work demands;
 - external customers who commission specific tasks;
 - regulatory bodies which commission the work to help enforce law, regulatory or licencing requirements;
 - funding bodies which commission large work programmes, within which specific tasks lie.

2. INTRODUCTION

2.2 ***What is Research and Development (R&D)?***

- 2.2.1 **Research** is a scientific investigation aimed at discovering and applying new facts, techniques and natural laws ^[7]. At its heart is inquiry into the unknown, addressing questions not previously asked. Research is done by a wide range of organisations: universities and colleges; government agencies; industry and contract organisations. Research projects vary widely in content and also in style, from open ended exploration of concepts to working towards specific targets.

Development in an industrial context is the work done to finalise the specification of a new project or new manufacturing process. It uses many of the methods of scientific inquiry, and may generate much new knowledge, but its aim is to create practicable economic solutions.

The combined term **Research and Development** can be seen as the work in an industrial or government context concentrating on finding new or improved processes, products *etc.*, and also on ways of introducing such innovations.

The use of the term **R&D** may not wholly encompass the activities intended to be covered by the Guidelines, but has been adopted by the authors as the most appropriate and convenient single term.

2.2.2 These guidelines are intended to cover analytical testing or measurements where for various reasons the work is non-routine or necessary procedures are not already in place, for example:

- methods already exists for the analytical problem, but have not previously been applied to the particular type of sample now encountered. The existing methods need to be evaluated and extended or adapted as necessary;
- the analytical problem is entirely new, but may be tackled by applying existing methods or techniques;
- the analytical problem is entirely new, there is no established method, and something has to be developed from the beginning.

Annex E provides some additional ideas for those carrying out R&D to develop analytical instrumentation.

2.3 ***Importance of QA***

2.3.1 The importance of quality assurance is well established and accepted for routine analysis. It is less well established for R&D.

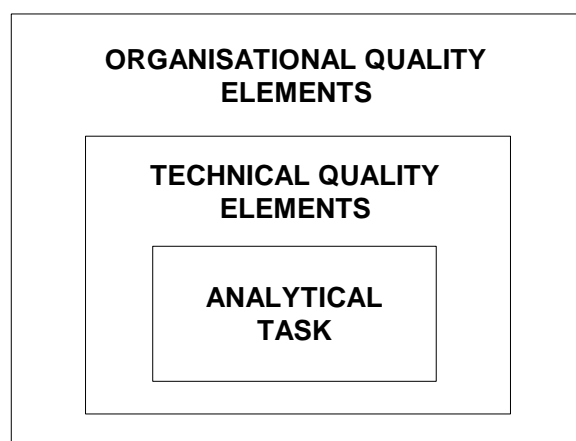


Figure 1: Nested Structure of activities

2.4 ***What needs to be controlled in R&D?***

2.4.1. Figure 1 shows a hierarchical approach to quality assurance within an organisation. The outer layer represents the elements of quality assurance that apply to all levels of activity within the organisation - so-called *organisational quality elements*. These are described in chapter 5. Examples at this level include a quality management structure with a defined role within the organisation; a quality system; documented procedures for key activities; a recruitment and training policy for all staff; *etc.*. The next layer, *technical quality elements*, described in chapter 6, forms a subset and comprises specific QA elements which apply to the technical activities of the organisation, such as policy and procedures for instrument calibration and performance

checks; use of calibrants and reference materials, and; use of statistical procedures. The inner layer, *analytical task quality elements*, described in chapter 7, represents the activities carried out for particular projects or individual analytical tasks. It includes the planning, control and reporting practices recommended at the start of, during, and at completion of R&D work.

3. DEFINITIONS

- 3.1 **Accreditation** - 'Procedure by which an authoritative body gives formal recognition that a body or person is competent to carry out specific tasks (ISO/CASCO 193 (Rev. 2), 1.11 ^[8], & ISO Guide 2:1996, 12.11) ^[9].
- 3.2 **Certification** - 'Procedure by which a third party gives written assurance that a product, process or service conforms to specified requirements (ISO/CASCO 193 (Rev. 2), 4.1.2 ^[8], & ISO Guide 2:1996, 15.1.2) ^[9].
- 3.3 **Contract** - An agreement made between two or more parties on specified terms. Typically as applied to analytical work it refers to an agreement between a laboratory (the contractor) to do work for the customer, at a specified price and within a specified timescale, with perhaps other conditions specified.
- 3.4 **Customer** - A purchaser of goods or services.
- 3.5 **Project** - 'a research or study assignment, a plan, scheme or proposal' ^[10]. In the analytical context a project refers to a discrete job starting with a particular problem and involving one or more tasks undertaken to solve the problem (see also study).
- 3.6 **Quality Assurance (QA)** - 'All the planned and systematic actions implemented within the quality system, and demonstrated as needed, to provide adequate confidence that an entity will fulfil requirements for quality.' (ISO 8402:1994, 3.5) ^[11].
- 3.7 **Quality Control (QC)** - 'Operational techniques and activities that are used to fulfil requirements for quality' (ISO 8402:1994, 3.4) ^[11].
- 3.8 **Registration** - 'Procedure by which a body indicates relevant characteristics of a product, process or service, or particulars of a body or person, in an appropriate, publicly available list (ISO/CASCO 193 (Rev. 2), 1.10 ^[8], & ISO Guide 2:1996, 12.10).
-

3.9 In ***routine*** analysis, the analytical problem will have been encountered before . A suitable validated method for solving the problem will exist and may be in regular use. The degree of associated staff training, calibration and quality control used with the method will depend on sample throughput.

3.10 ***Study*** - 'an attentive or detailed examination' ^[10].

N.B: use of the terms 'project' and 'study' in this guide do not mean that the guide is applicable only to GLP work

3.11.1 ***System (quality)*** - 'The organisational structure, procedures, processes and resources needed to implement quality management (ISO 8402:1994, 3.6) ^[11] .

3.11.2 *System* has been used in this guide to refer more generally to the infrastructure within which a laboratory undertakes analytical work and in this context does not necessarily constitute a quality system. This is entirely consistent with the ISO definition.

3.12 ***Task*** - No formal definition. The use of task in this guide denotes a small discrete piece of work, several tasks making up a project or study.

3.13 ***Validation*** - 'Confirmation by examination and provision of objective evidence that the particular requirement for a specified end use are fulfilled' (ISO 8402:1994, 2.18) ^[11] .

3.14 ***Verification*** - 'Confirmation by examination and provision of objective evidence that specified requirements have been fulfilled' (ISO 8402:1994, 2.17) ^[11] .

4. PRINCIPLES FOR MAKING VALID ANALYTICAL R&D MEASUREMENTS

4.1 Six basic principles have been identified as important for laboratories making measurements to follow ^[12].

- I. ***'Analytical measurements should be made to satisfy an agreed requirement'*** - In routine work it is usually a straightforward process to define the problem for which the analytical work is being carried out. In R&D specification of the problem is usually done as part of project definition. The customer may only have a vague idea of what the problem is and how chemical analysis can solve it, and will rely on the laboratory's technical expertise to design a suitable technical work-programme. Cost and time constraints will have to be considered as part of the

programme design. The programme will define how results will be reported and the importance of only using results in the appropriate context. Results can be badly misunderstood or misused if extrapolated outside the boundary conditions of the programme.

-
- II. ***‘Analytical measurements should be made using methods and equipment which have been tested to ensure they are fit for purpose’.*** Whatever type of measurements are made, suitable, well maintained and calibrated equipment is vital to ensure success. It is of the utmost importance that performance characteristics of methods should be evaluated to the extent necessary to show they are suitable for the measurements for which they are being used.
- III. ***‘Staff making analytical measurements should be both qualified and competent to undertake the task’.*** In R&D work it may not be possible to guarantee that the staff are totally competent as the full extent of the expertise required. The needs may not be fully appreciated when the work is started. It is possible that the analyst will not have much previous experience of the problem, but should have at least a basic knowledge of the underlying concepts involved in the work.
- IV. ***‘There should be regular independent assessment of the technical performance of a laboratory’.*** A laboratory's internal QC may indicate consistency in the measurements made within that laboratory. Independent assessment of the measurement capability by participation in proficiency testing schemes or measurement of well-characterised reference materials gives an idea of how well the laboratory's performance would compare with that of its peers. However it is recognised that the options for such independent assessment may be limited in an R&D environment.
- V. ***‘Analytical measurements made in one location should be consistent with those made elsewhere’.*** Use of reference materials (where available) and assessment of measurement uncertainty of the methods in use will help ensure traceability and compatibility with others making similar measurements.
- VI. ***‘Organisations making analytical measurements should have well defined quality control and quality assurance procedures’.*** All of the various measures taken to ensure quality of measurements within a laboratory should be incorporated into a quality system to ensure transparent and consistent implementation. If possible some sort of external audit is desirable to verify the working of this quality system.

5. ORGANISATIONAL QUALITY ELEMENTS

5.1 *Administrative and technical planning of the work- see also CITAC Guide CG1, section*

5.1.1 Laboratories which carry out analytical R&D need to have staff with suitable managerial and technical abilities to plan, control, deliver and report each project. This is considered in more detail in §7.1.3.

5.1.2 Where a laboratory is carrying out a number of projects simultaneously, coordination of the project management related to use of facilities is advised. Management needs to be aware of the different projects in progress in the laboratory at a given time and the corresponding risks of one project affecting another, both from a resource point of view but also from cross contamination. Similarly where projects are spread across several departments within a laboratory or involve input from external laboratories, suitable coordination is necessary to ensure coherent delivery of the work without any adverse effect on quality.

5.2 ***Quality management, corporate and local***

5.2.1 Regardless of whether the laboratory is formally recognised as compliant with a published quality management standard, it is recommended that it has a quality management system, whether formal or informal, through which its declared quality policy can be implemented. Typically this will involve staff with specific responsibilities for quality, who act as the focus and coordinators for quality matters within the laboratory. Quality also needs to be managed at various lower levels e.g. group, team or section. This may involve individuals having particular quality-related responsibilities as part of their duties and each member of staff should be aware of what role they have in the delivery of quality within the laboratory.

5.2.2 The management of quality in an R&D environment can be a delicate issue. A balance needs to be struck between maintaining a suitable level of control whilst at the same time not inhibiting creativity.

5.3 ***Record keeping and document control***

5.3.1 The purpose of keeping records is so that information and data held or gathered by the laboratory can be used to compile reports, make comparisons with other data (whether contemporary or historical), repeat work, and develop new or similar processes. Record keeping and document control are sufficiently important to justify a laboratory having a centralised policy, including relevant training for staff and competence assessment. The policy might typically cover:

- use of various types of media for record keeping;
- external considerations (such as recording requirements for patent applications);
- minimum levels of information for particular operations;
- use of forms and other approved formats;

- legibility, clarity, layout of information, and ease of data retrieval;
- traceability of records to time, date, analyst, sample, equipment, project;
- use of audit trails;
- authorisation of records by the use of signatures and other methods;
- methods for ensuring a record is complete;
- cross referencing copying restrictions;
- rules for amending and authorising amendments to records;
- rules for minimum retention of data, reports and other useful information.

5.3.2 Useful information should be recorded at the time or immediately after the work is completed.

5.3.3 Document control should be extended to all formal documents used in the analytical work, that is, those documents whose use is recognised within the quality system (as defined in the quality manual) and whose format, content and use has to be reviewed and authorised. It is not unusual for a laboratory to use a hierarchical approach for its quality system documentation. This ensures a maximum of flexibility as work patterns change. The table below shows four levels of formal document.

Level	Documentation	Subject / examples
1. (Highest)	Corporate quality policy	Quality manual
2.	<ul style="list-style-type: none"> • Formalised internal procedures operable across the laboratory • Other (external) normative documents 	Standard Operating Procedures (SOPs) Relevant laws, regulations, standards (ISO/CEN <i>etc.</i>), official methods (e.g. AOACI), Codes of Practice (COPs).
3.	Technical work instructions (specific applications)	In-house methods
4. (Lowest)	Records	Instrument logbooks, calibration records laboratory notebooks and other raw data, correspondence, reports

5.3.4 Clear responsibilities for document control should be assigned to staff. To maximise flexibility authorisation should be devolved as far down the management chain as possible, bearing in mind the need for those authorised to have sufficient expertise to make sound judgements.

5.3.5 For all controlled documents there should be a system for recalling and archiving versions of documents when they are upgraded or replaced. Suitable facilities for archiving information should be available and their use laid down within the document control policy. The use of

computer based systems is recommended to facilitate the control of documents but care is advised to ensure access to the system is only available to authorised staff.

5.4 ***Staff -qualifications, training and supervision of staff - see also CITAC CG1, section 10***^[1]

- 5.4.1 Analytical R&D must be carried out by staff having appropriate experience, knowledge and competence, consistent with the particular role they have in the work. Suitable qualifications may be academic, professional or technical, preferably with a specialisation in analytical chemistry and may also feature on-the-job training. For R&D leaders, a high level of qualifications and relevant experience is necessary. Published guidance is available ^[13]. The balance between academic qualifications and experience required for particular types of analytical work may vary from country to country.

Staff should receive relevant on-the-job training. The training programme should be assessed regularly and adjusted as necessary to ensure it continues to be relevant to the type of work carried out.

- 5.4.2 It is the responsibility of management to establish appropriate levels of supervision for each task, depending on the difficulty of the work and the capability of the analyst. It is recognised that analysts may be given unfamiliar tasks as part of their training; in such cases, management should take extra care to ensure that the level of supervision is appropriate.

- 5.4.3 Analysts involved with R&D will need to have or develop particular skills. For example they will have to exercise high levels of judgement about how to approach the analysis, about the selection of best methods, and about interpretation of results. They will occasionally encounter problems which are beyond their own experience and possibly also that of the laboratory, and so should have experience of literature searching and other information gathering techniques. They should maintain and develop their expertise by reading scientific literature, attending seminars and courses, participate in professional activities and be aware of colleagues who are experts in the various analytical subjects who might be able to give advice. They should also maintain an up-to-date awareness of quality assurance. Management is responsible for ensuring staff have the resources to maintain these professional skills.

- 5.4.4 Staff records are an important aspect of establishing the suitability of staff to undertake the analytical work. As a minimum, they should include:

- education leading to formal qualification e.g.: academic, professional, technical / vocational*;
- methodological / technical expertise;
- external and/or internal training courses attended;

-
- relevant on the job training;
 - previous R&D experience, in terms of subject areas covered;
 - list of scientific papers published, posters presented or lectures given.

** Vocational training is practical training related to a particular job, accompanied by study of the relevant theoretical knowledge. Part of the training may be provided within the laboratory, but the competence may be assessed independently and recognised via a formal qualification^[14-16].*

5.5 *Equipment - see CITAC CG1, section 12. For computer controlled equipment - see CITAC CG1 section 17 and App. C^[1] and GLP guidance^[17] .*

- 5.5.1 Equipment should be purchased against technical specifications derived from anticipated use and required performance capability. Where an instrument is sold on such a basis, there is an obligation on the agent or manufacturer to demonstrate to the purchaser, if required, that the instrument can meet that specification. Newly acquired items of equipment should be formally commissioned before being put into routine laboratory use, so that correct functioning and compliance with the appropriate specifications can be verified ^[18].
- 5.5.2 A list of equipment should be kept, indicating the equipment name, identification, records of commissioning, and related operating procedures, where appropriate. Records of calibration and maintenance should be kept.
- 5.5.3 It is not uncommon in R&D for a piece of equipment to be used by different persons, for a number of applications, perhaps in different projects, within a brief timescale. Where this is the case, special precautions for instrument cleaning and maintenance are advised, together with records detailing what the equipment has been used for, when, and by whom. This may help reduce unexpected observations which might have been caused by cross-contamination.
- 5.5.4 R&D may actually involve the modification of existing equipment or design of new equipment. Accepted engineering and scientific practices should be applied to design and construction. Method validation procedures and use of blanks, standards, old samples reference material can be used as part of the commissioning process.

5.6 *Monitoring quality - see CITAC CG1 section 18^[1] .*

- 5.6.1 Regular and systematic monitoring of quality is necessary to ensure that it is appropriate to the laboratory's needs and all aspects of it are functioning properly. Monitoring may be carried out by external bodies (different types of external assessment are described in more detail in section 8) or internally, using laboratory staff. Where there is a formal quality system internal assessment is conducted to formal procedures and known variously as audit or review ^[19-22] .
- 5.6.2 One approach to internal assessment is for a laboratory to train some of its own staff to act as internal auditors. The laboratory will benefit by involving its staff in monitoring the quality system. Assessors can be staff at any level in the organisation and should be independent of the work they are assessing, but have sufficient technical expertise and experience to be able to examine it critically.

- 5.6.3 All areas of the laboratory whose operations affect quality should be assessed in a systematic manner, typically at least once a year. Assessments should examine adequacy of procedures and ensure that these procedures are being followed, that suitable records are kept and appropriate actions are taken. Ideally a preplanned timetable should be followed, and over an agreed period should cover the whole quality system. It is unnecessary to examine the entire output of the laboratory - the assessment should be done on a 'sampling' basis. In the case of research it will be appropriate to select and examine entire projects or studies.
- 5.6.4 Even if a research laboratory's quality system is not fully documented to the requirements specified in quality standards, provided some form of work-plan is available an appropriate assessment can be made against this. For example, some of the questions which could be asked in assessment of a workplan could include:
- is the analytical task clearly described and understood?
 - is there an analytical working plan or study plan, and is there evidence of adequate experimental design?
 - are the task leader and other technical staff sufficiently competent?
 - are the applied procedures and equipment fit for purpose?
 - are calibration levels adequate and traceability suitable?
 - what measures are taken to confirm the reliability of results and are the results plausible (e.g. duplicate analysis, use of RM/CRM, spiked samples, cross-checking by other personnel, other internal and external quality control)?
 - has the work been completed and does the test report contain sufficient information (analytical results, interpretation, reference to customer requirements)?
 - is the level of record keeping sufficient for its purpose?
 - are scheduled milestones and deliverables being met?
 - are any relevant regulatory requirements being met?
- 5.6.5 Where changes to procedures are required staff should be identified to carry out them out over an agreed timescale. Subsequent completion of the changes should be confirmed.
- 5.6.6 In R&D it is not unusual to make ad-hoc deviations from procedures. These may adversely influence software or hardware performance, data collection, calculations, and interpretation of results. A simple system recording deviations as they occur and confirming that consequences have been evaluated and where appropriate corrective action has been taken should ensure that there is no inadvertant loss of quality arising from the deviations.

5.7 Subcontracting

- 5.7.1 The laboratory should consult with the customer before placing any part of a contract with subcontractors.
- 5.7.2 Where one laboratory (A) subcontracts work to a second laboratory (B), B should operate to at least equivalent levels of quality as A. A should put in place whatever procedures are appropriate to assure itself of the quality of the capabilities of B and the quality of the work it is producing. This might include:
- assessing the quality of subcontractors;
 - establishing a list of laboratories approved to act as subcontractors;
 - reviewing data and reports of subcontractors for scientific content;
 - limiting the scope for the subcontractor to work independently on the subcontract;
 - checking the subcontractor's work against the initial specification, and defining corrective action if necessary.

Note that the subcontractor and the laboratory placing the subcontract could be two different laboratories within the same organisation, i.e. the arrangement could be purely internal.

6. TECHNICAL QUALITY ELEMENTS**6.1 Unit operations**

- 6.1.1 R&D projects can be considered as a collection of discrete tasks or workpackages, each consisting of a number of unit processes, themselves composed of modules containing routine unit operations. The unit processes are characterised as being separated by natural dividing lines at which work can be interrupted and the test portion or extract can be stored without detriment before the next step. This is illustrated in Figure 2.
- 6.1.2 The benefit of this modular approach to defining R&D projects is that new R&D work is likely to contain at least some components which are familiar to the laboratory and may even be performed routinely. This approach offers benefits in terms of establishing staff competence and also in documentation of procedures.

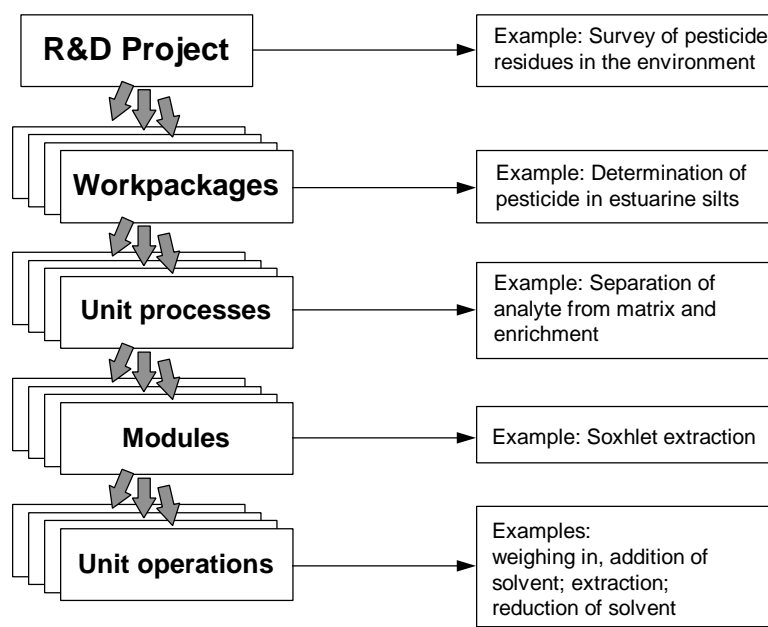


Figure 2 - Illustration of the breakdown of R&D projects into unit operations

6.2 **Technical capability of the laboratory**

6.2.1 It is common practice to allow the general acceptance of laboratory performances by a *type of test* approach. This means, if the laboratory has demonstrated its ability to perform a particular method, it is also accepted as fit to perform similar closely related methods. This logical, but knowledge- and experience-oriented approach, enables the demonstration of valid analytical measurements to external experts without the need for elaborate validation of every single unit operation or module or process.

6.3 **Methodology**

6.3.1 It is likely that procedures for carrying out *unit operations* and perhaps even *modules* (see Figure 2) will be sufficiently routine and/or common to other work to warrant full documentation as a written standard operation procedure (SOP). Using this principle, any new test procedure can be described by the appropriate combination of the SOPs of the relevant unit processes or modules, keeping new documentation to a minimum. Representation of new test methods by recombination of existing SOPs has a number of advantages in terms of using existing validation information and uncertainty contribution estimations. Validation of the whole workpackage or task will often be necessary but can be achieved using reference materials, *etc.*. In practice SOPs might even cover individual workpackages but care should be exercised in case this reduces the flexibility of operations.

- 6.3.2 SOPs provide a source of information to which analysts, carrying out a particular operation, can refer in order to ensure a consistent approach. A closely followed, well written SOP can improve the consistency of data produced for a particular process, between analysts, between laboratories, and over time intervals. Thus an SOP should contain whatever level of information is necessary to avoid ambiguity. A well written SOP also helps auditors to follow the course of the work done and so assess the validity of the data. In an R&D environment it is expected that as the science improves, so SOPs can be reviewed and changed to reflect the improvements (e.g. in speed, in material and money savings, in waste production, *etc.*) as long as the results are convincingly demonstrated to be comparable or better than those obtained with existing versions. Changes must be authorised, prior to use, in line with document control policy.
- 6.3.3 Where SOPs do not already exist or are inappropriate, contemporaneous notes should be made to describe the procedures used in the work. Sufficient detail should be recorded so that at some later time, the procedures used can be reconstructed, if necessary. Where a number of procedures were attempted before one was found that was satisfactory, records should be kept of the failures so that they can be avoided in future.
- 6.4 ***Reagents, reference materials, and calibrants - see CITAC CG1, sections 13 & 16***^[1]
- 6.4.1 Special attention should be given to chemical and physical properties of reagents, reference materials and calibrants (chemical and physical measurement standards). Careless preparation or poor storage may result in inadvertent degradation. This is particularly important where chemical metabolites, or chemicals about which little is known, are involved. Sometimes, the use of added preservatives or storage under inert atmospheres (e.g. Ar or N₂) may be appropriate.
- 6.4.2 Reagents, calibrants and reference materials prepared for specific R&D applications should be appropriately labelled and if appropriate, their use restricted, to prevent contamination through widespread use. Details of preparation *etc.* should be recorded in SOPs.
- 6.5 ***Calibration & traceability - see CITAC CG1, section 15***^[1]
- 6.5.1 Calibration establishes, for specified conditions, how the response of the measurement system relates to the parameter being measured. Calibration is usually performed using a reference material of established composition, or calibrant in which the property of interest (for example the chemical purity) is well characterised.
- 6.5.2 In R&D, one is more likely to encounter the situation where calibrants are absent or, if available, are poorly characterised. Where the stoichiometry of the calibrant is not known an approximate amount should be weighed and the exact amount of calibrant constituent determined with an
-

absolute method (coulometry, volumetry, gravimetry). Where no suitable calibrant is available the method for determining the response for the property analyte should be demonstrated.

- 6.5.3 Validation of the unit processes together with appropriate traceability is important to ensure that data produced is comparable with data for similar measurements made at different times, or by different analysts or laboratories, or using different methods and different samples. Traceability can be achieved by calibration using various calibrants, reference materials or even standardised procedures. Caution is advised when using standardised procedures as frequently they contain bias which may be poorly controlled.
- 6.5.4 Traceability to (the) SI is often possible in chemical analysis at some level of uncertainty. Traceability can be to a standard / calibrant, whether national or international, which has been accepted as the point of reference by the analytical community concerned and which all interested parties have access to, either directly, or indirectly, through a chain of subsidiary standards. Similarly traceability can also be established to a reference method.
- 6.5.5 Traceability is not to be confused with the traceability from the sample via the test procedure to the final test result. This has been tentatively termed "trackability" (from tracking back).

6.6 ***Instrument performance***

- 6.6.1 For instrumentation, design, installation, operational, and performance qualifications are of equal importance in R&D as they are in routine work. Design and operational qualifications are briefly dealt with in §5.5.1. This section deals with operational and performance qualifications - Does the instrument/system work in the specific application and what could be the interferences? Does the instrument continue to work in the manner intended (continuing fitness for purpose)?
- 6.6.2 In R&D it is not sufficient to adapt existing work without demonstrating that the instrumentation works properly with the new application. Care is also needed with novel or modified instrumentation; where the performance claims of the manufacturer may no longer be true because of the modification.
- 6.6.3 The ultimate performance test for any calibrated analytical instrument is to analyse a certified reference material (CRM) and obtain a result within the uncertainty range stated for the CRM. If the matrix of the CRM is similar to that for the samples, and the CRM is subjected to the whole analytical process then this serves to validate the entire procedure ^[23-25].
- 6.6.4 Often in R&D, no CRM is available and it is not possible to relate a property to an existing national or international standard or calibrant. Instead, in-house reference materials can be
-

used. It is advisable to specify one or two materials with characterised property values appropriate to the scope of the procedure which can be used for instrument performance checks, calibration or quality control. Specific mixtures of analytes can be contrived to test certain performance parameters, for example the resolution of two compounds in a separation process.

6.6.5 In critical instances the use of a different analytical procedure and/or technique, susceptible to different interferences, is advised to check results. This check is more valuable than, for example, interlaboratory comparisons involving only a limited number of laboratories using exactly the same overall procedure and measurement technique. However, interlaboratory comparisons involving larger numbers of laboratories and different techniques are more useful still.

6.6.6. Where R&D involves testing a large number of similar samples using a particular procedure, control samples and charts can be used to monitor the continuing stability of instrument performance.

6.7 ***Use of statistics***

6.7.1 Statistical techniques are an invaluable tool in the design or use of analytical methods. During the lifetime of an R&D method statistics can be used in four basic areas:

- I. experimental design of the method;
- II. characterisation of method performance, ruggedness and determination of uncertainty;
- III. quality control of the method (once the method is in use);
- IV. interpretation of populations of results.

6.7.2 In each of these areas a variety of statistical techniques may be applied or indeed are necessary, depending on the different parameters to be studied, and such chemometric approaches can also reduce time and costs. A detailed study of this area is beyond the scope of this guide; references to a number of suitable texts are provided in §9.

6.7.3.1 **Experimental design.** In any analytical procedure performance can be influenced by a number of different variables, such as: matrix interferences in the samples; reagent concentrations; temperature; derivatisation time; *etc.*. Experimental design is usually used to describe the stages of identifying the different factors that affect the result of an experiment, designing the experiment so that the effect of these factors is minimised, and using statistical analysis to separate the effects of the factors involved. For example, a ruggedness test will indicate firstly whether a particular method will stand up to everyday use, and will indicate which parts of the method are vulnerable to change and need to be subject to quality control. As part of the design

process regression or multiple regression analysis may be used, together with ANOVA (ANalysis Of VAriance) determinations and MANOVA (Multiple ANalysis Of VAriance)^[26, 27].

- 6.7.3.2 Statistical methods are very important in the design of sampling schemes. If used properly they can enable the desired results to be obtained with the minimum of samples and subsequent analysis. Internationally available standards have been published for the use of statistics in certain types of sampling ^[28]. However a broad knowledge of the history of the sample substantially helps to design a more intelligent sampling plan and reduces sampling time and costs.
- 6.7.3.3 SIMPLEX optimisation can be used for rapid method development where a number of factors affect method performance and to investigate all possible combinations would involve vast amounts of work ^[29]. Other specialised techniques which may be used in a similar way include: full factorial designs; fractions of factorial designs; Taguchi designs.
- 6.7.3.4 Where a large number of samples need to be processed and only a few are expected to yield “positive” results, screening techniques may be used for eliminating the large numbers of negative samples to leave the positive samples which can then be examined in more detail.
- 6.7.4 **Characterisation of method performance and determination of uncertainty.** This involves the evaluation of various parameters associated with the performance of the method, such as precision, trueness, *etc.*, followed by a judgement as to whether these performance capabilities are sufficient to meet the needs of the method. The process is generally referred to as method validation (see §6.8.5). Determination of measurement uncertainty use similar measures to those determined during method validation and involves identification, determination and final recombination of all the sources of uncertainty arising at all stages of the analytical procedure to give an overall measure (see §6.8.6). Both method validation and measurement uncertainty make use of simple statistical measures such as means, standard deviation, variance, *etc.*.
- 6.7.5 **Development of quality control.** The quality control procedures developed for a new method should concentrate on those parameters which have been identified as critically influencing the method. However for R&D work there may be problems in finding suitable samples for quality control purposes, and control charting techniques are less relevant in non-routine situations. Control charts can still be applied, for example to monitor instrument calibration, and the main thrust of quality control in the R&D situation is probably best directed towards ensuring instrumentation is working properly and calibrated, monitoring values from reference materials where available, and replicate analysis (consecutive and random, to monitor short and long term variation respectively).
-

6.7.6 **Interpretation of results.** The problems associated with validation of methods in R&D and the subsequent design of adequate quality control should be borne in mind when interpreting sets of data produced in R&D. Techniques used for the detection of outliers and measures of distribution of result populations, such as standard deviation, are particularly relevant in this case.

6.8 *Technical requirements related to particular unit processes:*

6.8.1 In most analytical R&D situations the following unit processes (which may or may not have subsidiary modules and unit operations) may be encountered: sampling; sample preparation; separation of the analyte from the matrix and enrichment; measurement; calculation and; presentation and interpretation of the result. Guidance is generally limited to information specific or more relevant to R&D.

6.8.2 *Sampling, - see also CITAC CG1, section 19*^[1]

6.8.2.1 Extensive guidance on sampling exists in the scientific literature^[28]. There is actually little advice on sampling in R&D that is not also applicable to routine measurements.

6.8.2.2 Where R&D involves the development of new test procedures for subsequent use on real samples, method development needs to consider practical sample sizes which will typically be available for testing. During the development stages it may be useful to have large quantities of real sample available for method validation, etc..

6.8.2.3 R&D may involve taking types of samples which have never been encountered before, with unknown or unfamiliar analyte contents or matrix types. The samples may present unknown hazards or problems with stability, handling, and storage. The sampling strategy should try to anticipate potential problems and if possible make suitable allowances. Customers' declarations of the expected contents of samples should be treated with caution. Sampling plans should be detailed even if some of the information recorded is subsequently not needed. The analytical staff involved with the R&D should use their scientific expertise to help ensure the sampling procedure is as appropriate as possible. Where appropriate, procedures should be recorded.

6.8.2.4 Similarly, for unfamiliar samples, storage conditions should err on the side of caution. In critical cases it is strongly advised that samples are retained after analysis at least until the validity of the tests results have been confirmed by suitable review.

6.8.2.5 With samples taken for R&D purposes little may be known about their homogeneity. It is particularly important to investigate this before any subsampling is carried out to reduce the effective bulk of the sample. Any means used to homogenise the sample must not compromise its integrity. It may be appropriate to separate phases in inhomogeneous samples and treat the separate phases as different samples. Conversely it may be appropriate to homogenise the samples. The uncertainty of subsampling which is determined by the level of homogeneity may be estimated by setting up a specific study and taking more subsamples and determining the uncertainty statistically.

6.8.2.6 It may be convenient to have a single SOP describing the variety of sample treatment methods (solvation; dissolution; digestion; extraction; surface cleaning; melting; combustion; *etc.*) used by the laboratory, and containing detail on the special precautions to be taken for the different analyte groups. It should also describe how the methods are applied to blanks (spiked and unspiked), reference materials and other calibrants, and other materials used for quality control purposes.

6.8.3 *Isolation of the analyte(s) using separation and enrichment*

6.8.3.1 Diverse techniques are available for separation and enrichment. The experience of the analyst will be an important factor in choosing the most appropriate for a particular application. For future reference, records should indicate the logic behind a particular choice.

6.8.4 *Measurements*

6.8.4.1 The measurement process consists of using a calibrated instrument to determine the net instrument signals of the test portions and various different blanks. Within run and between run changes in instrument response can be monitored using quality control samples and calibration standards.

6.8.4.2 Depending on the circumstances, this determination step may be repeated several times to allow a statistical data treatment of this single step. The determination of more than one test portion from the same sample can be used to determine (at least an estimate of) the overall repeatability of the analytical method. Where there is a suspicion that interferences are present, results obtained from test-portions using external standard calibration (using a calibration curve) can be checked by spiking test portions with known amounts of the analyte of interest.

6.8.4.3 Blank corrections for measurements should be made by calculating actual concentrations of sample and blank as indicated by the respective instrument signals and then subtracting one from the other. The practice of subtracting the blank signal from the sample signal and then calculating the result using the net signal is not recommended.

6.8.5 *Validation - see also CITAC CG1, section 22* ^[1]

6.8.5.1 There is a clear responsibility on the part of the test laboratory and its staff to justify the trust of the customer or data user by providing reliable data which can be used to solve the analytical problem. An implication of this is that methods developed in-house must be adequately validated, documented and authorised before use. Validation is normally quite straightforward for routine work but can be expensive and time consuming. For methods used or developed during the course of R&D, validation is equally important, but less straightforward. General guidance has been produced by EURACHEM ^[31].

- 6.8.5.2 Various options exist for characterisation of method performance. The trueness of a new method could be assessed against that of an established method, repeatability could be assessed using reference materials, and reproducibility through interlaboratory comparisons. In R&D, many of these options may not be available. Validation tools may be limited to the use of in-house reference materials, and uncertainty estimations based on error propagation principles relying on a solid understanding of the theoretical principles of the method and the practical experience of the research workers.
- 6.8.5.3 A suitable unit process for data treatment should include validation of the overall procedure. That means evaluation of various performance parameters of the method, and consideration of their adequacy relative to the analytical requirement. Parameters such as: limit of detection, limit of quantification, dynamic measuring range, sensitivity, repeatability (same analyst, same instrument, same laboratory, same day), reproducibility (different analyst, different instrument, different laboratory, different day), accuracy (difference from the true value) and other terms (e.g. robustness or ruggedness); will need to be considered.
- 6.8.5.4 The extent to which validation is needed, and the effort given to this task, depends on the use which will be made of the method or technique. At one limit, where new methods or techniques (or ones seldom applied) are being used, a customer requirement for durable methodology will justify extensive work on validation. In many situations, however, less than full validation is necessary or possible. Here the analysts' professional judgement will be introduced to decide those unit operations of the analysis which need to be investigated, and those whose performances can be estimated from comparable systems. The extent of validation, and the consequences in time and cost, are one of the key issues to be agreed between analyst and customer when commissioning method development.
- 6.8.5.5 It is generally assumed that R&D requires an increasing effort for validation since seldom applied or totally new techniques or methods are being used. The unit operation approach described above enables the possibility of recombination of the units into a large variety of testing methods. If these units can be individually validated it may be possible to estimate the overall performance capability of subsequent combinations of the modules which then require the minimum of further validation for verification. It is not necessary to define all unit operations for each possible analyte, but it might be sufficient for a group of analytes with a nearly similar matrix.
- 6.8.5.6 Ideally, individual recovery studies should be performed for each analyte. This can be done using a synthetic matrix similar to the sample matrix or by analyte addition (spiking) to subsample aliquots and determination of the increase of the measured concentration. Often the recovery factor depends strongly on the sample matrix. Guidance on acceptable recovery
-

ranges for similar analyte/matrix combinations may be available in the literature. Whether results should be corrected for non quantitative recoveries is the cause of much debate ^[32] and the client may have a preference. Reports should indicate clearly whether or not data has been changed to allow for non-quantitative recoveries.

6.8.5.7 Ideally the procedure should try to identify such a matrix effect so that any blank correction procedures can be performed properly. In analytical R&D the search for systematic errors is of greater importance since per se less is known in those fields. Wherever possible these systematic errors should be identified and if possible, eliminated.

6.8.5.8 It should be noted that methods can be validated at different levels. Analysis of CRM's with similar matrices to the test materials gives the highest confidence level for in-house validation. If the obtained results lies within the stated confidence range then the total analytical process is under control and all involved unit processes are automatically included in this validation. This means there is no need for any further method or instrument validation and no need for other more formal demands. Other mechanisms for validation are described below in order of decreasing confidence:

- taking part in inter-laboratory comparison tests;
- performing a limited number of control-analyses of the sample at a different test laboratory;
- employing several methods with different interferences possibility and obtaining only one and the same result;
- reanalysis of an in-house sample of known content.

6.8.6 ***Measurement uncertainty - see also CITAC CG1, section 21*** ^[1]

6.8.6.1 Uncertainty should be estimated and quoted in a way that is widely accepted, internally consistent and easy to interpret. More detailed guidance has been published by EURACHEM ^[32]. Where appropriate, uncertainty should be quoted with the analytical result, so that the user can be assured of the degree of confidence that can be placed on the result.

6.8.6.2 The most significant contributions to the overall uncertainty of a measurement are usually due to the sampling processes and the accuracy of the determination of recovery factors. Contributions due to instrument performance are generally less significant.

7. ANALYTICAL TASK QUALITY ELEMENTS

7.1 Preparation and planning before starting work:**7.1.1 Definition of task and project design**

7.1.1.1 Planning and preparation is a critical part of analytical R&D, especially where new analytical methods are generated or extensive validation of generic methods is required. The effort put into planning depends on the complexity and requirements of the work, previous experience, the extent to which the work is unfamiliar or novel in its character, the number of persons or organizations involved, expenditure for new equipment, consequences of wrong results, the duration of the work, deadlines *etc.*. A flowchart such as the one shown in annex B may assist planning. As a rule of thumb, proportionally more planning is needed for high risk work. When costing project work it is important to correctly estimate the resources needed in the planning or subsequent management stages. The structure of the project should be flexible enough to allow creative problem solving. The project management team is responsible for planning activities within the project and allocating resources to cover these activities. The sort of activities involved include:

- scoping;
- milestone planning;
- objective/goal setting;
- resource allocation and costing;
- contract control;
- financial control;
- change management;
- liaison with customers.

7.1.1.2 Task definition is the first stage of planning and should provide sufficient information to allow more detailed planning or indicate viability of proceeding. Go/no-go decision criteria should be incorporated in the project structure at the earliest opportunity. It is vital to establish a good link with the client to ensure work is defined adequately and thus maximise the chances of a productive outcome to the project. The sort of areas covered in task definition may include:

- nature of the problem that the work is intended to address, seeking clarifying from the client as necessary;
- objective, goals and expected information, purpose of results/data, intended use of information;
- type of material/product/matrix to be analysed/amount available/safety considerations;
- sampling procedures/sampling plans, statistical methods;
- element/species/determinand/property to be analysed/determined;
- methodology, generic methods to be used, destructive/non-destructive methods;
- required accuracy (or precision, bias, *etc.* as appropriate) and related equipment performance requirements;
- validation procedures and use of reference materials, standards, reference methods;
- required date of completion;

-
- available resources (personnel, equipment);
 - expected use of subcontracting;
 - success/failure criteria where appropriate;
 - expected/permissible costs and expenditures;
 - reference to exploratory work and review of literature required for definition and execution of the task;
 - degree of confidentiality necessary;
 - requirements and arrangements for archiving;
 - ownership of intellectual property;
 - possible strategy for dissemination and exploitation.

7.1.1.3 A questionnaire can be used to help define work. The example shown in annex C is adapted from one used for routine work. Note it is not exhaustive but illustrates some of the issues which should be addressed.

7.1.1.4 Where limited amounts of sample are available it is particularly critical to have a clear strategy in place before beginning work. Use of non destructive methods should be considered.

7.1.2 ***Project design and research plan***

7.1.2.1 Once task definition is complete the research plan(s) can be drawn up. The laboratory management should involve the client, and the laboratory staff from the very beginning in order to ensure that the finalised project as far as possible meets the client's requirements, is technically possible and suitable resources are available within the specified timescale. The project should be structured by a logical sequence of tasks or workpackages, points of decision where the work can change direction if necessary, and points of achievement. (milestones, target dates) which enable progress to be monitored. All contractual or technical issues should be resolved before the analytical work is begun. Particularly where operations may be complex, use of flowchart, such as that shown in annex B, a decision tree or other diagrams, may help to clarify the procedure.

7.1.2.2 The **research plan** defines:

- **Goals:** Set clear final (and if appropriate, intermediate) goals (measurable objectives including go/no-go decision points/acceptance criteria. Establish what questions need to be answered at each stage and the corresponding results/data required to answer them.
 - **Tactics:** Outline the strategy to be used at each stage. If necessary subdivide tasks into manageable, defined workpackages (unit operations) with discrete goals.
-

-
- Resources: Define the resources (personnel, equipment, facilities, consumables) needed at each stage.
 - Time schedule: Define start and end of project, dead lines for intermediate goals, and minimum critical path for completing work.

7.1.2.3 Research plans should contain as much detail as is necessary to define the tasks involved. For isolated tasks the plan may simply be an entry in a notebook or a form. A more detailed plan will be necessary for larger, more complex tasks or when time and cost constraints are to be closely controlled, or when high risk or significant investments depend on the outcome of the work. If there is significant doubt as to whether the work can be completed successfully by a single route, then alternative plans should be defined.

7.1.2.4 A workpackage typically consists of a discrete piece of work with: defined starting and finishing times/dates; necessary starting conditions (particularly if the workpackage is one in a sequence); a goal (achievement of which indicates successful completion of the workpackage); a budget indicating financial, time and other resource restrictions; a note of any particular resource requirements; a statement of the roles and responsibilities of the various staff involved with delivery at all levels from management to technician; a specification for reporting progress and the final goal.

7.1.2.5 Milestones are points of appraisal (usually) at the end of a workpackage. Their timing is normally fixed within the overall project timetable. They are points at which decisions can be made either to proceed with the project, to stop, or to select a particular path in the workplan for further action. Where appropriate the client should be involved in any important decisions.

7.1.2.6 A number of tools are available to assist project design and control ^[33]. They include:

- bar charts (Gantt chart);
- PERT chart (program evaluation and review technique);
- CPM (critical path method).

7.1.3 **Resource management of task**

7.1.3.1 Large or multitask projects may involve scientists from several departments of the laboratory and perhaps outside specialist subcontractors. The role of project management is particularly important in order to ensure the project team functions smoothly, with all members co-operating and aware of their roles and responsibilities. Particular attention should be given to:

- definition of the project management hierarchy, with leaders in particular areas, and defined authority and responsibility for all team members;

-
- involvement of all personnel pertinent to the project (including the client) in defining the task and assignments, and in planning the project;
 - setting clear tasks and goals which are challenging but achievable;
 - early consultation with the management of specialists in other departments or organisations, involved in the project. Unresolved questions concerning priorities and workload, and budget contributions often disrupt good team work;
 - communication. Hold meetings at appropriate intervals for exchange of information, problem solving, consultation, reporting, coordination and decision making.

For small, simple projects the same principles can and should be applied in a cut-down form.

7.1.3.2 Resource management at the planning stage may include:

- evaluation of the skills and facilities required for the project, comparing those against what is available, and plans to cover any shortfall. This includes special considerations such as environmental controls, special equipment and reagents, protective clothing, decontamination procedures;
- costing the planned deployment of personnel and facilities and set budgets for the various parts of the work (time and finance budget);
- establishment of a timetable for the work consistent with client requirements and the availability of personnel and facilities at each stage;
- availability and allocation of resources to defined tasks and/or appointed dates/ decision points (e.g. milestones) and including resource distribution in the project plans;
- definition of a system for monitoring time and resource expenditure in the project;
- identification of potential problems with disposal of samples, reagents and contaminated equipment, arising as a result of the work.

7.2 *While the work is in progress:*

7.2.1 *Progress review/monitoring analysis*

7.2.1.1 Progress of work and status of expenditure should be controlled by comparing achievements and use of resources against the planned budgets at convenient points within the work, typically at regular intervals or completion of milestones. Informal reviewing should be carried out individually by the laboratory staff as work progresses. Unexpected difficulties or results, or major deviations from goals may call for extraordinary reviews and interim reports with replanning of the work and reallocation of resources as necessary.

7.2.1.2 Progress should be reported to laboratory management or the client, in the format and at the time intervals agreed at the planning stage. Typically reports might cover: a review of the project plans; information on whether the work is running to schedule and will achieve its objectives - on-time/late/at all, an account of technical progress with achievements and failures/setbacks; and information on resources.

7.2.1.3 Effective project management requires records of laboratory data, observations, and reported progress against milestones or goals to be clear and comprehensive so that decisions made during the project and the underlying reasons are easily understood and laboratory work and results can be repeated if required. Records should include laboratory note books, computer print-outs, instrument charts indicating all activities, working conditions and instrument setting, observations during experimental work, as well as justification for tactics and/or changing plans.

7.2.1.4 Ultimately, the level of data recorded should comply with customer requirements, or those laid down for scientific papers, published standard methods, or other requirements such as patents or licences. It should be sufficient to enable other scientists to repeat the experiments and obtain data compatible with the original work. Thus:

- all experimental details, observations, and data necessary for possible replication of the work must be recorded;
- records should be made 'at the time' and kept as up-to-date as possible;
- records should be traceable to particular samples, tasks or projects, people, time;
- details of unsuccessful work should be recorded - In R&D it is worthwhile reporting failures as well as successes.

7.2.2 **Data verification**

7.2.2.1 Data verification should show that a new or adapted method gives consistent results with a particular sample. If results are not consistent with established data, the analytical procedure may need to be improved until the required consistency is achieved. Management should be aware that data and method validation costs form a significant part of the total costs of R&D.

7.2.2.2 The unit operations, as listed in §6.8.1, may influence one another, but contribute individually to variations in results. A step-by-step verification may often be impractical although it may be feasible and useful to study particular performance characteristics of particular stages of the sequence of operations. In R&D plausibility of data may be checked either using literature data, theoretical considerations, or using specially prepared reference materials and model substances.

7.2.3 **Changing direction**

7.2.3.1 Where a review of progress shows that a particular line of investigation is likely to be unsuccessful, goals or/and chosen tactics and tasks may have to be changed. Such a change may already have been anticipated during planning. Changes should be made in consultation with the client where appropriate and justified in reports.

7.3 **When the work is complete:**

7.3.1 **Achievement review**

7.3.1.1 The completed work should be reviewed by management to evaluate achievements. Experiences gained at all stages of the project may provide lessons for planning and carrying out similar work in the future. The review might typically cover:

-
- aspects of technical achievement such as differences between goals and results, problems encountered and how they were solved, usefulness of the results;
 - compliance with budgeted costs and timescales, with explanations for any deviations, correlation of expenditures and technical results;
 - quality of work of individual contributors;
 - consequences of project and results to the laboratory (organisation, personnel, equipment, methods and procedures, possibility of dissemination or exploitation);
 - satisfaction of client.

7.3.1.2 The achievement review may be supplemented by an external peer review, e.g. when data is published in scientific journals, or third party review (audit).

7.3.2 ***Reporting, technology transfer and publication:***

7.3.2.1 R&D may be reported in various ways. Primarily a report should be made to the client in the format previously agreed and be written in a language that the client can readily understand. The report should provide sufficient information to enable the client, any subsequent user, or assessor of the report to be able to follow any arguments, and if required, repeat any or all stages of the experimental work and obtain compatible results. In particular:

- the meaning of the test results should not be distorted by the reporting process;
- appropriate use should be made of conventions for rounding of numbers and expression of decimal places and significant figures;
- where appropriate, results should include an estimate of the associated uncertainty with its corresponding confidence level.

7.3.2.2 Compared to scientific publications, project reports typically contain project oriented information (technical, financial statements *etc.*), conclusions and recommendations, and usually present the findings in a less technical way.

7.3.2.3 If the work has yielded data, observations, new methods, techniques or new knowledge, of interest to the wider community, then dissemination or exploitation of the work is an important issue. Dissemination or exploitation can take a number of forms: lectures, publications in journals; patents; licences; standards; training material. Permission for dissemination or exploitation must be sought from the laboratory, the client or whoever else owns the intellectual property. Where it is hoped that new methods can be adopted more widely, further performance evaluation may be required, perhaps using collaborative study. Methodology must be described unambiguously, and in sufficient detail to allow others to be able to follow the arguments and replicate the work, otherwise its credibility may be adversely affected.

7.3.3 *Archiving*

7.3.3.1 Archiving primarily involves the secure storage of samples, analytical records, results, methods and other information for later retrieval and use. The method of archiving and the time for which material is kept depends on what is archived and why. It may be done for a number of reasons:

- legal or regulatory requirement;
- requirement of customer or some other external agency (e.g. accreditation body);
- verification of previous work and procedure at later stages of the project;
- validation of methods and results after completion of laboratory work and reporting/publication;
- proficiency testing or collaborative studies with samples;
- post-report questioning by client or peer review;
- problems associated with duplication of work/results; technology transfer;
- keeping the information benefits the laboratory.

7.3.3.2 Samples should normally be stored until the likelihood of their requiring retest has been ruled out or they have deteriorated to an extent where retest would be meaningless (unless study of their deterioration is part of the work).

7.3.3.3 An important feature of an effective archive system is knowing what it contains and being able to find things quickly. Use of a searchable data-base is recommended and offers some protection against illness, death, or transfer of expert staff and also helps to save time and money by providing a means of preventing the inadvertant duplication of earlier work.

7.3.3.4 Where space is important text based material can usually be archived in electronic or photographic form. Back-up copies should be kept in remote, flameproof storage. The use of different media may be preferred in different sectors, and use of others prohibited.

7.3.3.5 Retention of data, reports and other useful information should be consistent with regulatory and customer requirements.

8. **EXTERNAL VERIFICATION**

8.1 Whilst the laboratory may monitor the quality of its work by internal assessment, independent external assessment may be useful, in order to:

-
- demonstrate its quality to customers, regulatory bodies, funding bodies, or other external parties;
 - compare its level of quality with others in order to make improvements.

8.1.2 Whilst it is a straightforward process for a laboratory carrying out routine work to apply a structured quality assurance system and use it to regulate laboratory performance, the ever changing nature of work in an R&D laboratory demands a more flexible and less bureaucratic approach. It is a widely held opinion that the rigidity of conventional formal quality assurance systems and their associated means of external assessment restrict the creativity of thought and practice required in an R&D environment. A number of options are available for externally assessing R&D:

- formal assessment against conventional quality assurance standards (ISO Guide 25, ISO 9000, and Good Laboratory Practice);
- benchmarking;
- visiting groups and peer review of publications;
- ranking of laboratories;
- external quality assessment.

8.2 ***Formal Assessment against published quality assurance standards***

8.2.1 ***ISO Guide 25*** ^[3]

8.2.1.1 Traditionally the preferred route for routine laboratory environments, formal accreditation against standards derived from ISO Guide 25 provides an independent assessment against objective criteria that a laboratory is competent to perform specific calibration or testing measurements. The assessment is carried out by peers, that is specific measurement methods are assessed by colleagues from other organisations with expertise in those measurements, who can judge whether the procedures in use are technically valid. Accreditation is granted on the basis of the laboratory's ability to perform tests and does not cover peripheral issues, such as administrative procedures not related to the measurements, and perhaps more important, expert but subjective interpretation of the measurement data. Accreditation cannot guarantee the reliability of a measurement result. However it does provide recognition that the conditions under which the measurement was made maximises the probability of the measurement being verifiable. Even where there is no formal verification of compliance against ISO Guide 25, it remains a very useful technical quality assurance model for laboratories to refer to in order to regulate the quality of R&D.

8.2.1.2 Because accreditation is granted against a specified schedule of measurements, it is currently difficult and expensive to apply it to R&D. The 1998 revision of ISO Guide 25, now incorporates much of ISO 9001 ^[34]. However the definition of R&D used in ISO Guide 25 may not

necessarily correlate with its use in this document. In theory, R&D consisting of objective non-routine measurements, which could be fully documented and validated, could be accredited, provided the laboratory considered it to be cost-effective to do so.

8.2.1.3 It is sometimes possible for accreditation to be formally granted for groups of tests rather than specific tests, particularly where the laboratory in question has a proven quality system and has a high degree of established expertise in the technique relevant to the group of tests. It should be possible to extend this accreditation to whole types of test (see annex D). Whether or not accreditation could be granted for the unit operations described in §6 above is a matter for conjecture. Although a logical development of the principle of granting accreditation for test types, accreditation bodies currently only accredit the whole test. Some ideas of how accreditation of R&D might be achieved by type of test is given in annex D.

8.2.2 **ISO 9001**^[5]

8.2.2.1 ISO 9000 is unspecific about how technical work should be performed. The certification assessment is primarily aimed at the management of procedures and assessors are not normally from a relevant technical background. ISO 9000 requires no specific assessment of the validity of work and enables the laboratory to set its own level of quality. Certification thus has merits for assessment of how the overall work is managed but on its own does not assure its validity.

8.2.2.2 The main merit of applying ISO 9001 to an R&D environment lies in its use for controlling the organisation and project management aspects of work. There should be no reason why a laboratory cannot have certification to ISO 9001 to organise, manage and perform R&D work, using the more technically exacting requirements of ISO Guide 25 as a basis for the technical side of its work.

8.2.3 **Good Laboratory Practice (GLP)**^[6]

8.2.3.1 A laboratory operating to GLP (OECD Principles of Good Laboratory Practice) will have demonstrated that it has a management system and laboratory procedures which would enable a third party to reconstruct any GLP compliant study. GLP is concerned with traceability of the materials used, especially samples, and good descriptions of analytical methods. It is not, per se concerned with technical quality elements such as accuracy or precision, though many of the laboratory system elements required by GLP considerably assist in the delivery of technical quality. GLP traces its origins to testing in support of toxicological assessments carried out in support of product registration but in theory there is no reason why it cannot be applied to all areas of measurement. Eligibility of work for formal registration of compliance depends on the policy of the national bodies which administer GLP principles in each country.

8.3 **Benchmarking**

8.3.1 Benchmarking is a continuous, systematic process in which a laboratory/organisation compares its practices and procedures with comparable activities in other organisations in order to make improvements. It can be carried out at various levels with various partners (who need not be laboratories): internal; external; competitive; non-competitive; and *best-practice* (the acknowledged leaders of the process being benchmarked). When benchmarking with other organisations, an agreed Code of Conduct is vital to ensure an effective, efficient and ethical process, whilst protecting both parties. A typical benchmarking process is shown in Figure 3.

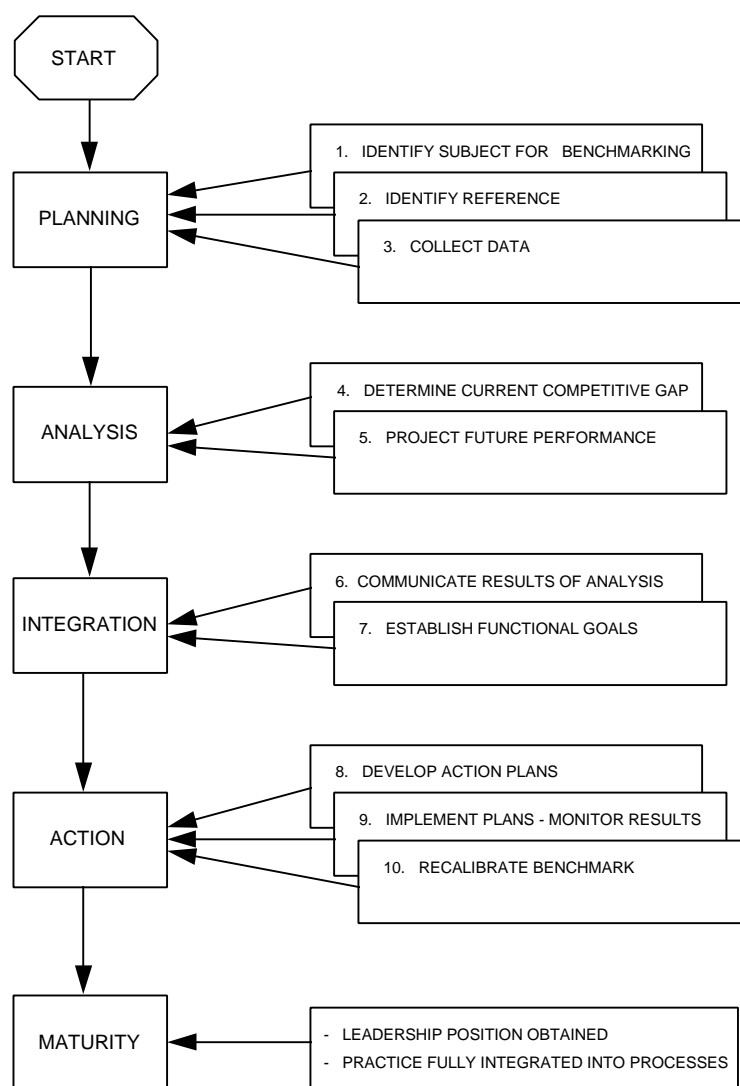


Figure 3: The Benchmarking Process

8.3.2 *Examples:*

-
1. External: A laboratory can assess its purchasing procedures by benchmarking with another organisation known to have very good purchasing procedures.
 2. Internal: Group A in a laboratory wins only 10% of possible contracts whilst group B in the same laboratory wins 50%. By benchmarking its bidding procedures against those of group B, group A ought to be able to improve its success rate at winning contracts.

8.4 Visiting groups and peer review.

8.4.1 These types of review involve the use of groups of senior level experts, probably from a wide range of sources, to evaluate a laboratory. The evaluation can be directed either at the laboratory itself or at the laboratory's scientific output.

8.4.2 In the former case the evaluation is likely to be against the laboratory's stated objectives, with a strong emphasis on the excellence of the science, staff, and facilities. Such groups typically act on behalf of R&D funding bodies and are a popular form of assessment in the academic world. The terms of reference of such groups may vary from group to group and there are no universally recognised criteria against which assessments are carried out. The sort of areas covered might include:

- whether staff have appropriate training and qualifications, and are fully conversant with the aims and objectives of their work;
- awareness of staff to published work in their subject areas;
- quality and availability of scientific support services;
- adequacy of resources;
- degree of scientific collaboration;
- effectiveness of technology transfer;
- management of the R&D programme;
- whether the organisation of projects effectively meets customer needs.

8.4.3 The strength of the visiting groups approach is that it concentrates on the quality of the science. However the way it is used at present makes it weak in several other respects:

- it lacks harmonised and transparent criteria;
- it tends to look at work retrospectively;
- it is subjective and susceptible to bias.

8.4.4 Assessment visits for accreditation/certification/registration purposes (see above) and visits by customers are a special subset of visiting groups / peer review. In the case of customers, those visiting may lack technical expertise in the areas concerned.

8.4.5 Peer review of publications, also known as citation analysis, involves:

- assessment of the number and quality of publications the laboratory under examination has published in the scientific press;
- assessment of how much those publications are being cited by colleagues within the same research field.

Citation analysis traces its origins to law but is now a widely used, significant research tool, adopted from the field of information science to a range of subject areas. The Science Citation Index (SCI) was first published in 1961. Four particular applications have been reported ^[35, 36]:

1. to assess the impact of individuals, institutions and journals;
2. to investigate hypotheses about the history and sociology of science;
3. to study performance characteristics of information search and retrieval;
4. evaluation tool

Increasingly it is used in the analysis of departmental output or as a measure of the value of the work of a department ^[37, 38].

8.4.6 Some journals will only accept papers for publication that have been the subject of satisfactory peer-review (this is the most common type of peer-review mechanism in use today). As a consequence it is more difficult to publish in these journals. From a citation analysis point-of-view, publication in a respected journal will score better than one in a less respected journal - the so called impact factor. Criteria, ranking journals in order of merit, are published annually by the Institute for Scientific Information. This system has some merit, as published work often reflects the competence and expertise of the publishing laboratory. A laboratory can deliberately raise the profile of its work by publishing as often as possible in the most highly regarded journals. However publication is not always an option and laboratories which do not publish are not necessarily producing poor quality work. One should also be aware that the status of journals sometimes change with time. Citation analysis has a number of other limitations, making it a dangerous technique to use in isolation:

- method papers are cited more often than empirical or theoretical papers, and tend to be referenced due to utility rather than innovation or novelty;
- work ahead of its time is not cited because there are no other scientists interested in the same field of work;
- citations are prone to discrepancies e.g. misspellings;
- citations are rarely complete or comprehensive. Citation counts need to be seen mainly as indicators, and comparisons can only be made if identical citable and citing pools are used;
- negative or contradictory citations tend to indicate a lack of value to the work.

8.4.7 Patents and licences are other forms of dissemination and exploitation that can be used as a measure of a laboratory's output.

8.5 ***Ranking of organisations***

8.5.1 This involves comparing laboratories against a set of common criteria and ranking them on the basis of the comparison.

8.6 **External Quality Assessment procedures (also known as Proficiency Testing)**

8.6.1 Participation in external quality assessment schemes provides an external measure of performance. In non-routine work or R&D, relevant schemes may be difficult to identify or may give an unrealistic impression of performance. Other types of interlaboratory comparison are perhaps more relevant to R&D, such as co-operative studies, but these do not give the same measure of laboratory performance. It should also be recognised that the proficiency testing schemes which give the most reliable measure of performance are those in which the participating laboratories receive the test samples blind.

8.7 **Conclusions**

8.7.1 No single method of assessment stands out as being the most suitable for monitoring the quality of non-routine and R&D work. It is recommended that where some kind of external assessment is required a combination of approaches should be taken and formal assessment should be confined wherever possible to those parts of the quality system that remain stable from project to project, e.g. the management levels and technical infrastructure. Typically this could be established for the 3-tier quality system approach as follows:

Quality Elements	Verification	
	Formal	Informal
Organisational	<ul style="list-style-type: none"> • Certification to ISO 9000 	<ul style="list-style-type: none"> • Follow ISO Guide 25 • Benchmarking • Self assessment
Technical	<ul style="list-style-type: none"> • Accreditation to ISO Guide 25 / EN 45001 	<ul style="list-style-type: none"> • Follow ISO Guide 25 • Visiting groups • Benchmarking • Peer review
Analytical task	<ul style="list-style-type: none"> • Registration to GLP • Proficiency testing 	<ul style="list-style-type: none"> • Follow GLP principles

8.7.2 The informal verification principles outlined above could be made more formal if required and the declared compliance with particular standards, guides or protocol could be independently assessed by a suitable outside body, e.g. a visiting group, or consultant, examining inputs, such as:

- existence of project plans where no elaborated methods are available;

-
- maintenance and calibration schedules;
 - record keeping.

and outputs, such as:

- reports and publications;
- satisfactory participation in relevant proficiency testing, external quality assessment or other intercomparisons.

8.7.3 A well functioning quality system need not stifle creativity in R&D, and is vital for ensuring the smooth transfer of technology from research to diagnostic or commercial environments. Research workers must have an appreciation of the quality requirements of clients and quality must be designed into every process.

9. BIBLIOGRAPHY & REFERENCES

(a) *References cited in the text:*

1. "International Guide to Quality in Analytical Chemistry - An Aid to Accreditation", CITAC CG1, ed. 1, 12/95, ISBN 0 948926 09 0.
2. EAL-G4 - "Accreditation for Chemical Laboratories - Guidance on the interpretation of the EN 45000 series of Standards and ISO/IEC Guide 25", ed. 1, 4/93 (originally EURACHEM/ WELAC GD1/WGD2).
3. ISO/IEC Guide 25:1990, 3rd edition, "General requirements for the competence of calibration and testing laboratories".
4. EN 45001:1989, "General criteria for the operating of test laboratories".

Note: CEN standard complying with [3].

5. ISO 9000 series of standards (primarily ISO 9001:1994, "Quality systems - model for quality assurance in design/development, production installation and servicing"; ISO 9002:1994, "Quality systems - Model for quality assurance in production and installation"; ISO 9003:1994, "Quality systems - model for quality assurance in production and installation"; and others).
 6. "The OECD Principles of Good Laboratory Practice", Environmental Monograph No. 45, OCDE/GD(92)32, Organisation for Economic Co-operation and Development, 1992.
 7. "McGraw-Hill Dictionary of Scientific and Technical Terms", 4th ed., Parker, 1989.
 8. ISO/CASCO 193 (Rev. 2) - ISO/CASCO discussion paper providing up-to-date references on terminology between full updates of ISO Guide 2.
 9. ISO/IEC Guide 2:1996, "Standardisation and related activities - General vocabulary".
 10. "Chambers 21st Century Dictionary", Chambers, 1996. ISBN: 0 550 10588 3.
-

-
11. ISO 8402:1994, "Quality - Vocabulary".
 12. "The Manager's Guide to VAM", Department of Trade and Industry, UK , 9/96, (available through LGC).
 13. Kellner, R., "The WPAC-EUROCURRICULUM on Analytical Chemistry", *Fresenius J. Anal. Chem.*, 1993, **347**, 1-2.
 14. "The Monitor (The Journal for NVQ Professionals)", National Council for Vocational Qualifications, 222 Euston Rd, London, NW1 2BZ.
 15. "Data News", National Council for Vocational Qualifications, 222 Euston Rd, London, NW1 2BZ.
 16. "Development of Evidence Requirements and Assessment Guidelines for an NVQ and SVQ in Analytical Chemistry", Report to the Royal Society of Chemistry, Sept. 1996, UK.
 17. "The Application of GLP Principles to Computer Systems, Advisory Leaflet No. 1", UK GLP Compliance Programme, London 1989.
 18. Bedson, P. J., "Guidance on equipment qualification for analytical instruments", *J. Accred. & QA.*, 1996, **1**, 265-274.
 19. ISO 10011-1:1990, "Guidelines for auditing quality systems. Part 1: Auditing".
 20. ISO 10011-2:1991, "Guidelines for auditing quality systems. Part 2: Qualification criteria for quality system auditors".
 21. ISO 10011-3:1991, "Guidelines for auditing quality systems. Part 3: Management of audit programmes".
 22. EAL-G3 - "Internal Quality Audits and Reviews", European Accreditation of Laboratories.
 23. ASTM D 3864 - 96: "Standard Guide for Continual On-line Monitoring Systems for Water Analysis".
 24. prEN 50228-1, "Performance requirements and test methods for control and monitoring pH analysers for use in the water industry. Part 1: Specification and test methods for assessing performance under laboratory conditions".
 25. prEN 50228-2, "Performance requirements and test methods for control and monitoring pH analysers for use in the water industry. Part 2: Procedures and test methods for assessing performance under field conditions".
 26. Aruajo, P. W., Brereton, R.G., "Experimental design II. Optimisation", *Trends in Analytical Chemistry*, 1996, **15(2)**, 63 - 70.
 27. Caulcutt, R., "Statistics in Research and Development", Chapman and Hall, London, 1983.
 28. Thomas, C. L. Paul, Schoefield, H., "Sampling Source Book - An indexed biography of the literature of sampling", Butterworth Heinemann, 1st ed. 1995. ISBN 0-7506-1947-3 (>2300 sampling references).
 29. Brereton, R.G., "Chemometrics - applications of mathematics and statistics to laboratory systems", Ellis Horwood, Chichester, 1990. ISBN 0-13-131350-9.
-

30. "Harmonised guidelines for the use of recovery information in analytical measurement", IUPAC, in preparation.
31. "The Fitness for Purpose of Analytical Methods - A Laboratory Guide to Method Validation and Related Topics", EURACHEM, 1st ed. 1998, ISBN 0-948926-12-0.
32. "Quantifying Uncertainty in Analytical Measurement", EURACHEM, 1st ed. 1995, ISBN 0-948926-08-2.
33. Lock, D. (Editor), "Gower Handbook of Project Management", Gower, 2nd ed., 1994, ISBN 0-566-07391-9.
34. ISO/IEC Guide 25, revision of 3rd edition, 6th draft 1996, circulated for public comment, ISO CASCO WG10, International Standards Organisation, Geneva.
35. Peritz, B.C., "On the objectives of citation analysis: problems of theory and method", *Journal of the American Society for Information Science*, 1992, **43(6)**, 448-451.
36. Snyder, H., Cronin, B., and Davenport, E., "What's the use of citation? Citation analysis as a literature topic in selected disciplines of the social sciences", *Journal of Information Science*, 1995, **21(2)**, 75-85.
37. Bradley, S.J., Willett, P. and Wood, F.E., "A publication and citation analysis of the department of information studies, University of Sheffield, 1980-1990", *Journal of Information Science*, 1992, **18**, 225-232.
38. Sandison, A., "Thinking about citation analysis", *Journal of Documentation*, 1989, **45(1)**, 59-64.

(b) Selected bibliography

Reference	Subject
Anderson, E.S., Grude, K.V., Hang, Y., "Goal Directed Project Management", Kogan Page, 2nd ed. 1995, ISBN 0 7494 1389 1.	Project management
ANSI Z1-13-19XX "Quality System Guidelines for Research (draft)".	R&D
"AOAC Peer Verified Methods Program - Manual on policies and procedures", AOAC International, Arlington VA 22201-3301, USA.	Food analysis
Association of Public Analysts, "A Protocol for Analytical Quality Assurance in Public Analysts' Laboratories", APA, London, 1986.	General QA
ASTM D 4210 - 89: "Standard Practice for Intra laboratory Quality Control Procedures and a Discussion on reporting Low-Level Data".	QC
ASTM D 4697 - 87: "Standard Guide for Maintaining Test Methods in the User's Laboratory".	
ASTM E 1212 - 92: "Standard Practice for Establishment and Maintenance of Quality Control Systems for non-destructive Testing Agencies".	General QA, QC

ASTM E 882 - 87: "Standard Guide for Accountability and Quality Control in the Chemical Analysis Laboratory".	QC
Bamfield, P., "Research and Development Management in the Chemical Industry", VCH Verlagsgesellschaft mbH, 1996, ISBN 3-527-28778-7.	Management of R&D
BS2846 part 4:1976. (ISO 2854) "Techniques of estimation relating to means and variances".	Statistics

Reference	Subject
BS2846 part 7:1984. "Tests for departure from normality".	Statistics
BS3534 part 1:1993. "Statistics - Vocabulary and symbols - Probability and general statistical terms".	Statistics
BS 4778:Part 2:1991, "Quality vocabulary, part 2, Quality concepts and related definitions".	Terminology
BS 5700:1984, + AMD 5480:1987, "Guide to process control using quality control chart methods and cusum techniques". (complements BS 5701:1980 and BS 5703:1980).	Statistics, QC
Carr-Brion, K. G., Clarke, J. R. P., "Sampling Systems for Process Analysis" Butterworth Heinemann, 2nd ed. 1996, ISBN 0-7506-1247-9.	Sampling
Caulcutt, R., Boddy, R., "Statistics for Analytical Chemists", Chapman and Hall, London, 1991.	Statistics
Chatfield, C., "Statistics for technology", Chapman and Hall, London 3rd ed. (revised) 1995, ISBN 0412 25340 2.	
"Criteria for evaluating acceptable methods of analysis for Codex purposes", Codex Alimentarius Commission on Methods of Analysis and Sampling (CX/MAS 95/3; 1995).	Food analysis
Cook S., "Practical Benchmarking - A manager's guide to creating a competitive advantage", Kogan Page, London, ISBN 0-7494-1551-7.	Benchmarking
DOE standard DOE-ER-STD-6001-92 "Implementation Guide for Quality Assurance Programs for Basic and Applied Research", US Department of Energy.	R&D
Dietrich, C. F., "Uncertainty, calibration and probability", Adam Hilger, London, 2nd ed. 1991.	Measurement uncertainty
Dux, J. P., "Handbook of Quality Assurance for the Analytical Chemistry Laboratory", Van Nostrand Reinhold, New York, 1986.	General QA
"Estimating uncertainties in testing - A guide to estimating and reporting uncertainties of measurement in testing", British Measurement and Testing Association (BMTA), 1st ed. October 1994.	Measurement uncertainty
Funk, W., Dammann, V., and Donnevert, G., "Quality Assurance in Analytical Chemistry", VCH Weinheim 1995, ISBN 3-527-28668-3.	
Garfield, F. M., "Quality Assurance Principles for Analytical Laboratories", Association of Official Analytical Chemists, Arlington VA, 1992.	General QA
"Guidelines for Collaborative Study - Procedures to Validate Characteristics of a Method of analysis", AOAC International, revised May 1994, originally published <i>J. Assoc. Off. Anal. Chem.</i> , 1989, 72 , 694-704.	Collaborative study
Günzler, H., "Accreditation and Quality Assurance in Analytical Chemistry", Springer Verlag, 1994, ISBN 3-540-60103-1.	General QA

Reference	Subject
"Harmonised guidelines for internal quality control in analytical chemistry laboratories", (IUPAC), <i>Pure & Appl. Chem.</i> , 1995, 67(4) , 649-666.	QC
"Harmonised Protocols for the Adoption of Standardised Analytical Methods and for the presentation of their Performance Characteristics", (IUPAC), <i>Pure & Appl. Chem.</i> , 1990, 62(1) , 149-162.	Collaborative study
ISO 3534-1:1993, "Statistics - Vocabulary and symbols - Part 1: Probability and general statistical terms".	Statistics, terminology
ISO 3534-2:1993, "Statistics - Vocabulary and symbols - Part 2: Statistical quality control".	Statistics, terminology
ISO 3534-3:1985, "Statistics - Vocabulary and symbols - Part 3: Design of experiments".	Statistics, terminology
ISO 5725-1:1994, "Accuracy (trueness and precision) of measurement methods and results - Part 1. General principles and definitions".	Terminology, accuracy
ISO 5725-2:1994, "Accuracy (trueness and precision) of measurement methods and results - Part 2. Basic methods for the determination of repeatability and reproducibility of a standard measurement method".	Accuracy
ISO 5725-3:1994, "Accuracy (trueness and precision) of measurement methods and results - Part 3. Intermediate measures of the precision of a standard measurement method".	Accuracy
ISO 5725-4:1994, "Accuracy (trueness and precision) of measurement methods and results - Part 4. Basic methods for the determination of the trueness of a standard measurement method".	Accuracy
ISO 5725-6:1994, "Accuracy (trueness and precision) of measurement methods and results - Use in practice of accuracy values".	Accuracy
ISO 78-2:1982, "Layouts for standards - Part 2: Standard for chemical analysis", (under revision).	Documentation of methods
ISO 7870:1993, "Control charts - General guide and introduction".	Statistics, QC
ISO 7873:1993, "Control charts for arithmetic average with warning limits".	Statistics, QC
ISO 7966:1993, "Acceptance control charts".	Statistics, QC
ISO 8258:1991, "Shewhart control charts".	Statistics, QC
"ISO VIM International vocabulary of basic and general terms in metrology", 2nd ed., 1993.	Terminology
ISO/CD 5725-5, "Accuracy (trueness and precision) of measurement methods and results - Part 5. Alternative methods for the determination of the precision of a standard measurement method".	Accuracy
ISO/DIS 11095 (April 1993), "Linear calibration using reference materials".	RMs, CRMs
ISO/IEC (with BIPM, IFCC, IUPAC, IUPAP, OIML): "Guide to the expression of uncertainty in measurement", 1994. ISBN 92 67 10188 9.	Measurement uncertainty

Reference	Subject
ISO/IEC Guide 30:1992, "Terms and definitions used in conjunction with reference materials".	Terminology
ISO/IEC Guide 33:1989, "Uses of certified reference materials".	CRMs
ISO/TAG4/WG3:1993, "Guide to the Expression of Uncertainty in Measurement".	Measurement uncertainty
Karlöf, B., and Östblom, S., "Das Benchmarking Konzept - Wegweiser zur Spitzenleistung in Qualität und Productivität", Verlag Vahlen, ISBN 3 8006 1831 1.	Benchmarking
Kateman, G., Buydens, L., "Quality Control in Analytical Chemistry", J Wiley & Sons, 1991.	General QA, QC
Litke, H. D., "Projekt management - Methoden, Techniken, Verhaltensweisen", Carl Hanser Verlag, Munchen, 1991, ISBN 3-446-15836-7.	Project management
Lock, D., "The Essentials of Project Management", Gower, 1996, ISBN 0-566-0774500.	Project management
McNair, C.J., and Leibfreid, K.H.J., "Benchmarking: A Tool for Continuous Improvement", Harper Business, London, 1992.	Benchmarking
Mesley, R. J., Pocklington, W. D., Walker R. F., "Analytical Quality Assurance - A Review", <i>Analyst</i> , 1991, 116 (10) , 975-1092.	General QA
Miller, J. C., and Miller, J. N., "Statistics for Analytical Chemistry", Ellis Horwood and Prentice Hall, London, 3rd ed. 1993, ISBN 0-13-030990-7.	Statistics
Murdoch, J., "Control Charts", Macmillan 1979.	Statistics, QC
"Nomenclature for the Presentation of Results of chemical analysis (IUPAC Recommendations 1994)", <i>Pure & Appl. Chem.</i> , 1994, 66(3) , 595-608.	Terminology, statistics
Parkany, M., "Quality Assurance for Analytical Laboratories". Special publication 130, Royal Society of Chemistry, UK. ISBN 0-85186-705-7.	General QA
Prichard, F.E., "Quality in the Analytical Chemistry Laboratory", ACOL, Wiley, 1995, ISBN 0-471-95541-8.	General QA
Prichard, E., Mackay, G.M., Points, J., "Trace Analysis - A structured approach to obtaining reliable results", Royal Society of Chemistry, 1996, ISBN 0-85404-417-5.	QA specifically for trace analysis
"Protocol for the Design, Conduct and Interpretation of Method-Performance Studies, (IUPAC)", <i>Pure & Appl. Chem.</i> , 1995, 67(2) , 331-343.	Collaborative study
Safety Series No. 50-SG-QA16 "Quality Assurance for Research and Development", International Atomic Energy Authority.	R&D
Taylor, B. N., Kuyatt, C. E., "Guidelines for evaluating and expressing uncertainty in NIST measurement results", NIST technical note 1297, National Institute of Standards and Technology, 1994.	Measurement uncertainty

Reference	Subject
Taylor, J. K., "Quality Assurance of Chemical Measurements", Lewis, Michigan, 1987, ISBN 0-87371-097-5.	General QA
"Report on Minimum Criteria to Assure Data Quality", USA - Environment Protection Agency, July 1989.	General QA Environmental analysis
"Validation of test methods - General principles and concepts (sic)", Committee paper EAL/GA(96)58, European Accreditation of Laboratories, Oct. 1996.	Method validation
Youden, W. J., and Steiner, E. H., "Statistical Manual of the AOAC", Association of Official Analytical Chemists, 1975.	Statistics
Zairi, DR M., and Leonard, P., "Practical Benchmarking: A Complete Guide", Chapman and Hall, Cambridge, 1994.	Benchmarking

Annex A Composition of EURACHEM / CITAC R&D / Non Routine Analysis Working Group

Name	Affiliation	Country	Background	Liaison Links
Prof C Adams	Unilever	UK	Industry/ Academic	DIRAG, EURACHEM UK
Prof K Cammann	ICBFhM	Germany	Academic	EURACHEM, EUROLAB, ISO/IUPAC/AOAC
ir HA Deckers	RvA	Netherlands	Industry	WOBAC, EUROLAB EAL
Prof Z Dobkowski	Ind. Chem. Res. Inst.	Poland	Government	Polish Chem. Soc., AOAC, EURACHEM
Mr D Holcombe	LGC	UK	Government	ISO/IUPAC/AOAC
Dr PD LaFleur	Kodak	USA	Industry	CITAC
Dr P Radvila	EMPA	Switzerland	Government	SAPUZ, EUROLAB- CH, EURACHEM
Dr C Rohrer	Lenzing AG	Austria	Industry	
Dr W Steck	BASF AG	Germany	Industry	CITAC, EURACHEM
ir P Vermaercke	S.C.K.	Belgium	Industry	BELTEST

Other inputs to the guide were made by :

- (a) CITAC Working Group with members from: Australia; Austria; Belgium; China; Germany; Hong Kong; Japan; Korea; Mexico; The Netherlands; Russia; Switzerland; United Kingdom; United States.
- (b) EURACHEM full, associate and observer members from: Austria; Belgium; Commission of the EC; Cyprus; Czech Rep.; Denmark; Finland; France; Germany; Greece; Hungary; Iceland; Ireland; Italy; Luxembourg; Malta; The Netherlands; Norway; Poland; Portugal; Russia; Slovakia; Slovenia; Spain; Sweden; Switzerland; Turkey; United Kingdom; United States; AOACI; FECS.
- (c) Miscellaneous inputs have been made by colleagues from Australia, Austria, Belgium, Canada, Cyprus, Czech Rep., Denmark, European Commission SMT Programme, Finland, Germany, Hungary, Iceland, Ireland, Israel, Italy, Malta, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States.

2.1 ***About CITAC and EURACHEM***

- 2.1.1 ***CITAC*** - *Co-operation on International Traceability in Analytical Chemistry* arose from an international workshop held in association with the Pittsburgh Conference in Atlanta, March

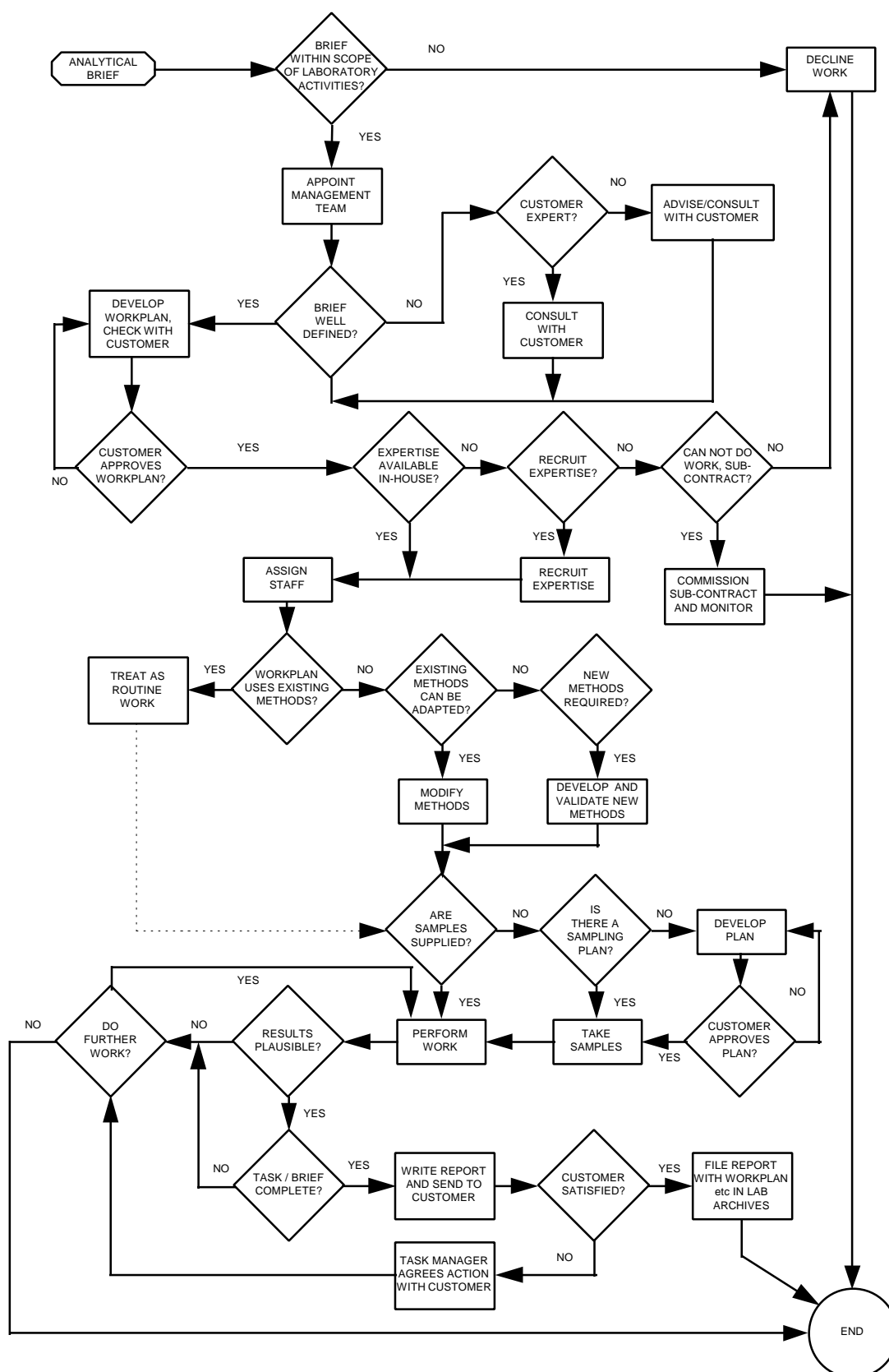
1993. CITAC aims to foster collaboration between existing organisations to improve the international comparability of chemical measurement. A working group co-ordinates activities which include: production of a directory of reference materials under development; preparation of quality system guidelines for the production of reference materials; preparation of a directory of chemical metrology activities; definition of criteria for establishing traceability to the mole; and preparation of an international guide to quality in analytical chemistry ^[1].

2.1.2 **EURACHEM** was established in 1989 to provide a focus for analytical chemistry and quality related issues in Europe. It is a network of European national laboratories which have an interest or responsibility for chemical analysis. It provides a framework which facilitates collaboration between analysts throughout Europe to improve the quality of analytical measurements and provides a forum for the discussion of common problems and for developing an informed and considered approach to both technical and policy issues.

Up-to-date information on EURACHEM activities is available from its twice yearly newsletter, or from its website <http://www.vtt.fi/ket/eurachem.html>.

Both CITAC and EURACHEM secretariats can be contacted via the drafting secretary at the address shown at the front of the guide.

Annex B: Flowchart showing typical lifecycle of R&D project



Annex C Questionnaire for Analytical Work

A. Client

Contact person:

Tel:/Fax:

Address:

B. Objective /goals/required information

Requested analysis:

☐ qualitative/semi-quantitative, limit of detection:

☐ quantitative, range of concentration:

Previous analysis/results:

C. Costs

• Expected costs:

• Cost limits:

D. Date of completion/schedule

Date of intermediate results/reports:

Deadline for final results/report:

E. Sampling☐ client☐ laboratory☐ other

Date of sampling:

Source/producer:

Responsible person:

Number of samples:

F. Description of sample(s)

Identification:

Approx. composition:

Main component:

Minor constituent:

Intended use:

Packaging/stability:

Special care for storage/
transport/ stabilisation:Pretreatment/
preconditioning:Reference materials/
reference sample**G. Methodology**

Description of methods used for sampling, sample preparation, measurement

Standard method:

Generic method:

New/adapted

R&D for new method:

method:

Validation for adopted
method:

Annex D Proposals for the accreditation of R&D tests by type.

D1 Purpose

The accreditation of types of tests serves to provide a flexible description for the scope of accreditation. This annex sets out proposals for possible conditions under which accreditation might be granted for tests by type. Note the responsibility for defining such conditions strictly lies with accreditation bodies.

D2 Area of application

These proposals should be applicable to all testing laboratories aiming for flexibility in their scope of accreditation, especially with regard to R&D work.

D3 Definitions

D3.1 *Type of test:*

“Sector (of a testing field) with similar technical-methodological features, with comparable calibration, validation and training principles.” Types of test may be defined on a technology or application related basis. For example:

- gas chromatography (or perhaps more broadly “separation techniques”);
- atomic spectroscopy;
- thermoanalysis;
- primary fire characteristics.

D3.2 *Testing field*

“Testing fields are sizable sectors distinguished by common fundamentals of a technical, methodological and training related nature.” For example:

- chemical and physio-chemical analysis;
- biological investigations;
- medical laboratory diagnostics.

D3.3 *Flexibilisation*

Flexibilisation of the scope of accreditation is understood to comprise all measures to be taken for accreditation not directed exclusively at the accreditation of individual test methods.

D4 General

The accreditation of types of tests means that the testing laboratories are given the opportunity to introduce new test methods within the approved type of test or of modifying existing methods without having to obtain approval from the accreditation authority in each individual case in advance. It also allows confirmation of the competence of R&D analytical activities on the basis of general work.

Accreditation of a type of test is granted under certain conditions and within the limits governed by the experience which has already been demonstrated by the laboratory for that type of test. Making the scope of accreditation flexible with respect to the methods used does not necessarily imply making it flexible with respect to the sample types under test.

D5 Recommended conditions for the accreditation of types of tests

For every type of test for which the laboratory requires accreditation it should submit to the accrediting body:

- a sufficient number of different test methods, SOPs or test reports;
- procedures for validation or verification as part of the type of test;
- corresponding records of validation and verification.

The methods submitted must reflect adequate operator competence (e.g. technical range) within the type of test applied for. For new or modified test methods, complete documentation and validation is required. For R&D, appropriate test reports and/or generic SOPs may be submitted instead of the test methods.

The laboratory should have available at all times a list of the methods currently covered by its accreditation. The list can be submitted to the accreditor as part of the monitoring procedure, with new or modified methods identified.

D6 Assessment of the scope of accreditation

In the accreditation of types of tests, the assessment is directed in particular towards:

- the organisational prerequisites the testing laboratory has to meet for it to validate or verify new or modified test methods;
 - the qualifications and experience of staff and management and the policies on further training;
 - the level of technical equipment;
 - the procedures for testing;
 - the quality management system;
-

-
- the records of validation and verification carried out.

The assessor has the responsibility for selecting and inspecting key test methods and equipment. The following criteria are amongst those that might be used as a basis for such selections:

- the technical complexity of the tests;
- the possible consequences of errors in performing the tests;
- the frequency of use of the test methods;
- the ratio of routine and non-routine tests.

The extent of the checks should be sufficient to allow the accrediting body to be confident of the capability of the laboratory to introduce new methods or to modify existing methods or to carry out R&D. At the same time the checks must not impose unreasonable costs on the laboratory. The assessor's report should indicate to which test items the respective types of test relate.

D7 Scope of accreditation of types of tests

The scope of accreditation may be specified in terms of:

1. testing field(s);
2. type(s) of test(s);
3. test method(s);
4. item(s) under test.

Annex E R&D to develop analytical instrumentation

E1 The following specific interpretation is recommended for R&D to develop analytical instrumentation.

E2 ***Introduction***

Instrumental R&D involves the improvement of existing analytical systems or development of entirely new systems. The basis for the R&D usually arises from the need for novel systems which are: faster; more sensitive; more accurate; more precise; more discriminating; simpler (and easier to use); more economic; more environmentally friendly; or applicable to different particular analyte(s)/sample matrix combinations. Occasionally it may be carried out on a purely speculative basis, *i.e.* with no particular end application in mind, for example, to investigate the practical potential of a particular measurement principle.

Instrumental R&D projects generally involve building and evaluating prototype instrumentation, making and evaluating changes until the prototype evolves either to a state where performance objectives have been met or further development is not viable. The prototype might be a whole new instrument or an accessory (such as a detector or a chromatography column) for an established instrument.

E3 ***Planning***

Instrumental R&D project planning involves objective setting as with conventional analytical R&D. The research plan effectively involves setting out the strategy for the project and defining the criteria against which the performance of the prototype can be assessed.

E4 ***Experimental design***

The project should include experiments to evaluate and validate instrument performance and to help define the behaviour of the instrument under calibration. Long term stability / acceptable performance should be monitored before the equipment is put into routine use. A means of controlling calibration should be established, either through external adjustment or fixed internally. Suitable standards, blanks, reference materials or check samples of known content can be used in these experiments.

The criteria which cause deterioration of instrument performance should be identified, and wherever possible routines established for controlling these criteria. Where instrument performance is particularly sensitive to operator skill, optimum operating procedures should be established. Checking procedures, using standards, check samples, test mixtures *etc.*, should be established as part of the monitoring process.

Where the instrument under development involves the processing of raw data or signal through some form of algorithm, access to the raw data/signal is advised so that the basic instrumental performance and signal processing can be checked independently.

A number of ways to evaluate and validate the novel instrumentation are possible. Where other techniques/ procedures/ instrumentation exist for the particular measurement application these could be used for the parallel evaluation and validation of the novel instrumentation. Collaborative trial could be used, either involving several laboratories each evaluating the novel instrumentation, or the developing laboratory comparing results generated by its own use of the novel instrumentation against other laboratories using other techniques.

E5 ***Data recording***

Data from instrument evaluation should include a record of conditions under which the instrument is, and is not, working satisfactorily. Typically this will include information on analyte and matrix condition, presence of particular chemical, spectral and physical interferences, temperature, humidity, electrical, magnetic settings. Sufficient data should be recorded over extended time periods and differing conditions to establish the reliability of the technology.

E6 ***Reports***

Where new instrumentation is successfully developed, the reports from the prototype evaluation and validation stages will form the basis for use of the instrumentation in more widespread use, *i.e.* the report is effectively the operating manual. It should include user-friendly instructions for operation of the instrument, applicability, information on storage, calibration and maintenance, and performance checks. Where appropriate, there should be an explanation of how the raw signal is processed by the algorithm for zeroing purposes, so that in routine use incorrect assumptions are not made in the subtraction of blanks. New instrumentation should be subject to equipment qualification procedures before being put into use.

E7 ***Evaluation***

Where the novel instrumentation performance overlaps with existing instrumentation, the success of the R&D can be evaluated by comparison of the two instruments against agreed performance criteria. Unless something is being developed for a particular end use, it is probably easier to test the instrumentation initially against simple problems and then more demanding problems as familiarity with the technique and the behaviour of the instrument improves. In general if the instrument appears to function correctly with one analyte in a single matrix this is not satisfactory evidence for the soundness of the technique *per se*. However it

may be acceptable where the that particular analyte/matrix pair are the main reason for the R&D work.

A Short Practical Introduction to TBCAT_S

Günther Meinrath (2006)

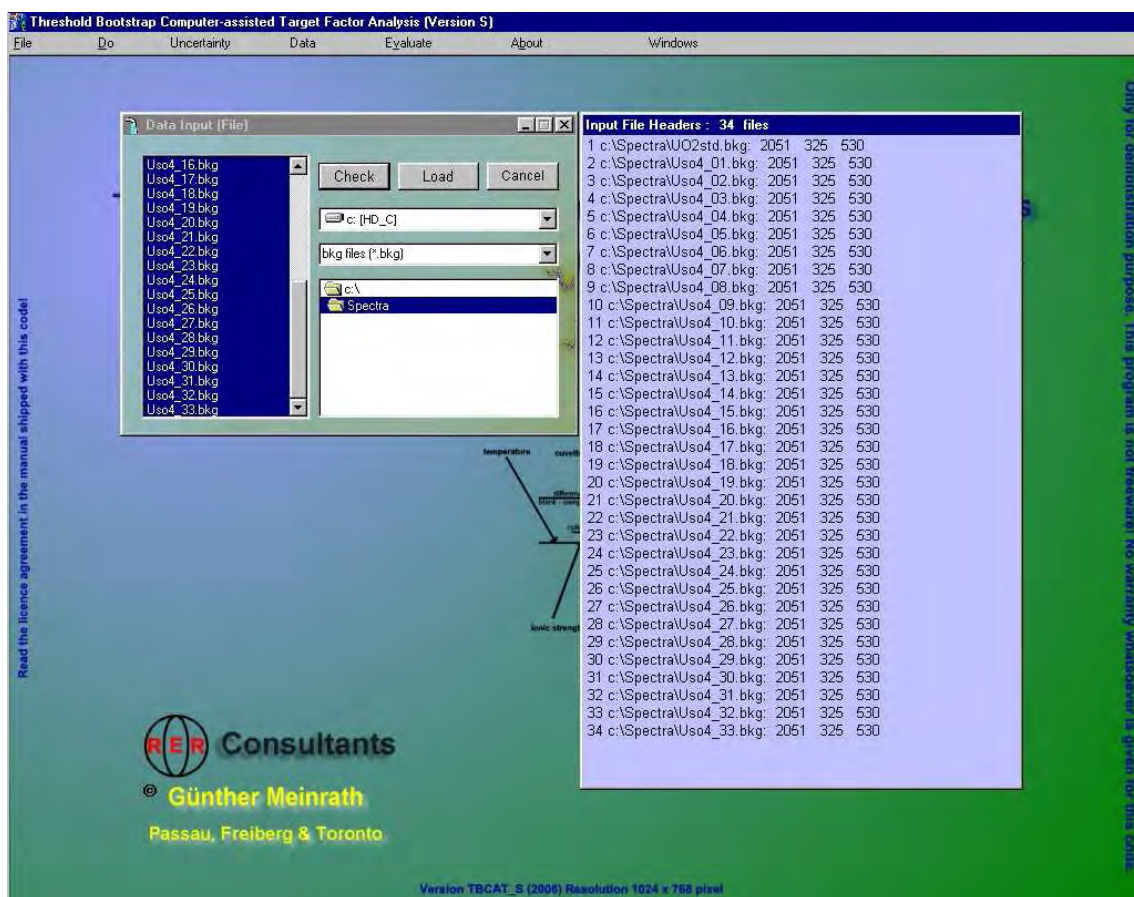


Provided as an addendum to 'Quality Assurance in Chemistry and Environmental Science - Metrology from pH Measurement to Nuclear Waste Disposal' (Springer-Verlag Heidelberg FRG) by G. Meinrath & S. Schneider

These pages give a short guide into the application of TBCAT_S. It is necessary that the manual has been studied and the numerical concepts are at least basically understood. It is further assumed that TBCAT_S has been installed on a suitable computer and the sub-folders of the "TBCAT_S" directory are saved on a hard disk (remove write protection if necessary!).

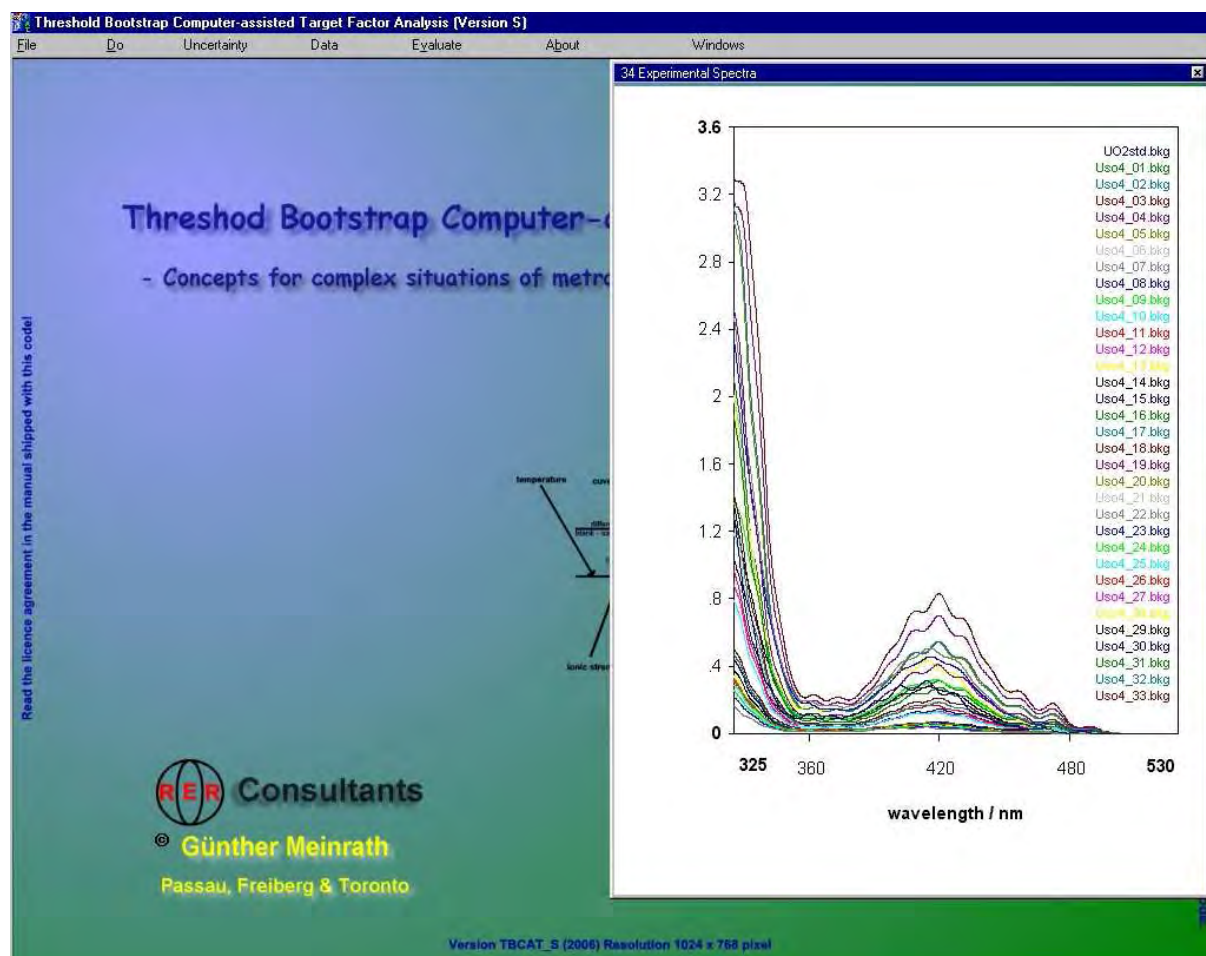
The folder "Spectra_Evaluation" holds the output of an interpretation of 34 UV-Vis absorption spectra by 75 TBCAT_S cycles. The metal ion is UO_2^{2+} , the ligands are OH^- (pH) and sulfate. There are 33 spectra collected at different total uranium concentrations, different SO_4^{2-} content and slightly varying pH. Ionic strength is assumed to be 0.1 M. These spectra have been discussed in Ref. [1]. There are 33 spectra plus a pure UO_2^{2+} spectrum: UO2std.bkg. This spectrum is used as a reference. It has to appear first in the list of spectra loaded to TBCAT_S. The measurement uncertainty values are default values of TBCAT_S.

The total of 34 spectra are loaded into TBCAT_S by selecting all spectra. It is essential that the spectra holding the 'pure component information' appears at first place in the list.



The light blue window on the right side appears if the 'Check' button is pressed. It gives some information of the spectra headers. Its purpose is to ensure that the correct spectra are selected and the basic properties of the spectra are identical. After clicking the 'Load' button, the spectra are loaded. This process takes some time, because a series of checks are done with the spectral data.

The spectral information is presented as a diagram as shown in the next figure. Note that these spectra are already background-corrected (extension is *.bkg). The wavelength range is set to 340 nm and 530 nm (the wavelength range differs from the values to be seen in the figures).



The first step in the analysis of the information provided by these spectra is the ANALYZE process. ANALYZE (in the DO menu) is searching for potential numerical interpretations (e.g. number of species giving rise to these spectra) on a purely numerical plausibility level. Negative absorption, misfit between observed and (re)calculated spectra, negative concentrations are criteria to eliminate certain interpretations. The process is based on factor analysis. Details are found in the manual. The following graph shows the ANALYZE window with reasonable input values. The file name of the file resulting from the process can be chosen freely. If the file name already has been used, a warning will pop up.

The working horse for fitting curves to data is the SIMPLEX. It is a reasonably fast non-linear fitting algorithms. It works without derivative information - which cannot be provided easily for data like that treated here. The SIMPLEX requires some starting input which is generated by TBCAT_S. The user has to provide the maximum number of iterations and a convergence criterion which should be in the range $5 \cdot 10^{-3}$ to $5 \cdot 10^{-5}$. The convergence criterion at occasions has an influence on the result - which, however, is very difficult to assess.

Analyze Mode Input Form

Search Interval

minimum variables: 2

maximum variables: 4

repetitions: 10

SIMPLEX Input Parameters

Iterations: 500

Convergence Criterion: 5e-4

Simplex Result (SOR): filled by program

Result Filename

g: [HD_G]

Buch_Springer

CD_Buch

TBCAT_S

Spectra_Evaluation

test1.lyz

ANALYZE result files (*.lyz)

your_name.lyz

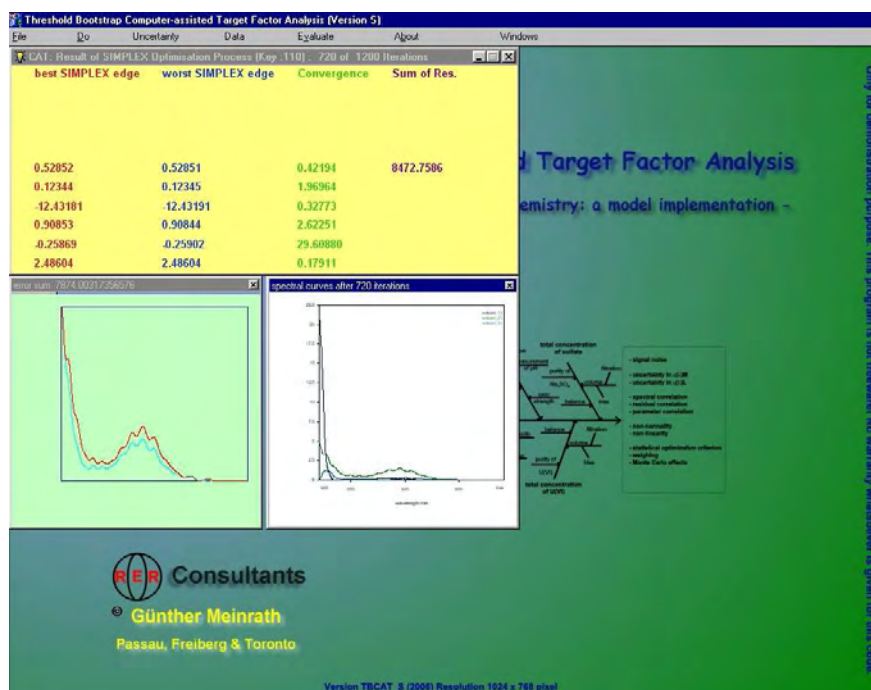
Cancel Start

Upon clicking the **START** button, the process runs automatically. The spectral data are analysed on the possible rank of the data matrix and the compatibility with certain general boundary conditions. The result is expressed by a sum of residuals (SOR). Several criteria apply. The user can influence these criteria under the "Evaluate" menu item "Penalties". The criteria, however, should be modified by experienced users only.

ANALYZE may take considerable time. The total number of runs necessary to test all possible permutations with some repetitions (at least 10 - the SIMPLEX starting values are selected randomly and a minimum sample size is necessary) is time-demanding.

After finishing, the keys and the respective SORs are sorted. Upon clicking on a field in the "ANLYZE: key vs. SOR" table a window pops up allowing to save the respective starting values having given rise to the SOR value in a default file. It is reasonable to keep the suggested file name. It is reasonable to save several default files corresponding to those keys with lowest SOR values. It is necessary to save different keys. It happens regularly that a set of spectra can be nicely interpreted with a certain key independent of the starting values. Then, all repetitions have the same, low SOR. It would be useless to save each of them. The default files will be helpful later because they provide a 'one click' access to reasonable test data in the search for quantitative interpretations. The folder "Spectra_Evaluation" already has some default files for illustration. Do not hesitate to overwrite them.

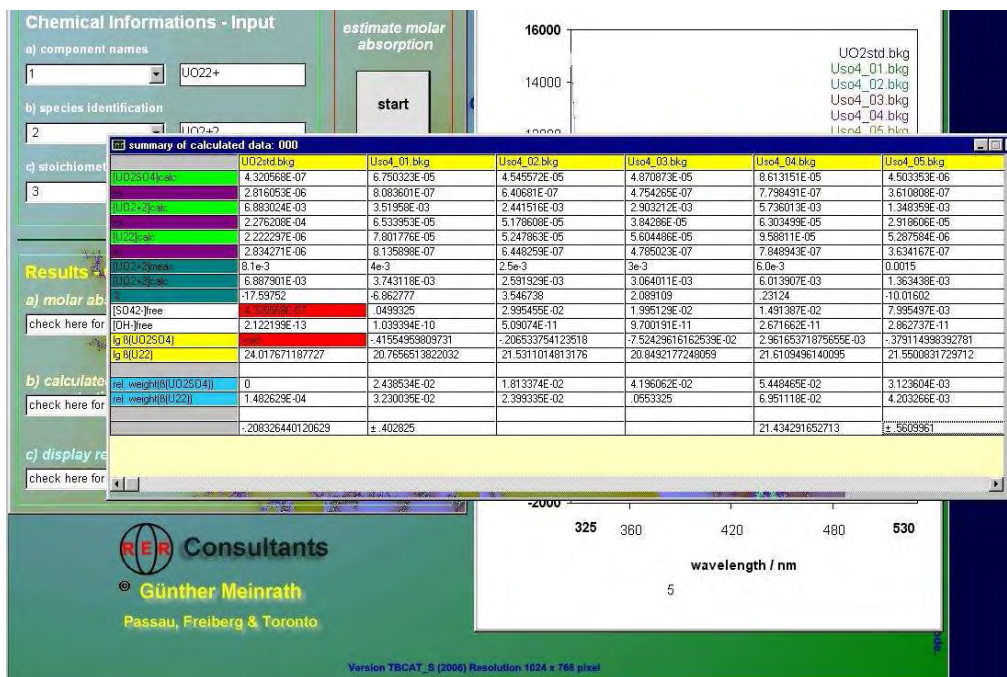
Clicking the 'run CAT' button starts the procedure. CAT is only looking for suitable spectral curves of the single components. The top yellow window gives some informations on the convergence process of the SIMPLEX. Important are the values under the 'Convergence' header and the value of the SOR. Furthermore, two graphics windows appear.



The green window shows the spectrum with known spectral information (here the UO2std.bkg file) with the red spectrum giving the best approximation CAT can obtain. The right side window shows all three (the length of the key "110" is three - therefore three species are assumed by CAT).

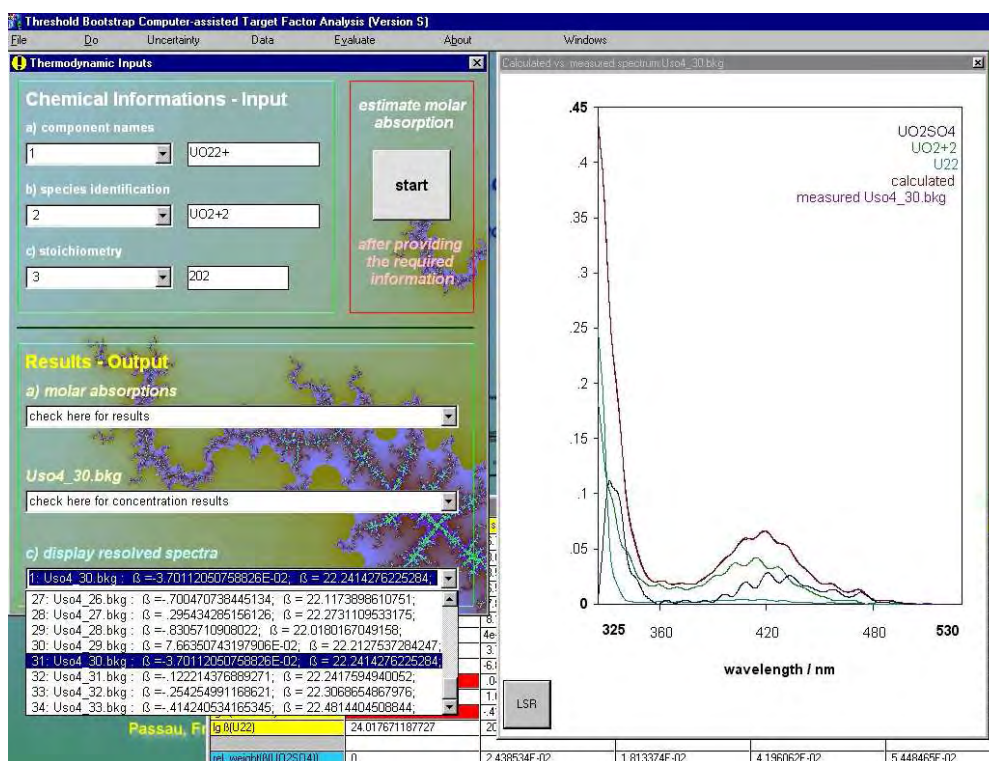
After termination of the CAT process (which often may end in a 'fatal error' message) the results are given. The graphical presentation of the single component spectral curves occasionally allows a general judgement on the feasibility of the interpretation and even a first guess on the species giving rise to the spectrum.

The scaling of these potential single components is done by the 'Molar absorption' item under the DO menu. Here, the chemist must provide good guesses on the possible species in solution and their likely stoichiometric composition. The solutions in this example have three components: UO_2^{2+} , SO_4^{2-} , and OH^- (expressed as pH). The component with the known spectral information (here: UO_2^{2+}) has always the second position. Its stoichiometric composition is 100 (one UO_2^{2+} , no SO_4^{2-} and no OH^-), while $(\text{UO}_2)_2(\text{OH})_2^{2+}$ has stoichiometric composition 202, while UO_2SO_4 is represented by 110. These values have to be entered in the drop-down fields of the "Chemical Information - Input" frame. The process starts with the big 'Start' button. An SIMPLEX procedure starts trying to scale the single component spectra that all spectra can be interpreted with minimum deviation between known total concentration(s) and minimum spectral deviations etc. The result is given graphically and as a table, as shown in the next graph.

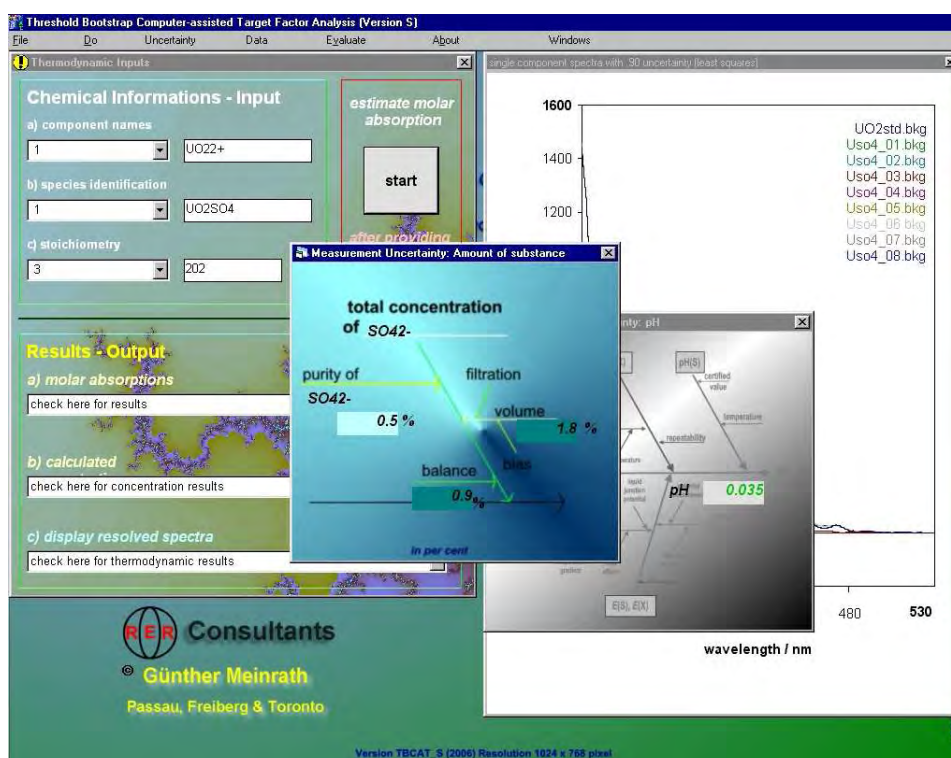


Red fields indicate physically impossible values (commonly negative concentrations or absorptions). Computers do not like the numerical value 'zero'. Small negative values are not significant - the meaning generally is 'zero'. Larger negative values, however, render an interpretation as questionable - or the input data as inconsistent. In the bottom line of the "Summary of calculated data" table, the formation constants and a first estimate of the variability are given. Some other information of interest characterising the fitting process and its result are given in this table.

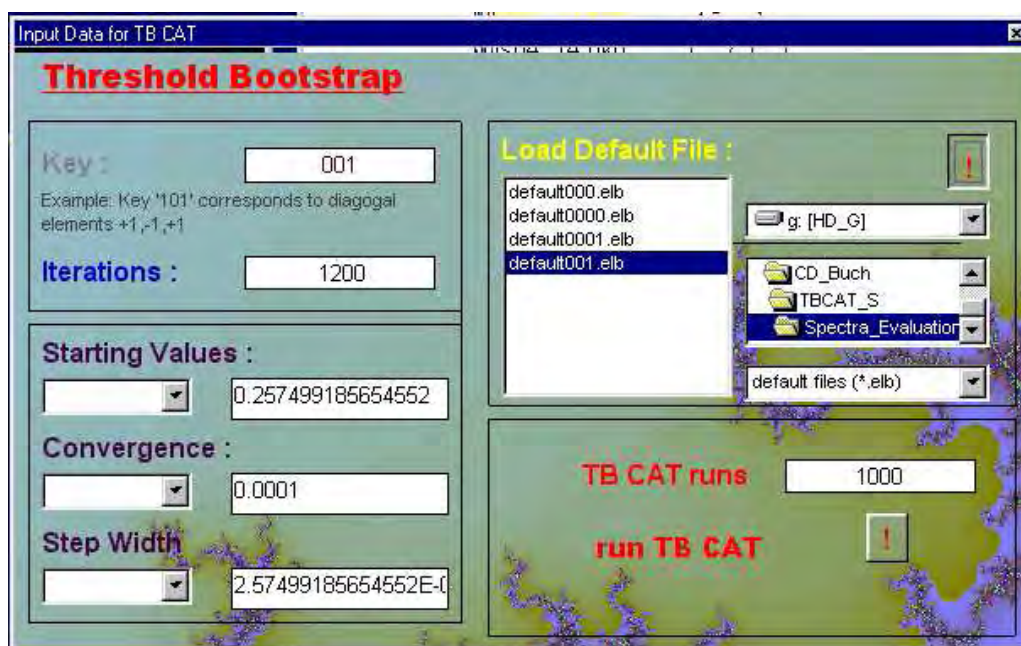
The 'Results - Output' section of the 'Thermodynamic Inputs' frame summarises further information. Selecting a spectrum under the 'display resolved spectra' drop-down list draws the spectrum, the contributions of the single components and the measured and calculated spectrum. This is shown in the following graph. The 'LSR' button performs a more detailed least-square analysis on basis of the QR algorithm. These calculations are quite time-consuming. The graphical data can be saved by selecting the SAVE option in the 'File' menu. The same is true for the summary table. The information is saved as ASCII file.



If an acceptable (and, hopefully, unique) interpretation is identified, the TBCAT process may start. TBCAT uses computer-intensive statistics to simulate the measurement process numerically. The result is a first estimate of the complete measurement uncertainty budget of the measurement process. The uncertainties for different influence quantities can be entered. The uncertainty input frames are accessed via the 'Uncertainty' menu item.



TBCAT requests the fitting information and the chemical information successively. In addition the 'TB CAT runs' item in the respective frame is enabled. A minimum of 1000 repetitions should be done.



After clicking the 'run TB CAT' button, the frame for the chemical informations appears. The required data should correspond to the values previously identified in the CAT runs. The TBCAT process is time-consuming. A careful analysis with 2000-5000 TBCAT cycles easily requires several hours even on a GHz CPU.

The directory 'Spectra_Evaluation' holds the output of a test run with just 75 repetitions for illustration purposes. TBCAT builds up a large number of output files holding the spectral and the numerical information from the individual runs. This bunch of information is transformed into empirical distributions and probability densities under the 'Evaluate' menu item. 'Spectral uncertainty' evaluates the empirical cumulative probability distribution for each wavelength, reads the confidence limits, and writes this information into an output file with defined file name.



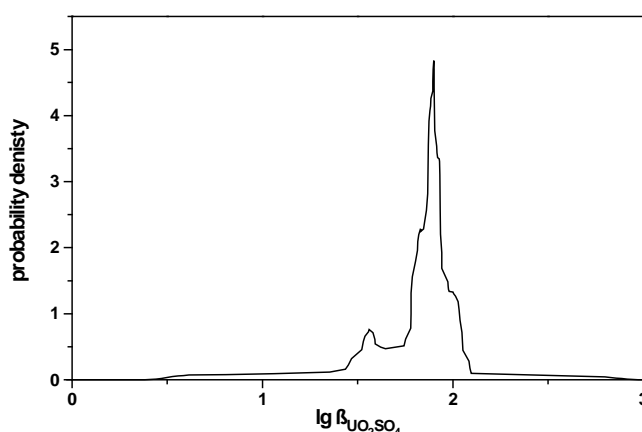
The process requires that the confidence limits are checked and the single component spectrum of interest is chosen. The 'Spectral Uncertainty' frame lists the first spectrum of all

available components from which one can be selected. The process may take some time depending on the total number of TBCAT cycles and the number of wavelengths in the files. The result is save under a name derived from the user input for the respective species. The defaults are 'species(1)', 'species(2)' and 'metal ion'. If more characteristic species names have been selected, e.g. 'UO₂SO₄' the result will appear as ' CDF_UVUO₂SO₄4001_3.dat'. As additional information, the filename holds the key and the position of this species.

From each TBCAT cycle, formation constants of the respective species are obtained. These are automatically saved after finishing the TBCAT process. The filename convention is 'cdf_UO₂SO₄4001.dat' for a species 'UO₂SO₄'. Selecting the 'Differentiate' item from the 'Evaluation' menu, the respective frame will appear.

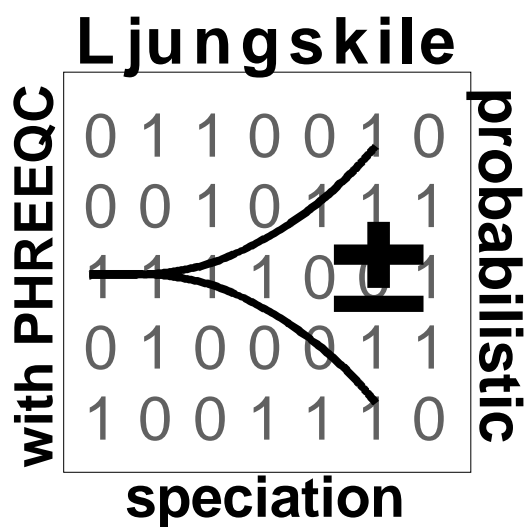


The probability density will be saved with the prefix 'dif'. On basis of the 75 test cycles, the following distribution for the formation constant $\lg \beta_{11}$ of the UO₂SO₄ species has been obtained:



Reference:

Meinrath G, Lis S, Piskula Z, Glatty Z (2006) An application of the total measurement uncertainty budget concept to the thermodynamic data of uranyl(VI) complexation by sulfate. J Chem Thermodynamics 38: 1274 - 1284.



LJUNGSKILE _S

**A computer program for investigation of
uncertainties in chemical speciation**

by Christian Ekberg, Arvid Ödegaard-Jensen and Günther Meinrath

last addition: March 2007

License

Please read the terms of this agreement and any provided supplemental license terms (collectively 'agreement') carefully before installing the respective software. By installing the software, the user agrees with the terms of this agreement. If the user is accessing the software electronically, he complies with this agreement by installing the software. If the user does not agree to all these terms, he must promptly deinstall this software and return the unused software or, if the software is accessed electronically, remove the software from the computer(s).

1. The copyright owner grants the user a non-exclusive and non-transferable right for the internal use only of the accompanying software and documentation. It is understood that the accompanying software is experimental – no warranty is included that the software may work on a certain computer.
2. The software is confidential and copyrighted. It is distributed free of charge but is not in the public domain. Title to software and all associated intellectual property rights is retained by the copyright owner. Any modification, especially by reversed engineering and decompilation, is forbidden except if explicitly authorized by the copyright owner. The user may not make copies, other than for archival purposes. No right, title or interest in or to any trademark, service mark, logo or trade name what so ever is granted under this agreement.
3. This software may be used, explicitly, for education purposes in an academic environment if it deems suitable. All necessary operations (except those requiring copyright infringement) may be performed with the software (e.g. installation of the software in a computer pool). There is any implicate or explicite claim that the software may be suitable for such a purpose. The sole responsibility remains with the user.
4. This agreement is effective until terminated. If any provision of this agreement is held to be unenforceable, this agreement will remain in effect with the provisions omitted, unless omission would frustrate the intent of the parties, in which this agreement will terminate immediately.

Citation

All reports of results obtained by application of the software, including illustrations, must be accompanied by the following citation:

Ödegaard-Jensen A, Ekberg C, Meinrath G (2005) LJUNGSKILE: a program for assessing uncertainties in speciation calculations. *Talanta* 63: 907 - 916.

Disclaimer of Warranty

Unless specified in this agreement, all expressed or implied conditions, representations and warranties, including any implied warranty of merchantability, fitness for a particular purpose or non-infringement are disclaimed, except to the extent that these disclaimers are held to be legally invalid.

To the extent not prohibited by the law, in no event will the author or the programmer of the code be liable for any lost revenue, profit or data, or for special, indirect, consequential, incidental or punitive damages, however caused regardless of the theory of liability, arising out of or related to the use of or inability to use this software, even if the authors have been advised of the possibility of such damages. In no event will the authors' liability to the user, whether in contract, tort (including negligence), or otherwise exceed the amount paid by the user for Software under this agreement. The foregoing limitations will apply even if the above stated warranty fails of its essential purpose. This agreement is effective until terminated. The user may terminate this agreement by destroying all copies of software. This agreement will terminate immediately without notice from the authors if the user fails to comply with any provision of this agreement. Upon termination, the user must destroy all copies of software.

Preface

It has never been the intention of the authors to generate a professional computer program. The computer code presented here merely intends to satisfy their curiosity on the likely variability introduced into speciation calculation by the uncertainties specified sometimes together with the respective formation constants. A code handling this issue in a general and reasonably convenient way has not been to the knowledge of the authors.

It is a general observation that satisfying a simple curiosity will give rise to more questions requiring more elaborate efforts. Thus it is important to keep the limited scope of this code in mind. It is more appropriate to consider it as a starting point of a discussion how to propagate uncertainties in scientific data onto the quantities derived from them. The authors have used several approaches to map uncertainties in thermodynamic quantities to the quantities calculated by using these thermodynamic data (Ekberg, 1999; Meinrath, 2000b). The Latin Hypercube approach by McKay, Beckman and Conover, discussed in 1979 in *Technometrics*, eventually emerged as the method of choice.

Chemical equilibria can be expressed by a series of thermodynamic equations. Most chemists have to solve such calculations for a simple system at least once during their academic education. Such calculations rapidly gain in complexity requiring sophisticated computing codes for their solution even if the number of chemical constituents in the system seems still quite handy. Computer codes solving these tasks have been developed in the past like MINEQL, EQ3/6 and PHREEQE - to name just a few of them. There was no need to reinvent the wheel another time. Thus, Parkhurst and Appelo's PHREEQC 2001 version has been selected as the working horse responsible for the number crunching.

The computer program evolving from combining PHREEQC with a Latin Hypercube strategy has been named "Ljungskile" for reasons specified below. It has a level of sophistication satisfactory for the intentions the authors had it designed for. The user may apply it to study the consequences of uncertainty in thermodynamic informations to the calculated species distributions.

With the need to communicate scientific results in a comparable way, normative guidelines have been issued in recent years demanding chemists to associate meaningful estimates of doubt together with the informations they provide. The past efforts of the authors to investigate the perspectives of these normative requirements for thermodynamic data have indicated an enormous challenge. Thus, the Ljungskile program may be seen as a mosaic stone in the effort to place aquatic chemistry in international framework of the measurement uncertainty concept and to profit from the world-wide traceability and comparability of measurements.

The authors have attempted to make the Ljungskile program as stable and versatile as possible. However, since the scope of the program has been limited, the results should not be overinterpreted. The authors explicitly state that there is no warranty whatsoever that the results calculated by this program represent something like truth or reliability. The Ljungskile code is intended as a tool to gain insight into our abilities and limitations to communicate doubt in scientific measurement.

Table of contents

Preface	2
Table of Contents	3
Lisencing Agreement	4
The LJUNGSKILE Code	5
Introduction	6
True Values and Uncertainties: The Necessity of Doubt	7
The Ljungskile Code: Motivation	9
Program description	10
Chemical calculation	10
Some Comments on Statistical and Programming Methods	11
Installation and Description of Files	15
Getting Familiar: A Step-by-Step Example	18
Description of Output	22
The LJUNGSKILE Display Program	24
References	26
Appendix I: Random Numbers	28
Appendix II: Latin Hypercube Sampling	30
Appendix III: The Virtual Species Im and Ip	33
Appendix IV: A Rough and Dirty Way to use LJUNGSKILE Output	34

LISENCING AGREEMENT

The **LJUNGSKILE** code is an extention of the public domain **PHREEQC** code. It is freeware which means that no monetary requests by the authors exist. However, it is a scientific contribution to be acknowledged by the users. Acknowledgement is done by inclusion of a reference to this report:

Ödegaard-Jensen A, Ekberg C, Meinrath G (2005) **LJUNGSKILE**: a program for assessing uncertainties in speciation calculations. *Talanta* 63: 907 - 916.

By installation of the code on a computer each user implicitly agrees with this lisencc agreement and admits that non-compliance is an unscientific behaviour.

The LJUNGSKILE Code

The LJUNGSKILE code was implemented in a Swedish summer house in July 2001 close to the village of Ljungskile in Bohuslän district.



Figure: Location of Ljungskile (black circle), north of Göteborg, Sweden

Ljungskile is about 70 km north of Göteborg and has all essentials in easy reach: a lake, the sea, a boat berth and a small airport. There is a small hut up there with no running water and otherwise minimal sanitary equipment in addition to the vast and generous spaces offered by nature. It is obviously a good place to create a computer program.

The same restraint luxury is provided by the LJUNGSKILE code. This reluctance to forward convenience and comfort has a good reason: it does not restrain the creativity of the user.

The user is able to generate his own data bases and input files, select species and associate uncertainties. He can choose among several distributions (i.e. normal distributed uncertainty or uniformly distributed uncertainties) for the thermodynamic data of each species. The complexity of a simulation is to his choice as are essential project

parameters like the composition of the water in which the reaction should be simulated.

The informations are returned by **LJUNGSKILE** code in a series of data files holding the concentrations of the relevant species. With the **LJUNGSKILE** Display Program as a supplement to the Ljungskile code, a convenient tool is provided to visualize the **LJUNGSKILE** calculation output, to modify and to save it.

INTRODUCTION

Natural aqueous systems commonly hold many components that may react with each others and the contacting geological materials. Substances enter solution and precipitate, gases dissolve from and are released into a gaseous phase, metal ions are hydrolysed, various constituents coordinate to each others. In fact, multiple processes in aqueous phases may occur in parallel on different time scales ranging from picoseconds to many years.

The microscopic turmoil in any aqueous material, however, has been rationalized for many decades by the concept of the chemical equilibrium. The rational of this point of view is provided by chemical thermodynamics and the three fundamental laws of chemical thermodynamics. On this basis, a many chemical reactions have been studied with the intention to derive quantitative information on their fundamental chemical properties expressed as enthalpies, entropies, heat capacities and formation constants. To make these heaps of individual bits of information work, however, powerful computers are indispensable as soon as more than two or three competing reactions in a solution have to be considered.

To properly understand, and thus being able to build models to mirror these phenomena, knowledge of the speciation of a certain element in a given environment is of fundamental importance. In a multi-component system, the speciation must be estimated conditional on all other interfering reactions in the system of interest. Hence, computer simulations of complex chemical systems are important for prediction and understanding of many phenomena in nature, e.g. sorption, chemical reactions and transport calculations, both physical and biological.

The speciation is usually obtained by using some thermodynamic equilibrium code, e.g. PHREEQE or EQ3/6 (Wolery, 1992). These codes need input data in the form of thermodynamic data such as stability constants and physical data such as water composition to work. All such input data are encumbered by uncertainties of different kinds (Ekberg, 2001; Meinrath, 2000a).

True Values and Uncertainties: The Necessity of Doubt

Quantitative modern science strives for true values. The idea of the 'true value' is fundamental for modern science. Its importance cannot be underestimated. However, even so being essential for the epistemological foundation of science a true value cannot be measured. The idea of true values known to nature, universally reproducible in time and space, is the basis of science.

All measurements are affected by unavoidable uncertainties. On a very fundamental level, the Heisenberg principle may be invoked to explain the persistence of uncertainty. However, less sophisticated effects can be named: the complexity of experimental set-ups and variety of procedures to investigate a quantity of interest, the impossibility to control all and every influence on the experiment under study, the lack of unique references to compare the obtained results. Last but not least, the imperfections of the human being must be acknowledged: *errare humanum est*. Human activity is not able to perceive an absolute truth.

All measurements are affected by a series of nuisance effects that reduce the accuracy of the results thus introducing doubt into the values obtained. It is usually only possible to give a range in which there is a certain probability of the true value being found.

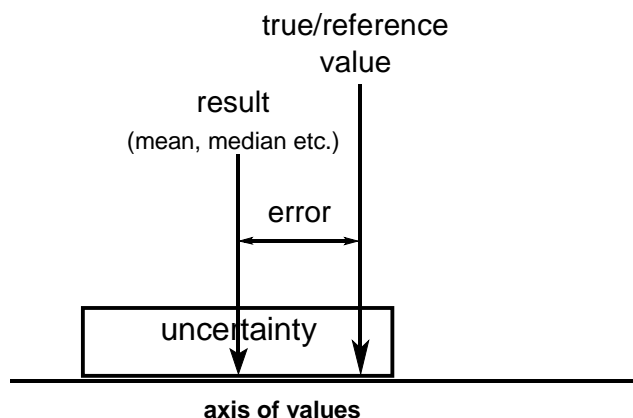


Figure 1: Metrological concept of true value, uncertainty and error (Meinrath, 2000a)

The uncertainty associated with a measurement is an interval that cannot be corrected for. While the measured value is desired to be as close as possible to the true value, the uncertainty does not have an equivalent 'true' counterpart. There is nothing like 'true uncertainty'. Uncertainty is an instrument for communication between humans (commonly scientists) about the likely discrepancy between true and measured value. The importance of the communication tool 'uncertainty' must not be underestimated (Ellison, 1997). It is an essential tool in transferring information about the quality of knowledge obtained and the doubt to be reasonably associated with this knowledge.

The complexity of our world and the activities taken place in it crucially depends on information with stated quality. Theoretical chemical speciation is only a minuscule element in this complexity. The decisions based in part on geochemical modelling for the performance assessment of nuclear waste repositories however have the potential to affect life on this world for some ten thousands of years.

Thermodynamic databases are often claimed to be consistent and as correct as possible. However, since stability constants generally are derived from experiment there will always be uncertainties in the determinations. The effect of these uncertainties on the speciation of an element in a given water may, in some cases, can be considerable (Ekberg, 1999). However, calculating such dependencies is very cumbersome to do by simple means even for a very simple case with only one ligand. Thus, for a more realistic case with several competing ligands, more elaborate methods are needed. There is also the problem of statistical sampling to make the obtained confidence interval good in a statistical sense. For these purposes the **LJUNGSKILE** program has been developed. It is an easy-to-use program with simple menus in a WINDOWS environment.

The LJUNGSKILE Code: Motivations

The abundance of fast computing power has largely simplified theoretical simulations of chemical equilibria even in rather complex solutions. A solubility is quickly calculated and a species diagram can be produced within a few minutes. Even those not well acquainted with computer programming languages may use freeware or commercial codes to produce such graphs. These graphs are, however, almost inevitably based in the mean values of certain thermodynamic quantities. It is not uncommon to find experimental observations interpreted within a tenth of a promille of a species read from a species distribution - based on uncertainty affected formation constants with up to a logarithmic unit of uncertainty (Ekberg, 2002).

In part, very far reaching conclusions are based on such theoretical simulations mostly without accounting for the uncertainties involved into such calculations. It has been shown elsewhere that existing thermodynamic data and data bases are neither consistent nor comparable. There is, however, no simple remedy in sight for this problem (Meinrath, 2002). The fundamental thermodynamic theory is available but our techniques are not yet powerful enough to warrant highest accuracy. If many uncertain parameters enter a calculation, the uncertainties accumulate and the calculation result will become comparatively vague. Uncertainty means doubt: the result of the calculation cannot be trusted arbitrarily. It is of outmost importance to evaluate a measure of the doubt involved. Despite this necessity, suitable tools for speciation calculations are not yet commonly applied.

A computer tool seemed appropriate that visualizes the effect of the figures behind a '±' symbol on the result of a speciation calculation. There are good reasons to base such a tool on Parkhurst's PHREEQC code (Parkhurst, 1995) because PHREEQC is widely used, comparatively easy to learn, freeware, well maintained, versatile with a good manual, stable in execution and manageable in size and use of computer memory.

The LJUNGSKILE code allows the user to select two approaches: the Monte Carlo (MC) approach and the Latin Hypercube Sampling (LHS) (McKey 1979). Latin Hypercube sampling (LHS) allows to produce a satisfactory statistics with a minimum of CPU time. It is, in general, possible to do a simple theoretical speciation calculation within seconds. There are, admittedly, alternatives to LHS and there is criticism towards the uncritical use of LHS output because commonly correlation between some of the input variables exists. LHS, like MC, is not capable to take these correlations into account. Such a correlation can, i.e. exist between the pH of a solution and the partial pressure of CO₂: higher pH solutions may absorb larger amounts of CO₂ and can reduce the CO₂ partial pressure. It is therefore of advantage to combine the both variables in a way that adequate pH/CO₂ partial pressure pairs are matched in the input vectors while LHS does not account for such correlations. Taken the generally rather poor degree of consistency, comparability and completeness in the existent thermodynamic data, however, such argumentation -justified as it is from a fundamental point of view- seems to be premature. It should be mentioned, however, that viable ways to account for correlation in input parameters have been reported in literature.

PROGRAM DESCRIPTION

The LJUNGSKILE Code

The **LJUNGSKILE** program is written in the programming language C++ using the Borland C++ builder (Borland). The program is divided into two major parts. One part containing the chemical calculations and one containing the statistical methods used. The basic idea is to make many runs with PHREEQC and in each run the stability constants of interest are changed slightly within their respective confidence interval according to the statistical method used (see below). Then, after each such cycle, if this option was selected, one of the factors may be changed by given increments to a given level, e.g. changing pH, and thus create a speciation diagram versus pH with uncertainty bands for each species concentration.

Chemical Calculations

The basic approach of the **LJUNGSKILE** program is to use different statistical sampling techniques to determine the effect of uncertain stability constants on a speciation calculation. The chemical speciation calculations are made by the well-known thermodynamic equilibrium program, PHREEQC. The chemical choices the user has to make are given in Figure 2.

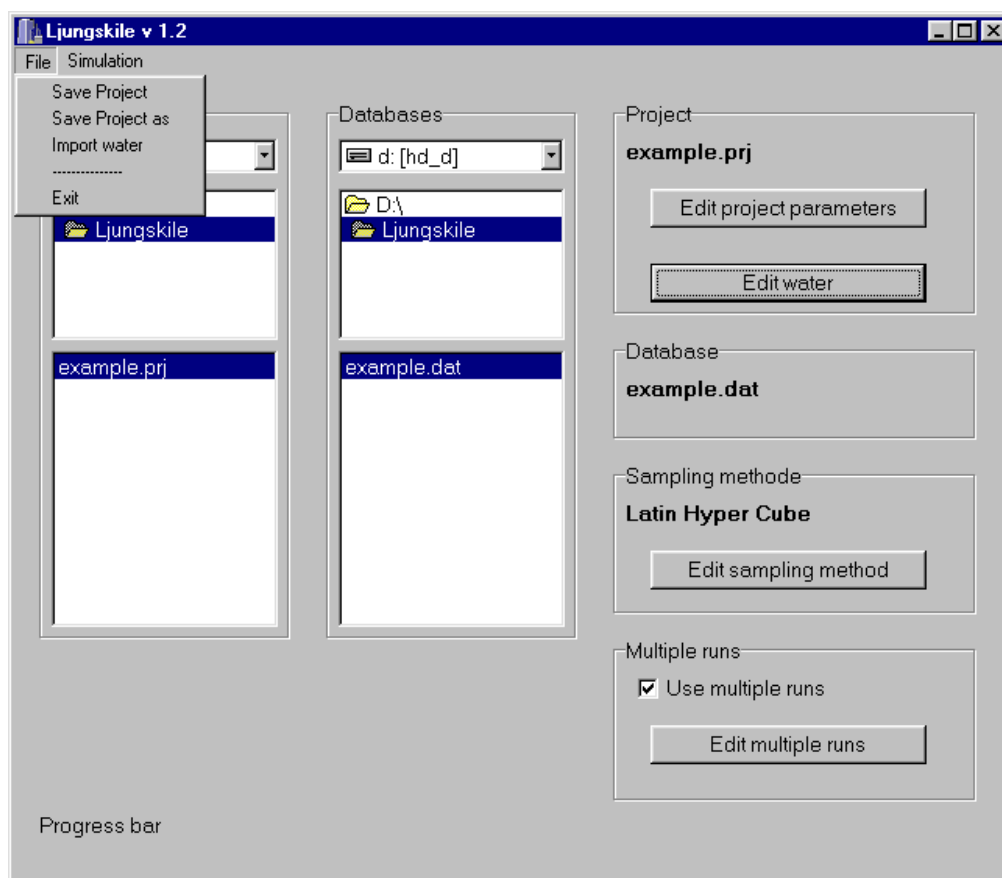


Figure 2: LJUNGSKILE main window

Selection of project file: The project file holds all relevant information of the task to be treated by LJUNGSKILE code. The project parameters are formation constants with respective uncertainties and a choice of uncertainty distribution. A solid phase may be specified and its amount in the solution can be given. CO₂ as a gas phase may be invoked and the partial pressure specified.

If the project parameters are set, the water is defined, the sampling method selected and the multiple run feature chosen or ignored, then Ljungskile is put to work by clicking "Simulation" in the menu. Before the code starts, the project file will have to be saved under a file name given by the user. If a project file is created from scratch, no name exists and a window will pop up ("Cannot create file"). The project file must be saved under a suitable name using "save as" in the File menu.

Selection of water: The user gives the chemical composition of a water together with other quantities such as pH, pe and temperature. This information is then written in the input file to PHREEQC. Its filename is set by default to "PHREEQC.IN" and cannot be changed. It is noteworthy that if the water is not charge balanced on input it will become so by the addition of an inert positive or inert negative element, Ip and Im, respectively. The concept of these species is briefly explained in Appendix IV.

Selection of database: The so-called thermodynamic databases are actually collections of some mean values for formation constants and solubility products of a certain number of chemical compounds. The format is depending on the very program designed to read this data file. We will nevertheless stick with the common term 'database' in this manual. The choice of a certain database may influence the calculated result considerably. Hence, the user may generate his own database. As long as the database is in agreement with the PHREEQC format, the LJUNGSKILE output will be based on the selected database. In some cases the desired species are not present in the general database and therefore there is a possibility to change the database used. Note, that the species Im and Ip must be included into any self-created database in the way they are given in the example database.

Selection of sampling method: The sampling methods offered are the Monte Carlo design and the Latin Hypercube design. The Latin Hypercube design is more efficient if a larger number of thermodynamic constants have been selected to vary in the project, while MC sampling is appropriate for just a handful of thermodynamic constants to vary. However, it is always helpful to have the possibility to compare different approaches and, hence, the MC approach is provided as a choice.

Selection of multiple run parameters: Sometimes it is also interesting to investigate the evolution of a situation as a function of a parameter. The parameter, its starting and stopping value and the step width can be specified here. Note, that the multiple step feature has to be used wisely. Especially if solid phase dissolution and/or precipitation occur, the solution variations may become considerable. The LJUNGSKILE program accounts for charge balancing by adding 'virtual' species Im and Ip but the result do not necessarily have to be meaningful.

PHREEQC settings

As described above the water composition is given by the user together with pH, pe and temperature. If the given water composition results in an charge imbalance this is

taken care of by the **LJUNGSKILE** program. The method used is to use the inert elements Ip and Im which are positive and negatively charged, respectively. This is made by first adding $1 \cdot 10^{-6}$ M of Im and then use the flag CHARGE. In case PHREEQC does not converge Ip is used in the same manner and thus charge balance is obtained, see Appendix III for a typical PHREEQC input file. Thus it is simple to change the concentration of one of the elements present in the water without affecting the other water properties. This is a necessity since in a speciation diagram there is generally, at least, one factor changing along the x-axes.

Other settings in the PHREEQC input file is the use of the PRINT keyword. The flags used here are: -reset false, -totals true, -species. This is to make the result file clearer and more structured (interested readers are directed to the PHREEQC manual).

Some Comments on Statistical and Programming Methods

Statistics is looking for objective criteria to extract informations from larger amounts of data. Data may be considered to be raw information in a qualitative or quantitative way. Data producers have the obligation to present all pertinent information that would impact on the use of it, to the extent possible. Of course, every possible use of data cannot be envisioned when it is produced, but the details of its production, its limitations and quantitative estimates of its reliability always can be presented. In an ideal situation, a reasonable measure of the doubt to be associated with given data will always be available.

There remains the task to act on basis of doubtful information. Within a computer program, handling and even creating uncertainty seems a contradiction to the foremost abilities of a computer being a deterministic machine for invariable execution of instructions (Knuth, 1981).

However, the computer is a very helpful tool - and its helpfulness is the larger as there is no other tool to perform the task - to handle doubt and uncertainty a) because it is a deterministic machine b) because it is fast and c) because suitable algorithms exist.

The use of computers for theoretical speciation calculations started rather early with the computer programs of Lars Gunnar Sillén in Sweden (Sillén, 1962; Sillén, 1964). Several major achievements have been included into Sillén's HALTAFALL modelling codes: application of suitable minimization routines for linear equations, suitable input formats for chemical information and programming techniques to accommodate all code in the tiny storage area of the early computers. Compared to these early applications, modern desktop machines represent the computing power of several computer centres in Sillén's time. Since then, storage capacity, CPU clock rates and algorithms have improved dramatically. The elements, however, have to be brought to work.

The sampling in the **LJUNGSKILE** program is made by either simple Monte Carlo (MC) sampling or Latin Hypercube sampling (LHS) (McKay, 1979), as described in detail in APPENDIX II. The basic step for LHS is to make a cumulative distribution function (cdf) for the sample based on the mean value and the standard deviation given for the respective variable (Meinrath, 2000a). Here the user may select how precise this cdf itself should be by selecting how many points are to be included in in

its construction. The samples are then drawn from this empirical cdf, as described in APPENDIX III. For the MC sampling there is no extra work since this method samples complete randomly within the given interval/distribution. Random number generators are further discussed in APPENDIX I.

In the current version of the **LJUNGSKILE** program there are two distributions and two sampling methods to select from. The distributions are the uniform distribution and the normal distribution. The sampling intervals are given slightly different for the two distributions. The uniform distribution requires a highest and a lowest value while the normal distribution uses a mean and a standard deviation.

The LJUNGSKILE Display Program

LDP is intended to graphically display the results obtained by the **LJUNGSKILE** 'multiple run' option. It is not a sophisticated graphics program but kept comparatively simple. LDP gets its information mainly from the respective project file in the \results\ directory. The **LJUNGSKILE** program copies these files to the results directory and renames the project file into an *.ldp file. Thus, LDP is set by default to open an *.ldp file and searches its respective input files in the same directory. From the different output files LDP can create a variety of diagrams. In fact, the diagram displayed may vary considerably depending on the task given to the **LJUNGSKILE** program. In general, the user can select three different ordinate presentations: the concentration, the log of the concentration and the percentual distribution.

If **LJUNGSKILE** has calculated a multiple run problem, e.g. a speciation as a function of pH, then the output is a speciation diagram as function of pH. The uncertainties are displayed as 68% confidence percentiles. The concentration presentations are calculated from the *.out files saved by the **LJUNGSKILE** program in the path\results*.prj directory. Species distribution is calculated from the phrout.* files. The program recalculates the species distribution for each run from the concentration informations, evaluates the empirical distribution function and chooses the value closest to the 0.16 percentile, the median and value closest to the .84 percentile value as uncertainty limits and center, respectively. It is obvious that a small number of runs provides a more variable uncertainty value than a larger number of runs.

The diagram has two further options: The user can choose the presentation of the abscissa in logarithmic format or linearly. A logarithmic presentation does not make sense if the abscissa variable is pH. The second option allows to display the location of the points calculated in the **LJUNGSKILE** run in the shape of crosses. A displayed data set can be saved as ASCII file to be converted to a presentation graphics by a suitable graphics program (e.g. ORIGIN).

If **LJUNGSKILE** calculated a single run option, e.g. the distribution of metal ion species at a given pH, then the results are displayed as simplified box plots. The box in the center gives the median value while the usually wider open box indicates .68 percentiles. The bar line indicates the total range of the property. Note that for small number of runs it is not uncommon that some values coincide. Again the user may select different presentation types. The LDP code recognises from the *.ldp file whether the single run option has been chosen and the respective graphical presentation is displayed as shown in fig. 2. The box plot presentation can be saved. The five respective characteristics are saved as numerical values.

It should be mentioned that most presentation programs or statistical data evaluation programs do have a box plot option corresponding to Tukey's original definition of box plots. The user may prefer to import the data calculated from the **LJUNGSKILE** code into such a commercial program.

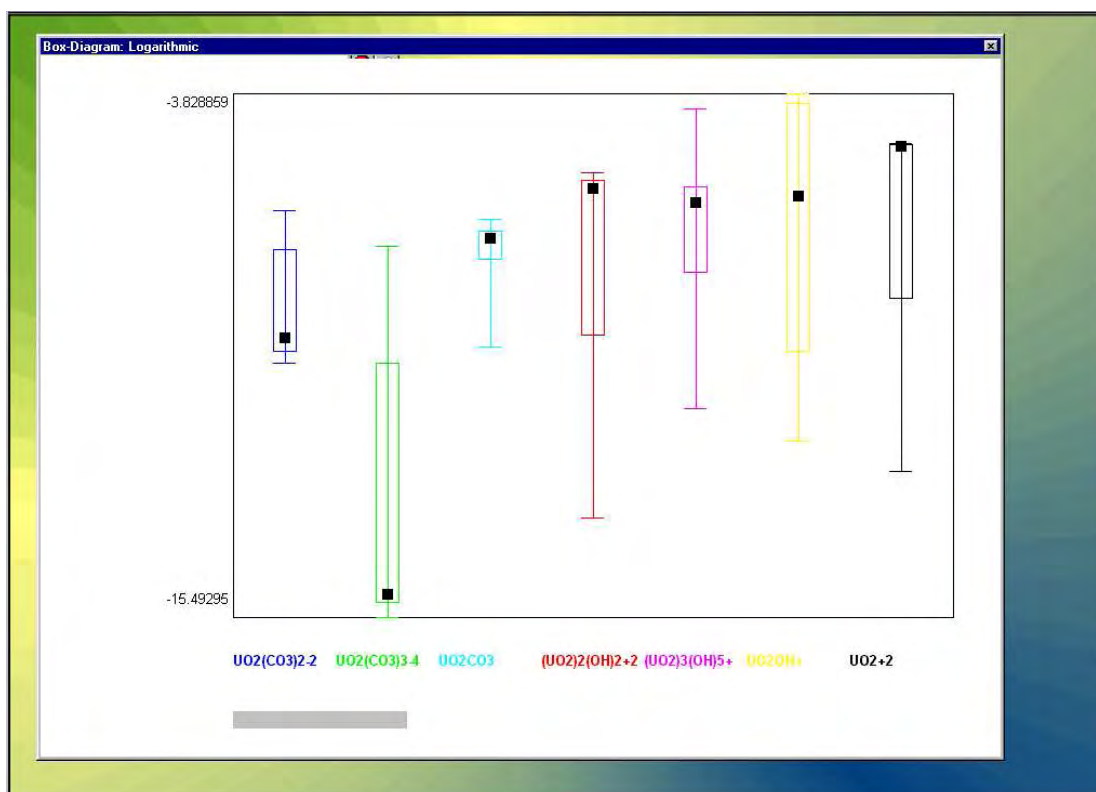


Figure 3: Box plot presentation of **LJUNGSKILE** single run results by LDP

Installation and Description of Files

Installation: LJUNGSKILE program and LJUNGSKILE Display Program (LDP)

The programs come with an installation routine. The install files are named "LDP20.cab". The install routine allows to change the installation location and the location in the Start menu.

A subdirectory \results has to be created before LJUNGSKILE is executed the first time. In the \results subdirectory, LJUNGSKILE generates a subdirectory with the name of the project file *.prj. Assuming a project file "Test.prj", the results calculated by LJUNGSKILE would be found in the directory D:\Ljungskile\results\Test.

The LJUNGSKILE Display Program is installed by an installation routine.

Description of Files: LJUNGSKILE code

The files made by or needed for the LJUNGSKILE programs are listed below.

PHREEQC.EXE	The thermodynamic equilibrium program used for the chemical calculations.
LJUNGSKILE20.EXE	Master program which runs all screen handling and user interface. This program also manipulates the database and the input file to PHREEQC according to the selections of the user
SIMULATION.EXE	Runs PHREEQC and checks its results
"project name".PRJ	Contains all information related to the selected project (created by the program).
PHROUT.#	Contains raw data from each simulation cycle (#) (created by the program)
PHROUT.AVR	Contains average concentrations of the selected species (created by the program)
PHROUT.SD	Contains the estimated standard deviation of the selected species concentrations obtained by the setting selected for this calculation (created by the program)

The following file names are either generated or used by the programs temporarily and therefore the user must not add or delete any of them manually unless clearly aware of what one is doing.

PHREEQC.DAT	(needed by the program)
PHREEQC.OUT	(created by the program)
PHREEQE.IN	(created by the program)
PHREOUT.AVR	(created by the program)
PHROUT.SD	(created by the program)

The database file "database".DAT must contain the inert elements Ip and Im as shown in APPENDIX III.

Description of Files: LJUNGSKILE Display Program (LDP)

After installing LDP, it starts by clicking the LDP20.exe file in the installation directory. From the "File" menu, the "Open" feature has to be selected and a suitable LJUNGSKILE project file (*.prj) must be selected. LDP reads the relevant information from the project file and displays the 'multiple run' calculations accordingly. Occasionally, LDP closes immediately. In this case, a file 'info.lju' can be found in the install directory which should be deleted.

Three different ordinate presentations can be selected from the "Diagram" menu: concentration, logarithmic concentration and relative distribution. In all cases, the mean values are given as solid lines while the upper and lower .68 percentile values are given as dashed lines. LDP uses 15 different colours to display the lines of different species. If more than 15 species are included into the LJUNGSKILE calculation, then some colours will be used twice. Experience shows that diagrams with more than ten relevant species already get messy, anyway.

Graphical presentation

The LDP graphical representation fits itself to the screen resolution. It is nevertheless possible to resize and to move the diagram. LDP does not create a grid to the diagram. Only minimum and maximum ordinate and abscissa values are given. But if the cursor moves over the diagram, its coordinates according to the diagram scale are given.

Diagram check boxes

The diagram offers two check boxes: "logarithmic abscissa" and "show data as crosses":

"Logarithmic abscissa": LJUNGSKILE saves the abscissa values linearly. If LJUNGSKILE's "logarithmic scale" feature is selected in the 'multiple runs' window, the value is not saved as, say, -5.5 but as 3.162278E-6. An exception is pH. Hence, checking and unchecking the box will allow to switch between the linear and the logarithmic scale.

"show data as crosses": LJUNGSKILE calculates values at fixed intervals. In generating the diagram, LDP interpolates linearly between the calculated values. By checking the box, the density of calculated points can be visualized.

Exporting diagram data

There is no "print" feature in LDP. Instead, the graphical data can be exported in ASCII format by selecting "save" from the "File" menu. The data is saved with the default extension *.psa. For each abscissa value, respective ordinate values are stored in a chain consisting of the sequence "upper, lower and median value" of each species. To allow identification, the date of the generated file is added as the final datum. The *.psa file can be imported into a suitable scientific presentation program and manipulated according to the needs of the user.

File information

To allow a quick overview on some basic data from the project file, the menu feature "Information" displays a brief list.

Getting Familiar: A Step-by-Step Example

Once the **LJUNGSKILE** code is installed, the directory where the code is installed should include the files (in the following "d:\Ljungskile\" is assumed).

Ljungskile.exe
Simulation.exe
Phreeqc.exe
Example_A1.prj
example.dat

The file example.dat holds a thermodynamic data base. This data base is part of the PHREEQC package (Parkhurst 1995) and further information can be found in the PHREEQC manual (download under <http://water.usgs.gov/software/phreeqc.html>). Advantageously the PHREEQC database can be extended and modified with comparative ease. The user is encouraged to display this database in a text processor program like Wordperfect, Word or StarWriter and study how to modify it.

Starting the ljungskile.exe file, the following window should appear.

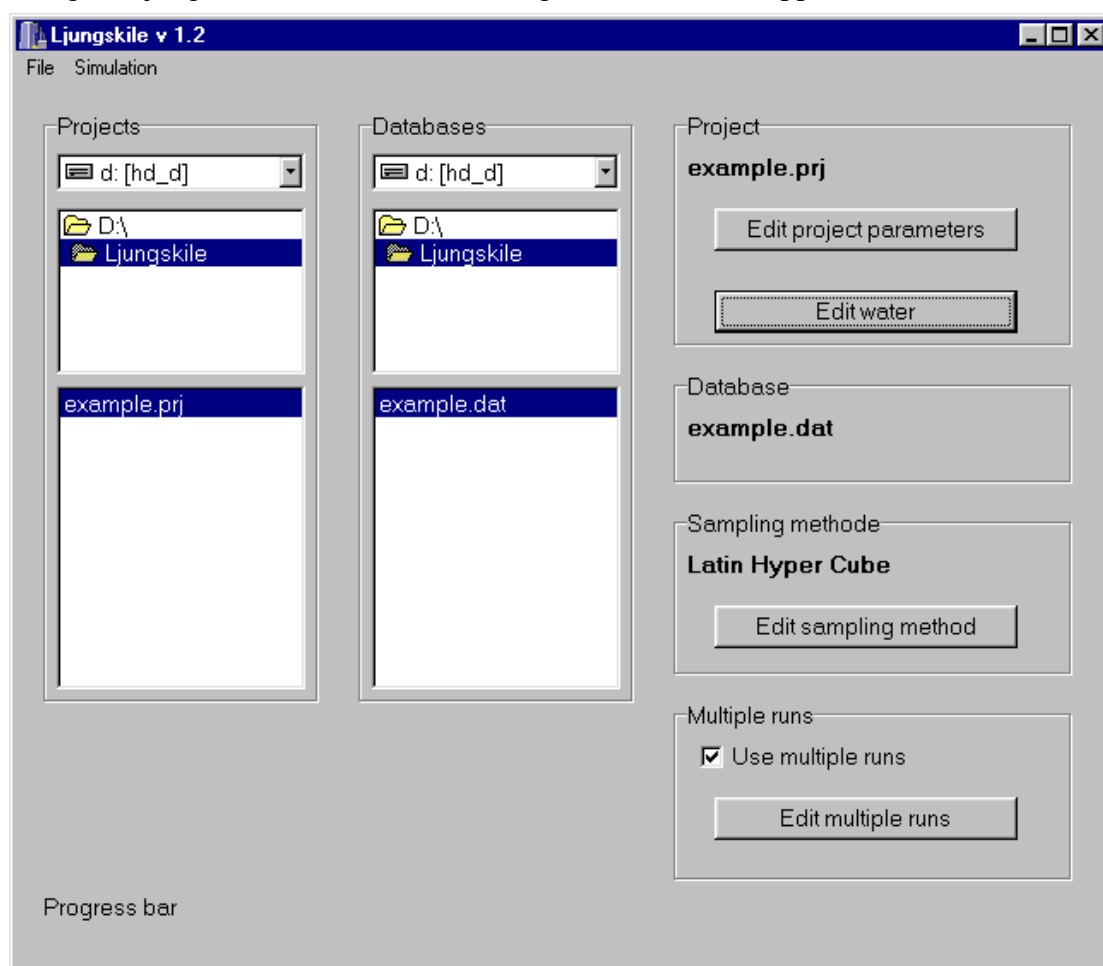


Figure 4: LJUNGSKILE main window

In order to run, the **LJUNGSKILE** code needs information: first on the data base (several data bases may be provided by the user) and second on the specific tasks summarized in a project file. Clicking on the project file `example.prj` will highlight it together with the Latin Hypercube feature in the field 'sampling method'. Clicking on "`example.dat`" in the right side selection box will make the project window "`example_Al.prj`" available for user editing. The user is advised to examine the editing and selection windows but is recommended not to make any changes at this level of the example.

☐ Solid phase Amount (mole)
☐ CO2(g) Partial pressure (bar)

Name	Mean value	SD or max value	Distribution
AlOH+2	-4.99	0.23	Normal
Al(OH)2+	-10.1	0.12	Normal
Al(OH)3	-16	0.23	Normal
Al+3	0	0	Master
Al(OH)4-	-23	0.4	Normal
AlSO4+	3.02	0.23	Normal

Close edit project parameters

Figure 5: Speciation window of **LJUNGSKILE**. The user may change the thermodynamic parameters, the uncertainties and the distribution for each species. Species can be added and deleted. The respective window pops up on clicking the right mouse button. The species name must be present in the database (case sensitive).

The first column of the project window table gives the species names. These names have to be in agreement with the PHREEQC data base conventions (see PHREEQC manual) and to be present in the selected database. A species Al(OH)_4^- can appear in the database as "`AlOH4-`", "`Al(OH)4-`" or even "`O4H4Al-`". The form given in the table must correspond to whatever form the database holds. In a similar way the formation constants have to comply with the PHREEQC sign conventions.

Depending on the uncertainty distribution of choice, the standard deviation (normal distribution) or a maximum value (uniform distribution) have to be specified. The distribution itself is specified in the last column.

It is possible to select a solid as solubility limiting phase in order to calculate mineral solubilities and to specify a CO_2 partial pressure. The name of the solid phase must correspond to a mineral name in the selected database. In the amount field a value can be entered (in units of mol L^{-1}). If no value is specified, the amount is assumed to be zero and only precipitation processes can be calculated. Otherwise, the amount specified may be dissolved and change the water composition. For the CO_2 partial pressure, an fixed value can be specified. However, handling the carbonate system correctly requires considerable experience with PHREEQC. The novice must be ready

to bear considerable frustration when starting to use this feature. Hence, these features will not be used in this example. Solid phases often change the physicochemical properties drastically. To warrant convergence in PHREEQC, the composition of the water should be close to the final composition. This can be achieved by calculating a single step run, checking the composition of the final solution in the PHREEQC.OUT file and modifying the water composition accordingly. The 'virtual' species Ip and Im will warrant charge balance otherwise. Similarly, a fixed CO₂ partial pressure requires PHREEQC to perform chemical reaction steps in order to balance the equations. These reactions may alter the solutions considerably causing non-convergence. Species can be added or deleted upon clicking the right mouse button and making the appropriate selection from the pop-up window. An added species must occur in the database and typed correctly (case sensitive). The button 'close edit project parameters' will close the project window.

An interesting feature of any speciation program is the calculation of equilibria in complex solutions. Laboratory solutions usually strive for a simple composition - the simple composition in turn allows a theoretical speciation without complex computational tools like the **LJUNGSKILE** code. However, the situation in groundwaters is usually less straightforward. The 'edit water' button in the 'project' frame allows to specify pH, pe, temperature and total concentrations of water components to be considered in the **LJUNGSKILE** speciation. The elemental composition can be modified by adding or deleting elements. Elements are added by clicking the right mouse button. Leave the water unchanged.

Groundwater properties

Description: test water

pH: 5

pe: 12

Temperature (C): 20

Name	Concentration
Al	1E-7
S	0.001

EXIT

Figure 6: Water window of **LJUNGSKILE**. The description should be chosen to allow clear identification of the file. The pH can be specified but it will be overridden if pH is chosen later as running variable in the 'Multiple run' window (see below).

The last window is the 'Method' window (Fig.7) where the upper left hand side box allows us to choose between a Monte Carlo method and the Latin Hypercube approach. For the present case, we will stick with the Latin Hypercube method. The seed for the random number generator can be given arbitrarily. The '#runs' field specifies the number of strata to be created for each variable, while '#points in the cdf' specifies the number of normally and uniformly, respectively, distributed random points used by the program in constructing a cumulative normal distribution. It is reasonable to have the '#runs' value smaller than the '#points in the cdf'. Otherwise the program issues a warning otherwise but it can be overridden.

Figure 7: Method window of LJUNGSKILE. Method for probabilistic calculation (Monte Carlo or Latin Hypercube Sampling), the starting value of the random number generator and some details of the cumulative distribution functions (cdf's) can be selected.

This example will follow the speciation over a certain pH range. Such a calculation will nicely illustrate what the LJUNGSKILE code can be used for. On checking the 'use multiple runs' box, the following window should open:

Figure 8: Multiple Runs window of LJUNGSKILE.

The running variable can be selected from the left hand side drop-down field. The start and end values of the running variable must be specified together with an interval

length. The selection field "Logarithmic scale" will be active by default as soon as a parameter other than pH or pe is selected. The user can set it inactive if he desires it for some reason. However, it would be unreasonable to vary, say, the Al total concentration between 10^{-7} M and 10^{-3} M in steps of 10^{-6} M - ending up in several thousand steps. A typical situation for inactivating this feature is the selection of temperature as variable.

In the left hand side selection box, the variable parameters are listed. The pH should be selected. The meaning of 'Start', 'Stop' and 'Interval length' explain themselves and should be filled with the values '3', '7' and '0.1'.

Now, the procedure can be started by clicking on the 'Simulation' menu item. PHREEQC runs in a DOS box. It is possible to run the calculations in the background and to use the computer for other purposes at the same time. The simulation takes some minutes to complete. The progress bar indicates the degree of completion.

Description of Output

After completion of the calculations a new subdirectory in the LJUNGSKILE installation directory has been created: \results. The project name specifies a subdirectory in \results. Hence, in the present example, the calculated information is found in D:\Ljungskile\results\example\. The top section of this subdirectory is shown in Fig. 9.

Dateiname	Größe	Typ
Al(OH)2+.out	21 KB	Datei OUT
Al(OH)3.out	21 KB	Datei OUT
Al(OH)4-.out	21 KB	Datei OUT
Al+3.out	21 KB	Datei OUT
AlOH+2.out	21 KB	Datei OUT
AlSO4+.out	21 KB	Datei OUT
example.ldp	1 KB	Datei LDP
phreeqc.out	5 KB	Datei OUT
PHROUT.0	4 KB	Datei 0
PHROUT.1	4 KB	Datei 1
PHROUT.10	4 KB	Datei 10
PHROUT.11	4 KB	Datei 11
PHROUT.12	4 KB	Datei 12
PHROUT.13	4 KB	Datei 13
PHROUT.14	4 KB	Datei 14
PHROUT.15	4 KB	Datei 15
PHROUT.16	4 KB	Datei 16
PHROUT.17	4 KB	Datei 17
PHROUT.18	4 KB	Datei 18
PHROUT.19	4 KB	Datei 19
PHROUT.2	4 KB	Datei 2
PHROUT.20	4 KB	Datei 20
PHROUT.21	4 KB	Datei 21
PHROUT.22	4 KB	Datei 22
PHROUT.23	4 KB	Datei 23
PHROUT.24	4 KB	Datei 24
PHROUT.25	4 KB	Datei 25
PHROUT.26	4 KB	Datei 26

Figure 9: Example of contents of the directory D:\Ljungskile\results\example

There are 41 Phrout.xx (xx= 1 to 41) files, one Phrout.avr and one Phrout.sd file. For each of the six species X.out files are available (X = $\text{Al}(\text{OH})_4^-$, Al^{+3} , AlSO_4^+ , $\text{Al}(\text{OH})_2^+$, $\text{Al}(\text{OH})_3$, AlOH^{+2}). These files can be inspected with a word processor and it is recommended to do so for illustrative purposes.

In the present situation, the files Phrout.avr and Phrout.sd are of major interest because they hold the mean values (Phrout.avr) and the standard deviations (Phrout.sd) of the calculated species concentrations:

The Phrout.avr file for the Al speciation example

```
pH 3.00000e+00 3.10000e+00 3.20000e+00 3.30000e+00 3.40000e+00 3.50000e+00 3.60000e+00
3.70000e+00 3.80000e+00 3.90000e+00 4.00000e+00 4.10000e+00 4.20000e+00 4.30000e+00
4.40000e+00 4.50000e+00 4.60000e+00 4.70000e+00 4.80000e+00 4.90000e+00 5.00000e+00
5.10000e+00 5.20000e+00 5.30000e+00 5.40000e+00 5.50000e+00 5.60000e+00 5.70000e+00
5.80000e+00 5.90000e+00 6.00000e+00 6.10000e+00 6.20000e+00 6.30000e+00 6.40000e+00
6.50000e+00 6.60000e+00 6.70000e+00 6.80000e+00 6.90000e+00 7.00000e+00
AlOH+2 1.37451e-09 1.55039e-09 1.75745e-09 1.99996e-09 2.28252e-09 2.60777e-09 2.97640e-09
3.38607e-09 3.83251e-09 4.30682e-09 4.79862e-09 5.29357e-09 5.77717e-09 6.23552e-09 6.65509e-
09 7.02376e-09 7.33232e-09 7.57352e-09 7.74210e-09 7.83411e-09 7.84555e-09 7.77333e-09
7.61357e-09 7.36324e-09 7.02051e-09 6.58635e-09 6.06730e-09 5.47719e-09 4.83857e-09 4.18147e-
09 3.53948e-09 2.94454e-09 2.41993e-09 1.97790e-09 1.61884e-09 1.33368e-09 1.10761e-09
9.23557e-10 7.65781e-10 6.22782e-10 4.89240e-10
Al(OH)2+ 6.81271e-13 1.09048e-12 1.73253e-12 2.73038e-12 4.26559e-12 6.60065e-12 1.01082e-11
1.53089e-11 2.29050e-11 3.38504e-11 4.93769e-11 7.10836e-11 1.00985e-10 1.41683e-10 1.96382e-
10 2.68909e-10 3.63663e-10 4.85250e-10 6.37578e-10 8.22895e-10 1.04086e-09 1.28900e-09
1.56331e-09 1.85882e-09 2.16903e-09 2.48293e-09 2.78453e-09 3.05224e-09 3.26141e-09 3.38853e-
09 3.41744e-09 3.34310e-09 3.17329e-09 2.92613e-09 2.62620e-09 2.29953e-09 1.96929e-09
1.65369e-09 1.36470e-09 1.10882e-09 8.88210e-10
Al(OH)3 5.98660e-16 1.21069e-15 2.42869e-15 4.83085e-15 9.51976e-15 1.85708e-14 3.58336e-14
6.83294e-14 1.28670e-13 2.39092e-13 4.38178e-13 7.91818e-13 1.41083e-12 2.47971e-12 4.30151e-
12 7.36521e-12 1.24501e-11 2.07690e-11 3.41612e-11 5.53159e-11 8.80574e-11 1.37667e-10
```

The .avr and .sd files provided by LDP can be used for getting a quick survey on the results using a short program in the style given in Appendix IV. The generated ASCII files may subsequently be imported into a presentation graphic program like SigmaPlot or ORIGIN.

The LJUNGSKILE Display Program (LDP)

A more convenient way to display the output of a LJUNGSKILE run is provided by a supplement, the LJUNGSKILE Display Program (LDP). Start LDP and select the example.prj file by selecting 'Open' from the file menu.

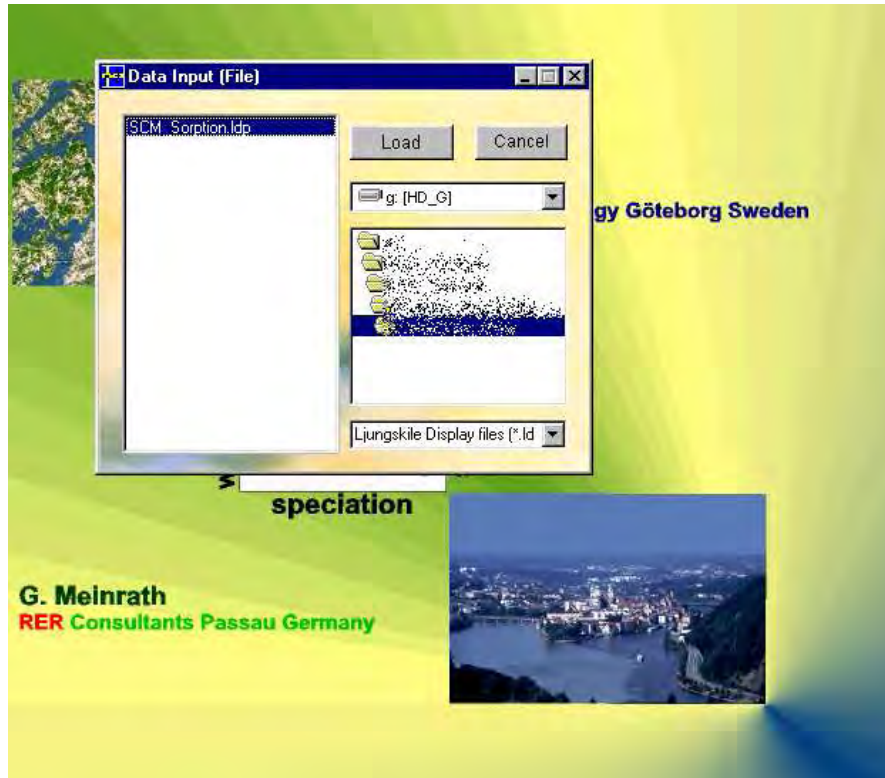


Figure 10: select 'example.ldp' from the respective \results\ subdirectory

Figure 10 shows the respective screen. In the diagram constructed by LDP (Fig. 11) the names of the species are listed and the colour of the species name corresponds to the respective colour of the calculated curve. The initial diagram is the log(concentration) display. The upper and lower .68 percentiles are shown as dashed lines. The diagram may be resized and moved around. The LDP filename are found in the diagram header and is also available from the Information menu item.

The x and y coordinates under the cursor are numerically given in two fields below the diagram. Click on the respective field to make LDP show the actually measured points.

Select 'distribution' from the 'Diagram' menu and see the changes. Select 'Save' from the file menu and save the data as ASCII files under the extension *.psa for postprocessing by a graphics program.

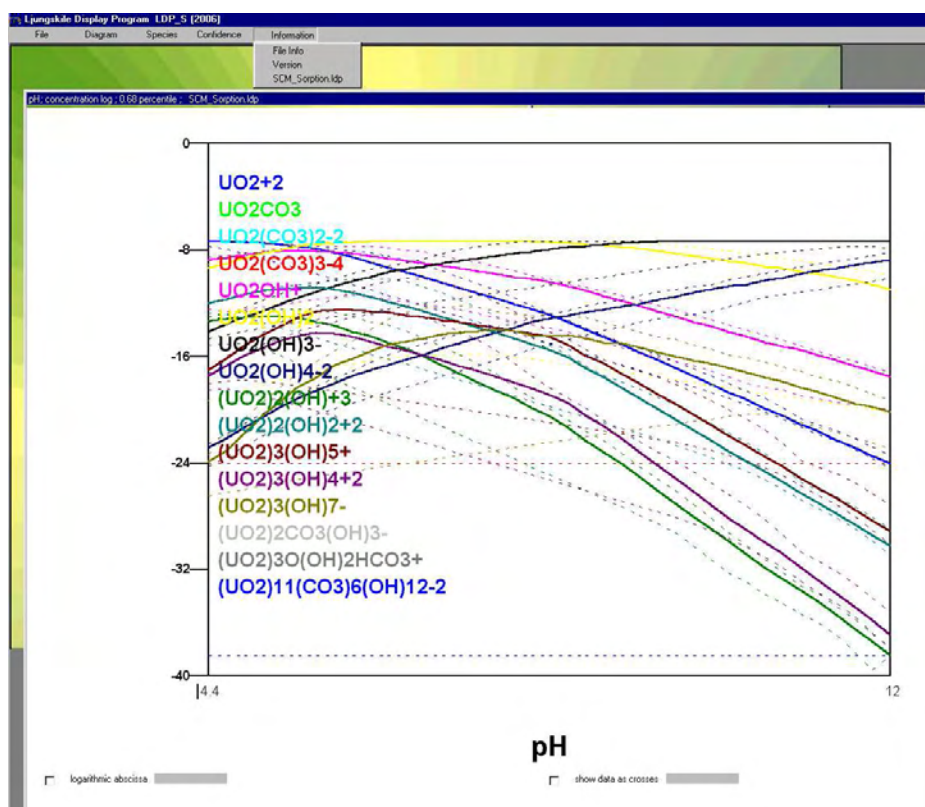


Figure 11: Species concentrations as a function of pH calculated for uranium speciation of the project file SCM_Sorption.prj.

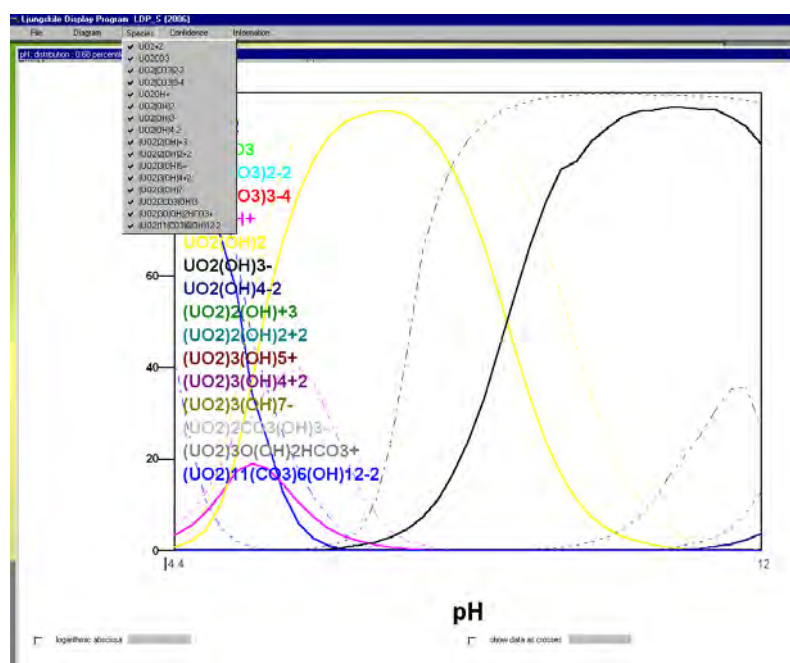


Figure 12: The species distribution of uranium species. The open menu item lists the available species. If a species is unchecked, its contribution is immediately removed from the diagram. The confidence limit may be changed from the 'Confidence' menu item.

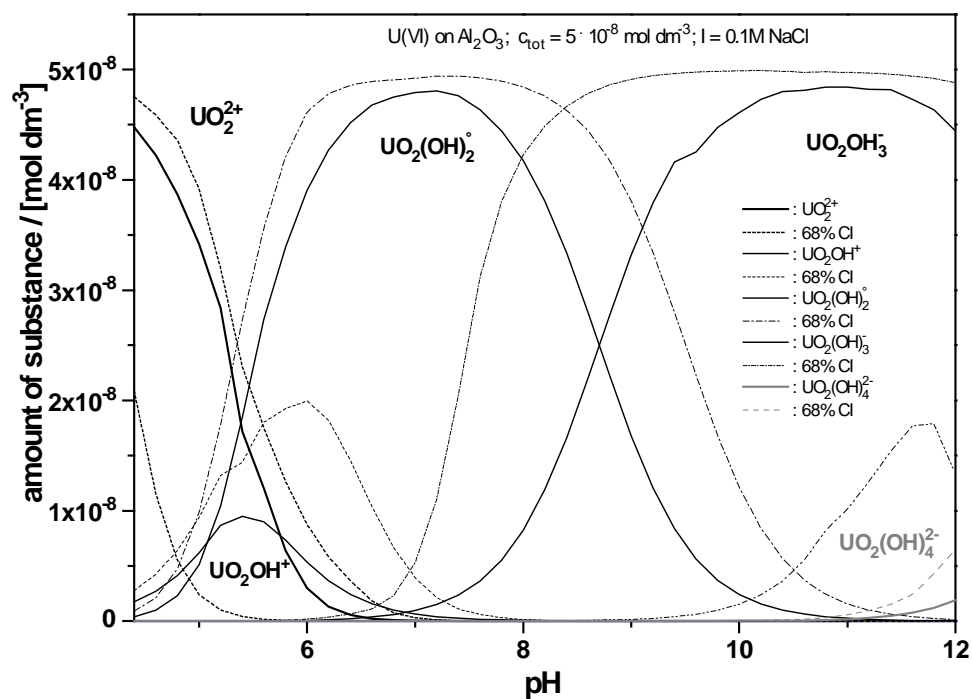


Figure 13: The 'uranium' speciation results in an ORIGIN presentation format

Relevant information on the LJUNGSKILE caculation giving rise to the respective diagram are found under the 'Information/File Info' menu item.

References

<http://www.borland.de>

Box G.E.P, Muller M.E. (1958) A Note on the Generation of Random Normal Deviates. *Ann. Math. Stat* 29: 610 - 611

Ekberg C., Emrén A.T. (1993) Finding and Correcting Calculation Errors in PHREEQE. *Proc. Int. Conf. Migration 93* Charleston/USA. Oldenbourg Verlag München/FRG.

Ekberg C. (1999) Uncertainty in Actinide Solubility Calculations Illustrated Using the Th-OH-PO₄ System. Thesis. Chalmers University of Technology.

Ekberg C., Emrén A.T. (2001) Uncertainties in Solubility Calculations. *Monatshefte Chemie* 132: 1171

Ekberg C. (2002) Uncertainties Connected with Geochemical Modelling of Waste Disposal in Mines. *Mine Water Environ* 21: 45 - 51

Ellison S., Wegscheider W. and Williams A. (1997) Measurement Uncertainty. *Anal Chem* 69: 607A - 613A

Knuth D.E. (1981) *The Art of Computer Programming* 2nd ed.. Vol II: Seminumerical Algorithms. Addison-Wesley, Reading/USA

McKay M.D., Conover W.J., Beckman R.J. (1979) A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output of a Computer Code. *Technometrics* 21: 239 - 245

Meinrath G., Ekberg C., Landgren A., Liljenzin JO (2000) Assessment of Uncertainty in Parameter Evaluation and Prediction. *Talanta* 51: 231 - 246

Meinrath G (2000a) Comparability of Thermodynamic Data - A Metrological Point of View. *Fresenius J Anal Chem* 368: 574 - 584

Meinrath G, Hurst S, Gatzweiler R (2000b) Aggravation of Licensing Procedures by Doubtful Thermodynamic Data. *Fresenius J Anal Chem* 368: 561 - 566

Meinrath G (2001) Measurement Uncertainty of Thermodynamic Data. *Fresenius J Anal Chem* 369: 690 - 697

Meinrath G, May PM (2002) Thermodynamic Prediction in the Mine Water Environment. *Mine Water Environ.* 21: 24 - 35

Parkhurst D.L. (1995) *Users Guide to PHREEQC - A Computer Program for Speciation, Reaction-Path, Advective-Transport, and Inverse Geochemical Calculations.* Water-Resources Investigations Report 95-4227. U.S. Geological Survey Lakewood, Colorado/USA.

Sillén L.G. (1962) Acta Chem Scand 16: 159

Sillén L.G. (1964) High-Speed Computers as a Supplement to Graphical Methods. III. Twist-Matrix Methods for Minimizing the Error-square Sum in Problems with Many Unknown Constants. Acta Chem Scand 18: 1085 - 1098

Wolery T.J. (1992) EQ3/6, A Software Package for Geochemical Modeling of Aqueous Solutions: Package Overview and Installation Guide (Version 7.0). REPORT UCRL-MA-110662 PT 1, LLNL, Livermore CA/USA

APPENDIX I

Random Numbers

Computer-generated random numbers are not truly random because they result from deterministic process. Theoretically and despite the seemingly perfect randomness of computer-generated random numbers, the deterministic law behind a computer-generated sequence of 'random numbers' can be resolved provided enough resources will be invested. However, by combining two random processes such efforts will be in vain even for very long sequences of random numbers and thus the differences between truly random processes and computer-simulated random processes becomes academic. The following procedures describe the random number generators used in the present implementation of **LJUNGSKILE** code. The algorithms are well known in literature on algorithms and are mostly summarized by Knuth (1981). Additional sources will be occasionally given below in order to direct the reader to further informations of interest.

Linear random number generator

A linear random generator returns a number from a specified range (commonly 0 and 1) where one number is equally likely to occur. The linear random number generator thus returns uniform random deviates (to be distinguished from the normally distributed random deviates discussed later). The algorithm of a linear random generator calculates an integer number I_{k+1} from a previous number I_k by

$$I_{k+1} = a I_k + b \pmod{c} \quad (1)$$

where I gives a random number, k the running index, a the multipliers, b the increment and c the modulus. The numbers a, b , and c are positive integer numbers. The modulus c gives the length of the chain after which the process (1) repeats itself. Therefore, the larger c the longer the chain. A random number between 0 and 1 is obtained by deviding I_{k+1}/c . The recurrence formula (1) requires a starting value, also termed 'seed'. The seed is allowed to be user-specified in **LJUNGSKILE** code.

Since the random number generator of type (1) may have many deficiencies (i.e. serial correlation and bad selection of the value for a, b and c), additional shuffling is done: the random number generator is used to fill an array of, say, 100 figures with random numbers. After this preparation, a random number is chosen based on a first random number generated (converted into an index of the array) and is replaced by a second random number. By coupling two suspicious processes of 'insufficient randomness', an almost perfect random number can be generated.

Normally distributed random deviates

Normally distributed random deviates with a mean of 0 and a standard deviation of 1 can be efficiently obtained by an algorithm of Box and Muller (1958) from a linear random generator.

On basis of two values LRG1 and LRG2

$$v_1 = 2 * \text{LRG1} - 1$$

$$v_2 = 2 * \text{LRG2} - 1$$

and the acceptance condition

$$q = v_1^2 + v_2^2$$

$$0 < q < 1$$

the normal deviates NRD are calculated by

$$f = \sqrt{\frac{-2 \cdot \ln(q)}{q}}$$

$$\text{NRD}_1 = f \cdot v_1$$

$$\text{NRD}_2 = f \cdot v_2$$

A cumulative normal distribution can be approximated by generating a larger series of N normally distributed random deviates, sorting them with increasing magnitude and putting a weight of 1/N on each value. Intermediate values of the thus generated cumulative normal distribution can be obtained by interpolation (Meinrath, Ekberg et al., 2000).

There are, admittedly, alternative and probably computationally more efficient ways in obtaining a normally distributed random deviate with defined uncertainty by using Chebyshev approximations of a scaled cumulative Gaussian distribution. In the present case, however, the computational burden imposed by the method selected for implementation in the **LJUNGSKILE** code is almost irrelevant compared to the CPU time required for the **PHREEQC** speciation calculations that further optimizations seem irrelevant under the overall aspect.

APPENDIX II

Latin Hypercube Sampling

Computer models are in common instruments for simulation of complex models in many areas of industry and science. Models allow to investigate intricate relationships on a broad level. Models summarize information obtained from data and may forward knowledge of highest complexity. Since each input variable is affected by uncertainty, the predictive power of a computer model is influenced by the evolvement of these uncertainties during a simulation run.

It is always possible to choose a Monte Carlo design for a study of the influence of input variable uncertainties on simulation output. There are, however, many input variables in even a simple theoretical speciation calculation and the Monte Carlo study consequently must involve a large number of runs. It must be ensured that for each input variable the probability distribution is sufficiently represented in the Monte Carlo runs. This representation is almost warranted for input values close to the mean value. But with larger distance from the mean value, inclusion becomes less and less likely. Consequently, the more uncertain input variables are to be considered, the larger the number of Monte Carlo runs has to be.

There is no need to delve into the question how to estimate the necessary number of runs conditional on the number of input variables. There are more efficient strategies. An especially clever and simple strategy has been suggested by McKay, Conover and Beckman (1979).

Latin Hypercube sampling (LHS) requires a division of the cumulative probability distribution of each of n input variables to be divided into n bins of equal probability $1/n$. Such a case is shown in Fig.1 for a cumulative normal distribution of mean 13.23 and $\pm .68$ percentiles of 0.2. The distribution is divided into 16 sections of equal probability $1/16$. The bottom arrows indicate a randomly chosen value from within the interval to represent the interval in the LHS procedure. In generating the intervals of equal probability, an arbitrary choice has to be made with respect to the position of the lower and upper boundary that is set in the present example to four times a .68 percentile. The probability of an event with a value larger than 14.03 is below 0.00001 and can safely be neglected in a LHS design on basis of 16 events. For very large designs it may be considered to set the lower/upper limit to ± 4 .68 percentiles but the likely variation most probably will be almost probably negligible.

Figure 2 shows the stratification of a single uncertainty affected input variable. Such a stratification, however, is necessary for all uncertainty affected input variables. The

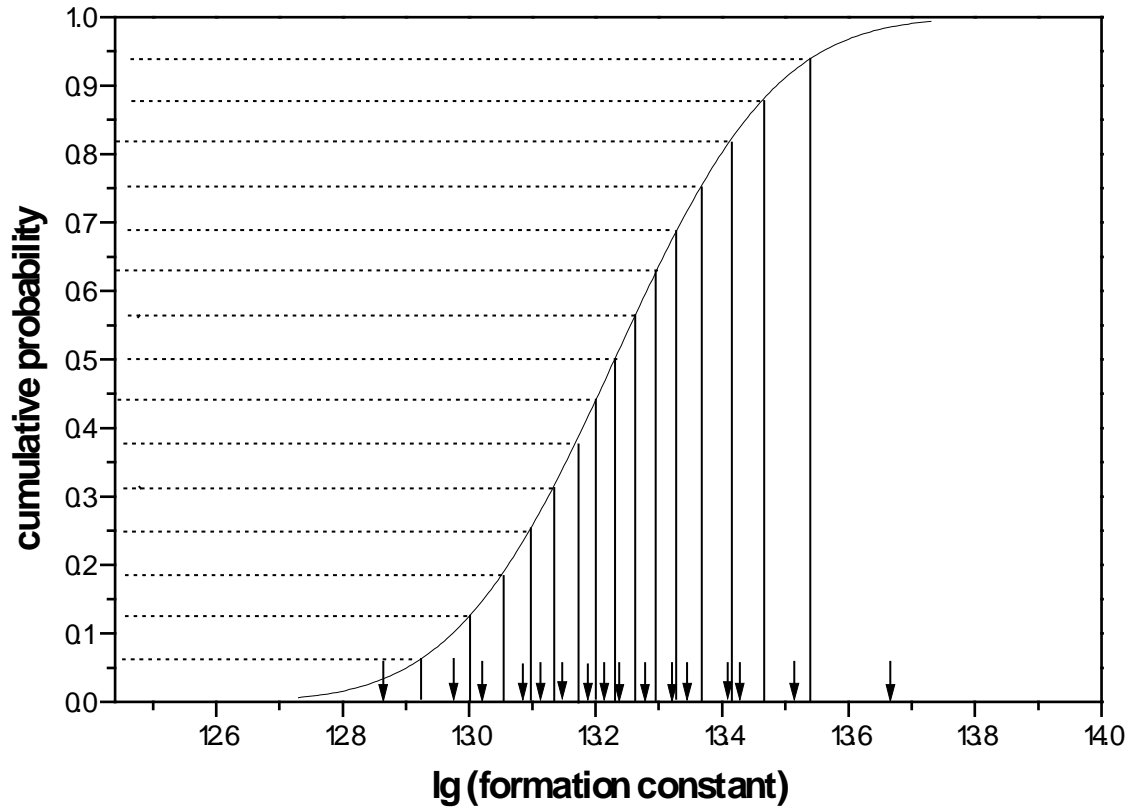


Figure A2: A cumulative normal distribution stratified into 16 strata of equal probability. The bottom arrows indicate the randomly drawn representatives for each stratum.

value of 16 strata was selected arbitrarily in the case given in fig. 2. In an LHS design, the number of strata is given by the total number of uncertainty affected input variables to be included into a simulation run. Thus, for each of the n input variables n values are selected that represent all parts of a distribution with the adequate probability. Hence, the risk to either overrepresent or underrepresent certain sections of a distribution is greatly reduced.

From the $n \times n$ input variable (representatives of n strata for n variables), input vectors into the computer simulation have to be constructed. An obvious choice is to randomly chose one representative for each variable, to run the simulation and the repeat the random selection until all representatives have been exhausted in n simulations. But even so an obvious option, random sampling is not optimal.

There is a more clever way to construct input vectors. This method is shown in fig. 3. Please note that in each row only one representative is to be found. It is not possible, for example, to have several representatives of the stratum 3 in one input vector while other strata are not considered in a certain input vector. Such a vector may occur in random sampling but not in Latin Hypercube sampling. It is necessary to generate the remaining 15 input vectors in accordance with the same construction principle.

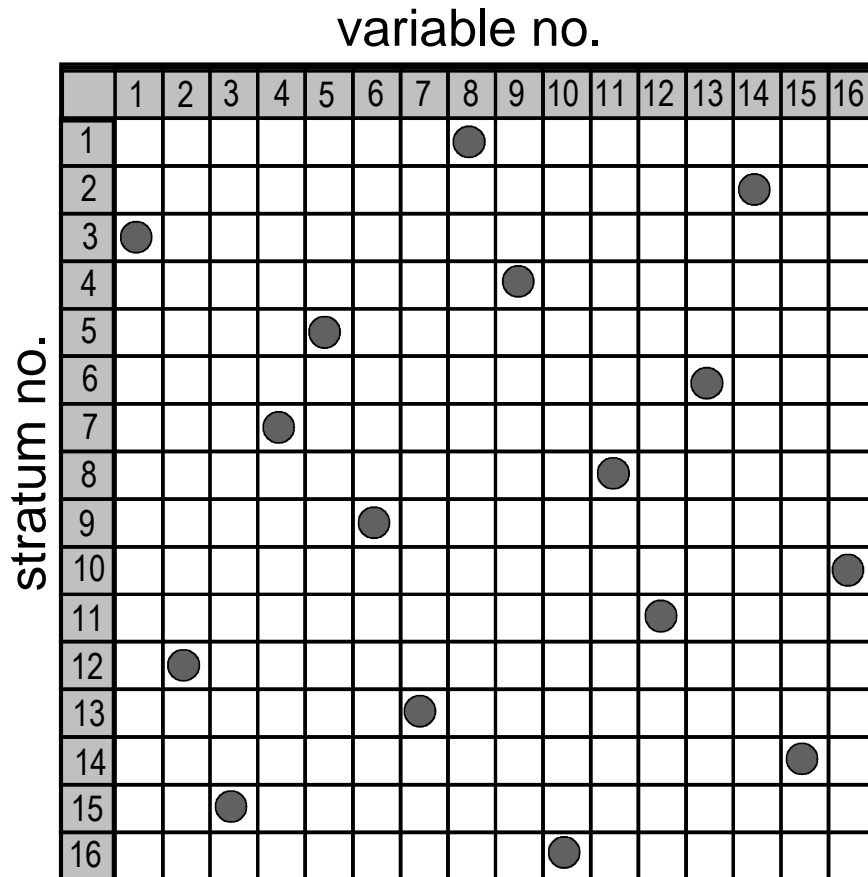


Figure A3: Example of one of 16 input vectors created from 16 strata of 16 input variables according to the LHS strategy.

Form the n runs, n realisations of the simulation output are obtained, from which mean values and standard deviations may be estimated according to the well known statistical formula on basis of a normal distribution assumption.

In case of larger samples, an alternative is to generate the empirical probability densities and to evaluate median and upper and lower percentiles (Meinrath, Ekberg et al. 2000). The **LJUNGSKILE** code gives the simulation output in a file PHROUT.AVR.

APPENDIX III

The Virtual Species Im and Ip

In many cases in geochemical modelling it may be desired to keep the water parameters pH and pe constant during a calculation. One such example is the solubility of some mineral at a given pH and pe. Without precautions, dissolution and/or precipitation may change the solution composition that drastically, that a fixed value for pH and/or pe cannot be reached by the PHREEQC code. The solution parameters may change due to the dissolution or the precipitation and thus the solution obtained is not any longer similar to the starting solution.

PHREEQC is designed to solve a precipitation/dissolution reaction and to print out composition of the equilibrium solution. Thus PHREEQC is capable of describing the process of putting a piece of mineral in a certain water and then calculate what has happened. But in the **LJUNGSKILE** program's "multiple run" feature, the goal is a different one. We would like to know what the solubility would be in a water at a given pH and pe. To facilitate this the inert elements Ip (inert positive) and Im (inert negative) have been introduced into the PHREEQC database. Im and Ip enter the calculation of ionic strength, but do not react, e.g. by forming complexes. The entry in the PHREEQC database is shown below:

a) add in the database head:

```
#29 Im
Im = Im
      logk      0.000000
      delta_h    0.000000 kcal

#30 Ip
Ip = Ip
      logk      0.000000
      delta_h    0.000000 kcal
```

and in the species definition section:

```
#300 Im-
Im + e- = Im-
      logk      20.000000
      delta_h    0.000000 kcal

#301 Ip+
Ip = Ip+ + e-
      logk      20.000000
      delta_h    0.000000 kcal
```

It should be noted that the numbers assigned here does not have to be the ones used in any given situation. With this modification it is possible with the aid of the keyword CHARGE to fulfil the charge balance condition, and thus keeping the pH or pe constant during dissolution or precipitation by addition of either Ip or Im. The same features could, naturally have been accomplished by addition of some other element but then the complexing behaviour of the solution would have changed.

The case of redox instabilities it has been described elsewhere (Ekberg 1993) for the older FORTRAN PHREEQE but the adoption to PHREEQC format is straight forward.

APPENDIX IV

A Rough and Dirty Way to Use LJUNGSKILE Output

The Ljungskile output can also be used without the Ljungskile Display Program. For that purpose, its needs to be extracted from the respective output files. An example is given in the following. The user has to specify the necessary information, e.g. the number of steps and the number of LHS samples used in the calculations. Depending on the program used to cast the numerical information into a graphical form, some rearrangement of the data may be necessary. The final graphical result is shown in fig. A4.

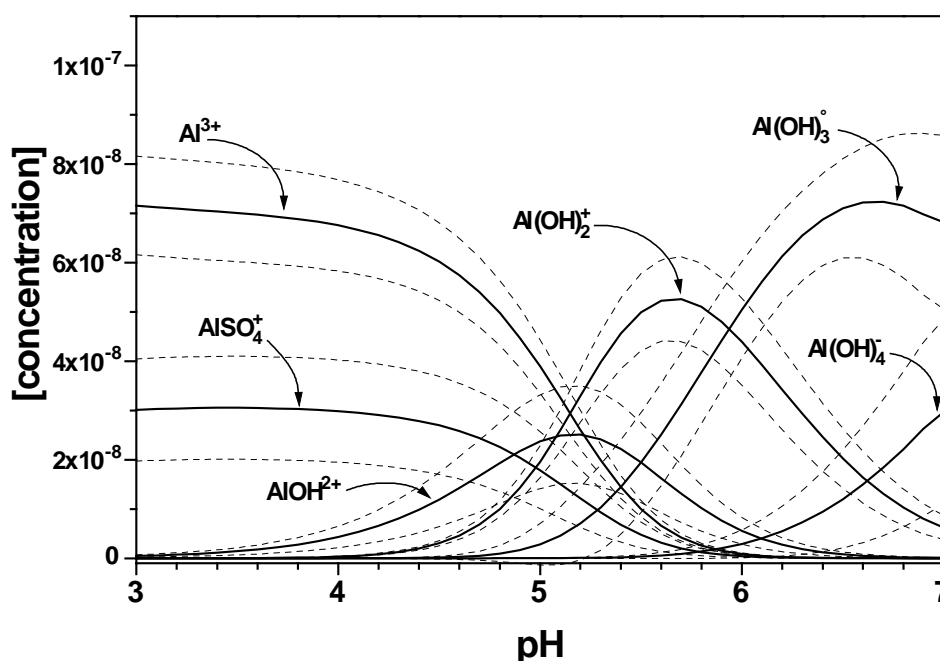


Figure A4: Distribution of Al(III) species in an aqueous solution of low mineralisation as a function of pH. Solid lines give mean concentrations while the dashed lines gives 1σ upper and lower standard deviations. Note the negative concentrations for the $Al(OH)_2^+$ species. LDP is using the empirical distributions and will not result in negative concentrations. (note: the Al.prj file is part of the Ljungskile installation).

It is also possible to calculate the information given in files Phrout.avr and Phrout.sd directly from the results in each of the files X.out. Depending on the post processing purpose this way may be even more convenient (the computer does the number crunching e.g. via a short QBasic program).

The code below opens the file X.out (the species specific part is given by the variable d\$), evaluates the mean and the standard deviation and writes the lower boundary (mean-sd), the mean and the upper boundary (mean+sd) to a X.dat file. Please note that the names of the PHREEQC output files cannot be read directly by QBasic and the non-ANSI characters (generally '+' characters in file names) must be eliminated.

Note that the upper and lower confidence limits are symmetric with respect to the mean value curves. This behaviour is caused by the simplified calculation in the Qbasic code based on the assumption of identically and independently distributed residuals with zero mean. The calculation of the calculated data however shows that in most cases the distributions are skewed. In the simplified form it may even be observed that lower uncertainties are below zero. LDP analyzes the distributions and evaluates approximations to the .68 percentiles about the median. These percentiles take skewness into account and present asymmetric upper and lower confidence bands.

```

***** Ljungskile Evaluation *****
SCREEN 12
CLS
d$ = "also4": REM "Al3", "AlOH2", "Al(OH)2+", "Al(OH)3", "Al(OH)4"
dname$ = "d:\eigene~1\" + d$ + ".out"
dout$ = "d:\eigene~1\" + d$ + ".dat"
OPEN dname$ FOR INPUT AS #1
OPEN dout$ FOR OUTPUT AS #2
test$ = "element species": REM input of text headers in the .out files
WHILE a$ <> test$
    LINE INPUT #1, a$
WEND
LINE INPUT #1, a$
i = 0
a = 0
INPUT #1, pH
WHILE NOT EOF(1)
    b = a
    INPUT #1, a
    IF a < 3 OR a > 7 THEN
        i = i + 1
        sum1 = sum1 + a
        sum2 = sum2 + a * a
    ELSE
        IF a >= 3 AND a <= 7 THEN
            sum1 = sum1 - b
            sum2 = sum2 - (b * b)
            i = i - 1
            sum1 = sum1 / i
            sum2 = sum2 / i
            sd = SQR(ABS(sum2 - (sum1 * sum1)))
            WRITE #2, pH, sum1, sum1 - sd, sum1 + sd
            i = 0
            pH = a
            sum1 = 0
            sum2 = 0
            sd = 0
        END IF
    END IF
WEND
sum1 = sum1 - b
sum2 = sum2 - (b * b)
i = i - 1
sum1 = sum1 / i
sum2 = sum2 / i
sd = SQR(ABS(sum2 - (sum1 * sum1)))
WRITE #2, pH, sum1, sum1 - sd, sum1 + sd
CLOSE #1
CLOSE #2
END
*****

```

Threshold Bootstrap Computer-assisted Target Factor Analysis (TBCAT)

version S

A self-modelling tool for resolution of convoluted spectral information

by
Günther Meinrath

copyright by RER Consultants Passau/Germany

Read the licence agreement in the manual shipped with this code!

Threshold Bootstrap Computer-assisted Target Factor Analysis
- Concepts for complex situations of metrology in chemistry: a model implementation -

temperature
cuvette cells
free OH- concentration
total concentration of sulfate
spectrometer settings
total concentration of UOVI

signal noise
- uncertainty in λ (nm)
- uncertainty in λ (Å)
- spectral correlation
- residual correlation
- parameter correlation
- non-normality
- non-linearity
- statistical optimization criterion
- weighting
- Monte Carlo effects

RER Consultants
© Günther Meinrath
Passau, Freiberg & Toronto

Version TBCAT_S (2006) Resolution 1024 x 768 pixel

Only for demonstration purpose. This program is not freeware. No warranty whatsoever is given for this code.

last addition: 10 March 2007

Table of Contents

License	3
Citation	3
Disclaimer of Warranty	4
Introduction	5
CAT Capabilities	7
1 Theoretical Basis	8
2 Statistical Basis	9
3 Computer-assisted Target Factor Analysis	15
4 Performing a CAT Analysis	18
5 Installation	19
6 Running CAT	20
<i>Data File Format Convention</i>	20
<i>Sequence of Spectra in the Input Procedure</i>	21
7 Basic Procedures	22
a) File	23
b) Do	26
c) Uncertainty	42
d) Data	42
e) Evaluate	44
f) About	46
g) Window	46
Example results	48
References	49

License

Please read the terms of this agreement and any provided supplemental license terms (collectively 'agreement') carefully before installing TBCAT_S. By installing the software, the user agrees with the terms of this agreement. If the user is accessing the software electronically, he complies with this agreement by installing the software. If the user does not agree to all these terms, he must promptly deinstall this software and return the unused software or, if the software is accessed electronically, remove the software from the computer(s).

1. The copyright owner grants the user a non-exclusive and non-transferable right for the internal use only of the accompanying software and documentation. It is understood that the accompanying software is experimental – no warranty is included that the software may work on a certain computer and is fit-for-purpose (whatever the purpose might be).
2. The software is confidential and copyrighted. It is distributed free of charge but is not in the public domain. Title to software and all associated intellectual property rights is retained by the copyright owner. Any modification, especially by reversed engineering and decompilation, is forbidden except if explicitly authorized by the copyright owner. The user may not make copies except for personal archival purposes. No right, title or interest in or to any trademark, service mark, logo or trade name what so ever is granted under this agreement.
3. This software may be used, explicitly, for education purposes in an academic environment if it deems suitable. All necessary operations (except those requiring copyright infringement) may be performed with the software (e.g. installation of the software in a computer pool). There is any implicate or explicite claim that the software may be suitable for such a purpose. The sole responsibility remains with the user.
4. This agreement is effective until terminated. If any provision of this agreement is held to be unenforceable, this agreement will remain in effect with the provisions omitted, unless omission would frustrate the intent of the parties, in which this agreement will terminate immediately.
5. Source code is only available in the framework of a collaboration.

Citation

All reports of results obtained by application of the software, including illustrations, must be accompanied by the following citations:

G. Meinrath, S. Lis "Quantitative resolution of spectroscopic systems using computer-assisted target factor analysis (CAT)", Fresenius Journal of Analytical Chemistry 369 (2001) 124 - 133

G. Meinrath, P.Schneider "Quality Assurance in Chemistry and Environmental Science: Metrology from pH Measurement to Nuclear Waste Disposal. Springer Verlag Heidelberg (2007)

Disclaimer of Warranty

Unless specified in this agreement, all expressed or implied conditions, representations and warranties, including any implied warranty of merchantability, fitness for a particular purpose or non-infringement are disclaimed, except to the extent that these disclaimers are held to be legally invalid.

To the extent not prohibited by the law, in no event will RER Consultants or the author or the programmer of the code be liable for any lost revenue, profit or data, or for special, indirect, consequential, incidental or punitive damages, however caused regardless of the theory of liability, arising out of or related to the use of or inability to use this software, even if RER Consultants has been advised of the possibility of such damages. In no event will RER Consultants' liability to the user, whether in contract, tort (including negligence), or otherwise exceed the amount paid by the user for Software under this agreement. The foregoing limitations will apply even if the above stated warranty fails of its essential purpose. This agreement is effective until terminated. The user may terminate this agreement by destroying all copies of software. This agreement will terminate immediately without notice from RER Consultants if the user fails to comply with any provision of this agreement. Upon termination, the user must destroy all copies of software.

Introduction

The computer code 'Computer-assisted Target Factor Analysis' (CAT) has been implemented with the intention to provide a model-free tool for extraction of information from spectral data. The theoretical basis of CAT is factor analysis and its algorithmic backbone is the singular value decomposition algorithm.

There are ample programs described in literature capable to decompose spectral matrices into the single components and their respective concentrations. Nevertheless, CAT is unique in several aspects:

- ◆ CAT resolves the spectra on basis of a few informations provided by the user. The resolving procedure is otherwise automatic. Hence, the user input decides upon CAT's performance. Therefore, the attribute '-assisted' has been chosen as an indication that CAT isn't a completely automatic procedure. The attribute '-assisted' emphasizes that CAT is not a black box. The user needs to understand profoundly what CAT is doing, how CAT is doing its work and which options can be manipulated in what manner.
- ◆ CAT allows to repeat the resolution procedure unassisted if the user has made her/his choices. Hence, as long as the input data doesn't vary too strongly the same set of user's parameters can be used to repeat spectral resolution without further need for interference. To the best knowledge of the author there is no other code available with this feature.
- ◆ Due to the capability to rerun slightly modified data sets without further need for user guidance, CAT is ideally suited to run in a bootstrap environment. Bootstrapping is a computer-intensive resampling technique to assess statistical properties of complex data sets without the need for complicated and often unavailable parametric statistical analysis. Hence, bootstrapping allows to extract location and dispersion estimates from data sets. Bootstrap methods can be applied to all kind of data sets provided the enormous amount of data generated from these methods. Fast modern computers and cheap digital storage render bootstrap methods to the ideal work horses for chemical data.

These three features are the key reason for the existence of CAT. But there is a fourth one, which at the time of writing these phrases is probably not aware to most chemists: CAT can be implemented into a code evaluating the complete measurement uncertainty of a spectroscopic measurement process. Thus, it is a very first implementation of a metrologically sound treatment of measurement uncertainty in complex chemical analysis. Metrology in chemistry, as metrology in general, is based on international treaties and agreements. Metrology is becoming the language of the international market place. The need to prove the quality of forwarded information will not exempt chemistry, least analytical chemistry. The need for comparable data of stated quality will change the way of analytical chemistry. CAT is an example how to realize certain aspects of the traceability chain in chemical analysis.

CAT is not a professional program. It is probably also not correct to say that CAT has been developed. CAT has been created by an evolutionary process. Users who expect the versatility and convenience of a modern commercial computer code will certainly not cheer the handling of this CAT implementation. The algorithms, however, are sound and tested. The algorithms are taken from available literature including Golub and Reinsch' Singular Value Decomposition (SVD), Box and Muller's algorithm to create normally distributed random deviates, some algorithms from the notable 'Numerical Recipes' and Nash's 'Compact Numerical Algorithms'.

The major motivation to start the implementation of CAT has been the search for a stable and easy-to-use method for deconvolution of overlapping spectra in rare-earth systems. The first tests have been made with an implementation in interpreted QuickBasic (QB). Because QB is limited to a memory of 64 KB an extended version was created in Visual Basic 5. As already mentioned before, there exist other proposals to deconvolute spectral information. But none of these proposals is simple enough to allow novices or even spectroscopists without any experience in chemometrics to handle the necessary input. It is believed that the algorithm implemented here has the required conceptional simplicity. However, since it is definitively not a black box program, the user must understand some mathematical and statistical concepts like matrices, correlation, and probability distributions. Experience with scientific programming is also recommended.

The code has been tested by applying it to problems where the approximate solution has been evaluated before. These test cases have been UV-Vis spectra of U(VI) carbonate and hydrolysis systems in aqueous solutions. U(VI) shows an extremely weak absorption in the UV-Vis range that increases continuously towards the UV region. Hence, these spectra do not provide a baseline towards the UV. Such spectra are especially challenging because baseline corrections depend solely on the spectral baseline information collected towards the IR side of the spectra. The code has been applied to evaluate a larger number of rare earth systems in aqueous and non-aqueous solutions. Some data have been published in the reviewed literature.

CAT Capabilities

CAT performs the following tasks:

- ◆ linear background correction using either data from right and left side of a spectral band or a single side selected by the user. The range of spectral information considered as background is also specified by the user. The spectra where background correction should be performed can be selected individually by the user.
- ◆ estimation of the number of species contributing to a spectral system. The selection criterion is a figure representing the sum of least squares residuals of some properties of the resolved matrices, i.e. negative absorption or species concentration values.
- ◆ automatic generation of default files with input data for the rank of matrices (= number of species) and starting values for the iterative optimisation process.
- ◆ estimation of singular values, abstract singular vectors of spectral and concentration matrices.
- ◆ estimation of single species spectra and concentrations.
- ◆ estimation of formation constants for the coordinated species based on user-supplied information about the likely stoichiometry. The concentration quotients are weighted and a weighted estimate of the formation constant is provided together with a detailed table of the concentration estimates.
- ◆ evaluation of molar absorption estimates and absorption maxima.
- ◆ user-specified weighing of residual components in the target iteration step.
- ◆ least-squares fit of experimental spectra using evaluated mean value single component spectra.
- ◆ threshold bootstrap computer-assisted target factor analysis (TB CAT) procedure with user-specified number of resampling cycles.
- ◆ generation of cumulative distributions for formation constants including a differentiation routine on basis of LOESS non-parametric regression algorithm.
- ◆ evaluation of empirical confidence bands for single component spectra for 68%, 90%, 95% and 99% confidence limits.

These features allow a complete analysis of spectral informations in a chemical system under study. In fact, in case of a TB CAT analysis the user obtains full information including confidence limits about a previously unknown chemical system studied by spectroscopy.

1 Theoretical Basics

Factor analysis decomposes a matrix of absorptions, say X , into two matrices E and C . Matrix E consists of n columns, where each column is the single component spectrum of one of the n relevant species. Matrix C consists of n rows, where each column holds n concentrations of the respective solution species. By matrix multiplication, elements x_{ik} of data matrix X can be calculated from the elements of the matrices E and C , resp. within the validity of Beer's Law

$$x_{ik} = \sum_{j=1}^n \epsilon_{ij} \cdot c_{jk} \quad (1)$$

x_{ik} : absorption observed at wavelength i in the k -th solution

ϵ_{ij} : molar absorption of the j -th species at wavelength i

c_{jk} : concentration of the j -th species in the k -th solution

or, expressed in matrix formulation

$$X = E C \quad (2)$$

First step in the analysis is to determine the number n of factors significantly contributing to the observed spectral variance by Abstract Factor Analysis (AFA). A variety of equivalent techniques are available for AFA, e.g. Jacobi Rotation, nonlinear iterative partial least squares algorithm (NIPALS) and singular value decomposition (SVD). By these techniques, observed variances are interpreted by a set of mutually orthogonal vectors (the eigenvectors of matrix X), where each vector is chosen to extract successively as much of the data variance as possible. In this work, SVD algorithm is used for determination of eigenvectors and their eigenvalues λ . A real data matrix X_{rc} , of r rows and c columns with $r \geq c$, is decomposed according to Eq. 3:

$$X_{rc} = U_{rc} S_{cc} V_{cc}^T \quad (3)$$

into a unitary matrix U of the column eigenvectors of X , a unitary transposed matrix V of the row eigenvectors of X and a diagonal matrix S , composed of the roots of the eigenvalues of X and elements $s_{ij} = 0$ ($i \neq j$). SVD extracts the roots of eigenvectors in decreasing relevance. The associated diagonal values s_{ii} with $s_{ii}^2 = \lambda_i$ ($\lambda_i = i$ -th eigenvalue) are ordered with decreasing magnitude. It is straightforward to identify

$$U_{rc} S_{cc} = E^\dagger \quad (4)$$

$$V_{cc} = C^\dagger \quad (5)$$

Data matrices E^\dagger and C^\dagger contain the requested information, however in a mathematical, abstract form and associated with random error and bias.

If experimental data could be obtained unaffected by random errors and bias, AFA would result in a limited number of non-zero eigenvalues in S_{cc} corresponding to the dimensionality of the data matrix, that is the number of factors contributing to the experimental data under investigation. However, experimental data can hardly be obtained without random errors and bias. Therefore, all eigenvalues λ_i are non-zero, albeit usually quickly approaching very small

values with increasing i . Decision on the dimensionality of the data space therefore has to be based on statistical tests. Only the n largest eigenvalues λ and the associated column and row eigenvectors are contributing significantly to the experimental variance.

The remaining $c-n$ eigenvalues and the associated eigenvectors λ° form the so-called null space or error space (indicated by $^\circ$) and are excluded from the further analysis. Forming matrices E^* and C^* from the first n row and column eigenvectors only allows calculation of a matrix X' , where random errors and bias are reduced by omitting summation over the null-space. Therefore, X' differs from X by the amount of removed random error and bias. The informations thus filtered from the original data has to be transformed from the abstract orthogonal eigenvectors into physically meaningful vectors by rotating the n eigenvectors. This transformation is called target rotation by rotation matrix T :

$$X' = E C = E^* T T^{-1} C^*. \quad (6)$$

The main task in this step is to identify a suitable rotation matrix T , that yields the required information.

In order to test a suspected single component spectrum s_i , a least square transformation vector t_i is calculated according to Eq. 7

$$t_i = E^{*+} s_i \quad (7)$$

where E^{*+} is known as the pseudo inverse of E^* . By Eq. 8, a vector x_i is predicted, being the best projection of s_i into the factor space of matrix E^* :

$$x_i = E^* t_i \quad (8)$$

The difference between s_i and x_i can be used as a measure for the acceptability of a suspected component vector s_i .

2 Statistical Basics

Least-squares, computer-intensive resampling and robust principal component analysis

The statistical treatment of factor analysis is a complex issue. Some references are given in the Appendix. It is fundamental to keep in mind that factor analysis is a method of linear regression. Linear regression is a common tool in interpreting data sets. Its availability on pocket calculators further contributed to its current position as the most often applied tool for interpreting univariate data sets with apparent linear relationship. The coefficient of correlation is commonly (and inadequately) considered as a measure for the quality of fit.

However, the concept of linear regression is based on several assumptions. The figures created via linear least-squares regression are meaningful only if these assumptions are valid:

1. the expectation function is correct
2. the response is expectation function plus disturbance
3. the disturbance is independent of the expectation function
4. each disturbance has normal distribution

5. each disturbance has zero mean
6. the disturbances have equal variances
7. the disturbances are independently distributed

It has been shown that some of the requirements are not fulfilled for spectroscopic data. The least-squares criterion has some optimality properties only in case the seven requirements hold. If the requirements do not hold, the optimality properties of the linear least-squares estimates are likewise not valid.

Multivariate analysis depends on the same criteria. A value obtained from a factor analysis is doubtful if the disturbances (often also termed residuals or noise) are not, for instance, identically and independently distributed (i.i.d.). Criterion 2, for instance, can be expressed in matrix formulation

$$A = X + N \quad (9)$$

where A is the matrix of spectral observations, X is the matrix of the true absorptions and N is the matrix of disturbances. After a factor analysis is performed, the matrix A' is formed being the estimate of X. Then, the matrix A-X = N', with N' being the estimate of the disturbances. An autocorrelation analysis of the disturbances will show immediately that the spectral residuals are not i.i.d. Consequently, the data in A' has no optimality criterion - it is in a certain degree arbitrary.

Here, certain statistical techniques can be introduced to estimate the range in which the optimal estimate can be found. These techniques do not depend on complicated mathematical analysis but on brute computing force. These techniques have been termed bootstrap techniques. In fact, the father of bootstrap techniques, B. Efron, intended to term the technique 'shotgun' because it "can blow the head off any problem if the statistician can stand the resulting mess". The resulting mess probably directed to the enormous amount of data generated by a bootstrap procedure. But in the times of GHz CPUs and Gbyte digital storage capacities, the mess can be handled easily.

Bootstrap methods are computer-intensive Monte Carlo methods. Bootstrap methods replace sophisticated and often unavailable statistical theory by the brute computing power. These techniques allow to concentrate on the answers that are asked for instead of tracing the questions to a small catalogue of mathematically solvable cases that itself are mostly only valid for large sample approximations, e.g. normal distributions..

A data set of experimental observations (x_i, y_i) is used as a basis of discussion. The data are sampled from an unknown probability distribution F conditional on x. The data are interpreted by a parametric model function eq. 10.

$$F(x)=F(x; X) \quad (10)$$

Due to unavoidable experimental errors, the total process is expressed actually by

$$y_i = F(x; X) + \xi_i \quad (11)$$

where ξ_i represents the random process while $F(x, X)$ represents the deterministic part. Hence, the information on the parameters X is enclosed in the data (x_i, y_i) and extracted via the model function $F(x, X)$. The presence of random noise ξ transfers to the parameters X ($X=K, \beta_{101}, \beta_{102}, \beta_{103}$). The job is to obtain the probability distribution of a statistics θ , say $\theta = t(X)$. Having evaluated an estimator $\hat{\theta} = t(X)$ conditional on the predictors x , the next task is to assess the accuracy of $\hat{\theta}$ as an estimator for the true value of θ . The standard error of $\hat{\theta}$, the square root of its variance,

$$\sigma(\hat{\theta}; F) = [\text{var}_F(t(X))]^{1/2} \quad (12)$$

is the most common measure of accuracy for estimators $\hat{\theta}$ that are unbiased, or nearly so. A formal expression exists if the statistic θ is the mean:

$$\sigma(\hat{\theta}; F) = [\sigma^2(F)/n]^{1/2} \quad (13)$$

An unbiased estimate for $\sigma^2(F)$ is available

$$\hat{\sigma}^2(F) = \frac{\sum_{i=1}^n (X_i - \hat{X})^2}{(n-1)}$$

Thus, the well-known estimate of a standard error of a mean is obtained.

$$\sigma(\hat{\theta}) = \left[\frac{\sum_{i=1}^n (X_i - \hat{X})^2}{(n-1)n} \right]^{1/2} \quad (14)$$

Note that the estimate of the standard error is valid independent of F . However, such a closed formula is available only for the sample mean. Other statistics, e.g., median, correlation coefficient, skewness or ratio of means, such formula are not available.

In order to apply eq. 12 an estimate for F , say \hat{F} , is required. It is common to circumvent this requirement (or better: to ignore the question completely in many cases related to chemistry) by using the normal distribution $N(\sigma, \hat{\theta})$ as an estimate for F . This plugging-in of $N(\sigma, \hat{\theta})$ for F will be termed as 'classical statistics' in the following.

Efron showed that a simple computer algorithm exists for estimating a distribution \hat{F} of the observed data (x_i, y_i) that gives the standard error estimate

$$\sigma(\hat{\theta}; \hat{F}) = ((\overline{x - \bar{x}})^2 / n)^{1/2}, \quad (15)$$

almost the same as the classical estimate eq. 14. This algorithm gives

$$\sigma_{\text{boot}}(\hat{\theta}) = \sigma(\hat{\theta}; \hat{F}) = (\text{var } \hat{F}(t(x^*)))^{1/2}, \quad (16)$$

where x^* is a hypothetical data vector generated by indendently and identically distributed sampling from the distribution \hat{F} . Hence, the empirical distribution \hat{F} is created by resampling from the original data (x_i, y_i) via Monte Carlo methods since the data represents a realization of the original but unknown distribution F . This procedure is called 'bootstrap' and the standard deviation σ_{boot} is termed 'bootstrap standard deviation'. \hat{F} is the maximum likelihood approximation of F . The necessary number of resamplings is large, e.g., 1000 resamplings. Therefore, bootstrap methods are computer-intensive Monte Carlo procedures that became feasible only by the increase of cheap computing power in the past twenty years.

It is straightforward to show that the linear least-squares approach gives confidence regions different from the bootstrap estimates.

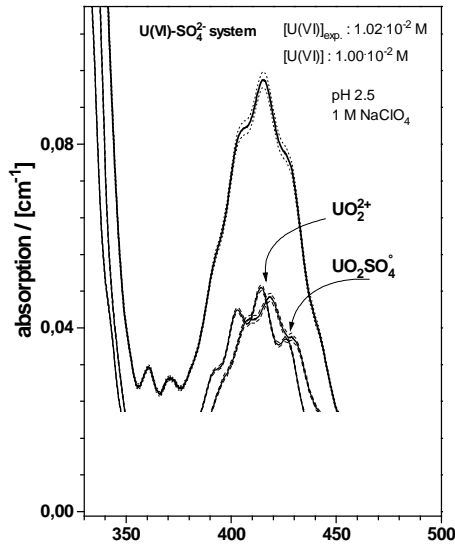


Figure a: Result of a MBB spectral analysis of an mixed component U(VI)-SO₄²⁻ UV-Vis absorption spectrum

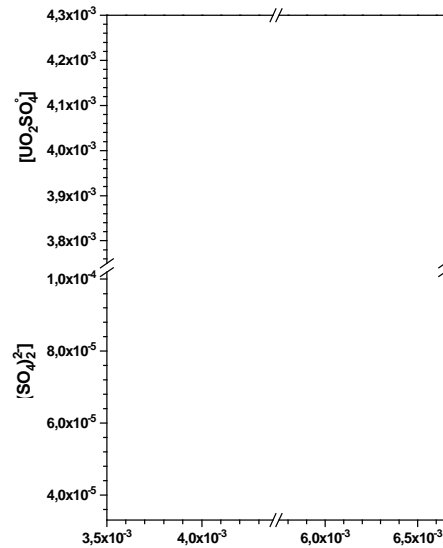


Figure b: Comparison of LSR 95% confidence ellipses with MBB results from 1000 resampling cycles for spectrum fig. a. The bias of LSR is obvious. Coverage of LSR ellipses by MBB results is almost negligible.

Figure a shows the interpreted three-component spectrum of a sulfate-containing U(VI) solution. Fig. b gives a comparison of a least-squares confidence region (LSR; dashed ellipses) and the corresponding distribution of 1000 bootstrap resamplings using a special bootstrap method called Moving Block Bootstrap (MBB). It is interesting to note that the MBB estimates are found almost consistently outside the range given by the 99% confidence ellipses. The LSR mean values are found outside the MBB cloud of data points.

To check the independence of the residuals, an autoregression analysis is performed. Autoregression (AR) assumes that there is a dependence of neighboring residuals over a distance of r neighbors, where r is called the 'lag'. A simple AR scheme is the a second order autoregressive scheme, AR(2), where the lag $r=2$. An AR(2) scheme estimates correlation by interpreting a residual z_t from spectral analysis due to a model

$$z_t = \beta z_{t-1} + \gamma z_{t-2} + \varepsilon_t, \quad (17)$$

where the t -th observation depends on the two previous observations z_{t-2} and z_{t-1} by the two parameters β and γ , resp. The disturbances ε_t are random contributions. Correlation is indicated by autoregression parameters β , γ significantly different from zero. By eq. 17, the seemingly random and uncorrelated residuals are interpreted by a model that explains the magnitude of the t -th residual by the magnitudes of the both closest neighbour residuals to the left. This is a kind of time-series analysis. A non-zero value of β and γ doesn't automatically indicate correlation. he magnitude of β and γ must be significantly diferent from zero. Again we are using bootstrapping to evaluate approximate confidence regions for the two parameters by sampling from the residuals ε_t . A bootstrap analysis gives AR(2) coefficients $\beta = 0.2867 \pm 0.03293$ and $\gamma = 0.66924 \pm 0.0329$. Both AR(2) coefficient estimates are significantly different from zero, as indicated by the standard deviations.

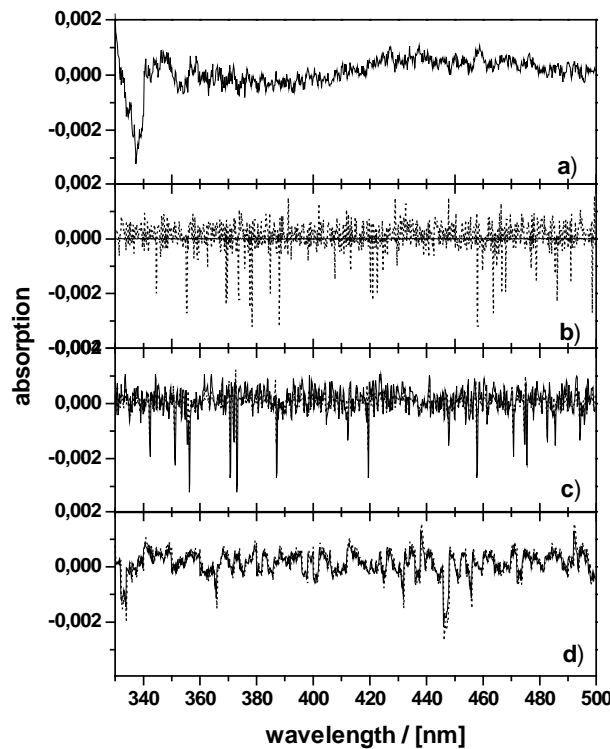


Figure c: Residuals (a) and results of AR(2) analysis for a Moving Block Bootstrap realization from the residuals with lag size $r=1$ (b), $r=2$ (c) and $r=10$ (d). MBB generated residuals are given as dotted lines, while AR(2) fits are given as straight lines.

Figure c shows the influence of the lag size on the reproduction of noise patterns generated by random sampling from the original residuals. It is obvious that the noise gets the more similar to the original noise the larger the lag size. It is very difficult to determine an optimal lag size. Hence, arbitrarily selecting a lag size introduces considerable subjectivity. If the lag size is inadequate, additional noise is introduced into the regenerated residuals. As a matter of fact, it has been shown that both a fixed lag size isn't appropriate in most cases and an inappropriate lag size may cause large breaks in the time series. A remedy is the threshold bootstrap.

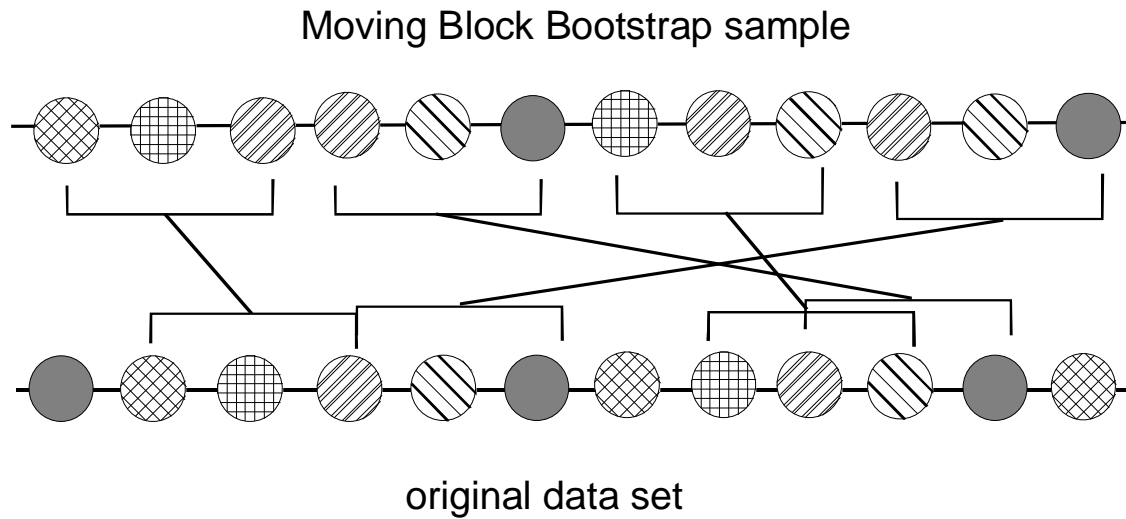


Figure d: Schematic representation of the Moving Block Bootstrap scheme. Twelve symbols representing the original data set are randomly redistributed in a Moving Block Bootstrap sample by choosing blocks of size r (here $r=3$).

Threshold bootstrapping residuals is a self-adaptive procedure. Hence, the noise itself defines the lag size. To perform a threshold bootstrap, the noise is mean-centered. Hence, the time series will run above and below the zero axis. A lag is now selected between points of a lag crossing the zero lines twice. Hence, the time series of noise is separated in individual regions of different length. When the different lags are assembled to form a new noise time series, no larger jumps at the connecting positions are introduced because in the original noise series, a transition above or below the mean value of the total series would have occurred, too. By using a threshold bootstrap strategy, we can take into account the correlation in the spectral residuals. In eq. 9, the matrix N is a random matrix. The randomness may influence the interpretation of the spectral data. By threshold bootstrapping, an optimal scheme is applied to investigate the influence of N on the possible interpretation of the spectral information.

It is necessary to give some comments about the difference between computer-intensive resampling methods and robust multivariate regression techniques like robust principal component analysis (RPCA). The robust regression techniques attempt to eliminate the effects of influential observations on the regression procedure. The term 'influential observations' refers to more than just 'outliers'. An 'outlier' can be easily spotted visually in case of

univariate data. In certain situations, simple and efficient tests (e.g. Dixon's test) are helpful in making a decision about a suspected outlier more objective. In univariate data (xy data), the definition of an outlying observation is much more difficult. Of course, there are data points which obviously deviate crudely from the remainder of the data set. However, the characterisation of a data point as an 'outlier' is more tricky than an untrained scientist may imagine. Just consider a situation where a large data set holds multiple 'outliers'. Due to the variety of situations where the structure of a data set may give rise to the assumption that one or even several members of the set are 'outlying', the term 'influential observation' has been coined. Set members with high leverage are an example of such influential data points. Robust regression techniques try to identify influential data points and to reduce their influence on the data analysis procedure. As an example of robust regression techniques, the 'least median of squares' method is mentioned here. Instead of minimising the sum of squares of residuals, as in common linear regression, the median of squared residuals is identified. This happens by collecting the residuals, ordering them according to size and identifying the residual at position $N/2+1$ of an uneven number N of data points, or the average value between $N/2$ and $N/2 + 1$ for an even number N of data points. Optimisation criterion is the minimisation of this median value. It is shown that almost 50% of data may be outlying without changing the position of this median value - and consequently the position of the regression line. In case of classical least sum of squared residuals regression, one single outlier changes the position of the regression line. Robust regression does have its disadvantages. For classical regression, a one-step algorithm exist to identify the optimum regression line. In case of, say, least median of squares regression, a complicated search algorithm must be implemented. There is no algorithm which warrants that the optimum robust regression line has been found.

What is true for univariate data sets is the more true for multivariate data sets. Recently, robust median of squared residual methods for principal component analysis have been made available. These techniques exclusively treat influential observations. Other effects, especially correlation effects, deviations from normality and non-linearity are not treated. Because no algorithm exists warranting optimal parameter extraction, RPCA does not have advantages in the present context. It is clear, however, that preparing samples and collecting spectra must be performed as careful as possible. Any incorrect procedure may spoil the results produced from CAT.

3 Computer-assisted Target Factor Analysis

There are a larger number of algorithms available in literature that make use of the fundamental mathematical relationships summarized in eq. 6. Examples are the SIMPLISMA approach, the convexity approach, orthogonal projection method and the Iterative Key Set Factor Analysis. Each of these methods does have its strengths. However, for the present situation they all have a fundamental disadvantage: They need user interaction. If, however, 1000 to 2000 repetitions of a data analysis are to be repeated, user-interaction in each step is impossible.

CAT takes another way. A matrix A' holds a series of informations about the chemical system under study, however affected by noise and error. The error is a general nuisance factor and must be minimised as far as possible. Factor analysis cannot remedy the inadequacies from experimentation. The noise shall be termed 'measurement uncertainty' from now on. Measurement uncertainty is an unavoidable part of all experimental data. Minimising measurement uncertainty is one of the major tasks of a scientist. However, all quantitative information concerning the composition of samples, the concentrations of components and the spectral absorptions carries measurement uncertainty. And uncertainty means doubt. Because neither humans nor computers are gods, we cannot know truth. Hence, we have to evaluate our information taking the doubt into account. Doubt needs not to stop us making conclusions. But it helps ourselves and others to assess the reliability of a conclusion and to direct us to possible other interpretations. Hence, the task of a numerical interpretation is not to take the burden of interpretation from a researcher's shoulders but to allow him to adequately balance the weights...

Estimating rank

While the threshold bootstrap algorithm allows to account for residual correlation, an estimate of the number of spectral contributions is required, that is the true rank of the experimental data matrix A . There are a larger number of techniques to estimate the rank of A . But all these techniques (scree test, Bartlett's χ^2 criterion, Durbin-Watson statistics, canonical correlation etc.) make statements about random properties and give only probabilistic comments about the likely rank of A . Even if the probability for a certain value of the rank should be high, it doesn't exclude the other ranks with low but non-zero probability to be correct anyway. Eventually, all possible ranks must be tested and the resulting interpretations of a system must be balanced. It occurs that factor analysis doesn't suggest a unique interpretation of a system but points out different possible models. It is the scientist's task to develop a strategy to come to a final conclusion.

For UV-Vis spectroscopic system, because the Bouguer-Lambert-Beer law allows to do so, a data matrix of different spectra of a chemical system can be interpreted with clear boundaries. The first boundary is the non-negativity of oth concentrations and absorptions. A furtherboundary can be selected if a component spectrum is known, for instance the uncoordinated, hydrated metal ion. Furthermore, the total concentrations of metal ion(s) and ligand(s) in the system are known, at least within the limits of measurement uncertainty. These informations are sufficient to resolve a chemical system. Hence, if a fast algorithm is available to resolve the mathematical equations at least semi-automatically, there is no need to rely on statistical tests in searching the true rank of an experimental data matrix A . Then, the total numerical process can be taken for any selected rank and the results can be compared. Often, one interpretation suggests itself while other ranks lead to numerical results with larger negative values in either the single component spectra and/or the species or ligand concentrations.

Performing the Target Rotation

The decisive step in target factor analysis is the target rotation step. If a rank is specified by the user, the mathematical-numerical steps to eigenvectors and eigenvalues are straightforwardly accomplished by efficient algorithms. Hence

$$X' = E C \quad (18)$$

is available without requiring further attention.

The essential step is the target rotation matrix T . T must fulfill several requirements, especially leading to non-negative values in E^* and C^* . There is no easy way to generate a matrix with these properties. However, target transformation matrix T has a helpful property: T is comparatively small. If the system under study is a two-component system, dimensions of T are 2×2 . For an n -component system, the dimensions of T are n^2 .

CAT estimates the elements t_{ij} of T via a SIMPLEX algorithm, using the sum of neative elements in the estimates of E^* and C^* as a criterion. Using the non-negativity as an optimisation criterion has, however, a special case: If the diagonal elements t_{jj} in T are zero, then all elements in E^* and C^* are zero, too and hence, non-negative. It is necessary to avoid this trivial solution.

Avoiding Trivial Solutions in Target Rotation Matrix T

It is not sufficient to fix the values of the diagonal elements in T to, say, $+1$ each. While the matrices A , E^* and C^* must be positive semi-definite, T needs not. All values between $-\infty$ and $+\infty$ are principally allowed. Hence, the diagonal values of T must be set to either $+1$ or -1 . And here, a problem arises. Which combination of $+1$ and -1 values is adequate?

Hence, all combinations of $+1$ and -1 values must be investigated. But the computational burden can be lessened because the first eigenvector is always either completely positive ($t_{11} = 1$) or negative ($t_{11} = -1$). Hence, the correct value for t_{11} can be taken directly from the properties of the first eigenvector in E . For all other diagonal elements in T , the combinations must be tested. Experience tells that there occur rare cases where two combinations of diagonal values give comparable but different results. CAT has a built-routine which allows to search for the adequate combination of diagonal elements.

The user will probably be surprised to be required to enter "0" instead of "-1". The diagonal elements are given as combinations of 1's and 0's and refered to as 'keys'. A 'key' 011 refers to diagonal elements $(-1, 1, 1)$, while a key 0100 refers to $(-1, 1, -1, -1)$.

SIMPLEX optimisation of the off-diagonal elements in T

The SIMPLEX routine is an iterative optimisation algorithm that does not need analytical or numerical derivatives. Hence, the SIMPLEX can be applied with optimisation criteria different from the squared residuals (termed L_2 criterion instatistics). The SIMPLEX also finds minima in the error space if the optimisation criterion is, e.g., the smallest sum of negative values. In case of implementation in CAT, the SIMPLEX algorithm starts with randomly chosen values in the $n(n-1)$ off-diagonal elements of T. Some precautions are necessary because there exist local minima in the error-space of the iteration procedure. Repeating an analysis with different starting values is often helpful. CAT has this feature built-in. In fact, CAT allows the resolution of a spectroscopic system by optimizing $n(n-1)$ values of a transformation matrix T. After exploring certain properties of the data space like rank and combination of diagonal values, CAT provides a solution without interference by the user. And it repeats this analysis with slightly modified values independently. Hence, CAT can advantageously used as a core element of a threshold bootstrap process.

This is the philosophy of Threshold Bootstrap Computer Assisted Target Factor Analysis.

4 Performing a CAT Analysis

A typical CAT analysis follows the following procedure:

1. collection of a set of UV-Vis spectra under multivariate conditions
2. modification of the spectra according to the CAT input conventions
3. background correction of the spectra and, if necessary, generation of a uniform wavelength range
4. analysis of the spectral information to estimate the likely number of species and the corresponding key of diagonal elements
5. CAT analysis to test possible species combinations and stoichiometries
6. Input of the Type B measurement uncertainties
7. evaluation of the complete measurement uncertainty budget via Threshold Bootstrap Computer Assisted Target Factor Analysis (TB CAT)
8. evaluation of the distributions of formation constants and spectral information

5 Installation

Requirements

Screen resolution of at least 1024 x 768 pixels

Windows OS 98SE or XP with at least 500 MHz processor

Installation

TB CAT is available as a compressed ZIP file. The ZIP file should be unpacked into a temporary directory. Upon clicking the SETUP.exe the installation routine starts. Follow the guidance of the installation routine in the way that is common for the installation of new software on Windows OS. The routine creates a desk top icon but does not make modifications in the registry.

Known bugs

This code is a work-in-progress. The code is not a commercial program. It has a number of incomplete sub routines and dead-end menus. However, the code is currently running stable on a variety of Win 98SE and XP machines.

Input Files

For its performance, TB CAT needs a few data files which must be prepared by the user. These files are

- a) spectral data files with component and concentration information in the header section
- b) input files with starting information for the target rotation procedure
- c) estimates of the ISO Type B uncertainties

Spectral data files are obtained from the experiments.

Header sections hold the information about the composition of experimental solutions.

Input files for the target rotation procedure are generated during the ANALYZE step of the CAT procedure. The user may request the system to generate several files in order to test different hypotheses.

The ISO Type B uncertainties are educated guesses or may be obtained from repeated experimentation

6 Running CAT

Data File Format Convention

The format of input data files is as follows (statement in *italics* are comments):

first line: beginn

beginn is a key word

second line: component 1, component 2,....

example: Co2+, NH3, Cl-:

1. Note: the component, whose spectral information is investigated, must be given in first place!

2. The component pH requires a special treatment, e.g. statement of ionic strength

third line: a,b,c

a,b,c are concentrations of the components in the given sample solution.

example: 2.1E-3, 0.0045, F2E-4

the key word F before a concentration indicates a free concentration. Otherwise, the free concentrations are obtained arithmetically as the difference between the total concentration of a component (e.g. a ligand) and the concentration found by CAT analysis to be bound to the metal ion. Values of pH obtained from glass electrode potentiometry must always be given with an F.

fourth line: ende

ende is a key word indicating end of header section

fifth line: d,e,f

d: number of lines (generally the number of (wavelengths, absorption) pairs

e: the shortest wavelength

f: the longest wavelength

all following lines: g,h

g: wavelength value

h: absorption value

```

beginn
UO22+,SO42-,pH, .1
0.01,0.0015,F3.25
ende
2051,325,530
325,1.99791648825377
325.1,1.98821712932768
325.2,1.97811777040159
325.3,1.9681184114755
325.4,1.9576190525494
325.5,1.94731969362331
325.6,1.93622033469722
325.7,1.92642097577113
325.8,1.91562161684504
325.9,1.90362225791894
326,1.89322289899285
326.1,1.88122354006676
326.2,1.86932418114067
326.3,1.85732482221458
326.4,1.84492546328848
326.5,1.83202610436239
326.6,1.8187267454363
326.7,1.80692738651021
326.8,1.79422802758411
326.9,1.78092866865802
327,1.76852930973193
327.1,1.75482995080584
327.2,1.74043059187975
327.3,1.72773123295365
327.4,1.71403187402756
327.5,1.69973251510147
327.6,1.68493315617538
327.7,1.67033379724929|

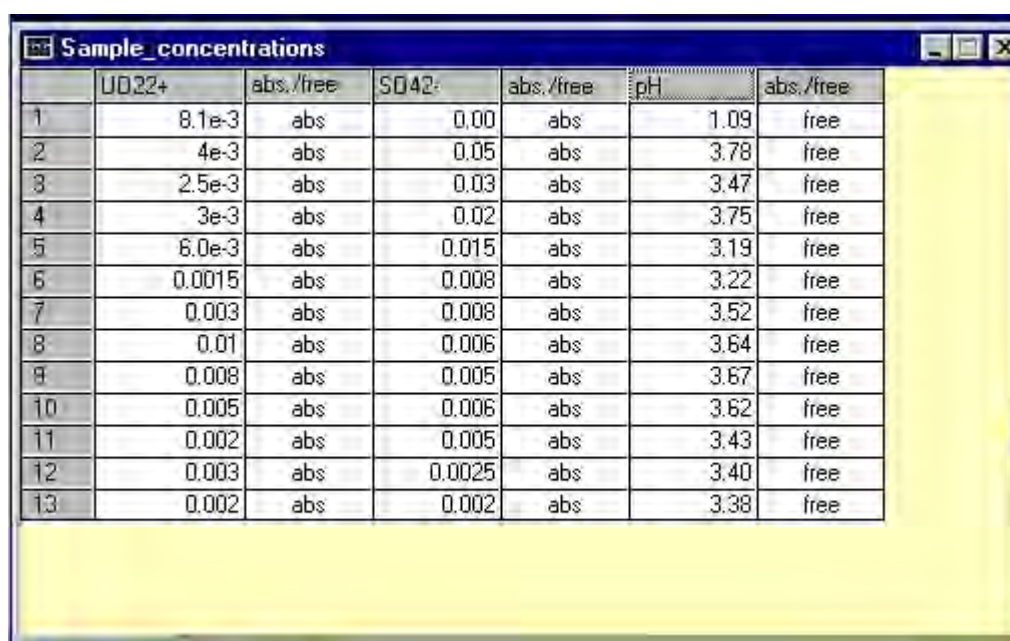
```

Figure e: First lines of a CAT input file of one UV-Vis spectrum

Sequence of spectra in the input procedure

With a single exception, the input sequence of the files can be selected arbitrarily. The exception is the spectrum of the component with the known single component spectrum. This spectrum **MUST** be loaded in first position. It is, therefore, helpful to name the data files appropriately to allow convenient sequential loading of the files.

Spectra are loaded into the program via the File menu. If the respective files are loaded, CAT displays the concentration information as follows:



	UO ₂₂₊	abs./free	SO ₄₂₋	abs./free	pH	abs./free
1	8.1e-3	abs	0.00	abs	1.09	free
2	4e-3	abs	0.05	abs	3.78	free
3	2.5e-3	abs	0.03	abs	3.47	free
4	3e-3	abs	0.02	abs	3.75	free
5	6.0e-3	abs	0.015	abs	3.19	free
6	0.0015	abs	0.008	abs	3.22	free
7	0.003	abs	0.008	abs	3.52	free
8	0.01	abs	0.006	abs	3.64	free
9	0.008	abs	0.005	abs	3.67	free
10	0.005	abs	0.006	abs	3.62	free
11	0.002	abs	0.005	abs	3.43	free
12	0.003	abs	0.0025	abs	3.40	free
13	0.002	abs	0.002	abs	3.38	free

Figure f: The sample concentrations window. CAT creates this window from the headers of the loaded spectrum files. The files are listed according to their input sequence. In the given case, the spectroscopic signal is generated by the metal ion UO_2^{2+} . Hence, the first spectrum holds only uncoordinated $\text{UO}_2^{2+}(\text{aq})$. The UV-Vis absorption spectrum of this species is therefore known in the subsequent analysis.

The pH values are transformed into OH^- concentrations inside the program using the information about ionic strength and the Davies equation. Therefore, in the spectrum headers, the user has to provide the ionic strength following the component name "pH"; e.g. '...,pH, 0.25'. If the equilibria under study depend on pH, the component pH should always given last. Figure e provides an example. The flag F informs CAT about the character of the data. The Sample Concentration window shows that the pH is understood as free concentrations.

7 Basic Procedures

TB CAT consists of a main window with seven principal menu items:

- a) File
- b) Do
- c) Uncertainty
- d) Data
- e) Evaluate
- f) About
- g) Windows

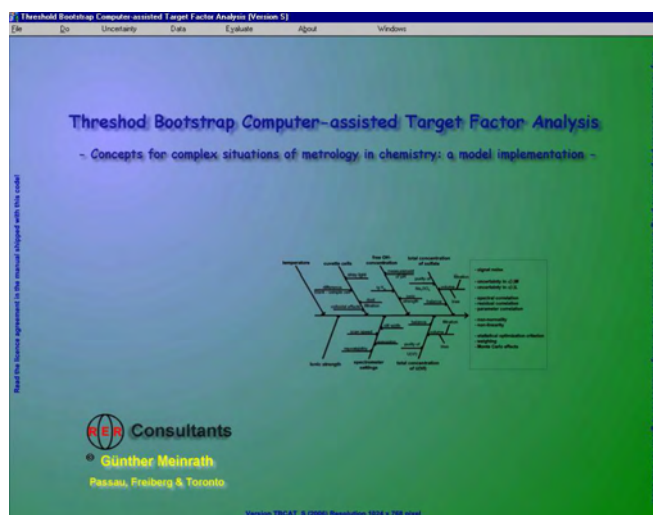


Figure g: The CAT main window.

a) File

The menu item FILE has the sub items

- (1) open
- (2) save
- (3) save as (inactive)
- (4) close
- (5) exit

(1) **Open** : load the data files composing the matrix of experimental data

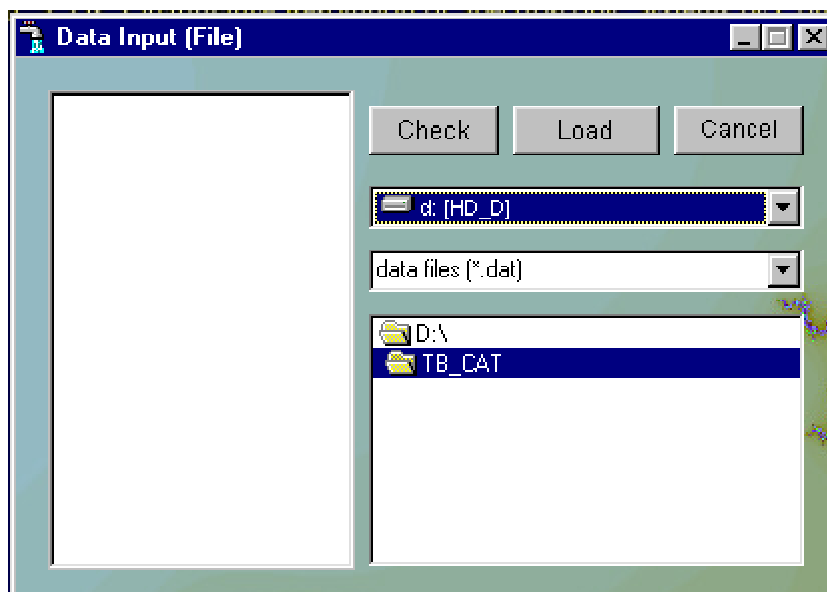


Figure h: Data Input screen of the **File/Open** menu item

In the left field, the files with extension specified by the user appear. The user may select the files which should have been adequately named before starting TB CAT. For raw data files, the extensions *.dat or *.asc are recommended.

Each file must follow the data file format convention explained in the previous §.

After selecting and before loading the files, the Check button allows to read the first line of each file. Thus the user can assure that the files do have the same wavelength range.

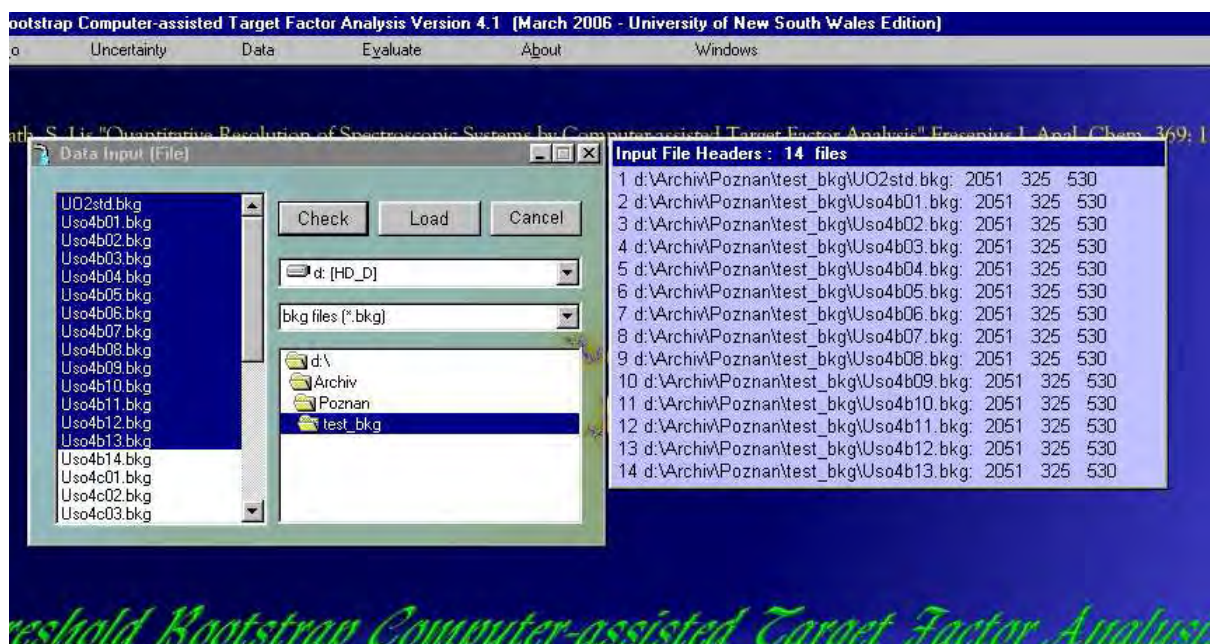


Figure i: Data Input Window with the File Header Checking Window

Upon loading, CAT performs a series of tests ensuring that the data format is appropriate for the subsequent operations. The spectral information is then displayed. This operations may take some time depending on the total amount of data.

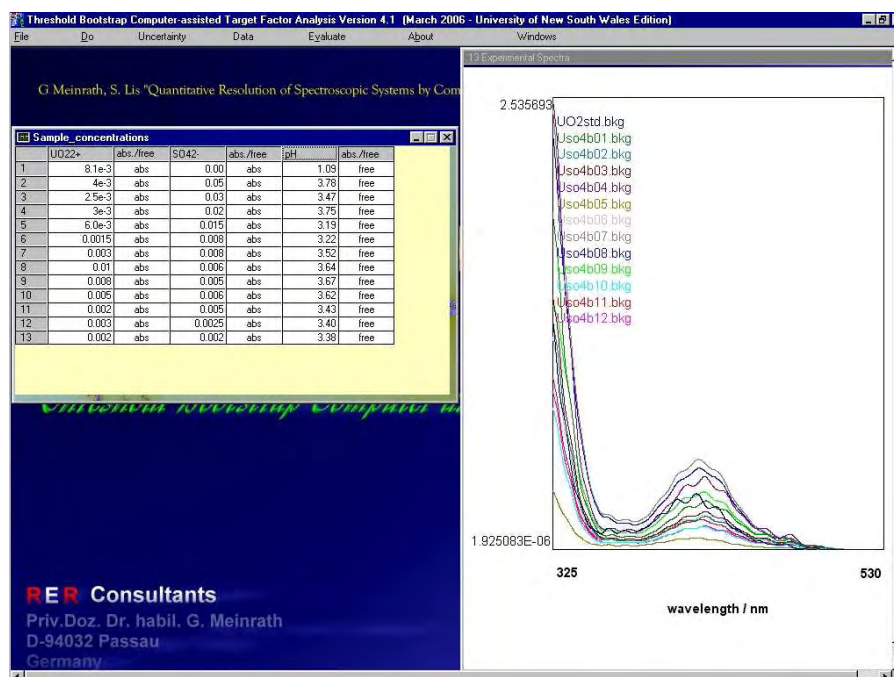


Figure j: Diagram with loaded spectral information. The position of the cursor inside the diagram is given interactively below the x-axis caption.

(2) Save

Information in the active window can be saved if the relevant information is available. Upon selecting **Save** a dialog window opens. Data in graphic windows is saved as ASCII data of the information displayed in the window. There is no way to save graphics data as a graphics format (e.g. JPG). The default extension is *.lis. The user may select own extensions.

(3) Save as

not implemented

(4) Close

not implemented

(5) Exit

Immediately shuts down TB CAT without further notice. No information will be saved.

b) Do

The DO menu is the working center of the TB CAT code. It offers the following items:

- (1) *Background correction*
- (2) *analyze*
- (3) *CAT*
- (4) *molar absorption*
- (5) *TB CAT*

(1) Baseline correction

Raw spectra from a UV-Vis spectrometer usually require to be normalised to a common background. Occasionally, it is also required to limit the spectral information to a more narrow wavelength range. For this procedure, the user selects the range to the left and right of the absorption band representing the baseline. TB CAT uses a linear least squares regression fit of a straight line through the specified intervals to determine the baseline for each spectral file. With this information the files are corrected to a common baseline. The result will be displayed in the graphics window immediately. This process can be applied to a subsection of the loaded spectra. The user can select individual files in the *Select Files* window.

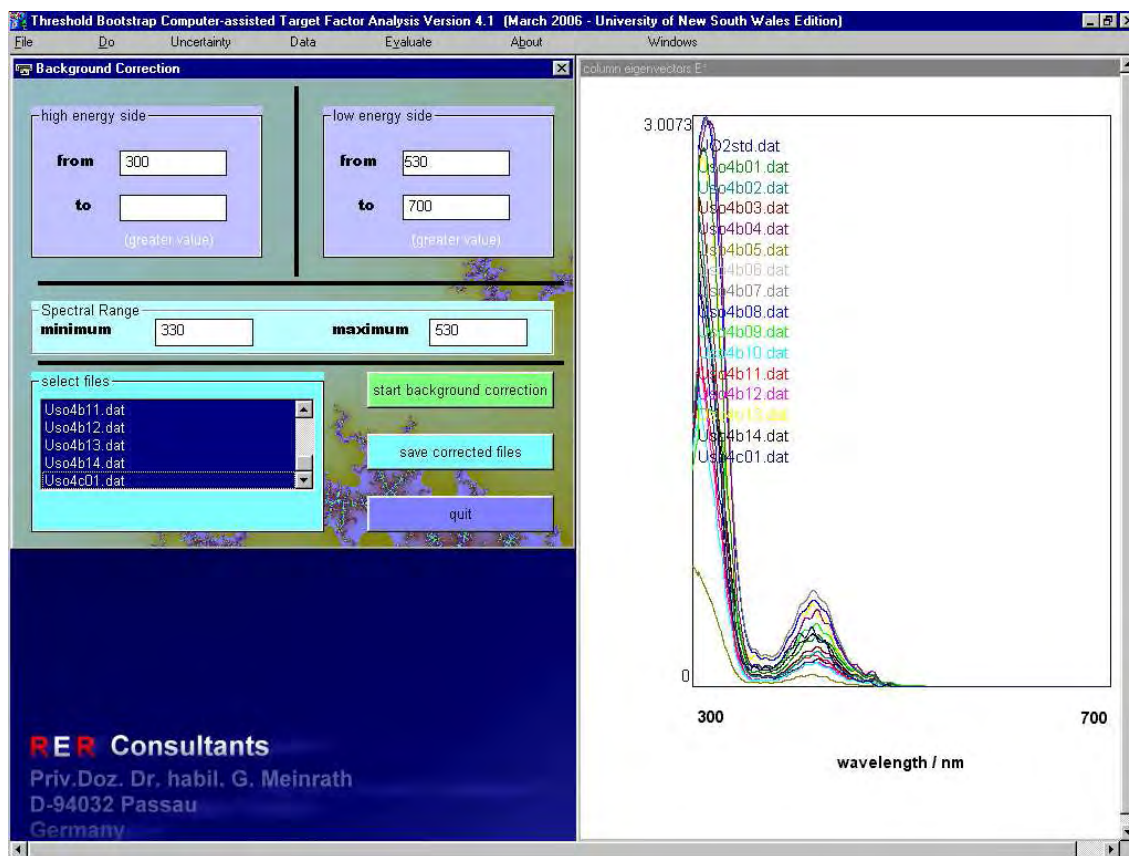


Figure k: Background correction window

Upon loading CAT sorts the imported data in ascending order of the dependent variable. It is hence possible to import data with disordered wavelength values or to import data ordered in descending order.

Specifying the baseline for least-squares correction

The both top windows allow specification of a baseline range at the high energy and low energy side. The left side is the low energy side, where the high wavelength appear. The right side is the high energy side where the low wavelength appear. It is for the sake of this procedure that the user should keep with the data input file convention. If this convention is reversed, the settings in this window must be reversed, too.

The shortest wavelength appear in the bottom field of the right top frame while the longest wavelength appear in the top field of the left top frame. The user is required to modify this position if the side should be used for baseline correction. If no modification is made and the empty field in the frame remains empty, CAT will not use this side in the regression leading to the definition of a baseline. This gives the user the possibility to exclude one side from the baseline calculation, e.g. if such a baseline is not available due to spectroscopic properties (continuous absorption to the UV) or if a baseline hasn't been recorded. The process of baseline correction can be repeated unlimited.

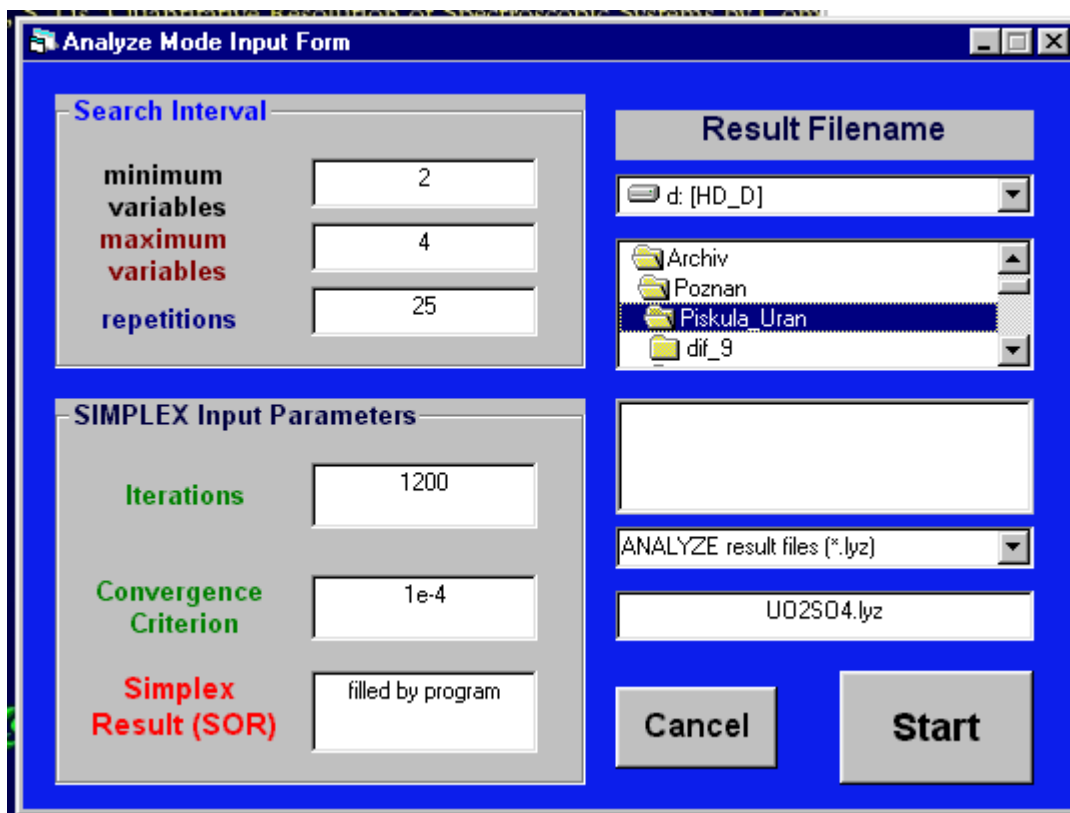
It is recommended to use the cursor position indicator fields of the graphical display for locating the baseline range to be applied in the baseline correction procedure.

Starting baseline correction and saving corrected file

The green button "start baseline correction" atarts the procedure. The button "save corrected files" opens a dialog window. The files will be saved by default with the same filename as the input files but an extention *.bkg. The user can also create a different directory clicking the "new dir" button. If a directory already holds files with the same file name, the user is asked to confirm the deletion of the old files.

(2) Analyze

The **analyze** item opens the Analyze Mode Input Form. It controls the determination of both the diagonal values of the target transformation matrix and the best starting values for the CAT routine.



The image shows a Windows-style dialog box titled "Analyze Mode Input Form". It has a blue border and standard window controls (minimize, maximize, close) in the top right corner. The dialog is divided into several sections:

- Search Interval**: A section on the left with three input fields:
 - minimum variables**: A text box containing the value "2".
 - maximum variables**: A text box containing the value "4".
 - repetitions**: A text box containing the value "25".
- SIMPLEX Input Parameters**: A section below the first one with three input fields:
 - Iterations**: A text box containing the value "1200".
 - Convergence Criterion**: A text box containing the value "1e-4".
 - Simplex Result (SOR)**: A text box containing the text "filled by program".
- Result Filename**: A section on the right with a file selection interface:
 - A drive dropdown menu showing "d: [HD_D]".
 - A file list showing folders: "Archiv", "Poznan", "Piskula Uran" (which is selected), and "dif_9".
 - A text box for the filename, currently containing "U02S04.lyz".
 - A dropdown menu showing the file type "ANALYZE result files (*.lyz)".
- Buttons**: At the bottom right, there are two buttons: "Cancel" and "Start".

Figure 1: Analyze Mode Input Form

Specifying the upper and lower limits of the rank and the repetition of the search

The left side frames *Search Interval* and *SIMPLEX Input Parameters* accept user input to explore the basic properties of the data matrix. The *minimum variables* field accepts an integer value for the lowest possible rank of the matrix, while the *maximum variables* field accepts an integer value for the upper limit of the true rank. The field *repetitions* specifies how often CAT creates a set of random starting values and repeats the interpretation of the matrix on assumption of a given rank and a given set of diagonal values +1 or -1. In fact, CAT repeats the interpretation for each possible permutation of diagonal elements. Each combination will be tested as often as specified under *repetitions*. Hence, if the minimum rank is 2 and the maximum rank is 4, CAT performs 100 tests if each permutation is tested 10 times.

SIMPLEX Input Parameters

The SIMPLEX algorithm is not as popular as it should be. The SIMPLEX algorithm is an efficient method to fit models to data which does not require derivatives. The concept of SIMPLEX is easy to grasp with a more detailed explanation. Such explanations are available in literature. There is no need to repeat such an explanation here. A SIMPLEX can be understood as a mesh of parameter values that moves over the parameter space according to detailed instructions. Its direction of movement is the minimum of the parameter surface where the surface is defined by the sum of negative values in E and C (L_1 criterion) or the sum of squared negative values in E and C (L_2 criterion). At present, only the L_2 criterion is implemented (L_2 is more commonly termed 'least sum of squared residuals'). There are two conditions when the SIMPLEX algorithm stops. First condition is fulfilled if the number of iterations specified in the field **Iterations** is satisfied. Second condition occurs if the convergence criterion is satisfied. The convergence criterion is satisfied if the difference in the mutual distances of the parameter values forming the SIMPLEX mesh is below a given critical value. The critical value is set in the **Convergence Criterion** field.

Rules of thumb for setting SIMPLEX parameters

In fact it is not necessary to have a detailed understanding of the SIMPLEX to use TB CAT. Some experimentation is more teaching than theory. As a rule of thumb, 1000 iterations and a convergence criterion of $1 \text{ E-}4$ is adequate for most situations.

Result filename

The Analyze process creates a table of informations which preferably should be saved for later inspection. The default extension for these files is *.lyz. A filename must be specified and confirmed by clicking the Return key. Then the **Start** button will become active and the Analyze procedure can be started.

CAT will calculate the possible permutations in the diagonal elements of T and the total number of runs given the minimum and maximum rank to be tested. The sum-of-residuals (SOR) indicate how well the single component spectra obtained from the experimental spectra under a certain hypothesis on the matrix rank (= number of species in solution) and the diagonal elements in T are able to explain the experimental spectra.

The Analyze procedure windows shows all central features of the TB CAT window. The yellow top window displays information about the SIMPLEX procedure. It shows the widest and narrowest mesh of the SIMPLEX. Of central importance are the values listed under convergence criterion. The SIMPLEX has converged if all figures are equal or below 1.0. The second essential element is the lowest right-hand side figure. It gives the sum of squared negative values in E and C (as will be explained later, a weighing factor is included).



Figure m: The Analyze procedure windows

Graphics windows

Two graphics windows are on display during the ANALYZE procedure. These windows are also available during the CAT and the TB CAT step. The light green left-hand-side window compares the estimate of the first calculated spectral component with the experimental spectrum of the known component (note that CAT assumes this spectrum to be loaded in first position). Inside the algorithm, the spectrum of the known component and the first estimated component spectrum are normalized. Therefore, only the shape of the spectral information is of interest. The magnitude doesn't play a role at this step of the procedure. The experimental spectral curve is given in cyan, while the calculated spectrum is given in red colour. CAT tries to find a spectrum that optimally fits the cyan spectrum. The sum of squared differences between the both spectral curves are one of three components summing to the total sum of residuals (SOR). This value is specified in the header of the spectrum graphics window.

The white background right-hand-side graphics window shows the current best estimates for the single component spectral curves. Because the diagonal elements of T are fixed, the relative magnitudes of these spectra are irrelevant – only the spectral shape is of importance here. Arranging rapidly changing windows on screen can be tiresome. The both graphics windows

are coupled. The left window will always assume the same size as the right window. While the right window can be moved freely, the left window will always stay with the left border of the TB CAT main window and the lower border of the SIMPLEX window. If the right window is enlarged by dragging at the borders of the diagram, the left window will follow.

	key	SOR
1	00	31003.04
2	00	31003.04
3	00	31003.04
4	00	31003.04
5	00	31003.04
6	00	31003.04
7	00	31003.04
8	00	31003.04
9	00	31003.04
10	00	31003.04
11	00	31003.04
12	00	31003.04
13	00	31003.04
14	00	31003.04
15	00	31003.04
16	00	31003.04
17	00	31003.04
18	00	31003.04
19	00	31003.04
20	01	16026.08
21	01	16026.08
22	01	16026.08
23	01	16026.08
24	01	16026.08
25	01	16026.08
26	01	16026.08
27	01	16026.08
28	01	16026.08
29	01	16026.08
30	01	16026.08
31	01	16026.08
32	01	16026.08
33	01	16026.08
34	01	16026.08
35	01	16026.08
36	01	16026.08
37	01	16026.08
38	02	16026.09

Analyze window (Key vs. SOR)

Central element of the Analyze process is the Analyze table. The left column gives the diagonal elements of the T matrix. A 1 represents a value +1 while a 0 represents a value -1. Because the value of the first element is given by the values of the first abstract factor, its value will not be permuted (this factor has either only positive or only negative values) and is always given as a zero in the keys. Internally, its value is taken into account by CAT. The sequence of 1's and 0's is termed a key. In the left uppermost field the total number of calculations is displayed. This figure allows a rough estimation of time consumption until completion of the total ANALYZE procedure.

During the Analyze procedure, the Analyze window will be filled with the key values and the respective SOR. Each key will be tested as often as specified by the user in the 'Repetitions' field of Analyze Mode Input Form. After termination of the Analyze procedure, the data list in the Analyze window will be sorted with increasing SOR. Hence, the smallest SOR representing the best fit will be displayed on top.

Figure n: The ANALYZE table

Creating default input files

Upon clicking a field in the SOR column of the Analyze window after termination of the procedure, a dialog box opens asking whether the user wants the program to generate a default input file with the relevant informations for subsequent program steps. The default input file will be saved as 'defaultxxx.elb'. (The extension *.elb honours the deceased former head of the Rare Earths Department at Faculty of Chemistry of Adam Mickiewicz University at Poznan/Poland, Prof. Marian Elbanowski). The 'xxx' stands for the key. Hence, it is possible to create an *.elb file with the best fitting input data sets of each possible key. Note

that the length of the key is not limited to three. Longer and shorter keys are also acceptable. The *.elb files are ASCII files.

An example of a default000.elb file is given below, which holds the following informations:

```
"key", "000"  
"niter", 400  
.269645854203445, 2.69645854203445E-02, .0001  
.213553310743779, 2.13553310743779E-02, .0001  
.987070096896006, 9.87070096896006E-02, .0001  
.335093279935166, 3.35093279935166E-02, .0001  
3.41121586767416, .341121586767416, .0001  
7.88138696376839E-02, 7.88138696376839E-03, .0001
```

The first line gives the key word "key", followed by the key in string format. The second line holds the key word "niter" which specifies the total number of iterations. The total number of iterations is fixed to 400. For the time being, modifications in total number of iterations must be made either by manipulating the ASCII file outside CAT or is handled within the CAT code.

The following lines define the parameters of the SIMPLEX. Because the key is currently 000, indicating diagonal values of the target rotation matrix to be -1, -1, -1, the dimension of the target matrix is 3 x 3. The 3 diagonal elements are set to -1, -1, -1. Hence, six variable matrix elements remain. The values of these matrix elements are given in the first column. SIMPLEX tries to optimise these elements further by creating a seven-dimensional polygon around the starting values. The polygon is built using the informations in the second column. It is easy to see that the values in the second column are just 10% of the first column. The last column specifies the stopping condition of the SIMPLEX iterations. The stopping condition is reached if the seven-dimensional SIMPLEX net has contracted in that way, that the largest difference is 0.0001 times the smallest value in that dimension. For those unfamiliar with the SIMPLEX algorithm, the following readings by M.S. Caceci are recommended.

In fact, the user is not obliged to be familiar with the SIMPLEX because the 'Analyze' procedure generates the necessary files. The user has to specify the range of ranks he wants to investigate and to specify the number of repetitions. Because starting values are generated by a random procedure, each run has some different starting values. This random feature tries to overcome local minima, which would mislead the procedure. A minimum number of 10 repetitions is therefore recommended.

It is possible to create default input files with different keys. For example, it is recommended to select the best result for each key within the first 5% to 10% of total results generated in the 'key vs. SOR' list. If for one key several default files should be created, the already existing default file with this key must be renamed. Otherwise, the existing key will be overwritten without any further notice. Note that default files should be stored in the same directory as the spectra files.

(3) CAT

CAT is a mean square principal component analysis procedure. It returns best fit values for sample composition, single component spectra and formation quotients.

CAT proceeds in two steps. In the first step, the abstract factor analysis is performed. For this purpose, CAT needs the informations which have previously been collected into the default input files, defaultxxx.elb. By filling the fields 'Starting Values', 'Step Width' and 'Convergence', the user is free to enter values by hand – a tiresome procedure.

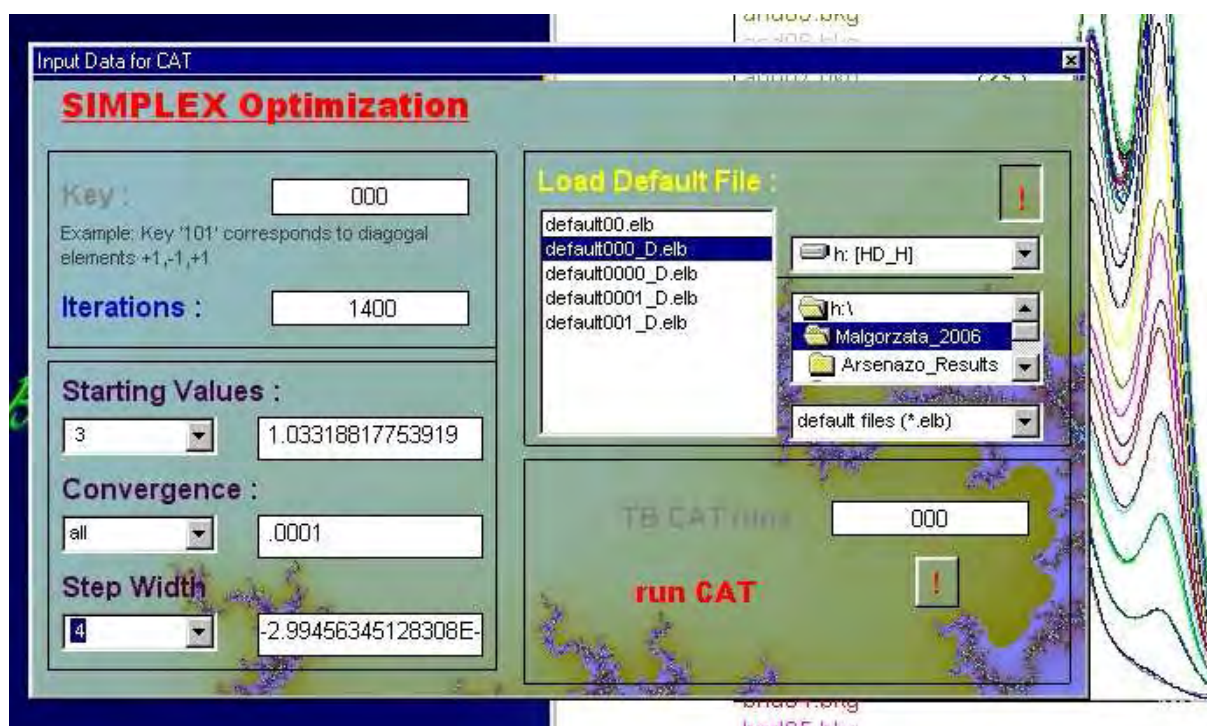


Figure p: CAT input window. The default file default000.elb has already been loaded and the number of iterations is set to 1000. In the next step, the concentration information file example.dat will be loaded. The values from the default000_D.elb file have been copied to the lower left hand side fields with caption 'Starting Values', 'Step Width' and 'Concergence'. These fields can be manipulated by the user.

The available default files are specified in the directory list box bottom right of the data input window. Loading a default file and specifying the number of iterations is sufficient to start the optimization procedure.

By clicking the respective "run CAT" button, CAT is searching for the rotation matrix values t_{ij} and displays the results. The result may look as in fig. q.

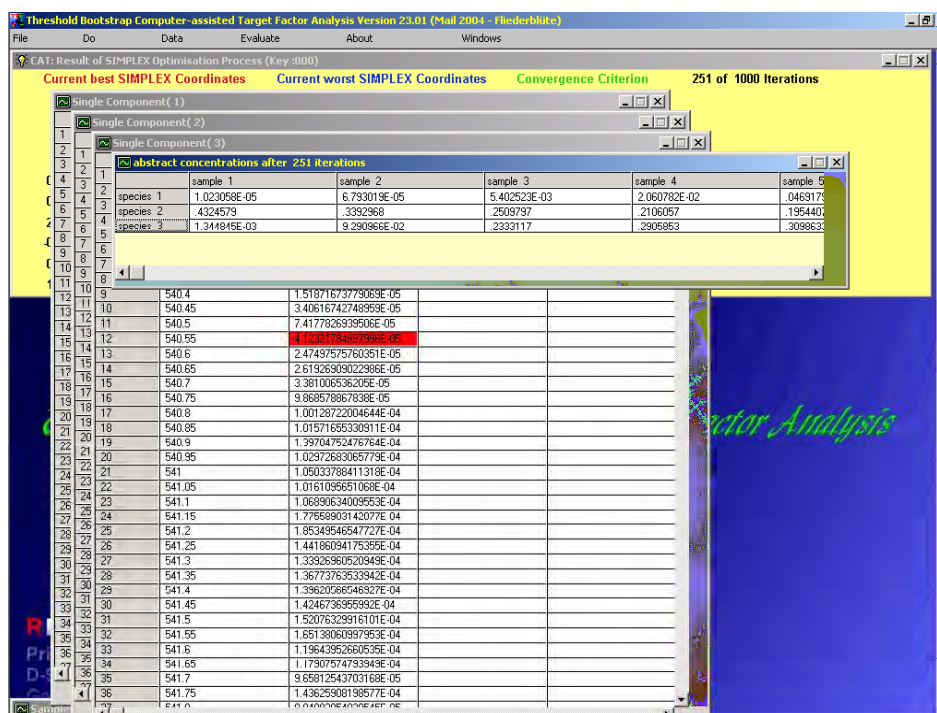


Figure q: Result of the CAT procedure. The abstract concentrations and the normalised single component spectra are given.

The estimated single component spectra are given numerically together with the concentration estimates. The user has the possibility to study these results which already represent the final spectral curves. But at this stage, there is no chemical interpretation because the rotation matrix T is normalised by the diagonal elements -1 or 1. It is necessary to transfer the normalised curves and concentrations into concentrations and molar absorptions consistent with the chemical composition of the samples. This is done by the 'Molar Absorption' menu item. In most cases the user will therefore will proceed with the next menu item - the 'Molar Absorption' step.

(4) Molar Absorption

The Molar Absorption menu item opens with a window collecting basic chemical informations about the system under study in the upper section and returning calculated information in the lower section.

CAT needs informations on the chemical system under study, e.g. the probable composition of the species formed, their names (even though 'species 1' and 'metal ion' are also acceptable).

The component names must already be given in the headers of the spectrum files. CAT automatically puts these names under 'components names'. 'Species identification' allows the user to change the default names 'species 1', 'metal ion', 'species 2', into meaningful names which allow an easier identification for the given system. The 'stoichiometry' section is the most essential input.

In the 'stoichiometry' section the user must give the chemical composition of the species. If component 1 is the metal ion M and component 2 is a ligand, say X, then a species MX is represented by the stoichiometric coefficient '11'. The metal ion alone is represented by the coefficient '10'. Note that the left hand field under 'stoichiometry' identifies the species, while the position of a stoichiometry coefficient represents a component. The sequence of stoichiometric coefficients must comply with the sequence of components given in the spectrum file headers – to be found under 'Chemical Informations - a) component names'.

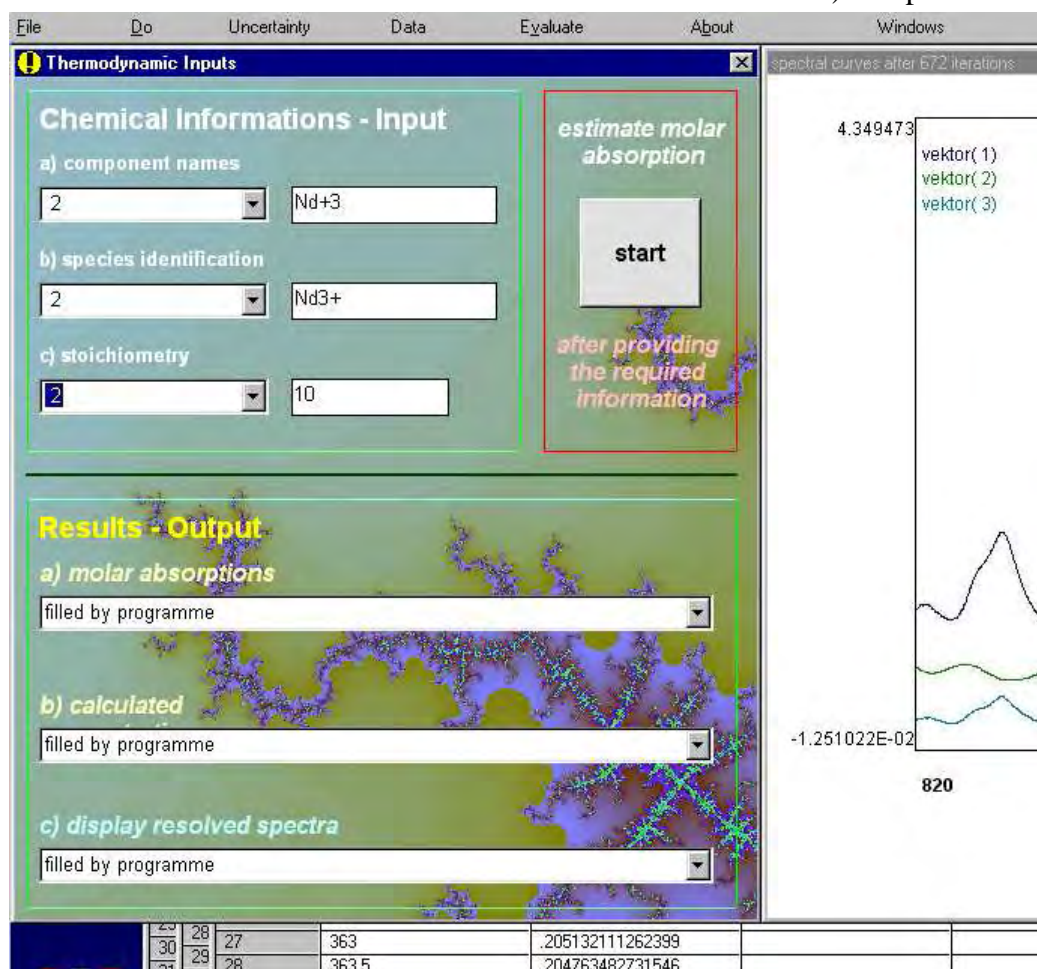


Figure r: The Basic Thermodynamic Information window of the Molar Absorption menu item. Note that the species where the single component spectrum is known (here Nd^{3+}) (cf. the 'species identification' field), always appears in second place!

It is important to give stoichiometric coefficients for all components, even if the species is not holding a component. Having a system with components Fe^{3+} , Cl^- and pH (for OH^-), a species $\text{Fe}(\text{OH})_2^+$ is characterised by the stoichiometric coefficients '102', while FeCl^{2+} is characterised by '110'. If only '11' is given for FeCl^{2+} , CAT will crash quickly.

Upon clicking the right hand side 'GO' button, CAT searches for the best fitting agreement between the metal ion concentration in the species and the total metal ion concentration. Free ligand concentrations are calculated from difference (if absolute concentrations have been given) or takes the free concentrations to calculate formation quotients for each solution. If negative concentrations follow, the calculation is stopped. Sample solutions with negative concentrations will be highlighted in red colour as shown below.

	NdNic000.bkg	NdNic015.bkg	NdNic030.bkg	NdNic045.bkg	NdNic060.bkg	NdNic090.bkg
[Nd(NicNO)]calc	1.775262E-10	2.69348E-05	7.943659E-05	1.288067E-04	1.908852E-04	2.751428E-04
[Nd+3]calc	3.03055E-06	2.743347E-06	3.216998E-06	2.928844E-06	2.518523E-06	4.322782E-06
[Nd+3]meas	4.70394E-03	4.295681E-03	4.21941E-03	4.036007E-03	3.786236E-03	3.610186E-03
[Nd(NicNO)2]calc	5.098715E-05	4.615514E-05	5.412403E-05	4.927602E-05	4.237262E-05	7.272818E-05
[Nd(NicNO)2]meas	5.0782E-04	4.634314E-04	5.318805E-04	7.350508E-04	1.094918E-03	1.327352E-03
[Nd+3]free	5.778433E-05	5.230814E-05	6.133938E-05	5.584508E-05	4.802137E-05	8.242368E-05
[Nd+3]meas	0.005	0.005	0.005	0.005	0.005	0.005
[Nd+3]calc	5.21176E-03	4.786048E-03	4.830727E-03	4.899865E-03	5.07204E-03	5.212681E-03
%	4.063124	-4.470335	-3.504082	-2.043618	1.420327	4.080068
[NicNO]free	1.01664E-03	2.03975E-04	3.568024E-04	6.510918E-04	6.192783E-04	1.570154E-03
lg B(Nd(NicNO))	void	void	1.72234114452085	1.69034431596093	1.91067871453729	1.68608640632186
lg B(Nd(NicNO)2)	void	void	5.99570686491748	5.63308100438717	5.87740219969984	5.17357216984344
rel. weight(B(Nd(NicNO)))	0	0	5.246523E-02	.0981745	.1845818	8.611163E-02
rel. weight(B(Nd(NicNO)2))	0	0	6.290929E-02	.1003292	.1896042	7.439426E-02
	1.82950434694669	± .1503942			5.39630441894693	± .4072311

Figure s: The 'summary of calculated data' window. The system has been interpreted by two coordinated species, termed NdNicNO^{2+} and $\text{Nd}(\text{NicNO})_2^+$. The first solution holds only Nd^{3+} , hence no formation quotient has been calculated. For the second solution, a negative free ligand concentration, $[\text{NicNO}]_{\text{free}}$, is obtained. Hence, the field is highlighted red and the formation quotient holds the word 'void'. The violet fields hold uncertainty estimates on basis of Clifford's method (see References). These uncertainties represent misfit and are solely used to assign a weight to each calculated quantity. The weights of all components of a formation quotient are multiplied. After all weights are calculated, the weights of all non-void samples are normalised and given in the row 'rel. weight' for each species. Given the weights and the formation quotients in each sample solution the formation quotients in the lowest row are calculated together with an uncertainty estimate.

The 'Summary of Calculated Data' window can be saved. If the window has the focus, it is the format to be saved if 'Save' is selected in the 'File' menu. The data in the file are put in an ASCII file which can be evaluated externally.

Further informations may be obtained from the 'Basic Thermodynamic Information' window. Here the 'Results' section is now filled with additional information. The molar absorptions are given in the top section. The middle section allows to inspect the agreement between

specified total metal ion concentration and the calculated sum of metal ion. The difference is given in per cent. In the bottom section, the user can select a spectrum, where experimental spectrum, fitted total spectrum and the contributions of the species as calculated from the given model are specified. This spectrum appears in a new graphics window.

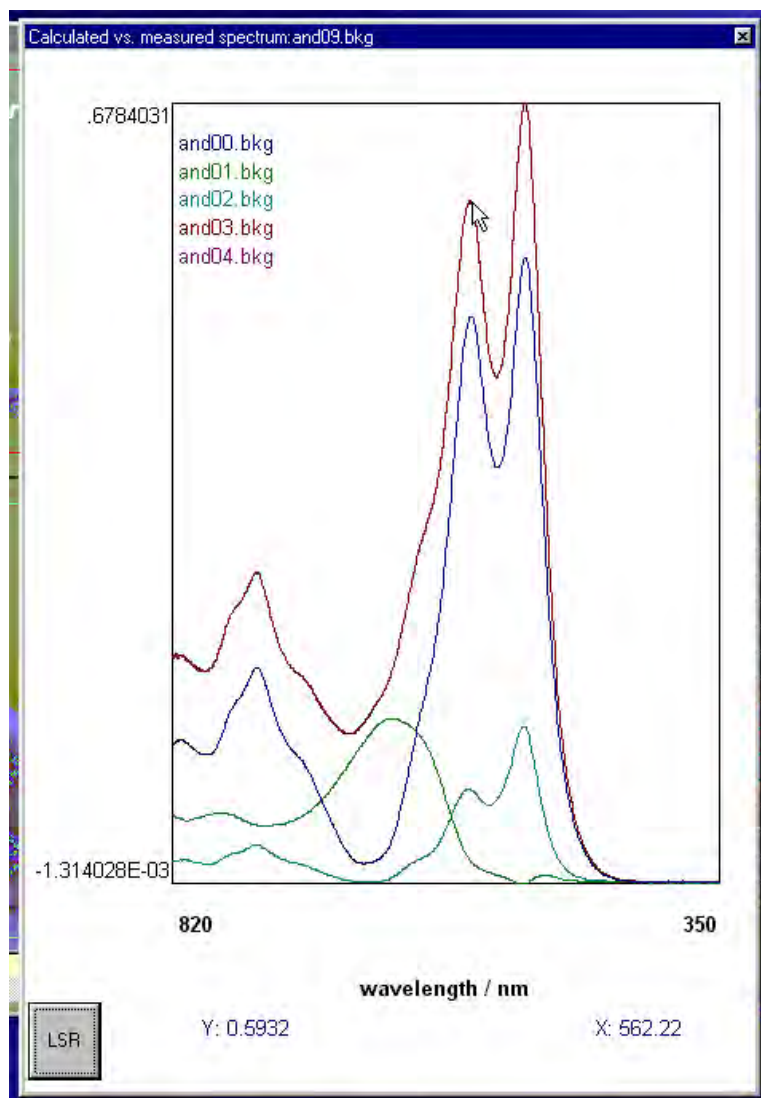


Figure t: Example of an interpreted spectrum and its single species contributions – available from the 'display resolved spectra' field (cf. figure r: under 'Results – Output'). Pointing with the mouse into the diagram allows to read the spectroscopic informations: here the absorption peak is at 562.2 nm with an absorption of 0.593. The button LSR at the lower left hand side allows to display a least squares interpretation. The underlying algorithm is QR decomposition. The calculations may become quite time-consuming but provides the variance-covariance matrix.

The displayed spectrum can be saved in an ASCII file and loaded into an appropriate graphics program. Thus, all spectra can be transferred into a presentation graphics.

(5) TB CAT

The simple CAT procedure returns formation quotient(s) and single component spectra for the system under study. The constraints are the key (the specific permutation of the diagonal elements in the target transformation matrix), the single component spectrum of the free metal ion and the values for the composition of the sample solutions. CAT gives values but without information concerning the stability of the values. As long as these values are considered as results per se, there is no need to inquire into their reliability and stability. Without an understanding of reliability and stability of a result, no further conclusions and comparisons, e.g. with similar values from other sources or separate experimentation, should be made.

The menu items CAT and Molar Absorption evaluate mean value based data interpretations on the assumption that the informations about component concentrations are perfectly true. Hence, if a concentration is given as $0.0002 \text{ mol dm}^{-3}$, the algorithm assumes $0.000200000000... \text{ mol dm}^{-3}$ of that component. However, can the user be sure that the concentration isn't, say, $0.0001998 \text{ mol dm}^{-3}$? Or $0.00020314 \text{ mol dm}^{-3}$? There is almost no certainty. Each volume operation introduces some uncertainty. Mixing samples from several standard (standardised by what? or to what?) solution is prepared with a small but non-negligible uncertainty. Uncertainties accumulate. Further uncertainties, e.g. in spectral residual correlation, add into the procedure.

Threshold bootstrap computer-assisted target factor analysis tries to approach the problem of limited certainty of the knowledge about a system by computer-intensive resampling methods. In fact, TB CAT creates a large number of new input files with input quantities varying within specified limits from run to run. The residuals may vary, the input concentrations may vary, the volumes may vary.

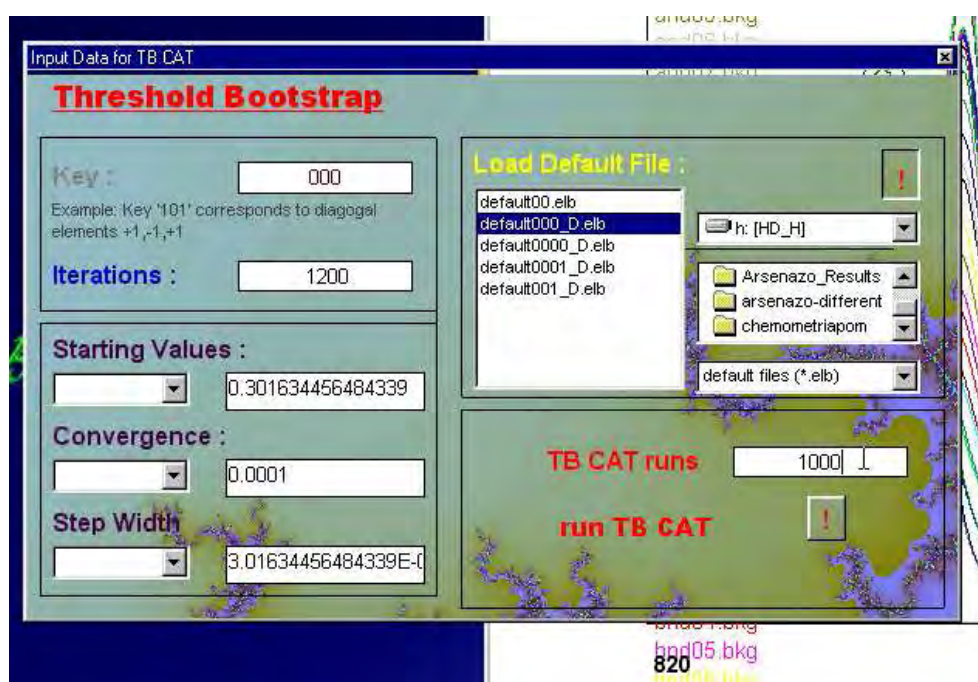


Figure u: The TB CAT input window. The window is the same as figure p, but now the 'TB CAT runs' field is enabled.

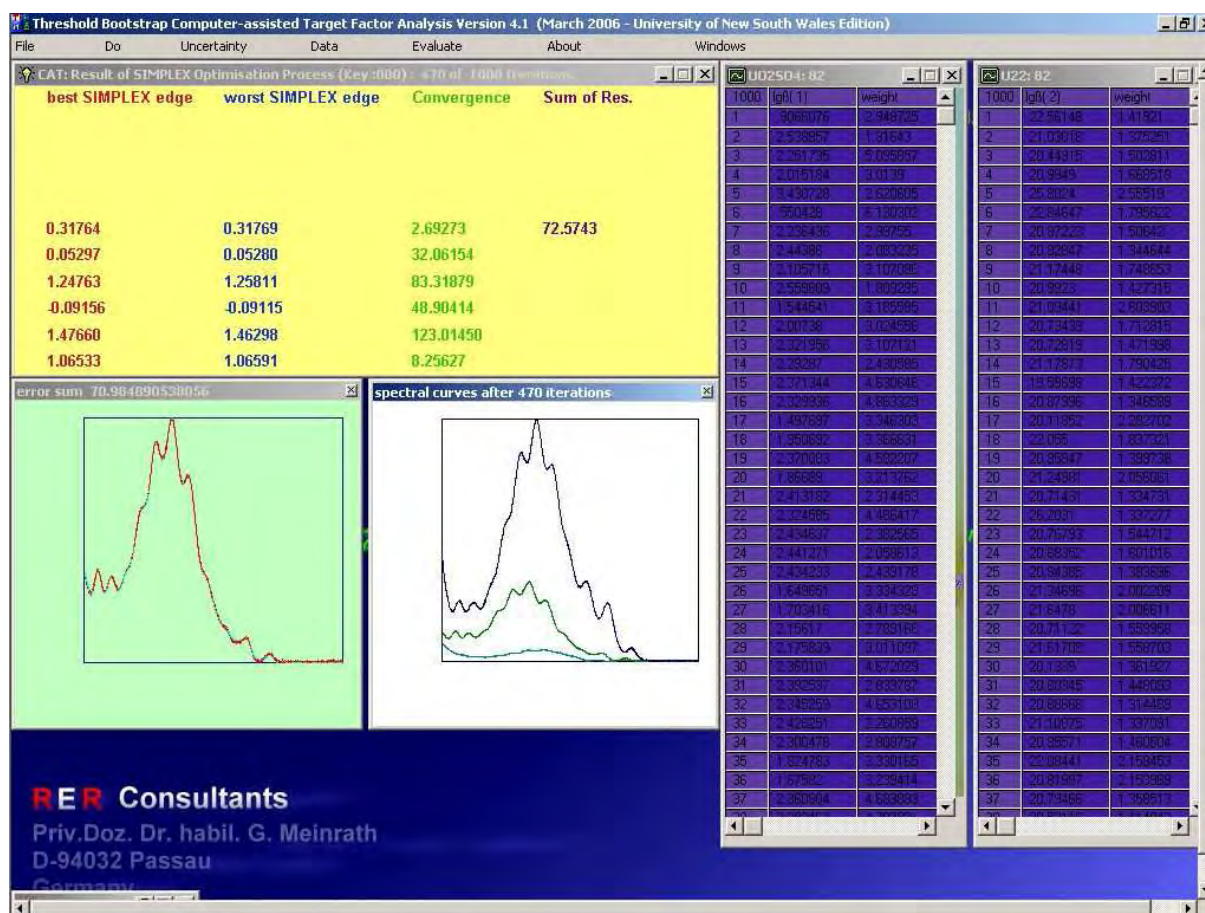


Figure v: TB CAT at work. The formation constants obtained in each TB CAT run are displayed in the two windows at right. CAT automatically generates the required number of result windows, with names taken from the species information given by the user. In the top left field, the total number of runs is given. The current status is shown in the header of each column. Upon termination, TB CAT sorts each column and writes the information to the disk using a naming convention specified below.

The user proceeds as if a CAT would be calculated. In the 'Input Data for CAT' window, the field 'TB CAT runs' is enabled. The user specifies the total replicate runs. All informations are requested by the program as has been shown in the previous section. TB CAT reruns the data the specified amounts of times (at least 1000 TB CAT runs should be performed). For each species, a table is created where the calculated formation constants and a weight factor are tabled. From these tables, probability densities and spectral uncertainties are derived.

TB CAT should procede without further need of user interaction. After having finished the calculations, the list of formation quotients will be sorted and the weights will be transformed into probability densities. At the same time, TB CAT creates a large number of spectral data files in the directory from which the experimental spectra have been loaded. The name convention is as follows:

AAAANNN_B.iii

AAAA: name of the species (from previous user input in the Thermodynamic Data Window)

NNN : key (may have more or less than three positions)

B : number of species (pure spectral component is always no. 2)

iii : no. of TB CAT run

example: "metal ion000_2.123"

Furthermore, a file 'cdf*.dat' is created holding the cumulative probability density for each formation quotient. Example: cdf_species(2)000.dat

From these files, the user may create various probability densities as described in the section 'Evaluate'.

c) Uncertainty

This menu item allows the user to communicate the uncertainties to be associated with relevant influence quantities to the TB CAT routine. This menu item has two fixed elements:

(1) Load/Save

For the convenience of the user, a ASCII file may be either created or loaded holding the relevant values for the uncertainties to be associated with the components.

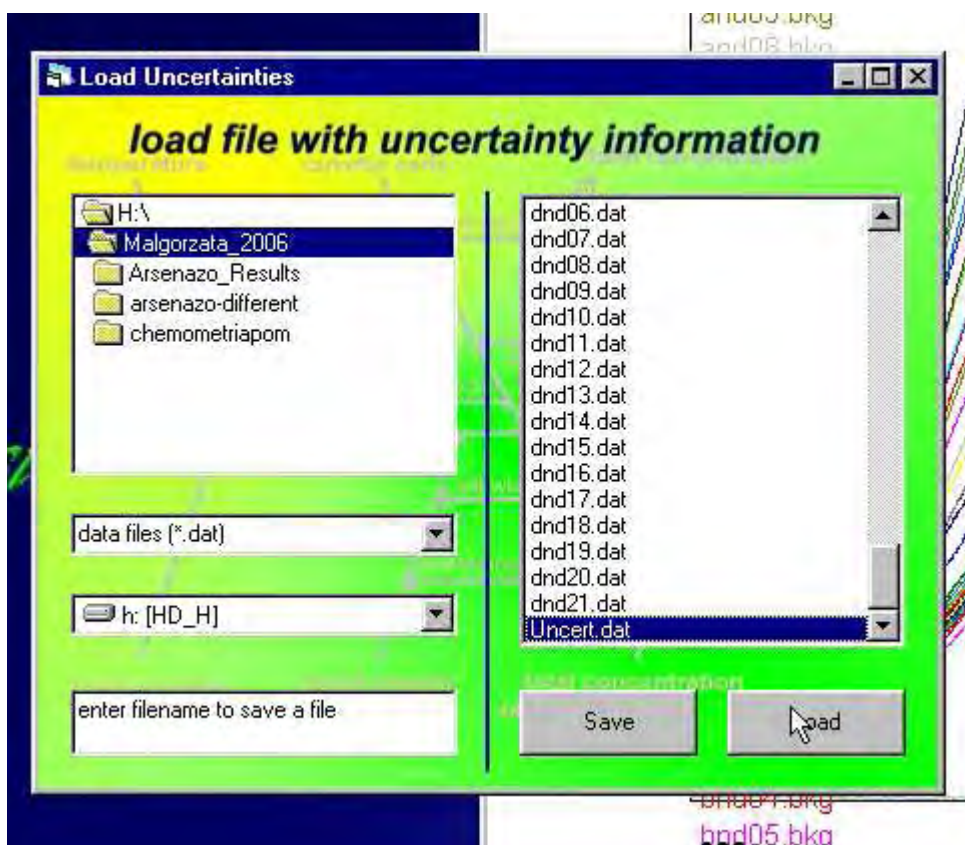


Figure w: Window for loading and saving information on measurement uncertainty.

(2) Repeatability

Upon clicking on this item, an input window will open allowing to specify the repeatability. Repeatability tries to account for random variations in recording spectral files. It is a common observation that the spectral curve of a sample is quite precisely reproducible, but the maximum of the spectrum may show some variability.

The other menu items are created during run-time on basis of the informations given in the file headers of the spectrum files. The component names will appear in the menu. Upon clicking a component name, a window opens allowing to specify uncertainties for volume operation (pipetting), balance (referring to the certainty that the weighted amount also arrives in the sample cuvette) and purity. Note that the uncertainties in these quantities, not the quantity itself must be specified. The purity of common laboratory chemical can be nominally 99.5%. After some time, it is reasonable to assume an uncertainty of 0.5% to 1% in this value.

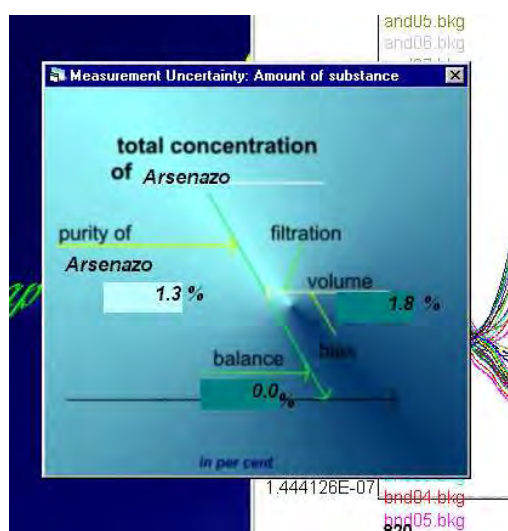


Figure x: An input window for the measurement uncertainty. All information are transfered to the code by upon closing the window.

If a component is pH, a special window opens, allowing to state the gross uncertainty in pH measurement. It is not possible to state the measurement uncertainty for each individual sample.

If all values are given, these may be saved into an ASCII file (see above). TB CAT has default values. Information given in the Uncertainty menu is only used during TB CAT procedures.

(d) Data

Under this menu item those data are listed that can be scrutinized by the user. This includes mainly the information about component concentrations in the different sample solutions and the spectral input data.

Opening the 'sample concentrations' wondow during a TB CAT analysis allows to follow the modifications made by the TB CAT algorithm on the sample concentrations on basis of the uncertainty information provided by the user under the 'Uncertainty' menu item.

(e) Evaluate

The menu item 'Evaluate' has possible selections 'Penalties', 'list Eigenvalues', 'plot Eigenvectors', 'Spectral Uncertainty' and 'Differentiate'

(1) Penalties

In searching the best target transformation matrix T , CAT relies on the common least-sum-of-squared-residuals criterion. However, there are several criteria which CAT must satisfy simultaneously. First, an optimum agreement between a known spectrum and a calculated spectrum is desirable. Second, the number of negative values in the estimated single component spectra and the species concentrations is strived for. Third, the difference between the measured spectra and those calculated from the numerical procedure should be a minimum. CAT and TB CAT balance these three components of the total sum-of-residuals value by three values, termed 'penalties'. Upon clicking the respective menu item, the current values can be modified. There never has been any need to do so.

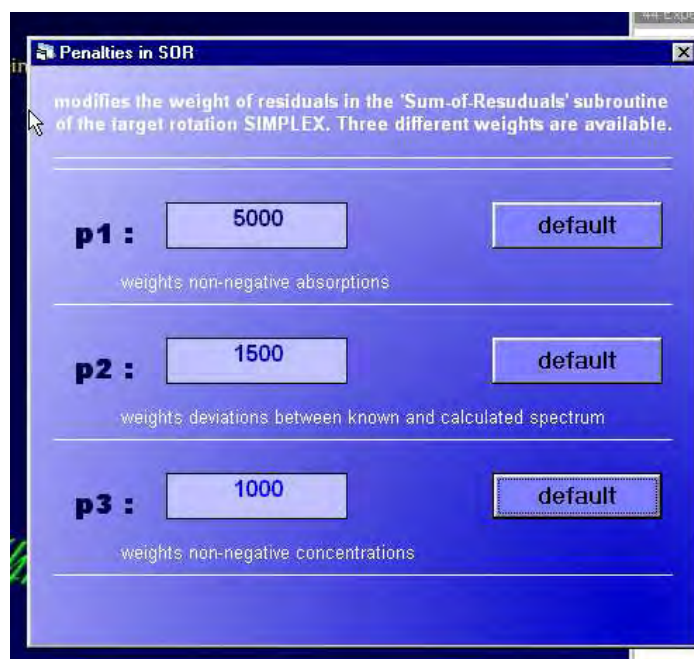


Figure y: The Penalty window. The default values have never been modified in practice.

(2) list Eigenvalues

Clicking this menu item provides a ordered list holding the singular values obtained from the abstract factor analysis step of CAT. Putting the focus to this window allows to save the list as an ASCII file.

(3) Plot Eigenvectors

Clicking this menu item plots the column eigenvectors into a diagram. Setting the focus to this diagram allows to save the graphical data into an ASCII file for subsequent analysis or processing by an external graphics program.

(4) Spectral Uncertainty

The spectral uncertainty window displays the first single component files (*.1) of the species available. These files are created during the TB CAT procedure. The user may choose four different confidence limits. If the respective selections have been made, TB CAT starts to generate a cumulative distribution function at each(!) wavelength. This process takes some time. If the process has finished, it is indicated in the title bar ('CDF done').

The generated ASCII file has the following name convention:

CDF_UVAAAANN_B.dat

example: CDF_UVspecies(1)000_1.dat.

The letter code is the same as above. These files may be loaded into any graphics program and manipulated.



Figure z: Spectral Uncertainty window

(5) Differentiate

The menu item 'differentiate' transforms a cumulative distribution function of a formation constant into a probability density. A cumulative density function is by default saved as

`cdf_AAAANNN.dat`.

If a file is selected, the GO button starts to transfer the cumulative distribution data in file `cdf_AAAANNN.dat` into a probability density ASCII file `dif_AAAANNN.dat`. For this purpose a stepwise weighted linear regression algorithm (LOESS smoother) is used. Example: `dif_species(2)000.dat`.

A window opens showing the differentiated curve. However, the display may be misleading. To judge the probability density curve, it is necessary to open the data file in an external graphics program.

(f) About

opens a window giving some informations about the code. It mainly holds information on the person to which the respective code was personalized. It is understood that this code is NOT distributed without the explicite consent of the copyright owner.

The major practical purpose of this menu item however is to stop CAT temporarily. CAT is programmed to make full use of the computer resources. It doesn't like to share the CPU time. On the other hand, a TB CAT evaluation of, say 35, spectra with, say, three species may take several hours up to several days (large spectral data sets and slow CPU). To temporarily stop CAT, just open the 'About' window. All activity stops while the program waits to close this window. After closing it, CAT proceeds.

(g) Windows

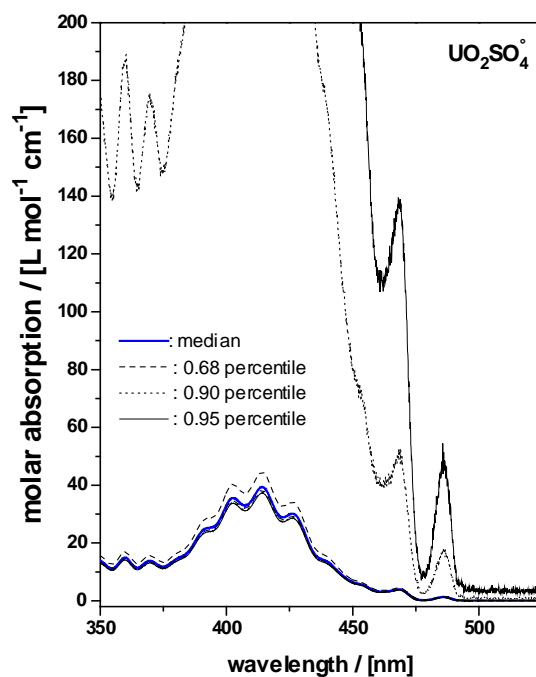
The Window list is a special item of the VB programming environment. Under this menu item all currently accessible windows are listed and can be conveniently displayed even if a certain window should be covered completely by other windows.

A Summary of Recommendations

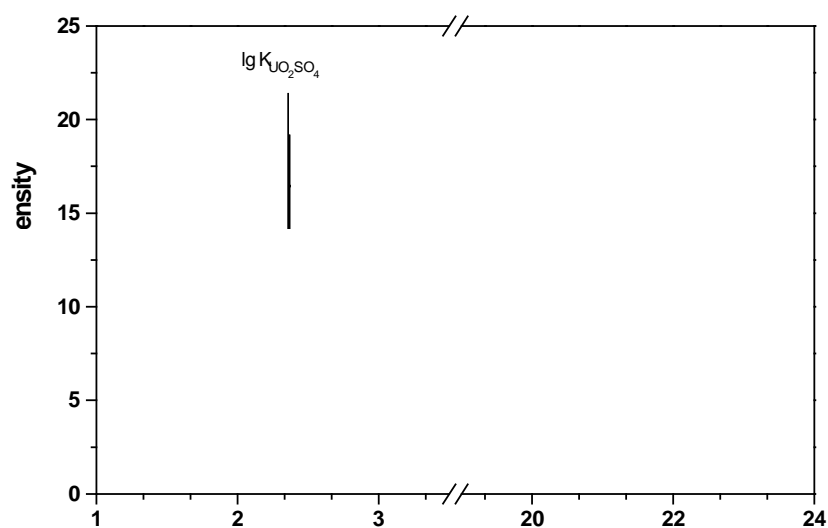
To get happy with TBCAT_S, the following suggestions may be helpful.

- 1) Do not expect to get a perfect code. TBCAT_S somehow runs – but it also crashes often. Often it is helpful just to restart the procedure.
- 2) Allow at least 24 hours before you engage into physical action against your computer.
- 3) After each procedure (Background correction, Analyze, CAT and TBCAT), save the results and restart the program. Some efforts have been made to allow just a modification in the 'basic thermodynamic data' window and restart the fitting from the same window. In this procedure the likely chemical composition of different species is commonly tested. It works often – but likes to fail in other cases. Be patient. At this stage of the analysis, where different species models must be tested to assess the likely meaning of the numerical data, the chemist's intuition and knowledge is required. No computer can replace that (fortunately). If the chemistry of the system is unclear, manual documentation of the already tested species models is required because TB CAT does not store or otherwise documents the user input.
- 4) An advantage of factor analysis is the large number of data to be handled simultaneously. Please remind that TBCAT_S, especially in the TBCAT subroutine, generates huge amount of data. Visual Basic commonly efficiently cares for the memory – but the memory handling of its operation system has several weak points. The "unexpected error" message has often been observed.
- 5) Familiarize yourself with the program on basis of synthetic data sets or simple spectroscopic systems. CAT has handled already quite complex systems with up to seven components. But experience shows that such systems should be split into smaller units to reduce correlation between the spectra.
- 6) Remind the input conventions: There must be always one component whose single component spectrum is known. This spectrum must be loaded in first place.
- 7) TBCAT_S is designed to resolve spectra of metal ions in solutions with ligands. Therefore, it is assumed that the free metal ion's spectrum is the known component. But TBCAT_S also works if the known spectrum belongs to a ligand, e.g. arsenazo III. Then, the absorption spectrum of arsenazo III takes the role of the known component and all input has to be modified accordingly. But there is no fundamental problem for CAT to evaluate such systems. A limitation is the situation where two components absorb. CAT cannot resolve this situation. A feasible way is to limit the analysis to wavelength ranges where only one component absorbs.
- 8) TBCAT_S is work-in-progress. Feel free to make suggestions.

Example results



Example 1: Single component spectrum of $\text{UO}_2\text{SO}_4^\circ$. The blue curve gives the median. Note the extreme values for the upper 0.90 and 0.95 percentile spectral curves. The underlying distributions are highly non-Normal and skewed.



Example 2: Probability densities for formation constants of solution species.

References

There are a large number of publications having been influential in developing TBCAT_S. The following small selection is meant as a starter. The methods described in these references are of a general interest for the application of computers in chemistry.

M.S. Caceci "Estimating Error Limits in Parametric Curve Fitting". Anal. Chem. 61 (1989), 2324

M.S. Caceci, W.P. Cacheris; "Fitting Curves to Data: The SIMPLEX Algorithm is the Answer". Byte (1984) 340

A.A. Clifford, "Multivariate Error Analysis". Applied Science, London/UK (1973)

G.H. Golub, C. Reinsch, "Singular Value Decomposition and Least Squares Solutions". Numer. Math. 14 (1970) 403

J.A. Nelder, R. Mead, "A SIMPLEX Method for Function Minimization". Computer J. 7 (1965) 308

J.C. Nash, "Compact Numerical Methods for Computers". Adam Hilger Bristol/UK (1981)

G.E.P. Box, M.E. Muller, "A Note on the Generation of Random Normal Deviates". Ann. Math. Stat. 29 (1958) 610

B. Efron, "Computers and the Theory of Statistics: Thinking the Unthinkable" SIAM Review 21 (1979) 460

B. Efron, R. Tibshirani, "Statistical Analysis in the Computer Age" Science 253 (1991) 390

P.J. Rousseeuw, B.C. van Zomeren, "Unmasking Multivariate Outliers and Leverage Points" J. Am. Stat. Assoc. 85 (1990) 633

General comments

When we got access to the draft of the NEA-TDB review on the thermodynamics of Np and Pu, of course, our first interest was directed towards the review of our extensive work in the field of the thermodynamics of Np(V). Therefore most of the comments worked out in the present manuscript concern the following publications: the RCM and KfK reports (in German) [91KIM/KLE, 94RUN/KIM] and [94NEC/KIM], respectively, the papers [94NEC/RUN, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC, 96FAN/NEC, 97NEC/FAN], and two papers of our former coworkers G. Meinrath [94MEI] and W. Runde [96RUN/NEU].

We noticed very soon that the reviewer of this section (P. Vitorge) favors the studies from his own laboratory in an incredible way and disregards results from other authors, in particular those of our group. The reviewer takes each opportunity to discredit the papers from “Kim’s group”. Two of the most often cited pretended “shortcomings” in our papers are “problems with pH electrode calibration” and “inconsistencies” among our results. In order to achieve an objective judgement of our pH measurements, we would appreciate a second reviewer to read the paper [96FAN/NEC], which clearly shows that in all the papers mentioned above, correct measurements were made to determine the quantities of interest, i.e. the concentrations of H^+ , OH^- and CO_3^{2-} . Some of the arguments documented in this paper are also mentioned in the present manuscript. More detailed information could be available from our earlier report [Neck, Fanghänel, Kim, FZKA 5599 (1995)] or by personal communication.

The second argument (“inconsistencies” among our results) is exclusively based on untrue statements, on obscure calculations of the reviewer himself, leading to nonsensical data and conclusions, which were never published or stated in our papers. In some of the “reviews” important information from our papers is just omitted, not even mentioned. We have the strong impression that some of the reviewer’s mistakes or incorrect statements were not accidentally, but on purpose, with the intention to discredit our work. Possibly, the same “review style” was applied as well on papers of other authors - we did not have the time to check this.

In order to demonstrate these severe incriminations, some unambiguous examples will be shown in this manuscript. We have already recognized numerous further examples, including obvious mistakes of the reviewer or ridiculous recalculations concerning papers, which we know very well. However, we did not yet have the time to work out all these examples.

Moreover, the review does not represent the present state of the art. At the end of this

manuscript, we will point out some problems, which complicate the understanding and the interpretation of the available experimental data in the Np(V) carbonate systems. Although these problems were already discussed in papers of Kim and coworkers [95NEC/FAN, 95FAN/NEC, 96FAN/NEC] and also by other authors [95NOV/ROB, 96RUN/KIM, 97NOV/ALM, 98ALM/NOV], the reviewer might not have understood these papers. Some of these problems are closely related to the shortcomings and simplifications of the SIT approach used in the NEA-TDB reviews for the calculation of activity coefficients. This concerns also auxiliary data used in the NEA-TDB for the calculation of carbonate ion trace activity coefficients in NaClO₄ solutions.

Example 1

This first example demonstrates the attempt of the reviewer (P. Vitorge) to discredit our work by claiming that there are inconsistencies and the published data are therefore unacceptable. This statement is repeated several times throughout the review. The example concerns aqueous Np(V) carbonate complexes, and our measurements with two different methods differ 0.2 log units, which is in general considered as fair agreement and not as inconsistency.

In the review of [91KIM/KLE] (Appendix A), the reviewer writes:

page 808, lines 1 - 7:

The spectra of Np(V) carbonate complexes were also reported and were used later [94NEC/RUN] to calculate the formation constants of two Np(V) carbonate complexes. The spectra were in accord with those previously found by Riglet [90RIG]. The reported value of β_1 was not consistent (within the uncertainties estimated by the authors) with the value measured by the solubility technique, and the same was true for the β_2 values reported in [94NEC/RUN]. This inconsistency was pointed out by the authors, but was not explained.

An almost identical comment is given in section 12.1.2 (page 255, lines 1 - 7) the reviewer then concludes:

page 255, lines 7 - 10:

There appears to have been a problem with this work, but there is not enough experimental information to reinterpret these spectrophotometric measurements. These values are not used in the present review in selection of the values for the complexation constants.

In order to emphasize the “miserable” work in [94NEC/RUN], the reviewer repeats in the review of paper [96CLA/CON] (Appendix A):

page 841, lines 29-32:

.... possibly as described by Neck et al. [94NEC/RUN] (However, the spectrophotometric results of that publication were not particularly accurate and were not used in the final data assessment in the present review. See the discussion of [94NEC/RUN])

Thus criticized, the reader must expect strong inconsistencies and inaccuracies among the $\log \beta_1$ and $\log \beta_2$ values reported in [94NEC/RUN] for 0.1 M NaClO₄ solution. Actually the review concerns the following data:

solubility:	$\log K_1 = 4.58 \pm 0.04$	$\log K_2 = 6.60 \pm 0.07$	[94NEC/RUN]
spectroscopy:	$\log K_1 = 4.38 \pm 0.04$	$\log K_2 = 6.4 \pm 0.3$	[94NEC/RUN]

Indeed, the $\log K_1$ values do not agree within the underestimated range of uncertainty. However, readers with experimental experience in the field of aquatic actinide chemistry might consider differences of 0.2 log units (for two different experimental methods) as fair agreement, not as inconsistencies. In order to prevent these readers from doing so, the reviewer does not give the numerical values for comparison, well knowing that only few readers will look for these values in the original paper or in the comprehensive Table 12.4 (p.245 - 251).

On the other hand, it is somewhat surprising that there is no corresponding critical comment on the solubility and spectroscopic studies of Vitorge and coworkers [86GRE/ROB, 90RIG], which include 2 - 3 times larger inconsistencies (c.f. Table 12.4 and Figs. 12.2 - 12.4), e.g. for the consecutive complexation equilibrium $\text{NpO}_2(\text{CO}_3)_2^{3-} + \text{CO}_3^{2-} \rightleftharpoons \text{NpO}_2(\text{CO}_3)_3^{5-}$ in 3 M NaClO_4 (from Table 12.4, p.247):

solubility:	$\log K_3 = 2.31 \pm 0.14$	[86GRE/ROB]
spectroscopy:	$\log K_3 = 2.9 \pm 0.2$	[90RIG]

Quite in contrast to the disregarded “inaccurate” data reported in [94NEC/RUN], the results in [86GRE/ROB, 90RIG] are considered as most reliable and selected for the evaluation of ion interaction coefficients and equilibrium constants at $I = 0$. This reflects the reviewer’s (Vitorge’s) very personal interpretation of the guidelines for NEA-TDB reviews.

Example 2

Np(V) carbonate complexes; statement on page 256

page 256, lines 27 - 29:

Two different values of K°_3 (based on the same data) were reported by Kim et al [91KIM/KLE, 94NEC/KIM], and no explanation was offered for this inconsistency.

This short sentence, which is again nothing but an attempt to discredit our work, includes the following incorrect citations or statements:

- 1.) We never evaluated consecutive constants $\log K_n$, neither from our solubility data in [91KIM/KLE, 94NEC/KIM, 94NEC/RUN] nor in our later review paper [95FAN/NEC], but formation constants $\log \beta_n$. (Since there are generally enough and accurate data in the pH-range, where NpO_2^+ is the predominant species, it is much more convenient to evaluate the solubility product and $\log \beta_n$ instead of $\log K_n$ values - analog to the methodology applied in the NEA review on the U(VI)-carbonate system [92GRE/FUG])
- 2.) The cited report [94NEC/KIM] does not include any extrapolation to $I = 0$, but only experimental data in 5 M NaCl and in 5 M NaClO₄. Extrapolations to $I = 0$, either with the SIT or Pitzer equations, are given in other papers [94NEC/RUN, 95FAN/NEC].
- 3.) The $\log \beta_n$ values evaluated in [91KIM/KLE, 94NEC/RUN, 95FAN/NEC] are not based on the same data as stated by the reviewer. In all our papers, it is well noted, which data are used for the evaluation of constants at $I = 0$ (c.f. table below). And it is most trivial that a different set of experimental data consequently leads to different values at $I = 0$. Further, it is well-known, that particularly for highly charged ions like $\text{NpO}_2(\text{CO}_3)_3^{5-}$, SIT and Pitzer equations may give different activity coefficients in the very low ionic strength range.
- 4.) In order to obtain an impression on the magnitude of the “inconsistencies” (which are of course not given numerically by Vitorge), the evaluated constants $\log \beta_n$ are listed in the table below. Moreover, from the comparison with the values selected by Vitorge, the following question arises: If all the papers of the “Kim group” are that inconsistent and contain so many mistakes, how is it then possible that Vitorge selects the same constants (within the range of uncertainty of those given in [91KIM/KLE, 94NEC/RUN, 95FAN/NEC]) ?

Table 1. Formation constants $\log \beta_n$ of Np(V) carbonate complexes

Ref.	Method	$\log \beta_1$	$\log \beta_2$	$\log \beta_3$	exp. data used
[91KIM/KLE]	SIT	5.04 ± 0.06	6.59 ± 0.09	5.73 ± 0.17	NaClO_4 ^{a)}
[94NEC/RUN, 95FAN/NEC]	SIT	4.83 ± 0.15	6.55 ± 0.23	5.54 ± 0.09	NaClO_4 ^{b)} ;
[95FAN/NEC]	Pitzer	5.03 ± 0.06	6.47 ± 0.14	5.37 ± 0.36	NaClO_4 + NaCl ^{c)}
Vitorge's review	SIT	4.96 ± 0.06	6.53 ± 0.10	5.50 ± 0.15	NaClO_4

^{a)} exp. data (0.1 - 3.5 m NaClO_4) from [91KIM/KLE]

^{b)} exp. data (0.1 - 3.5 m NaClO_4) from [83MAY, 86GRE/ROB, 85BID/TAN, 85INO/TOC, 90NIT/STA, 90RIG, 91KIM/KLE, 94NEC/RUN, 94MEI]

^{c)} exp. data (0.1 - 6.5 m NaClO_4 and 0.1 - 5.6 m NaCl) from [83MAY, 86GRE/ROB, 85BID/TAN, 85INO/TOC, 90NIT/STA, 90RIG, 91KIM/KLE, 94NEC/RUN, 94MEI, 94NEC/KIM, 94RUN/KIM]

Example 3

Mixed Np(V) hydroxide-carbonate complexes

page 241, lines 3 - 7:

Varlashkin, Begun and Hobart [84VAR/BEG] reported ... These results were later confirmed by Riglet [90RIG] and Vitorge and Capdevila [98VIT/CAP], and interpreted as evidence of the formation of mixed Np(V) hydroxide-carbonate complexes.

In this section the reviewer does not mention that, extending the former qualitative studies, we identified and quantified two ternary complexes, $\text{NpO}_2(\text{OH})(\text{CO}_3)_2^{4-}$ and $\text{NpO}_2(\text{OH})_2(\text{CO}_3)^{3-}$ [97NEC/FAN]. Our paper is cited later on page 270 and reviewed in Appendix A (discussion of selected references). However, neither the formation constant for the complex $\text{NpO}_2(\text{OH})_2(\text{CO}_3)^{3-}$ evaluated from the absorption spectra nor the estimate given for $\text{NpO}_2(\text{OH})(\text{CO}_3)_2^{4-}$ are mentioned. The following comments are given instead:

page 270, lines 1 - 4:

Neck, Fanghänel and Kim [97NEC/FAN] also recently reported a very similar spectrophotometric study of the dissociation of the carbonate limiting complex in alkaline media; but for the reasons explained in Appendix A, the present review does not rely on their conclusions.

Appendix A; page 850, line 37 - page 851, line 4:

The predominance diagram shown in this paper (Figure 8 of [97NEC/FAN]) indicated that $\text{NpO}_2(\text{OH})(\text{CO}_3)^{2-}$ was less than one per cent for the experimental condition of the measurements. This is below the detection limit of their experimental method, and well below the concentration needed to determine its stoichiometry. This is a standard error when stoichiometries are determined only from curve fitting;

Indeed, it would have been most nonsensical, to evaluate the formation constant for the complex $\text{NpO}_2(\text{OH})(\text{CO}_3)^{2-}$ from the recorded absorption spectra. However, in our paper ([97NEC/FAN], p.173) it is very clearly pointed out that this complex has not been observed and the constant given for this complex is not derived from absorption spectra, but estimated by interpolation between the known constants in the system Np(V)-OH-CO_3 in 3 M $\text{Na(OH/CO}_3/\text{ClO}_4)$.

The impertinent intention of the reviewer is evident: The reader of this NEA review, who does not read the original paper [97NEC/FAN], should conclude that Neck, Fanghänel and Kim are as stupid to evaluate formation constants from spectroscopic data, even though the contribution of the corresponding species is less than 1 %.

Similarly, the reviewer gives another completely incorrect citation:

page 270, lines 17 - 21:

As suggested by Neck, Fanghänel and Kim [97NEC/FAN] in the qualitative discussion of their spectrophotometric results, formation of $\text{NpO}_2(\text{OH})_2^-$ and precipitation of a hydroxide compound might be sufficient to explain most of the experimental observations in more alkaline media

No such statement is given in [97NEC/FAN] ! Just the opposite is the case. It would be absolutely erroneous to assume the formation of $\text{NpO}_2(\text{OH})_2^-$ under the conditions of our study, i.e. at carbonate concentrations > 0.01 M (c.f. speciation scheme Fig.8 in [97NEC/FAN])

Example 4

Aqueous Np(V) chlorides. In the discussion of aqueous Np(V) chlorides (section 9.2.2.3), the reviewer writes:

page 179, line 34 - page 181, line 2:

Neck, Kim and Kanellakopulos [94NEC/KIM] reported formation constants for the 1:1 and 1:2 complexes, but the ionic strength is high ($I = 5\text{ M}$) and in the absence of precise estimates of the interaction coefficients, the error in the extrapolation to $I = 0$ will be dominated by the error of our estimates of the ionic interactions. Giffaut [94GIF] observed no significant changes in the visible absorption spectrum of NpO_2^+ in 1 M HClO_4 and 4 M NaCl . Hence it would be highly speculative to make a selection on the basis of the existing data.

The studies at $I = 5\text{ M}$ (in the cited report [94NEC/KIM]) are extended to $I = 1, 2$ and 3 M by analogous experiments in [95NEC/FAN]. This paper also includes the SIT extrapolation to $I = 0$. All these data are neither cited in the text nor in Table 9.7 (p.181). Further, the paper [95NEC/FAN] shows absorption spectra in $0, 1, 2, 3$ and 5 M NaCl and in $5, 8,$ and 10 M LiCl (all at $\text{pH } 3$), i.e. it provides much more spectroscopic information than the paper cited in the review. Finally, the conclusion is given in [95NEC/FAN] (already in the abstract), that there is no evidence for the formation of Np(V) inner sphere chloro complexes, and that the interaction between NpO_2^+ and Cl^- ions should be considered as strong ion-ion interaction without invoking chloro complexes. The quantification of this problem is also reported in [95NEC/FAN]. This paper is well known and the results are frequently cited in the literature.

It appears somewhat strange, that the paper [95NEC/FAN] is not mentioned in the section 9.2.2.3 (Aqueous Np(V) chlorides), although it is cited several times in other sections. The “review“ in Appendix A (discussion of selected references) gives rise to the assumption that the reviewer did not even read this paper. It is restricted to the short statement below, referring to the review of other papers, in which the results discussed above are not included:

page 836, lines 10-11:

[95NEC/FAN] See the discussion of [91KIM/KLE, 94NEC/KIM, 94NEC/RUN] in this appendix.

Example 5

Solubility of Np(V) carbonates - comment on page 285

page 285, lines 23 - 25

Kim and co-workers [91KIM/KLE, 94MEI, 94NEC/KIM, 94NEC/RUN, 95FAN/NEC, 95NEC/FAN, 95NEC/RUN] reported measurements, that are not in accord with previous work.

This statement gives the impression that our work differs largely from the results of other authors. A closer look demonstrates that the “disagreement“ is generally very small (~ 0.2 log units, which is within the range of usual experimental uncertainties), although the data from different authors were obtained under different conditions. In the review of [91KIM/KLE] the reviewer writes:

page 826, lines 23 - 27

The solubility values in 1 M NaClO₄ aqueous solutions were the same as those in [83MAY], and this may be coincidental ... For 3 M NaClO₄ aqueous solutions the results are clearly different from those of Grenthe, Robouch and Vitorge [86GRE/ROB]. The difference is too large to have resulted from pH electrode calibration problems discussed above, but might have resulted from different solid phases being present in the two studies.

In the figure on the next page (taken from [95FAN/NEC] and [86GRE/ROB]), the solubility data of Kim and co-workers are shown in comparison with previous and later work of other authors. Such a direct comparison of solubility data, which would certainly be more helpful for the reader than the reviewer’s incorrect and contradictory statements, is not shown in the review.

The data shown in the figure below refer to the following papers and experimental conditions:

exp. conditions	pCO ₂ = 10 ^{-3.5} atm	pCO ₂ = 0.01 atm or batch experiments
0.1 M NaClO ₄	[91KIM/KLE, 94NEC/RUN]	[94MEI]
1 M NaClO ₄	[91KIM/KLE, 94NEC/RUN]	[83MAY]
3 M NaClO ₄	[91KIM/KLE, 94NEC/RUN]	[86GRE/VIT]
5 M NaCl	[94NEC/KIM]	[94RUN/KIM]

Example 6

Solubility of Np(V) carbonates - review of [94MEI] in Appendix A

page 829, lines 21 - 28 (review of [94MEI] in Appendix A)

As in a previous publication from Kim's laboratory [91KIM/KLE], Meinrath reproduced Vitorge's measurements ... Since original solubility values were reported in the earlier report [91KIM/KLE], and not in the later publication [94MEI], no attempt was made to evaluate Meinrath's results in the present review.

It is to note that Meinrath was not involved in the Np(V) solubility studies performed in Kim's laboratories (c.f. authors' names of the corresponding papers). The solubility study published in [94MEI] was done independently, at the Japan Atomic Energy Research Institute. And the question must be allowed: How can Meinrath's experiments in 0.1 M NaClO₄ reproduce Vitorge's measurements in 3 M NaClO₄?

These comments are absolutely inadequate and reveal the reviewer's ridiculous vanity. They clearly demonstrate the reviewer's disinterest to do an objective scientific review:

Example 7

In the review of the paper [94NEC/RUN] in Appendix A, it is stated:

page 830, line 32 - page 831, line 2

In this publication Neck et al. added some new solubility measurements (in 5 M NaCl-carbonate aqueous solutions) to those that had appeared in previous reports ...

The presentation of new experimental results included figures showing the transformation of the initial hydrated phase, NaNpO₂CO₃(s) into a Na₃NpO₂(CO₃)₂(s) phase.

Unfortunately the reviewed paper [94NEC/RUN] does neither contain solubility data in NaCl solutions nor solubility data for Na₃NpO₂(CO₃)₂(s). These results are published in the paper [95NEC/RUN], which is however not listed in Appendix A (discussion of selected references).

Example 8

Table 12.4.

Besides wrong citations, e.g. of uncertainties in studies, where originally no uncertainties are given [86GRE/VIT] or vice versa [96RUN/NEU], Table 12.4 contains erroneous citations of experimental data or references (Ref. [91KIM/KLE] does not include data in 3 M NaCl), abstruse recalculations, which change the original published data by more than 0.5 log units (and for some of these recalculations no plausible reasons is given, neither in section 12.1.2 nor in Appendix A). Further, an incorrect citation is found for the solubility product of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O(s)}$:

Table 12.4, page 248

5 M NaCl $\log_{10} K(I_c) = -9.52^{(h)}$ [94NEC/KIM]

^(h) corrected for chloride complexation (p. 246, bottom).

Actually this is a direct experimental value, which is by no means corrected for chloride complexation. May be, the notation ^(h) refers to the value of $\log_{10} K(I_c) = -10.54$ recalculated by the reviewer. However if so, the question would arise, which chloride complexation constants are used for this correction. In section 9.2.2.3 (Aqueous Np(V) chlorides), there was a clear statement that no chloride complexation constant could be selected.

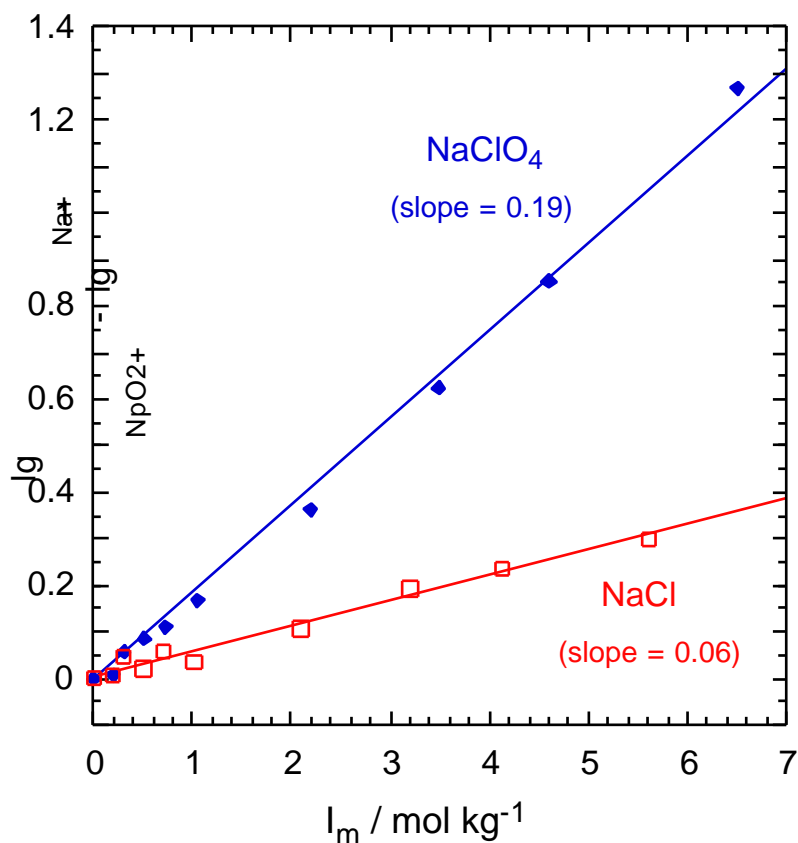
These are just a few examples, which we could recognize at once. Possibly Table 12.4 contains more mistakes.

Example 9

Ion interaction coefficients for the NpO_2^+ ion

In the report [94NEC/KIM], we published solvent extraction studies, from which the activity coefficient ratio $(\text{NpO}_2^+)/(\text{Na}^+)$ in 0.2 - 5 M NaCl and NaClO_4 solutions was determined very accurately. The same results were reported again in the paper [95NEC/FAN], and used to evaluate ion interaction Pitzer coefficients. These Pitzer coefficients were transformed by Vitorge into SIT coefficients (Table A24, p.832 and review of [96RUN/NEU], p.849-850). The more direct and hence more reasonable way to determine $(\text{NpO}_2^+/\text{Cl}^-) - (\text{Na}^+/\text{Cl}^-)$ and $(\text{NpO}_2^+/\text{ClO}_4^-) - (\text{Na}^+/\text{ClO}_4^-)$ is shown in the figure below. With the known interaction coefficients for Na^+ (from the NEA-TDB), those for NpO_2^+ can be calculated:

$$(\text{NpO}_2^+/\text{Cl}^-) = 0.09 \pm 0.02 \quad \text{and} \quad (\text{NpO}_2^+/\text{ClO}_4^-) = 0.20 \pm 0.03$$



SIT plot of $\log \left\{ \frac{(\text{NpO}_2^+)}{(\text{Na}^+)} \right\} = (\text{NpO}_2^+/\text{X}^-) - (\text{Na}^+/\text{X}^-)$, for $\text{X}^- = \text{ClO}_4^-$ and Cl^-

These very important experimental data were “reviewed” (and of course not accepted by the reviewer, because they are from Kim and coworkers) with the following sentences:

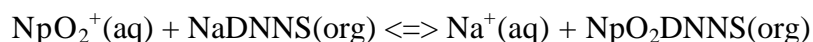
page 830, lines 21 - 26 (Appendix A, review of [94NEC/KIM])

Experimental data of Np(V) liquid-liquid extraction by NaDNNS from 0.2 - 5 M Na(ClO₄, Cl) aqueous carbonate solutions were also tabulated in this report. These were used by Neck, Kim and Kanellakopulos to estimate Np(V) activity coefficients, however there are difficulties with the interpretation of the authors as explained in the discussion of [94NEC/RUN] (work based on the same experimental data of the same group).

page 831, lines 27 - 29 (Appendix A, review of [94NEC/RUN])

The interpretation of the liquid-liquid extraction study neglected activity coefficient corrections for the organic phase. These corrections and systematic errors were quite important ...

Both sentences demonstrate the reviewer’s superficial work and his intention to discredit our work. First of all the studies were of course not performed in aqueous carbonate solutions, which would have been nonsensical, but at constant pH 3, where there is neither hydrolysis nor carbonate complexation. This is clearly documented in both papers [94NEC/KIM, 95NEC/FAN]. Within the liquid-liquid extraction equilibrium studies:



the composition of the organic phase was kept constant, in order to keep constant the activity coefficients of the neutral species in the organic phase. They do not play any role for the evaluation of Pitzer or SIT coefficients for NpO_2^+ in the aqueous phases.

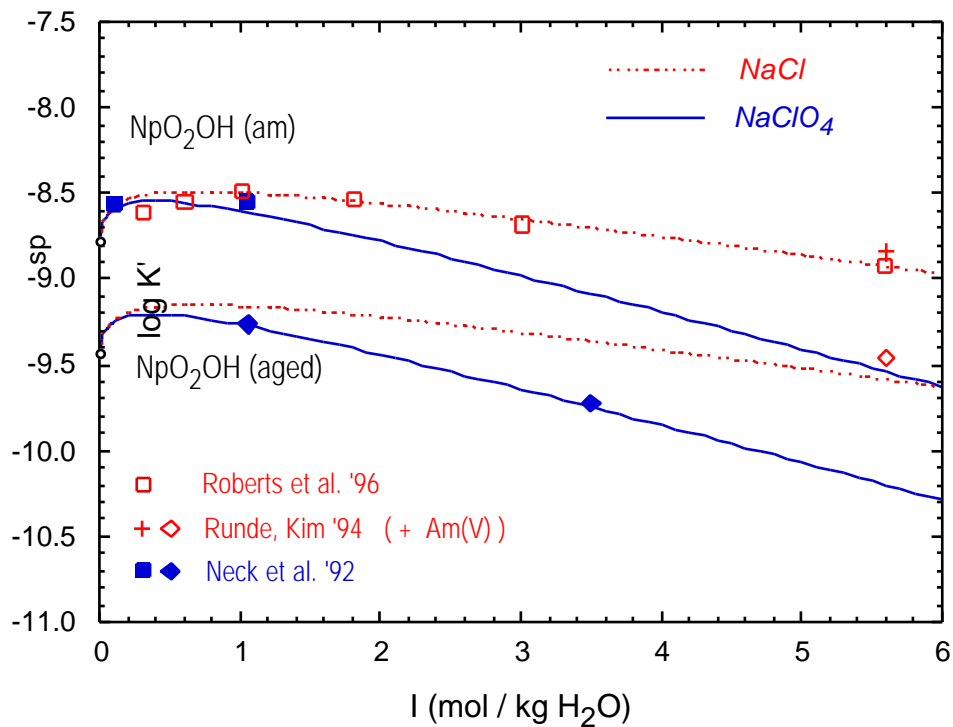
It is to note that the values of $(\text{NpO}_2^+/\text{Cl}^-)$ and $(\text{NpO}_2^+/\text{ClO}_4^-)$, together with $(\text{Na}^+/\text{OH}^-)$ and the solubility product for $\text{NpO}_2\text{OH}(\text{am})$ selected in the present NEA review, excellently describe the solubility product of Np(V) hydroxide as a function of the NaCl and NaClO₄ concentrations, respectively (c.f. review of [96ROB/SIL], p.849 and the Figure below).

The value of $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.20 \pm 0.03$ is in fair agreement with the value selected by Vitorge from measurements on the quasi-reversible redox potential Np(V)/Np(VI):

$(\text{NpO}_2^+/\text{ClO}_4^-) = 0.25 \pm 0.05$. Vitorge’s value is based on the reasonable assumption that interaction coefficients for UO_2^{2+} and NpO_2^{2+} should be the same. The same has to be expected for interaction coefficients for AnO_2^+ ions. The inaccuracy of the redox potential method becomes evident, when the corresponding results for An = U, Np and Pu are compared:

$(\text{UO}_2^+/\text{ClO}_4^-) = 0.26 \pm 0.03$, $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.25 \pm 0.05$ $(\text{PuO}_2^+/\text{ClO}_4^-) = 0.17 \pm 0.05$
(NEA review [95SIL/BID], page 325; data for Np and Pu from Vitorge’s research group)

Of course, the question, which of the two methods leads to more accurate interaction coefficients ($\text{NpO}_2^+/\text{ClO}_4^-$) might be discussed controversially. However, just disregarding the data from the solvent extraction study in [95NEC/FAN] is certainly not in accord with the guidelines for NEA-TDB reviews.



Solubility products of amorphous and aged Np(V) hydroxide at 25°C as a function of the NaCl and NaClO₄ molality (exp. data from [92NEC/KIM, 94RUN/KIM, 96ROB/SIL]). The ionic strength dependence is calculated (predicted) with ($\text{NpO}_2^+/\text{Cl}^-$) = 0.09, ($\text{NpO}_2^+/\text{ClO}_4^-$) = 0.20 and (Na^+/OH^-) = 0.04 [NEA-TDB]

Example 10

Ion interaction coefficients for Np(V) species.

In Table A.24 (page 832), the reviewer deduces “*values of $\epsilon(\text{NpO}_2^+/\text{ClO}_4^-)$ from the measurements of Kim’s group*”. The only reasonable values given in this table are those derived from the Pitzer parameters reported in [95NEC/FAN, 96RUN/NEU]:

$(\text{NpO}_2^+/\text{ClO}_4^-) = 0.20 \pm 0.03$ and 0.18 ± 0.03 , respectively, which are comparable to the value of 0.25 ± 0.05 from Vitorge (c.f. Example 9). All other values given in this table do not make any sense. We never claimed nonsensical interaction coefficients of $(\text{NpO}_2^+/\text{ClO}_4^-)$ in the range 0.34 - 0.42, which are “recalculated” by the referee. We published values for the dissolution reaction of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ and $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$. These values are not based on “measurements of Kim’s group” alone. They also include the data from other authors. And the reviewer knows very well that analogous calculations, which were based exclusively on the studies declared as reliable by himself (those of Maya [83MAY] and Vitorge [86GRE/VIT] in 1 and 3 M NaClO_4), would lead to almost the same value of . This is of course not mentioned by the reviewer. The reviewer’s subtle intention to discredit our work bare of any scientific reason becomes evident in the following “conclusions”:

p. 806, lines 17-19

The reported parameters [91KIM/KLE, 94NEC/RUN] were also not consistent with those in a later paper [95FAN/NEC] from this group.

p. 832, lines 11-16

Thus the ionic strength corrections from Kim’s group are not self consistent, and the ones used for solubility products are not consistent with currently accepted e value. These two inconsistencies could result from several causes: the small systematic errors noted in this appendix, incorrect theoretical methodology, errors in the experimental determination of free carbonate concentrations at high ionic strength or chemical problems with the solid phases.

We recognized very well the problem of the inconsistencies, when values calculated from $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ solubility products at different NaClO_4 concentrations or from auxiliary values. And from the perfect consistency concerning the results in the Np(V)-hydroxide system (c.f. Example 9), we also knew very well that this problem does not arise from the activity coefficients of the NpO_2^+ ion. In our papers [95NEC/FAN, 95FAN/NEC, 95FAN/NEC] it is clearly pointed out that this problem arises from incorrect auxiliary data for the CO_3^{2-} ion in NaClO_4 solutions. (This question will be discussed later in the present manuscript). And the fact that we had several personal discussions on this problem with the reviewer (P. Vitorge) himself, provides further evidence for his impertinent intentions.

A similar, even worse example is included in Table 12.5 (page 253):

Extremely outstanding and unrealistic SIT coefficients β and γ are reported for the papers [95FAN/NEC, 95NOV/ROB]. Since in the original papers, no SIT coefficients are reported at all, these β and γ values are obviously “calculated” by the reviewer. However this is not indicated by footnotes. The experimental data in NaClO₄ solution applied in [95FAN/NEC] and the evaluated $\log \gamma_{\pm}$ values are essentially the same or similar to those in [94NEC/RUN]. Consequently, the same must hold for the β values, i.e. the reviewer’s calculations (possibly transforming Pitzer into SIT coefficients) must include severe errors. If the reviewer is unable to do such calculations, he should let it be! At least, he should indicate that these “results” were evaluated by the reviewer and not by the authors of the original papers!

Example 11

Solubility product of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$;

(discussed on page 284, line 26 - page 288, line 11 and in Fig.12.6, page 289)

From the large number of solubility studies, the reviewer selected those of Maya [83MAY] and Vitorge [86GRE/ROB] (in 1 and 3 M NaClO_4 , respectively) as the most reliable ones. The results of Kim and coworkers were generally criticized to be not reliable. Solely their solubility product value at $I = 0$ was included in the reviewer's data selection (which seems to be somewhat arbitrary, if all other data are disregarded). Although the reviewer consedes in another section of the book:

p.256, lines 9 - 11:

... it is possible that the present reviewer did not completely understand the calibration procedure used by Kim et al. ...

he takes each opportunity to repeat the following reasons for disregarding the results of Kim and coworkers, and also those of former coworkers of Kim (c.f. section 12.1.2 and the corresponding paper reviews in Appendix A):

- (a) systematic errors in pH calibration,
- (b) calculation of $\log [\text{CO}_3^{2-}]$ or $\log [\text{OH}^-]$ from measured $\log [\text{H}^+]$ with auxiliary data, which are inconsistent with the NEA-TDB
- (c) chemical problems with the solid phase

Non of these reasons holds for the studies in 0.1 M NaClO_4 and 0.1, 1, 3 and 5 M NaCl . The dissociation constants of H_2O and H_2CO_3 determined by the authors themselves and used to calculate $\log [\text{CO}_3^{2-}]$ or $\log [\text{OH}^-]$ from measured $\log [\text{H}^+]$ are in reasonable agreement (within < 0.05 and 0.1 log units, respectively) with well-known and generally accepted literature values and also with the auxiliary data of the NEA-TDB. This was already clearly pointed out in our paper [96FAN/NEC]. (It must be emphasized that reasons (a) and (c) do not hold either for our other studies in 1, 3 and 5 M NaClO_4 , which will be discussed later.)

The table below (next page) shows the results of 7 solubility experiments, which were performed by 3 different investigators at 4 different institutes: Neck [91KLE/KIM, 94NEC/RUN] (FZK Karlsruhe), Runde (Techn. Univ. Munich and Los Alamos Nat. Lab.) [94RUN/KIM, 96RUN/NEU] and Meinrath [94MEI] (Japan Atomic Energy Research Institute). We now use auxiliary data given in the NEA-TDB for $(\text{Na}^+/\text{Cl}^-)$, $(\text{Na}^+/\text{ClO}_4^-)$ and

($\text{Na}^+/\text{CO}_3^{2-}$), together with ($\text{NpO}_2^+/\text{Cl}^-$) = 0.09 and ($\text{NpO}_2^+/\text{ClO}_4^-$) = 0.20 or 0.25 discussed in example 9, to calculate the solubility product of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ at $I = 0$ (see table below). It becomes evident that (within the range of experimental uncertainties) all these studies lead to a consistent values of $\log K_{\text{sp}}^\circ$, in particular if we assume the hydration number of $x = 3.5$ given in [83MAY]. This demonstrates that in all these studies, even in those at high NaCl concentrations, the solubility data refer to a well-defined unique solid phase.

A serious review would have included these calculations, but the reviewer disregarded all these results for untrue reasons (or for personal animosities concerning the investigators?).

Solubility products of hydrated $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ from studies in 0.1 M NaClO_4 and in 0.1, 1, 3 and 5 M NaCl (at 20 -25°C)

Ref.	Medium	$\log K_{\text{sp}}$ (molar)	$\log K_{\text{sp}}^\circ (I = 0)$ ^{a)}	
			x = 0	x = 3.5
[91KIM/KLE, 94NEC/RUN]	0.1 M NaClO_4	-10.28	-10.92	-10.92
[94MEI]	0.1 M NaClO_4	-10.22	-10.86	-10.86
[96RUN/NEU]	0.1 M NaCl	-10.40	-11.05	-11.05
[94RUN/KIM, 96RUN/NEU]	1.0 M NaCl	- 9.77	-10.93	-10.98
[96RUN/NEU]	3.0 M NaCl	- 9.40	-10.67	-10.86
[94RUN/KIM, 95NEC/RUN]	5.0 M NaCl	- 9.61	(-10.83	-11.21) ^{b)}
[94NEC/KIM]	5.0 M NaCl	- 9.52	(-10.74	-11.12) ^{b)}

^{a)} calculated with (Na^+/Cl^-) = 0.03, ($\text{Na}^+/\text{ClO}_4^-$) = 0.01, ($\text{Na}^+/\text{CO}_3^{2-}$) = -0.08 (from [95SIL/BID]), ($\text{NpO}_2^+/\text{Cl}^-$) = 0.09 and ($\text{NpO}_2^+/\text{ClO}_4^-$) = 0.20, for x = 0 and 3.5 hydration water molecules, respectively

^{b)} at this NaCl concentration the SIT may become inaccurate

Problems concerning the evaluation of thermodynamic data for solid Np(V) carbonates from experimental studies in NaClO₄ solution

In the table below, analogous calculations are done with the published solubility products in NaClO₄ solution. Again the auxiliary interaction coefficients from the NEA-TDB are used (with the exception that $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.20$ is used, instead of 0.25 as proposed by Vitorge, but this has only a rather limited impact on the calculations.)

In contrast to the observations above in Example 11, the calculated $\log K_{\text{sp}}^\circ$ values are not consistent. They evidently decrease. The reviewer explains this effect to more or less aged solid phases with different numbers of crystal water molecules. Of course, this possibility cannot generally be ruled out.

Solubility products of hydrated $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ from studies in 0.1 - 5 M NaClO₄

Ref.	Medium	$\log K_{\text{sp}}$ (molar)	$\log K_{\text{sp}}^\circ (I = 0)$ ^{a)}	
			x = 0	x = 3.5
[91KIM/KLE, 94NEC/RUN]	0.1 M NaClO ₄	-10.28	-10.92	-10.92
[94MEI]	0.1 M NaClO ₄	-10.22	-10.86	-10.86
[83MAY]	1.0 M NaClO ₄	-10.14	-11.18	-11.23
[91KIM/KLE, 94NEC/RUN]	1.0 M NaClO ₄	-10.10	-11.14	-11.19
[86GRE/ROB]	3.0 M NaClO ₄	-10.56	-11.41	-11.60
[91KIM/KLE, 94NEC/RUN]	3.0 M NaClO ₄	-10.45	-11.30	-11.49
[94NEC/RUN]	5.0 M NaClO ₄	-11.06	(-11.49	-11.87) ^{b)}

^{a)} calculated with $(\text{Na}^+/\text{ClO}_4^-) = 0.01$, $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$ (from [95SIL/BID]), and $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.20$, for x = 0 and 3.5 hydration water molecules, respectively

^{b)} at this NaClO₄ concentration the SIT may become inaccurate

On the other hand, because of (a) the systematic decrease of the $\log K_{\text{sp}}^\circ$ values and (b) the fact that a comparable decrease of $\log K_{\text{sp}}^\circ$ was not observed for the different NaCl solutions, it is more likely that the auxiliary data used for the calculation of activity coefficients are not correct. Actually, in our report [Neck, Fanghänel, Kim, FZKA 5599 (1995)], which was later

summarized in the paper [96FAN/NEC], this suspicion was confirmed.

In contrast to our H_2CO_3 dissociation constants in 0.1 M NaClO_4 and in 0.1 - 5 M NaCl , determined as accompanying work within our Np(V) solubility studies, our corresponding H_2CO_3 dissociation constants in 1, 3 and 5 M NaClO_4 do not agree with literature values accepted or in accord with the NEA-TDB auxiliary data. It is to note that the same experimental procedure was used in all our experiments, and that the determined dissociation constants of H_2O , were generally in excellent agreement with well-accepted literature values at $I = 0.1 - 5 \text{ M}$, in both NaCl and NaClO_4 solution. The latter results confirm that our pH electrode calibration was certainly correct, and not related to systematic errors, as claimed and repeated again and again by the reviewer. We also explained, why so many literature data are incorrect: in all these studies the pH electrode was calibrated only in acidic solutions and then extrapolated to the alkaline range assuming ideal Nernst slopes (59.16 mV / pH unit), whereas the real slopes of commercial glass electrodes are generally somewhat smaller (58.0 - 58.8 mV / pH unit) [96FAN/NEC].

We finally evaluated the activity coefficients and ternary ion interaction Pitzer parameters for carbonate ions in NaClO_4 solution from the H_2CO_3 dissociation constants [96FAN/NEC]. Combining these parameters with those evaluated in [95NEC/FAN] for the activity coefficients of the NpO_2^+ ion, the value $\log K_{\text{sp}}^\circ$ for $\text{NaNpO}_2\text{CO}_3 \cdot 3.5 \text{ H}_2\text{O}(\text{s})$ is again calculated from the experimental solubility products (see table below, next page). Now, consistent values at $I = 0$ are obtained for 12 solubility experiments from 5 different investigators with the solution composition widely varied ($I = 0.1, 1, 3$ and 5 M in both, NaCl and NaClO_4 solutions). The average value is found to be

$$\log K_{\text{sp}}^\circ = -11.08 \pm 0.20 \text{ (2 } \sigma \text{) for } \text{NaNpO}_2\text{CO}_3 \cdot 3.5 \text{ H}_2\text{O}(\text{s})$$

Using the same activity coefficients, consistent $\log K_{\text{sp}}^\circ$ values are obtained as well for $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$ (see [95NEC/FAN]), from experimental results in 1, 3 and 5 M NaClO_4 and in 5 M NaCl [86GRE/ROB, 91KIM/KLE, 94RUN/KIM, 95NEC/RUN].

These calculations provide strong evidence that all the solubility studies listed in the table below refer to a unique, well-defined solid phase. Additional evidence for this arises from several experimental reasons given in [91KIM/KLE, 94NEC/RUN, 95NEC/RUN]. Such experimental reasons are e.g. the reproducibility of the solubility data, when the equilibrium carbonate concentration was increased and decreased again, and the reproducibility when $\text{NaNpO}_2\text{CO}_3 \cdot x \text{ H}_2\text{O}(\text{s})$ had transformed into $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$ and then back into $\text{NaNpO}_2\text{CO}_3 \cdot x \text{ H}_2\text{O}(\text{s})$, or the long total duration time of about half a year for each experiment.

Solubility products of hydrated $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$
in 0.1 - 5 M NaCl and NaClO_4 (at 20 -25°C)

Ref.	Medium	$\log K_{\text{sp}}$ (molar)	$\log K_{\text{sp}}^{\circ} (\text{I} = 0)^{\text{a)}}$ $x = 3.5$
[96RUN/NEU]	0.1 M NaCl	-10.40	-11.08
[94RUN/KIM, 96RUN/NEU]	1.0 M NaCl	- 9.77	-11.10
[96RUN/NEU]	3.0 M NaCl	- 9.40	-11.00
[94RUN/KIM, 95NEC/RUN]	5.0 M NaCl	- 9.61	-11.15
[94NEC/KIM]	5.0 M NaCl	- 9.52	-11.06
[91KIM/KLE, 94NEC/RUN]	0.1 M NaClO_4	-10.28	-10.94
[94MEI]	0.1 M NaClO_4	-10.22	-10.88
[83MAY]	1.0 M NaClO_4	-10.14	-11.18
[91KIM/KLE, 94NEC/RUN]	1.0 M NaClO_4	-10.10	-11.14
[86GRE/ROB]	3.0 M NaClO_4	-10.56	-11.25
[91KIM/KLE, 94NEC/RUN]	3.0 M NaClO_4	-10.45	-11.14
[94NEC/RUN]	5.0 M NaClO_4	-11.06	-11.08

^{a)} calculated with Pitzer parameters given in [95NEC/FAN, 96FAN/NEC]

As a consequence, the conclusions drawn by the reviewer, that the results of Maya [83MAY] and the Kim group refer to a more hydrated solid phase with $\log K_{\text{sp}}^{\circ} = -11.16 \pm 0.35$ and those of Vitorge [86GRE/ROB] to an aged, less hydrated solid phase with $\log K_{\text{sp}}^{\circ} = -11.66 \pm 0.50$ (page 287) should urgently be overthought. Even the reviewer states:

p.286, lines 5 - 7:

... However, it is also possible that this work may indicate a need to eventually revise auxiliary data used in the present review (see discussion of [96FAN/NEC] in Appendix A)

but he does not give any hint, what would be the alternative conclusion.

Problems due to shortcomings of the SIT approach

But now, if we accept the results and interpretation given in [95NEC/FAN, 96FAN/NEC] another problem arises. This problem is directly related to shortcomings of the SIT approach. According to the results in [96FAN/NEC], the trace activity coefficients of the CO_3^{2-} in NaCl and NaClO_4 solutions above 1 molal are considerably different (see figure, next page). However, these differences cannot be described with the simple SIT approach used in the NEA reviews, because (in contrast to the Pitzer equations) interaction coefficients between ions of the same charge sign are generally set equal to zero. As a consequence, with the SIT approach equal CO_3^{2-} trace activity coefficients are calculated for NaCl and NaClO_4 solutions of equal molality.

In order to demonstrate this kind of shortcoming of the SIT, we worked out two further corresponding examples of well-known different trace activity coefficients, which cannot be described with the SIT:

1) Trace activity coefficients of the SO_4^{2-} ion in NaCl and NaTcO_4 solution, which is quite analog to the problem of CO_3^{2-} in NaCl and NaClO_4 solution.

Unfortunately there are no data for the system Na- SO_4 - ClO_4 , but the TcO_4^- ion has properties very similar to those of the ClO_4^- ion. The Pitzer parameters in the system Na- SO_4 - TcO_4 are very well known and ascertained by numerous isopiestic and solubility data [Neck, Könnecke, Fanghänel, Kim, J. Solution Chem. 27 (1998), 107 and Neck, Könnecke, Fanghänel, Kim, Radiochim. Acta 83 (1998), 75].

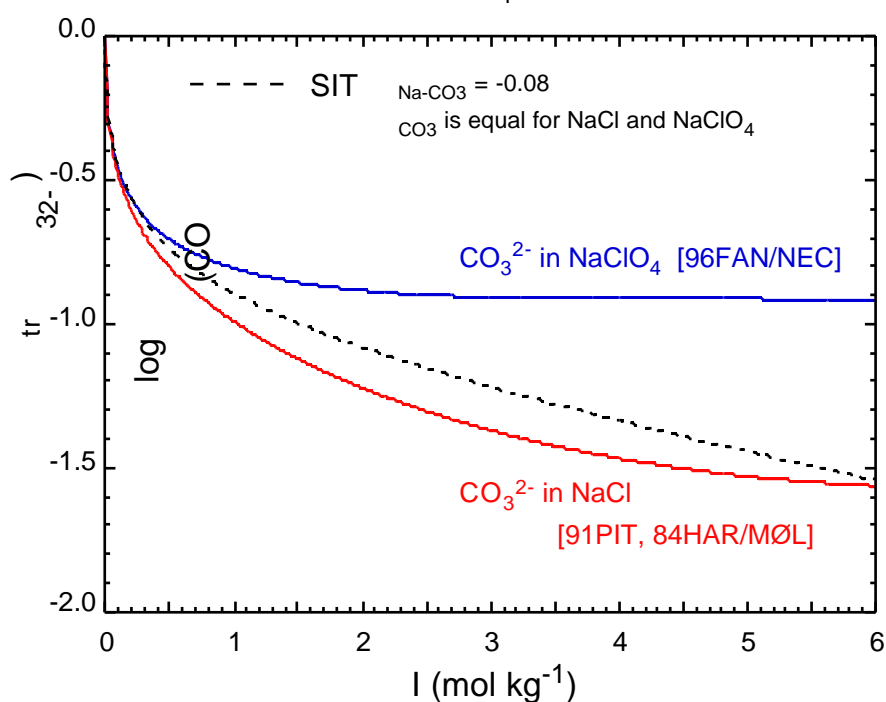
The results are shown on the next page. They show that the corresponding differences in the trace activity coefficients of CO_3^{2-} and SO_4^{2-} are comparable, not only qualitatively but also quantitatively.

2) Trace activity coefficients of the H^+ ion in NaCl and CsCl solution.

According to the SIT, they should be equal, but from emf measurements it is well-known that they are different. The values shown in the corresponding figure are calculated with the Pitzer parameters [91PIT] evaluated from the experimental emf data.

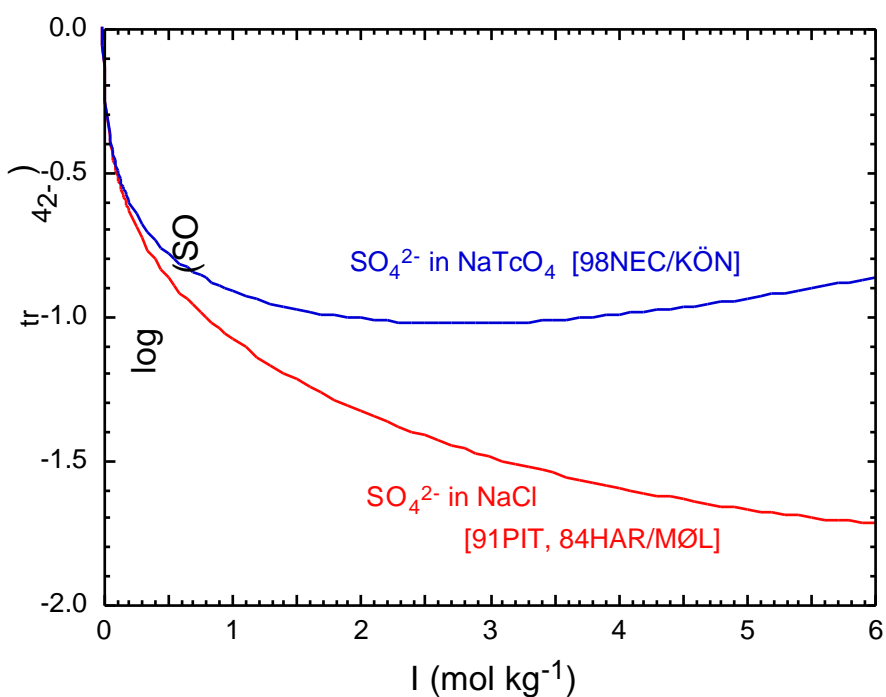
Trace activity coefficients of the CO_3^{2-} ion

Comparison between SIT and Pitzer calculations
for NaCl and NaClO_4 solutions, 25°C

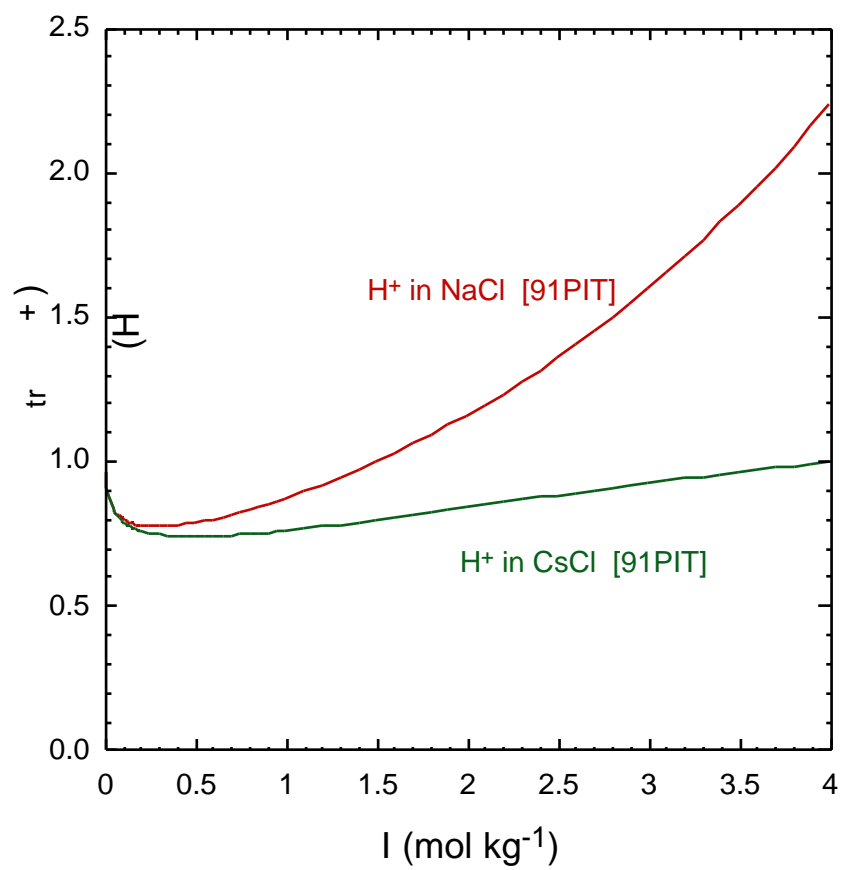


Trace activity coefficients of the SO_4^{2-} ion

Pitzer calculations for NaCl and NaTcO_4 solutions



Trace activity coefficients of the H^+ ion
in NaCl and CsCl solutions



Problems in interpreting experimental data on aqueous Np(V) complexes in NaCl solution

Analog problems arise, when the reviewer considered the results of Runde et al.

[96RUN/NEU] on Np(V) carbonate complexation constants in 0.1 - 5 M NaCl.

In the review of [96RUN/NEU] (Appendix A, page 849-850), the reviewer evaluates SIT coefficients between Na^+ and negatively charged Np(V) carbonate complexes from Runde's data in NaCl solution:

... which do not seem reliable.... (page 850, lines 11-12)

The reviewer then blames Runde for applying the Pitzer equations and writes:

page 850, lines 21-23

... The usual scatter of experimental data is in the same order of magnitude as the ionic strength corrections...

In contrast to the reviewer's statements, the data in [96RUN/NEU] are very well correct. However, again the trace activity coefficients of negatively charged Np(V) complexes are different in NaCl and NaClO_4 solution, which cannot be described with the simple SIT approach. And as already pointed out in [95FAN/NEC, 96RUN/NEU], this has nothing to do with chloride complexation as supposed in our earlier work [94NEC/KIM, 94RUN/KIM, 95NEC/RUN].

One possibility to overcome this problem would be to calculate only values for the reactions, which are then different for NaClO_4 and NaCl solution. Another possibility would be to introduce interaction coefficients between ions of the same charge sign, e.g.

$(\text{NpO}_2(\text{CO}_3)_2^{3-}/\text{Cl}^-)$. However, the latter method would have a severe impact on previous NEA-TDB reviews.

In the past, studies on the ionic strength dependence of dissolution or complexation equilibria were almost exclusively performed in NaClO_4 solutions. The Np(V)-carbonate system is the first one, for which a large number of experimental data are available in two different electrolyte media. Meanwhile, much more studies on actinide solution thermodynamics are focussed on chloride solutions, because of their relevance for natural systems. This means: the described problems will appear again for other systems with negatively charged complexes (e.g. for U(VI) or Am(III) carbonate and sulphate complexes). Therefore, it does not seem reasonable, just to ignore the problems in the case of Np(V) by disregarding the set of experimental data in NaCl solution.

Dr. Robert J. Lemire
AECL Research
Whiteshell Laboratories

**Pinawa, Minitoba R0E 1L0
Canada**

Date: 15.10.99
Persons in charge: Dr. Thomas Fanghänel
Dr. Volker Neck
Phone: +49 7247 82 5485 / 2592
Fax: +49 7247 82 4308
email: thomas@ine.fzk.de
neck@ine.fzk.de

NEA TDB review: Chemical Thermodynamics of Neptunium and Plutonium (draft of September 16, 1999)

Dear Dr. Lemire,

as members of the review team in the NEA-TDBII project we recently got access to the draft of the Np/Pu volume. First we were impressed by the tremendous work the reviewers have carried out to assess the thermodynamic data of Np and Pu. However, when we took a closer look we were shocked about the obvious and impertinent attempt of one of the reviewers (we have strong indications that this reviewer is P. Vitorge) to discredit our work on aqueous chemistry and thermodynamics of Np(V). Besides substantially erroneous conclusions, these parts of the review represent inadequate review work including numerous mistakes. Further, they contain misquotations, incorrect abstruse statements, and even untruths, which suggest to the reader that all data and publications which were carried out in Prof. Kim's laboratories in Munich and Karlsruhe (the reviewer refers to "Kim's group") during the last decade are useless for the scientific community and of very low scientific quality. Some striking examples are documented in the attached comments. Working out more in detail numerous other examples would require more time. As we know about the planned time schedule for the publication of this volume, we have decided to inform you immediately about this attempt to abuse the NEA-TDB project.

A major revision of these parts, by another expert, appears to be urgently necessary in order to prevent the whole review team involved in this book or even the whole NEA-TDB project

from being brought into discredit. Of course we know that this will cause a further delay of the publication, but it is not at all convenient to do all the necessary revisions in the NEA-TDBII project.

So far the NEA-TDB has a high reputation in the scientific community. It is accepted as the most important and most reliable database on thermodynamics of actinides. Therefore an enormous responsibility is burdened on the shoulders of the review team for a high quality thermodynamic database on Np and Pu. We are convinced that you together with your review team, the management board of the NEA-TDB project and the NEA team will take action to keep the standards of NEA-TDB and to prevent the misuse of the NEA TDB as a platform for personal animosities.

A soon answer in this matter is highly appreciated.

Sincerely

Volker Neck

Thomas Fanghänel

Enclosure

cc Mehdi Askarieh (Chairman of the Management Board of the OECD/NEA TDB Project)
Kastriot Spahiu, Cynthia Palmer, Jordi Bruno, Hans Wanner (Members of the Executive Group of the OECD/NEA TDB Project)
Eric Östhols (OECD/NEA TDB Project Coordinator)
Jean Fuger, Heino Nitsche, Paul Potter, Malcom Rand, Jan Rydberg, James Sullivan, William Ullman, Pierre Vitorge (Members of the review team)
Ingmar Grenthe, Robert Guillaumont

Comments on actinide carbonate compounds concerning

- **NEA-TDB draft of the Np/Pu review**
- **NEA-TDB II (update of U, Am, Np, Pu, Tc)**

V. Neck and Th. Fanghänel

Institut für Nukleare Entsorgungstechnik, Forschungszentrum Karlsruhe,
Postfach 3640, D-76021 Karlsruhe, Germany

November 1999

Contents

1. Comments on the experimental studies of Neck and Runde
 - 1.1. On the solubility studies with Np(V) carbonates
 - 1.2. On the determination of the H^+ , OH^- and CO_3^{2-} concentrations
2. Erroneous conclusions on actinide carbonates, which arise from limitations/shortcomings of the SIT and erroneous NEA-TDB auxiliary data on trace activity coefficients of the carbonate ion (in $NaClO_4$ solutions above 1 molal)
 - 2.1. General comments on shortcomings of the SIT
 - 2.2. Actinide carbonate solids of Np(V), Am(III), U(VI); consequences for solubility constants at $I = 0$
 - 2.3. Proposal to solve the problem of carbonate trace activity coefficients
 - 2.4. Consequences for NEA-TDB reviews
3. References
4. Appendix

4.3.3 Additional conditions for Users outside the European Community: Springer-Verlag and authors will not be liable for any damages, including any lost profits, lost savings, or other incidental or consequential damages arising from the use of, or inability to use, this software and its accompanying documentation, even if Springer-Verlag and/or the authors have been advised of the possibility of such damages.

4.4) Liabilities of the User

4.4.1 The User agrees to comply with the rules outlined herein. Violations of these rules may be criminal offences and may also give rise to claims for damages against the User from the licensors of Springer-Verlag.

4.4.2 In the case of serious violations committed by the User, Springer-Verlag may revoke the license.

4.4.3 Upon installation of the codes the User(s) implicitly accept the licensing agreements provided with the manuals of the respective code(s). Therefore, the manuals are to be read immediately after installation of the respective program(s).

5) Final Provisions

5.1) If any provision of the entire Agreement in chap. 4) is or becomes invalid or if this Agreement is incomplete, the remainder of the Agreement is not affected. The invalid provision shall then be replaced by a legally valid provision, which comes as close as possible to the invalid provision as far as its economic effect is concerned. The same applies to possible gaps in the Agreement.

5.2) This Agreement falls under the jurisdiction of the courts at Heidelberg, if the User is a merchant who has been entered in the commercial register, a legal entity under public law, or a public special fund, or if the User has no residence or place of business in Germany.

5.3) This Agreement is subject to the laws of Germany to the exclusion of the Uncitral Trading Rules.

Additional conditions for Users outside the European Community see chap. 4.3.3.

1. Comments on the experimental studies of Neck and Runde

In the draft of the NEA review on Np/Pu, the results of Kim and coworkers on the solid and aqueous Np(V) carbonates [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC, 96FAN/NEC], were generally criticized to be not reliable. In section 12.1.2 and the corresponding paper reviews in Appendix A (not only in the papers cited above, but also in related papers from other authors) the reviewer takes each opportunity to repeat the following reasons for disregarding the results of Kim and coworkers, and also those of former coworkers of Kim [94MEI, 96RUN/NEU]:

- (a) chemical problems with the solid phase
(insufficient knowledge on the equilibrium solid phase)
- (b) systematic errors in pH calibration,
- (c) calculation of $\log [\text{CO}_3^{2-}]$ from measured $\log [\text{H}^+]$ with auxiliary data, which are inconsistent with the NEA-TDB

Statement (a) is absolutely incorrect. In contrast to the reviewer's study (c.f. discussion of [90RIG] in the Np/Pu draft) we had sufficient accurate experimental data and unambiguous experimental proofs for the solubility limiting solid phases (see section 1.1).

Statement (b) is also incorrect as will be shown in section 1.2.

Statement (c) holds only for a part of our studies: for those in 1, 3 and 5 M NaClO_4 , where the auxiliary data of the NEA-TDB are incorrect as will be shown in the present manuscript. In the studies in 0.1 M NaClO_4 and 0.1, 1, 3 and 5 M NaCl , the H_2CO_3 dissociation constants determined and used to calculate $\log [\text{CO}_3^{2-}]$ from measured $\log [\text{H}^+]$ agree with well-known and generally accepted literature values and also with the auxiliary data of the NEA-TDB.

Nevertheless, the reviewer disregarded all these results.

Careful reading of our original papers might be sufficient to agree with our arguments.

Several members of the NEA-TDB project groups required additional information and explanations, which hopefully will now be given in this manuscript.

In addition, it has to be stated that the primary intention of any scientific study should be to obtain correct results. The consistency with recommendations of other people or organisations is desirable, but only of secondary importance.

1.1. On the solubility studies with Np(V) carbonates

For the discussion in section 2 of this manuscript, it is important to know that the solubility data in [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC, 96RUN/NEU] refer to well-defined solid phases.

- a) hydrated $\text{NaNpO}_2\text{CO}_3 \cdot 3.5\text{H}_2\text{O}(\text{s})$ characterised in [77VOL/VIS, 83MAY]
- b) $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$, formed at high carbonate concentration and $[\text{Na}^+] = 1 \text{ mol/l}$.

Experimental procedure

All solubility experiments were performed in titration cells (see Appendix 1) under a given CO_2 partial pressure ($10^{-3.5}$ or $10^{-2.0}$ atm) and constant Na^+ concentration of the background solution (0.1, 1, 3 and 5 M NaClO_4 or NaCl). The Np(V) solid was precipitated in the titration vessel, and left aging 1 - 2 weeks before the experiment was started. Within the solubility studies in NaClO_4 solutions, the pH was varied by adding $\text{HClO}_4/\text{NaClO}_4$ or $\text{NaHCO}_3/\text{NaClO}_4$. Equilibration was achieved by bubbling the CO_2/Ar gas mixture through the solution. (The CO_2/Ar stream was preequilibrated with water vapor by bubbling it through a corresponding background solution.) The equilibration between $\text{CO}_2(\text{g})$, aqueous carbonate and solid Np(V) carbonate was monitored as a function of time by measuring the H^+ and Np concentrations until these concentrations remained constant. This could last a few days up to 3 weeks, depending primarily on the time needed for the equilibrium between HCO_3^- and CO_3^{2-} in the aqueous phase and $p\text{CO}_2$ in the gas phase.

Maya [83MAY] reported in the experimental section of his paper:

Solid phase $\text{NaNpO}_2\text{CO}_3 \cdot 3.5\text{H}_2\text{O}(\text{s})$. The crystalline compound was prepared ... by addition of Na_2CO_3 in a 1:1 mole ratio to a NpO_2^+ solution. The gelatinous precipitate initially formed was aged at 25°C for 8 days with slow stirring in the presence of excess Na_2CO_3 (0.025 M). This treatment produced a crystalline solid that settled by gravity within a few minutes.

Exactly the same observation was made in our studies and therefore the solubility experiments were started with a precipitate aged for 1 - 2 weeks.

Solubility limiting solid phases

1) X-ray powder diffraction

The X-ray powder diffraction pattern reported in [94NEC/RUN, 94RUN/KIM, 95NEC/RUN] were taken from a part of the solids used in our solubility studies. The X-ray pattern observed for $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ were all consistent with those reported by Volkov et al. [77VOL/VIS] for $\text{NaNpO}_2\text{CO}_3 \cdot 3.5\text{H}_2\text{O}(\text{s})$. The patterns obtained for $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$ were consistent with the results of Volkov et al. [81VOL/VIS] for that compound. In addition, the report [94RUN/KIM] contains a series of x-ray pattern recorded during the time of solid phase transformation from $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ to $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$ in 5 M NaCl (see Appendix 2).

In contrast to the results given in [94NEC/RUN, 94RUN/KIM, 95NEC/RUN], Meinrath [94MEI]

reported an other (hexagonal) modification of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$, with a different x-ray pattern. However, he obtained comparable solubility data: the reported solubility constant $\log K_s$ in 0.1 M NaClO_4 is about 0.1 - 0.2 log units higher than those of [94NEC/RUN] in the same medium and those of [96RUN/NEU] in 0.1 M NaCl .

2) Slope analysis

As shown in [94NEC/RUN], the slopes of the solubility curves ($\log[\text{Np}]$ vs. $\log[\text{CO}_3^{2-}]$) are exactly -1 (not -0.9 ± 0.2 or something like that), over a range of 2 - 2.5 orders of magnitude where NpO_2^+ is the predominant aqueous species. The ratio $\text{Np} : \text{CO}_3^{2-}$ in the solid was exactly 1:1. The presence of solids like $\text{Na}_{0.6}\text{NpO}_2(\text{CO}_3)_{0.8}(\text{s})$ or $\text{Na}_{0.72}\text{NpO}_2(\text{CO}_3)_{0.86}(\text{s})$, as discussed by Vitorge in the NEA review, (p.284 lines 26ff and in Appendix A, discussion of [84VIT, 90RIG]), can certainly be excluded in all our studies. May be such solids are formed as intermediates or as $\text{Np}(\text{V})$ -hydroxide-carbonate solid mixtures in the early state of precipitation (where the precipitate is not fine crystalline but hydroxide-like gelatinous) but they are certainly instable and rapidly (within a few day) transformed into a well-defined crystalline $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$, with $x = 3.5$ according to [77VOL/VIS, 83MAY].

In addition, the slope analysis at high carbonate concentration, where the the complex $\text{NpO}_2(\text{CO}_3)_3^{5-}$ is confirmed spectroscopically as the predominant solution species, again demonstrates that the equilibrium solids are either $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ (slope +2) or $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$ (slope +1) (see discussion in [91KIM/KLE, 95NEC/RUN]).

3) Solubility data for $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ in 0.1, 1.0 and 3.0 M NaClO_4

(from [91KIM/KLE], also reported in [94NEC/RUN])

At first, experiments in 0.1 M NaClO_4 were performed in parallel in two titration vessels. After finishing the experiments at $I = 0.1$ M, the solids in the two vessels (meanwhile about half a year old) were further used for the solubility experiments in 1 and 3 M NaClO_4 . For this purpose the overstanding solution ($I = 0.1$ M) was replaced by $\text{HCO}_3/\text{NaClO}_4$ solutions of $I = 1$ and $I = 3$ M. Hence, the solubility data at $I = 0.1$, 1 and 3 M refer to the same solid phase. And since the solubility data at $I = 1$ and $I = 3$ are practically the same as those obtained by Maya [83MAY] and Vitorge [86GRE/ROB] in the corresponding media (c.f. Fig. 1.1), it is evident that the solubility data of these authors refer as well to the same solid, and not to different (more or less aged or hydrated) solids as concluded in the NEA review.

Solubility of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$

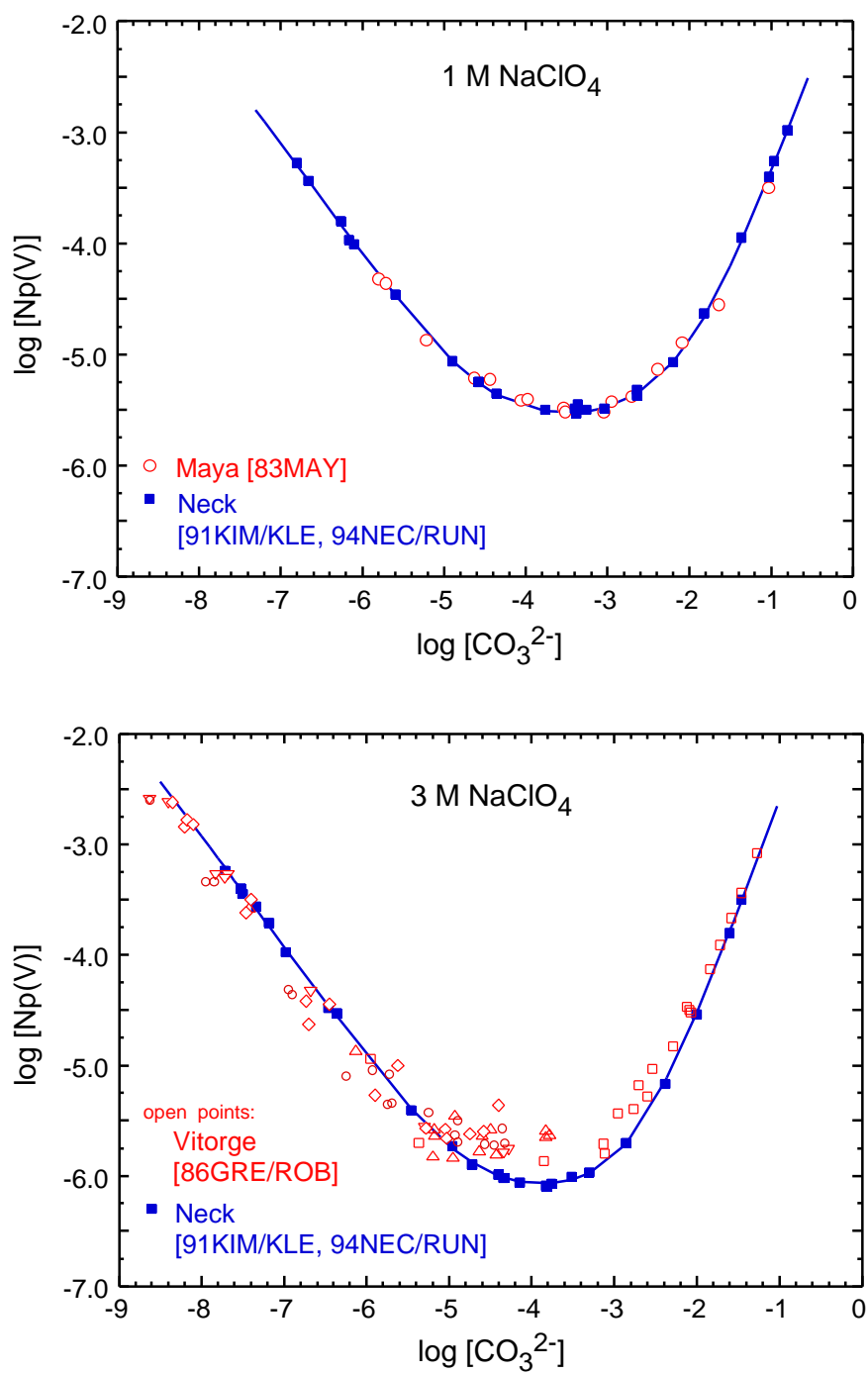


Fig. 1.1

4) Reproducibility of the solubility data

In the report [91KIM/KLE], the solubility data in 0.1, 1 and 3 M NaClO₄ are listed in the sequence of the measurements (copies of the tables are shown in Appendix 3). From the following observations the formation of different (more or less aged or hydrated) solids of NaNpO₂CO₃·xH₂O(s) can definitely be ruled out:

- a) The results in 0.1 M NaClO₄, measured in parallel in two titration cells are in excellent agreement with each other, i.e. the solids must have been the same.
- b) The results in 0.1, 1 and 3 M NaClO₄ are reproducible, when pH and carbonate concentration are increased and decreased again, i.e. the solid was not affected.
- c) The results for NaNpO₂CO₃·xH₂O(s) in 1 and 3 M NaClO₄ are reproducible, even after transformation into Na₃NpO₂(CO₃)₂(s) and retransformation into NaNpO₂CO₃·xH₂O(s), when pH and carbonate concentration were decreased again, i.e. the retransformation lead again to the same solid NaNpO₂CO₃·xH₂O(s),.

These are 4 unambiguous proofs that the solubility data refer to well-defined solids.

In Vitorge's review of the Np(V) carbonates, not a single sentence was written on these efforts and solid phase characterizations, neither in chapter 12.1.2 nor in the Appendix A (discussion of references). Quite in contrast, he states and repeats several times that Kim and coworkers might have had chemical problems with the solid phase.

Accuracy of solubility measurements

In Vitorge's review and also in the comments of Robert Lemire (in his reply to our letter), there were doubts on the solubility data given in our papers [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN], because they are less scattered than expected.

In order to achieve such accurate data, great efforts are necessary. And actually great efforts were made, not only in our studies but also by other authors (e.g. [83MAY, 94MEI]), who also reported Np(V) carbonate solubility data of comparable high accuracy. Some of these efforts (considered as self-evident and not reported in detail in our papers) are given below.

In general, there are two important points:

- 1) The analytical methods to determine the concentrations of interest. The analytical method to determine the Np-237 concentration will be discussed below. The methods used to determine the H⁺ and CO₃²⁻ concentrations will be described separately in section 1.2
- 2) In a phase equilibrium study, it is urgently necessary to ascertain that the measured data actually refer to the equilibrium state. Non-equilibrium data are useless for the determination of thermodynamic quantities.

Determination of the Np-237 concentration

The accurate determination of the Np-237 concentration requires certain efforts and is not trivial. In the NEA review possible uncertainties are underestimated:

p.255, lines 27-31

Another potential problem is that ... can result in a constant minimum total $^{237}\text{Np(V)}$ solubility, which is not far from the usual analytical detection limits (for α - or γ -spectrometry or liquid scintillation methods); however this does not seem to have caused any difficulties in any of the publications cited here.

In contrast to this statement, insufficient analytical facilities might very well have lead to inaccurate data or data scattering. It is a more probable reason for the data scattering in the solubility experiments of Vitorge [86GRE/ROB] or Lemire et al.[93LEM/BOY] than the presence of different solid phases (which is the explanation given by the reviewer).

Lemire et al. [93LEM/BOY] mainly applied α -spectrometry (efficiency < 20%). The analytical procedure included several steps (reduction to Np(IV), extraction, back-extraction) to avoid further reduction of the efficiency due to sorption effects, if the dried aliquots contain large amounts of salt from the electrolyte medium. It is not surprising, if this analytical procedure leads to scattered data, in particular for Np concentrations < 10^{-5} M, close to or at the detection limit of α - and γ -spectrometry.

Liquid scintillation β -counting (efficiency: 100%) is certainly much more appropriate to determine the concentration of Np-237. However, the LSC β -spectrum of Np-237 overlaps with the β -spectrum of the short-lived daughter nuclide Pa-233 ($t_{1/2} = 27$ days, β -decays with maximum energies of 0.3 and 0.6 MeV) - see Appendix 4. The accurate determination of the Np-237 concentration requires an additional β/γ discrimination for the counts from Pa-233. The ratio of the counts from Np-237 and Pa-233 can vary in a wide range and is not reproducible. The extremely small traces of Pa can be sorbed on the glass surface, on the Np(V) solid or dissolved as carbonate complexes. From the experience in our studies we know that particularly in the solubility minimum range, the ratio of ^{233}Pa - β -counts : ^{237}Np - β -counts can be very high. If the LSC measurements are not corrected by discriminating the β -radiation from Pa-233, the Np(V) solubility will be considerably overestimated. The analytical method used by Vitorge et al. is neither mentioned in the paper [86GRE/ROB] nor in Riglet's thesis [90RIG]. May be, at the time Vitorge performed the solubility experiments shown in Fig.1.1 (before 1984), he did not have the analytical facilities necessary to record LSC spectra for β/γ discrimination.

Procedure to ascertain the equilibrium state

Within the solubility studies in NaClO_4 solutions, the pH was varied by adding $\text{HClO}_4/\text{NaClO}_4$ or $\text{NaHCO}_3/\text{NaClO}_4$. The excess of total carbonate (compared to that in equilibrium with $\text{pCO}_2 = 10^{-3.5}$

atm) was driven out of the solution by bubbling the CO₂/Ar gas mixture through the solution: $\text{HCO}_3^- + \text{H}^+ \rightarrow \text{CO}_2(\text{g}) + \text{H}_2\text{O}$

When the equilibrium between HCO_3^- and CO_3^{2-} in the aqueous phase and pCO₂ in the gas phase is reached, the H⁺ concentration remains constant. And when the dissolution equilibrium of the Np(V) carbonate solid is reached, the Np concentration remains constant as well. In our studies, the equilibration procedure was continued until the H⁺ and Np concentrations remained constant within about ±0.02 logarithmic units.

That means: each of the given solubility data represents the final, asymptotically reached equilibrium value of a series of H⁺ and Np measurements as a function of time! This explains why our data show such a small scattering. Of course, such an experimental procedure requires large efforts and a very long time compared to other methods like closed system batch experiments at given total carbonate concentration. However, it has the following advantages:

- 1) One can very well ascertain that the equilibrium state (solid phase \rightleftharpoons aqueous phase \rightleftharpoons gas phase) is reached.
- 2) One can very well recognize, from jumps in the Np concentration, when the solid phase is aging (e.g. $\text{NpO}_2\text{OH}(\text{am}) \rightarrow \text{aged}$), c.f. Figs. in [91KIM/KLE, 92NEC/KIM] or transformed (e.g. $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s}) \rightarrow \text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$, c.f. Figs. in [91KIM/KLE, 94RUN/KIM, 95NEC/RUN]).

1.2. On the determination of the H^+ , OH^- and CO_3^{2-} concentrations

At first it is to state that for the evaluation of conditional equilibrium constants in solutions of constant background medium, the correct concentrations of the reactands involved (H^+ , OH^- and CO_3^{2-}) are needed (not pH or activity coefficients). All our experimental studies of concern [92NEC/KIM, 91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 96FAN/NEC] were performed in a constant background medium (0.1, 1, 3 and 5 M NaClO_4 or NaCl). In order to avoid repetition for all these background media in the following sections, the methods used to determine the concentrations of H^+ , OH^- and CO_3^{2-} are explained for the example of 3 M NaClO_4 as background medium. The same methods are applied in all other cases.

Calibration of pH electrodes and determination of the H^+ and OH^- concentrations

Method A

The pH electrodes are calibrated in the acidic and alkaline range, with solutions of known H^+ and OH^- concentrations in the same background medium as used for the experiments.

Example:

Experiment at constant ionic strength (3 M NaClO_4):

Calibration with a) x M HClO_4 / $(3 - x)$ M NaClO_4 ; $x = 0.1 - 0.001$

b) x M NaOH / $(3 - x)$ M NaClO_4 ; $x = 0.1 - 0.001$

The ion products of water needed to calculate $\log[\text{H}^+]$ from $\log[\text{OH}^-]$ (or vice versa) in NaCl or NaClO_4 solutions are well known (e.g. auxiliary data of the NEA-TDB). However, the electrode calibration has to be permanently checked during the experiment, and the handling of NaOH solutions under inert gas atmosphere requires certain efforts. Therefore we applied also Method B.

Method B

We use standard pH buffer solutions (related to NIST or NBS) for calibration. This requires an additional correction to obtain the concentration $\log[\text{H}^+]$, because the liquid junction potential when calibrating with standard buffer solutions of low ionic strength differs from the liquid junction potential when measuring the test solutions (e.g. in 3 M NaClO_4). The value of $\text{pH}_{(\text{exp})}$ measured in the test solutions is related to the H^+ concentration by

$$-\log[\text{H}^+] = \text{pH}_{(\text{exp})} + A$$

with

$$A = \text{pH} + \log \text{H}^+$$

The term A represents a correction for the differences in liquid junction potentials (pH) and the trace activity coefficient of H^+ . It has to be determined experimentally by measuring $\text{pH}_{(\text{exp})}$ in solutions with known H^+ concentration (x M HClO_4 / $(3 - x)$ M NaClO_4). After that correction,

Methods A and B lead to the same values of $\log[H^+]$.

It is to note that Method B is applied by very many research groups, not only in the studies of Neck and Runde, but also by Rai et al. [91FEL/RAI, 97RAI/FEL] and by other groups in the USA and in Japan.

Unfortunately, our earlier publications [92NEC/KIM, 91KIM/KLE, 94NEC/RUN] contain illustrations of solubilities as a function of pH ($-\log$ of the H^+ activity), which is calculated from the concentration $\log[H^+]$ by assuming equal trace activity coefficients for H^+ and OH^- ions. Similarly this assumption is used to estimate the “real” shift pH caused by the liquid junction potentials. Indeed this assumption is not consistent with the SIT or Pitzer splitting conventions, which might be confusing to the reader and to the NEA reviewers. However, it must be emphasized that this assumption or convention is completely irrelevant for the calculation of the conditional equilibrium constants, because the equilibrium constants are, of course, calculated from the concentrations $\log[H^+]$ and $\log[OH^-]$, not from activities.

Question: correct or incorrect values of $\log [H^+]$ and $\log [OH^-]$?

1) Combining calibration Methods A and B, it is possible to evaluate the ion product of water ($\log K'_w$) in the given background solution, because the value of pH is given by the background electrolyte concentration and the same for acidic and alkaline solutions. The details are given in our papers (e.g. in [92NEC/KIM, 96FAN/NEC]).

It is to note that all our values of $\log K_w$ (in 0.1, 1, 3 and 5 M $NaClO_4$ and in 0.1, 1, 3 and 5 M $NaCl$) agree well with generally accepted literature data, with those calculated from Pitzer parameters [91PIT] and with $\log K_w$ according to SIT and NEA-TDB. It is hence obvious that we applied a correct methodology to determine the concentrations of H^+ and OH^- .

2) Example: $\log[OH^-]$ in the studies on Np(V) hydrolysis

If we use our calibrated pH electrode to measure a solution of 0.01 M $NaOH$ / 2.99 M $NaClO_4$, then the final result must be: $\log[OH^-] = -2.00$ (\pm uncertainty). And if we use the same electrode and the same pH-meter and measure the same emf value in a solution of a solubility experiment in 3 M $NaClO_4$, it certainly follows again that $\log[OH^-] = -2.00$ (\pm uncertainty). We cannot accept any other interpretation.

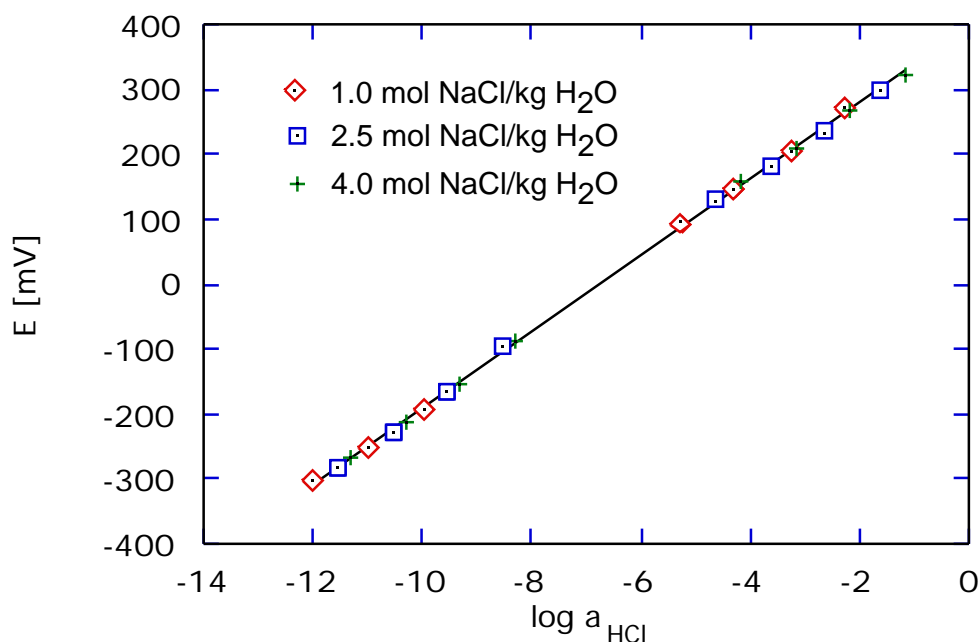
3) Comment on objections (Pierre Vitorge, Robert Lemire) concerning possible carbonate contamination in NaOH calibration solutions.

As demonstrated in our experiments on Np(V) hydrolysis [92NEC/KIM, 91KIM/KLE, 94RUN/KIM], we are certainly able to keep a possible carbonate contamination below 10^{-5} mol/l. (At this carbonate concentration the Np(V) speciation is already significantly affected by carbonate complexation.) And it is a simple calculation exercise, how $\log[\text{OH}^-]$ is affected by carbonate contamination due to uptake of $\text{CO}_2(\text{g})$. Example: even if we assume a considerably overestimated carbonate contamination of 10^{-4} mol/l in a 0.01 M NaOH / 2.99 M NaClO_4 solution, then $\log[\text{OH}^-]$ is decreased from -2.00 to -2.01. Even in this “worst case“, the error would be less than the general uncertainty of pH measurements.

Comments on other pH calibration procedures

In the literature, pH glass electrodes are often calibrated according to Method A, however, only in the acidic range, and then extrapolated to the alkaline range by assuming an ideal Nernst slope of (59.16 mV/pH unit). In the neutral and alkaline range this can lead to errors up to 0.1 - 0.2 units in $\log[\text{H}^+]$, because commercial glass electrodes do not have ideal slopes. Their slopes are slightly lower (58.0 - 58.8 mV/pH unit). We observed this deviation for ROSS electrodes, independent of using calibration method A or B. We also asked Orion Co., and they confirmed non-ideal slopes of their glass electrodes. It is to note that this deviation is also observed, if other glass electrodes and half cells without liquid junction are used (Fanghänel et al.[94FAN/KIM] and unpublished results of Grambow et al.). Fig.1.2 shows a typical example of Fanghänel's investigations. The activity ($\log a_{\text{H}^+} + \log a_{\text{Cl}^-}$) of 10^{-2} to 10^{-5} m HCl or NaOH in 1.0, 2.5 and 4 m NaCl is measured with a glass electrode and a chloride sensitive electrode without liquid junction. All measured emf data are represented by one straight line with a slope of 58.8 mV / $\log a_{\text{HCl}}$.

As discussed in [96FAN/NEC], an incorrect electrode calibration procedure has a severe impact on the NEA-TDB auxiliary data concerning H_2CO_3 dissociation constants and carbonate trace activity coefficients in concentrated NaClO_4 solution. Numerous authors (see refs. in [96FAN/NEC]) determined H_2CO_3 dissociation constants in 0.3 - 3 NaClO_4 solution, calibrating their glass electrodes only in the acidic range and extrapolating the calibration to the alkaline range under the assumption of an ideal Nernst slope. These constants are consistent with the NEA-TDB auxiliary data. However, they are incorrect by the shift in $\log [\text{H}^+]$ caused by the non-ideal Nernst slope.



Measured potential of the liquid junction free cell vs. calculated $\log(a_{\text{HCl}})$ for 3 series of standard solutions (slope = 58.8 mV/ $\log(a_{\text{HCl}})$; intercept = 397 mV)

Fig. 1.2
(from [94FAN/KIM])

Vitorge et al. [86GRE/ROB, 90RIG] calibrate their electrodes in the acidic range with $\text{HClO}_4/\text{NaClO}_4$ solutions and in the neutral to alkaline range with carbonate buffers (accepting the NEA-TDB auxiliary data for H_2CO_3 dissociation constants in 3 M NaClO_4). By this way, of course they observe an apparently ideal Nernst slope. But actually they just turn in a circle, because the auxiliary data refer to incorrect literature data determined with glass electrodes calibrated only in the acidic range and assuming an ideal Nernst slope.

Vitorge's comment, that we should have checked our electrodes before use or that we should have used other types of electrodes, and his statement that our pH measurements include systematic errors, are completely inadequate.

H₂CO₃ dissociation constants in NaCl and NaClO₄ solution

In the literature, there are very careful experimental studies on the dissociation constants of carbonic acid, usually in diluted chloride solutions (c.f. refs. of Harned et al and others given in [96FAN/NEC]). In these studies, pH or H⁺ concentrations are usually measured with Pt/H₂ and chloride electrodes in cells without liquid junction. Based on these studies, the equilibrium constants at I = 0 are well known and accepted as standard values (c.f. NBS tables [82WAG/EVA], NEA-TDB [92GRE/FUG, 95SIL/BID], model of Harvie, Møller and Weare for the seawater salt system [84HAR/MØL]).

In NaCl solution, there are numerous experimental data of different type, which allow the determination of the Pitzer parameters in the system Na-H-OH-HCO₃-CO₃-Cl-H₂O at 25°C [91PIT, 84HAR/MØL]. In Fig.1.3 these data are used to calculate the equilibrium constants for the reactions



and



as a function of the NaCl molality. Fig.1.3 shows also that the corresponding calculation with the SIT parameters given in the NEA-TDB is almost identical (if we disregard the deviations at high ionic strength). In addition, Fig.1.3 shows that the corresponding conditional constants determined in [94RUN/KIM, 94NEC/KIM] as accompanying work within the Np(V) carbonate solubility studies are in reasonable agreement with the known and generally accepted literature data. This again confirms the experimental procedures (including pH calibration) applied by Neck and Runde.

NaCl solution

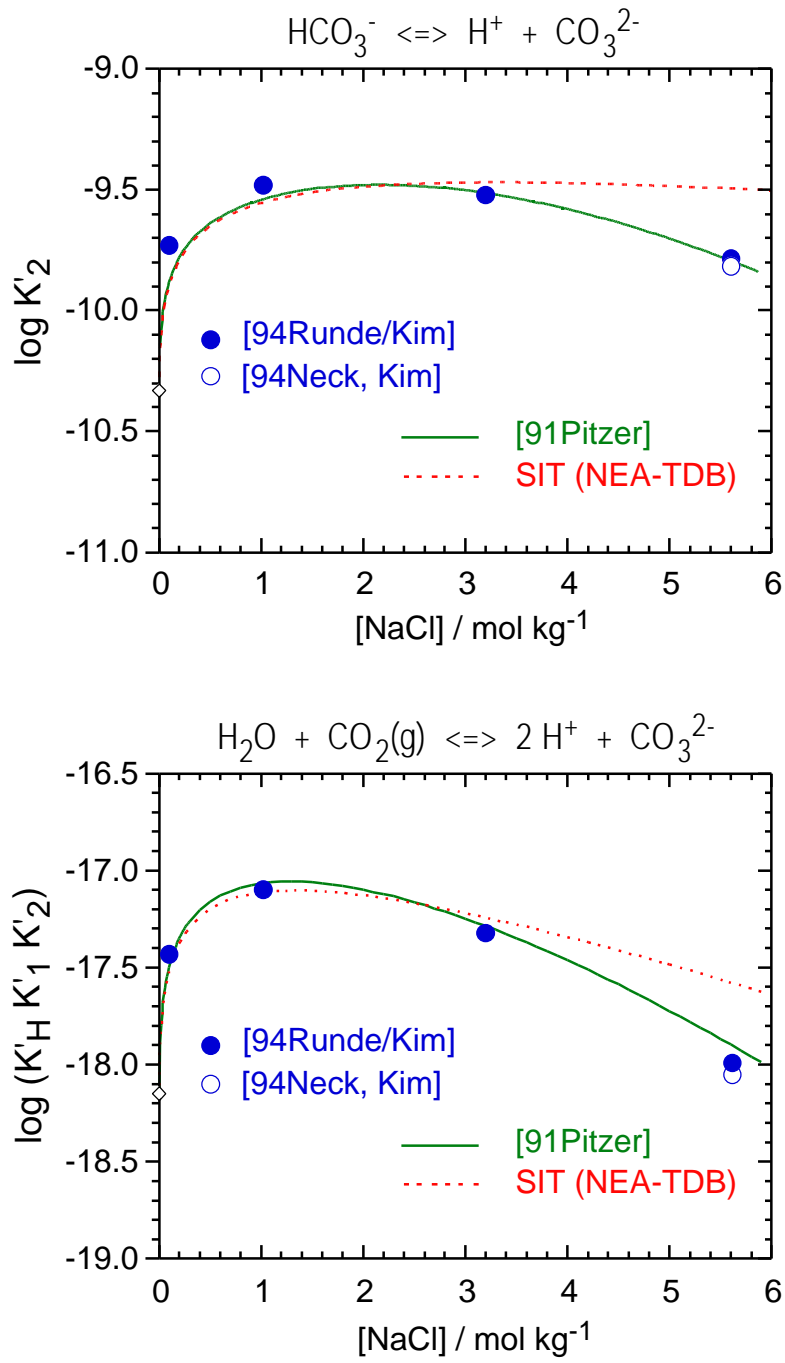


Fig. 1.3

Fig.1.4. shows the corresponding equilibrium constants in NaClO₄ solution. In contrast to the good agreement in NaCl, the experimental results and the Pitzer modeling in [94NEC/RUN, 96FAN/NEC] are strongly conflicting with literature values and with the SIT calculation based on the NEA-TDB parameters (Solely the values in 0.1 M NaClO₄, where the effect of SIT coefficients is negligible, are in agreement). However, as already mentioned above and discussed in [96FAN/NEC], except of one log K₂ value, all literature data shown in Fig.1.4 (the references are given in [96FAN/NEC]) were determined with glass electrodes calibrated only in the acidic range, and extrapolated to the alkaline range by assuming ideal Nernst slopes. That means: all these data are incorrect. And even if so many authors obtain comparable results by making the same mistake, a mistake will always remain a mistake.

(The consequences will be discussed in section 2 of the present manuscript.)

This mistake is illustrated in Figs.1.5 and 1.6. The shown real example is taken from our FZKA report (Neck, Fanghänel, Kim, FZKA 5599, June 1995).

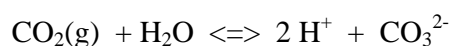
Fig.1.5 shows the two different ways of electrode calibration in 3 M NaClO₄, either the calibration with both, acidic HClO₄/NaClO₄ and alkaline NaOH/NaClO₄ solutions (solid line), or the calibration only in the acidic range extrapolated with the ideal Nernst slope (dashed line). In Fig.1.6, the value of log K₂ is determined in a most simple way, by measuring log[H⁺] in a solution with the composition:

0.05 M NaHCO₃ / 0.05 M Na₂CO₃ / 2.85 M NaClO₄

i.e. with log [HCO₃⁻] = log [CO₃²⁻] and hence: log [H⁺] = log K₂

If we would have accepted the dashed line for electrode calibration, the value of log K₂ = -9.64 would have been obtained, which would have been very well consistent with the literature data and with the auxiliary value of -9.62 from the NEA-TDB. However, the dashed calibration line is not correct - the solid line represents the correct calibration and hence it follows that log K₂ = -9.81.

Due to the incorrect pH calibration, the errors in log K₂ are about 0.1 - 0.2 log units. The errors in log K_HK₁K₂ are about the double, because two H⁺ ions are involved in the equilibrium



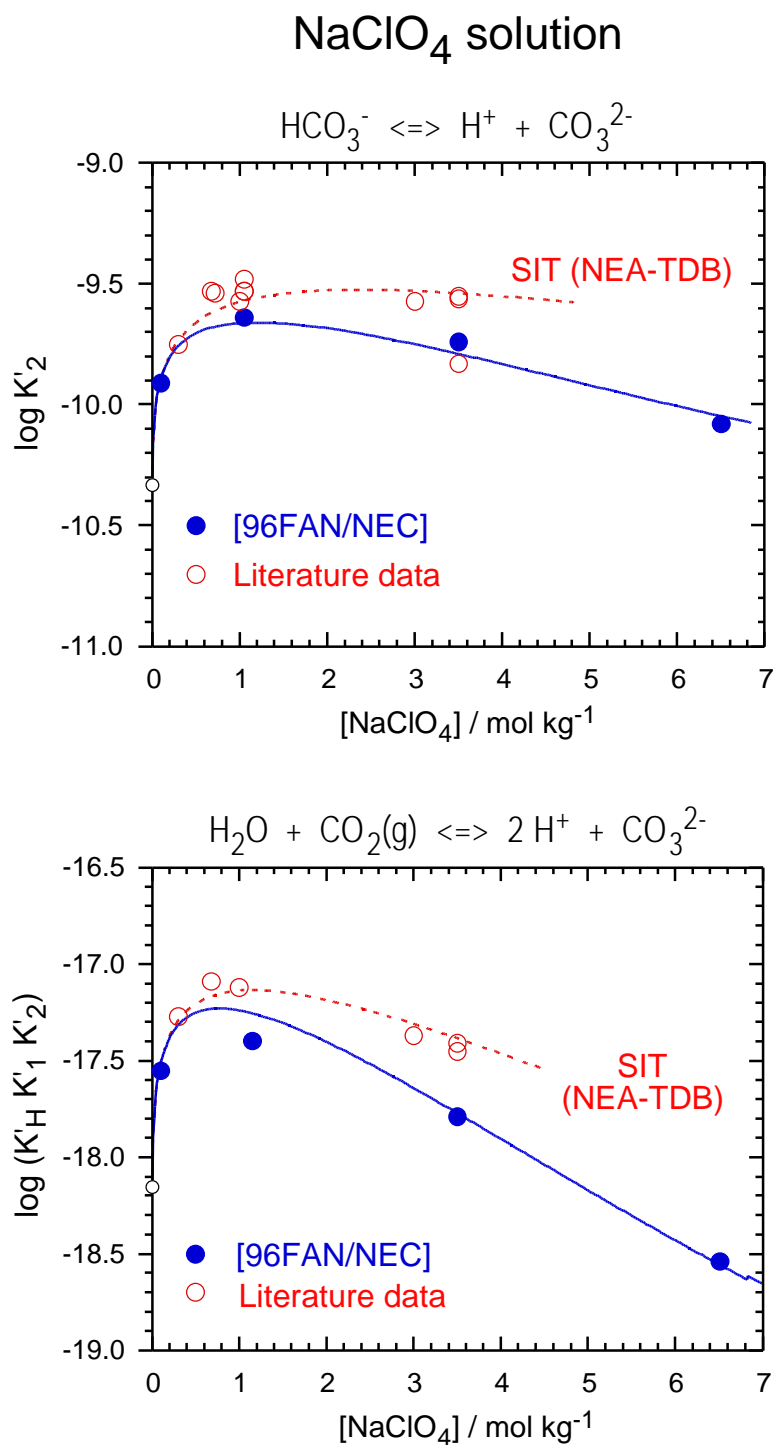


Fig. 1.4

Calibration of the pH glass electrode

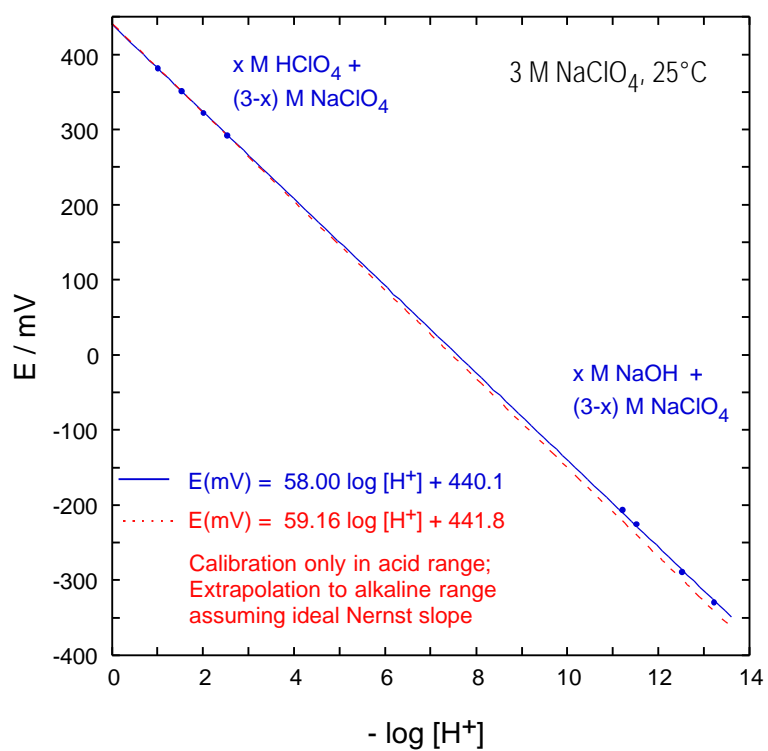


Fig. 1.5

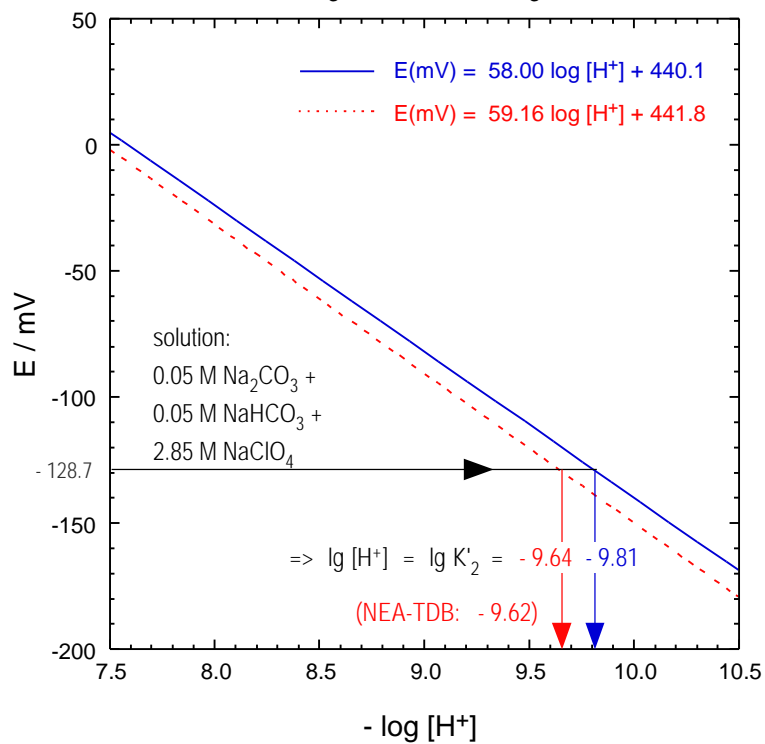
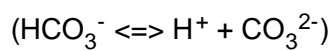
Determination of $\log K'_2$ in 3 M NaClO₄

Fig.1.6

Question: are the carbonate concentrations determined in experimental studies on actinide carbonates correct or incorrect ?

In the Np/Pu draft, the reviewer is irritated, because the Np(V) carbonate solubilities of [91KIM/KLE, 94NEC/RUN] and those of [83MAY] are in good agreement, although considerably different constants are used to calculate the carbonate concentration from the measured values of H^+ concentration:

page 806, lines 23 - 27

The solubility values in 1 M NaClO₄ aqueous solutions were the same as those in [83MAY], and this may be coincidental as different values were used for the protonation constants for carbonate ion.

The agreement between the Np(V) carbonate solubilities in [94NEC/RUN] and those in [83MAY] (in 1 M NaClO₄) and [86GRE/ROB] (in 3 M NaClO₄) is of course not accidentally as supposed by Vitorge in the NEA review. This is demonstrated below for a carbonate solution in 3 M NaClO₄. Because of the different pH calibration methods, Vitorge et al. (ideal Nernst slope) and Neck (58.0 mV/pH, Fig.1.6) determine a considerably different H^+ concentration. However, because they apply different H₂CO₃ dissociation constants in 3 M NaClO₄ (log $K_H K_1 K_2 = -17.62$ [86GRE/ROB, 90RIG] and -17.99 [94NEC/RUN]) they finally obtain practically the same (correct) values of log $[CO_3^{2-}]$.

Example: Solution in 3 M NaClO₄
equilibrated with atmospherical CO₂(g) partial pressure (log $pCO_2 = -3.52$)
 $[HCO_3^-] = [CO_3^{2-}]$

Vitorge et al.: log $[H^+] = \log K_2 = -9.62$
log $[CO_3^{2-}] = -17.62 - 3.52 - 2 \log [H^+] = -1.90$

Neck: log $[H^+] = \log K_2 = -9.81$
log $[CO_3^{2-}] = -17.99 - 3.52 - 2 \log [H^+] = -1.89$

Of course, in acidic solutions the calculated carbonate concentrations would be different, e.g. in 0.01 M HClO₄/ 2.99 M NaClO₄, Vitorge and Neck would measure the same value of log $[H^+] = -2.00$ but different values of log $[CO_3^{2-}] = -17.14$ and -17.51 , respectively. However, experiments on actinide carbonates are usually restricted to the limited range of log $[H^+] = -6.5$ to -10.5 , and the errors in the carbonate concentration remain within the range of other experimental uncertainties.

=> General comment on “recalculations“ of solubility or complexation constants for actinide carbonates

In the literature and also in the NEA reviews there are often recalculations of original data on solid

or aqueous actinide carbonates, and the following reason is given:

The authors XY used the constant of $\log x$ to calculate the CO_3^{2-} concentration from the H^+ concentration. However the correct or updated value would be $\log x'$. Therefore the solubility or complexation constant given in the original paper is recalculated using the value of $\log x'$ and the following values are obtained ...

In many cases such recalculations do not correct the original data. Just in contrast - they make them incorrect,

a) if, in a first step, the authors use carbonate auxiliary data to calibrate their pH electrode with carbonate buffers and then, in the second step, they use the same auxiliary data again to calculate $\log[\text{CO}_3^{2-}]$ from their measured $\log[\text{H}^+]$ concentration.

b) if, in a first step, the authors determine themselves (with their method of pH calibration) the relation between $\log[\text{CO}_3^{2-}]$ and $\log[\text{H}^+]$ or pH and then, in a second step, they again use this relation to calculate $\log[\text{CO}_3^{2-}]$ from their measured values of $\log[\text{H}^+]$ or pH.

In both cases, the procedure is internally consistent. The measured values of $\log[\text{H}^+]$ may be incorrect (e.g. those of Vitorge et al. in 3 M NaClO_4 , c.f. example above), but the values of $\log[\text{CO}_3^{2-}]$ are correct. The mentioned recalculations are always restricted to the correction of the second step, they never correct the first step, which would be necessary as well. And hence they lead to incorrect results !

2. Erroneous conclusions on actinide carbonates, which arise from limitations/shortcomings of the SIT and erroneous NEA-TDB auxiliary data on trace activity coefficients of the carbonate ion (in NaClO₄ solutions above 1 molal)

2.1. General comments on shortcomings of the SIT

2.2. Actinide carbonates: consequences for solubility constants at $I = 0$

Np(V) $\text{NaNpO}_2\text{CO}_3(\text{s})$, $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$,

Am(III) $\text{Am}_2(\text{CO}_3)_3(\text{s})$,

U(VI) $\text{UO}_2\text{CO}_3(\text{s})$

2.3. Proposal to solve the problem of carbonate trace activity coefficients

2.4. Consequences for NEA-TDB reviews

2.1. General comments on shortcomings of the SIT

First of all, it must be emphasized that we do not at all intend to replace the SIT procedure manifested in the NEA-TDB by introducing Pitzer modelling. (It seems that members of the NEA review groups misinterpreted our objections that way.) However, it is unacceptable that shortcomings and limitations of the SIT are simply ignored, with the consequences that:

- (1) erroneous chemical conclusions are drawn and incorrect thermodynamic data are selected because of these shortcomings
- (2) correct experimental data are ignored or, even worse, criticized as not reliable, because they are not consistent (or better: cannot be explained) with the simplified SIT approach used in the NEA-TDB.

Limitations of the SIT

- 1) No triple ion interactions (as included in the Pitzer equations)
=> inaccuracies at high ionic strength ($I > 4$ m)
- 2) Debye-Hückel equation with a fixed value of $B\alpha = 1.5$
=> inaccuracies at low ionic strength ($I \rightarrow 0$) for ions with high charge $|z| > 3$
- 3) Simplification: negligible anion-anion and cation-cation interactions
=> general problem, which makes it impossible to use the same SIT coefficients for carbonate trace activity coefficients in different electrolyte media

The shortcoming 1) is well-known, and the resulting inaccuracies may be acceptable. The shortcoming 2) can lead to problems and needs further discussion in the future. However these two shortcomings are not the issue of our objections and not further discussed or criticized in the present manuscript. The mentioned objections, which lead to misinterpretations and errors (not inaccuracies, but actually errors!) are exclusively based on shortcoming 3).

Therefore it is now clearly stated once again: the scientific problems do not arise as a question of SIT or Pitzer modelling - they arise from experimental data! In some cases we refer to the Pitzer modelling performed in our papers [95NEC/FAN, 95FAN/NEC, 96FAN/NEC], but this is only done in order to demonstrate the errors coming from the oversimplification of the SIT procedure. At the end of this manuscript, a possible way is shown, how these problems could be solved as well by using an extended SIT formalism, which is not in contradiction to the NEA-TDB guidelines.

According to the results in [96FAN/NEC], where the mixing parameters for CO_3^{2-} and HCO_3^- in NaClO_4 solution were evaluated from H_2CO_3 dissociation constants, the trace activity coefficients of the CO_3^{2-} in NaCl and NaClO_4 solutions above 1 molal are considerably different (Fig.2.1, next page). These differences cannot be described with the simple SIT approach used in the NEA

reviews, because (in contrast to the Pitzer equations) interaction coefficients between ions of the same charge sign are generally set equal to zero. (However, these interactions are not zero, and hence included in the binary cation/anion coefficients.) As a consequence, with the SIT approach equal CO_3^{2-} trace activity coefficients are calculated for NaCl and NaClO_4 solutions of equal molality.

In order to demonstrate that different carbonate trace activity coefficients in NaCl and NaClO_4 solutions are not an artefact arising from erroneous experimental data, we worked out two further corresponding examples of well-known and experimentally well ascertained different trace activity coefficients for an ion dissolved in different media.:

1) Trace activity coefficients of the SO_4^{2-} ion in NaCl and NaTcO_4 solution, which is quite analog to the problem of CO_3^{2-} in NaCl and NaClO_4 solution (Fig.2.2).

Unfortunately there are no data for the system Na- SO_4 - ClO_4 , but the TcO_4^- ion has properties very similar to those of the ClO_4^- ion. The Pitzer parameters in the systems Na- SO_4 -Cl^{a)} and Na- SO_4 - TcO_4 ^{b)} are very well known and ascertained by numerous isopiestic and solubility data. The results are shown on the next page. They show that the corresponding differences in the trace activity coefficients of CO_3^{2-} and SO_4^{2-} are comparable, not only qualitatively but also quantitatively.

^{a)} Harvie, Møller, Weare, *Geochim. Cosmochim. Acta.* 48 (1984), 723

^{b)} Neck, Könnecke, Fanghänel, Kim, J. *Solution Chem.* 27 (1998), 107 and
Neck, Könnecke, Fanghänel, Kim, *Radiochim. Acta* 83 (1998), 75.

2) Trace activity coefficients of the H^+ ion in NaCl and CsCl solution.

The values shown in Fig.2.3 are calculated with the Pitzer parameters in [91PIT] evaluated from the experimental emf data.

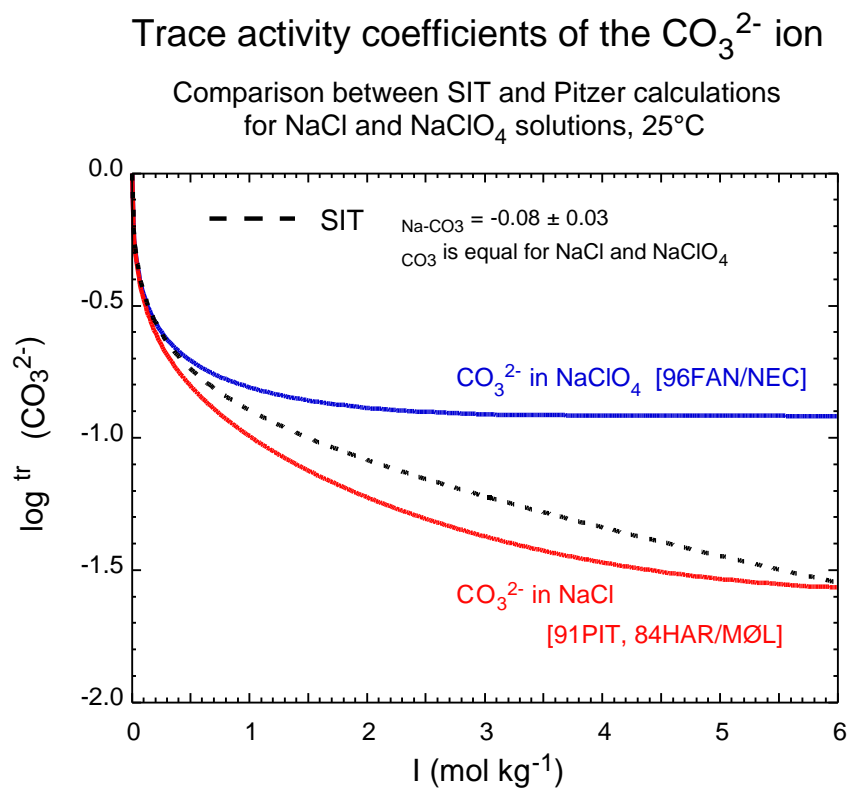


Fig. 2.1

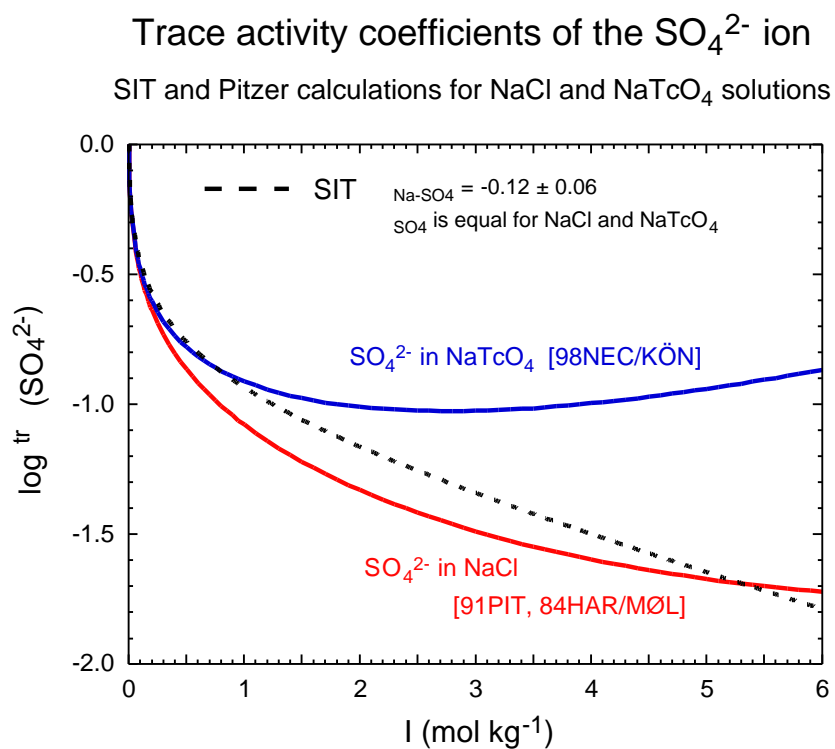


Fig. 2.2

Trace activity coefficients of the H^+ ion in NaCl and CsCl solutions

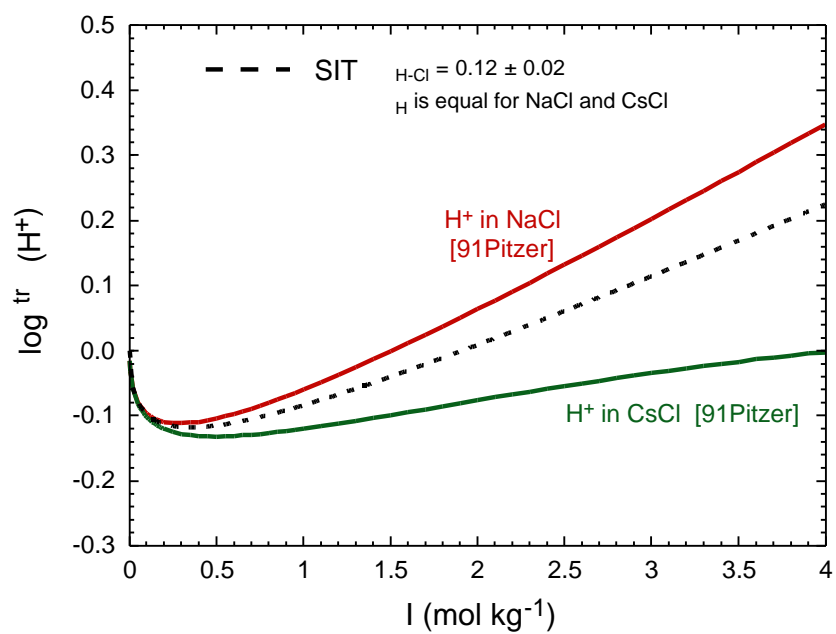
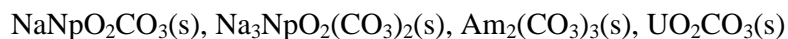


Fig. 2.3

2.2. Actinide carbonate solids of Np(V), Am(III), U(VI); consequences for solubility constants at I = 0



Since the carbonate trace activity coefficient is directly involved in the solubility constants for actinide carbonate solid phases (and of course also in the formation constants of aqueous carbonate complexes) at I = 0, the consequences of the different set of auxiliary data is shown in the following sections. Because of the limited time, the present calculations are restricted to the solid phases $\text{NaNpO}_2\text{CO}_3(\text{s})$, $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$, $\text{Am}_2(\text{CO}_3)_3(\text{s})$ and $\text{UO}_2\text{CO}_3(\text{s})$

Neptunium(V)

Solubility constant for $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$

In the NEA review, the results of Kim and coworkers [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC] as well as those of former coworkers of this group [94MEI, 96RUN/NEU] were generally criticized to be not reliable and disregarded (open points in Fig.2.4). Particularly in the case of the data in NaCl solution and in 0.1 M NaClO_4 , this is a pure arbitrary act of the reviewer, since all auxiliary data used in these studies (ion product of water, H_2CO_3 dissociation constants) are consistent with the NEA-TDB.

The evaluation of the solubility constant at I = 0 is primarily based on solubility studies of Maya [83MAY] (in 1 M NaClO_4) and Vitorge (in 3 M NaClO_4) reported in [86GRE/ROB] and later again in [90RIG]. Further, the reviewer applied the SIT coefficients from the NEA-TDB ($(\text{Na}^+/\text{ClO}_4^-) = 0.01$, $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$ and $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.25$ [85SIL/BID]), and hence $\gamma_{\text{NpO}_2^+} = 0.18$) to calculate the constants at I = 0.

As a consequence, the reviewer concludes that the results of Maya [83MAY] refer to a hydrated solid phase and those of Vitorge [86GRE/ROB] to an aged, less hydrated solid phase with

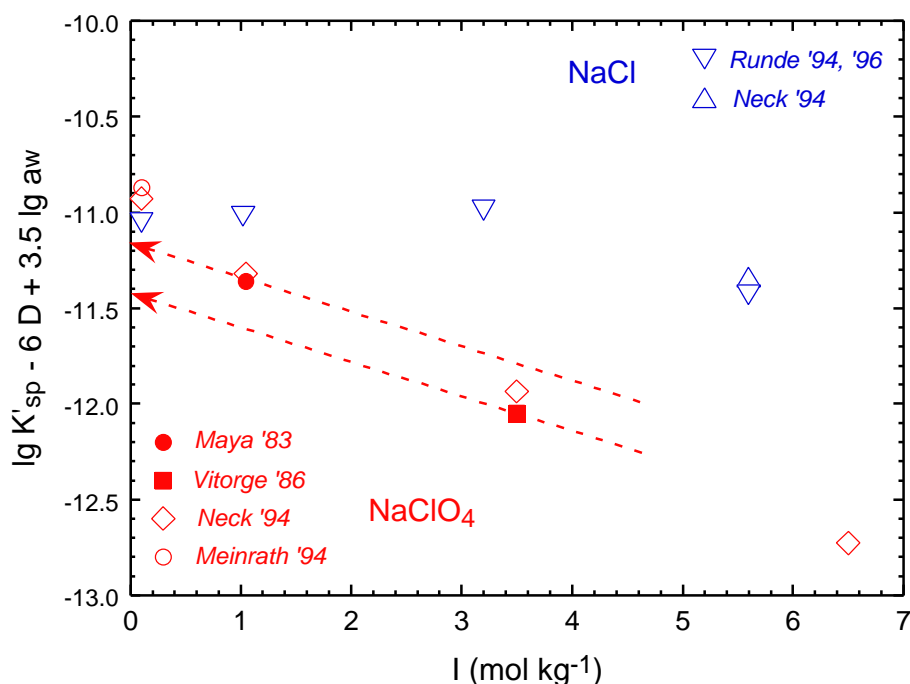
$$\log K_s^\circ = -11.16 \pm 0.35 \text{ for } \text{NaNpO}_2\text{CO}_3 \cdot 3.5\text{H}_2\text{O}(\text{s})$$

and

$$\log K_s^\circ = -11.66 \pm 0.50 \text{ for } \text{NaNpO}_2\text{CO}_3(\text{s}), \text{ aged}$$

NEA review (VITORGE)

Solubility product of $\text{NaNpO}_2\text{CO}_3 \cdot x \text{H}_2\text{O}$ ($x = 3.5$ and $x = 0$)



- Vitorge: 1) Data of Kim's group (open points) are not considered !
 2) SIT coefficients are taken from the NEA-TDB

---- NEA-TDB: $\text{NpO}_2^+ / \text{ClO}_4^- = 0.25$
 $= 0.18 \pm 0.06$ $\text{Na}^+ / \text{ClO}_4^- = 0.01$
 $\text{Na}^+ / \text{CO}_3^{2-} = -0.08$

Fig. 2.4

In Fig.2.4, the original $\log K_s$ value given in [86GRE/ROB, 90RIG] is used to extrapolate Vitorge's data in 3.5 m NaClO_4 to $I = 0$. In the NEA review a lower value is evaluated (c.f. discussion of [90RIG], p.795-798, Appendix A). However, this recalculation is somewhat speculative. It is based on the assumptions that the scattering of the experimental data is due to the presence of different solid phases and that the lowest solubility data refer to aged $\text{NaNpO}_2\text{CO}_3(\text{s})$.

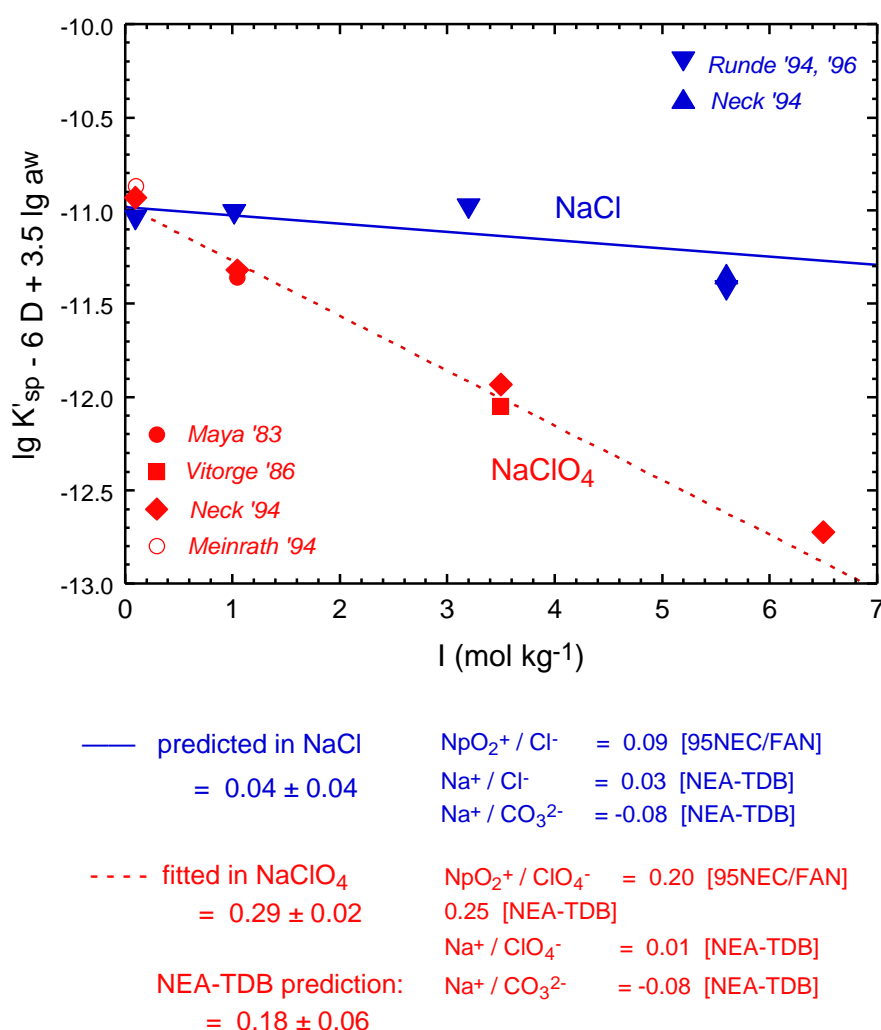
Solubility product of $\text{NaNpO}_2\text{CO}_3 \cdot 3.5 \text{H}_2\text{O}$ Application of the SIT: $\log K_{\text{sp}}^\circ = -11.0 \pm 0.2$ (2)

Fig. 2.5

Fig. 2.5 shows the SIT extrapolation to $I = 0$, including as well the results of Kim and coworkers. It is to note that the solid line (for the data in NaCl solution) is predicted by independent SIT coefficients: $(\text{Na}^+/\text{Cl}^-) = 0.03$ and $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$ from the NEA-TDB [85SIL/BID], $(\text{NpO}_2^+/\text{Cl}^-) = 0.09$ from a solvent extraction study in [95NEC/FAN], and hence $= 0.04$. For both media (NaCl and NaClO_4) the extrapolation to $I = 0$ leads to a consistent value of $\log K_{\text{sp}}^\circ = -11.0 \pm 0.2$.

As already documented in section 1.1 of this manuscript, the solubility data in [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC, 96RUN/NEU] refer to the same solid phase, the hydrated $\text{NaNpO}_2\text{CO}_3 \cdot 3.5\text{H}_2\text{O}(\text{s})$ described by Volkov et al. [77VOL/VIS] and Maya [83MAY]. (Solely Meinrath [94MEI] reported an other

(hexagonal) modification of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$, with different x-ray pattern).

However, as expected according to section 1.2, the value of $\gamma_{\text{Na}^+} = 0.29$ in NaClO_4 solution (dashed line) deviates significantly from the NEA-TDB prediction, because the value of $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$ is appropriate for the carbonate trace activity coefficients in NaCl , but not for those in NaClO_4 solution (c.f. discussion in 2.1). The dashed line would be predicted with $(\text{Na}^+/\text{CO}_3^{2-}) = +0.04$ (in NaClO_4 solution), evaluated in section 2.3 from the H_2CO_3 dissociation constants determined in [96NEC/FAN].

In the table below (next page) auxiliary data given in the NEA-TDB for $(\text{Na}^+/\text{Cl}^-)$, $(\text{Na}^+/\text{ClO}_4^-)$ and $(\text{Na}^+/\text{CO}_3^{2-})$, together with $(\text{NpO}_2^+/\text{Cl}^-) = 0.09$ and $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.20$ are used to calculate the solubility product of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ at $I = 0$. It becomes evident that (within the range of experimental uncertainties) all studies in NaCl solution lead to a consistent value of $\log K_s^\circ$, in particular if we assume the hydration number of $x = 3.5$ given in [83MAY].

In the same table, analogous calculations are done with the data published in NaClO_4 solution. Again the auxiliary interaction coefficients from the NEA-TDB are used (with the exception that $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.20$ is used, instead of 0.25 as proposed by Vitorge, but this has only a rather limited impact on the calculations). In contrast to the observations in NaCl solution, the calculated $\log K_s^\circ$ values are not consistent. They decrease systematically with increasing NaClO_4 concentration.

Remember: $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$, the SIT coefficient of the NEA-TDB was not appropriate for the carbonate trace activity coefficients in NaClO_4 solution above 1 molal (Fig. 2.1).

In the next table, we use the Pitzer parameters for the carbonate ion (from [91PIT, 96FAN/NEC] and for the NpO_2^+ ion from [95NEC/FAN] to calculate $\log K_s^\circ$.

Now, consistent values at $I = 0$ are obtained for 12 solubility experiments from 5 different investigators with the solution composition widely varied ($I = 0.1, 1, 3$ and 5 M in both, NaCl and NaClO_4 solutions). The average value is found to be

$$\log K_s^\circ = -11.08 \pm 0.20 \quad (2) \quad \text{for } \text{NaNpO}_2\text{CO}_3 \cdot 3.5\text{H}_2\text{O}(\text{s})$$

It is to note that the SIT or Pitzer coefficients for the NpO_2^+ ion are based on the same experimental input data. That means: the only essential difference between SIT and Pitzer calculations is that for CO_3^{2-} !

The consistency of the values calculated for $I = 0$ may be considered as a corroboration of the activity coefficients used, which justifies the SIT extrapolation in Fig.2.5.

Solubility products of hydrated $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ at $20 - 25^\circ\text{C}$

SIT calculation with auxiliary data from the NEA-TDB

Ref.	Medium	$\log K_s$ (molar)	$\log K_s^\circ (I = 0)$ ^{a)}	
			$x = 0$	$x = 3.5$

a) in 0.1 - 5 M NaCl

[96RUN/NEU]	0.1 M NaCl	-10.40	-11.05	-11.05
[94RUN/KIM, 96RUN/NEU]	1.0 M NaCl	- 9.77	-10.93	-10.98
[96RUN/NEU]	3.0 M NaCl	- 9.40	-10.67	-10.86
[94RUN/KIM, 95NEC/RUN]	5.0 M NaCl	- 9.61	(-10.83	-11.21) ^{b)}
[94NEC/KIM]	5.0 M NaCl	- 9.52	(-10.74	-11.12) ^{b)}

b) in 0.1 - 5 M NaClO₄

[91KIM/KLE, 94NEC/RUN]	0.1 M NaClO ₄ -10.28	-10.92	-10.92
[94MEI]	0.1 M NaClO ₄ -10.22	-10.86	-10.86
[83MAY]	1.0 M NaClO ₄ -10.14	-11.18	-11.23
[91KIM/KLE, 94NEC/RUN]	1.0 M NaClO ₄ -10.10	-11.14	-11.19
[86GRE/ROB]	3.0 M NaClO ₄ -10.56	-11.41	-11.60
[91KIM/KLE, 94NEC/RUN]	3.0 M NaClO ₄ -10.45	-11.30	-11.49
[94NEC/RUN]	5.0 M NaClO ₄ -11.06	(-11.49	-11.87) ^{b)}

^{a)} NEA-TDB [95SIL/BID]: ($\text{Na}^+/\text{ClO}_4^-$) = 0.01, (Na^+/Cl^-) = 0.03, ($\text{Na}^+/\text{CO}_3^{2-}$) = -0.08
($\text{NpO}_2^+/\text{ClO}_4^-$) = 0.20, and ($\text{NpO}_2^+/\text{Cl}^-$) = 0.09

^{b)} at this concentration the SIT may become inaccurate

Solubility products of hydrated $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$ at 20 - 25°C

Pitzer calculation (Fanghänel, Neck)

Ref.	Medium	$\log K_s$ (molar)	$\log K_s^\circ (I = 0)$ ^{a)} x = 3.5
[96RUN/NEU]	0.1 M NaCl	-10.40	-11.08
[94RUN/KIM, 96RUN/NEU]	1.0 M NaCl	- 9.77	-11.10
[96RUN/NEU]	3.0 M NaCl	- 9.40	-11.00
[94RUN/KIM, 95NEC/RUN]	5.0 M NaCl	- 9.61	-11.15
[94NEC/KIM]	5.0 M NaCl	- 9.52	-11.06
[91KIM/KLE, 94NEC/RUN]	0.1 M NaClO ₄ -10.28	-10.94	
[94MEI]	0.1 M NaClO ₄ -10.22	-10.88	
[83MAY]	1.0 M NaClO ₄ -10.14	-11.18	
[91KIM/KLE, 94NEC/RUN]	1.0 M NaClO ₄ -10.10	-11.14	
[86GRE/ROB]	3.0 M NaClO ₄ -10.56	-11.25	
[91KIM/KLE, 94NEC/RUN]	3.0 M NaClO ₄ -10.45	-11.14	
[94NEC/RUN]	5.0 M NaClO ₄ -11.06	-11.08	

^{a)} calculated with Pitzer parameters given in [95NEC/FAN, 96FAN/NEC]

Solubility constant for $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$

In the table below, the experimental solubility constants reported for $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$ in 1, 3 and 5 M NaClO_4 and in 5 M NaCl [86GRE/ROB, 91KIM/KLE, 94RUN/KIM, 95NEC/RUN] are extrapolated to $I = 0$. The results are comparable with those observed for $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O}(\text{s})$. If SIT coefficients are used, with CO_3^{2-} according to the NEA-TDB, the calculated $\log K_s^\circ$ values decrease systematically with increasing NaClO_4 concentration, whereas the Pitzer parameters given in [95NEC/FAN, 96FAN/NEC] lead to consistent values of $\log K_s^\circ$.

Again it is to note that the SIT or Pitzer coefficients for the NpO_2^+ ion are based on the same experimental input data. That means: the only essential difference between SIT and Pitzer calculations is that for CO_3^{2-} !!!

Solubility product of $\text{Na}_3\text{NpO}_2(\text{CO}_3)_2(\text{s})$ at 20 - 25°C

Ref.	Medium	$\log K_s$ (molar)	$\log K_s^\circ (I = 0)$	
			SIT ^{a)}	Pitzer ^{c)}
[94NEC/RUN]	1.0 M NaClO_4 -12.23	-14.5	-14.5	
[86GRE/ROB]	3.0 M NaClO_4 -12.44	-14.8	-14.2	
[94NEC/RUN]	3.0 M NaClO_4 -12.59	-14.9	-14.4	
[94NEC/RUN]	5.0 M NaClO_4 -13.57	(-15.7) ^{b)}	-14.3	
[94RUN/KIM] [95NEC/RUN]	5.0 M NaCl	-11.46	(-14.2) ^{b)}	-14.2

^{a)} NEA-TDB [95SIL/BID]: $(\text{Na}^+/\text{ClO}_4^-) = 0.01$, $(\text{Na}^+/\text{Cl}^-) = 0.03$, $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$
 $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.20$, and $(\text{NpO}_2^+/\text{Cl}^-) = 0.09$

^{b)} at this concentration the SIT may become inaccurate

^{c)} calculated with Pitzer parameters given in [95NEC/FAN, 96FAN/NEC]

Note: The only essential difference between SIT and Pitzer calculations is that for CO_3^{2-} !!!

Americium(III)

For the hydrated $\text{Am}_2(\text{CO}_3)_3 \cdot x\text{H}_2\text{O}(\text{s})$ there are experimental results from Meinrath and Runde (Kim's research group) in 0.1 M NaClO_4 and from Robouch (Vitorge's research group) in 3.0 M NaClO_4 . These data were extrapolated to $I = 0$ in the NEA review [95SIL/BID], and the resulting values were found to be strongly inconsistent. Since it is not clear whether more or less crystalline solid phases were obtained in these studies, the NEA review recommended an average value with a large uncertainty:

$$\log K^\circ_s = -16.7 \pm 1.1 \text{ [95SIL/BID]}$$

If we use the CO_3^{2-} values from [96FAN/NEC] to calculate the equilibrium constants at $I = 0$, the consistency is considerably increased, and the mean value of the two research groups has a significantly smaller uncertainty:

$$\log K^\circ_s = -16.6 \pm 0.4 \text{ (2)}$$

Hydrated $\text{Am}_2(\text{CO}_3)_3 \cdot x\text{H}_2\text{O}(\text{s})$ at 20 - 25°C;

Solubility constant $\log K_s$ for the reaction: $0.5 \text{ Am}_2(\text{CO}_3)_3(\text{s}) \rightleftharpoons \text{Am}^{3+} + 1.5 \text{ CO}_3^{2-}$

Ref.	Medium	$\log K_s$ (molar)	$\log K^\circ_s (I = 0)$	
			SIT NEA-TDB ^{a)}	CO_3^{2-} from [96FAN/NEC] ^{b)}
[92RUN/MEI]	0.1 M NaClO_4 - 14.73	- 16.33	- 16.3	
[91MEI/KIM]	0.1 M NaClO_4 - 14.90	- 16.50	- 16.5	
[91MEI/KIM2]	0.1 M NaClO_4 - 14.79	- 16.39	- 16.4	
[89ROB]	3.0 M NaClO_4 - 15.08	- 17.36	- 16.8	
		(-15.27) ^{c)}	(-17.54) ^{c)}	

^{a)} calculated with $(\text{Am}^{3+}/\text{ClO}_4^-) = 0.49$ and $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$ [95SIL/BID], for $x = 0$

^{b)} Am^{3+} is the same as in a), whereas CO_3^{2-} is taken from [96FAN/NEC]

^{c)} recalculated in [95SIL/BID]. The original values in [89ROB] are corrected by using a somewhat different H_2CO_3 dissociation constant to calculate $\log [\text{CO}_3^{2-}]$ from $\log [\text{H}^+]$. However, since the calibration of Robouch is internally consistent, such a recalculation does not "correct" the data of Robouch. It makes them incorrect.

Uranium(VI)

In the case of the solubility product of $\text{UO}_2\text{CO}_3(\text{s})$ at 20 -25°C, the experimental data published before 1992 are discussed and extrapolated to $I = 0$ in the previous NEA review [92GRE/FUG]. In the literature published later, there are numerous studies in 0.1 M NaClO_4 , which can be divided into two groups. One set of results leads to $\log K_s^\circ = -14.2 \pm 0.2$, which is in the range of the previous results discussed in [92GRE/FUG]. The other set of results, from laboratories in Japan (Meinrath, Kimura, Kato) lead to a significant lower solubility constant of $\log K_s^\circ = -14.9 \pm 0.2$.

Independent of the auxiliary data used to calculate CO_3^{2-} in NaClO_4 solution (either SIT coefficients from [95SIL/BID] or Pitzer parameters from [96FAN/NEC]), the calculated values of $\log K_s^\circ$ are ranging from -14.1 to -14.5, or even to -14.9 ± 0.2 if we include the data from [93MEI/KIM, 93MEI/KIM2, 96MEI/KAT, 96KAT/KIM], respectively.

From the data reported in [72/SER/NIK, 76NIK2, 84GRE/FER, 92KRA/BIS, 93PAS/RUN, 96MEI/KLE], the following unweighted overall mean values (± 2) are obtained:

$$\log K_s^\circ = -14.32 \pm 0.30 \quad (\text{with } \text{CO}_3^{2-} \text{ according to [95SIL/BID]})$$

or

$$\log K_s^\circ = -14.27 \pm 0.28 \quad (\text{with } \text{CO}_3^{2-} \text{ according to [96FAN/NEC]})$$

The consistency obtained with CO_3^{2-} according to [95SIL/BID] or according to [96FAN/NEC] is approximately the same.

Solubility product of $\text{UO}_2\text{CO}_3(\text{s})$ at 20 -25°C

Ref.	Medium	log K _s (molar)	log K ^o _s (I = 0)	
			SIT NEA-TDB ^{a)}	CO ₃ ²⁻ from [96FAN/NEC] ^{b)}
[92GRE/FUG] NEA-TDB review			-14.47 ± 0.04	
[72/SER/NIK]	I = 0.0002 - 0.02 M		-14.26 ± 0.3	-14.26 ± 0.3
[76NIK2]	I = 0.01 M	-14.15	-14.50	-14.50
[92KRA/BIS.]	0.1 M NaClO ₄	-13.29	-14.11	-14.12
[93PAS/RUN]		-13.35	-14.18	-14.19
[96MEI/KLE]		-13.50	-14.33	-14.34
[84GRE/FER]	0.5 M NaClO ₄	-13.21	-14.40	-14.37
[84GRE/FER]	3.0 M NaClO ₄	-13.94	-14.48	-14.12

New data

[92KRA/BIS]	0.1 M NaClO_4	-13.29	-14.11	
[93PAS/RUN]	"	-13.35	-14.18	
[96MEI/KLE]	"	-13.50	-14.33	
mean value (± 2)		-13.38 ± 0.22	-14.21 ± 0.22	-14.22 ± 0.22
[93MEI/KIM]	0.1 M NaClO_4	-13.89	-14.72	
[93MEI/KIM2]	"	-14.18	-15.01	
[96MEI/KAT]	"	-14.05	-14.88	
[96KAT/KIM]	"	-14.10	-14.93	
mean value (± 2)		-14.06 ± 0.24	-14.89 ± 0.24	-14.90 ± 0.24

^{a)} calculated with $(\text{UO}_2^{2+}/\text{ClO}_4^-) = 0.46$ [92GRE/FUG], $(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$ [95SIL/BID]

^{b)} UO_2^{2+} is the same as in a), whereas CO_3^{2-} is taken from [96FAN/NEC]

2.3. Proposal to solve the problem of carbonate trace activity coefficients:

In the tables above, the solubility constants at $I = 0$ for actinide carbonate solids are partly based on SIT activity coefficients and partly on Pitzer activity coefficients. Such a mixing is certainly not desirable and not acceptable for the NEA-TDB. Therefore, it is necessary to find a solution, which is based exclusively on the SIT formalism.

1) At first, we have to recognize that interactions anion/anion and cation/cation are not generally negligible as experimentally demonstrated (c.f. examples H^+ in NaCl / H^+ in CsCl and SO_4^{2-} in NaCl / SO_4^{2-} in NaTcO₄; Figs. 2.1 - 2.3.). If we consider traces of cation C or anion A in an electrolyte solution MX, the values of “ (C/X) ” and “ (A/M) ” actually represent the sums:

$$\begin{aligned} \text{“ (C/X) ”} &= (C/X) + (C/M) \\ \text{and} \\ \text{“ (A/M) ”} &= (A/M) + (A/X) \end{aligned}$$

2) In the original SIT equation neither these anion/anion and cation/cation interactions nor interaction between neutral solutes and ions are explicitly excluded. Setting (C/M) and (A/X) equal to zero is an (over)simplification, which fairly works in many cases. E.g. in the case of OH⁻ trace activity coefficients in NaCl and NaClO₄ solution. Therefore, fortunately, we do not have problems to describe the ion product of water or solubility constants for actinide hydroxides with the same value of “ (OH/Na⁺) ”. However, anion/anion interactions generally become important if the charge of an anion is -2 or larger, e.g. for SO_4^{2-} , CO_3^{2-} and the Np(V) carbonate complexes $NpO_2(CO_3)_n^{1-2n}$.

In order to avoid the explicit evaluation of values for anion/anion and cation/cation interactions, we could indicate that “ (C/X) ” and “ (A/M) ” values are valid for the medium MX by introducing the definitions:

$$\begin{aligned} (C/X)_{MX} &= (C/X) + (C/M) \\ (A/M)_{MX} &= (A/M) + (A/X) \end{aligned}$$

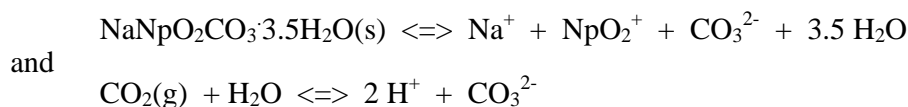
3) It is to note that in [80CIA] and the NEA-TDB, there are many SIT coefficients derived from binary system osmotic coefficients. In these cases, they actually represent (C/A) values. We neglect cation/cation interactions (C/M) and use (C/A) as (C/A)_{MA} to calculate trace activity coefficients of C in the medium MA (e.g. for C = UO₂²⁺, Nd³⁺ as analog for Am³⁺, and MA = NaCl or NaClO₄). If the medium cation is always M = Na⁺, we will probably not run into trouble, but we have to take care with data in other media, e.g. MgCl₂.

Trace activity coefficients of the carbonate ion in NaCl solution

In 0 - 3 M NaCl solution, the SIT coefficient proposed in the NEA-TDB:

$$(\text{Na}^+/\text{CO}_3^{2-})_{\text{NaCl}} = (\text{Na}^+/\text{CO}_3^{2-}) + (\text{Cl}^-/\text{CO}_3^{2-}) = \underline{-(0.08 \pm 0.03)} \text{ [95SIL/BID]}$$

accurately describes the experimental data for the equilibria



In Fig.2.1, there are certain differences, if (CO_3^{2-}) is calculated with this SIT coefficient or the known Pitzer parameters. The Pitzer activity coefficients would be better described with a SIT coefficient of $(\text{Na}^+/\text{CO}_3^{2-})_{\text{NaCl}} = -0.11$. These discrepancies are probably due to different exp. input data. (To a certain extent they might also be due to small differences in the splitting conventions because of the triple ion interactions or higher order terms in the Pitzer equations). However combined with the differences in (H^+) they cancel out for the equilibrium $\text{CO}_2(\text{g}) + \text{H}_2\text{O} \rightleftharpoons 2 \text{H}^+ + \text{CO}_3^{2-}$. For reasons of consistency the value of $(\text{Na}^+/\text{CO}_3^{2-})_{\text{NaCl}} = -(0.08 \pm 0.03)$ given in the NEA-TDB [95SIL/BID] should not be changed.

Trace activity coefficients of the carbonate ion in NaClO_4 solution

$$(\text{Na}^+/\text{CO}_3^{2-})_{\text{NaClO}_4} = (\text{Na}^+/\text{CO}_3^{2-}) + (\text{ClO}_4^-/\text{CO}_3^{2-})$$

For the two equilibria given above there are sufficient experimental data to determine the activity and SIT coefficients of the carbonate ion in NaClO_4 solution:

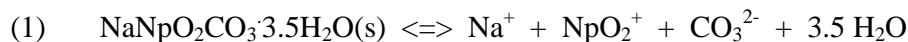


Fig. 2.5: SIT plot with exp. data from [83MAY, 86GRE/ROB, 94NEC/RUN, 94MEI]
This includes the assumption that the reported solubility data refer to the same solid phase. For the data from [83MAY, 94NEC/RUN] this is well ascertained.

$$\Rightarrow \log K^\circ = -11.0 \pm 0.2; (\quad)_{\text{NaClO}_4} = 0.29 \pm 0.02$$

$$\begin{aligned} (\text{Na}^+/\text{CO}_3^{2-})_{\text{NaClO}_4} &= (\quad)_{\text{NaClO}_4} - (\text{Na}^+/\text{ClO}_4^-) - (\text{NpO}_2^+/\text{ClO}_4^-) \\ &= (0.29 \pm 0.02) - (0.01 \pm 0.01) - (0.22 \pm 0.03)^* = \underline{+(0.06 \pm 0.04)} \end{aligned}$$

* mean value of $(\text{NpO}_2^+/\text{ClO}_4^-) = 0.25 \pm 0.05$ (NEA-TDB, from redox measurements of Vitorge et al.) and 0.20 ± 0.03 (from solvent extraction study in [95NEC/FAN])

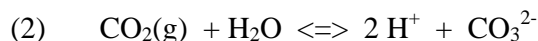


Fig. 2.6: SIT plot with exp. data from [96FAN/NEC]; $\log K^\circ$ fixed from NEA-TDB

$$\begin{aligned} (\text{Na}^+/\text{CO}_3^{2-})_{\text{NaClO}_4} &= (\quad)_{\text{NaClO}_4} - 2 (\text{H}^+/\text{ClO}_4^-) \\ &= (0.32 \pm 0.03) - 2 \cdot (0.14 \pm 0.02) = \underline{+(0.04 \pm 0.05)} \end{aligned}$$

The values derived from equilibria (1) and (2) are consistent and a mean value of $(\text{Na}^+/\text{CO}_3^{2-})_{\text{NaClO}_4} = +(0.05 \pm 0.05)$ could be proposed.

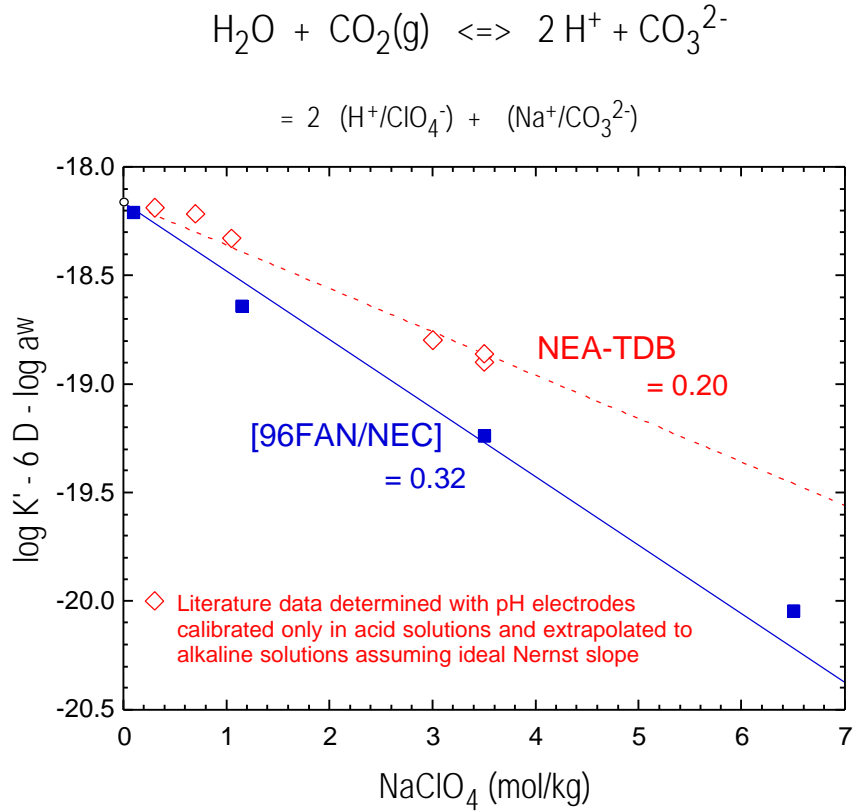


Fig. 2.6

Another possibility to overcome the problem of different SIT coefficients in different media would be to calculate only values for the reactions, which are then different for NaClO₄ and NaCl solution: ()_{NaClO₄}. ()_{NaCl}.

2.4. Consequences for NEA-TDB reviews

- (1) Limitations of the SIT have to be pointed out.
- (2) The SIT has to be extended for anion-anion interactions (c.f. proposal), at least in cases where this is necessary to avoid erroneous conclusions
- (3) a) New (correct) trace activity coefficients for CO_3^{2-} in NaClO_4 should be used.
 => All coefficients of actinide carbonate complexes, derived from exp. data at high NaClO_4 concentrations should be reevaluated. The existing values are of course not affected

If possible, the experimental results on H_2CO_3 dissociation constants given in [96FAN/NEC] should be checked in an independent laboratory.

b) As long as there is no final decision on the problem in 3a), only unambiguous data (in NaCl solution or at low NaClO_4 concentration, where the NEA-TDB auxiliary data on the carbonate ion are free of any doubt) should be used to evaluate SIT coefficients and equilibrium constants at $I = 0$.

c) Similar problems as in the case of trace activity coefficients for CO_3^{2-} in NaClO_4 solution have to be expected for the SO_4^{2-} ion.

(4) Np/Pu review

a) Conclusions for solid Np(V) carbonates are incorrect. The selected $\log K_s^\circ$ values have to be changed.

b) \log° , and SIT parameters (and) have to be reevaluated, including data in NaCl solution for the equilibria $\text{NpO}_2^+ + n \text{CO}_3^{2-} \rightleftharpoons \text{NpO}_2(\text{CO}_3)_n^{1-2n}$ (or for the stepwise constants), because NaCl is an important medium with respect to natural aquatic systems. If necessary, the \log° values can be fixed from the corresponding extrapolation with data in NaClO_4 solution (c.f. Fig.2.7, next page). The known values of $(\text{NpO}_2^+/\text{Cl}^-)_{\text{NaCl}} = 0.09 \pm 0.02$ [95NEC/FAN] and $(\text{Na}^+/\text{CO}_3^{2-})_{\text{NaCl}} = -(0.08 \pm 0.03)$ can then be used to evaluate SIT coefficients $(\text{Na}^+/\text{NpO}_2(\text{CO}_3)_n^{1-2n})_{\text{NaCl}}$ for the Np(V) carbonate complexes in NaCl solution.

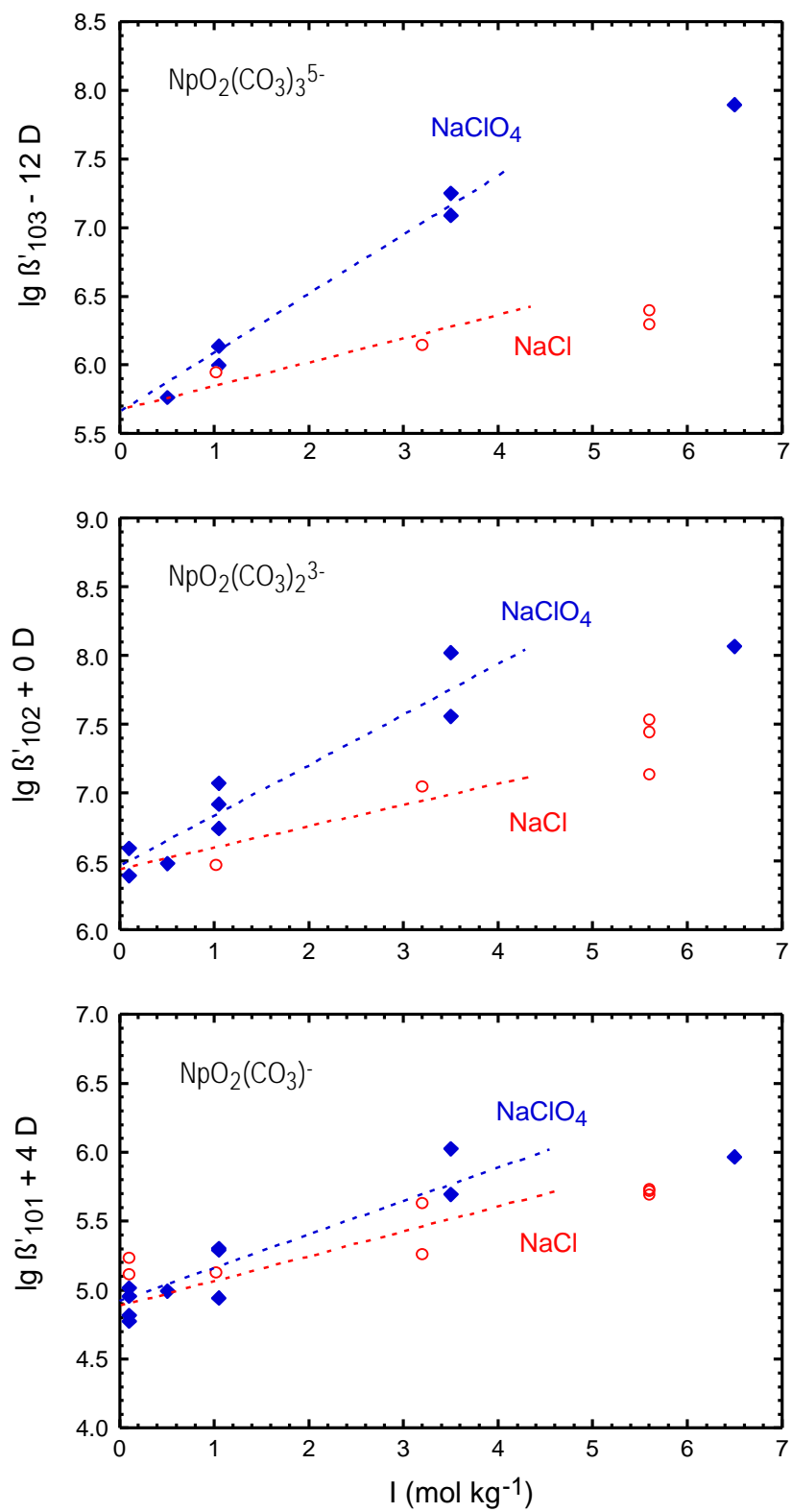


Fig. 2.7

3. References

[72/SER/NIK] cited in [92GRE/FUG]

[76NIK2] cited in [92GRE/FUG]

[77VOL/VIS]

Volkov, Yu.F., Visyashcheva, G.I., Kapshukov, I.I.;
Study of Carbonate Compounds of Pentavalent Actinides with Alkali Metal Cations.
V. Production and Identification of Hydrate Forms of Sodium Monocarbonato-
neptunate, Soviet Radiochem. **19**, 263 (1977)

[79VOL/VIS]

Volkov, Yu.F., Visyashcheva, G.I., Tomilin, S.V., Spiryaev, V.I., Kapshukov, I.I.,
Rykov, A.G.; Carbonate Compounds of Pentavalent Actinides with Alkali Metal
Cations, VII. Synthesis and Crystal Structure Hydrate Compounds with the
Composition $\text{Na}_{0.6}\text{NpO}_2(\text{CO}_3)_{0.8}$, Soviet Radiochem. **21**, 583 (1979)

[81VOL/VIS]

Volkov, Yu.F., Visyashcheva, G.I., Tomilin, S.V., Kapshukov, I.I., Rykov, A.G.;
Study of Carbonate Compounds of Pentavalent Actinides with Alkali Metal Cations.
VIII. Synthesis and X-Ray Diffraction Investigation of Several Compounds of
Neptunium(V) with Sodium and Rubidium, Soviet Radiochem. **23**, 191 (1981)

[81VOL/VIS2]

Volkov, Yu.F., Visyashcheva, G.I., Tomilin, S.V., Kapshukov, I.I., Rykov, A.G.;
Study of Carbonate Compounds of Pentavalent Actinides with Alkali Metal Cations.
IX. Determination of the Crystal Structure of $\text{M}_3\text{AnO}_2(\text{CO}_3)_2 \cdot n\text{H}_2\text{O}$,
Soviet Radiochem. **23**, 195 (1981)

[82WAG/EVA]

Wagmann, D.D., Evans, W.H., Parker, V.B., Schumm, R.H., Halow, I., Bailey, S.M.,
Churney, K.L., Nuttal, R.H.; The NBS tables of chemical thermodynamic properties,
J. Phys. Chem. Ref. Data **11**, Suppl.2 (1982)

[83MAY]

Maya, L.; Hydrolysis and Carbonate Complexation of Dioxoneptunium(V) in 1.0 M
 NaClO_4 at 25°C, Inorg. Chem. **22**, 2093 (1983)

[84GRE/FER]

Grenthe, I., Ferri, D., Salvatore, F., Riccio, G.;
Studies on metal carbonate equilibria. Part 10. A solubility study of the complex
formation in the uranium(VI) water-carbon dioxide system at 25°C;
J. Chem. Soc. Dalton Trans., 2439 (1984)

[84HAR/MØL]

Harvie, C.F., Møller, N., Weare, J.H., The prediction of mineral solubilities in natural
waters: The Na-K-Mg-Ca-H-Cl-SO₄-OH-HCO₃-CO₂-H₂O System to high ionic
strengths at 25°C, Geochim. Cosmochim. Acta **48**, 723 (1984)

[86GRE/ROB]

Grenthe, I., Robouch, P., Vitorge, P.; Chemical Equilibria in Actinide Carbonate Systems, *J. Less-Common Met.* **122**, 225 (1986)

[89ROB]

Robouch, P., Contribution à la prevision du comportement de l'américium, du plutonium et du neptunium dans la géosphère; données chimiques, Thesis, Strasbourg (1987),
Report CEA-R-5473, Commissariat à l'Energie Atomique, Gif-sur-Yvette, France (1989)

[90RIG]

Riglet, C.; Chimie du Neptunium et autres Actinides en milieu carbonate,
Rapport CEA-R-5535, CEN Fontenay aux Roses, France, 1990

[91FEL/RAI]

Felmy, A.R., Rai, D., Mason, M.J.;
The Solubility of Hydrated Thorium(IV) Oxide in Chloride Media: Development of an
Aqueous Ion-Interaction Model; *Radiochim. Acta* **55**, 177 (1991)

[91KIM/KLE]

Kim, J.I., Klenze, R., Neck, V., Sekine, T., Kanellakopulos, B.;
Hydrolyse, Carbonat- und Humatkomplexierung von Np(V),
Report RCM 01091, Technische Universität München, 1991

[91MEI/KIM]

Meinrath, G., Kim, J.I., The Carbonate Complexation of the Am(III) Ion,
Radiochim. Acta **52/53**, 29 (1991)

[91MEI/KIM2]

Meinrath, G., Kim, J.I., Solubility products of different Am(III) and Nd(III) carbonates,
Eur. J. Solid State Inorg. Chem. **28**, 383 (1991)

[91PIT]

Pitzer, K. S., Activity Coefficients in Electrolyte Solutions, Chap.3,
CRC Press, Boca Raton, Florida, 1991

[92GRE/FUG]

Grenthe, I., Fuger, J., Konings, R.J.M., Lemire, R.J., Muller, A.B., Nguyen-Trung, C.
Wanner, H. (OECD, NEA-TDB), Chemical Thermodynamics Vol. 1. Chemical
Thermodynamics of Uranium, Elsevier Science Publ., North-Holland, Amsterdam (1992)

[92KRA/BIS]

Kramer-Schnabel, U., Bischoff, H., Xi, R.H., Marx, G.;
Solubility Products and Complex Formation Equilibria in the Systems Uranyl Hydroxide
and Uranyl Carbonate at 25°C and I = 0.1 M; *Radiochim. Acta* **56**, 183 (1992)

[92NEC/KIM]

Neck, V., Kim, J.I., Kanellakopulos, B.; Solubility and Hydrolysis Behaviour of
Neptunium(V), *Radiochim. Acta* **56**, 25 (1992)

[92RUN/MEI]

Runde, W., Meinrath, G., Kim, J.I., A Study of Solid-Liquid Phase Equilibria of Trivalent
Lanthanide and Actinide Ions in Carbonate Systems, *Radiochim. Acta* **58/59**, 93 (1992)

[93LEM/BOY]

Lemire, R.J., Boyer, G.D., Campbell, A.B.;
The Solubilities of Sodium and Potassium Dioxoneptunium(V) Carbonate Hydrates

at 30, 50 and 75°C; *Radiochim. Acta* **61**, 57 (1993)

[93MEI/KIM]

Meinrath, G., Kimura, T.; Behaviour of U(VI) solids under conditions of natural aquatic systems; *Inorg. Chim. Acta* **204**, 79 (1993)

[93MEI/KIM2]

Meinrath, G., Kimura, T.; Carbonate complexation of the uranyl(VI) ion; *J. Alloys Comp.* **202**, 89 (1993)

[93PAS/RUN]

Pashalidis, I., Runde, W., Kim, J.I.;
A Study of Solid-Liquid Phase Equilibria of Pu(VI) and U(VI) in Aqueous Carbonate Systems; *Radiochim. Acta* **61**, 141 (1993)

[94FAN/KIM]

Fanghänel, Th., Kim, J.I., Paviet, P., Klenze, R., Hauser, W.,
Thermodynamics of Radioactive Trace Elements in Concentrated Electrolyte Solutions: Hydrolysis of Cm^{3+} in NaCl-Solutions, *Radiochim. Acta* **66/67**, 81 (1994)

[94MEI]

Meinrath, G.; Np(V) Carbonates in Solid State and Aqueous Solution,
J. Radioanal. Nucl. Chem. Letters **186**, 257 (1994)

[94NEC/RUN]

Neck, V., Runde, W., Kim, J.I., Kanellakopulos, B.;
Solid-Liquid Equilibrium Reactions of Neptunium(V) in Carbonate Solution at Different Ionic Strength, *Radiochimica Acta* **65**, 29 (1994)

[94NEC/KIM]

Neck, V., Kim, J.I., Kanellakopulos, B.;
Thermodynamisches Verhalten von Neptunium(V) in konzentrierten NaCl- und NaClO_4 -Lösungen; Report KfK 5301 (1994)

[94RUN/KIM]

Runde, W., Kim, J.I.; Report RCM 01094, Technische Universität München, 1994
Chemisches Verhalten von drei- und fünfwertigem Americium in salinen NaCl-Lösungen, Runde, W.; Thesis, Technische Universität München, 1993

[95FAN/NEC]

Fanghänel, Th., Neck, V., Kim, J.I.; Thermodynamics of Neptunium(V) in Concentrated Salt Solutions. II. Ion Interaction (Pitzer) Parameters for Np(V) Hydrolysis Species and Carbonate Complexes, *Radiochim. Acta* **69**, 169 (1995)

[95NEC/FAN]

Neck, V., Fanghänel, Th., Rudolph, G., Kim, J.I.;
Thermodynamics of Neptunium(V) in Concentrated Salt Solutions: Chloride Complexation and Ion Interaction (Pitzer) Parameters for the NpO_2^+ Ion, *Radiochim. Acta* **69**, 39 (1995)

[95NEC/FAN2]

Neck, V., Fanghänel, Th., Kim, J.I.;
Dissoziationskonstanten von H_2O und H_2CO_3 in NaClO_4 -Lösung und Pitzer-Parameter im System $\text{Na}^+/\text{H}^+/\text{OH}^-/\text{HCO}_3^-/\text{CO}_3^{2-}/\text{ClO}_4^-/\text{H}_2\text{O}$, 25°C, Report FZKA 5599, Forschungszentrum Karlsruhe, 1995

[95NEC/RUN]

Neck, V., Runde, W., Kim, J.I.;
Solid-liquid equilibria of neptunium(V) in carbonate solutions of different ionic strengths: II. Stability of the solid phases, *J. Alloys and Compounds* **225**, 295 (1995)

[95SIL/BID]

Silva, R.J., Bidoglio, G., Rand, M.H., Robouch, P., Wanner, H., Puigdomenech, I.. (OECD, NEA-TDB), *Chemical Thermodynamics Vol. 2. Chemical Thermodynamics of Americium*; Elsevier Science Publ., North-Holland, Amsterdam (1995)

[96FAN/NEC]

Fanghänel, Th., Neck, V., Kim, J.I.;
The Ion Product of H_2O , Dissociation Constants of H_2CO_3 and Pitzer parameters in the system $\text{Na}^+/\text{H}^+/\text{OH}^-/\text{HCO}_3^-/\text{CO}_3^{2-}/\text{ClO}_4^-/\text{H}_2\text{O}$ at 25°C , *J. Solution Chem.* **25**, 327 (1996)

[96KAT/KIM]

Kato, Y., Kimura, T., Yoshida, Z., Nitani, N.;
Solid-Liquid Phase Equilibria of Np(VI) and U(VI) under Controlled CO_2 Partial Pressures; *Radiochim. Acta* **74**, 21 (1996)

[96MEI/KAT]

Meinrath, G., Kato, Y., Kimura, T., Yoshida, Z.;
Solid-Aqueous Phase Equilibria of Uranium(VI) under Ambient Conditions; *Radiochim. Acta* **75**, 159 (1996)

[96MEI/KLE]

Meinrath, G., Klenze, R., Kim, J.I.;
Direct Spectroscopic Speciation of Uranium(VI) in Carbonate Solutions; *Radiochim. Acta* **74**, 81 (1996)

[96RUN/NEU]

Runde, W., Neu, M.P., Clark, D.L.;
Neptunium(V) hydrolysis and carbonate complexation: Experimental and predicted neptunyl solubility in concentrated NaCl using the Pitzer approach; *Geochim. Cosmochim. Acta* **60**, 2065 (1996)

[97RAI/FEL]

Rai, D., Felmy, A.R., Sterner, S.M., Moore, D.A., Mason, M.J., Novak, C.F.;
The Solubility of Th(IV) and U(IV) Hydrated Oxides in Concentrated NaCl and MgCl_2 Solutions; *Radiochim. Acta* **79**, 239 (1997)

4. Appendix

(Only available as hardcopy)

Appendix 1: Schematic illustration of the titration vessels used in the solubility experiments of Neck and Runde

Appendix 2: X-ray diffraction pattern from [94RUN/KIM]

Appendix 3: Experimental solubility data in 0.1, 1 and 3 M NaClO₄ tabulated in the sequence of the measurements (Tables from [91KIM/KLE])

Appendix 4: LSC spectra of Np-237 and daughter nuclide Pa-233

I have made remarks page by page on your last 44 page comments. It appears the main problem was wording in our draft. Your last comments do not really bring new qualitative or important information, but rather more details on what you previously wrote, hence the needed changes had been performed in the draft. As already said you proposed several interpretations that could be interesting; but for which more experimental confirmation are needed, hence our review did not rely on them. This mainly resulted in non considering your activity coefficients, which is quite a minor problem since for the system under discussion (Np(V) carbonate solubility product) the main goal of our review is to provide stoichiometries and equilibrium constants. The main subjects I commented are the following:

1. Despite what you wrote and repeated, we did not disregard your work. We even used data from your laboratory to select equilibrium constants and solubility product for all the components of the Np(V) carbonate system (this was already done before you started this polemic) with the same weight as similar works from other laboratories.
2. We originally found inconsistency between the values of NpO_2^+ activity coefficients extracted from your work using two different techniques (liquid-liquid extraction and solubility), hence on experimental data at high I. In the course of this discussion it appears that you propose to resolve this self inconsistency by adding a new term in the SIT formula used to calculate both CO_3^{2-} and NpO_2^+ activity coefficients. This results in a consistent interpretation of your data; but the values of NpO_2^+ activity coefficients extracted from your liquid-liquid extraction study are still poorly consistent with the corresponding value selected from the beginning of this series of reviews, a selection we confirmed. Anyhow I agree with you when you wrote (page 21 the paragraph before the last one) the main arises from experimental data and not in the choice of Pitzer versus SIT formula: it is first needed to confirm experimentally this problem with CO_3^{2-} activity coefficient: (i.e different value for pair parameter measured in NaCl and NaClO_4 media, while it should only depend on Na^+ concentration according to SIT hypothesis used in TDB review) at least because your interpretation is that most of the existing published data on this system in NaClO_4 aqueous solution would have systematic error, while your measurements would not. In other word the reverse explanation cannot be ruled out (see my comment page by page on your interpretation). It was actually the position of our review from the beginning (i.e before your starting this polemic): we recognised that auxiliary data used in our review (namely to calculate CO_3^{2-} activity coefficient) might need revision; but more experimental evidence is still needed. NpO_2^+ activity coefficients should not need revision at the moment (despite what you proposed which would) because emf measurements (on which the present value relies) are more reliable than solubility and liquid-liquid extraction techniques, and anyhow the new number you proposed is rather to fit inconsistency with one (yours) liquid-liquid extraction set of measurements. Anyhow these systematic deviations has no influence on the standard (i.e at $I=0$) values selected in our review, because we used a selection procedure that minimised this type of possible systematic error (whatever it is originated in your pH calibration or in TDB auxiliary value).
3. You also claimed that most published data in NaClO_4 media used wrong pH calibration (non nernstian slope of the glass electrode), you must be very sure of your own data to write this. You claimed this induced considerable difference in measured pH and hence in solubility product determinations that used this pH measurements. You said this difference could be of up to 0.2 unit \log_{10} which is not so big and is in contradiction with what you wrote in a previous comments where you said that 0.2 unit \log_{10} difference is not important in two of your determination of the first Np(V) carbonate complex constant using two different techniques (solubility and spectrophotometry: actually we agreed with your opinion, and we had written inconsistency was within your stated uncertainty that was certainly overoptimistic). So your supposed systematic error is within usual uncertainty. Actually it is even less for the solubility product under discussion: (following your reasoning) this 0.2 unit \log_{10} is the maximum possible value at high pH, hence in chemical condition where Np(V) carbonate are formed, while the solubility product under discussion was fitted on data in more acidic (non complexing) media where the possible systematic deviation should be less. Typically in my own experimental determination from the number you provided (for the worst non nernstian glass electrode) I estimated the corresponding possible systematic deviation should be of 0.06 which is within uncertainty (anyhow slope was independently checked in these measurements, and titration data were provided which allow a supplementary and they indeed confirmed the above estimation).
4. For the solid phase under discussion, we found experimental evidence of ageing, but this information is not much reported in your publications because, as you wrote pages 6 you usually did not provide transient experimental measurement in the course of solubility measurement: you only reported the final, asymptotically reached equilibrium value (page 8) at equilibrium to deduce values for the equilibrium constants (we suspected it from the beginning). This is perfectly correct; but as a consequence we did not rely on your work to discuss solid phase ripening. You did not much discuss the published experimental evidence of solid phase transformation, nevertheless you claimed there is no problem with ageing of the solid. Statistical analysis of the standard values for those of published solubility products you provided, rather confirm our interpretation. You also calculated that it was possible to fit these published solubility products, this is neither completely convincing (for the above statistical reason and) because you added empirical terms and fitted parameters for activity coefficients in perchlorate media which are neither completely validated (above comment 2). In other words you might have seen evidence of interesting new phenomenon (typically anion-anion and triple interaction contribution to activity coefficient); but it cannot be ruled out that you simply added a new empirical term or a fitted parameter to compensate possible systematic error.

5. Instead of discussing published information on scattering of solubility data during the initial precipitation (in term of ageing: above comment 4) you claimed your solubility data are the most accurate among all the published ones, while (as recalled just here above) you wrote you eliminated from your reported data, the intermediary scattered ones. You used this artefact to claim (page 7) my results were not accurate. You also used similar trick to draw a figure in the same goal. No comment.
6. You proposed to add to the SIT formula, the above supplementary terms. This should be indeed useful if, as already said, your experimental observations are confirmed. To keep consistency with TDB, pair empirical parameters should keep the same values as in existing data set when possible, hence at least for a reference electrolyte. You typically used these extra terms in Pitzer formula for the $\text{NpO}_2(\text{CO}_3)_2^{3-}$ and $\text{NpO}_2(\text{CO}_3)_3^{5-}$ complexes while there is experimental evidence they are the same in chloride and perchlorate media, hence it should be possible to set them to zero in another most consistent (with existing SIT parameters) fitting exercise. Nevertheless, your treatment is mathematically perfectly correct. This does not seem possible for $\text{NpO}_2\text{CO}_3^-$ which can either be evidence of possible formation of a weak mixed complex, typically $\text{NpO}_2\text{ClCO}_3(\text{aq})$.

I really considered all your arguments. Actually I completely rejected quite few of them: In most case I only asked for more experimental confirmations for the reasons explained or suggested in my page by page comments. I believe the (scientific) discussion is quite over:

1. Comments on the experimental studies of Neck and Runde

In the draft of the NEA review on Np/Pu, the results of Kim and coworkers on the solid and aqueous Np(V) carbonates [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC, 96FAN/NEC], were generally criticized to be not reliable.

PV 2a

Not generally, only on a specific minor problem in activity coefficient, generally our review found these works reliable enough to rely on in the selection of the standard complexation constants and solubility product. Only ionic strength corrections were not used. They were not used only on the following minor problem in activity coefficient calculations : possible inconsistency was found between the values of the interactions parameters $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-)$ calculated from your experimental studies using either solubility or liquid-liquid extraction techniques respectively, and anyhow inconsistency with TDB value. The origin of this inconsistency is still not resolved.

This comment is certainly rather due to an impression, than to what was actually written in an anyhow draft version.

the reviewer takes each opportunity to repeat the following reasons for disregarding the results of Kim and coworkers,

PV 2b

Yes, this was exaggerated and is now changed (see PV 2a concerning rewording)

(a) chemical problems with the solid phase (insufficient knowledge on the equilibrium solid phase)

PV 2c

Problems in solid phase is discussed or suspected for all the published studies and supported by experimental results: x-ray diffraction patterns (see the figure) and kinetic observations. Your study is not used to discuss these problems, just because you only paid attention to obtain a stable reproducible solid phase, and not to discuss its ripening and evolution, hence you provided fewer information than in other studies : nothing is wrong with this. I still do not know whether you performed x-ray diffraction before solubility equilibration, or after (from your publications it rather seems it was characterisation of the phase initially used for solubility measurements).

This comment is certainly rather due to an impression, than to what was actually written in an anyhow draft version : we did not criticised your solid phase.

(b) systematic errors in pH calibration,

PV 2d

Not exactly, possible systematic error

This comment is certainly rather due to an impression, than to what was actually written in an anyhow draft version, and anyhow rewording of the draft have now been performed to change this impression.

(c) calculation of $\log [\text{CO}_3^{2-}]$ from measured $\log [\text{H}^+]$ with auxiliary data, which are inconsistent with the NEA-TDB

PV 2e

Where is the problem in this statement? You also wrote it

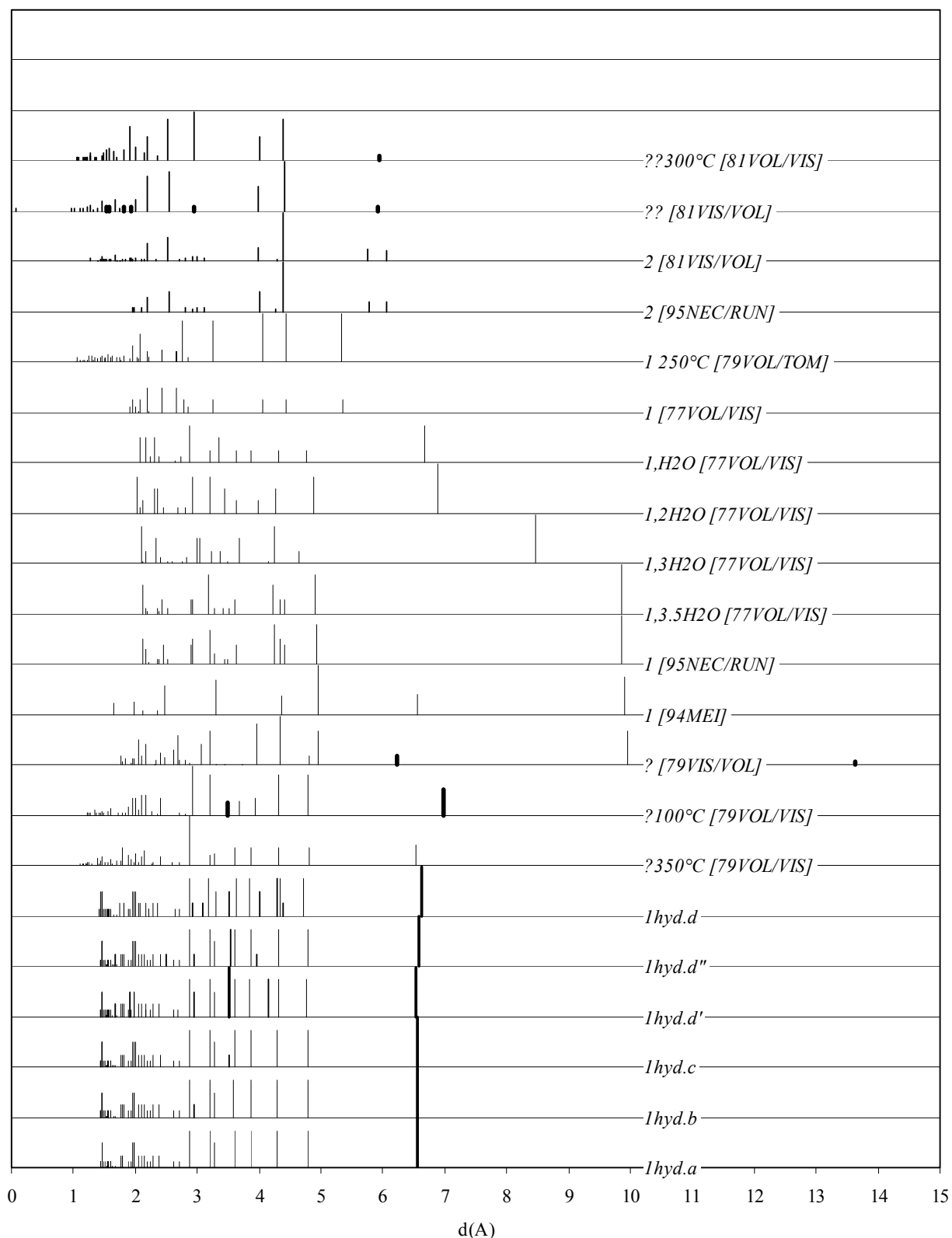
This comment is certainly rather due to an impression, than to what was actually written in an anyhow draft version, and anyhow rewording of the draft have now been performed to change this impression (see also PV 2a).

Statement (a) is absolutely incorrect. In contrast to the reviewer's study (c.f. discussion of [90RIG] in the Np/Pu draft) we had sufficient accurate experimental data and unambiguous experimental proofs for the solubility limiting solid phases (see section 1.1).

PV 2f

The author of ref.[90RIG] is Riglet (Martial-Riglet) who is not a reviewer. She did her work in Vitorge's laboratory, who is a reviewer. The solubility work already appeared in ref. [86GRE/ROB] and is better documented in typically references [84VIT, 98VIT/CAP]. In references [84VIT, 86GRE/ROB, 98VIT/CAP] all measurements are reported including those which were a priori disregarded for chemical reasons (pH not stabilised, too short equilibration time, X-

ray analysis showing extra lines) prior to the treatment of the data, while they were not in Kim's et al. publications. Both presentation are correct; this was already said in our review, it has already been recognised in the course of the present discussion (PV 2f 5, 6e, 6c and 8a).



This comment is one of the aspect already discussed (see PV 2c): there are less experimental information and discussion on solid phase characterisation and evolution in your published solubility works, than in the studies used in our review to discuss solid phase formation.

Statement (c) holds only for a part of our studies: for those in 1, 3 and 5 M NaClO₄,

PV 2g

Yes, there is an increasing systematic deviation with ionic strength, which is certainly within uncertainty for data at 0.1 M, (you did not published data in the range 0.1-1M, nevertheless, the systematic deviation should not be neglected in this range ($I=0.1-1M$)).

where the auxiliary data of the NEA-TDB are incorrect as will be shown in the present manuscript.

PV 2h

Yes, you pointed out, that you used data not much consistent with our auxiliary data, which resulted in different values for carbonic acid constants at high ionic strength, our review wrote (before receiving your comments) it might indicate our auxiliary data should be revised; but other possibilities are proposed to explain the deviation, typically possible liquid junction potential despite you measured it together with the influence of activity coefficients: it seems there are not enough redundant measurements for checking (making difference between liquid junction and activity coefficient effects), and anyhow, since this problem arose it might be better to check again the measurements using potentiometric cell of zero (or negligible) junction potential, and to compare with published data using such experimental set up. For these reasons, up to now our review did not adopt your values. Anyhow one cannot disregard, without further examination, the experimental values (some from famous old laboratories) used to established the TDB auxiliary values you criticise. The auxiliary values used in our review, are those adopted from the beginning (U review as already corrected in this particular case for $\epsilon(\text{Na}^+, \text{HCO}_3^-)$ and $\epsilon(\text{Na}^+, \text{CO}_3^{2-})$ as published in the Am book).

Even if they are finally revised (and again this might be again necessary because they do not rely on enough and consistent experimental determinations in my opinion), this would anyhow not be enough to resolve the inconsistency found in our review in some activity coefficients for Np(V) (PV 2a and e).

In the studies in 0.1 M NaClO_4 and 0.1, 1, 3 and 5 M NaCl , the H_2CO_3 dissociation constants determined and used to calculate $\log [\text{CO}_3^{2-}]$ from measured $\log [\text{H}^+]$ agree with well-known and generally accepted literature values and also with the auxiliary data of the NEA-TDB.

PV 2i

This means that published values for K in NaCl and NaClO_4 media are in poor agreement. In that case this should be first reviewed (again) and submitted to eventual future TDB specialists in charge of eventually changing these auxiliary values, possibly extra experimental measurement would be needed (see PV2 a, e and h).

Your reasoning seems to be: (i) $\epsilon(\text{Na}^+, \text{HCO}_3^-)$ and $\epsilon(\text{Na}^+, \text{CO}_3^{2-})$ values used in our review reproduce K experimental values in NaCl , but not in NaClO_4 media; (ii) pair interaction parameters are not enough to model at least one of the two media (since $\epsilon(\text{H}^+, \text{ClO}_4^-)$ and $\epsilon(\text{H}^+, \text{Cl}^-)$ values are reliable). This is a possible physical explanation, but it is not clear why these extra parameters would be needed specifically for this (or these) system(s), and not on many other. Anyhow this discussion is on (ii), while we must first agree on experimental evidence on (i).

It is now quite well established water dimer, trimer etc are formed in liquid water through (networks of) hydrogen bonds. Adding an ion locally modify it, and the first sphere hydration (water) molecules (of the added ion) can possibly be a starting points of co-operative hydrogen bonding, where (specially at high I) counterions could possibly participate to such network of (hydrogen bounds), this should result in ion pair which is usually taken into account at the macroscopic level by activity coefficients, specially in this particular case through second virial terms (ion pair empirical parameters). This could equally be modelled with equilibrium constant (hence for a weak complex), and Pitzer already pointed out this (and there is also a calculation in Riglet's thesis [90RIG] to link ion pair equilibrium constant with empirical pair interaction coefficient, i.e a tentative link between the two models). On the other hand several effect are certainly taken into account in the pair parameter, for this reason it is rather called empirical. Anyhow this gives a picture of the geometry on which are based the calculation of physics which gave SIT and Pitzer formula (the same type of calculations was used: resolving Boltzmann and Poisson equations, which gives the Debye-Hückel term, and using Taylor series expansion which give the virial terms). With this picture you can as well add a third cation in hydrogen bond cycles; but this addition should in principle modify the structure, or at least the energetics of each of its component, hence interaction terms for three ions should rather be used in place of (at least) some of the pair interaction contributions, and not added to it. In other words the competition between pairing equilibria is more appropriate than the ion multi (and eventually non interacting) layer picture that gives the virial expansion. On the other hand the enrgetics of this structure is of the order of magnitude of hydrogen bonding in water, hence very weak complex or activity coefficient, so this difference between the 2 approaches (competition versus addition of interaction) should not be detected experimentally. This means several sets of interaction parameters should in principle allow the description of the same system, and this indeed was published by Pitzer who typically showed correlation between the numerical values of its two empirical pair parameters. This finally should allow consistency between second toward third virial expansion: typically by fitting the pair parameters on single electrolyte (without any higher term, hence the pair parameter(s) does not only account for pair interactions) in a first step. In a second step you use experimental data involving exactly three ions, and you fit the three ion empirical parameters of higher terms, keeping the fixed values for the previous pair parameters, hence again the new fitted parameters have a mixed origin: they are not for three ions

interactions, but some of the previous two ion interactions are implicitly subtracted. The same procedure can be used also if you need to use empirical pair parameters for ions of the same charge: typically by choosing an arbitrary reference single electrolyte.

The previous discussion is only to say it is possible to add consistently new terms to the SIT formula. This terms are already added to the Pitzer formula, hence they are not automatically consistent with simplification of Pitzer formula into a second order (i.e SIT like) one (unless the above typical procedure have been used to obtain the Pitzer parameters, in other word unless a strong enough weight is given in the fit, to data were pair parameters are enough to model the data; and, when pair interaction parameters are used for ions of same charge, including data of the reference single electrolyte). We tested this (and the result was briefly reported in the Np/Pu draft) on your published Pitzer parameters: many of them are consistent with the SIT (i.e a third to second order simplification) while it seems some are not, specially for anionic Np(V) complexes. This does not mean your Pitzer parameters are incorrect: it means they are not consistent with the SIT formula. They could be at least for $\text{NpO}_2(\text{CO}_3)_i^{1-2i}$ for $i=2$ and 3 for which there is no need to use different anion-anion parameters in chloride and perchlorate media, hence they can be set to zero (in other words NaCl and NaClO₄ are what I called above a reference electrolyte for this system).

Nevertheless, the reviewer disregarded all these results.

PV 2j

Our review relied on your work in the selection of the standard complexation constants and solubility products for the $\text{Na}^+/\text{Np(V)}/\text{CO}_3^{2-}$ system (you are not discussing these numerical values). Only ionic strength corrections were not used (see PV 1a, b, g, h and i).

This comment is certainly rather due to an impression, than to what was actually written in an anyhow draft version, and anyhow rewording of the draft have now been performed to change this impression.

Careful reading of our original papers might be sufficient to agree with our arguments.

PV 2k

Do you suggest we did not read carefully your papers?

This comment is certainly rather due to an impression, than to what was actually written: when we do not accept all your conclusions it does not mean we did not carefully examined all your arguments (since you are now repeating almost the same arguments, we have spent much time examining the same original experimental information many times). Despite your repeating of the same arguments give the bad impression you cannot produce new ones, neither you can answer ours, I can insure you we are re-examining all your publications in view of your comments each time we receive them.

Several members of the NEA-TDB project groups required additional information and explanations, which hopefully will now be given in this manuscript.

PV 2l

What do you mean? Somebody is hiding information? Who? Several members are not able to read the information you are repeating?

In addition, it has to be stated that the primary intention of any scientific study should be to obtain correct results. The consistency with recommendations of other people or organisations is desirable, but only of secondary importance.

PV 2m

Yes, hence you should not be annoyed that for checking and consistency our review performed a few reinterpretation and recalculation on your papers, and disregard some of your proposed values (actually only for activity coefficients) for which too many assumptions should be made to obtain consistency, while other consistent data are available

Page 3

1.1. On the solubility studies with Np(V) carbonates

For the discussion in section 2 of this manuscript, it is important to know that the solubility data in [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC, 96RUN/NEU] refer to well-defined solid phases.

PV 3a

Yes, we did not write it was not, or it was worst then in any other published study on the same subject (see PV 2 c).

Experimental procedure

All solubility experiments were performed in titration cells (see Appendix 1)... The Np(V) solid was precipitated in the titration vessel, and left aging 1 - 2 weeks before the experiment was started.

PV 3b

Hence it seems, it is the same experimental set up as I previously used with the help of Ingmar Grenthe and Diego Ferri [84VIT, 98VIT/CAP], but with different CO₂ partial pressure and several ionic strengths: Initial equilibration time seems shorter here, than in some other studies (but see PV 3c).

The equilibration between CO₂(g), aqueous carbonate and solid Np(V) carbonate was monitored as a function of time by measuring the H⁺ and Np concentrations until these concentrations remained constant. This could last a few days up to 3 weeks, depending primarily on the time needed for the equilibrium between HCO₃⁻ and CO₃²⁻ in the aqueous phase and pCO₂ in the gas phase.

PV 3c

In my own experiments, we used 10 % to 100 % carbonic gas mixture, equilibration was usually achieved within 10 minutes, while rather a few days were needed to achieved solid liquid equilibration. I obtained the solid of correct stoichiometry within at least about two months in similar experimental set up [98VIT/CAP].

Maya [83MAY] reported in the experimental section of his paper:... Exactly the same observation was made in our studies and therefore the solubility experiments were started with a precipitate aged for 1 - 2 weeks.

PV 3d

Maya used batch experiments while you (and I) used open cells. I also used batch experiments and found equilibration was not yet achieved within 2 weeks (I used 4 and more weeks).

Solubility limiting solid phases

1) X-ray powder diffraction

The X-ray powder diffraction pattern reported in [94NEC/RUN, 94RUN/KIM, 95NEC/RUN] were taken from a part of the solids used in our solubility studies.

PV 3e

After or before solid-liquid equilibration? Mine were from the filter used for sampling solubility measurements, hence usually after solid-liquid equilibration (see PV2 c and the corresponding figure).

In contrast to the results given in [94NEC/RUN, 94RUN/KIM, 95NEC/RUN], Meinrath [94MEI] reported... a different x-ray pattern. However, he obtained comparable solubility data.

PV 3,4

This rather confirms what is written in our review: very similar (but not identical) diffraction patterns were obtained for compound apparently prepared in similar ways, and their solubilities (hence thermodynamic stabilities) were found to be quite similar. Nevertheless according to the solubility product values you extrapolated to I=0 (at the end of your report under discussion), Meinrath's data can be considered as outlier (2σ threshold), which does not automatically mean his results should be rejected, but that he possibly had a different solid phase (which is consistent with x-ray diffraction pattern as you noticed). In principle it should be a fresh solid because it is more soluble (hence less stable, less aged) than others, however the difference is quite small and can partially be also attributed to systematic deviation between different laboratories. Let us assume he really used a fresh compound, its x-ray diffraction pattern (PV 2c and the corresponding figure) can be attributed to a mixture of the solid initially prepared by Maya (and probably you), and mine: this can be consistent with ripening of the solid phase as discussed in our review (anyhow we did not rely on this observation in our review to draw any conclusion).

Finally your remark can induce a discussion that ends up with supporting our decision to select 2 solubility products, decision you are criticising below (without any more referring to this problem).

(Page 4)

2) Slope analysis

As shown in [94NEC/RUN], the slopes of the solubility curves (log[Np] vs. log[CO₃²⁻]) are exactly -1

PV 4a

Certainly, not exactly: all experimental determinations have an associated uncertainty. Anyhow this confirms what is written in our review: smaller slope is not obtained when the solid phase is equilibrated a long enough time with the aqueous solution.

...The presence of solids like $\text{Na}_{0.6}\text{NpO}_2(\text{CO}_3)_{0.8}(\text{s})$ or $\text{Na}_{0.72}\text{NpO}_2(\text{CO}_3)_{0.86}(\text{s})$, as discussed by Vitorge in the NEA review, (p.284 lines 26ff and in Appendix A, discussion of [84VIT, 90RIG]), can certainly be excluded in all our studies. May be such solids are formed as intermediates or as Np(V)-hydroxide-carbonate solid mixtures in the early state of precipitation (where the precipitate is not fine crystalline but hydroxide-like gelatinous) but they are certainly instable

PV 4b

Yes, we never wrote neither thought it was a problem to extract equilibrium constants from your study. Evidence of compound with lower stoichiometry was indeed observed as intermediate in our study, while you did not report this information from yours hence both observations are correct and consistent. This might indicate that solid phases prepared by Volkov et al. have not been equilibrated a long time with aqueous solutions.

3)...After finishing the experiments at $I = 0.1$ M, the solids in the two vessels (meanwhile about half a year old) were further used for the solubility experiments in 1 and 3 M NaClO_4 . Hence, the solubility data at $I = 0.1$, 1 and 3 M refer to the same solid phase.

PV 4c

Have you got the x ray characterisation? Over such a long period of time ripening of the solid phase cannot completely be excluded. If the time of equilibration is correlated with the ionic strength (0.1, 1 then 3 M) this could, in principle, induce systematic error on activity corrections from the solubility product.

And since the solubility data at $I = 1$ and $I = 3$ are practically the same as those obtained by Maya [83MAY]

PV 4d

Yes (despite different equilibration time and auxiliary data were used)

and Vitorge [86GRE/ROB]

PV 4e

Not exactly.

in the corresponding media (c.f. Fig. 1.1),

PV 4f

See also similar figures reported in [98VIT/CAP].

it is evident that the solubility data of these authors refer as well to the same solid, and not to different (more or less aged or hydrated) solids as concluded in the NEA review,

PV 4g

Your solid phase is identical to the one reported by Maya, or by Vitorge? In principle it cannot be to both of them because they do not refer to the same x ray diffraction pattern. On the other hand Maya did not published its diffraction pattern.

Anyhow our review did not rely on such discussion for selection of thermodynamic data because (i) enough solid phase characterisation information was available only from my own work (we did not, and still do not know when your solid phase was characterised and how often) and (ii) anyhow x ray gives an information on the structure of the bulk solid phase which might be different from the surface or any other undetected phase controlling the solubility. Discussion in our review is also based on kinetic observations.

and not to different (more or less aged or hydrated) solids as concluded in the NEA review.

PV 4h

Not exactly, another conclusion can be deduced from your above remark and other observations (PV 3-4 4g and 5). Anyhow your affirmation is not a conclusion but a possible explanation for the (finally rather limited) differences in solubility product values between different laboratories. This type of difference is quite usual. Anyhow our review did not rely on such discussion for selection of thermodynamic data.

Page 5

Fig1.1

PV 5

As in figures given in [98CAP/VIT], it is better to use different colours to point out data a priori excluded (see PV 2f, 5, 6c, 6e, and 8a) for extraction of equilibrium constants (figure at I=3M), and to plot all the experimental data. Otherwise, one might imagine you wanted to point out with this figure one set of data is more scattered than the other one, while in addition (to the previous remark) in your set you said you have excluded scattered data (which is of course correct, to extract equilibrium constants). Honestly did you mean this?

The graphical presentation of the figure at I=1M is neither satisfactory because it seems Maya's data are on a single curve, while he showed a second series of slightly shifted for data after bubbling CO₂ in batches of the first series. He certainly attributed this to CO₂ because he also fitted an Np(V) hydrolysis constant of value not much consistent with later works. Hence the shift can also be attributed to ripening of the solid (the solid in the second series have been equilibrated twice). See PV 6d

Page 6

In Vitorge's review

PV 6a

I will have problem with the co-authors!

not a single sentence was written on... solid phase characterizations... Quite in contrast, he states and repeats several times that Kim and coworkers might have had chemical problems with the solid phase.

PV 6b

Our review did not criticise your solid phase (see PV 2c, 3a, e, 3-4 and 5), and relied on your work in the selection of the standard complexation constants and solubility products for the Na⁺/Np(V)/CO₃²⁻ system. Only ionic strength corrections were not used (see PV 2a, b, g, h, i, j and 4 g)

This comment is certainly rather due to an impression, than to what was actually written in an anyhow draft version, and anyhow rewording of the draft have now been performed to change this impression..

Accuracy of solubility measurements

In Vitorge's review and also in the comments of Robert Lemire (in his reply to our letter), there were doubts on the solubility data given in our papers [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN], because they are less scattered than expected.

PV 6c

This is only a matter of wording, this has been changed. We mainly wanted to say that all your published solubility data were at equilibrium (which is perfectly correct): you did not provide much information on the transient measurements, hence we cannot much discuss ripening of the solid phase from your publications. In the ongoing discussion you confirmed this, nevertheless you also claimed you did not had such problems with your solid phase: again it is a matter of wording (what do we call a problem? Hence I am not using this word in this sentence), in the ongoing discussion you gave extra information on the time needed to reach solubility equilibrium, this time is possibly quite consistent with previous experimental observations, hence with ripening of the solid phase (see also PV 2f 5, 6e, and 8a)

...other authors (e.g. [83MAY, 94MEI]), who also reported Np(V) carbonate solubility data of comparable high accuracy.

PV 6d

Data reported in ref. [83MAY] are certainly less accurate than yours: Maya measured twice the solubility in some batches after a new equilibration time, he found slightly different solubility, and attributed this small difference to the effect of adding acid (CO₂ carbonic gas bubbled a short time in the batch) in the batch before the second equilibration period. I calculated from the information given in ref. [83MAY] and latter publish work on Np(V) hydrolysis, that this interpretation is certainly not correct, hence the small difference in solubility measured in Maya's publication is probably due to ripening of the solid phase, and this is consistent with the (relatively short) equilibration time he used (see PV 5).

Accuracy of Meinrath's data is less clear, because there is not enough information to know whether and how he reproduced measurements.

Non-equilibrium data are useless for the determination of thermodynamic quantities.

PV 6e

Yes; but they are of course useful to obtain kinetics information (typically ripening of the solid phase). Anyhow we never thought or wrote there was an equilibrium problem with your solid phase, and we used your data to select equilibrium constants.

Page 7

insufficient analytical facilities might very well have lead to inaccurate data or data scattering. It is a more probable reason for the data scattering in the solubility experiments of Vitorge [86GRE/ROB] or Lemire et al.[93LEM/BOY]

PV 7a

This assumption is in contradiction with experimental results: they are not particularly more scattered at low solubility (near the detection limit of ^{237}Np). In data from ref. [86GRE/ROB], see also more information on point in ref. [84VIT] and [98VIT/CAP], the scattering of data is clearly correlated with chemical information: equilibration time, quicker for dissolution than precipitation, much slower for the initial precipitation.

than the presence of different solid phases (which is the explanation given by the reviewer).

PV 7b

This explanation is supported by experimental data: kinetics (see PV 3a, 3c, 3-4, 4h, 5 and 7a) and X ray diffraction analysis, while your explanation is only an assumption.

Lemire et al. [93LEM/BOY] mainly applied α -spectrometry... It is not surprising, if this analytical procedure leads to scattered data, in particular for Np concentrations $< 10^{-5}$ M, close to or at the detection limit of α - and γ -spectrometry.

PV 7c

I do not think so, I currently used α -spectrometry to measure $^{237}\text{Np(V)}$ concentration with a detection limit down to about 10^{-6} M, with no particular problem, with the type of procedure you described and 2 other ones. This is not particularly difficult: published works show detection limits down to less than 10^{-8} M. I never thought this could be a problem in Lemire's work, you must be very sure of what you are saying to put it in the forum discussion, and anyhow you could ask directly to the author on this specific points before initiating a public polemic relying only on assumptions. All along the review I never had such problems with Lemire: he never tried to extract more from his data, than should be, he rather had the reverse attitude.

The accurate determination of the Np-237 concentration requires an additional α/β discrimination for the counts from Pa-233... If the LSC measurements are not corrected by discriminating the β -radiation from Pa-233,

PV 7d

Which is done by commercial apparatus.

The analytical method used by Vitorge et al. is neither mentioned in the paper [86GRE/ROB] nor in Riglet's thesis [90RIG]. May be, at the time Vitorge performed the solubility experiments shown in Fig.1.1 (before 1984), he did not have the analytical facilities necessary to record LSC spectra for α/β discrimination.

PV 7e

Again "maybe". Again your assumption is incorrect. I had the necessary experimental set up. The α/β discrimination (by electronic discrimination of the form of the electric pulse, not by software which is usually not reliable) is very well known for a very long time (much before software and micro-computer existed) in CEA. Anyhow I did not use liquid scintillation but γ spectrometry at 29 keV (Pa has a pic at 83 keV). Anyhow Kim have all this information which is in CCE reports, because he was responsible of the contract, I do not know whether he showed you my semi annual reports, but I assume he read it in details because he was responsible of the contract and possibly because the same procedure as mines were reproduce in his lab. (the procedure is from Ingmar Grenthe for the cell experiments).

Procedure to ascertain the equilibrium state

Page 8

each of the given solubility data represents the final, asymptotically reached equilibrium value of a series of H^+ and Np measurements as a function of time! This explains why our data show such a small scattering.

PV 8a

This is perfectly correct, and confirms you did not provide the transient information as I wrote above (PV 2f 5, 6c and 6e).

One can very well recognize, from jumps in the Np concentration, when the solid phase is aging

PV 8b

Solid phase transformation can be quite slow, specially with the time of at least several days (and actually several weeks) to transform the mono carbonato to the bicarbonato solid phase, I would not call this a jump. It is a matter of vocabulary.

Ripening of the initial phase after the first minutes, is quite slow, not by jumps.

Page 9

1.2. On the determination of the H^+ , OH^- and CO_3^{2-} concentrations

...

Calibration of pH electrodes and determination of the H^+ and OH^- concentrations

...

The term A represents a correction for the differences in liquid junction potentials (ΔpH) and the trace activity coefficient of H^+ .

PV 9a

It would have been better to have information on activity coefficients, independent from the liquid junction potential, for this you could have filled the reference compartment of your combined pH electrode with typically Na^+ 3M, Cl^- 0.01 M, ClO_4^- solution + AgCl solid. The slope is checked with a series of buffers at the same I (any I). 0.01 M H^+ , Na^+ , ClO_4^- 3M solution can typically be used for the calibration, additional checking with $CO_2(g)/HCO_3^-$ and HCO_3^-/CO_3^{2-} buffers.

Page 10

It is to note that Method B is applied by very many research groups... but also by Rai et al. [91FEL/RAI, 97RAI/FEL] and by other groups in the USA and in Japan.

PV 10a

In his last publication on Pu(IV) solubility in carbonate media, Rai wrote he considered his pH measurements at high I, not reliable, for this reason he used calculated pH. I calculated a shift of about 0.5 pH unit between calculated and measured pH.

Unfortunately, our earlier publications [92NEC/KIM, 91KIM/KLE, 94NEC/RUN] contain illustrations of solubilities as a function of pH (-log of the H^+ activity), Indeed this assumption is not consistent with the SIT or Pitzer splitting conventions, which might be confusing to the reader and to the NEA reviewers.

PV 10b

We used the numbers given in tables of your reports, and we fitted virtually the same equilibrium constants as the one you published. So it seemed we finally correctly understood your data.

Question: correct or incorrect values of $\log [H^+]$ and $\log [OH^-]$?

...2) Example: $\log[\text{OH}^-]$ in the studies on Np(V) hydrolysis If we use our calibrated pH electrode to measure a solution of 0.01 M NaOH / 2.99 M NaClO₄, then the final result must be: $\log[\text{OH}^-] = -2.00$

PV 10c

This example is not particularly well chosen because most glass electrodes should not be used (without further correction for alkaline error) at pH more than 10, and usually cannot be used at pH = 11 or 12 and above. at $\lg[\text{OH}^-] = -2$, potentiometric titration is certainly the best way to obtain speciation. Similarly in carbonate studies speciation is in principle known as accurately from the concentrations of the buffer components (excepted when radiolysis or side reactions modify pH) hence pH measurement can also be considered as checking.

This was also done in my Np(V) studies and it is in the reports Kim had (and the primary information is now in [98VIT/CAP]: the titration points ($-\lg[\text{H}^+]$ vs. added acid) are on the calculated line, hence error in calibration, if any, is quite small in the cell experiments (bubbling 100 % or 10% CO₂) this is consistent with your remark on non nernstian slope for the glass electrode: possible systematic deviation due to non nernstian slope is expected at higher pH.

Since (at the high CO₂ partial pressure I used) NpO_2^+ is predominating in neutral to acidic media, possible error in the slope of the electrode in my work (which is not expected in this pH range) should have negligible consequence on the determination of my solubility product. This is further confirmed since after 2 a month equilibration, I obtained a slope of -1.04 ± 0.07 (during the rest of the year) [98VIT/CAP], hence uncertainty on the pH measurement is within uncertainty of solubility data (I estimated independently reproducibility of my pH measurements to be of about ± 0.06). This is further confirmed by the numbers you gave below (PV 11b and 15 e) for typical non nernstian slope for the glass electrode.

Page 11

3) Comment on objections (Pierre Vitorge, Robert Lemire) concerning possible carbonate contamination in NaOH calibration solutions.

PV 11a

I do not remember we made such an objection: please give the exact citation and where you found it (it might be an error or unclear wording to correct in the Np/Pu draft)

Comments on other pH calibration procedures

....

As discussed in [96FAN/NEC], an incorrect electrode calibration procedure has a severe impact on the NEA-TDB auxiliary data concerning H₂CO₃ dissociation constants and carbonate trace activity coefficients in concentrated NaClO₄ solution.

PV 11b

Again assumptions: other laboratories (among them very old and well known ones in the field of solution chemistry) are supposed to have made an error in calibrating their glass electrode. This is a possibility, you need more cross verification (redundant measurements) and inter laboratory comparison to be more convincing (even if you are right).

However $-\lg[\text{H}^+]$ of CO₃²⁻/HCO₃⁻ buffers are in the range of about 9-11, hence possibly 7 log unit from calibration fixed point (actually I usually did not use this procedure, but, after verification of the slope, I rather checked reproducibility with buffers in the pH range of my measurements); while for HCO₃⁻/CO₂(g) buffers $-\lg[\text{H}^+]$ are in the range of about 6-7 hence closest to the possible calibration point. Hence, possible systematic error due to non nernstian slope on the corresponding data is less important, which finally probably had negligible influence on solubility product determination (see PV 10c).

You wrote the slope of the electrode can be of 58.0 to 58.8 mV/pH unit instead of 59.16 mV/pH unit, hence a shift of 1.16-0.4 mV/pH unit error for the slope of one of your electrode, hence 0.0196-0.0068 /pH unit in CO₃²⁻/HCO₃⁻ buffers the error should then be of

0.14-0.05 log unit,

while in HCO₃⁻/CO₂(g) buffers the error should then be of

0.10-0.03 log unit.

This confirms (PV 10c) that this error should be within uncertainty for experimental determinations in cell under CO₂(g) 100 % bubbling as I used to measure my solubility product (see PV10 c and 15 e)

Page 12

Vitorge et al. [86GRE/ROB, 90RIG] calibrate their electrodes in the acidic range with HClO₄/NaClO₄ solutions and in the neutral to alkaline range with carbonate buffers (acceting

the NEA-TDB auxiliary data for H_2CO_3 dissociation constants in 3 M NaClO_4). By this way, of course they observe an apparently ideal Nernst slope.

PV 12a

Again assumptions: ask me what I did before starting a polemic. The slope of the electrode used in my Np(V) solubility study, was also checked with independent buffers at low I. The difference with the theoretical slope was within the reproducibility (0.06 pH), which finally seems quite consistent with your above discussion on slope, and anyhow would induce a maximum possible systematic error within uncertainty for my experimental determination of the solubility product (PV 10c and 11b).

Vitorge's comment, that we should have checked our electrodes before use or that we should have used other types of electrodes, and his statement that our pH measurements include systematic errors, are completely inadequate.

PV 12b

This was only discussion in the case you wanted to change TDB auxiliary values: in that case redundant measurements, cells with no liquid junction and inter laboratory comparison would of course be required.

Page 13

H_2CO_3 dissociation constants in NaCl and NaClO_4 solution

Page 14

Page 15

all literature data shown ... were determined with glass electrodes calibrated only in the acidic range, and extrapolated to the alkaline range by assuming ideal Nernst slopes.

PV 15a

This is your basic assumption. you wrote a 44 pages document, where you basic assumption is in less than a sentence, within 2 lines. You could have given more details, typically as we do in the TDB review in Appendix A (typically give the slope you imagine the electrode had, and possibly try to find out whether it is realistic for material eventually used at that time). Your assumption is possible.

That means: all these data are incorrect. And even if so many authors obtain comparable results by making the same mistake, a mistake will always remain a mistake.

PV 15b

You must be very sure, to say this. More verification of your above assumption, is certainly required, before deducing such conclusion.

Repeating 4 times the word "mistake" give the impression of insisting, repeating your judgement, rather than be able to give clear enough argument, that the reader will deduce the conclusion by himself. This style is sometime needed for reports given to persons that take care of science but are not doing themselves science, it is not convincing for scientist of the field

This mistake

PV 15c

You decide it is a mistake, while it is only an assumption to explain inconsistent data between your laboratory and much published work (as you wrote), hence it is certainly more appropriate to write "possible systematic deviation" This wording also avoid moral judgement.

is illustrated... example is taken from our FZKA report

PV 15d

You are not showing what was done in other laboratories; but modelling what you assume was their systematic error. Hence this does not demonstrate they actually made or not this approximation; but only that your explanation is possible. You do not need to convince us of this, because before you comments, we had already written in the draft it might be possible auxiliary data selected in our review, need reevaluation. I also evaluated above (PV 11b) what would be the maximum possible systematic error according to your interpretation: this is rather more useful to discuss whether such possible deviation can account for observed discrepancy... and this depend on the pH range and the methodology used.

...two different ways of electrode calibration in 3 M NaClO₄, either the calibration with both, acidic HClO₄/NaClO₄ and alkaline NaOH/NaClO₄ solutions (solid line), or the calibration only in the acidic range extrapolated with the ideal Nernst slope (dashed line)... the value of log K₂ is determined in a solution with the composition:

0.05 M NaHCO₃ / 0.05 M Na₂CO₃ / 2.85 M NaClO₄

...the dashed line for electrode calibration, the value of log K₂ = -9.64

...consistent with the literature data and with the auxiliary value of -9.62 from the NEA-TDB.

- the solid line ...log K₂ = -9.81.

the errors in log K₂ are about 0.1 - 0.2 log units.

PV 15e

OK, this is interesting and confirm you have a possible explanation. Nevertheless try more confirmation: again find a way to have redundant measurements, typically use commercial buffer in the whole pH range between your calibration and working points (lg[H⁺] = -2? and -9.8). Do together the calibration in NaCl and NaClO₄ media (you possibly have already done this) in that way pollution of the NaOH is less probable, look for alkaline error (in that case the slope would not be constant: it decrease at high pH)...

You found a shift of

0.17 pH unit

which is consistent with twice the error I evaluated above (PV 11b) from your non nernstain slope information.

Similarly the deviation for lgK_{p1} with the same electrode should be of

0.06

which is of course consistent with what I wrote above (PV 10c and 11b): this is within uncertainty for my determination of the solubility product.

Page 18

Question: are the carbonate concentrations determined in experimental studies on actinide carbonates correct or incorrect ?

In the Np/Pu draft, the reviewer is irritated,

PV 18a

Our review wrote a remark, it is your interpretation to deduce a reviewer was irritated which is not supported by the way this remark was done and written. Anyhow within our review team, irritation was not an argument accepted to write a remark.

because the Np(V) carbonate solubilities of [91KIM/KLE, 94NEC/RUN] and those of [83MAY] are in good agreement, although considerably different constants

PV 18b

Different, not considerably different.

are used to calculate the carbonate concentration from the measured values of H⁺ concentration:

page 806, lines 23 - 27

The solubility values in 1 M NaClO₄ aqueous solutions were the same as those in [83MAY], and this may be coincidental as different values were used for the protonation constants for carbonate ion.

The agreement between the Np(V) carbonate solubilities in [94NEC/RUN] and those in [83MAY] (in 1 M NaClO₄) and [86GRE/ROB] (in 3 M NaClO₄)

PV 18c

There is very good agreement with those in ref. [83MAY]; but not with those in ref. [86GRE/ROB]. The error in pH calibration might cancel as you wrote below depending on the way Maya actually checked his electrode, beside this (from the information you gave in the course of the present discussion) its equilibration time might have been shorter than yours (see PV 3d, 3-4, 5, 6d, 18c and d).

is of course not accidently as supposed

PV 18d

Not exactly supposed which rather means we adopted this hypothesis in contradiction with what is written in our draft and you reproduced above (PV 18c).

by Vitorge in the NEA review.

PV 18e

I am not the only author of the NEA review. I have published some of the draft of my contributions, so you should rather refer to it, despite I might have changed my mind (typically after discussion within the review team) or do new calculations since.

This is demonstrated below

PV 18f

When you make assumptions (as in most parts of this document) what you can usually demonstrate is that it does not lead to any contradiction with available information, hence that it is a possible explanation; but you usually cannot validate it.

for a carbonate solution in 3 M NaClO₄....Vitorge et al. (ideal Nernst slope)

PV 18g

Again an assumption: I indeed checked the slope and found that in my working domain, it was the theoretical slope within uncertainty, but I rather stuck on the NEA value the closest to my working conditions. Anyhow the possible systematic error as estimated from your number should be within reproducibility of pH measurements and experimental uncertainty for the determination of the solubility product under discussion (PV 10c, 11b and 15b)

and Neck determine a considerably different H⁺ concentration.

PV 18g

Different, not considerably different. You calculated above this difference could be of about 0.2. Elsewhere in the draft we wrote that two of your experimental determinations for the log₁₀ of an equilibrium constant were inconsistent within your stated uncertainty, you complained because you said they only differed by 0.2 log₁₀ units. So in one case this is considerable, and in the other one it is not (actually it is not).

However, because they apply different H₂CO₃ dissociation constants... they finally obtain practically the same (correct) values of log [CO₃²⁻].

PV 18h

In other words, possible systematic deviation should cancel, I agree with this; nevertheless it, in principle, depends on the procedure used for calibration. Anyhow I made several assumptions (where was the fixed calibration point? Which slope was used?) I came to the same conclusion: the influence of this systematic deviation on solubility product under discussion, is within uncertainty (PV 10c, 11b and 15b).

=> General comment on “recalculations” of solubility or complexation constants for actinide carbonates

...in the NEA reviews there are often recalculations of original data...

Page 19

...The authors XY used the constant of log x to calculate the CO₃²⁻ concentration...

In many cases such recalculations do not correct the original data. Just in contrast - they make them incorrect,

PV 19a

Recalculation is usually performed, not specially to use other auxiliary data; but to check the numerical treatment and in a few cases, for sensitivity analysis. As written above (PV 18h), I rather used the assumption that the systematic error should cancel. You can typically see that the equilibrium constants I recalculated from your experimental measurements, are usually virtually the same as the data you published

The measured values of $\log[H^+]$ may be incorrect (e.g. those of Vitorge et al. in 3 M NaClO_4 , c.f. example above),

PV 19b

This is already discussed above, this is one possibility among others. From what you wrote above, it actually comes out that all the pH measurement in this media are incorrect, excepted yours, pointing out only one of these measurement, rather gives the impression you are not able to objectively write on the subject, and that your main goal is personal resentment, this actually is not much convincing from this point of view. My results are, in addition, a bad example for your demonstration for 2 reasons: (i) I rather used published auxiliary data hence the discussion should be for the corresponding publications (not mine) (ii) I also calculated the potentiometric titration curves for the experiments at 0.10 to 1 atm of carbonic gas, which model quite well the measured pH this is shown in CCE reports that Kim have (or at least had) and the needed information is anyhow published in a CEA report. This is also a checking of my pH calibration. This checking does not contradict your above discussion on slopes (PV 10c, 11b, 15c).

The mentioned recalculations are always restricted to the correction of the second step, they never correct the first step, which would be necessary as well.

PV 19c

Again you are making assumption on what was done, in that case it is clearly incorrect, see PV 19a.

And hence they lead to incorrect results !

PV 19d

Give numerical examples. the values I recalculated are usually virtually the same as the ones originally published.

Page 21

2.1. General comments on shortcomings of the SIT

...However, it is unacceptable that shortcomings and limitations of the SIT are simply ignored,...

PV 21a

Limitations of the SIT formula are well known from the beginning of its using.

(1) erroneous chemical conclusions are drawn and incorrect thermodynamic data are selected because of these shortcomings

PV 21b

Activity coefficients are usually a small correction, usually even smaller than the scattering of experimental data, specially when from different laboratories, only in special difficult cases it can induce such errors.

Beside this, the possible systematic deviation on the solubility product under discussion is within uncertainty as estimated from the numerical information you are giving in your comments (PV 10c, 11b and 15c).

(2) correct experimental data are ignored or, even worse, criticized as not reliable, because they are not consistent (or better: cannot be explained) with the simplified SIT approach used in the NEA-TDB.

PV 21c

Your experimental data were not ignored, they were used with the same weight as data from other laboratories to select thermodynamic values in our review. Only your activity coefficients were rejected, this has nothing to do with SIT or Pitzer formula: just because we extracted internally inconsistent values for activity coefficients, from your experimental data: your liquid-liquid extraction data are inconsistent with your solubility product data, when using TDB methodology: $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-)$ values in, at least, the range 0.18 to 0.34 kg.mol^{-1} was calculated in our review, the mean is 0.25, the TDB value is 0.25 ± 0.05 , the value you propose is 0.20. To resolve your contradiction you had to develop a new methodology, adding new empirical terms and fitted parameters for data at high I, while the aim of thermodynamic data bases is usually to provide values at $I=0$. Anyhow the new methodology you proposed does not seem to rely on enough experimental confirmations, and the way you used it up to now is not enough consistent with the TDB methodology (see PV 2i)

Limitations of the SIT

1) No triple ion interactions (as included in the Pitzer equations)

=> inaccuracies at high ionic strength ($I > 4 \text{ m}$)

PV 21d

This is known, used and written in the TDB guidelines and books from the beginning. This I range is enough for most of the published data. Anyhow the aim of the review is to select data at $I=0$, not to provide a formula to extrapolate them to very high I .

2) Debye-Hückel equation with a fixed value of $B\alpha = 1.5$

=> inaccuracies at low ionic strength ($I \rightarrow 0$) for ions with high charge $|z| > 3$

PV 21e

This has practical low importance. One does not know accurately the behaviour at low I , because accurate measurements are more difficult. The idea of the SIT formula is to give an explicit and reproducible way to extrapolate data to $I=0$, not to give the good corrections to pure water which nobody knows up to now. Pitzer formula is not better, and this is already published: you need data at low ionic strength to fit one of the second virial parameter, unfortunately you often do not have such data for complexes, hence the fitted low I parameter is not well defined, neither is the result at $I=0$. Some authors who recognised this type of limitation (not enough accurate data in the good I range to fit all the parameters) usually fix one parameter or use other approximations, some of them are perfectly correct; but you end up with a formula no more accurate than the SIT formula. Note that both formulae are empirical, their fixed parameters or functions were fitted on experimental data (usually isopiestic ones). Finally there is no good universal approach (see also PV 2i).

3) Simplification: negligible anion-anion and cation-cation interactions

=> general problem, which makes it impossible to use the same SIT coefficients for carbonate trace activity coefficients in different electrolyte media

PV 21f

There are not (or very few) direct experimental evidence of such interactions. And your treatment is not enough consistent with the SIT (see PV 2i and 21c).

the scientific problems do not arise as a question of SIT or Pitzer modelling - they arise from experimental data!

PV 21g

This was already written in our draft, before your comments. Please avoid writing 44 pages on "objections" on which we agree.

Pitzer modelling performed ...only ... in order to demonstrate the errors coming from the oversimplification of the SIT procedure.

PV 21h

See above, each formula has its drawback. One advantage of Pitzer formula is it has three times more parameter to fit, among which even one at low I where there is no experimental evidence that there is any difference between different electrolytes of same stoichiometry. This advantage can become a disadvantage since you can simply fit the errors, which you possibly did since, as mentioned above, you published data from which inconsistent activity coefficients can be extracted, nevertheless I did not try to calculate whether this propagates error on other activity coefficients. (see also PV 2i and 21e)

According to the results in [96FAN/NEC], where the mixing parameters for CO_3^{2-} and HCO_3^- in NaClO_4 solution were evaluated from H_2CO_3 dissociation constants, the trace activity coefficients of the CO_3^{2-} in NaCl and NaClO_4 solutions above 1 molal are considerably different

PV 21i

Again, not considerably different: different. This can also be seen directly by comparing the values for the acidic constants in both media.

These differences cannot be described with the simple SIT approach used in the NEA

PV 21j

You are right: there is only one empirical parameter for each pair of anion-cation. You could as well add new terms to the SIT formula (PV 2i), typically using a third (instead of second) virial development which would result in empirical parameters for 3 species (hence at least 2 of the same charge), for consistency it would then also be needed to use pair empirical parameters for ions of the same charge. This was not done for a series of reasons: one is simplicity (typically the present SIT formula is symmetrical for each pair) hence avoiding too many parameters which would require more experimental data than available; a second series of reason is that there is no experimental evidence that such parameters are needed, excepted in the 2 cases you pointed out. Hence the discussion should first be on the experimental results corresponding to these two cases. The first case is the above difference in your measurement of the carbonic acid constant between NaCl and NaClO₄ media, as discussed above, you might be right, but experimental confirmation is still needed. The second case is discussed below, it is for Np(V) system, you claimed such mixed empirical parameters are needed for all the anionic Np(V) carbonate complexes, while there is experimental indication only for the first one and again experimental confirmation is still needed, at least because inconsistent values for the activity coefficient of the aquo cation can be extracted from your experimental measurements (solubility and liquid-liquid extraction).

If these experimental confirmation will be obtained, it would be straightforward to add the extra terms and fit the corresponding parameters (PV 2i).

Page 22

(However, these interactions are not zero,

PV 22a

Not exactly, experimental confirmation is needed: hence "might not be zero..."

two further corresponding examples

PV 22b

This is a good idea to look for further examples. As you wrote it is better using data for which reliable experimental information is available. Nevertheless I will not discuss your examples, because it would first be needed to review the experimental data which is out of the field of the present discussion. Just one comment: it is better to compare the experimental data than Pitzer parameters, because you usually obtain triple interaction parameters even if they finally cancel when using them for modelling the original experimental results. This because there is not usually a unique set of Pitzer parameters than can be fitted on a single system: Pitzer parameters are correlated as pointed by Pitzer in his earlier publication, this is the drawback of having so many empirical parameters (see also PV 2i).

Page 23

Fig.2.1

PV 23a

This figure essentially provide new information as compared to figures 1.4 (it is not "a new proof", but a combination of figures 1.4)

Page 25

Neptunium(V)

Solubility constant for NaNpO₂CO₃ · xH₂O(s)

In the NEA review, the results of Kim and coworkers... were generally criticized to be not reliable and disregarded

PV 25a

These data were used to select the standard equilibrium constant of the Np(V) carbonate data, hence they are not disregarded. Only your activity coefficients were disregarded. As already said from the beginning and in my e-mail to Phanganel in 1997, the reason is that inconsistent values for the activity coefficient of NpO₂⁺ are calculated from your own data using TDB methodology, hence we neither used the other data.

...this is a pure arbitrary act of the reviewer,

PV 25b

A review is always no more than the opinion of the reviewers (not reviewer actually), in this particular case it is not arbitrary (see the reason just above: PV25a).

The evaluation of the solubility constant at $I = 0$ is primarily based on solubility studies of Maya [83MAY] (in 1 M NaClO₄) and Vitorge (in 3 M NaClO₄)

PV 25c

You forgot one study: your own solubility data extrapolated to $I=0$

As a consequence, the reviewer concludes that the results of Maya [83MAY] refer to a hydrated solid phase and those of Vitorge [86GRE/ROB] to an aged, less hydrated solid phase

PV 25d

Not only, the main reason is difference in equilibration time. Maya did not published its X-ray diffraction pattern; but if we accept what he claimed, his diffraction pattern was not the same as in my work. So the conclusion is not specially based on the values of the solubility products; but rather on kinetics observation (see also PV 3d, 3-4, 4d, 5, 6d, 18c and d).

Page 26

Fig. 2.4

PV 26a

Several interpretations can be given for this figures, but it is clear that all data in NaClO₄ media are consistent within uncertainty considering usual possible errors for such systems: ripening of the solid and different calibrations procedures between different laboratories. The figure is not well chosen to compare data in NaCl and NaClO₄ media, because the activity coefficient of NpO₂⁺ cation, hence $\Delta\epsilon$ values in both media are expected to be different.

However, this recalculation is somewhat speculative. It is based on the assumptions that the scattering of the experimental data is due to the presence of different solid phases

PV 26b

Again, it is not an assumption, but the result of x-ray diffraction study.

I performed the same calculation from your data: at $I=0.1$ to 3 M

$$\lg K_s^\circ = -10,954 \pm 0,14740043$$

$$\Delta\epsilon' = 0,233 \pm 0,070$$

$$\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0,335 \pm 0,100$$

where uncertainty is 1.96 standard deviation

$\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = \Delta\epsilon' - \epsilon(\text{Na}^+, \text{ClO}_4^-) - \epsilon(\text{Na}^+, \text{CO}_3^{2-}) - 3.5 [a(\text{H}_2\text{O}) / m]$, the last term is an approximation for water activity

$$\epsilon(\text{Na}^+, \text{ClO}_4^-) = -0.01 \pm 0.01$$

$$\epsilon(\text{Na}^+, \text{CO}_3^{2-}) = -0.06 \pm 0.06$$

$$a(\text{H}_2\text{O}) / m = -0.015 \pm 0.01$$

These three last values are calculated from Pitzer parameters given in your publications or Pitzer's; but with no mixing term, the idea was only to be as much consistent as possible with your data (hence not using TDB auxiliary data). The only difference with TDB auxiliary value is $\epsilon(\text{Na}^+, \text{CO}_3^{2-})$ is bigger by 0.02 kg.mol⁻¹, hence you can decrease $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-)$ by 0.02 kg.mol⁻¹, which anyhow is within uncertainty.

On your Fig2.5 page 27, you obtained a little more than -11.0 for $\lg K_s^\circ$ which is consistent with the value I calculated. You wrote below this figure $\Delta\epsilon = 0.29$. From my above values I calculate $\Delta\epsilon = \epsilon(\text{NpO}_2^+, \text{ClO}_4^-) + \epsilon(\text{Na}^+, \text{ClO}_4^-) + \epsilon(\text{Na}^+, \text{CO}_3^{2-}) = \Delta\epsilon' - 3.5 [a(\text{H}_2\text{O}) / m] = 0,285$ which is the same value as yours. You wrote on the figure $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0.20$ [95NEC/FAN], while from $\Delta\epsilon'$ I obtained

$$\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0.335 \pm 0.099$$

which is different. It should not be since our $\Delta\epsilon$ values are consistent. This is a bit misleading: the choice of 0.20 is not much consistent with TDB methodology as used for your solubility product.

From your liquid-liquid extraction data, I obtained

$$\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0,182 \pm 0,013$$

which is neither consistent with TDB methodology, nor with you solubility data.

Page 28 you wrote this inconsistency in your own data can be resolved by using $\epsilon(\text{Na}^+, \text{CO}_3^{2-}) = +0.04$ instead of -0.06 (as I used above, or -0.08 which is the TDB auxiliary value): this results in the value of

$$\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0.24 (\pm 0.07 \text{ at least: or more depending the uncertainty you attributes to } \epsilon(\text{Na}^+, \text{CO}_3^{2-}) = +0.04)$$

(and not 0.20) instead of 0.335 ± 0.099 . 0.24 is virtually the TDB value (0.25); is still different from the value

$$\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0,182 \pm 0,013 \text{ extracted from your liquid-liquid extraction data. This last inconsistency will be}$$

discussed below; hence the discussion is now on $\epsilon(\text{Na}^+, \text{CO}_3^{2-}) = +0.04$ instead of -0.06 . Your explanation might be correct, but there are still two apparent inconsistencies:

1. isopiestic measurements rather gives a negative value for $\epsilon(\text{Na}^+, \text{CO}_3^{2-})$
2. from your Pitzer parameters I also obtained a negative value for $\epsilon(\text{Na}^+, \text{CO}_3^{2-})$

Both inconsistencies can be resolved; but quite in a contradicting way unless you add a new parameter. The first one refers to high carbonate concentration where activity coefficients could be different from trace activity coefficient (as it seems you suggested). To deduce trace activity coefficients from Pitzer parameters obtained at high concentration, one possibility should be to use only the ion pair empirical terms and parameters (PV 2i). It is what I did and end up with the second inconsistency. Conversely you can object I only used the ion pair empirical terms and parameter to calculate $\epsilon(\text{Na}^+, \text{CO}_3^{2-})$ from your Pitzer parameters; but in that case you cannot resolve the first inconsistency... Finally it seems you added an empirical term and a parameter for the $(\text{ClO}_4^-, \text{CO}_3^{2-})$ pair. This is mathematically correct, but this can also be interpreted as adding a new parameter because the existing model cannot predict a new set of experimental observations: your conclusion is the existing model is indeed incorrect, the opinion of our review is either the existing model is too limited, or the new sets of experimental measurements are incorrect. For this reason we did not use your activity coefficients.

You provided further liquid-liquid extraction experimental results which is a good idea to have independent confirmation: From your liquid-liquid extraction data in 0.2 to 3 M NaClO_4 , I calculated $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0,182 \pm 0,013$

this value is inconsistent with value I calculated above (0.335 ± 0.099) from your solubility data; but is consistent with the value $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-) = 0.24 (\pm 0.07 \text{ at least})$; but is not consistent with the TDB value (0.25 ± 0.05). In your liquid-liquid extraction data, there was only one anion, hence no eventual anion-anion parameters. The same was true for the electrochemical and isopiestic measurements used to select the TDB value, hence I see no way to resolve this last inconsistency.

However since we have a data (actually 2 inconsistent ones) with no anion-anion parameters we should use them to select a value for $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-)$ as explained above (see also PV 2i) to keep consistency with simplified formula, i.e without triple or anion-anion term, i.e the actual SIT formula (that you want to extend with these extra terms). First possibility you take the TDB value (typically because all this is to obtain consistency with the TDB), hence you decide there is an experimental error or an over simplification in the data treatment of your liquid-liquid extraction study, the drawback of this, is since you published eventually erroneous activity coefficient, our review might prefer to avoid using all your activity coefficients. An alternative would be to use the value extracted from the liquid-liquid extraction data; but since it is not consistent with the TDB value, our review would reject this interpretation. Alternatively you proposed a sort of compromise, a medium value of 0.20: this is a possible solution, but it is finally supported by less clear experimental evidence, than the existing value. It seems new experimental determinations are needed for binary system. I see two of them: isopiestic measurements and $\text{NpO}_2\text{OH(s)}$ or $\text{Np}_2\text{O}_5\text{(s)}$ solubility product. I think you have data for this system: can you extract a value for $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-)$ from the hydrated hydroxide (or oxide) solubility product? On the other hand the solid phase is even less well defined than from the carbonate solid, hence the effect of activity coefficient might be less than the scattering of the data.

Anyhow, the influence of triple terms for the activity coefficient of typically the limiting carbonate (anionic) complex is not negligible, while there is no difference in its activity coefficient in sodium chloride and perchlorate media. As explained above, to keep consistency with the (pair) SIT formula the triple terms should be set to zero, as they should in all the similar cases of your interpretation to improve consistency with TDB.

Note that to resolve your contradictions, you need to add a new empirical term and a fitted parameter, even adding this you end up with values for $\epsilon(\text{NpO}_2^+, \text{ClO}_4^-)$ with associated uncertainty bigger than the value selected in our review (0.25 ± 0.05). In addition you introduced this new empirical term and a fitted parameter to fit experimental determinations of carbonic acid constants in NaClO_4 media discussed elsewhere in your comments. More experimental evidence is also still needed on this specific problem, i.e to confirm experimental evidence of the need of anion-anion or triple terms and empirical fitted parameters. For these reasons, we cited your interpretation and said from the beginning (before your started this polemic) it might possibly be correct: some TDB auxiliary values might need revision; but more experimental evidence is needed, conversely it cannot be ruled out there is an experimental error in your measurements.

Page 27

It is to note that the solid line (for the data in NaCl solution) is predicted by independent SIT coefficients: $\epsilon(\text{Na}^+/\text{Cl}^-) = 0.03$ and $\epsilon(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$ from the NEA-TDB [85SIL/BID], $\epsilon(\text{NpO}_2^+/\text{Cl}^-) = 0.09$ from a solvent extraction study in [95NEC/FAN], and hence $\Delta\epsilon = 0.04$. For both media (NaCl and NaClO_4) the extrapolation to $I = 0$ leads to a consistent value of $\log K_s^\circ = -11.0 \pm 0.2$.

PV 27a

As above, uncertainty is 1.96σ . In the calculation below I entirely attributed to activity coefficient, shift from NaCl media to ideal pure water. This is consistent with the decision of our review not to consider any Np(V) chloride complex, while it might not be with your other publications where you determined chloride complexation constants for Np(V), nevertheless it seems you also used the same assumption as in our review (hence possibly inconsistent with your published interpretation for your other measurements), therefore we used the same assumption and finding the same numbers will indeed mean we agree on these numbers.

From your liquid-liquid extraction data in 0.2 to 3 M NaCl, I calculated

$$\varepsilon(\text{NpO}_2^+, \text{Cl}^-) = 0,067 \pm 0,013$$

which is rather consistent with your value

From your solubility products in 0.1 to 3 M NaCl, I calculated (as above in NaClO₄ media)

$$\lg K_{\text{so}}^\circ = -11,01 \pm 0,26 \text{ which is consistent with your value } (-11.0), \text{ and the value I calculated above in NaClO}_4 \text{ media } (-10,95 \pm 0,15)$$

$$\Delta\varepsilon = -0,00 \pm 0,07$$

From $\Delta\varepsilon = -0,00 \pm 0,07$, I calculated:

$$\varepsilon(\text{NpO}_2^+, \text{Cl}^-) = 0,09 \pm 0,10$$

which is consistent with your above value (0.09), and the value I calculated from your liquid-liquid extraction data ($0,07 \pm 0,01$)

Hence we agree on the calculation (even if we eventually do not present them in the same way) in NaCl media.

I found inconsistency between data extracted from the results of your measurements using two different experimental techniques in NaClO₄, but not in NaCl media. Different explanations can be given for this; typically (non of them can be proven (excepted eventually the first one)):

-I made an error in my calculations or in interpreting your data in NaClO₄ media,

-activity corrections are less important in NaCl than in NaClO₄ media, consequently possible systematic error is within uncertainty in NaCl but not in NaClO₄ media,

-there is a true chemical or experimental difference in both media (in typically the liquid-liquid extraction systems, or the solubility studies, typically non negligible difference in ripening of the solid phase in the NaClO₄ work, but not in the NaCl one, impurity in NaClO₄...)

-and of course your above explanation (need of extra empirical term and anion-anion fitted parameter) but it does not seem it provides consistency for your liquid-liquid extraction data in NaClO₄ media.

the solubility data in [91KIM/KLE, 94NEC/RUN, 94NEC/KIM, 94RUN/KIM, 95NEC/RUN, 95NEC/FAN, 95FAN/NEC, 96RUN/NEU] refer to the same solid phase, the hydrated $\text{NaNpO}_2\text{CO}_3 \cdot 3.5\text{H}_2\text{O}(\text{s})$ described by Volkov et al. [77VOL/VIS] and Maya [83MAY].

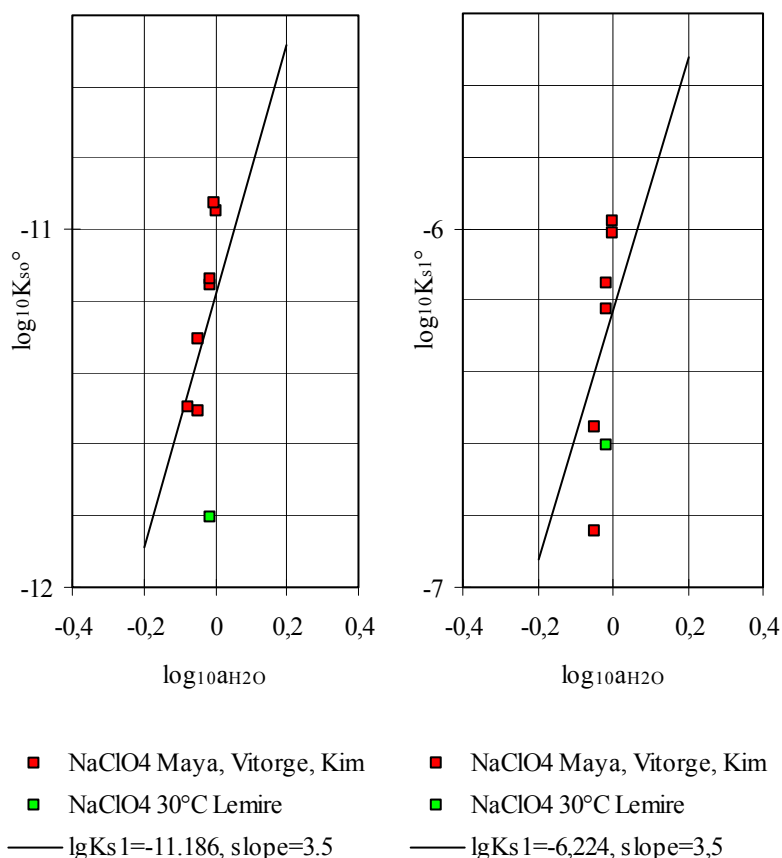
PV 27b

The x-ray diffraction patterns published in [77VOL/VIS] and [95NEC/RUN] are the same, but I do not know whether they correspond to samples prior or after the achievement of equilibrium solubility

and Maya [83MAY].

PV 27c

That is what he said in his publication, but the x-ray diffraction patterns were not published (see also PV 3d, 3-4, 4d, 5, 6d, 18c and d), and it seems it was for an initial solid phase, while equilibration time was relatively short, and his two series of measurements can be interpreted as evidence of ageing of the solid phase (but this interpretation cannot be proven. I also tried correlation with water activity because Volkov et al. interpreted similar variations in the x-ray spectra as observed from those published with solubility data, with variation of the hydration of the solid compounds; but this effect on the solubility is probably less important than other reasons for the scattering of the data (see the figures)



(Solely Meinrath [94MEI] reported an other (hexagonal) modification of $\text{NaNpO}_2\text{CO}_3 \cdot x\text{H}_2\text{O(s)}$, with different x-ray pattern).

[PV 27d](#)

X-ray spectra published by Meinrath can be interpreted with a mixture of the solid phase you prepared (or obtained after solubility equilibration?), and the one I obtained after solubility equilibration. (see the figure referred in PV 2c)

Page 28

The dashed line would be predicted with $\varepsilon(\text{Na}^+/\text{CO}_3^{2-}) = +0.04$ (in NaClO_4 solution), evaluated in section 2.3 from the H_2CO_3 dissociation constants determined in [96NEC/FAN].

[PV 28a](#)

This is a possibility; but only one possible interpretation (see PV 27a)

It becomes evident that (within the range of experimental uncertainties) all studies in NaCl solution lead to a consistent value of $\log K_s$, in particular if we assume the hydration number of $x = 3.5$ given in [83MAY].

[PV 28b](#)

There is no determination of $x = 3.5$ in ref.[83MAY], only the indication that he prepared a solid compound of x-ray diffraction pattern consistent to a published one of $x = 3.5$. Realistic structures have been proposed for this type of solids; but they cannot be proven.

(with the exception that $\varepsilon(\text{NpO}_2^+/ \text{ClO}_4^-) = 0.20$ is used, instead of 0.25 as proposed by Vitorge,

[PV 28c](#)

The value of 0.25 is the TDB auxiliary value, it was selected in the Uranium book (because the methodology had been first validated on Uranium), and confirmed in our review, it is actually from Riglet's thesis (who was my student), see also PV 27a.

Remember: $\varepsilon(\text{Na}^+/\text{CO}_3^{2-}) = -0.08$, the SIT coefficient of the NEA-TDB was not appropriate for the carbonate trace activity coefficients in NaClO_4 solution above 1 molal (Fig. 2.1).

PV 28d

Not exactly, it is your proposition; but, as discussed above, more confirmation is needed.

Now, consistent values at $I = 0$ are obtained for 12 solubility experiments from 5 different investigators with the solution composition widely varied ($I = 0.1, 1, 3$ and 5 M in both, NaCl and NaClO_4 solutions). The average value is found to be

Pages 28-29

PV 28-29

From your calculations it appears that SIT and Pitzer treatments gave standard values for the solubility product in a range of ± 0.19 and 0.10 respectively (-10.96 ± 0.19 and -11.06 ± 0.10 , uncertainty is 1.96σ), hence Pitzer's treatment gives smaller uncertainty and smaller range of values, but more empirical parameters (than in the SIT formula) were used. However, all the data are from the same laboratory: this is an important (for our discussion) difference with the data in NaClO_4 media.

In NaClO_4 media the range is ± 0.74 and 0.37 respectively (-11.215 ± 0.58 and -11.09 ± 0.28), as for data in NaCl media Pitzer's treatment gives smaller uncertainty and smaller range of values, more empirical parameters were again used, in addition you also had to add an extra fitted parameter as compared to the treatment of data in NaCl media. Note that anyhow, while there are more experimental data in NaClO_4 media and more fitted parameters in the Pitzer treatment, the range of values is bigger than in NaCl media which is quite expected (data are no more from the same laboratory here), but statistical uncertainty is also bigger, while the contrary would be expected from a statistical point of view, if in both media uncertainty of each experimental determination was the same. One logical assumption (which cannot really be demonstrated) is of course that uncertainty is less important in NaCl media because all the data are originated from the same laboratory... In other word typically differences in auxiliary values and/or calibration procedures are usual. Nevertheless it cannot also be ruled out from a statistical point of view, that the extra fitted parameter you introduced (for "anion-anion interactions") actually fit this type of inter laboratory systematic deviation or other chemical problem.

Anyhow X-ray diffraction patterns (figure attached to comment PV 2c) indicate the solids used in 2 of the studies were, or might have been different from the other ones. These two solids should be closer, while they surprisingly gave the maximum and minimum solubility products. Disregarding these two data in NaClO_4 media, one obtain -11.21 ± 0.46 and -11.10 ± 0.21 in NaClO_4 media. This time all the data are from the same laboratory (yours), excepted Maya's one which is in agreement with yours, there are still more data than in NaCl media, nevertheless statistical uncertainty is still more important.

You have a consistent interpretation for all the data; but these discussion indicate a small statistical problem, hence possibly originated in chemical or other problems (or maybe there is not enough data to support the statistical discussion). This was in part expected because inconsistent results were at high ionic strength, and you added an extra empirical term and a fitted parameter to solve this problem.

Your treatment also showed there is no difference (within statistical uncertainty) between the standard values of the solubility products determined from data in chloride and perchlorate media. Again this is not really new because we saw from the beginning at $I=0$ (or as you pointed out at low I) there is no problem. Your point is that a single solubility product is enough to interpret all the data, this is correct: there is no statistic outlier in the overall series ($\text{NaCl} + \text{NaClO}_4$) using a 1.96σ threshold, nevertheless in the NaClO_4 series (treated alone) there is one outlier ([94MEI]). I also checked this on the values I reinterpreted, the numbers are not much different, Meinrath's value can still be considered as a statistical outlier; but this time mine too, and since its solubility product is less than the others, this would mean in principle that it corresponds to the most stable solid, hence the value could be selected (instead of the mean of less aged solids); but I would not rely on this analysis because the 1.96σ threshold is quite arbitrary. If now we add the values reported in ref. [93LEM/BOY] at 30° (or the result of reinterpretation of this study, in our review), it is an outlier: from a statistical point of view, and it correspond to an even more stable compound than the one I prepared, and for chemical reasons it corresponds to a very interesting study that showed evidence of solid phase transformation.

Now what is the consequence of this: our review had a first possibility that was to choose a single solubility product. Actually this was really discussed, and for a time proposed in the course of the review. The value could be any of the values discussed above (typically -11.15 ± 0.59): they are all within statistical uncertainty. According to the methodology adopted in our review for estimation of uncertainty, we should not select in this case the statistical uncertainty, because there is experimental evidence of changes in the solid phase, hence uncertainty should have

covered the whole range of published values and possible associated uncertainty and systematic errors. This was of course not much satisfactory, because one of the main reason for this relatively big uncertainty was not much consistent with thermodynamics: if the solid phase was changing, the selected value should correspond to the most stable phase, hence the less soluble: the smallest solubility products. This type of discussion was not specific to this solid/liquid system, and we choose one of the usually adopted solution in our and previous TDB review. Our solution has also its drawback, but is more consistent with experimental observations..

Anyhow this decision of selecting two solubility products is finally quite independent from the above discussion on activity coefficients and pH calibration, because its rely on direct experimental qualitative evidence. It might even be reflected in your numbers because, as discussed in our review, several solubility product values seem to be logically correlated with their time of equilibration: from your numbers, as from the ones used in our review the most stable solid at room temperature is the one I prepared, and anyhow the one prepared by Lemire at slightly higher temperature (30°C) is even more stable. On the other hand it is not clear whether this difference between solubility product values, are more than possible systematic deviation between methodologies used in different laboratories.

Page 31

PV 31

The discussion of the second Np(V) solid phase is quite similar as above, hence I do not repeat it.

I have the additional comments:

- Instead of discussing K_{s0} value, it is better to discuss K_{s3} value which is directly determined experimentally, while K_{s0} value has to be deduced from thermodynamic cycle.
- another advantage of K_{s3} is that it involves only one Np(V) aquo ion, and you more easily discuss ionic strength corrections, specially you see there is no difference between chloride and perchlorate media for the limiting complex (you also see it directly on the solubility curves). For this reason it is quite artificial to discuss (again) NpO_2^+ activity coefficient in this part: this species is only introduced in the thermodynamic cycle you implicitly used to deduce K_{s0} from K_{s3} , hence it is certainly better to performed this part of the cycle at $I=0$ avoiding propagation of the errors originated in the uncertainty of the activity coefficient for NpO_2^+ . This is more or less what you wrote, but did not use in your calculation.
- you did not mentioned the very important work of Simakin.

Pages 32-34

I am not discussing Am and U: the review is already published

Page 35

2.3. Proposal to solve the problem of carbonate trace activity coefficients:

In the tables above, the solubility constants at $I = 0$ for actinide carbonate solids are partly based on SIT activity coefficients and partly on Pitzer activity coefficients.

PV 35a

We used another procedure in our review: simply excluding activity coefficients originated from studies were systematic errors could not be ruled out... and we end up we similar standard values as you. Hence it might not be so important for TDB to add more sophistication in its methodologies, anyhow the SIT formula was chosen to have a reliable way to extrapolate data to $I=0$, not really to model electrolyte solutions.

However, anion/anion interactions generally become important if the charge of an anion is -2 or larger,

PV 35b

In the case under discussion it is rather the contrary: there is clearly a problem for $\text{NpO}_2\text{CO}_3^-$ whose activity coefficient indeed seems different in NaCl and NaClO_4 media (unless this difference is due to an undetected mixed carbonate chloride complex), while there is no experimental evidence of this difference for the more negatively charged complexes typically K_{s3} ionic strength corrections are the same in both media (see PV 31).

Page 36

This includes the assumption that the reported solubility data refer to the same solid phase. For the data from [83MAY, 94NEC/RUN] this is well ascertained.

PV 36a

Not really (see discussion above) the solid was possibly slightly changing during then measurement ,and it seems it was not characterised or control after the achievement of equilibrium solubility in ref. [83MAY]

PV 36b

It is not a very good idea to select a correction on the correction of activity, hence a very small difference, from Np(V) carbonate solubility data, because ripening of the solid phase during solubility measurements cannot be ruled out in the existing available experimental information.

Page 37

PV 37a

The results showed on the figure were essentially discussed above, as was a part of the rest of this Section: more experimental confirmation is needed to check your interpretation.

Page 38

2.4. Consequences for NEA-TDB reviews

(1) Limitations of the SIT have to be pointed out.

PV 38a

they are already known and documented

(2) The SIT has to be extended for anion-anion interactions (c.f. proposal), at least in cases where this is necessary to avoid erroneous conclusions

PV 38b

Experimental confirmation is needed

(3) a) New (correct) trace activity coefficients for CO_3^{2-} in NaClO_4 should be used.

=> All ϵ coefficients of actinide carbonate complexes, derived from exp. data at high NaClO_4 concentrations should be reevaluated.

PV 38c

Only if your interpretation is correct, which needs confirmation

If possible, the experimental results on H_2CO_3 dissociation constants given in [96FAN/NEC] should be checked in an independent laboratory.

PV 38d

This is the first thing to do, it will then be time for your other proposals if confirmation is obtained

b) As long as there is no final decision on the problem in 3a), only unambiguous data (in NaCl solution or at low NaClO_4 concentration, where the NEA-TDB auxiliary data on the carbonate ion are free of any doubt) should be used to evaluate SIT coefficients and equilibrium constants at $I = 0$.

PV 38e

You will then come into the debate weak complex/activity coefficient to model actinide cations interactions with Cl^-

(4) Np/Pu review

a) Conclusions for solid Np(V) carbonates are incorrect. The selected $\log K_s$ values have to be changed.

PV 38f

From the above discussion, the qualitative discussion may be reworded, but the problem was handle in our review in a way to limit such systematic deviation, finally the numbers you calculated are usually consistent with the ones calculated in our review. The solid phase problem is quite independent from the activity coefficient discussion: it rely on x-ray studies and kinetics observation, both not much originated from your work.

b) $\log \beta^\circ$, and SIT parameters ($\Delta\epsilon$ and ϵ) have to be reevaluated, including data in NaCl solution for the equilibria $\text{NpO}_2 + n \text{CO}_3^{2-} \rightleftharpoons \text{NpO}_2(\text{CO}_3)_n$ (or for the stepwise constants),

because NaCl is an important medium with respect to natural aquatic systems. If necessary, the $\log \beta^\circ$ values can be fixed from the corresponding extrapolation with data in NaClO₄ solution (c.f. Fig.2.7, next page). The known values of $\varepsilon(\text{NpO}_2 + / \text{Cl}^-)_{\text{NaCl}} = 0.09 \pm 0.02$ [95NEC/FAN] and $\varepsilon(\text{Na} + / \text{CO}_3^{2-})_{\text{NaCl}} = -(0.08 \pm 0.03)$ can then be used to evaluate SIT coefficients $\varepsilon(\text{Na} + / \text{NpO}_2(\text{CO}_3)_n^{1-2n})_{\text{NaCl}}$ for the Np(V) carbonate complexes in NaCl solution.

PV 38g

Taking data only in NaCl media might end up with taking data only from your laboratory, which is not satisfactory. It seems the solution you propose to avoid this is more or less the same as the one we used in our review; but giving the reverse role to chloride and perchlorate media, I do not think this is worthwhile to restart all the work this way because of course practically the same standard values would be obtained.

Page 39

Fig. 2.7

PV 39

These figures might be misleading because it seems that the activity coefficient difference in chloride and perchlorate media is more important for the 1-3 complex than for the 1-2 than for the 1-1, while experimental solubility directly show that there is a difference only for the aquo cation (which is expected) and with the first complex (which is not expected because it is an anion)n using K_{st} better shows this.

I received today the 44 pages document of comments on the Np/Pu draft by Neck and Fanghänel. I first looked for new piece of information that would require quick action to further update the Np/Pu draft; it appears these comments are not really part of a needed dialogues on specific scientific questions and numbers pointed out in the course of the ongoing discussion; but rather more details on their opinion concerning their own work and a few others. It seems to me, this type of comments on the Np(V) sections have already induced proposal to modify the draft, other comments are clearly for TDB II project. I will continue examining possible new changes in the draft, or performing new calculations and checking if I am asked to, as usually done within our TDB group at this very stage of its duty; but I will not repeat here previous discussion. To this point I am not much interested in the opinion of the authors; but to their arguments: there are not really new ideas, but rather new examples in which they wrote a few errors concerning the description of experimental methodologies used in them.

Typical:

The only point (again not new in this discussion) where I thought they should be right, is the activity coefficient of the highly negatively charged Np(V) limiting complex in sodium chloride and perchlorate media. Figures they published and Fig.2.7 in their last comments show differences between both media, they interpret this as specific anion-anion interactions to taken into account in activity coefficient. Actually non of these figures is enough to discuss this point; hence I rather used similar figure for the corresponding K_{s3} values measured in their laboratory... there is no evidence of any difference between both media: I did not really expect this, I was actually expecting to add something about it in the selection of the corresponding interaction coefficient..

This is a typical:

- the problem is not important (just an activity coefficient for a species to be used in media that do not really correspond to conditions for waste disposal),
- it is an interpretation of a global curve fitting exercise, where they did not realised an (other) specific treatment of the data should show direct qualitative evidence of their interpretation. Beside this, if I did not make any error in my figure, it seems the interpretation is finally incorrect: just check it by yourself (see figure below).

An advice to Neck

By working with TDB specialist groups, I (very progressively) learnt to handle problems I found in other's work. I did not write above "a first year student will, at first glance, see your demonstration is, completely stupid beside several errors you made on other problems";

but only: I am just suggesting you to plot a modified SIT figure for K_{s3} , I already plotted it myself; but I prefer you do the same work on your side to check my calculations.

This avoid being ridiculous, otherwise the one who did not think of the special way to perform the calculation, or the one who made an error in the calculation will, and it would be impossible to run a true working team this way.

Figure:

For those who are not really familiar with SIT using, the figure to plot is (after corrections to molal units when needed)

$(\log_{10} K_{s3} - 18 D)$ v.s m_{Na^+}

its theoretical is $\log_{10} K_{s3}^{\circ} - \Delta \epsilon m$,

where $\Delta \epsilon m = (\epsilon(NpO_2(CO_3)_3^{5-}, Na^+) - 2 \epsilon(CO_3^{2-}, Na^+)) m_{Na^+} + \epsilon(X^-, Na^+) m_{X^-} + 3.5 \log_{10} a(H_2O)$

where X^- is chloride or perchlorate anion,

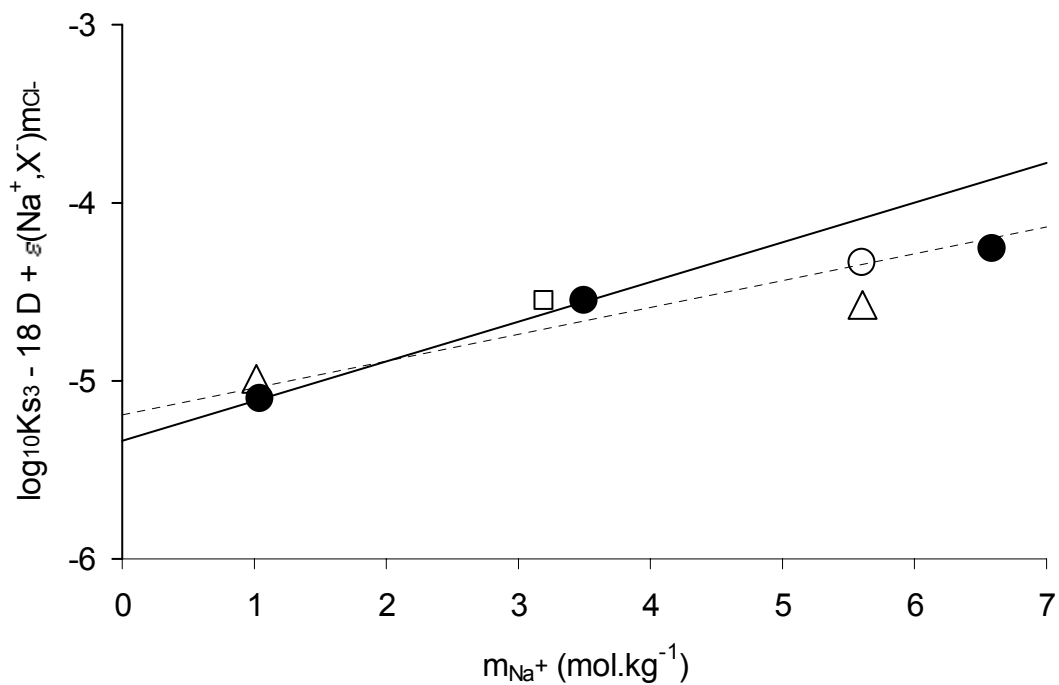
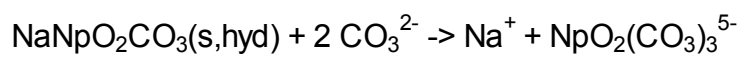
as a first approximation the major contribution to the slope, $-\Delta \epsilon$, is $\epsilon(NpO_2(CO_3)_3^{5-}, Na^+)$ and this can show whether it has or not different values in chloride and perchlorate media.

Instead of this approximation you can plot a figure similar to the above (classical SIT) one:

$(\log_{10} K_{s3} - 18 D + \epsilon(X^-, Na^+) m_{Cl^-} + 3.5 \log_{10} a(H_2O))$ v.s m_{Na^+}

Figures added later to this file. They evidence no difference between Cl^- and ClO_4^- media for equilibrium constants written with negatively charged aqueous complexes (typically $NpO_2(CO_3)_3^{5-}$), as determined by the authors, while they used different empirical coefficients for anions in these two media.

Ks₃ for



● NaClO₄ [91KIM/KLE, 94NEC/RUN, 95NEC/FAN]

○ NaCl [94NEC/KIM]

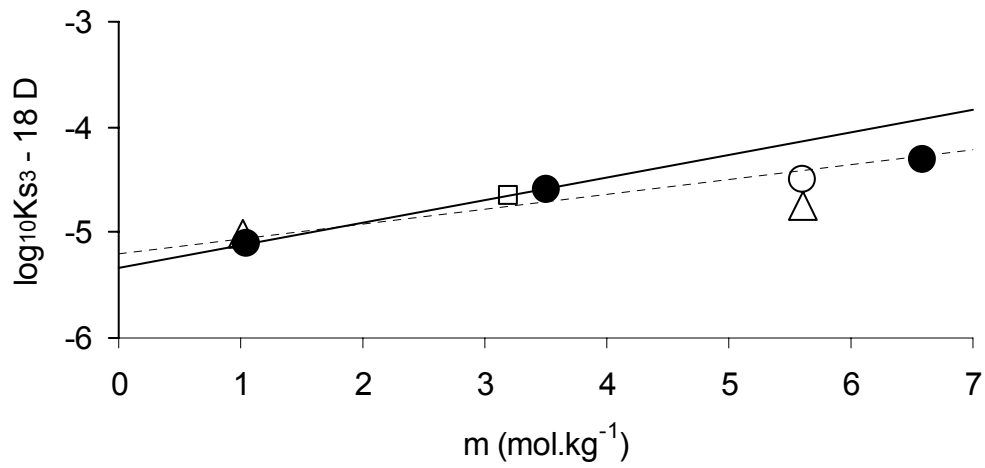
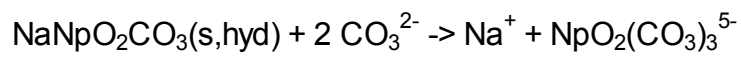
△ NaCl [94RUN/KIM]?

----- 1 to 5 M NaClO₄

———— 1 and 3 M NaClO₄

□ [96RUN/NEU]

Ks₃ for



● NaClO₄ [91KIM/KLE, 94NEC/RUN, 95NEC/FAN]

○ NaCl [94NEC/KIM]

△ NaCl [94RUN/KIM]?

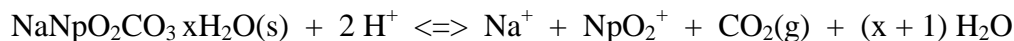
----- 1 to 5 M NaClO₄

—— 1 and 3 M NaClO₄

□ [96RUN/NEU]

Solubility of $\text{NaNpO}_2\text{CO}_3\text{xH}_2\text{O(s)}$ in NaClO_4 solution [91KIM/KLE, 94NEC/RUN]

The dissolution equilibrium of $\text{NaNpO}_2\text{CO}_3\text{xH}_2\text{O(s)}$ can be written as



This formulation corresponds directly to the experimental data. It is independent of carbonic acid dissociation constants or carbonate activity coefficients, on which there is still ongoing discussion within the NEA-TDB project.

The H^+ concentration is obtained from the measured $\text{pH}_{(\text{obs})}$ (the experimental relations between $\log [\text{H}^+]$ and $\text{pH}_{(\text{obs})}$ are given in [91KIM/KLE, 94NEC/RUN]), the NpO_2^+ concentration is measured, Na^+ is given by the NaClO_4 concentration and $\text{pCO}_2 = 10^{-3.52}$ bar in the solubility experiments discussed.

The following equilibrium constants $\log K_s$ for the reaction above can be derived directly from the experimental data:

$$0.1\text{ M NaClO}_4: \log K_s = 7.28 \quad (a_w = 0.997)$$

$$1.0\text{ M NaClO}_4: \log K_s = 7.37 \quad (a_w = 0.966)$$

$$3.0\text{ M NaClO}_4: \log K_s = 7.54 \quad (a_w = 0.884)$$

$$5.0\text{ M NaClO}_4: \log K_s = 7.82 \quad (a_w = 0.777)$$

If we apply the SIT extrapolation to these constants

$$\log K_s^\circ = \log K_s(I) + (\text{x} + 1) \log a_{\text{H}_2\text{O}} - 0\text{ D} + \Delta\epsilon\text{ I}$$

and use the interaction coefficients for H^+ and Na^+ from the NEA-TDB, we obtain

$$\log K_s^\circ = 7.28 \text{ (corresponding to } \log K_{\text{sp}}^\circ = -10.88) \text{ and}$$

$$\Delta\epsilon = -0.06 \pm 0.02 \text{ and } \epsilon(\text{NpO}_2^+/\text{ClO}_4^-) = 0.21 \pm 0.05$$

(if we assume an unhydrated solid with $\text{x} = 0$)

or

$$\Delta\epsilon = -0.01 \pm 0.01 \text{ and } \epsilon(\text{NpO}_2^+/\text{ClO}_4^-) = 0.26 \pm 0.04$$

(if we assume the hydrated solid with $\text{x} = 3.5$)

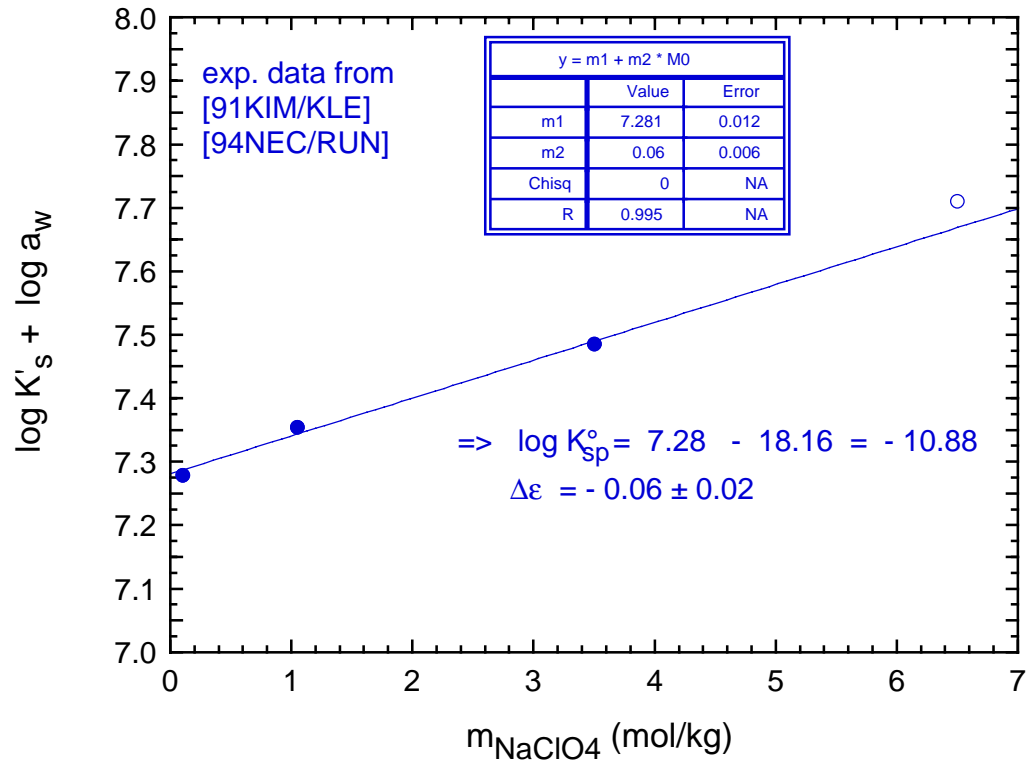
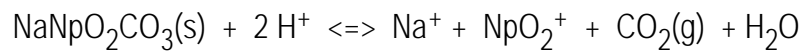
These values are in good agreement with the SIT coefficient selected in the NEA-TDB (0.25 ± 0.05), i.e. the experimental data could, vice versa, be interpreted with a fixed value of $\Delta\epsilon$ taken from the NEA-TDB and then all values of $\log K_s$ in 0.1 to 3.5 and even in 6.5 m NaClO_4 would lead to a consistent value at $I = 0$.

The values for $\epsilon(\text{NpO}_2^+/\text{ClO}_4^-)$ evaluated here from our Np(V) carbonate solubilities agree also with those from our solvent extraction study in [95NEC/FAN] (0.20 ± 0.03) and from our study on the solubility of $\text{NpO}_2\text{OH(s)}$ [92NEC/KIM] (0.19 ± 0.04).

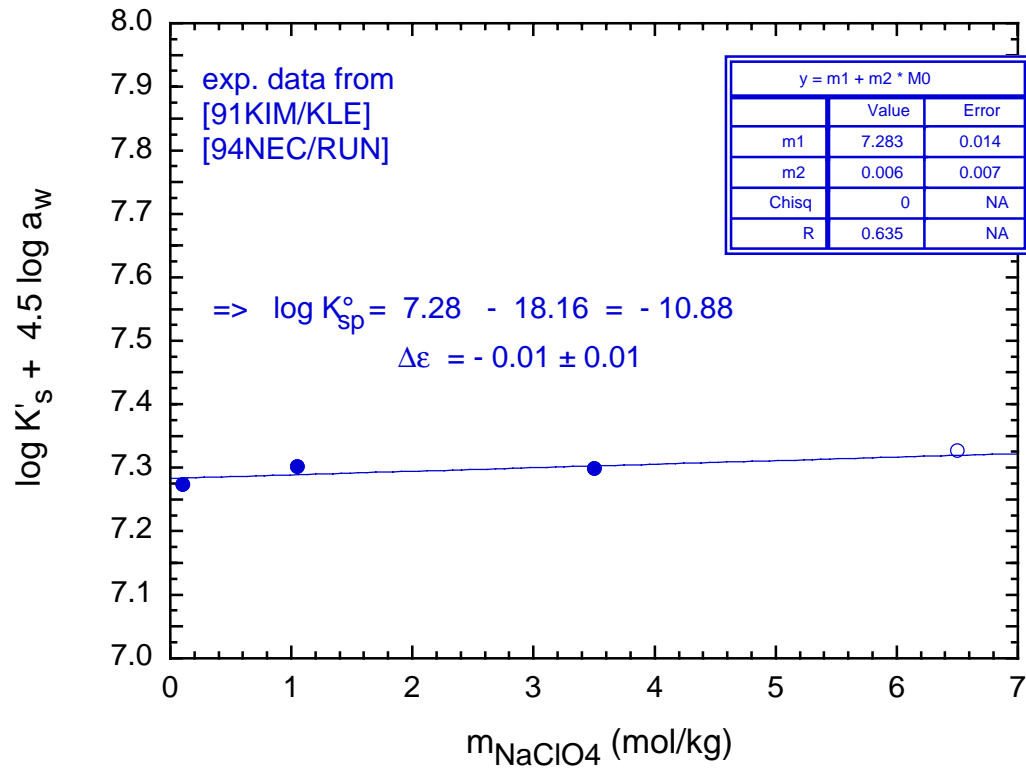
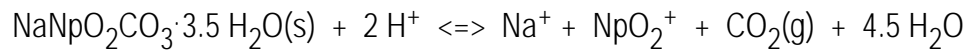
The application of the SIT extrapolation is illustrated in the Figures on the next pages. Formulating the dissolution reaction in this way (excluding the uncertainties from experimental carbonate dissociation constants) the scattering of our experimental data is very small (< 0.05 log units). This demonstrates that the solid phase must have been the same in these experiments.

We hope that these considerations will help to convince you that

- 1) our results are self-consistent and
- 2) reasonable SIT coefficients for the NpO_2^+ ion can be derived from the solubility experiments in [91KIM/KLE, 94NEC/RUN], but not those given in the NEA-TDB review (Table A.24, p.842). We would appreciate very much, if this table could be changed in the final version.



$$\begin{aligned} \epsilon(\text{NpO}_2^+/\text{ClO}_4^-) &= \Delta\epsilon - \epsilon(\text{Na}^+/\text{ClO}_4^-) + 2 \epsilon(\text{H}^+/\text{ClO}_4^-) \\ &= (-0.06 \pm 0.02) - (0.01 \pm 0.01) + 2 (0.14 \pm 0.02) = 0.21 \pm 0.05 \end{aligned}$$



$$\begin{aligned} \epsilon(\text{NpO}_2^+/\text{ClO}_4^-) &= \Delta\epsilon - \epsilon(\text{Na}^+/\text{ClO}_4^-) + 2 \epsilon(\text{H}^+/\text{ClO}_4^-) \\ &= (-0.01 \pm 0.01) - (0.01 \pm 0.01) + 2 (0.14 \pm 0.02) = 0.26 \pm 0.04 \end{aligned}$$